



HAL
open science

Sur des aspects algébriques et combinatoires de l'Analyse Relationnelle.

Julien Ah-Pine

► **To cite this version:**

Julien Ah-Pine. Sur des aspects algébriques et combinatoires de l'Analyse Relationnelle.. Mathématiques [math]. Université de Paris VI, Pierre et Marie Curie, 2007. Français. NNT: . tel-01515257

HAL Id: tel-01515257

<https://hal.science/tel-01515257v1>

Submitted on 27 Apr 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

THÈSE DE DOCTORAT DE L'UNIVERSITÉ PARIS VI

Spécialité :

MATHÉMATIQUES

Présentée par :

M. Julien AH-PINE

Pour obtenir le grade de :

DOCTEUR DE L'UNIVERSITÉ PARIS VI

Sujet de la thèse :

**SUR DES ASPECTS ALGÈBRIQUES ET COMBINATOIRES
DE L'ANALYSE RELATIONNELLE**

**APPLICATIONS EN CLASSIFICATION AUTOMATIQUE, EN THÉORIE
DU CHOIX SOCIAL ET EN THÉORIE DES TRESSSES**

Soutenue le 26 Octobre 2007, devant le jury composé de :

MM.	Paul DEHEUVELS	Président
	Jean-François MARCOTORCHINO	Directeur de thèse
	Pierre CAZES	Rapporteur
	Israël-César LERMAN	Rapporteur
	Younes BENNANI	Examineur
	Jean MAIRESSE	Examineur
	Gilbert SAPORTA	Examineur

PREAMBULE

Cette thèse s'est déroulée selon une "Convention Industrielle de Formation par la Recherche" (CIFRE) proposée par l'Agence Nationale de la Recherche Technique (ANRT). Ce mode de financement consiste en un partenariat entre une entreprise et une université. Dans le cas de la présente thèse CIFRE, les partenaires ont été d'une part, la société Thales Communications et d'autre part, l'université Pierre et Marie Curie (Paris VI). Notamment, J.F. Marcotorchino, responsable du Centre des Nouvelles Technologies d'Analyse de l'Information (CeNTAI) de Thales Communications et P. Deheuvels, directeur du Laboratoire de Statistique Théorique et Appliquée (LSTA) de l'université Paris VI, ont été respectivement les responsables en entreprise et en université de cette thèse.

Les travaux menés *in situ* chez Thales Communications, ont été essentiellement supervisés par J.F. Marcotorchino. Ce-dernier, au travers de la théorie de l'Analyse Relationnelle, s'intéresse entre autre, à de nombreux problèmes faisant intervenir l'étude des relations binaires. Ces structures mathématiques telles que les partitions ou les relations d'ordre se retrouvent dans de nombreux domaines. Mes activités de recherche se sont naturellement orientées vers cette théorie. Toutefois, elles ont été structurées autour de "challenges" que J.F. Marcotorchino, avait identifiés et qu'il m'a incités à relever.

J.F. Marcotorchino m'a ainsi proposé des problèmes relevant de domaines aussi divers que l'analyse de données, la théorie du choix social et la théorie des tresses. Ces domaines appartiennent "traditionnellement" à des spécialités différentes et quasi-étanches : l'analyse des données relève des statistiques et des mathématiques appliquées; la théorie du choix social, des sciences économiques et politiques; et la théorie des tresses, de la topologie algébrique et des mathématiques pures. Aussi, il paraît "surprenant" à première vue de voir cohabiter ces thèmes d'apparence aussi distincts, dans le titre d'un seul et même mémoire de thèse. Et pourtant, ces domaines font intervenir chacun un type de relation binaire particulier. En effet, l'analyse des données et plus particulièrement le problème de la classification automatique fait intervenir des relations d'équivalence; la théorie du choix social et l'agrégation de classements individuels raisonne en termes de relation d'ordre; et le groupe des tresses vu comme une extension du groupe des permutations qui est le groupe sous-jacent aux relations d'ordre strict et total, est, c'est ce que nous nous efforcerons à montrer, une sorte de relation d'ordre "bi-dimensionnelle".

La présence donc de relations structurées dans chacun de ces domaines permet dans une certaine mesure à l'Analyse Relationnelle de s'y intéresser. C'est le constat qu'a fait J.F. Marcotorchino et c'est sur cette base que se situe au départ le travail présenté ici. Ainsi, au delà des thématiques diverses exposées, le fil conducteur de ce mémoire de thèse est l'étude de différents types de relations binaires.

TABLE DES MATIÈRES

Préambule	i
Table des matières	ii
1 Introduction, Rappels et Contributions	1
1.1 L'Analyse Relationnelle et quelques unes de ses applications	1
1.2 Introduction succincte à la théorie des tresses et des noeuds	20
1.3 Contributions	25
I Aspects algébriques et combinatoires de l'AR des relations d'équivalence et applications en classification automatique	30
2 Etude des codages de l'ensemble des partitions possibles d'un ensemble d'objets	31
2.1 Etude de l'ensemble des partitions d'un ensemble d'objets	32
2.2 Les méthodes de classification automatique vues sous un angle algébrique	50
3 Présentation unifiée des mesures de similarité et des critères de partitionnement	68
3.1 Extension des indices de similarité entre vecteurs binaires à des vecteurs non binaires	69
3.2 Extension des critères d'association maximale à des variables quantitatives et la notion d'écart à une tendance centrale	88

3.3	Les notions d'indépendance et d'indétermination dans le cas de la mesure d'association entre variables qualitatives	99
4	Les enveloppes de Condorcet et le calcul consensuel	113
4.1	Structures algébriques booléennes des matrices relationnelles	114
4.2	Les enveloppes de Condorcet	124
4.3	Le calcul consensuel ou l'application du calcul de l'enveloppe de Condorcet à des données non binaires	142
II	Aspects algébriques et combinatoires de l'AR des relations d'ordre et applications en théorie du choix social et en théorie des tresses	150
5	Etudes des codages de l'ensemble des classements possibles d'un ensemble d'objets	151
5.1	Définitions et rappels	152
5.2	Action des transpositions sur les classements	170
6	Sur l'association, l'agrégation et le consensus de relations d'ordre	180
6.1	Sur les critères d'association entre deux relations d'ordre	181
6.2	L'effet Condorcet, le théorème d'impossibilité d'Arrow et les conditions de restriction sur le domaine des préférences individuelles	192
7	Modélisation des tresses et des noeuds par approche relationnelle	215
7.1	Problèmes autour du groupe des tresses et quelques méthodes de résolution . . .	217
7.2	Modélisation des tresses par approche relationnelle	224
7.3	Les problèmes du mot du groupe des tresses, du mot de tresse minimal et du mot de tresse "shortlex" dans le cadre de la modélisation par paires de brins nommés	253
7.4	Modélisation des noeuds par approche relationnelle	271
8	Conclusion et perspectives	284

Bibliographie	287
Liste des tableaux	296
Liste des figures	299
Notations	300
Résumé, Abstract	305

CHAPITRE 1

Introduction, Rappels et Contributions

Sommaire

1.1 L'Analyse Relationnelle et quelques unes de ses applications	1
1.1.1 L'agrégation des similarités	5
1.1.1.1 Classification automatique ou non supervisée, sans fixation de nombre de classes	5
1.1.1.2 Statistiques des contingences	9
1.1.2 L'agrégation des préférences	15
1.1.2.1 Théorie du choix social	16
1.1.2.2 Critères d'association entre relations d'ordre	17
1.2 Introduction succincte à la théorie des tresses et des noeuds	20
1.3 Contributions	25

1.1 L'Analyse Relationnelle et quelques unes de ses applications

L'essentiel des idées présentées dans le cadre de ce mémoire sont issues et/ou inspirées d'une théorie développée à la fin des années 70 par deux chercheurs/ingénieurs de la compagnie IBM à l'époque : Jean-François Marcotorchino et Pierre Michaud. Cette théorie, l'Analyse Relationnelle (AR), provient de la rencontre de plusieurs cultures mathématiques que sont : la théorie des graphes, la logique, l'optimisation mathématique, la statistique, l'analyse de données et la combinatoire. D'un point de vue théorique, l'AR étudie les relations binaires et particulièrement les problèmes qui ont trait à leurs mesures d'association, leurs agrégations et à la détermination d'une relation consensuelle. D'un point de vue pratique, l'AR a permis de modéliser et de contribuer à la résolution de nombreux problèmes issus de la théorie des votes, de l'analyse de données et de la statistique contingentielle.

Supposons que nous disposons d'un ensemble d'objet \mathbb{O} de cardinal N décrit par un ensemble de variables \mathbb{V} de cardinal M . Chaque variable V^k , $k = 1, \dots, M$, engendre des rela-

tions binaires sur l'ensemble des N objets O^i , $i = 1, \dots, N$. Selon le type de variable, nous pouvons obtenir différents types de relations binaires. En ce qui nous concerne, nous nous intéressons à des variables numériques¹ qui sont de deux types : les variables qualitatives et les variables quantitatives. Les variables qualitatives sont constituées de modalités et définissent des relations d'équivalence. Si les modalités sont ordonnées, elles peuvent aussi déterminer des relations de préordre. Les variables quantitatives quant à elles, engendrent naturellement des relations d'ordre. Toutefois, elles peuvent être transformées en des variables qualitatives par un processus de discrétisation à partir duquel nous pouvons engendrer des relations d'équivalence ou de préordre. Le type d'une relation binaire dépend des propriétés relationnelles que celle-ci respecte et notamment parmi les suivantes : la réflexivité, la symétrie, l'antisymétrie forte, la totalité et la transitivité².

Nous rappelons en premier lieu la définition d'une relation binaire sur un ensemble d'objets ainsi que les différentes propriétés relationnelles mentionnées ci-dessus.

Définition 1.1.1 Relation binaire sur un ensemble

Une relation binaire \mathcal{R} sur un ensemble \mathbb{O} , est un couple $(\mathbb{O}, G(\mathcal{R}))$ où $G(\mathcal{R})$, appelé graphe de la relation \mathcal{R} , est un sous-ensemble du produit cartésien $\mathbb{O} \times \mathbb{O}$. Si nous avons $(O^i, O^{i'}) \in G(\mathcal{R})$ alors nous dirons que O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R} , ce que l'on notera par $O^i \mathcal{R} O^{i'}$.

Définition 1.1.2 Complémentaire d'une relation binaire sur un ensemble

Soit \mathcal{R} une relation binaire sur un ensemble \mathbb{O} . Nous pouvons lui associer une relation binaire complémentaire notée $\overline{\mathcal{R}}$ constituée des paires $(O^i, O^{i'})$ du produit cartésien $\mathbb{O} \times \mathbb{O}$ tels que $(O^i, O^{i'}) \notin G(\mathcal{R})$. Cette relation complémentaire sera notée $\overline{\mathcal{R}}$.

Nous rappelons dans le tableau tab-1.1, la définition des différentes propriétés relationnelles que peut respecter une relation binaire \mathcal{R} .

L'étude des données de type relationnel a engendré de nombreux travaux. Nous citerons par exemple, les références suivantes concernant les travaux de recherches suivants [Ler81], [Mon85], [HA85], [BGH89], [BL95]. Mais l'Analyse Relationnelle telle que nous la développerons ici concerne la branche de ce domaine définie par J.F. Marcotorchino et P. Michaud. Celle-ci se distingue des autres approches du fait des points suivants.

Dans un premier temps, l'AR représente les relations binaires engendrées par les variables V^k , $k = 1, \dots, M$ au sein des objets O^i , $i = 1, \dots, N$, par des matrices binaires de comparaisons par paires de ces objets. Ces matrices sont appelées matrices relationnelles ou matrices de Condorcet en hommage au Marquis de Condorcet qui fut le premier à représenter les relations sous cette forme [MM79], [Mic82], [Mic85a], [Mic85b]. De manière générale, soit \mathcal{R} une relation binaire sur l'ensemble des objets \mathbb{O} , nous notons alors par C sa matrice relationnelle. Le terme général de cette matrice, $C_{ii'}$, $i, i' = 1, \dots, N$, est donné par la relation suivante :

$$C_{ii'} = \begin{cases} 1 & \text{si } O^i \mathcal{R} O^{i'} \\ 0 & \text{sinon} \end{cases} \quad (1.1)$$

La matrice relationnelle représentative du complémentaire $\overline{\mathcal{R}}$ d'une relation \mathcal{R} est notée \overline{C} et est définie par :

$$\overline{C}_{ii'} = \begin{cases} 1 & \text{si } O^i \overline{\mathcal{R}} O^{i'} \\ 0 & \text{sinon} \end{cases}$$

1. au sens des bases de données structurées
 2. d'autres propriétés existent telle que l'intermédiarité mais nous nous sommes pour l'instant restreint aux propriétés les plus fondamentales

Propriété relationnelle	Définition logique
Réflexivité	$O^i \mathcal{R} O^i \quad \forall O^i \in \mathbb{O}$
Irréflexivité	$O^i \overline{\mathcal{R}}^k O^i \quad \forall O^i \in \mathbb{O}$
Symétrie	$O^i \mathcal{R} O^{i'} \Rightarrow O^{i'} \mathcal{R} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O}$
Antisymétrie forte (asymétrie)	$O^i \mathcal{R} O^{i'} \Rightarrow O^{i'} \overline{\mathcal{R}}^k O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} : O^i \neq O^{i'}$
Transitivité	$O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''} \quad \forall (O^i, O^{i'}, O^{i''}) \in \mathbb{O} \times \mathbb{O} \times \mathbb{O}$
Totalité	$O^i \mathcal{R} O^{i'} \vee O^{i'} \mathcal{R} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} : O^i \neq O^{i'}$

TABLE 1.1 – Propriétés relationnelles exprimées sous forme de règles logiques

Le plus souvent, nous avons la relation suivante qui lie une matrice relationnelle et son complémentaire :

$$\overline{C}_{ii'} = 1 - C_{ii'}$$

Si nous notons U_N la matrice carrée de dimension N remplie de 1, nous avons l'expression matricielle suivante :

$$\overline{C} = U_N - C \tag{1.2}$$

Cette représentation simple mais riche permet en fait, d'exprimer les propriétés relationnelles qui définissent le type de la relation binaire, par des équations linéaires d'égalités ou d'inégalités qui sont fonctions des termes de la matrice relationnelle. Cette propriété du codage relationnel est fondamentale en AR et constitue l'un de ces principes majeurs puisqu'elle permet notamment une meilleure résolution (en terme d'efficacité) de certains problèmes combinatoires que nous présenterons par la suite en utilisant les outils de la programmation linéaire en nombres bivalents. Nous donnons dans le tableau tab-1.2, les expressions en termes relationnels des différentes propriétés d'une relation binaire \mathcal{R} représentée par sa matrice relationnelle associée C [MM79].

A l'aide de ce codage, l'AR a permis, entre autre, de résoudre d'une manière efficace, le problème d'agrégation multicritère que nous rappelons ci-dessous.

Supposons que nous souhaitions déterminer l'ordre collectif ou consensuel d'un ensemble d'objets issu des ordres donnés par deux variables telles que le poids exprimé en kilogrammes et la taille exprimée en centimètres.

Il faut pour cela "agrèger" les relations d'ordres induites par chaque variable.

Lorsqu'on raisonne sur les objets d'une variable, il est possible de déterminer des statistiques telle que la moyenne en additionnant les valeurs prises par chaque objet.

Lorsqu'on raisonne sur les variables descriptives d'un objet, il est, *a contrario*, délicat d'agrèger pour chaque objet les valeurs prises par chaque variable : il n'y a pas de sens à additionner des kilogrammes et des centimètres.

Propriété relationnelle	Expression sous forme de contraintes linéaires fonctions des termes des éléments d'une matrice relationnelle
Réflexivité	$C_{ii} = 1 \quad \forall i = 1, \dots, N$
Irréflexivité	$C_{ii} = 0 \quad \forall i = 1, \dots, N$
Symétrie	$C_{ii'} = C_{i'i} \quad \forall i, i' = 1, \dots, N$
Antisymétrie forte (asymétrie)	$C_{ii'} + C_{i'i} \leq 1 \quad \forall i, i' = 1, \dots, N : i \neq i'$
Transitivité	$C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 \quad \forall i, i', i'' = 1, \dots, N$
Totalité	$C_{ii'} + C_{i'i} \geq 1 \quad \forall i, i' = 1, \dots, N : i \neq i'$

TABLE 1.2 – Propriétés relationnelles exprimées sous forme de contraintes linéaires dans le cadre de l'AR

Si nous déterminons pour chaque variable l'ordre induit et que nous le représentons sous forme de matrice relationnelle, l'agrégation devient possible : elle est déterminée par l'addition des deux matrices relationnelles.

Pour chaque paire d'objets $(O^i, O^{i'})$, nous obtenons alors le nombre de variables indiquant que le rang de l'objet O^i est inférieur à celui de l'objet $O^{i'}$.

Supposons désormais que nous ayons M variables numériques telles que la taille et que nous calculions la matrice relationnelle représentative de l'ordre induit pour chacune d'elle. L'addition de ces M matrices relationnelles individuelles définit la matrice relationnelle (ou de Condorcet) collective C dont le terme général est donné par :

$$C_{ii'} = \sum_{k=1}^M C_{ii'}^k \quad (1.3)$$

La matrice relationnelle collective est ainsi le résultat de l'agrégation des relations binaires individuelles. Le principe d'agrégation que nous venons de décrire dans le cadre des relations d'ordre, qui est fondé sur la représentation sous forme de matrices de comparaisons par paires des relations est ainsi, un deuxième aspect spécifique à l'AR.

Suite à cette agrégation, il est alors possible de proposer une solution efficace au problème de la détermination d'une relation collective ou consensuelle. Ce problème consiste à calculer une relation qui résume au mieux l'information contenue dans un ensemble de relations.

L'AR se base alors sur la maximisation du critère de Condorcet qui donne une mesure d'association entre d'une part, la relation consensuelle recherchée notée X et d'autre part, l'ensemble des relations que l'on cherche à résumer.

Dans le cadre spécifique de l'agrégation multicritère qui traite de la détermination d'un ordre collectif au sein de plusieurs juges, ce critère permet d'ailleurs de déterminer le vainqueur de Condorcet ou le classement consensuel global [MM79], [Mic85b].

Le critère de Condorcet est donné par la formule suivante :

$$\text{Condorcet}(\mathbf{C}, X) = \sum_{i=1}^N \sum_{i'=1}^N \mathbf{C}_{ii'} X_{ii'} + \sum_{i=1}^N \sum_{i'=1}^N \bar{\mathbf{C}}_{ii'} \bar{X}_{ii'} \quad (1.4)$$

où X est une matrice relationnelle, \bar{X} son complémentaire, \mathbf{C} la matrice relationnelle collective des M relations étudiées et $\bar{\mathbf{C}} = \sum_{k=1}^M \bar{\mathbf{C}}^k$ la somme des matrices relationnelles complémentaires qui vaut, dans le cas où il n'y a pas de données manquantes :

$$\bar{\mathbf{C}} = M U_N - \mathbf{C} \quad (1.5)$$

En remplaçant $\bar{X}_{ii'}$ par $1 - X_{ii'}$ et $\mathbf{C}_{ii'}$ par $M - \mathbf{C}_{ii'}$, nous aboutissons à la forme simplifiée du critère de Condorcet :

$$\text{Condorcet}(\mathbf{C}, X) = \sum_{i=1}^N \sum_{i'=1}^N \left(\mathbf{C}_{ii'} - \frac{M}{2} \right) X_{ii'} \quad (1.6)$$

Ecrite sous cette forme, la maximisation du critère de Condorcet s'exprime de la manière suivante : les objets O^i et $O^{i'}$ sont susceptibles d'être en relation³, si une majorité de variables ou de juges supporte la relation entre ces deux objets⁴.

Nous venons d'introduire les concepts de base qui ont fait naître l'AR ainsi que les notations générales associées. Nous présentons, dans ce qui suit, de manière plus précise, certaines applications de cette théorie en analyse de données et en agrégation multicritère. Ces applications concernent, entre autre, celles qui intéressent le présent mémoire de thèse :

- les problèmes et traitements des relations d'équivalence et leurs applications en **classification automatique**⁵ de données
- les problèmes et traitements des relations d'ordre et leurs applications en **agrégation multicritère, en théorie du choix social et de consensus d'opinions**
- l'extension de l'AR à l'étude de nouveaux types de relations que sont les **tresses**

1.1.1 L'agrégation des similarités

L'agrégation des similarités regroupe l'étude des relations d'équivalence dans le cadre de l'AR et de leurs contributions au domaine de l'analyse de données. Nous présenterons succinctement, les apports significatifs de l'AR dans deux sous-branches de la statistique appliquée : la classification automatique et les statistiques contingentielles.

1.1.1.1 Classification automatique ou non supervisée, sans fixation de nombre de classes

Le problème consiste à déterminer automatiquement une partition, c'est à dire une relation d'équivalence, d'un ensemble d'objets. Cette partition est, en principe, telle que deux objets appartenant à une même classe sont les plus similaires possible alors que deux objets appartenant à deux classes distinctes sont les plus différents possible.

Il existe une approche "naïve" qui permet de déterminer de manière exacte la solution à ce problème. Il s'agit d'énumérer toutes les partitions possibles de l'ensemble d'objets et de

3. ie $X_{ii'} = 1$

4. ie $\mathbf{C}_{ii'} - \frac{M}{2} \geq 0$

5. ou non supervisée, le terme anglo-saxon étant "clustering"

calculer pour chacune d'elle la valeur d'un critère à optimiser. Une fois que toutes les partitions ont été parcourues, il suffit alors de sélectionner la ou les meilleures partitions au sens du critère utilisé. Toutefois, cette approche est combinatoire car le nombre de partitions possibles d'un ensemble de N objets, donné par le nombre de Bell noté $Bell(N)$, croît de manière exponentielle :

$$Bell(N) = \frac{1}{e} \sum_{i=1}^{\infty} \frac{i^N}{i!}$$

A titre d'exemple, le nombre de Bell pour $N = 71$ vaut approximativement 4×10^{74} partitions possibles!! Enumérer un tel nombre de partitions à l'aide d'un ordinateur actuel, nécessiterait plusieurs centaines d'années!

Le problème de la classification automatique est donc un problème hautement combinatoire. En théorie de la complexité, il est classé parmi les problèmes \mathcal{NP} -difficiles [Wak98].

Même si des solutions exactes ont été obtenues par programmation linéaire "relaxée"⁶ jusqu'à $N = 120$, il n'en reste pas moins vrai que l'approche nécessaire et pratique dans ce cas sera une approche heuristique. Dans ce contexte, nous discuterons plus particulièrement, des algorithmes spécifiques qui sont fondés sur une compréhension approfondie des structures manipulées et non pas sur une approche généraliste telles que le recuit simulé, la recherche avec liste "tabou", les colonies de fourmis...

Plusieurs méthodes spécifiques au problème ont été en effet développées afin de déterminer dans des temps de traitements raisonnables, une partition solution du problème de classification automatique. Ces heuristiques n'atteignent pas toujours l'optimum global. De plus, elles comportent certains inconvénients. En particulier, certaines nécessitent la fixation *a priori* ou *a posteriori* d'un nombre de classes. Cet inconvénient peut être important dans le cadre de la "découverte de connaissances", dans la mesure où nous avons rarement *a priori* une idée du nombre de classes que comporte la partition cherchée. Parmi les méthodes les plus employées, nous citerons les méthodes de type centres mobiles et les méthodes hiérarchiques. Nous présenterons plus précisément ces méthodes ultérieurement.

Sur un plan théorique, l'AR a permis de modéliser le problème de la classification automatique sous la forme d'un programme linéaire en nombres bivalents qui est basé sur la maximisation du critère de Condorcet⁷ (1.4).

Dans le cadre des relations d'équivalence, nous avons les termes $C_{ii'} = \sum_k C_{ii'}^k$ qui donnent le nombre de variables indiquant que O^i est dans la même classe que $O^{i'}$. Cette quantité peut s'apparenter à une **mesure de la similarité** entre ces deux objets.

En ce qui concerne les termes $\bar{C}_{ii'}$, qui donnent donc une **mesure de la dissimilarité** entre objets, nous considérerons, dans le contexte des relations d'équivalence, la formulation plus générale suivante :

$$\bar{C}_{ii'} = \mathcal{M}^t(C_{ii}, C_{i'i'}) - C_{ii'} \quad (1.7)$$

où $\mathcal{M}^t(A, B)$ est une moyenne d'ordre t avec $-\infty < t < +\infty$ donnée par :

$$\mathcal{M}^t(A, B) = \left[\frac{1}{2} (A^t + B^t) \right]^{\frac{1}{t}} \quad (1.8)$$

La dissimilarité entre deux objets doit être inversement proportionnelle à leur similarité. Nous considérerons ainsi qu'elle est donnée par une moyenne des auto-similarités moins la similarité dans l'hypothèse courante qu'un objet ne peut être davantage similaire qu'à lui-même ce qui implique que $C_{ii'} \leq C_{ii}$ et $\bar{C}_{ii'} \geq 0, \forall i, i' = 1, \dots, N$.

6. en utilisant l'approche "dual du dual" cf [MM78]

7. ou sur d'autres critères linéaires relationnels que nous présenterons ultérieurement

Dans le cas où il n’y a pas de données manquantes, nous constatons que $C_{ii} = M, \forall i = 1, \dots, N$, ce qui nous ramène à $\bar{C}_{ii'} = M - C_{ii'}$. En effet, nous avons dans ce cas, $\mathcal{M}^t(C_{ii}, C_{i'i'}) = M$. En revanche, s’il y a des données manquantes la dissimilarité sera plus ou moins grande selon la moyenne utilisée. Nous citerons pour exemple les cas suivants, qui peuvent être pris en compte dans le cadre des variables qualitatives ou de variables binaires :

- $t = 1 \quad \leftrightarrow \quad \mathcal{M}^1(A, B) = \text{Moyenne arithmétique entre } A \text{ et } B$
- $t \rightarrow -\infty \quad \leftrightarrow \quad \mathcal{M}^t(A, B) = \min(A, B)$

Par défaut, nous prendrons la définition suivante du complémentaire qui possède des propriétés de distance comme nous aurons l’occasion de le constater ultérieurement :

$$\bar{C}_{ii'} = \frac{C_{ii} + C_{i'i'}}{2} - C_{ii'} \quad (1.9)$$

Dans ce cas, nous obtenons la simplification suivante⁸ du critère de Condorcet que nous considérerons par la suite dans le contexte de la classification automatique :

$$\text{Condorcet}(\mathbf{C}, X) = \sum_{i=1}^N \sum_{i'=1}^N \left(C_{ii'} - \frac{1}{2} \frac{C_{ii} + C_{i'i'}}{2} \right) X_{ii'} \quad (1.10)$$

La modélisation du problème de la classification automatique par programmation linéaire, repose alors d’une part, sur la linéarité⁹ du critère de Condorcet lorsque la matrice relationnelle collective¹⁰ est préalablement calculée et d’autre part, sur la propriété que les conditions vérifiées par une relation d’équivalence s’expriment, dans le codage relationnel, sous la forme de contraintes linéaires.

Nous rappelons qu’une partition est une relation binaire qui vérifie les propriétés de réflexivité, de symétrie, et de transitivité. Le programme linéaire que l’on peut résoudre à l’aide d’un “solver” et qui permet de déterminer la partition maximisant le critère de Condorcet précédemment présenté, sans fixer un nombre de classe, est le suivant [Mar81], [MM79] :

$$\begin{array}{llll} \max_X \text{Condorcet}(\mathbf{C}, X) & \text{s.t.} & X_{ii'} \in \{0, 1\} & (\text{binarité}) \\ & & X_{ii} = 1 & \forall i = 1, \dots, N \quad (\text{réflexivité}) \\ & & X_{i'i} = X_{ii'} & \forall i, i' = 1, \dots, N \quad (\text{symétrie}) \\ & & X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 & \forall i, i', i'' = 1, \dots, N \quad (\text{transitivité}) \end{array} \quad (1.11)$$

La mesure de similarité entre deux objets est une notion fondamentale dans un problème de classification automatique. Indépendamment de la méthode de classification employée, une mauvaise définition de celle-ci aboutira à une partition solution de qualité médiocre. L’application de l’AR en classification automatique a donc naturellement engendré des travaux autour de la notion de similarité. Dans ce cadre, nous pouvons citer les travaux de H. Benhadda et de J.F. Marcotorchino qui ont défini différentes notions de **similarité régularisée**. Cette approche consiste à raffiner les mesures de similarités entre objets décrits notamment par des variables qualitatives [Ben98] [BM98]. H. Benhadda et J.F. Marcotorchino définissent des pondérations qui tiennent compte des caractéristiques intrinsèques de chaque variable ou de chaque modalité.

De manière concrète, soit \mathcal{T} le tableau de données où on “mesure” les modalités prises par un ensemble de N objets pour un ensemble de M variables qualitatives. Supposons que chaque variable $V^k, k = 1, \dots, M$, possède p_k modalités différentes. Pour chaque variable, nous considérons alors la matrice disjonctive K^k , de dimensions $(N \times p_k)$, et dont le terme principal est

8. en suivant la même démarche que celle permettant d’obtenir (1.6)
 9. en $X_{ii'}$
 10. ie matrice de similarités entre objets \mathbf{C}

donné par, $\forall i = 1, \dots, N, \forall j = 1, \dots, p_k$:

$$K_{ij}^k = \begin{cases} 1 & \text{si } O^i \text{ possède la modalité } D^j \in V^k \\ 0 & \text{sinon} \end{cases} \quad (1.12)$$

Nous pouvons considérer l'ensemble des modalités de l'ensemble des variables qualitatives. Notons alors par \mathbb{D} cet ensemble de modalités. Celui-ci est de cardinal $P = \sum_{k=1}^M p_k$. Dans ce contexte, il est coutume de définir en analyse de données le tableau disjonctif complet \mathbf{K} qui est la juxtaposition de l'ensemble des matrices disjonctives $K^k, \forall k = 1, \dots, M$. Le tableau disjonctif complet associé à \mathcal{T} est ainsi de dimension $(N \times P)$. Son terme général est donné par, $\forall i = 1, \dots, N, \forall j = 1, \dots, P$:

$$\mathbf{K}_{ij} = \begin{cases} 1 & \text{si } O^i \text{ possède la modalité } D^j \in \mathbb{D} \\ 0 & \text{sinon} \end{cases} \quad (1.13)$$

Dans le contexte du traitement de variables qualitatives, la similarité $\mathbf{S}_{ii'}$ entre les objets O^i et $O^{i'}$, appelée similarité logique, est égale au terme $\mathbf{C}_{ii'}$ de la matrice relationnelle collective issue de l'agrégation des relations d'équivalence induite par les variables qualitatives. Nous avons en fait la relation suivante qui lie matrice relationnelle collective et tableau disjonctif complet :

$$\mathbf{S}_{ii'} = \mathbf{C}_{ii'} = \sum_{j=1}^P \mathbf{K}_{ij} \mathbf{K}_{i'j} \quad (1.14)$$

La similarité régularisée $\mathbf{S}_{ii'}^\alpha$ vise alors à pondérer le "matching¹¹" de deux objets pour une modalité D^j par un poids α_j :

$$\mathbf{S}_{ii'}^\alpha = \sum_{j=1}^P \alpha_j \mathbf{K}_{ij} \mathbf{K}_{i'j} \quad (1.15)$$

La définition particulière d'un poids provient d'une interprétation ayant trait aux notions de rareté et de difficulté. Nous rappelons dans tab-1.3, les différents types de similarités régularisées définis pour des variables qualitatives dans [BM98].

Dans ce tableau, nous notons par p_k^j le nombre de modalités de la variable qualitative V^k à laquelle le descripteur D^j appartient :

$$p_k^j = p_k \quad \forall D^j \in V^k \quad (1.16)$$

La similarité de rareté cherche à favoriser des descripteurs dont les fréquences sont faibles.

Les similarités de difficulté et empirique se basent sur des interprétations probabilistes de la difficulté que deux objets aient la même modalité d'une variable.

Dans le cas de la similarité de difficulté, la probabilité de "matching¹²" est donnée, sous l'hypothèse d'uniformité, par $1/(p_k^j)^2$. En effet, si toutes les modalités $D^j \in V^k$ sont équiprobables, alors la probabilité qu'un objet pris au hasard ait la modalité D^j de V^k est $(1/p_k^j)$ où $p_k^j = p_k$ est le nombre de modalités D^j de V^k . Ainsi, sous l'hypothèse d'indépendance, la probabilité que deux objets O^i et $O^{i'}$ pris au hasard aient la même modalité de V^k est $\sum_{D^j \in V^k} (1/p_k^j)^2 = 1/p_k^j$ et la difficulté de "matching" est donnée par le complémentaire de cette probabilité.

Dans le cas de la similarité empirique, la probabilité de "matching" est estimée par $\frac{1}{N^2} \sum_{D^j \in V^k} (\mathbf{K}_{.j})^2$. Cette dernière expression appelée densité de la variable V^k peut aussi être écrite de la manière

11. ie lorsque les deux objets possèdent la même modalité

12. ie la probabilité pour que deux objets aient la même modalité d'une variable

Similarité régularisée	Poids α_j	Notations
Similarité de rareté	$\alpha_j = \frac{1}{\mathbf{K}_j}$	$\hat{\mathbf{C}}_{ii'} = \sum_j \frac{1}{\mathbf{K}_j} \mathbf{K}_{ij} \mathbf{K}_{i'j}$
Similarité de difficulté	$\alpha_j = 1 - \frac{1}{p_k^j}$	$\check{\mathbf{C}}_{ii'} = \sum_j (1 - \frac{1}{p_k^j}) \mathbf{K}_{ij} \mathbf{K}_{i'j}$
Similarité empirique	$\alpha_j = 1 - \frac{1}{N^2} \sum_{D^{j'} \in V^k} (\mathbf{K}_{.j'})^2$ où $D^j \in V^k$	$\tilde{\mathbf{C}}_{ii'} = \sum_j (1 - \frac{\sum_{i,i'} C_{ii'}^k}{N^2}) \mathbf{K}_{ij} \mathbf{K}_{i'j}$ où $D^j \in V^k$

TABLE 1.3 – Rappels des différentes similarités régularisées dans le cas de variables qualitatives

suivante : $\frac{1}{N^2} \sum_{i,i'} C_{ii'}^k$. Elle correspond à l'estimation empirique de la probabilité que deux objets pris au hasard aient la même modalité de V^k . De ce fait $(1 - \frac{1}{N^2} \sum_{D^j \in V^k} (\mathbf{K}_{.j})^2)$ donne la difficulté de "matching" relative à une variable qualitative.

Les similarités de difficulté et empirique cherchent donc à favoriser le "matching" d'objets pour des variables pour lesquelles il est difficile à deux objets pris au hasard, de partager la même modalité.

1.1.1.2 Statistiques des contingences

Les statistiques des contingences cherchent à définir des critères d'association permettant de mesurer puis de tester le lien entre deux ou plusieurs variables qualitatives. Parmi les critères d'association les plus employés, nous pouvons citer le critère du Chi-deux de Pearson, le critère du Chi-deux de Tchuprow, le critère de Belson, le critère de Rand ... Ces critères sont définis et étudiés à partir d'un tableau nommé tableau de contingence qui croise deux variables qualitatives selon leurs modalités. Soient deux variables qualitatives V^k et V^l constituées respectivement de p_k et p_l modalités. Notons K^k et K^l leurs matrices disjonctives de dimensions respectives $(N \times p_k)$ et $(N \times p_l)$. Le tableau de contingence croisant les variables V^k et V^l noté \mathbf{n}^{kl} et de dimension $(p_k \times p_l)$ est défini par :

$$\mathbf{n}^{kl} = {}^t K^k \cdot K^l \quad (1.17)$$

où ${}^t K^k$ est la transposée de la matrice K^k .

Son terme général s'interprète de la manière suivante :

$$\mathbf{n}_{uv}^{kl} : \text{nombre d'objets ayant à la fois la modalité } D^u \text{ de } V^k \text{ et la modalité } D^v \text{ de } V^l \quad (1.18)$$

Nous avons par ailleurs les définitions suivantes :

$$\begin{aligned} \mathbf{n}_{.u}^{kl} = \sum_v \mathbf{n}_{uv}^{kl} & : \text{nombre d'objets ayant la modalité } D^u \text{ de } V^k \\ \mathbf{n}_{.v}^{kl} = \sum_u \mathbf{n}_{uv}^{kl} & : \text{nombre d'objets ayant la modalité } D^v \text{ de } V^l \\ N = \sum_{u,v} \mathbf{n}_{uv}^{kl} & : \text{nombre total d'objets} \end{aligned} \quad (1.19)$$

Nous rappelons au préalable, quelques uns des critères d'association entre variables qualitatives. Nous nous basons essentiellement sur ceux étudiés par J.F. Marcotorchino dans [Mar84a],

[Mar84b], [Mar85] et [Mar86]. Dans le souci d'alléger les notations, nous omettrons les indices supérieurs du terme général du tableau de contingence.

Définition 1.1.3 Critère de Belson

$$\text{Belson}(V^k, V^l) = \sum_{u=1}^{p_k} \sum_{v=1}^{p_l} \left(\mathbf{n}_{uv} - \frac{\mathbf{n}_u \cdot \mathbf{n}_v}{N} \right)^2 \quad (1.20)$$

Définition 1.1.4 Critère de Rand

$$\text{Rand}(V^k, V^l) = \frac{2 \sum_{u,v} \mathbf{n}_{uv}^2 - \sum_u \mathbf{n}_u^2 - \sum_v \mathbf{n}_v^2 + N^2}{N^2} \quad (1.21)$$

Définition 1.1.5 Critère de Janson et Vegelius

$$\text{JV}(V^k, V^l) = \frac{p_k p_l \sum_{u,v} \mathbf{n}_{uv}^2 - p_k \sum_u \mathbf{n}_u^2 - p_l \sum_v \mathbf{n}_v^2 + N^2}{\sqrt{(p_k(p_k - 2) \sum_u \mathbf{n}_u^2 + N^2) (p_l(p_l - 2) \sum_v \mathbf{n}_v^2 + N^2)}} \quad (1.22)$$

Nous donnons une deuxième expression du critère de Janson et Vegelius qui a été définie par J.F. Marcotorchino dans [Mar84a] et [Mar84b].

$$\text{JV}(V^k, V^l) = \frac{p_k p_l \sum_{u,v} \left(\mathbf{n}_{uv} - \left(\frac{\mathbf{n}_u}{p_l} + \frac{\mathbf{n}_v}{p_k} - \frac{N}{p_k p_l} \right) \right)^2}{\sqrt{(p_k(p_k - 2) \sum_u \mathbf{n}_u^2 + N^2) (p_l(p_l - 2) \sum_v \mathbf{n}_v^2 + N^2)}} \quad (1.23)$$

L'expression (1.23) du critère de Janson et Vegelius permet de mettre en évidence le lien de ce critère avec la notion d'indétermination au sens de J.F. Marcotorchino et sur laquelle nous reviendrons ultérieurement.

Définition 1.1.6 Critère de Lerman

Lerman(V^k, V^l) =

$$\frac{\sum_{u,v} \mathbf{n}_{uv} (\mathbf{n}_{uv} - 1) - \frac{\sum_u \mathbf{n}_u (\mathbf{n}_u - 1) \sum_v \mathbf{n}_v (\mathbf{n}_v - 1)}{N(N-1)}}{\sqrt{\left(\sum_u \mathbf{n}_u (\mathbf{n}_u - 1) \left(1 - \sum_u \frac{\mathbf{n}_u (\mathbf{n}_u - 1)}{N(N-1)} \right) \right) \left(\sum_v \mathbf{n}_v (\mathbf{n}_v - 1) \left(1 - \sum_v \frac{\mathbf{n}_v (\mathbf{n}_v - 1)}{N(N-1)} \right) \right)}} \quad (1.24)$$

Nous donnons par ailleurs, la formulation modifiée du critère de Lerman noté Lerman'(V^k, V^l) due à A. Najah Idrissi [NI00]. Nous pouvons par ailleurs, rapprocher cette formulation du coefficient tétrachorique¹³ :

$$\text{Lerman}'(V^k, V^l) = \frac{\sum_{u,v} \mathbf{n}_{uv}^2 - \frac{\sum_u \mathbf{n}_u^2 \sum_v \mathbf{n}_v^2}{N^2}}{\sqrt{\left(\sum_u \mathbf{n}_u^2 \left(1 - \sum_u \frac{\mathbf{n}_u^2}{N^2} \right) \right) \left(\sum_v \mathbf{n}_v^2 \left(1 - \sum_v \frac{\mathbf{n}_v^2}{N^2} \right) \right)}} \quad (1.25)$$

Définition 1.1.7 Critère du Chi-deux de Tchuprow

$$\chi_T^2(V^k, V^l) = \frac{\sum_{u,v} \frac{1}{\mathbf{n}_u \cdot \mathbf{n}_v} \left(\mathbf{n}_{uv} - \frac{\mathbf{n}_u \cdot \mathbf{n}_v}{N} \right)^2}{\sqrt{(p_k - 1)(p_l - 1)}} \quad (1.26)$$

Définition 1.1.8 Critère (dérivé¹⁴) de Jordan

$$\text{Jordan}(V^k, V^l) = \frac{1}{N} \sum_{u,v} \left(\mathbf{n}_{uv} \left(\mathbf{n}_{uv} - \frac{\mathbf{n}_u \cdot \mathbf{n}_v}{N} \right) \right) \quad (1.27)$$

13. cf chapitre 3 section 3.3.1.1.2 page 100

14. ce critère a été défini par J.F. Marcotorchino dans [Mar85] d'après une idée de Ch. Jordan [Jor27]

Définition 1.1.9 *Critère de Light-Margolin*¹⁵

$$\text{LM}(V^k, V^l) = \sum_{u,v} \frac{\mathbf{n}_{uv}^2}{\mathbf{n}_u} - \frac{1}{N} \sum_v \mathbf{n}_{\cdot v}^2 \quad (1.28)$$

Définition 1.1.10 *Critère d'écart carré à l'indépendance*¹⁶

$$\text{E}(V^k, V^l) = \sum_{u,v} \left(\mathbf{n}_{uv}^2 - \frac{\mathbf{n}_u \cdot \mathbf{n}_{\cdot v}}{N^2} \right) \quad (1.29)$$

Dans une série de trois articles [Mar84a], [Mar84b], et [Mar85], J.F. Marcotorchino reprend les études de M.G. Kendall sur les liens entre les tableaux de contingence et les matrices de comparaisons par paires et développe des formules de passage entre ces deux codages. En donnant l'écriture en termes de matrices relationnelles, les travaux de J.F. Marcotorchino ont permis une meilleure compréhension et interprétation des critères d'association.

Les principales formules de passage contingence/relationnel utilisées sont données dans le tableau tab-1.4 où C^k et C^l sont respectivement les matrices relationnelles associées aux variables V^k et V^l .

Ces formules de passage ont permis d'exprimer les critères d'association dans le codage relationnel que nous rappelons ci-dessous.

Propriété 1.1.1 *Expression relationnelle du critère de Belson*

$$\text{Belson}(C^k, C^l) = \sum_{i=1}^N \sum_{i'=1}^N \left(C_{ii'}^k - \frac{C_{i\cdot}^k + C_{\cdot i'}^k}{N} + \frac{C_{\cdot\cdot}^k}{N^2} \right) C_{ii'}^l \quad (1.30)$$

Propriété 1.1.2 *Expression relationnelle du critère de Rand*

$$\text{Rand}(C^k, C^l) = \frac{1}{N^2} \sum_{i,i'} \left(C_{ii'}^k C_{ii'}^l + \bar{C}_{ii'}^k \bar{C}_{ii'}^l \right) \quad (1.31)$$

Nous avons par ailleurs :

$$2\text{Rand}(C^k, C^l) - 1 = \frac{\sum_{i,i'} \left(C_{ii'}^k - \frac{1}{2} \right) \left(C_{ii'}^l - \frac{1}{2} \right)}{\sqrt{\sum_{i,i'} \left(C_{ii'}^k - \frac{1}{2} \right)^2 \sum_{i,i'} \left(C_{ii'}^l - \frac{1}{2} \right)^2}} \quad (1.32)$$

Propriété 1.1.3 *Expression relationnelle du critère de Janson et Vegelius*

$$\text{JV}(C^k, C^l) = \frac{\sum_{i,i'} \left(C_{ii'}^k - \frac{1}{p_k} \right) \left(C_{ii'}^l - \frac{1}{p_l} \right)}{\sqrt{\sum_{i,i'} \left(C_{ii'}^k - \frac{1}{p_k} \right)^2 \sum_{i,i'} \left(C_{ii'}^l - \frac{1}{p_l} \right)^2}} \quad (1.33)$$

Propriété 1.1.4 *Expression relationnelle du critère de Lerman*

$$\text{Lerman}(C^k, C^l) = \frac{\sum_{i,i':i \neq i'} \left(C_{ii'}^k - \sum_{i,i':i \neq i'} \frac{C_{ii'}^k}{N(N-1)} \right) \left(C_{ii'}^l - \sum_{i,i':i \neq i'} \frac{C_{ii'}^l}{N(N-1)} \right)}{\sqrt{\sum_{i,i':i \neq i'} \left(C_{ii'}^k - \sum_{i,i':i \neq i'} \frac{C_{ii'}^k}{N(N-1)} \right)^2 \sum_{i,i':i \neq i'} \left(C_{ii'}^l - \sum_{i,i':i \neq i'} \frac{C_{ii'}^l}{N(N-1)} \right)^2}} \quad (1.34)$$

15. il correspond au numérateur du critère connu sous le nom de τ_b de Goodman et Kruskal

16. il est lié au numérateur du critère de Morey-Agresti

Contingence ↔ Relationnel
$\sum_{u=1}^{p_k} \sum_{v=1}^{p_l} (\mathbf{n}_{uv}^{kl})^2 = \sum_{i=1}^N \sum_{i'=1}^N C_{ii'}^k C_{ii'}^l$
$\sum_u (\mathbf{n}_{u.}^{kl})^2 = \sum_{i,i'} C_{ii'}^k$ $\sum_v (\mathbf{n}_{.v}^{kl})^2 = \sum_{i,i'} C_{ii'}^l$
$\sum_{u,v} \frac{(\mathbf{n}_{uv}^{kl})^2}{\mathbf{n}_{u.}^{kl} \mathbf{n}_{.v}^{kl}} = \sum_{i,i'} \frac{C_{ii'}^k C_{ii'}^l}{C_{i.}^k C_{i.}^l}$
$\sum_{u,v} \mathbf{n}_{uv}^{kl} \mathbf{n}_{u.}^{kl} \mathbf{n}_{.v}^{kl} = \sum_{i,i'} \frac{C_{i.}^k + C_{i.}^l}{2} C_{ii'}^l$
$\sum_{u,v} (\mathbf{n}_{uv}^{kl})^2 \mathbf{n}_{u.}^{kl} = \sum_{i,i'} \frac{C_{i.}^k + C_{i.}^l}{2} C_{ii'}^k C_{ii'}^l$
$\sum_{u,v} \frac{(\mathbf{n}_{uv}^{kl})^2}{\mathbf{n}_{u.}^{kl}} = \sum_{i,i'} \frac{C_{ij}^k}{C_{i.}^k} C_{ii'}^l$
$\sum_v (\sum_u \mathbf{n}_{u.}^{kl} \mathbf{n}_{uv}^{kl})^2 = \sum_{i,i'} C_{i.}^k C_{i.}^l C_{ii'}^l$
$\sum_{u,v} (\mathbf{n}_{u.}^{kl})^2 (\mathbf{n}_{.v}^{kl})^2 = \sum_{i,i'} C_{i.}^k C_{i.}^l$
où $\mathbf{n}_{u.}^{kl} = \sum_v \mathbf{n}_{uv}^{kl}$ et $C_{i.}^k = \sum_{i'} C_{ii'}^k$

TABLE 1.4 – Formules de passage contingence/relationnel

Dans le cas de la formulation modifiée de Lerman, nous avons [NI00] :

$$\text{Lerman}'(C^k, C^l) = \frac{\sum_{i,i'} \left(C_{ii'}^k - \sum_{i,i'} \frac{C_{ii'}^k}{N^2} \right) \left(C_{ii'}^l - \sum_{i,i'} \frac{C_{ii'}^l}{N^2} \right)}{\sqrt{\sum_{i,i'} \left(C_{ii'}^k - \sum_{i,i'} \frac{C_{ii'}^k}{N^2} \right)^2 \sum_{i,i'} \left(C_{ii'}^l - \sum_{i,i'} \frac{C_{ii'}^l}{N^2} \right)^2}} \quad (1.35)$$

Propriété 1.1.5 *Expression relationnelle du critère du Chi-deux de Tchuprow*

$$\chi_T^2(C^k, C^l) = \frac{\sum_{i,i'} \left(\frac{C_{ii'}^k}{C_{i.}^k} - \frac{1}{N} \right) \left(\frac{C_{ii'}^l}{C_{i.}^l} - \frac{1}{N} \right)}{\sqrt{\sum_{i,i'} \left(\frac{C_{ii'}^k}{C_{i.}^k} - \frac{1}{N} \right)^2 \sum_{i,i'} \left(\frac{C_{ii'}^l}{C_{i.}^l} - \frac{1}{N} \right)^2}} \quad (1.36)$$

Propriété 1.1.6 *Expression relationnelle du critère de Jordan*

$$\text{Jordan}(C^k, C^l) = \frac{1}{N} \sum_{i,i'} \left(C_{ii'}^k - \frac{C_{i.}^k + C_{i.}^l}{2N} \right) C_{ii'}^l \quad (1.37)$$

Propriété 1.1.7 Expression relationnelle du critère de Light-Margolin

$$LM(C^k, C^l) = \sum_{i,i'} \left(\frac{C_{ii'}^k}{C_{i.}^k} - \frac{1}{N} \right) C_{ii'}^l \quad (1.38)$$

Propriété 1.1.8 Expression relationnelle du critère d'écart carré à l'indépendance

$$E(C^k, C^l) = \sum_{i,i'} \left(C_{ii'}^k - \frac{C_{i.}^k}{N^2} \right) C_{ii'}^l \quad (1.39)$$

Nous faisons remarquer la relation suivante :

$$Rand(C^k, C^l) = \frac{Condorcet(C^k, C^l)}{N^2} \quad (1.40)$$

Le codage relationnel a permis à A. Najah Idrissi d' "unifier" partiellement certains critères d'association qui s'expriment sous la forme d'un coefficient de corrélation entre matrices relationnelles [NI00]. C'est le cas des critères de Rand, de Janson et Vegelius, de Lerman et du Chi-deux de Tchuprow.

A ce propos, nous précisons que le résultat relatif à ce dernier critère a aussi été démontré par B. Mirkin dans [Mir01]. Dans cette approche, les matrices de comparaisons par paires correspondantes sont interprétées tels des opérateurs de projection.

De manière plus générale, nous citons les avantages du passage du codage contingentiel vers le codage relationnel :

- certains des critères déviennent linéaires en codage relationnel, on peut alors les employer dans le cadre de la modélisation par programmation linéaire
- tous les critères sont séparables en codage relationnel¹⁷. Autrement dit, celui-ci permet de séparer l'influence de la variable V^k de celle de la variable V^l
- les critères sous forme relationnelle peuvent être employés pour la mesure d'association entre graphes quelconques à condition que ceux-ci portent sur le même ensemble d'objets¹⁸. En effet, ces critères sont initialement étudiés pour des variables qualitatives que nous codons sous forme de relation d'équivalence. Cette dernière est un graphe qui possède des propriétés particulières. Ainsi, nous pouvons aussi étendre ces critères à des graphes non nécessairement structurés qui codent des relations quelconques¹⁹

Nous donnons un petit exemple pour illustrer ce troisième point. Considérons les deux graphes représentés sur la figure fig-1.1 représentant deux relations non structurées au sein d'un ensemble de 4 objets.

Les matrices relationnelles de ces deux graphes sont respectivement :

$$C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix} ; \quad C'' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

¹⁷. cet aspect est fondamental au niveau de l'agrégation des critères et de leur association maximale comme nous le verrons plus loin

¹⁸. ie même ensemble de noeuds

¹⁹. d'ailleurs, nous retrouvons certains de ces critères tels que Condorcet dans le cadre des graphes ou matrices relationnelles, codant des relations d'ordre cf équation (1.44). Le critère de Condorcet est, pour deux graphes quelconques, équivalent à leur distance de la différence symétrique à quelques modifications près

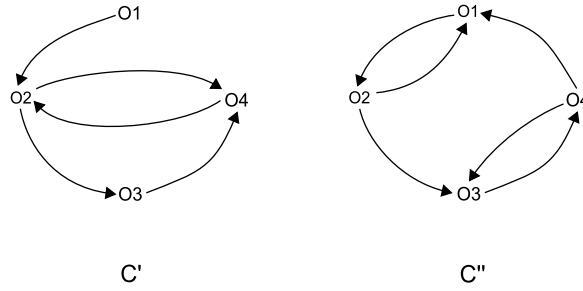


FIGURE 1.1 – Exemple de deux graphes quelconques

Nous pouvons mesurer l'association entre ces deux graphes en utilisant les critères définis précédemment dans leur formalisme relationnel. Par exemple, nous avons : $\text{Rand}(C', C'') = \frac{1}{N^2} \sum_{i,i'} (C'_{ii'} C''_{ii'} + \bar{C}'_{ii'} \bar{C}''_{ii'}) = \frac{3+8}{16} = \frac{11}{16}$.

De cet ensemble de résultats, il en ressort de plus, qu'il existe des liens étroits entre les approches logiques de l'AR et des concepts géométriques plus répandus en analyse des données. En effet, des critères qui s'expriment initialement en termes du tableau contingenciel, se trouvent être des coefficients de corrélation dans le formalisme relationnel.

Une autre contribution notable en statistiques des contingences de l'AR, exploite le fait que les expressions relationnelles des critères permettent de séparer les variables qualitatives : les deux matrices relationnelles C^k et C^l se factorisent. Ceci a permis de définir des critères de partitionnement issus de certains critères d'association à partir d'un modèle général, **le modèle d'association maximale**, développé par J.F. Marcotorchino et repris par G. Saporta :

$$\max_X \sum_{k=1}^M \Delta(V^k, X) \quad (1.41)$$

où $\Delta(V^k, X)$ est une mesure d'association ou de corrélation entre une variable V^k et une "structure" X .

Dans le cadre de l'AR, V^k est une variable qualitative représentée par sa matrice relationnelle C^k et X représente une matrice relationnelle d'équivalence. Dans ce cas, Δ est un critère d'association présenté précédemment et la solution recherchée est une partition [Mar86].

Dans le cadre de l'analyse en composantes principales, V^k est un vecteur relatif à une variable quantitative centrée-réduite et X est un vecteur appelé composante principale qui contient les coordonnées de la projection du nuage des points²⁰ sur un axe tel que l'inertie du nuage est maximale. Dans ce cas, Δ est le carré du coefficient de corrélation et l'objectif recherché est une représentation optimale d'un nuage de points dans un espace à dimensions réduites [Sap88].

Nous donnons dans tab-1.5 les critères de partitionnement obtenus pour des données qualitatives à partir de certains des critères présentés ci-dessus.

En ce qui concerne le critère de Light et Margolin, il est nécessaire de rappeler la propriété

20. ie les objets

Critère d'association	→ Critere d'association maximale
$\sum_k \text{Belson}(C^k, X)$	→ $\text{Belson}(\mathbf{C}, X) = \sum_{i,i'} \left(C_{ii'} - \frac{C_{i.} + C_{.i'}}{N} + \frac{C}{N^2} \right) X_{ii'}$
$\sum_k \text{Rand}(C^k, X)$	→ $\text{Rand}(\mathbf{C}, X) = \frac{2}{N^2} \sum_{i,i'} \left(C_{ii'} - \frac{1}{2} \left(\frac{C_{ii} + C_{i'i'}}{2} \right) \right) X_{ii'}$ $- \sum_{i,i'} \left(\frac{C_{ii} + C_{i'i'}}{2} - C_{ii'} \right)$
$\sum_k \text{Jordan}(C^k, X)$	→ $\text{Jordan}(\mathbf{C}, X) = \frac{1}{N} \sum_{i,i'} \left(C_{ii'} - \frac{C_{i.} + C_{.i'}}{2N} \right) X_{ii'}$
$\sum_k \text{LM}(C^k, X)$	→ $\text{LM}(\mathbf{C}, X) = \sum_{i,i'} \left(\hat{C}_{ii'} - \frac{M}{N} \right) X_{ii'}$
$\sum_k \text{E}(C^k, X)$	→ $\text{E}(\mathbf{C}, X) = \sum_{i,i'} \left(C_{ii'} - \frac{C}{N^2} \right) X_{ii'}$

TABLE 1.5 – Critères de partitionnement issus de l'association maximale

suivante²¹ qui permet d'obtenir le résultat présenté dans le tableau tab-1.5 :

Propriété 1.1.9

$$\hat{C}_{ii'} = \sum_k \frac{2C_{ii'}^k}{C_{i.}^k + C_{.i'}^k} \tag{1.42}$$

La matrice \hat{C} est aussi appelée matrice collective de Condorcet pondérée²² [Mar91a].

Il existe ainsi, en AR, de nombreux critères de partitionnement d'objets décrits par des variables qualitatives. Nous obtenons la partition optimisant un critère particulier en appliquant le modèle de programmation linéaire (1.11) pour celui-ci.

1.1.2 L'agrégation des préférences

L'agrégation des préférences concerne l'étude de l'agrégation de plusieurs relations d'ordre et de la détermination d'une relation d'ordre collective. Nous présentons ci-dessous les contributions de l'AR dans plusieurs aspects de ce domaine : la théorie du choix social et les critères d'association statistiques de deux relations d'ordre.

21. dans le cas des variables qualitatives, nous avons l'expression suivante en fonction du tableau disjonctif complet \mathbf{K} : $\hat{C}_{ii'} = \sum_k \frac{C_{ii'}^k}{C_{i.}^k} = \sum_j \frac{\mathbf{K}_{ij}\mathbf{K}_{i'j}}{\mathbf{K}_{.j}}$

22. qui est aussi valable pour des graphes quelconques

1.1.2.1 Théorie du choix social

Cette théorie rattachée notamment aux sciences économiques, cherche à définir des fonctions dites de choix social permettant d'agrèger plusieurs relations de préférences individuelles afin de déterminer une relation de préférence collective appelée également consensuelle. Cette problématique s'inscrit aussi dans le cadre de la théorie des votes : à partir des préférences prononcées par plusieurs votants sur plusieurs candidats, comment déterminer la préférence globale qui satisfait "au mieux" l'ensemble des préférences des votants ?

Il s'agit historiquement de la première application de l'AR. Elle a été initialement étudiée comme réponse à un article de E. Jacquet-Lagrèze de 1969 [JL69] par J.F. Marcotorchino et P. Michaud et réétudié ensuite axiomatiquement par P. Michaud en 1982 [Mic82]. Cet axe de recherche est issu de nombreuses questions posées autour du fameux théorème d'impossibilité d'Arrow [Arr74].

Ce théorème, stipule qu'il n'existe pas de fonction de choix social indiscutable²³ permettant de donner une préférence collective à partir de plusieurs préférences individuelles. En particulier, les règles majoritaires simples à l'instar de celle de Condorcet que nous rappelons ci-dessous, ne donnent pas de relation collective qui soit transitive ce qu'on appelle aussi effet Condorcet.

Définition 1.1.11 Règle majoritaire (par paire) de Condorcet et représentation relationnelle

Supposons qu'il y ait M votants. Un candidat O^i sera collectivement préféré ou indifférent à un candidat $O^{i'}$ si le nombre total de voix supportant " O^i préféré ou indifférent à $O^{i'}$ " est au moins égal à la majorité $\frac{M}{2}$ des voix²⁴.

Dans le formalisme relationnel, notons par C^k les matrices relationnelles individuelles représentatives des préférences exprimées par chaque votant V^k . Nous avons $\mathbf{C} = \sum_k C^k$ qui est la matrice relationnelle collective aussi appelée matrice de Condorcet collective.

Soit X la matrice relationnelle représentative de la relation obtenue en appliquant la règle majoritaire (par paire) de Condorcet. Nous avons alors :

$$X_{ii'} = \begin{cases} 1 & \text{si } C_{ii'} \geq \frac{M}{2} \\ 0 & \text{sinon} \end{cases} \quad (1.43)$$

La règle majoritaire de Condorcet telle qu'énoncée ci-dessus vérifie le théorème d'Arrow : dès qu'il y a plus de deux votants et plus de trois candidats, cette règle aboutit souvent à un effet Condorcet. Cet effet exprime une incohérence que M.J.A. Condorcet a été le premier à constater en 1785 [Con85]. Il indiquait que l'agrégation, selon cette règle majoritaire, de plusieurs relations transitives pouvait conduire à une relation collective qui est non transitive et pour laquelle il n'y a pas de vainqueur.

Dans [Mic82], [Mic85a] et [Mic85b], P. Michaud se fait le défenseur des travaux mal interprétés, voire ignorés, du Marquis de Condorcet dans ce domaine.

Tout d'abord, il indique qu'il existe une seconde règle d'agrégation méconnue de Condorcet qu'il nomme "règle de la majorité sous contraintes" et qui est la suivante : de toutes les opinions collectives possibles, c'est à dire celles représentant effectivement des classements, retenir pour choix collectif celle supportée par le plus grand nombre de voix.

Le problème de l'effet Condorcet est alors résolu dans la mesure où les seules opinions collectives évaluées sont celles correspondant à des classements qui sont par définition des relations transitives.

Puis P. Michaud, revisite les axiomes sur lesquels se fonde le théorème d'impossibilité d'Arrow. Il énonce les axiomes vérifiés par le critère de vote de Condorcet présenté page 5 et qui est en fait issu de la règle majoritaire (de Condorcet). P. Michaud conclut qu'il s'agit d'un très bon critère et que celui-ci vérifie de nombreux axiomes.

23. ie qui repose sur des axiomes raisonnables telle que la non-dictature

24. $\frac{M}{2}$ dans le cas d'un nombre pair de voix et $\lfloor \frac{M}{2} \rfloor + 1$ dans le cas d'un nombre impair de voix où $\lfloor a \rfloor$ est le plus grand entier relatif inférieur ou égal à a

Si nous prenons en compte la version contrainte de la règle majoritaire de Condorcet, il en ressort finalement que l'effet Condorcet présenté souvent comme étant le paradoxe de la "transitivité perdue", n'en n'est pas un : la propriété de transitivité que doit vérifier la relation consensuelle doit être explicitement imposée.

Ces travaux de J.F. Marcotorchino et P. Michaud ont permis de modéliser la règle de la majorité sous contraintes de Condorcet par programmation linéaire en nombres bivalents que nous rappelons ci-dessous et qui donne donc une solution au problème d'agrégation des préférences [MM79] :

$$\max_X \text{Condorcet}(\mathbf{C}, X) \quad \text{s.t.} \quad \begin{array}{ll} X_{ii'} \in \{0, 1\} & \text{(binarité)} \\ X_{ii'} + X_{i'i} \leq 1 & \forall i, i' = 1, \dots, N : i \neq i' \quad \text{(asymétrie)} \\ X_{ii'} + X_{i'i} \geq 1 & \forall i, i' = 1, \dots, N : i \neq i' \quad \text{(totalité)} \\ X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 & \forall i, i', i'' = 1, \dots, N \quad \text{(transitivité)} \end{array} \quad (1.44)$$

$\text{Condorcet}(\mathbf{C}, X)$ est le critère de Condorcet où \mathbf{C} est la matrice de Condorcet collective préalablement calculée. La relation binaire recherchée X étant une relation d'ordre strict et total, elle doit vérifier les conditions d'asymétrie, de totalité et de transitivité. A l'aide de ce programme linéaire, il est ainsi possible de déterminer parmi l'ensemble des $N!$ ordres stricts et totaux possibles celui qui supporte au mieux l'ensemble des préférences énoncées par un ensemble de M votants.

Remarque 1.1.1 Effet Condorcet

Le principe de l'effet de Condorcet se retrouve dans d'autres observations que celles issues de la théorie des votes. Par exemple, considérons plusieurs triangles rectangles ABC vérifiant chacun le théorème de Pythagore, $BC^2 = AB^2 + AC^2$, où BC est l'hypothénuse. Considérons un triangle moyen de ces triangles rectangles en calculant respectivement la longueur moyenne des côtés AB , AC et BC . Il apparaît alors que ce triangle moyen ne vérifie pas de façon générale le théorème de Pythagore.

1.1.2.2 Critères d'association entre relations d'ordre

Comme pour les relations d'équivalence, l'étude des relations d'ordre dans le cadre de l'AR trouve aussi des applications en statistiques [MM78].

Plus précisément, des travaux ont été entrepris sur les critères de mesure d'association entre deux relations d'ordre. Ces critères ont été définis en statistiques des rangs, domaine des statistiques non paramétriques, qui permet d'avoir des mesures de corrélation robustes entre variables quantitatives. Utiliser les rangs permet de se rendre indépendant des transformations monotones croissantes des variables quantitatives ce qui explique la robustesse accrue de ces critères en comparaison avec d'autres critères de mesure de corrélation comme le coefficient ρ de Bravais-Pearson.

Ces travaux ont été entrepris par S. Ghashghaie dans sa thèse [Gha90] où y sont données notamment, les formulations relationnelles des critères de Goodman et Kruskal [GK63], de Somers [Som62], de Kendall [Ken70], de Deuchler et de Kim [Kim71].

Soient deux variables ordinales²⁵ V^k et V^l . Notons par V_i^k la valeur prise par l'objet O^i pour la variable V^k . Nous avons alors les définitions suivantes :

$$\text{Conc}(V^k, V^l) = \# \{(i, i'), i < i' : \{V_i^k < V_{i'}^k \cap V_i^l < V_{i'}^l\} \cup \{V_i^k > V_{i'}^k \cap V_i^l > V_{i'}^l\}\} \quad (1.45)$$

$$\text{Disc}(V^k, V^l) = \# \{(i, i'), i < i' : \{V_i^k < V_{i'}^k \cap V_i^l > V_{i'}^l\} \cup \{V_i^k > V_{i'}^k \cap V_i^l < V_{i'}^l\}\} \quad (1.46)$$

$$\text{Tie}(V^k) = \# \{(i, i'), i < i' : \{V_i^k = V_{i'}^k\}\} \quad (1.47)$$

$$\text{Tie}(V^k, V^l) = \# \{(i, i'), i < i' : \{V_i^k = V_{i'}^k \cap V_i^l = V_{i'}^l\}\} \quad (1.48)$$

25. variables quantitatives ou variables qualitatives dont les modalités sont ordonnées

où $\#\mathbb{A}$ désigne le cardinal de l'ensemble \mathbb{A} .

$\text{Conc}(V^k, V^l)$ représente le nombre de paires concordantes; $\text{Disc}(V^k, V^l)$, le nombre de paires discordantes. $\text{Tie}(V^k)$ représente le nombre de paires dont les objets ont la même valeur pour V^k . $\text{Tie}(V^k, V^l) = \text{Tie}(V^k) \cap \text{Tie}(V^l)$ représente le nombre de paires dont les objets ont la même valeur à la fois pour V^k et V^l .

Plusieurs critères d'association de relations d'ordre ont été définis à partir des quantités ci-dessus.

Définition 1.1.12 Critère de Goodman et Kruskal²⁶

$$\text{GK}(V^k, V^l) = \frac{\text{Conc}(V^k, V^l) - \text{Disc}(V^k, V^l)}{\text{Conc}(V^k, V^l) + \text{Disc}(V^k, V^l)} \quad (1.49)$$

Définition 1.1.13 Critère asymétrique de Somers

$$\text{Somers}(V^k, V^l) = \frac{\text{Conc}(V^k, V^l) - \text{Disc}(V^k, V^l)}{\frac{N(N-1)}{2} - \text{Tie}(V^k)} \quad (1.50)$$

$$\text{Somers}(V^l, V^k) = \frac{\text{Conc}(V^k, V^l) - \text{Disc}(V^k, V^l)}{\frac{N(N-1)}{2} - \text{Tie}(V^l)} \quad (1.51)$$

Définition 1.1.14 Critère de Kendall²⁷

$$\text{Kendall}(V^k, V^l) = \frac{\text{Conc}(V^k, V^l) - \text{Disc}(V^k, V^l)}{\sqrt{\left(\frac{N(N-1)}{2} - \text{Tie}(V^k)\right) \left(\frac{N(N-1)}{2} - \text{Tie}(V^l)\right)}} \quad (1.52)$$

Définition 1.1.15 Critère de Deuchler

$$\text{Deuchler}(V^k, V^l) = \frac{\text{Conc}(V^k, V^l) - \text{Disc}(V^k, V^l)}{\frac{N(N-1)}{2} - \text{Tie}(V^k, V^l)} \quad (1.53)$$

Définition 1.1.16 Critère de Kim

$$\text{Kim}(V^k, V^l) = \frac{\text{Conc}(V^k, V^l) - \text{Disc}(V^k, V^l)}{\frac{N(N-1)}{2} - \frac{\text{Tie}(V^k) + \text{Tie}(V^l)}{2}} \quad (1.54)$$

S. Ghashghaie a établi des formules de passages qui expriment les quantités de concordances en termes relationnels. Ces formules sont rappelées dans le tableau tab-1.6.

Il est nécessaire de préciser, au préalable, que les matrices relationnelles codant des relations d'ordre impliquées dans le tableau tab-1.6 sont réflexives. Ainsi, un ordre strict sera étendu à un ordre large si bien que nous avons ici $C_{ii}^k = 1, \forall i = 1, \dots, N$ pour tout type d'ordre considéré. Il est très utile d'établir cette convention afin que les quantités de concordance et de discordance données dans ce tableau soient générales.

A l'aide de ces formules, on obtient l'expression relationnelle des critères d'association mentionnés, que nous donnons ci-dessous :

26. connu aussi sous la dénomination de γ de Goodman et Kruskal

27. connu aussi sous la dénomination du τ de Kendall

Concordance \leftrightarrow Relationnel
$\text{Conc}(C^k, C^l) = \sum_{i=1}^N \sum_{i'=1}^N (1 - C_{i'i}^k)(1 - C_{i'i}^l)$
$\text{Disc}(C^k, C^l) = \sum_{i,i'} (1 - C_{ii'}^k)(1 - C_{i'i}^l)$
$\text{Tie}(C^k) = C_{..}^k - \frac{N(N+1)}{2}$
$\text{Tie}(C^k, C^l) = \sum_{i,i'} (C_{ii'}^k + C_{i'i}^k)C_{ii'}^l - C_{..}^k - C_{..}^l + \frac{N}{2}$

TABLE 1.6 – Formules de passage “concordance” /relationnel de S. Ghashghaie

Propriété 1.1.10 Critère de Goodman et Kruskal en termes relationnels

$$\text{GK}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)C_{ii'}^l}{\sum_{i,i'} (2 - C_{ii'}^k - C_{i'i}^k)(1 - C_{ii'}^l)} \quad (1.55)$$

Propriété 1.1.11 Critère asymétrique de Somers en termes relationnels

$$\text{Somers}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)C_{ii'}^l}{N^2 - C_{..}^k} \quad (1.56)$$

$$\text{Somers}(C^l, C^k) = \frac{\sum_{i,i'} (C_{ii'}^l - C_{i'i}^l)C_{ii'}^k}{N^2 - C_{..}^l} \quad (1.57)$$

Propriété 1.1.12 Critère de Kendall en termes relationnels

$$\text{Kendall}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)C_{ii'}^l}{\sqrt{(N^2 - C_{..}^k)(N^2 - C_{..}^l)}} \quad (1.58)$$

Propriété 1.1.13 Critère de Deuchler en termes relationnels

$$\text{Deuchler}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)C_{ii'}^l}{N^2 - \sum_{i,i'} (C_{ii'}^k + C_{i'i}^k - 1)(C_{ii'}^l + C_{i'i}^l - 1)} \quad (1.59)$$

Propriété 1.1.14 Critère de Kim en termes relationnels

$$\text{Kim}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)C_{ii'}^l}{N^2 - \frac{1}{2}(C_{..}^k + C_{..}^l)} \quad (1.60)$$

Il ressort au terme de ce rappel sur les bases de l'AR, qu'il s'agit d'une théorie combinant plusieurs approches reposant sur des bases mathématiques provenant de la logique, des graphes, des structures algébriques basiques et de l'optimisation en programmation linéaire. L'AR permet de contribuer dans de nombreux domaines où interviennent des relations binaires structurées²⁸. Elle trouve ainsi de nombreuses applications en statistiques et en agrégation multicritère qui sont à première vue des domaines relativement éloignés. Or, **dans le cadre de l'AR, le**

28. ie relations vérifiant des propriétés relationnelles telles que les relations d'équivalence et les relations d'ordre

problème de classification automatique et le problème du choix social sont modélisés de la même manière, et ceci constitue un résultat majeur.

Dans la section suivante, nous introduirons un domaine qui peut, de la même manière, paraître assez éloigné des problèmes précédemment cités mais la réalité est toute autre. Il s'agit de la théorie des tresses. Nous nous intéresserons dans le cadre de ce mémoire de thèse à une représentation nouvelle des tresses qui est inspirée de l'AR.

1.2 Introduction succincte à la théorie des tresses et des noeuds

Dans ce mémoire, nous proposons une modélisation nouvelle des tresses et des noeuds. Ces objets mathématiques ont des propriétés qui rappellent celles des relations binaires. En particulier, le groupe des tresses appelé aussi groupe d'Artin, peut être vu comme une extension du groupe des permutations qui est aussi le groupe sous-jacent aux relations d'ordre strict et total. Nous faisons ici, quelques rappels simples sur ces objets mathématiques. De manière générale, une introduction "accessible" présentant, ces objets et leurs propriétés, pourra être trouvée dans [Sos99] et [Ada04].

Les tresses qui font l'objet d'études en topologie algébrique notamment, correspondent à ce que nous entendons dans la vie courante. Il s'agit d'un ensemble de brins qui se croisent les uns avec les autres. Ces brins ont toutefois une position initiale qui est "figée" de sorte que le tressage, c'est à dire les croisements progressifs d'un brin au-dessous ou au-dessus d'un autre, se fasse dans un sens unique : de haut en bas par exemple. Dans la figure fig-1.2, nous avons représenté une tresse géométrique à trois brins de couleurs distinctes : le brin bleu est initialement en position 1, le brin rouge en position 2 et le brin vert en position 3. Cette représentation graphique est aussi appelée diagramme d'une tresse .

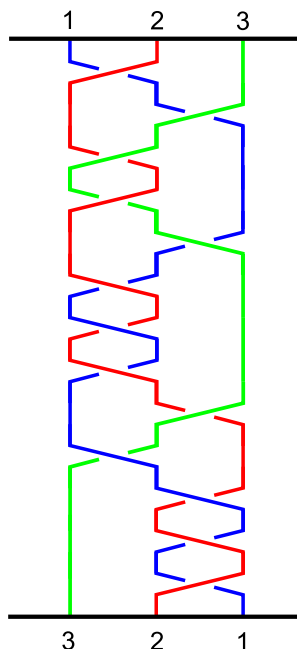


FIGURE 1.2 – Exemple d'une tresse à 3 brins

Nous pouvons facilement mettre en évidence le lien qu'entretiennent les tresses avec les permutations. Nous voyons en effet sur l'exemple précédent, qu'à l'arrivée, les brins ont permuté : le brin vert est à la place du brin bleu et *vice versa*. En fait, une permutation est équivalente à une tresse pour laquelle, nous ne faisons pas la distinction entre les informations topologiques "au-dessus" et "au-dessous". **Nous pouvons dire que contrairement aux permutations, c'est l'"histoire" des croisements élémentaires entre deux brins ainsi que l'information topologique "au-dessus" et "au-dessous" que ces croisements générateurs renferment, qui sont essentiels dans une tresse et non pas la permutation finale sur les brins que la tresse engendre.**

Nous faisons la distinction entre une tresse géométrique que nous notons symboliquement par la lettre grecque β et sa représentation algébrique que nous présentons ci-dessous et qu'on appelle mot de tresse et qui sera noté symboliquement par la lettre w . Les tresses peuvent effectivement être étudiées dans un cadre algébrique où elles définissent en fait, une extension du groupe des permutations.

Notons par \mathbb{B}_N l'ensemble des tresses à N brins dans le cadre de la représentation algébrique. Les générateurs de cet ensemble qui forment un groupe pour la loi notée \circ sont au nombre de $N - 1$ et sont les croisements élémentaires suivants :

Définition 1.2.1 Générateurs d'Artin

$$\sigma^j = \text{"le brin en position } j \text{ passe au-dessous du brin en position } j + 1\text{"} \quad (1.61)$$

$$(\sigma^j)^{-1} = \text{"le brin en position } j \text{ passe au-dessus du brin en position } j + 1\text{"} \quad (1.62)$$

La loi de composition des générateurs consiste graphiquement à juxtaposer progressivement les générateurs les uns au-dessous des autres²⁹. Nous voyons ainsi que la tresse β' de l'exemple représenté sur la figure fig-1.2 peut être représentée par le mot de tresse w' suivant :

$$w' = \sigma^1 \circ \sigma^2 \circ \sigma^1 \circ \sigma^1 \circ (\sigma^2)^{-1} \circ (\sigma^1)^{-1} \circ (\sigma^1)^{-1} \circ (\sigma^1)^{-1} \circ \sigma^2 \circ (\sigma^1)^{-1} \circ (\sigma^2)^{-1} \circ (\sigma^2)^{-1} \circ \sigma^2$$

Les générateurs d'Artin vérifient les relations suivantes que nous appelons respectivement $R1$, $R2$ et $R3$:

Définition 1.2.2 Relations d'Artin

$$\begin{aligned} R1 : \quad & \sigma^j \circ (\sigma^j)^{-1} \equiv (\sigma^j)^{-1} \circ \sigma^j \equiv \sigma^\epsilon \\ R2 : \quad & \sigma^j \circ \sigma^{j'} \equiv \sigma^{j'} \circ \sigma^j \quad \forall |j - j'| > 1 \\ R3 : \quad & \sigma^j \circ \sigma^{j+1} \circ \sigma^j \equiv \sigma^{j+1} \circ \sigma^j \circ \sigma^{j+1} \end{aligned} \quad (1.63)$$

où \equiv indique une équivalence topologique c'est à dire pour laquelle nous pouvons passer de l'argument à droite de \equiv vers l'argument à gauche par des déplacements de brins sans que ces derniers soient coupés, et *vice versa*.

Nous faisons remarquer que $R1$ est appelée réduction libre. Cette relation consiste à supprimer des croisements redondants comme il est montré dans la figure fig-1.3. Nous voyons que la recherche³⁰ de simplifications par le biais de cette relation permet en fait de démêler une tresse en supprimant petit à petit des croisements non significatifs.

29. dans le cas où notre sens de parcours est du haut vers le bas

30. en déplaçant des brins

Soit w un mot de tresse, élément du groupe \mathbb{B}_N . Nous noterons par \tilde{w} la tresse géométrique que représente le mot de tresse w . De ce fait, nous avons $\tilde{w}' = \beta'$. Toutefois, il existe pour une tresse β une infinité de mots de tresse³¹ w qui est la classe d'équivalence de \tilde{w} . Si deux mots de tresses w et w' tels que $w \neq w'$ représentent la même tresse géométrique β , nous dirons qu'ils sont topologiquement équivalents ce que nous noterons par $w \equiv w'$.

Nous illustrons les générateurs et les relations d'Artin dans la figure fig-1.3.

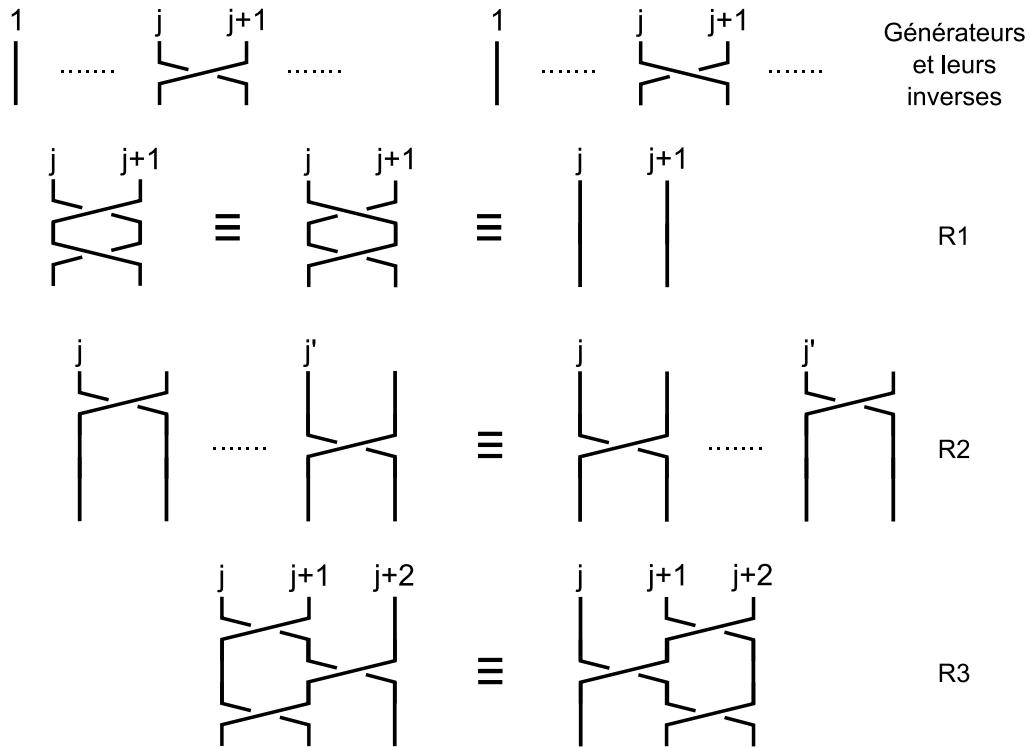


FIGURE 1.3 – Relations d'Artin du groupe des tresses

L'inverse d'un mot de tresse w , noté w^{-1} est donné par l'inverse des générateurs de w pris dans l'ordre opposé. Nous avons par exemple pour la figure fig-1.2 :

$$(w')^{-1} = (\sigma^2)^{-1} \circ \sigma^2 \circ \sigma^2 \circ \sigma^1 \circ (\sigma^2)^{-1} \circ \sigma^1 \circ \sigma^1 \circ \sigma^1 \circ \sigma^2 \circ (\sigma^1)^{-1} \circ (\sigma^1)^{-1} \circ (\sigma^2)^{-1} \circ (\sigma^1)^{-1}$$

Par ailleurs, la longueur d'un mot de tresse w , notée $|w|$, est le nombre de générateurs composant ce mot de tresse. La largeur d'un mot de tresse est quant à elle la quantité $j^+ - j^- + 2$ où j^+ et j^- sont respectivement l'indice maximal et l'indice minimal de l'ensemble des générateurs composant ce mot de tresse. La largeur d'un mot de tresse donne le nombre de brins appartenant au domaine non trivial de w .

Si nous prenons pour exemple le mot de tresse w' , celui-ci a une longueur égale à 13 et une largeur égale³² à 3.

Nous dirons qu'un mot de tresse est librement réduit (au sens de la relation R1) s'il ne présente pas de motif $\sigma^j \circ (\sigma^j)^{-1}$ ou $(\sigma^j)^{-1} \circ \sigma^j$.

Souvent, on note les générateurs par des lettres de l'alphabet en employant la correspondance suivante : $a = \sigma^1, A = (\sigma^1)^{-1}, b = \sigma^2, B = (\sigma^2)^{-1} \dots$ Par ailleurs, l'opération de compo-

31. la tresse triviale notée β^e qui ne présente aucun croisement de brins dans sa forme simplifiée est par exemple représentée par les mots de tresses suivants $(\sigma^j \circ (\sigma^j)^{-1})^k \quad k \leq 0$

32. en effet, l'indice maximal est $j^+ = 2$ et l'indice minimal est $j^- = 1$ ce qui nous donne $j^+ - j^- + 2 = 3$

tion \circ est omis. Le mot de tresse w' de β' est alors le suivant :

$$\beta' = abaaBAAAbABBb$$

Ce sont initialement les travaux d'E. Artin [Art47] qui ont permis de modéliser les tresses sous la forme d'un groupe. Depuis, les tresses suggèrent de nombreux problèmes d'ordre algébrique et algorithmique dont les suivants sont les plus fréquemment rencontrés :

- le problème d'isotopie appelé aussi le problème des mots de tresses qui consiste à déterminer si un premier mot de tresse est topologiquement équivalent à un deuxième mot de tresse
- le problème de conjugaison entre deux mots de tresse w' et w'' qui consiste à déterminer une tresse w''' telle que $w' = w''' \circ w'' \circ (w''')^{-1}$
- le problème de la tresse minimale qui consiste à déterminer le mot de tresse w^* qui est équivalent à w' mais dont le nombre de générateurs est minimal

L'aspect combinatoire de ces objets mathématiques font qu'ils trouvent des applications en cryptographie [Deh04a], [Mah]. Par ailleurs, les tresses sont intimement liées aux noeuds. De ce fait, elles trouvent aussi des applications en biologie pour la modélisation de la molécule d'ADN [Sos99], et en physique tels que certains modèles en physique statistique [Jon05], [Kau05].

Cette mise en relation fondamentale entre tresses et noeuds est due à J.W. Alexander [Ale28]. Ce dernier montra que lorsque nous relient, sur le diagramme d'une tresse, les brins au départ aux brins d'arrivée, nous obtenons un noeud ou un entrelac³³. Cette opération est appelée clotûre d'une tresse et ainsi, de manière réciproque, tout noeud peut être obtenu à partir de la clotûre d'une tresse. Nous noterons par $\hat{\beta}$ le noeud obtenu à partir de la clotûre de la tresse β .

Nous illustrons dans la figure fig-1.4, l'opération de clotûre du mot de tresse aaa qui donne le noeud de trèfle.

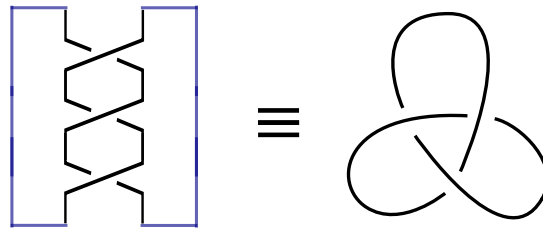


FIGURE 1.4 – Noeud de trèfle

Les noeuds sont des objets qui fascinent notamment mathématiciens et physiciens. Leur étude à permis par exemple à V.F.R. Jones d'obtenir une médaille Fields en 1990 pour sa contribution à la définition d'un invariant complet des noeuds et de la mise en évidence de liens insoupçonnés entre certains modèles physiques et les noeuds [Jon05], [Kau88]. Le problème de classification des noeuds est l'un des problèmes majeurs de cette théorie de nombreux

^{33.} ie un ensemble de noeuds possiblement liés les uns aux autres

chercheurs ont étudié. Il consiste à déterminer une méthode ou un calcul permettant de distinguer les noeuds à partir de leur diagramme. A l'instar du problème d'isotopie des tresses, un même noeud peut être représenté par différents diagrammes en utilisant des mouvements isotopiques de brins : comment savoir alors si deux diagrammes représentent en fait un seul et même noeud ? Si ce problème est aujourd'hui résolu efficacement pour les tresses, il reste davantage difficile pour les noeuds.

Le polynôme de Jones, comme d'autres polynômes tel que le polynôme d'Alexander, est fondé sur la représentation des noeuds sous forme de tresses. L'approche de V.F.R. Jones utilise une représentation matricielle particulière des générateurs des tresses et, en utilisant de plus les propriétés d'un opérateur de type trace, il définit un polynôme caractéristique du noeud. Cet opérateur trace possède en fait des caractéristiques qui permettent de tenir compte des mouvements de Markov [MK99]. Nous présentons ces mouvements et le théorème qui en découle qui précisent davantage les liens entre tresses et noeuds.

Définition 1.2.3 *Mouvements de Markov*

Les mouvements de Markov sont de deux types et seront notés respectivement $M1$ et $M2$:

- $M1$ consiste à remplacer un mot de tresse de \mathbb{B}_N , w , par sa conjuguée $w' \circ w \circ (w')^{-1}$ où w' est un mot représentatif d'une tresse à N brins quelconque. En d'autres termes, nous avons $w \equiv w' \circ w \circ (w')^{-1}$ où \equiv représente ici l'équivalence topologique entre deux mots de tresse cloturés définissant un même noeud
- $M2$ consiste à remplacer un mot de tresse de \mathbb{B}_N , w , par un mot de tresse de \mathbb{B}_{N+1} , $w \circ \sigma^N$ ou $w \circ (\sigma^N)^{-1}$ où w est alors vu tel un mot de tresse du groupe \mathbb{B}_{N+1} et où σ^N est le N -ème générateur d'Artin de \mathbb{B}_{N+1} . Autrement dit nous avons : $w \equiv w \circ (\sigma^N)^{\pm 1}$

Théorème 1.2.1 *Théorème de Markov*

Soient β et β' deux tresses qui n'ont pas nécessairement le même nombre de brins. Alors les clotures respectives de β et de β' représentent le même noeud ou le même entrelac, si et seulement si β peut être déformé jusqu'à obtenir β' en appliquant un nombre fini de mouvements de Markov $M1$ et / ou $M2$ ou leurs inverses.

Nous illustrons ces deux mouvements dans la figure fig-1.5 ainsi qu'une représentation graphique de l'équivalence topologique.

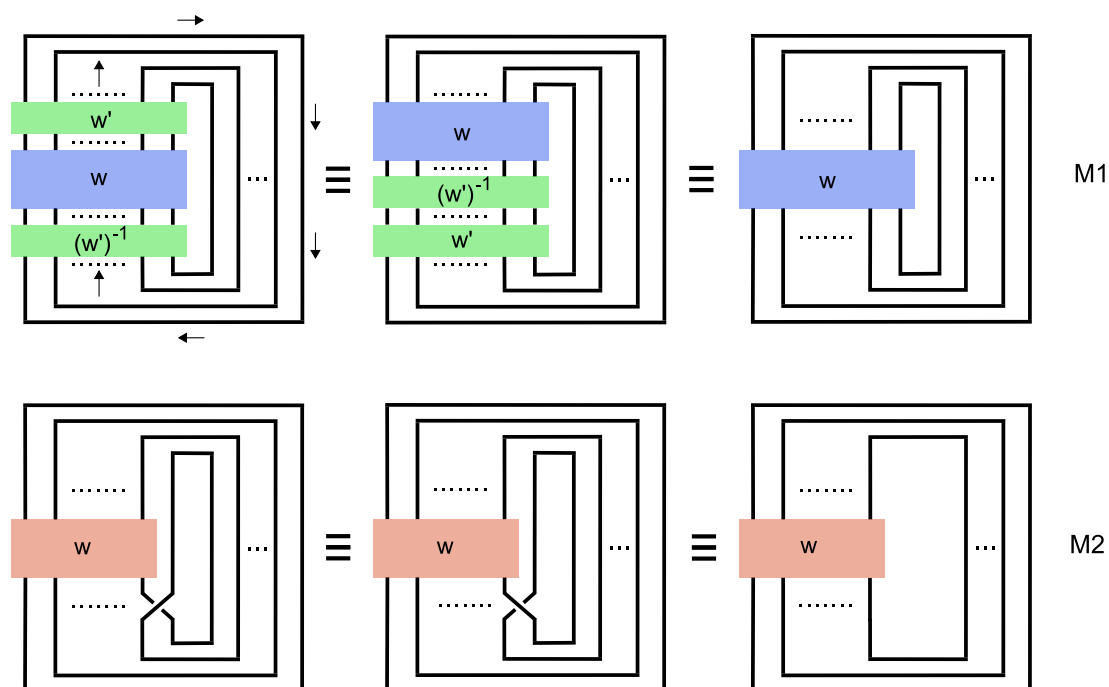


FIGURE 1.5 – Les mouvements de Markov

1.3 Contributions

L'introduction précédente avait pour objectif de rappeler les différents domaines que nous allons aborder dans ce mémoire de thèse.

Ce travail est composé de deux grandes parties : l'étude des relations d'équivalence et l'étude des relations d'ordre. Chacune de ces deux types de relations binaires structurées ont engendré des axes de recherches et des domaines d'applications distincts que l'AR permet, dans une certaine mesure, de traiter dans un même cadre.

En ce qui concerne l'étude des **relations d'équivalence**, nous allons étudier les points suivants :

- le chapitre 2 aborde le **problème de la classification automatique sous un angle algébrique**. Dans un premier temps, nous rappelons différents types de codage de ces relations : la manière de représenter des objets mathématiques joue un rôle important car chaque codage ne possède pas en effet, les mêmes propriétés. Nous nous intéressons ensuite à la structure sous forme de treillis des partitions. Puis, en nous inspirant du groupe des permutations, et plus particulièrement de la représentation géométrique sous forme de graphe de Cayley de celui-ci, nous définissons des opérations dites de transferts permettant de transformer la structure des partitions. L'ensemble des partitions d'un ensemble d'objets associé à l'ensemble des opérations de transferts permet de définir un automate à états finis donnant une structure algébrique particulière aux relations d'équivalence. Dès lors, **nous montrons dans la section 2.2, que certaines heuristiques de classification automatique miment différents types de chemins dans cet automate, c'est le cas notamment des algorithmes de centres mobiles et de classifications hiérarchiques ascendantes. Cette approche "graphique", nous permet de mettre en évidence les limites de ces algorithmes : ils permettent d'agréger des classes mais pas de les scinder en plusieurs. A l'inverse, d'autres algorithmes permettent de diviser**

successivement des classes en deux mais ne permettent pas de les agréger, il s'agit des algorithmes hiérarchiques descendants par exemple. *Nous proposons un nouvel algorithme de classification qui constitue à notre connaissance, la meilleure heuristique cherchant à résoudre le modèle en programmation linéaire (1.11) issu de l'AR pour de grandes bases de données avec variables qualitatives. Cet algorithme utilise des opérations plus évoluées permettant soit d'agréger soit de diviser des classes ce qui permet d'explorer de manière plus efficace et plus exhaustive le graphe des partitions d'un ensemble d'objets.*

- dans le chapitre 3, nous cherchons à **contribuer aux travaux engendrés par l'AR autour des mesures de similarité entre objets et des mesures d'association entre variables dans une perspective d'unification**. Les précédents travaux ont montré qu'il existait des liens étroits entre les mesures d'ordre géométrique définies en analyse de données en général et les mesures d'ordre logique³⁴ définies en AR. Tout au long de ce chapitre, nous cherchons à exploiter cette dualité "géométrie / logique" pour unifier traitements de données qualitatives et traitements de données quantitatives.
 - Ainsi, nous cherchons tout d'abord à étendre les indices de similarité d'ordre logique définis pour des vecteurs d'objets décrits par des données binaires à des vecteurs d'objets décrits par des données non binaires (quantitatives ou autres). Nous définissons lors de la section 3.1, une classe d'indices fondée sur 2 paramètres géométriques, le cosinus de l'angle et le rapport des normes de deux vecteurs. **Cette classe dite des indices d'ordre t généralise les indices de Dice et d'Ochiaï et est valable pour des données de présence / absence ou quantitatives.**
 - Après avoir traité la mesure de similarité entre vecteurs d'objets, nous nous intéressons à la mesure d'association entre variables qui sont des vecteurs duaux aux vecteurs d'objets. D'une part dans la section 3.2, nous étendons les critères d'association maximale définis dans le cas des variables qualitatives aux cas des variables quantitatives. Nous poursuivons notre interprétation géométrique en montrant que ces critères sont intimement liés aux critères inertiels : dans le cas de données centrées, **la grande majorité des critères d'association maximale, qu'ils soient définis sur des données qualitatives ou quantitatives, sont à une constante multiplicative près, le critère de la différence inertielle.**
 - D'autre part dans la section 3.3, nous revenons sur la notion d'indépendance entre variables qualitatives. Nous enrichissons les résultats d'unification de A. Najah Idrissi [NI00], en montrant que d'autres critères d'association comme celui de Belson et celui de Jordan s'expriment aussi dans le codage relationnel tel une covariance. Nous étudions aussi la **notion d'indétermination** qui est une indépendance d'ordre logique étudiée par J.F. Marcotorchino [Mar84a], [Mar84b]. **Nous clarifions certains aspects de ce concept et l'étendons à la notion d'indétermination pondérée.** Finalement, nous définissons deux classes d'indices de mesure du lien entre deux variables qualitatives : les indices de corrélation et les indices d'indétermination. Ces deux classes définissent des indices qui sont similaires aux indices de similarité d'ordre t mais pour des vecteurs "matrices relationnelles" de variables qualitatives.
- le chapitre 4 est davantage combinatoire que les précédents.
 - Nous commençons dans la section 4.1, par introduire les matrices relationnelles comme étant une algèbre de Boole particulière. Nous rappelons d'ailleurs, d'autres opérations que l'union et l'intersection booléenne qui sont issues de l'algèbre relationnelle de C.S. Peirce : l'addition et la multiplication relative entre deux relations ainsi que la converse d'une relation. Ces considérations sont importantes pour la suite. En effet, l'intersection entre deux matrices relationnelles permet de définir les enveloppes de Condorcet et les opérations peirciennes joueront un rôle majeur dans nos contributions du chapitre 6. Nous introduisons succinctement par ailleurs les "t-norms" qui sont des opérations de conjonction étendues à des logiques multivalentes. En particulier, **nous montrons les fondements logiques des contraintes linéaires des propriétés relationnelles** que nous avons rappelées dans le tableau tab-1.2. En effet, ces dernières peuvent toutes être obtenues en utilisant les conjonctions et disjonctions de Lukasiewicz [KMP00]. Par ailleurs, cette démarche nous permet de dégager des

34. ie de comptage

contraintes linéaires de diverses autres propriétés relationnelles.

- Dans la section 4.2, **nous définissons une méthode combinatoire d'agrégation de données non binaires dont les interprétations sont en terme de règle majoritaire : nous calculons la mesure consensuelle à une majorité m donnée parmi M variables, pour que deux objets soient dans une même classe.** La matrice de comparaisons par paires des données consensuelles est appelée enveloppe de Condorcet. Nous montrons que celle-ci s'exprime sous la forme d'une combinaison linéaire où les coefficients sont issus de l'inversion du triangle de Pascal. Nous aboutissons à une formulation qui est en fait de même type que les formules combinatoires de H. Poincaré et de Ch. Jordan. Toutefois, nous enrichissons l'étude des coefficients obtenus en montrant leurs liens avec les nombres de Lah et les coefficients binomiaux négatifs. De plus, nous montrons une propriété de décomposition des enveloppes de Condorcet selon la majorité m . **Nous proposons cette méthode pour la classification automatique de données hétérogènes.**

Nous développerons les points suivants dans le cadre des **relations d'ordre** :

- de manière symétrique aux relations d'équivalence, nous étudions les aspects algébriques des relations d'ordre strict et total. Celles-ci sont des classements sans *ex-aequo* et sans données manquantes sur un ensemble d'objets. Nous utilisons le même formalisme que celui des relations d'équivalence, ainsi un ordre strict et total est représenté par une relation d'affectation d'un ensemble d'objets à un ensemble de rangs. Les opérations qui agissent sur ces relations sont des permutations qui sont définies soit sur les objets soit sur les rangs. **Nous aboutissons à une structure sous forme d'automate qui est semblable à celle des relations d'équivalence : l'objectif est de montrer le parallélisme entre le traitement de différents types de relations binaires structurées.** Toutefois les propriétés particulières des ordres stricts et totaux engendrent des propriétés algébriques plus "fortes" que celles des partitions. En effet, il existe une bijection entre l'ensemble des permutations d'un ensemble d'objets et l'ensemble de ses classements. Ces propriétés, nous permettent de définir un **nouveau type de codage basé sur des comparaisons par paires appelés matrices de paires d'objets nommés** et qui sera à l'origine de notre **modélisation relationnelle des tresses.**
- le chapitre 6 traite des ordres larges et non nécessairement totaux qui sont donc plus laxistes que les ordres stricts et totaux.
 - dans la section 6.1, la problématique étudiée est en continuité avec celle de la section 3.2. Nous redéfinissons l'ordre dual d'un ordre par celui donné par l'opération unaire de converse de l'algèbre relationnelle et non son complémentaire comme il est coutume de faire en AR. Cette opération se traduit au niveau des matrices relationnelles par l'opération transposée. Nous donnons aussi de nouvelles formules de passage entre espace de contingence, à partir duquel sont définis les critères d'association, et espace relationnel. Ces formules font jouer aux matrices relationnelles des rôles symétriques. **En prenant en considération ces nouvelles formules de passages ainsi que la converse d'une relation d'ordre comme étant l'ordre dual naturel d'un ordre quelconque, nous aboutissons à divers résultats d'unification : le concept d'indétermination étudiée en section 3.3 dans le cas des relations d'équivalence, correspond au concept d'écart entre concordance et discordance utilisé en statistiques non paramétriques; de plus, les critères de Kendall et de Kim s'interprètent comme étant des cas particuliers des indices d'ordre t d'écart à l'indétermination présentés dans le cadre des relations d'équivalence étudiés dans cette même section 3.3.** Il ressort donc une certaine unification des mesures d'association entre relation d'équivalence et relations d'ordres par l'intermédiaire du codage relationnel.
 - dans la section 6.2, nous abordons une thématique issue des sciences économiques mais qui peut être appréhendée par l'AR. Le problème étudié est l'agrégation des relations de préférence qui sont en fait des relations d'ordre. Un résultat important est le théorème d'impossibilité d'Arrow qui met en évidence l'impossibilité d'avoir une règle majoritaire raisonnable qui puisse donner un ordre collectif à partir de plusieurs ordres individuels. En particulier, la règle de la majorité simple aboutit très souvent à une relation collective qui est cyclique et pour laquelle, nous ne pouvons détermi-

ner un vainqueur : il s'agit de l'effet Condorcet. De nombreux travaux ont consisté à déterminer des conditions restrictives sur l'ensemble des préférences individuelles initiales afin d'éviter l'effet Condorcet au niveau du consensus final. Les conditions suffisantes de Black, d'Inada ou de Sen en sont des résultats importants. **Nous abordons cet aspect du problème dans un cadre plus général : en raisonnant à partir des comparaisons par paires, en utilisant les opérations de l'algèbre relationnelle et en nous plaçant dans un dioïde particulier, nous donnons des conditions de nature algébrique sur les matrices relationnelles collectives qui englobent celles citées précédemment dans la mesure où elles sont nécessaires et suffisantes et qu'elles concernent plusieurs règles majoritaires.** Nous terminons cette section en montrant que les règles d'agrégation définissent des critères de vote qui sont en lien avec les critères d'association entre relation d'ordre par l'intermédiaire du principe d'agrégation maximale. Nous donnons une version "contraintes" de ces règles dont les solutions sont obtenues par programmation linéaire en nombres bivalents.

- le chapitre 7 enfin, aborde les tresses et certains problèmes algorithmiques autour du groupe des tresses. **Nous définissons une modélisation nouvelle des tresses basée sur des matrices de paires de brins.** Celle-ci est une extension du codage présentée lors de l'étude des ordres stricts et totaux du chapitre 5. **Nous cherchons ensuite, à l'aide de ce codage, à modéliser les déplacements isotopiques des brins d'une tresse : ces déplacements peuvent être vus comme une généralisation de la réduction de poignée d'une tresse définie par P. Dehornoy.** Fort de cette généralisation, nous proposons une nouvelle modélisation des problèmes algorithmiques qui s'apparentent alors davantage à des problèmes d'optimisation combinatoire.

Nous nous sommes appliqués à donner une cohérence globale à ce mémoire, malgré la diversité des domaines abordés. Nous souhaitons mettre l'accent sur les différents points suivants afin d'illustrer notre effort d'"unification" tant sur la forme que sur le fond :

- nous avons cherché à présenter des aspects algébriques des relations d'équivalence, relations d'ordre, et "relations" de tresse selon le même formalisme d'automate à états finis selon une définition de ces derniers se rapprochant davantage des groupes automatiques définis par W.P. Thurston notamment dans [ES00]. Par ailleurs, les problèmes auxquels nous nous intéressons pour chacun de ces types de relations relèvent tous de l'optimisation combinatoire
- les mesures de similarité que nous avons définies sont valables à la fois pour des vecteurs d'objets décrits par des variables qualitatives et des variables quantitatives
- les mesures d'association basées sur le concept d'indétermination³⁵ entre relations d'équivalence d'une part et relations d'ordre d'autre part, coïncident par l'intermédiaire du codage relationnel. Pour ce faire, l'inverse d'une relation d'équivalence est donnée par sa relation complémentaire alors que nous définissons l'inverse d'une relation d'ordre par sa relation converse (opérateur unaire de l'algèbre relationnel)
- la méthode d'agrégation de données non binaires ou floues que nous proposons peut être employée pour la modélisation de similarité entre objets décrits à la fois par des variables qualitatives et quantitatives. Par ailleurs cette même méthode d'agrégation a des liens avec les règles majoritaires qui définissent les critères de vote en théorie du choix social

Ci-dessous nous donnons un schéma montrant les liens qu'entretiennent les chapitres entre eux.

Dans le corps du texte, **nous mettons en gras les passages qui expliquent plus particulièrement la démarche que nous poursuivons ou les objectifs que nous souhaitons atteindre** dans chaque paragraphe.

35. écart entre le nombre de concordance et le nombre de discordance

Les résultats présentant un caractère relativement nouveau seront **encadrés**.

Parmi ces derniers, certains sont volontairement mis davantage en valeur. Ceux-ci seront alors **encadrés deux fois**.

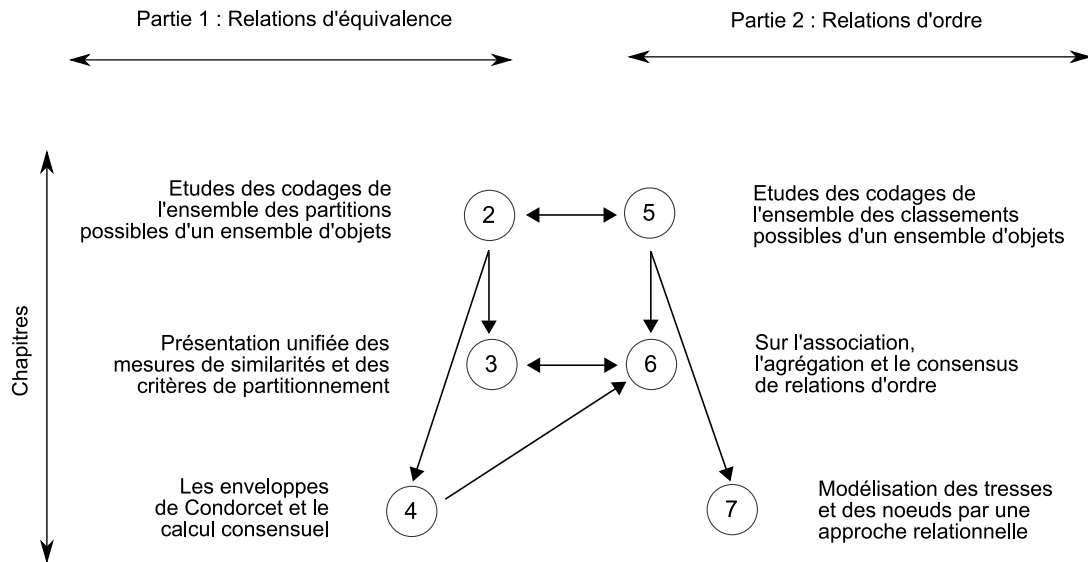


FIGURE 1.6 – Relations entre les différents chapitres

Première partie

Aspects algébriques et combinatoires de l'AR des relations d'équivalence et applications en classification automatique

CHAPITRE 2

Etude des codages de l'ensemble des partitions possibles d'un ensemble d'objets

Sommaire

2.1 Etude de l'ensemble des partitions d'un ensemble d'objets	32
2.1.1 Définitions et rappels	32
2.1.1.1 Les différents formalismes matriciels des relations d'équivalence et les formules de passage correspondantes	33
2.1.1.1.1 Le codage algébrique	33
2.1.1.1.2 Le codage par affectation	34
2.1.1.1.3 Le codage relationnel	36
2.1.1.1.4 Lien entre le codage par affectation et le codage relationnel	37
2.1.1.2 Le treillis des partitions	38
2.1.2 Opérations de transferts et structures algébriques associées	39
2.1.2.1 Le monoïde des opérations de transferts	39
2.1.2.2 Le sous-monoïde des opérations de transferts simples	42
2.1.2.3 Le groupe des opérations de transferts doubles	43
2.1.3 Action des opérations de transferts sur les partitions	44
2.1.3.1 Action des opérations de transferts sur les matrices disjointives	45
2.1.3.1.1 Action sur les objets	45
2.1.3.1.2 Action sur les classes	46
2.1.3.1.3 Tableau récapitulatif	47
2.1.3.2 Action des opérations de transferts sur les matrices relationnelles	47
2.1.3.2.1 Action sur les objets	47
2.1.3.2.2 Action sur les classes	48
2.1.3.2.3 Dualité "Affectation / Relationnel"	49
2.2 Les méthodes de classification automatique vues sous un angle algébrique	50
2.2.1 Graphe orienté de l'ensemble des partitions d'un ensemble d'objet	51
2.2.2 Les heuristiques classiques de classification automatique et leur interprétation au vu du graphe orienté des partitions	54

2.2.2.1	Les méthodes de type centres mobiles	55
2.2.2.2	Les méthodes hiérarchiques	56
2.2.2.3	Les méthodes inspirées de l'AR et de la similarité régularisée	57
2.2.3	Les nouvelles heuristiques de type centres mobiles et l'opération de "cas- sage" des classes	59
2.2.4	Une nouvelle heuristique de l'application de l'AR en classification auto- matique	61

Dans ce chapitre, nous étudions l'ensemble des relations d'équivalence possibles d'un ensemble d'éléments \mathbb{O} que nous notons par $\mathbb{RE}(\mathbb{O})$. La définition d'une relation d'équivalence en termes des propriétés relationnelles, correspond à la définition d'une partition qui est en fait une interprétation ensembliste de celle-ci. Ainsi, par la suite, nous emploierons indifféremment ces deux termes.

Après avoir présenté différents types de représentation des partitions, nous définirons des opérations dites de transfert dont l'action sur une partition modifie sa structure. Ces opérations permettent de définir un graphe orienté sur l'ensemble des partitions de \mathbb{O} . Une partition \mathcal{R}^k pointerait vers une partition \mathcal{R}^l s'il existe une opération de transfert permettant de transformer la partition \mathcal{R}^k en la partition \mathcal{R}^l .

L'objectif de ce chapitre est dans un premier temps de rappeler l'intérêt de pouvoir représenter les relations d'équivalence selon plusieurs formalismes. En effet, chaque codage possède des propriétés particulières.

Dans un second temps, nous cherchons à définir des groupes d'action sur des relations d'équivalence pouvant soit agir sur les classes, soit sur les objets. Cette démarche nous permet de donner finalement une approche algébrique aux heuristiques classiques de la classification automatique tels que les centres mobiles et les classifications hiérarchiques ascendantes. Nous montrons que chaque méthode correspond à un type particulier de combinaisons d'actions. Ceci nous permet par ailleurs de montrer les limites théoriques de ces algorithmes.

Fort de ces études, nous présentons un nouvel algorithme de classification automatique pour variables qualitatives qui permet de dépasser les limites identifiées. Celui-ci cherche à approcher au mieux la solution donnée par le modèle de programmation linéaire issu de l'AR et défini par (1.11) page 7. D'un point de vue théorique, il permet de mettre en oeuvre des opérations qui permettent d'agréger ou de casser des classes ce qui, contrairement aux autres heuristiques classiques, permet une meilleure exploration du graphe des partitions.

2.1 Etude de l'ensemble des partitions d'un ensemble d'objets

Il existe plusieurs manières de représenter une partition. Nous rappelons différents codages, leurs propriétés particulières ainsi que les liens qu'ils entretiennent entre eux. Nous définissons ensuite des opérations qui permettent de modifier les classes d'une partition en réaffectant des objets dans des classes distinctes de leurs classes d'origine. Nous montrons comment ces opérations agissent sur chacun des codages que nous présentons.

2.1.1 Définitions et rappels

Les définitions et rappels concernent l'ensemble des partitions possibles d'un ensemble d'objets. Nous noterons par \mathbb{D} l'ensemble des classes ou des descripteurs auxquels les objets

de \mathbb{O} sont affectés. Le nombre de classes est propre à chaque relation d'équivalence considérée, toutefois, il ne peut excéder N qui est le nombre total d'objets.

Remarque 2.1.1 Signification des éléments de l'ensemble des descripteurs \mathbb{D}

Dans le contexte de l'analyse de données, les différentes relations d'équivalence peuvent être représentées par des variables qualitatives V^k dont les modalités $D^j \in V^k$, constituent les classes ou descripteurs.

Les modalités sont supposées non ordonnées : même si une variable possède des modalités ordonnées¹, ce qui nous intéresse c'est l'appartenance des objets à ces différentes modalités et non pas l'ordre que cette appartenance induit sur les objets.

Le cas où nous nous intéressons à l'ordre induit par une variable ordonnée correspond au traitement des relations d'ordre large et sera étudié au chapitre 6.

2.1.1.1 Les différents formalismes matriciels des relations d'équivalence et les formules de passage correspondantes

Définition 2.1.1 Relation d'équivalence ou partition

Une relation d'équivalence, \mathcal{R} , d'un ensemble $\mathbb{O} = \{O^i, i = 1, \dots, N\}$, est une relation binaire sur cet ensemble d'éléments qui respecte les propriétés suivantes :

- *Réflexivité : $O^i \mathcal{R} O^i \quad \forall O^i \in \mathbb{O}$*
- *Symétrie : $O^i \mathcal{R} O^{i'} \Rightarrow O^{i'} \mathcal{R} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O}$*
- *Transitivité : $O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''} \quad \forall (O^i, O^{i'}, O^{i''}) \in \mathbb{O} \times \mathbb{O} \times \mathbb{O}$*

En d'autres termes, une partition est une famille de classes disjointes des éléments de l'ensemble considéré. Le nombre de partitions possibles d'un ensemble de N objets est donné par le nombre de Bell que nous avons rappelé en introduction page 6.

Il existe plusieurs manières de représenter une relation d'équivalence. Nous présenterons les trois types de codage suivant :

- Codage algébrique
- Codage par affectation
- Codage relationnel

Afin d'illustrer la définition de chaque type de codage, nous prendrons un exemple de 4 objets. La partition que nous noterons par \mathcal{R}' , sera constituée de trois classes : la première rassemble les objets O^1 et O^2 , la seconde et la troisième sont constituées respectivement des objets O^3 et O^4 . En termes ensemblistes, nous notons cette partition par : $\mathcal{R}' = \{\{O^1, O^2\}; \{O^3\}; \{O^4\}\}$.

2.1.1.1.1 Le codage algébrique

Le codage algébrique consiste simplement à préciser pour chaque objet la classe à laquelle il appartient.

Définition 2.1.2 Codage algébrique

Soit une relation d'équivalence \mathcal{R} de N objets constituée de p classes notées : $D^j, j = 1, \dots, p$ avec

1. exemple $V^k =$ taille et $D^1 =$ petit, $D^2 =$ moyen, $D^3 =$ grand

$p \leq N$. Le codage algébrique représentant \mathcal{R} est noté κ et est défini selon le formalisme suivant :

$$\kappa = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ \kappa(O^1) & \kappa(O^2) & \dots & \kappa(O^N) \end{pmatrix} \quad (2.1.1)$$

où $\kappa(O^i) \in \{D^1, \dots, D^p\}$, représente la classe à laquelle appartient l'objet O^i selon \mathcal{R} . κ est une **application**² de \mathbb{O} dans \mathbb{D} .

D'après ce codage, nous avons la représentation suivante de \mathcal{R}' :

$$\kappa' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^1 & D^1 & D^2 & D^3 \end{pmatrix}$$

où nous avons nommé les trois classes $\{O^1, O^2\}$, $\{O^3\}$, $\{O^4\}$ par respectivement, D^1 , D^2 et D^3 .

2.1.1.1.2 Le codage par affectation

Le codage par affectation est une représentation matricielle. Il s'agit d'un codage binaire ou encore de présence / absence.

Définition 2.1.3 Codage par affectation

Soit une partition \mathcal{R} de N objets constituée de p classes notées D^j , $j = 1, \dots, p$. Le codage par affectation de cette relation d'équivalence consiste en une matrice carrée notée K de dimension N dont le terme général K_{ij} , $i, j = 1, \dots, N$; est donnée par la relation suivante :

$$K_{ij} = \begin{cases} 1 & \text{si } O^i \text{ est affecté à la classe } D^j \\ 0 & \text{sinon} \end{cases} \quad (2.1.2)$$

K sera appelée matrice d'affectation.

Le codage par affectation peut être vue comme une représentation matricielle du codage algébrique présentée précédemment.

La définition classique donnée au codage matriciel d'une partition est la **matrice disjonctive** présentée en introduction page 8. Celle-ci est une matrice rectangulaire non nécessairement carrée et en particulier, le nombre de colonnes est égal au nombre de classes de la partition. Ainsi, les matrices disjonctives K sont généralement de dimension $(N \times p)$.

Nous décidons ici de représenter les relations d'équivalence par des matrices carrées de taille fixe et égale à N . Ces matrices d'affectation sont ainsi de dimensions maximales $(N \times N)$. L'intérêt de ce codage est qu'il nous permettra par la suite, d'exprimer des actions sur les partitions représentées par ces matrices, par l'intermédiaire de multiplications matricielles³.

Nous donnons à présent les propriétés de la matrice K .

Propriété 2.1.1 Propriétés de la matrice K

En cas de données complètes⁴, nous avons les propriétés suivantes :

- $\sum_j K_{ij} = K_i = 1 \quad \forall i = 1, \dots, N$

2. ni injective, ni surjective
 3. par ailleurs, ce codage permet l'agrégation de matrices d'affectation qui pourrait déboucher sur de nouvelles modélisations en analyse de données
 4. ie pas de données manquantes

- $\sum_i K_{ij} = K_{.j} = \#D^j \quad \forall j = 1, \dots, N$
- $\sum_{i,j} K_{ij} = K_{..} = N$

où $\#\mathbb{A}$ représente le cardinal de l'ensemble \mathbb{A} .

En revanche, une même partition peut être représentée par plusieurs matrices d'affectation. Il est utile de définir la notion d'équivalence entre matrices d'affectation codant une même partition. Il est nécessaire pour cela, de rappeler au préalable les propriétés d'une matrice de permutation.

Définition 2.1.4 Matrice de permutation

Soit P une matrice binaire carrée de dimension N vérifiant les propriétés suivantes :

- $P_{ij} \in \{0, 1\} \quad \forall i, j = 1, \dots, N$
- $\sum_j P_{ij} = P_{i.} = 1 \quad \forall i = 1, \dots, N$
- $\sum_i P_{ij} = P_{.j} = 1 \quad \forall j = 1, \dots, N$

Alors P ainsi définie est appelée matrice de permutation.

Propriété 2.1.2 Inverse d'une matrice de permutation

Soit P une matrice de permutation, et soit P^{-1} son inverse, nous avons alors la propriété suivante :

$$P^{-1} = {}^tP \tag{2.1.3}$$

P est une matrice orthogonale.

Définition 2.1.5 Equivalence entre deux matrices d'affectation

Nous dirons que deux partitions \mathcal{R}^k et \mathcal{R}^l , représentées par leurs matrices d'affectation K^k et K^l , sont équivalentes si nous avons la relation suivante :

$$\mathcal{R}^k \equiv \mathcal{R}^l \Leftrightarrow K^k \equiv K^l \Leftrightarrow \exists P : \{K^k = K^l \cdot P\} \vee \{K^l = K^k \cdot {}^tP\} \tag{2.1.4}$$

où P est une matrice de permutation.

La relation \equiv ainsi définie, engendre une relation d'équivalence sur l'ensemble des partitions.

A titre illustratif, la matrice d'affectation K' , représentative de la partition \mathcal{R}' est la suivante :

$$K' = \begin{matrix} & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

La matrice K'' suivante est une matrice d'affectation équivalente à K' . En effet, la matrice de permutation P donnée ci-dessous permet de passer de l'une à l'autre :

$$K'' = \begin{matrix} & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix} ; \quad P = \begin{matrix} & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} D^1 \\ D^2 \\ D^3 \\ D^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

2.1.1.1.3 Le codage relationnel

La codage relationnel est fondé sur les comparaisons par paires des objets sur lesquels porte la relation d'équivalence. Il s'agit aussi d'une représentation binaire mais elle est **davantage proche de la définition même des relations**. On retrouve aussi cette représentation dans le contexte de la théorie des graphes. Les matrices relationnelles que définit le codage relationnel sont aussi appelées matrices d'adjacence du graphe de la relation \mathcal{R} , noté $G(\mathcal{R})$.

Définition 2.1.6 Codage relationnel

Soit une partition \mathcal{R} de N objets constituée de p classes notées : $D^j, j = 1, \dots, p$. Le codage relationnel de cette relation d'équivalence consiste en une matrice carrée notée C de dimension N dont le terme général $C_{ii'}$ où $i, i' = 1, \dots, N$; est donnée par la relation suivante :

$$C_{ii'} = \begin{cases} 1 & \text{si } O^i \text{ et } O^{i'} \text{ appartiennent à la même classe} \\ 0 & \text{sinon} \end{cases} \quad (2.1.5)$$

Nous donnons les propriétés caractéristiques du type de la relation qu'une matrice relationnelle de partition vérifie.

Propriété 2.1.3 Propriétés relationnelles ou structurales de la matrice C

La matrice relationnelle vérifie les contraintes suivantes qui retranscrivent les propriétés relationnelles d'une relation d'équivalence :

- *Réflexivité* : $C_{ii} = 1 \quad \forall i = 1, \dots, N$
- *Symétrie* : $C_{ii'} = C_{i'i} \quad \forall i, i' = 1, \dots, N$
- *Transitivité* : $C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 \quad \forall i, i', i'' = 1, \dots, N$

Nous donnons de plus des propriétés relatives à la partition codée que permet d'exprimer le codage relationnel.

Propriété 2.1.4 Autres propriétés de la matrice C

- $\sum_{i'} C_{ii'} = C_{i.} = \text{nombre d'objets dans la classe de l'objet } O^i \quad \forall i = 1, \dots, N$
- $\sum_{i,i'} \left(\frac{C_{ii'}}{C_{i.}}\right)^2 = p$, nombre de classes de la relation d'équivalence
- $\sum_{i,i'} C_{ii'}(C_{i.} - C_{i'}) = 0$,

Démonstration

Nous démontrons les deuxième et troisième points.

$$\begin{aligned} \sum_{i,i'} \left(\frac{C_{ii'}}{C_{i.}}\right)^2 &= \sum_{i,i'} \frac{C_{ii'}^2}{C_{i.}^2} \\ &= \sum_{i,i'} \frac{C_{ii'}}{C_{i.}^2} \\ &= \sum_i \frac{C_{i.}}{C_{i.}^2} \\ &= \sum_i \frac{1}{C_{i.}} \\ &= p \end{aligned}$$

Par ailleurs, de deux choses l'une : soit $C_{ii'} = 0$; soit $C_{ii'} = 1$ et donc $C_{i.} = C_{i'}$ puisque O^i et $O^{i'}$ sont dans la même classe. De cette manière $\sum_{i,i'} C_{ii'}(C_{i.} - C_{i'}) = 0$.

□

Un autre intérêt du codage relationnel réside en la propriété étudiée ci-dessous.

Propriété 2.1.5 Complémentaire d'une relation d'équivalence

Soit une partition \mathcal{R} représentée par sa matrice relationnelle C . Soit \bar{C} la matrice relationnelle du complémentaire de \mathcal{R} défini par $\bar{C} = U_N - C$ où U_N est la matrice carrée de dimension N remplie de 1. Le terme général de \bar{C} vaut donc, $\forall i, i' = 1, \dots, N$:

$$\bar{C}_{ii'} = 1 - C_{ii'} \quad (2.1.6)$$

Alors, \bar{C} possède les propriétés d'une relation de distance :

- Irréflexivité : $\bar{C}_{ii} = 0 \quad \forall i = 1, \dots, N$
- Symétrie : $\bar{C}_{ii'} = \bar{C}_{i'i} \quad \forall i, i' = 1, \dots, N$
- Inégalité triangulaire : $\bar{C}_{ii''} \leq \bar{C}_{ii'} + \bar{C}_{i'i''} \quad \forall i, i', i'' = 1, \dots, N$

Démonstration

Les résultats sur l'irréflexivité et la symétrie sont immédiats. La propriété d'inégalité triangulaire est le complémentaire de la propriété de transitivité :

$$\begin{aligned} C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 &\Leftrightarrow -C_{ii''} \leq 1 - (C_{ii'} + C_{i'i''}) \\ &\Leftrightarrow 1 - C_{ii''} \leq 2 - (C_{ii'} + C_{i'i''}) \\ &\Leftrightarrow \bar{C}_{ii''} \leq \bar{C}_{ii'} + \bar{C}_{i'i''} \end{aligned}$$

□

A titre d'exemple, nous donnons la matrice représentative de la relation \mathcal{R}' :

$$C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

2.1.1.1.4 Lien entre le codage par affectation et le codage relationnel

Ces différents codages sont liés. Nous rappelons les formules de passage permettant de déterminer notamment le codage relationnel à partir du codage par affectation.

Propriété 2.1.6 Du codage par affectation au codage relationnel

Soit une relation d'équivalence \mathcal{R} et soit K sa matrice d'affectation associée. A partir de K , nous obtenons la matrice relationnelle C de \mathcal{R} , en appliquant la relation suivante :

$$C = K \cdot I_N \cdot {}^t K \quad (2.1.7)$$

où I_N est la matrice identité.

Nous mettons volontairement en évidence la matrice identité I_N dans l'équation précédente. Nous verrons au paragraphe 2.1.3.1.2, qu'il s'agit d'une matrice de passage qui est susceptible de changer lors de l'action d'opérations de transfert agissant sur les classes d'une partition.

Propriété 2.1.7 Matrices d'affectation équivalentes et matrice relationnelle

Soit \mathcal{R}^k une partition, K^k sa matrice d'affectation et C^k sa matrice relationnelle associée. Toute matrice d'affectation K^l , représentative d'une partition \mathcal{R}^l , appartenant à la classe d'équivalence de K^k possède la même représentation relationnelle :

$$\boxed{\mathcal{R}^k \equiv \mathcal{R}^l \Leftrightarrow K^l \equiv K^k \Leftrightarrow C^l = C^k} \quad (2.1.8)$$

Démonstration

$$\begin{aligned} C^l &= K^l \cdot {}^t K^l \\ &= (K^k \cdot P) \cdot {}^t (K^k \cdot P) \quad \text{car } K^k \text{ et } K^l \text{ sont équivalentes} \\ &= K^k \cdot P \cdot {}^t P \cdot {}^t K^k \\ &= K^k \cdot {}^t K^k \quad \text{car } {}^t P = P^{-1} \\ &= C^k \end{aligned}$$

□

2.1.1.2 Le treillis des partitions

Définition 2.1.7 Ordre partiel sur l'ensemble des partitions

Nous dirons que la partition \mathcal{R}^k est inférieure ou égale à la partition \mathcal{R}^l ce que nous noterons par $\mathcal{R}^k \preceq \mathcal{R}^l$ si leurs représentations relationnelles respectives, C^k et C^l , respectent la propriété suivante :

$$\mathcal{R}^k \preceq \mathcal{R}^l \Leftrightarrow C_{ii'}^k \leq C_{ii'}^l \quad \forall i, i' = 1, \dots, N \quad (2.1.9)$$

En d'autres termes, une première partition est inférieure ou égale à une seconde partition, si les classes de la première sont incluses dans les classes de la seconde.

Propriété 2.1.8 Structure de treillis

L'ensemble des partitions d'un ensemble d'objets, muni de la relation d'ordre partiel \preceq , respecte une structure de treillis pour laquelle :

- la borne supérieure est la partition d'une classe regroupant tous les objets
- la borne inférieure est la partition de N classes où chaque objet définit une classe singleton notée κ^ϵ dont la représentation par affectation notée K^ϵ est égale à I_N

Nous représentons dans la figure fig-2.1, le treillis des partitions de l'ensemble d'objets $\{O^1, O^2, O^3, O^4\}$. Une partition \mathcal{R}^l pointe vers une partition \mathcal{R}^k si $\mathcal{R}^k \preceq \mathcal{R}^l$.

Nous constatons qu'il y a incomparabilité pour des partitions de même nombre de classes. Nous rappelons au passage, que le nombre de partitions à k classes d'un ensemble de N objets est donné par le nombre de Stirling de seconde espèce noté $S(N, k)$ dont la formule est :

$$S(N, k) = \frac{1}{k!} \sum_{j=0}^k (-1)^j \binom{k}{j} (k-j)^N \quad (2.1.10)$$

Les formules de récurrence à partir desquelles, nous pouvons aussi obtenir les $S(N, k)$ sont les suivantes :

$$\begin{cases} S(N, k) &= S(N-1, k-1) + kS(N-1, k) \\ S(N, 1) &= 1 \\ S(N, N) &= 1 \end{cases} \quad (2.1.11)$$

Ainsi dans l'exemple présenté nous avons $S(4, 1) = 1$, $S(4, 2) = 7$, $S(4, 3) = 6$, et $S(4, 4) = 1$.

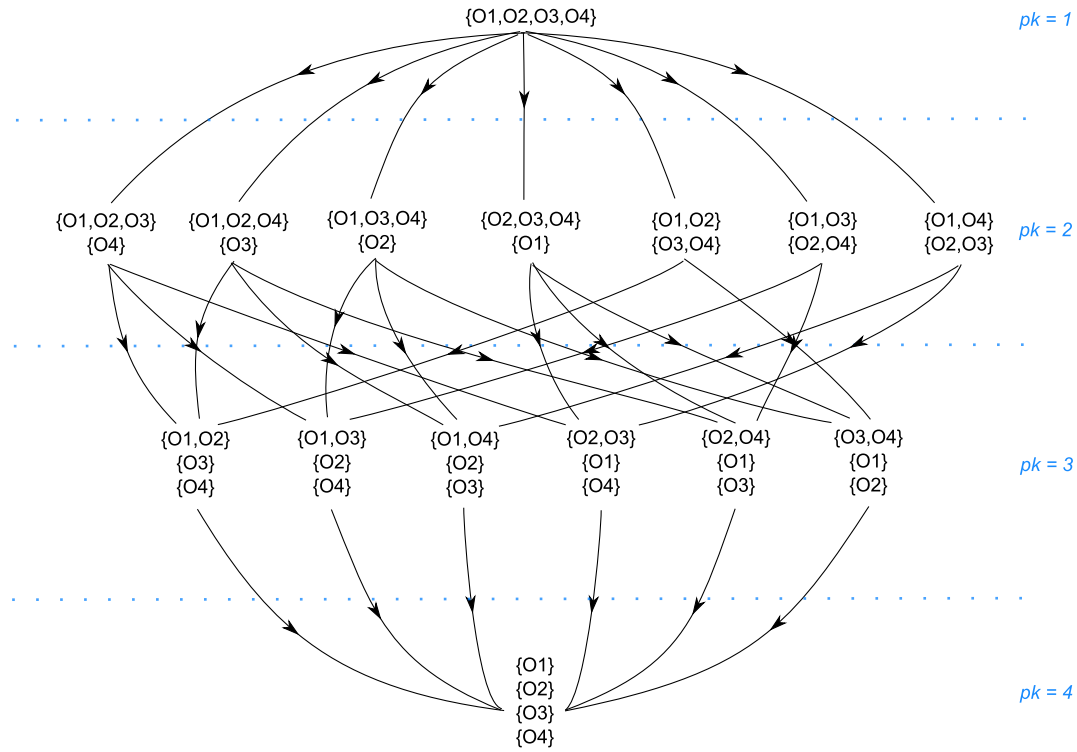


FIGURE 2.1 – Exemple de treillis de partitions

2.1.2 Opérations de transferts et structures algébriques associées

Nous définissons des opérations de transferts qui agissent sur les partitions et permettent de modifier la structure de ces dernières. Nous étudions ensuite de manière plus précise, deux types de transferts particuliers que nous appellerons transferts simples et transferts doubles.

Cette sous-section, bien que définissant d'un point de vue algébrique des éléments simples que sont les transferts, s'inscrit dans une perspective avant tout algorithmique que nous souhaitons souligner dès à présent afin d'expliquer la démarche entreprise. L'objectif est de modéliser par une approche théorique certains algorithmes de classification automatique et de montrer⁵ ainsi leurs limites. De plus, ceci nous permet de mettre en valeur, un nouvel algorithme de classification automatique que nous présentons en section 2.2.4 et qui utilise des opérations davantage évoluées que les transferts.

2.1.2.1 Le monoïde des opérations de transferts

Définition 2.1.8 Opérations de transfert

Un transfert sur un ensemble d'objets $O^i, i = 1, \dots, N$; consiste à associer de manière unique chaque objet à un autre objet. Soit τ une opération de transfert, celle-ci sera notée de la façon suivante :

$$\tau = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ \tau(O^1) & \tau(O^2) & \dots & \tau(O^N) \end{pmatrix} \quad (2.1.12)$$

où $\tau(O^i) \in \{O^1, \dots, O^N\}$, représente l'objet auquel est associé l'objet O^i . Il s'agit d'une application⁶ de \mathbb{O} dans lui-même.

5. en section 2.2

6. ni injective, ni surjective

Définition 2.1.9 Représentation matricielle des opérations de transfert

Soit τ une opération de transfert entre les objets $O^i, i = 1, \dots, N$. Nous pouvons coder cette opération par une matrice carrée binaire T , de dimension N dont le terme général, $T_{ii'}$ où $i, i' = 1, \dots, N$; est donné par :

$$T_{ii'} = \begin{cases} 1 & \text{si } \tau(O^i) = O^{i'} \\ 0 & \text{sinon} \end{cases} \quad (2.1.13)$$

Propriété 2.1.9 Propriété structurelle des matrices de transfert

La propriété qui caractérise les matrices de transfert est la suivante :

$$\sum_{i'=1}^N T_{ii'} = T_{i.} = 1 \quad \forall i = 1, \dots, N \quad (2.1.14)$$

Remarque 2.1.2 Liens entre matrices d'affectation et matrices de transfert

Les matrices d'affectation et les matrices de transfert possèdent la même propriété structurelle : pour chaque ligne la somme sur les colonnes vaut 1.

La différence entre ces matrices se situe au niveau des ensembles qu'elles représentent. Pour ces deux types de matrices, l'ensemble des objets est représenté par leurs lignes. Mais en ce qui concerne les colonnes, les matrices d'affectation représentent l'ensemble des classes ou descripteurs alors que les matrices de transfert représentent à nouveau l'ensemble des objets.

En d'autres termes, les matrices d'affectation codent des applications d'un ensemble vers un autre ensemble distinct alors que les matrices de transfert codent des applications d'un ensemble dans lui-même. Il en résulte une signification bien distincte malgré les propriétés communes que partagent ces matrices.

Remarque 2.1.3 Bijection entre un transfert et sa matrice associée

Il y a une bijection entre chaque transfert et sa matrice associée. Ainsi, par abus de langage, nous associerons l'ensemble $\mathbb{T}(\mathbb{O})$ aux opérations définies par déf-2.1.8 mais aussi aux matrices définies en déf-2.1.9.

Il existe des lois internes aux transferts. Celles-ci sont des compositions d'applications. **Nous verrons deux types de compositions : la composition à gauche et la composition à droite.** Nous précisons ces deux définitions dans ce qui suit. Puis, **nous examinerons de manière plus détaillée en quoi cette distinction est utile en sous-section 2.1.3.1 et en section 2.2 : en effet, elle permet de séparer les actions des transferts sur les objets, des actions de transferts sur les classes de la même manière que l'algorithme des centre mobiles se distingue de l'algorithme de classification hiérarchique ascendante.**

Définition 2.1.10 Composition à gauche d'opérations de transfert

Soient τ^k, τ^l et τ^m trois transferts de $\mathbb{T}(\mathbb{O})$. Nous dirons que τ^m est le transfert obtenu par la composition à gauche de τ^k par τ^l si celui-ci est défini de la manière suivante :

$$\tau^m(O^i) = (\tau^l \circ \tau^k)(O^i) = \tau^l(\tau^k(O^i)) \quad \forall i = 1, \dots, N \quad (2.1.15)$$

Soient T^k, T^l et T^m les représentations matricielles respectives de chaque transfert. La composition à gauche se traduit matriciellement par la post-multiplication :

$$\tau^m = \tau^l \circ \tau^k \Leftrightarrow T^m = T^k \cdot T^l \quad (2.1.16)$$

Définition 2.1.11 Composition à droite d'opérations de transfert

Soient τ^k, τ^l et τ^m trois transferts de $\mathbb{T}(\mathbb{O})$. Nous dirons que τ^m est le transfert obtenu par la composition à droite de τ^k par τ^l si celui-ci est défini de la manière suivante :

$$\tau^m(O^i) = (\tau^k \circ \tau^l)(O^i) = \tau^k(\tau^l(O^i)) \quad \forall i = 1, \dots, N \quad (2.1.17)$$

Soient T^k, T^l et T^m les représentations matricielles respectives de chaque transfert. La composition à droite se traduit matriciellement par la pré-multiplication :

$$\tau^m = \tau^k \circ \tau^l \Leftrightarrow T^m = T^l \cdot T^k \quad (2.1.18)$$

Pour fixer les idées, nous donnons l'exemple de la composition à droite du transfert τ' par le transfert τ'' suivant :

$$\tau' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^1 & O^1 & O^2 & O^3 \end{pmatrix} \quad ; \quad \tau'' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^4 & O^2 & O^1 \end{pmatrix}$$

Nous obtenons :

$$\tau' \circ \tau'' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^1 & O^3 & O^1 & O^1 \end{pmatrix}$$

Les représentations matricielles sont les suivantes :

$$T' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix} \quad ; \quad T'' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

Nous vérifions que nous avons bien le même résultat :

$$T'' \cdot T' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

L'ensemble de ces opérations possèdent des structures algébriques que nous mettons en évidence. Pour cela nous rappelons dans un premier temps la définition d'un monoïde.

Définition 2.1.12 Monoïde (\mathbb{A}, \oplus)

(\mathbb{A}, \oplus) est un monoïde si \mathbb{A} est un ensemble d'éléments muni d'une loi de composition interne \oplus et qui vérifie les propriétés suivantes :

- \oplus possède un élément neutre ϵ
- \oplus est associative

Propriété 2.1.10 Monoïde des opérations de transfert

L'ensemble des matrices de transfert sur l'ensemble des objets \mathbb{O} , noté $\mathbb{T}(\mathbb{O})$, forme un monoïde pour la multiplication matricielle que l'on notera $(\mathbb{T}(\mathbb{O}), \cdot)$.

L'élément neutre de ce monoïde, noté T^ϵ est la matrice identité d'ordre $N : I_N$.

Nous venons de présenter l'ensemble $\mathbb{T}(\mathbb{O})$, des opérations de transfert définies sur un ensemble \mathbb{O} . Nous pouvons distinguer deux sous-ensembles particuliers que nous nous proposons d'étudier : **les transferts simples et les transferts doubles. Nous verrons en sous-sections 2.1.3.1 et 2.1.3.2, en quoi l'action sur les partitions de ces deux types de transfert diffèrent.**

Nous présentons dans les paragraphes suivants, ces deux sous-ensembles de transferts ainsi que les structures algébriques qu'ils vérifient.

2.1.2.2 Le sous-monoïde des opérations de transferts simples

Définition 2.1.13 Opération de transfert simple

Une opération de transfert τ de $\mathbb{T}(\mathbb{O})$ sera dite simple si sa matrice de transfert associée, T , vérifie la propriété additionnelle suivante :

$$\tau \text{ est simple} \Leftrightarrow \exists i' \in \{1, \dots, N\} : \sum_i T_{ii'} = T_{.i'} = 0 \quad (2.1.19)$$

Une matrice de transfert simple est une matrice pour laquelle, il existe au moins un objet en colonne, auquel aucun autre objet en ligne n'est associé.

Le transfert suivant est simple par exemple :

$$\tau''' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^2 & O^4 & O^4 \end{pmatrix} \Leftrightarrow T''' = \begin{matrix} O^1 & O^2 & O^3 & O^4 \\ \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

Propriété 2.1.11 Monoïde des opérations de transferts simples

Le sous-ensemble des matrices de transfert simple défini sur un ensemble \mathbb{O} , noté $\mathbb{T}^s(\mathbb{O})$ auquel nous joignons la matrice identité $T^\epsilon = I_N$, forme un monoïde pour la multiplication matricielle .

Démonstration

La propriété structurelle des matrices de transfert simple est stable par multiplication matricielle.

$$\begin{aligned} \sum_i [T^k \cdot T^l]_{ii'} &= \sum_i (\sum_{i''} T_{ii''}^k T_{i''i'}^l) \\ &= \sum_{i''} \sum_i T_{ii''}^k T_{i''i'}^l \\ &= \sum_{i''} T_{.i''}^k T_{i''i'}^l \end{aligned}$$

Or, $\exists i'' : T_{.i''}^k = 0$ car T^k est une matrice de transfert simple. Ainsi, $\exists i' : [T^k \cdot T^l]_{.i'} = 0$.

□

Parmi ces opérations de transferts simples, nous nous intéressons aux transferts générateurs, c'est à dire au sous-ensemble des transferts simples dont les différentes combinaisons de composition permettent de définir tout l'ensemble des transferts simples. Ce sous-ensemble sera noté $\mathbb{T}^{sg}(\mathbb{O})$.

Définition 2.1.14 Transfert simple générateur

Un transfert simple générateur, $\tau_i^{i'}$ de $\mathbb{T}^{sg}(\mathbb{O})$, est celui qui associe l'objet O^i à l'objet $O^{i'}$ et tous les autres objets à eux même.

Définition 2.1.15 Représentation matricielle des opérations de transferts simples générateurs

Soit $\tau_i^{i'}$ une opération de transfert simple générateur. Le terme général de sa matrice associée, est défini de la manière suivante, $\forall \iota, \iota' = 1, \dots, N$:

$$[T_i^{i'}]_{\iota \iota'} = \begin{cases} 1 & \text{si } \{\iota = i \wedge \iota' = i'\} \vee \{\iota = \iota' \wedge \iota \neq i\} \\ 0 & \text{sinon} \end{cases} \quad (2.1.20)$$

Propriété 2.1.12 Propriétés structurelles des matrices $T_i^{i'}$

- $\sum_\iota [T_i^{i'}]_{\iota i} = [T_i^{i'}]_{.i} = 0$

- $\sum_{i'} [T_i^{i'}]_{i,i'} = [T_i^{i'}]_{i,i'} = 2$

Propriété 2.1.13 Propriétés algébriques des matrices $T_i^{i'}$

- Le nombre de transferts simples générateurs distincts du transfert identité est de $N(N - 1)$
- $T_i^i = T^\epsilon = I_N$, la matrice identité
- $T_i^{i'} \cdot T_i^{i'} = T_i^{i'}$: ce sont des matrices de projection
- $T_{i'}^i \cdot T_i^{i'} = T_i^{i'}$

Nous donnons pour illustration, l'exemple du transfert simple τ_1^2 , dans le cas où $N = 4$:

$$\tau_1^2 = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^2 & O^3 & O^4 \end{pmatrix} \Leftrightarrow T_1^2 = \begin{matrix} O^1 & O^2 & O^3 & O^4 \\ O^1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ O^2 \begin{pmatrix} 0 & 1 & 0 & 0 \\ O^3 \begin{pmatrix} 0 & 0 & 1 & 0 \\ O^4 \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{matrix}$$

2.1.2.3 Le groupe des opérations de transferts doubles

Définition 2.1.16 Opération de transfert double

Une opération de transfert τ de $\mathbb{T}(\mathbb{O})$ sera dite double si sa matrice associée, T , vérifie la propriété additionnelle suivante :

$$\tau \text{ est double} \Leftrightarrow \sum_i T_{ii'} = T_{i,i'} = 1 \quad \forall i' = 1, \dots, N \quad (2.1.21)$$

Autrement dit un transfert est double s'il n'est pas simple et vice versa.

Un transfert double est une opération pour laquelle, tous les objets en colonne sont pourvus d'une unique association. Il s'agit d'applications bijectives contrairement aux transferts simples.

L'exemple suivant est un transfert double :

$$\tau'''' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^4 & O^1 \end{pmatrix} \Leftrightarrow T'''' = \begin{matrix} O^1 & O^2 & O^3 & O^4 \\ O^1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ O^2 \begin{pmatrix} 0 & 0 & 1 & 0 \\ O^3 \begin{pmatrix} 0 & 0 & 0 & 1 \\ O^4 \begin{pmatrix} 1 & 0 & 0 & 0 \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{pmatrix} \end{matrix}$$

La définition des matrices de transfert double aboutit à une structure identique à celle des matrices de permutation⁷. Nous en déduisons la propriété suivante :

Propriété 2.1.14 Groupe des opérations de transferts doubles

Le sous-ensemble des matrices de transfert double définies sur un ensemble \mathbb{O} , noté $\mathbb{T}^d(\mathbb{O})$ forme un groupe pour la multiplication matricielle que l'on notera $(\mathbb{T}^d(\mathbb{O}), \cdot)$.

L'élément neutre de ce groupe, noté T^ϵ est la matrice identité d'ordre N : $T^\epsilon = I_N$.

Comme précédemment, nous allons présenter le sous-ensemble des transferts doubles générateurs. Nous notons ce sous-ensemble $\mathbb{T}^{dg}(\mathbb{O})$.

7. cf page 35

Définition 2.1.17 Transfert double générateur

Un transfert double générateur $\tau^{ii'}$ de $\mathbb{T}^{dg}(\mathbb{O})$, est celui qui associe de manière simultanée l'objet O^i à l'objet $O^{i'}$ et l'objet $O^{i'}$ à l'objet O^i et tous les autres objets à eux même⁸.

Définition 2.1.18 Représentation matricielle des opérations de transfert double générateur

Soit $\tau^{ii'}$ une opération de transfert double générateur. Le terme général de sa matrice de transfert associée, est définie de la manière suivante, $\forall \iota, \iota' = 1, \dots, N$:

$$[T^{ii'}]_{\iota\iota'} = \begin{cases} 1 & \text{si } \{(\iota, \iota') = (i, i')\} \vee \{(\iota, \iota') = (i', i)\} \vee \{\iota = \iota' \wedge \iota \neq i \wedge \iota \neq i'\} \\ 0 & \text{sinon} \end{cases} \quad (2.1.22)$$

Propriété 2.1.15 Propriétés algébriques des matrices $T^{ii'}$

- Le nombre de transferts doubles générateurs distincts du transfert identité est de $N(N - 1)/2$
- $T^{ii'} = T^{i'i} = {}^tT^{ii'}$: ce sont des matrices symétriques
- ${}^tT^{ii'} = (T^{ii'})^{-1}$: ce sont des matrices orthogonales par ailleurs, nous avons $T^{ii'} \cdot T^{ii'} = I_N$
- $T^{ii'} \cdot T^{i''i'''} = T^{i''i'''} \cdot T^{ii'}$ $\forall i \neq i' \neq i'' \neq i'''$: propriété de commutativité partielle

Nous donnons pour illustration, l'exemple du transfert double τ^{12} , dans le cas où $N = 4$:

$$\tau^{12} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^1 & O^3 & O^4 \end{pmatrix} \Leftrightarrow T^{12} = \begin{matrix} O^1 & O^2 & O^3 & O^4 \\ O^1 \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ O^3 \begin{pmatrix} 0 & 0 & 1 & 0 \\ O^4 \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \end{pmatrix} \end{matrix}$$

2.1.3 Action des opérations de transferts sur les partitions

Après avoir défini les opérations de transferts, nous nous intéressons à leurs actions sur les partitions. **Nous étudions l'action des transferts sur les partitions représentées tout d'abord par des matrices d'affectation puis par des matrices relationnelles.** Ces actions sont de deux types⁹ : elles agissent soit sur les objets, soit sur les classes.

L'intérêt des études entreprises précédemment et dans cette sous-section prend tout son sens pratique dans la section 2.2 suivante, où nous montrons que les algorithmes de classification de centres mobiles et de classification hiérarchique ascendante sont modélisés par les différentes actions présentées ici. Un deuxième intérêt est d'enrichir le "dictionnaire" des formules relatives aux matrices relationnelles dans le cadre de l'AR. En particulier, le paragraphe 2.1.3.2.3 montre un nouveau résultat qui met en évidence une dualité entre codage relationnel et codage par affectation lors de l'action d'un transfert sur une matrice relationnelle.

Nous supposons dans ce qui suit, que nous disposons d'une partition \mathcal{R} représentée soit par κ son codage algébrique, soit par K sa matrice d'affectation, soit encore par C sa matrice relationnelle.

8. un élément $\tau^{i, i+1}$ sera aussi appelé transposition d'ordre 1 étant donné qu'il correspond à ces éléments dans le cadre du groupe des permutations comme nous le verrons au chapitre 5 en sous-section 5.1.2

9. cette distinction est à rapprocher de la différence de stratégie existante entre l'algorithme des centres mobiles et celui de la classification hiérarchique ascendante que nous rappelons en section 2.2. En effet, le premier "raisonne" en termes de transferts simples d'objets alors que le second "raisonne" en terme de transferts simples de classes

Remarque 2.1.4 Notations

Nous allons distinguer deux types d'opérations de transferts : ceux qui agissent sur les objets $\{O^1, \dots, O^N\}$ et ceux qui agissent sur les descripteurs $\{D^1, \dots, D^N\}$. Les premiers seront notés τ et sont des éléments de $\mathbb{T}(\mathbb{O})$ et les seconds seront notés σ et sont des éléments de $\mathbb{T}(\mathbb{D})$.

De la même manière, nous noterons par T les matrices de transfert représentatives des éléments τ et par S les matrices de transfert représentatives des éléments σ .

Ces deux types de matrices possèdent les mêmes propriétés. La seule différence est qu'elles ne sont pas définies sur les mêmes ensembles. Ainsi, dans un souci de clarté, nous avons préféré adopter des notations différentes.

2.1.3.1 Action des opérations de transferts sur les matrices disjointives

2.1.3.1.1 Action sur les objets

L'action de τ sur les objets de κ consistera à transférer un objet dans la classe de l'objet auquel celui-ci est associé.

Autrement dit, supposons que l'objet O^i soit associé à l'objet $O^{i'}$ selon τ . L'action de τ sur κ localement à l'objet O^i , consiste donc à transférer celui-ci dans la classe de l'objet $O^{i'}$.

Définition 2.1.19 Action d'un transfert sur les objets d'une partition \mathcal{R} dans le cadre du codage par affectation

Soit κ le codage algébrique d'une relation d'équivalence \mathcal{R} , et soit τ un élément de $\mathbb{T}(\mathbb{O})$. L'action de τ sur l'espace des objets \mathbb{O} de κ est donnée par la **composition à droite** de κ par τ :

$$\tau(\mathcal{R}) = \kappa \circ \tau \tag{2.1.23}$$

où

$$(\kappa \circ \tau)(O^i) = \kappa(\tau(O^i)) \quad \forall i = 1, \dots, N \tag{2.1.24}$$

Soit T la matrice de transfert associée à τ . L'action du transfert sur la partition se traduit matriciellement par la relation suivante :

$$\boxed{\tau(\mathcal{R}) = T \cdot K} \tag{2.1.25}$$

Nous donnons l'exemple de l'action de τ''' donné page 42 sur \mathcal{R}' dont le codage algébrique est donné page 34 :

$$\tau'''(\mathcal{R}') = \kappa' \circ \tau''' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^1 & D^1 & D^2 & D^3 \end{pmatrix} \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^2 & O^4 & O^4 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^1 & D^1 & D^3 & D^3 \end{pmatrix}$$

Dans le cadre des représentations matricielles, l'action de T''' sur K' est donnée par :

$$T''' \cdot K' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & & D^1 & D^2 & D^3 & D^4 & & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} & \cdot & \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} & = & \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

La matrice d'affectation résultante est modifiée localement au niveau de l'objet O^3 . En effet, celui-ci est transféré dans la classe de l'objet O^4 selon τ''' . Or, l'objet O^4 a pour classe D^3 . Ainsi, l'objet O^3 passe de la classe D^2 à la classe D^3 .

2.1.3.1.2 Action sur les classes

Soit σ un élément de $\mathbb{T}(\mathbb{D})$ représenté par sa matrice de transfert S . A la différence de ce qui précède, les transferts sont définis sur l'ensemble des éléments de \mathbb{D} qui est l'ensemble des classes¹⁰.

De manière symétrique, l'action sur les classes de σ sur κ consistera à transférer tous les objets appartenant à une classe dans la classe à laquelle celle-ci est associée.

Supposons que la classe D^j soit associée à la classe $D^{j'}$ selon σ . L'action de σ sur κ localement à la classe D^j , consiste à transférer tous les objets de celle-ci dans la classe $D^{j'}$.

Définition 2.1.20 Action d'un transfert sur les classes d'une partition \mathcal{R} dans le cadre du codage par affectation

Soit κ le codage algébrique d'une relation d'équivalence \mathcal{R} , et soit σ un élément de $\mathbb{T}(\mathbb{D})$. L'action de σ sur l'espace des classes \mathbb{D} de κ est donnée par la **composition à gauche** de κ par σ :

$$\sigma(\mathcal{R}) = \sigma \circ \kappa \quad (2.1.26)$$

où,

$$(\sigma \circ \kappa)(O^i) = \sigma(\kappa(O^i)) \quad \forall i = 1, \dots, N \quad (2.1.27)$$

Soit K la matrice d'affectation représentative de κ et soit S la matrice de transfert associée à σ . L'action du transfert sur la partition se traduit matriciellement par la relation suivante :

$$\boxed{\sigma(\mathcal{R}) = K \cdot S} \quad (2.1.28)$$

Parmi les actions des transferts sur l'ensemble des descripteurs d'une partition, celles du sous-ensemble des transferts doubles permettent de changer la classe à laquelle est affecté un ensemble d'objets. Cependant, une matrice de transfert double étant telle une matrice de permutation, son action sur les colonnes d'une matrice d'affectation ne change pas la structure de la partition. En effet, l'action d'un transfert double sur les descripteurs d'une partition K^k donne une partition K^l équivalente à K^k au sens de la définition déf-2.1.5. Nous avons donc la propriété suivante :

Propriété 2.1.16 Action invariante d'un transfert double sur les classes d'une partition \mathcal{R}

Soit σ une matrice de transfert double élément de $\mathbb{T}^d(\mathbb{D})$. L'action de σ sur les descripteurs d'une partition ne change pas la structure de cette dernière.

$$\boxed{\sigma \in \mathbb{T}^d(\mathbb{D}) \Rightarrow \sigma(\mathcal{R}) \equiv \mathcal{R}} \quad (2.1.29)$$

Pour illustrer l'action sur les classes d'une partition, nous donnons l'exemple de l'action de σ_1^2 donné page 43 sur \mathcal{R}' :

$$\sigma_1^2(\mathcal{R}') = \sigma_1^2 \circ \kappa' = \begin{pmatrix} D^1 & D^2 & D^3 & D^4 \\ D^2 & D^2 & D^3 & D^4 \end{pmatrix} \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^1 & D^1 & D^2 & D^3 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^2 & D^2 & D^2 & D^3 \end{pmatrix}$$

Dans le cadre des représentations matricielles, l'action de S_1^2 sur K' est donnée par :

$$K' \cdot S_1^2 = \begin{matrix} & D^1 & D^2 & D^3 & D^4 & & D^1 & D^2 & D^3 & D^4 & & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} & \cdot & \begin{matrix} D^1 \\ D^2 \\ D^3 \\ D^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & = & \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

10. descripteurs ou modalités dans le cas de variables qualitatives

2.1.3.1.3 Tableau récapitulatif

Nous récapitulons les actions des opérations de transferts sur les relations d'équivalence codées par des matrices d'affectation dans le tableau tab-2.1. \mathcal{R} est une relation d'équivalence représentée soit par κ , son codage algébrique, soit par K , sa matrice d'affectation.

Type de transferts	Codage algébrique	Codage par affectation
$\tau \in \mathbb{T}(\mathbb{O})$	$\tau(\mathcal{R}) = \kappa \circ \tau$	$\tau(\mathcal{R}) = T \cdot K$
$\sigma \in \mathbb{T}(\mathbb{D})$	$\sigma(\mathcal{R}) = \sigma \circ \kappa$	$\sigma(\mathcal{R}) = K \cdot S$

TABLE 2.1 – Action des opérations de transferts sur les matrices d'affectation

2.1.3.2 Action des opérations de transferts sur les matrices relationnelles

2.1.3.2.1 Action sur les objets

Dans le cadre des matrices relationnelles, l'action d'un transfert sur les objets d'une partition s'effectue de manière différente de celle que pour les matrices d'affectation. Il s'agit en fait d'une action similaire à une **action par conjugaison**.

Définition 2.1.21 *Action d'un transfert sur les objets d'une partition \mathcal{R} dans le cadre du codage relationnel*

Soit C la matrice relationnelle représentative de la partition \mathcal{R} et soit T la matrice de transfert associée à τ élément de $\mathbb{T}(\mathbb{O})$. L'action du transfert T sur les objets de la partition C se traduit par la relation suivante :

$$\tau(\mathcal{R}) = C' = T \cdot C \cdot {}^tT \quad (2.1.30)$$

où C' est la matrice relationnelle représentative de $\tau(\mathcal{R})$.

Démonstration

Dans le cas où κ est représentée par une matrice d'affectation K , nous avons :

$$\tau(\mathcal{R}) = T \cdot K$$

Appelons K' la matrice d'affectation résultante de l'action de τ sur K :

$$K' = T \cdot K.$$

Or, la matrice relationnelle C' représentative de la partition codée par K' est obtenue à partir de cette dernière de la manière suivante, en raison de la propriété prop-6.1.6 :

$$C' = K' \cdot {}^tK'$$

En développant cette dernière relation nous obtenons le résultat :

$$\begin{aligned} C' &= (T \cdot K) \cdot {}^t(T \cdot K) \\ &= T \cdot (K \cdot {}^tK) \cdot {}^tT \\ &= T \cdot C \cdot {}^tT \end{aligned}$$

□

Nous pouvons vérifier rapidement que nous obtenons, pour l'exemple de la page 45, la même partition mais dans son codage relationnel. Nous rappelons que la matrice relationnelle C' représentative de la relation d'équivalence \mathcal{R}' a été donnée page 37.

$$\begin{aligned} \tau'''(\mathcal{R}') &= T''' \cdot C' \cdot {}^tT''' \\ &= \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \\ O^3 & \\ O^4 & \end{pmatrix} \cdot \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \\ &= \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{aligned}$$

2.1.3.2.2 Action sur les classes

L'action sur les classes dans le cadre des matrices relationnelles n'est pas aussi immédiate que pour l'action sur les objets. Comme nous le verrons, les matrices de transfert n'agissent pas directement sur une matrice relationnelle représentative de la partition subissant l'action. **Cette action se traduit en fait par la modification d'une matrice de passage qui est une matrice relationnelle de partition sur l'ensemble des classes.**

A la différence du paragraphe précédent, l'opération de transfert considérée est un élément de $\mathbb{T}(\mathbb{D})$ où \mathbb{D} est l'ensemble des classes. Elle sera donc notée σ et représentée matriciellement par S .

Définition 2.1.22 Action d'un transfert sur les classes d'une partition \mathcal{R} dans le cadre du codage relationnel

Soit K la matrice d'affectation représentative de la partition \mathcal{R} et soit S la matrice de transfert associée à σ élément de $\mathbb{T}(\mathbb{D})$. L'action du transfert S sur les classes de la partition représentée par K dans le cadre du codage relationnel se traduit par la relation suivante :

$$\boxed{\sigma(\mathcal{R}) = C' = K \cdot \mathbf{R} \cdot {}^tK} \quad (2.1.31)$$

où C' est la matrice relationnelle représentative de $\sigma(\mathcal{R})$ et où \mathbf{R} est une matrice relationnelle de passage égale ici à $S \cdot I_N \cdot {}^tS$ telle que, $\forall j, j' = 1, \dots, N$:

$$\mathbf{R}_{jj'} = \begin{cases} 1 & \text{si } D^j \text{ est associée à } D^{j'} \text{ selon le transfert } S \\ 0 & \text{sinon} \end{cases} \quad (2.1.32)$$

Démonstration

Appelons K' la matrice d'affectation résultant de l'action de σ sur les descripteurs de K :

$$K' = K \cdot S$$

Or, la matrice relationnelle C' associée à la matrice d'affectation K' est obtenue à partir de cette dernière de la manière suivante, en raison de la propriété 6.1.6 :

$$C' = K' \cdot {}^tK'$$

En développant cette dernière relation nous obtenons le résultat :

$$\begin{aligned} C' &= (K \cdot S) \cdot {}^t(K \cdot S) \\ &= K \cdot (S \cdot {}^tS) \cdot {}^tK \\ &= K \cdot \mathbf{R} \cdot {}^tK \end{aligned}$$

□

Nous vérifions que nous obtenons, pour l'exemple donné page 46, le même résultat mais où la partition est représentée par son codage relationnel.

$$\begin{aligned}
 \sigma_1^2(\mathcal{R}') &= K' \cdot (S_1^2 \cdot {}^t S_1^2) \cdot {}^t K' \\
 &= \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \left(\begin{matrix} D^1 & D^2 & D^3 & D^4 \\ D^1 \\ D^2 \\ D^3 \\ D^4 \end{matrix} \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{matrix} D^1 & D^2 & D^3 & D^4 \\ D^1 \\ D^2 \\ D^3 \\ D^4 \end{matrix} \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \right) \cdot \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 &= \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} \begin{pmatrix} D^1 & D^2 & D^3 & D^4 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \cdot \left(\begin{matrix} D^1 & D^2 & D^3 & D^4 \\ D^1 \\ D^2 \\ D^3 \\ D^4 \end{matrix} \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} \right) \cdot \begin{matrix} D^1 \\ D^2 \\ D^3 \\ D^4 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \\
 &= \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}
 \end{aligned}$$

2.1.3.2.3 Dualité "Affectation / Relationnel"

Lorsque nous étudions l'action des transferts sur les partitions dans le cadre des matrices relationnelles, il en ressort une dualité entre les matrices de type d'affectation ou de transfert et les matrices relationnelles selon que l'action porte sur les objets ou sur les classes.

Nous présentons ci-dessous cette propriété que nous appelons "affectation / relationnel". Nous précisons que par abus de langage, nous parlons de codage par affectation pour faire référence à la fois à des matrices d'affectation et à des matrices de transfert. Ceci est dû aux propriétés communes que ces deux types de matrices vérifient comme nous l'avons déjà fait remarquer en rq-2.1.2 page 40.

Propriété 2.1.17 Dualité "Affectation / Relationnel"

Soit une relation d'équivalence \mathcal{R} représentée soit par sa matrice d'affectation K , soit par sa matrice relationnelle C .

Soit τ un transfert agissant sur l'ensemble des objets \mathbb{O} , et soit σ un transfert agissant sur l'ensemble des descripteurs \mathbb{D} qui sont respectivement représentées par leurs matrices de transfert associées T et S .

Soit \mathcal{R}' la partition obtenue après l'action de τ ou de σ sur \mathcal{R} et soit C' la matrice relationnelle associée à la partition \mathcal{R}' .

Soit \mathbf{R} une matrice relationnelle codant une relation d'équivalence sur les éléments de \mathbb{D} dite matrice de passage. Soit enfin, \mathbf{D} une matrice d'affectation ou de transfert : les deux types de matrices vérifiant les mêmes propriétés structurelles.

Les actions de τ sur \mathcal{R} dans le cadre du codage relationnel sont de la forme générale suivante :

$$C' = \mathbf{D} \cdot \mathbf{R} \cdot {}^t \mathbf{D} \quad (2.1.33)$$

que nous exprimons par la formulation suivante :

$$\text{"Relationnel"} = \text{Affectation} \cdot \text{Relationnel} \cdot {}^t \text{Affectation"}$$

Nous avons les différents cas suivants que nous illustrons dans le tableau tab-2.2. Dans celui-ci I_N est la matrice identité qui peut être vue ici comme étant la matrice de passage standard : I_N représente la partition triviale où chaque élément de \mathbb{D} constitue une classe singleton.

Type de transferts	Définition de \mathbf{R}	Définition de \mathbf{D}
$\tau \in \mathbb{T}(\mathbb{O})$	$\mathbf{R} = K \cdot I_N \cdot {}^t K = C$	$\mathbf{D} = T$
$\sigma \in \mathbb{T}(\mathbb{D})$	$\mathbf{R} = S \cdot I_N \cdot {}^t S$	$\mathbf{D} = K$

TABLE 2.2 – Dualité “affectation / relationnel” lors de l’action des opérations de transferts sur des matrices relationnelles codant des partitions

2.2 Les méthodes de classification automatique vues sous un angle algébrique

Le problème de la classification automatique d’un ensemble d’objets \mathbb{O} , qui nous intéresse, consiste à déterminer une partition de cet ensemble de sorte que les classes constituées soient les plus homogènes possible. Une classe est homogène si les objets qui la constituent se ressemblent mutuellement. Ainsi la notion d’homogénéité est donnée par la définition d’un critère Δ dit de partitionnement qui est fondé sur l’utilisation d’une mesure de similarité (ou de distance) entre objets et entre classes.

Les critères de partitionnement et les mesures de similarités sont des éléments fondamentaux en classification automatique. Nous nous intéressons d’ailleurs à ces points au chapitre 3.

Pour l’heure, nous nous consacrons au troisième aspect essentiel du problème qui est celui de la définition d’une heuristique qui permet de construire une partition approchée de la solution optimale au sens du critère Δ utilisé. En effet, nous avons rappelé en introduction page 6 que le problème de la classification automatique est un problème hautement combinatoire et qu’il fallait en pratique définir des heuristiques afin de traiter le problème pour des bases de données conséquentes malgré la modélisation par programmation linéaire permise par l’AR.

Dans le cadre de cette section, nous aborderons l’étude des méthodes de classification automatique selon une approche théorique.

Dans un premier temps, à partir des éléments développés dans la section précédente, nous définissons un automate à états finis qui met en valeur l’action des opérations de transferts sur les partitions d’un ensemble d’objets.

Nous rappelons ensuite trois heuristiques classiques de classification automatique : les centres mobiles, les classifications hiérarchiques ascendantes et une heuristique nommée AR1 inspirée de l’AR. **Nous montrons que celles-ci peuvent être appréhendées sous l’angle de l’automate des partitions : ces heuristiques utilisent des opérations de transferts simples et chacune correspond à une stratégie particulière d’exploration de l’automate dans la recherche de la solution optimale.**

Cette modélisation des heuristiques étudiées permet de mettre en évidence les limites théoriques de celles-ci. En particulier, nous montrons qu’elles effectuent des chemins qui ne permettent pas d’explorer de manière exhaustive et efficace les états de l’automate. **Par conséquent, il leur est difficile d’approcher la solution optimale.**

Par ailleurs, les centres mobiles nécessitent la fixation *a priori* d’un nombre de classes. Ceci

peut être un inconvénient pratique important si on a aucune idée de cette quantité. Des travaux récents ont cherché à améliorer ce point en ajoutant des opérations de “cassage” permettant de scinder une classe en deux. Nous exposons ainsi les travaux suivants dans un second temps : [PM00], [AE03]. Ces méthodes permettent de relaxer la contrainte rigide du nombre de classes en lui permettant d'appartenir à un intervalle.

Malgré des améliorations, il persiste des inconvénients d'ordre théorique à chacune des approches étudiées. **Ainsi, dans un troisième temps, nous définissons, dans le cadre de variables qualitatives, un nouvel algorithme de classification automatique dont les caractéristiques sont plus avantageuses que les méthodes précédemment énoncées. Cet algorithme est fondé sur une modélisation distincte de ces dernières. En particulier, il utilise une opération d'intersection / union relationnelle¹¹ entre deux classes qui est davantage performante que les transferts puisqu'elle permet soit de fusionner soit de scinder des classes. Cet algorithme que nous nommerons AR2 est la deuxième heuristique inspirée de l'AR. Elle donne, dans la majorité des cas, une meilleure solution approchée au modèle sous forme de programmation linéaire (1.11) que l'algorithme AR1.**

2.2.1 Graphe orienté de l'ensemble des partitions d'un ensemble d'objet

Les opérations de transferts que nous avons définies agissent sur des partitions et permettent de modifier la structure de celles-ci. Nous montrons ici, comment les transferts générateurs permettent en particulier, de définir un graphe orienté sur l'ensemble des partitions $\mathbb{RE}(\mathbb{O})$ d'un ensemble d'objets \mathbb{O} .

La représentation sous forme de graphe orienté que nous donnons s'apparente à la définition d'un automate à états finis dont nous rappelons la définition ci-dessous.

Définition 2.2.1 Automate à états finis

Un automate à états finis A est un quintuplé $(\mathbb{E}, \Sigma, \mu, \mathbb{A}, E^\epsilon)$ où :

- \mathbb{E} est un ensemble fini appelé ensemble des états
- Σ est un ensemble fini appelé alphabet
- $\mu : \mathbb{E} \times \Sigma \rightarrow \mathbb{E}$ est une fonction appelée fonction de transition
- \mathbb{A} est un sous-ensemble de \mathbb{E} (qui peut être vide) appelé ensemble des états acceptant
- E^ϵ est un élément de \mathbb{E} appelé état initial

Les automates à états finis peuvent donc être vus comme étant une classe de graphes orientés dont les arcs sont des éléments d'un ensemble distinct des noeuds. Dans le cas qui nous concerne ici, nous avons les définitions particulières suivantes qui permettent d'introduire le graphe orienté de l'ensemble des partitions :

Définition 2.2.2 Graphe orienté de l'ensemble des partitions d'un ensemble \mathbb{O}

Le graphe orienté de l'ensemble des partitions de \mathbb{O} est un automate $A(\mathbb{RE}(\mathbb{O})) = (\mathbb{E}, \Sigma, \mu, \mathbb{A}, E^\epsilon)$ où :

- $\mathbb{E} = \mathbb{RE}(\mathbb{O}) =$ Ensemble des partitions de l'ensemble \mathbb{O}
- $\Sigma :$
 - $\Sigma = \mathbb{T}(\mathbb{O}) =$ ensemble des opérations de transferts agissant sur les objets
 - $\Sigma = \mathbb{T}(\mathbb{D}) =$ ensemble des opérations de transferts agissant sur les classes d'objets

11. cf chapitre 4, section 4.1 page 115 pour une présentation plus théorique de ces opérations

- μ :
 - $\mu = \circ$: composition à droite si les transferts agissent sur les objets
 - $\mu = \circ$: composition à gauche si les transferts agissent sur les classes d'objets
- $\mathbb{A} = \emptyset$
- $E^\epsilon = \kappa^\epsilon$ la partition triviale à N singletons,

Remarque 2.2.1 A propos des ensembles Σ

L'ensemble Σ des opérations de transferts peut être restreint aux sous-ensembles générateurs.

Remarque 2.2.2 Graphe orienté dans le cadre du codage par affectation

Si les partitions et les transferts sont représentés dans leurs codages par affectation ou de transfert, alors la définition précédente reste valable en prenant :

- $\mu = \cdot$: pré-multiplication matricielle si les transferts agissent sur les objets
- $\mu = \cdot$: post-multiplication matricielle si les transferts agissent sur les classes d'objets

Remarque 2.2.3 Graphe orienté dans le cadre du codage relationnel

Si les partitions sont représentées dans leur codage relationnel, alors la définition précédente reste valable uniquement pour les transferts agissant sur les objets. Dans ce cas, nous avons :

$$\mu(C^k, T^l) = T^l \cdot C^k \cdot {}^t T^l$$

Dans le cas où les transferts agissent sur les classes, l'équivalence n'est plus immédiate. En effet, comme nous l'avons vu en paragraphe 2.1.3.2.3, il existe une dualité "affectation / relationnel". Celle-ci, dans le cas de l'action d'un transfert sur les classes d'une partition, nécessite la connaissance de la matrice d'affectation correspondante à la matrice relationnelle que nous cherchons à modifier. Or ceci n'est pas a priori possible d'une manière rigoureuse dans la mesure où à une matrice relationnelle peut correspondre à une multitude de matrices d'affectation.

Toutefois, la modélisation sous forme d'automate à états finis du graphe des partitions fait intervenir un état initial E^ϵ qui est la matrice identité : $C^\epsilon = I_N$. Nous pouvons alors passer d'une matrice relationnelle C' à une matrice relationnelle C'' de la manière suivante.

Supposons que nous ayons $C' = \mu(C^\epsilon, S^1)$ et $C'' = \mu(C', S^2)$ où $S^1, S^2 \in \mathbb{T}(\mathbb{D})$. Comme nous avons $K' = K^\epsilon \cdot S^1$, nous avons $K'' = K' \cdot S^2 = K \cdot S^1 \cdot S^2$. Nous obtenons C'' ainsi :

$$\begin{aligned} C'' &= K'' \cdot {}^t K'' \\ &= S^1 \cdot S^2 \cdot {}^t(S^1 \cdot S^2) \\ &= S^1 \cdot (S^2 \cdot {}^t S^2) \cdot {}^t S^1 \end{aligned}$$

Nous donnons à présent des résultats simples sur la contiguïté dans l'automate des partitions définis ci-dessus. La démonstration de ces propriétés est relativement simple et s'inscrit dans le cadre de la théorie des graphes.

Propriété 2.2.1 Chemins de transferts de $\mathbb{T}^{sg}(\mathbb{O}) \cup \mathbb{T}^{sg}(\mathbb{D})$ reliant deux partitions

Soient deux partitions \mathcal{R}^k et \mathcal{R}^l d'un même ensemble d'objets \mathbb{O} telles que $\mathcal{R}^k \preceq \mathcal{R}^l$ au sens de l'ordre partiel défini par déf-5.1.12. Il existe alors au moins un chemin constitué de transferts simples générateurs de $\mathbb{T}^{sg}(\mathbb{O})$ ou de $\mathbb{T}^{sg}(\mathbb{D})$ permettant d'aller de la partition \mathcal{R}^k à la partition \mathcal{R}^l .

A contrario, il n'existe aucun chemin constitué d'éléments de $\mathbb{T}^{sg}(\mathbb{O})$ ou de $\mathbb{T}^{sg}(\mathbb{D})$ permettant d'aller de la partition \mathcal{R}^l à la partition \mathcal{R}^k .

Propriété 2.2.2 Chemins de transferts de $\mathbb{T}^{sg}(\mathbb{O})$ reliant deux partitions

Soient deux partitions \mathcal{R}^k et \mathcal{R}^l d'un même ensemble d'objets \mathbb{O} telles que p_l et p_k , leurs nombres de classes respectifs, vérifient $p_l \leq p_k$. Il existe alors au moins un chemin constitué de transferts simples générateurs de $\mathbb{T}^{sg}(\mathbb{O})$ permettant d'aller de la partition \mathcal{R}^k à la partition \mathcal{R}^l .

A contrario, il n'existe aucun chemin constitué d'éléments de $\mathbb{T}^{sg}(\mathbb{O})$ permettant d'aller de la partition \mathcal{R}^l à la partition \mathcal{R}^k .

Propriété 2.2.3 Chemins de transferts de $\mathbb{T}^{dg}(\mathbb{O})$ reliant deux partitions de même distribution
 Soient deux partitions \mathcal{R}^k et \mathcal{R}^l ayant toutes deux la même distribution de classes¹². Il existe alors au moins un chemin constitué de transferts doubles générateurs de $\mathbb{T}^{dg}(\mathbb{O})$, permettant d'aller de la partition \mathcal{R}^k à la partition \mathcal{R}^l et vice versa.

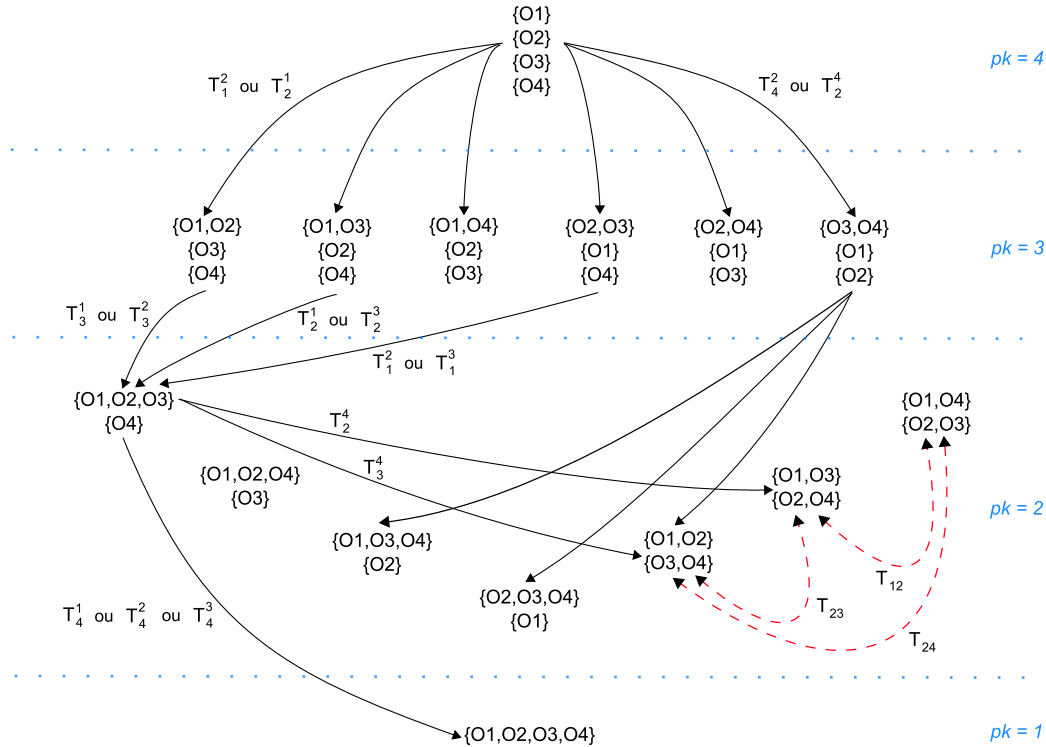


FIGURE 2.2 – Graphe orienté des partitions d'un ensemble d'objets reliées par des transferts générateurs de $\mathbb{T}(\mathbb{O})$

Sur la figure fig-2.2, nous représentons partiellement le graphe orienté des partitions d'un ensemble de 4 objets $\{O^1, O^2, O^3, O^4\}$. Il existe 15 partitions possibles : 1 partition à 4 classes (première partition triviale), 6 partitions à 3 classes, 7 partitions à 2 classes et 1 partition à 1 classe (deuxième partition triviale). Ces partitions sont représentées dans l'ordre croissant de leurs nombres de classes et constituent les noeuds du graphe ou encore les états de l'automate. Deux noeuds sont reliés s'il existe un transfert générateur de $\mathbb{T}^{sg}(\mathbb{O})$ permettant de passer de l'une à l'autre. Nous distinguons deux types de transferts : les transferts simples générateurs représentés par des flèches pleines et les transferts doubles générateurs¹³ représentés par des flèches en pointillées. Compte tenu du nombre important d'arcs, nous ne représentons que quelques uns d'entre eux.

Nous avons toutefois représenté tous les transferts simples générateurs arrivant et partant de la partition en 2 classes $\{\{O^1, O^2, O^3\}; \{O^4\}\}$. De plus, nous avons représenté l'ensemble des transferts doubles générateurs permettant de relier les 3 partitions à 2 classes dont la répartition au sein des classes est uniforme¹⁴.

Nous rappelons qu'il s'agit d'un graphe où sont représentés des transferts générateurs agissant uniquement sur l'ensemble des objets. Ainsi, nous représentons dans la figure fig-2.3, le graphe des partitions du même ensemble d'objets mais dont les arcs correspondent aux transferts générateurs de $\mathbb{T}^{sg}(\mathbb{D})$. Afin de tenir compte de la notion d'équivalence entre deux partitions d'objets définie en déf-2.1.5, nous représentons les partitions par leurs codages algébriques. En effet, il existe plusieurs opérations de transferts permettant de passer d'une partition à une autre qui dépendent en fait de la dénomination des classes auxquelles les ensembles

12. ie même nombre de classes et même répartition du nombre d'objets dans chaque classe
 13. ou transposition d'ordre 1 comme nous le verrons ultérieurement
 14. ie deux objets dans chaque classe

d'objets sont affectés. Ainsi, en raison des nombreuses possibilités de codage d'une partition, nous représenterons sur la figure fig-2.3, un exemple **cohérent** de graphe mettant en évidence l'action des transferts de $\mathbb{T}^{sg}(\mathbb{D})$ sur les partitions. Un transfert d'une classe à une autre peut être vue comme étant l'ensemble des transferts des objets de la première classe vers la deuxième. Nous voyons qu'il y a en pratique moins de transferts agissant sur les classes que sur les objets. En conséquence, nous représentons sur fig-2.3 la totalité des arcs possibles correspondants à l'exemple particulier que nous avons choisi.

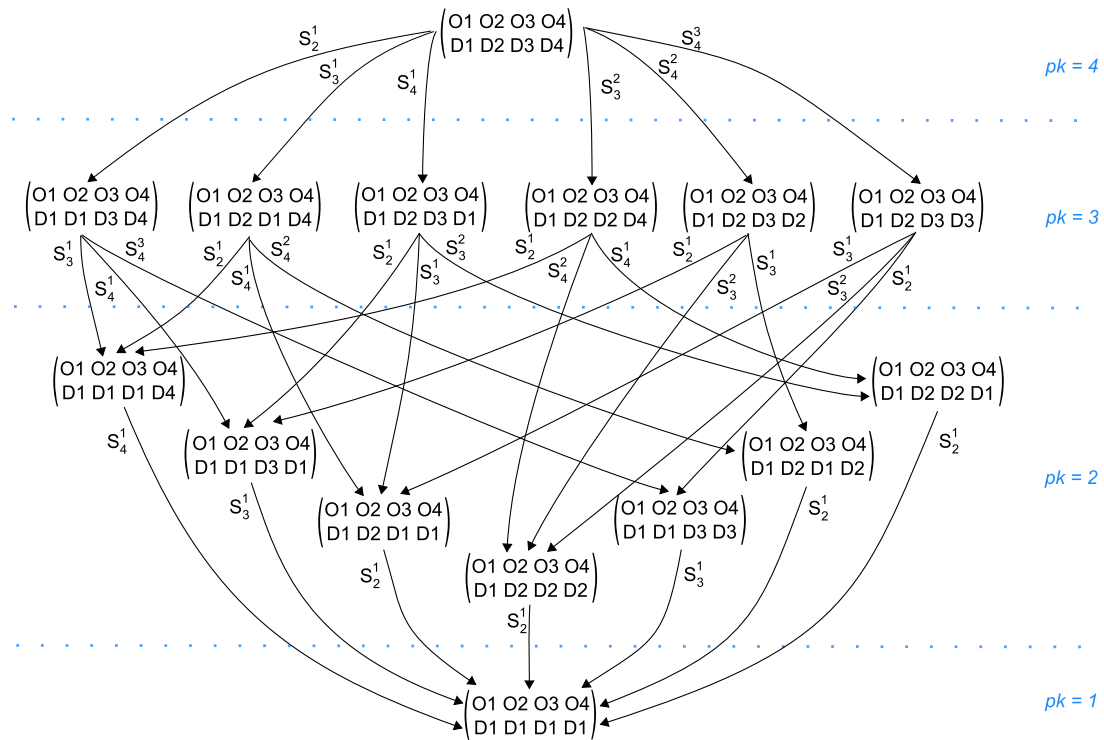


FIGURE 2.3 – Graphe orienté des partitions d'un ensemble d'objets reliées par des transferts générateurs de $\mathbb{T}(\mathbb{D})$

2.2.2 Les heuristiques classiques de classification automatique et leur interprétation au vu du graphe orienté des partitions

Nous présentons quelques heuristiques des plus employées en classification automatique qui sont les méthodes de centres mobiles et les méthodes hiérarchiques ascendantes [Har75]. Nous rappelons aussi une troisième heuristique nommée AR1, qui est basée sur l'AR et qui est proche de l'algorithme "first-leader" [Har75].

Après avoir rappelé brièvement leurs fonctionnements et leurs caractéristiques, nous montrons les liens que ces méthodes ont avec l'automate à états finis défini précédemment pour les partitions. En effet, chaque heuristique correspond à des types de chemins bien particulier dans la recherche de la partition optimale au sein des états¹⁵ de l'automate. Cette modélisation va nous permettre de mettre en évidence les limites théoriques des heuristiques étudiées. En particulier, nous exposerons la difficulté de "casser" une classe existante en deux sous-classes. En effet, comme nous le verrons ci-dessous, les opérations de transfert sur lesquels sont fondées ces approches permettent de réduire le nombre de classes mais pas de l'augmenter.

15. ie les partitions

2.2.2.1 Les méthodes de type centres mobiles

Les méthodes de type centres mobiles ou “ k -means” ont été initialement proposées par E. Forgy [For65]. Elles nécessitent la fixation *a priori* du nombre de classes noté k . Le procédé est itératif et en pratique, s’arrête après un nombre d’itérations fixé par l’utilisateur.

A chaque étape, on calcule la distance séparant chaque objet à chaque centre de gravité des k classes. On affecte ensuite chaque objet à la classe dont le centre de gravité est à distance minimale. On calcule alors les nouveaux centres de gravité de chacune des k classes. On répète ces opérations autant de fois que le nombre d’itérations donné.

On obtient ainsi une partition des objets en k classes. Le critère de partitionnement Δ repose sur l’inertie intra-classe¹⁶. Il s’agit de la moyenne pondérée sur chaque classe des moyennes pondérées des distances deux à deux des objets appartenant à une même classe. En d’autres termes, plus l’inertie intra-classe est faible, plus les classes sous-jacentes sont homogènes puisque les objets appartenant à une même classe sont proches au sens de la distance utilisée. Ainsi, à chaque étape, l’algorithme des centres mobiles permet de diminuer l’inertie intra-classe de la partition dont la qualité au sens de Δ devient meilleure en affectant chaque objet dans la classe qui lui est la plus proche.

Le calcul de base de cet algorithme est celui de la distance entre deux vecteurs qui sont d’une part un objet, d’autre part, un barycentre d’une classe existante. La distance la plus utilisée est la distance euclidienne dont le calcul est de complexité $\mathcal{O}(P)$ où P est le nombre de variables ou descripteurs¹⁷.

Ainsi, si le nombre de variables P est très inférieur au nombre d’objets N à classifier et si de plus le nombre d’itérations est relativement faible ainsi que le nombre de classes fixé, nous pouvons alors considérer, en pratique, que l’algorithme est de complexité $\mathcal{O}(N)$.

Nous donnons l’algorithme de la méthode ci-dessous.

Algorithm 1 Centres mobiles

Require: k = nombre de classes et itr = nombre d’itérations

- Choisir k objets comme premiers centres de gravité de chaque classe

for $l = 1$ to itr **do**

for $i = 1$ to N **do**

for $j = 1$ to k **do**

 - Calculer la distance de l’objet O^i au centre de gravité G^j de la classe courante

end for

 - Affecter O^i à la classe dont le centre de gravité est à distance minimale

 - Mettre à jour le centre de gravité de la classe à laquelle est affectée O^i ainsi que celle à laquelle appartenait O^i avant son nouveau transfert

end for

end for

Au vu du graphe orienté des partitions d’un ensemble d’objets, nous constatons que l’algorithme consiste à effectuer successivement N transferts simples générateurs de $\mathbb{T}^{sg}(\mathbb{O})$ où chaque objet est soit affecté à une nouvelle classe¹⁸, soit réside dans sa classe courante¹⁹. L’algorithme converge si bien que le nombre de transferts d’objets dans des classes distinctes de leur classe courante diminue au fur et à mesure.

Le chemin parcouru dans le graphe est un transfert simple de $\mathbb{T}^s(\mathbb{O})$ qui résulte de la composition de l’ensemble des N transferts simples générateurs de chaque objet. Ce chemin est à chaque étape optimal au sens où la partition d’arrivée est, parmi l’ensemble des partitions

16. nous étudions ultérieurement ces critères dits inertiels au chapitre suivant en sous-section 3.2.2 page 91

17. nombre de colonnes du tableau

18. en l’affectant à un objet appartenant déjà à cette nouvelle classe

19. transfert identité

accessibles à partir de la partition courante par un chemin de longueur N , celle qui est de meilleure qualité²⁰.

Toutefois, comme le nombre de classes est fixé à k , les chemins parcourus par l'algorithme au sein du graphe des partitions se situent uniquement au niveau des partitions à k classes : la méthode ne parcourt pas l'ensemble du graphe. De plus, il existe des chemins que l'algorithme ne peut effectuer de manière efficace. Ce sont des chemins constitués de transferts doubles. Il s'agit d'ailleurs d'une remarque faite par J.F. Marcotorchino dans [MM81]. Celui-ci propose donc d'enrichir les heuristiques à base de transferts d'objets par des opérations de transferts doubles générateurs. Nous précisons toutefois, qu'il s'agit d'opérations qui, si nous les effectuons toutes à une itération donnée, augmenteraient la complexité de l'algorithme à $\mathcal{O}(N^2)$ dans la mesure où elles sont au nombre de $\binom{N}{2}$. De cette façon les transferts doubles ne sont envisageables que pour un petit sous-ensemble d'objets (notion de pivot).

2.2.2.2 Les méthodes hiérarchiques

Parmi les méthodes hiérarchiques, nous nous intéressons aux méthodes ascendantes qui sont celles les plus couramment utilisées. Partant de la partition triviale à N classes, elles consistent à regrouper successivement les deux classes qui sont à distance minimale jusqu'à l'obtention de la partition triviale à 1 classe. A chaque étape, il faut calculer l'ensemble des couples de distances entre les classes constituées.

On obtient un ensemble de partitions emboîtées dont la représentation sous forme d'arbre est appelé dendrogramme ou arbre de classification. C'est alors à l'utilisateur, aidé de différents indicateurs statistiques, de "couper" l'arbre et d'en déduire une partition particulière.

Les distances entre classes qui sont les calculs de base de la classification hiérarchique ascendante, sont le plus souvent des cas particuliers de la formule de Lance et Williams [LW67], ci-dessous :

$$d(D^k, D^j \cup D^{j'}) = \alpha d(D^k, D^j) + \beta d(D^k, D^{j'}) + \gamma d(D^j, D^{j'}) + \delta |d(D^k, D^j) - d(D^k, D^{j'})| \quad (2.2.1)$$

$d(D^k, D^j \cup D^{j'})$ représente la distance entre la classe D^k et la classe constituée par la réunion des classes D^j et $D^{j'}$; $d(D^k, D^j)$ est la distance entre les classes D^k et D^j ; enfin, $\alpha, \beta, \gamma, \delta$ sont quatre réels dont les variations définissent autant de distances différentes.

En particulier, lorsque la distance $d(D^k, D^j)$ est la distance euclidienne entre les centres de gravité des classes respectives, et lorsque nous prenons la configuration suivante, nous obtenons la distance de Ward qui est la plus utilisée et qui permet de diminuer à chaque étape l'inertie intra-classe :

$$\begin{aligned} \alpha &= \frac{\#D^k + \#D^j}{\#D^k + \#D^j + \#D^{j'}} \\ \beta &= \frac{\#D^k + \#D^{j'}}{\#D^k + \#D^j + \#D^{j'}} \\ \gamma &= -\frac{\#D^k}{\#D^k + \#D^j + \#D^{j'}} \\ \delta &= 0 \end{aligned} \quad (2.2.2)$$

où $\#\mathbb{A}$ est le cardinal de l'ensemble \mathbb{A}

Supposons que le calcul de distance soit de complexité $\mathcal{O}(P)$ où P est le nombre de variables ou descripteurs, ce qui est en général le cas. Alors si P est très inférieur à N , le nombre d'objets à classer, l'algorithme de classification hiérarchique ascendante a une complexité en $\mathcal{O}(N^2)$ étant donné qu'il est nécessaire de calculer à chaque étape les distances deux à deux entre classes constituées.

20. au sens de l'inertie intra-classe que nous rappellerons page 91

Algorithm 2 Hiérarchie ascendante

```

-  $k = N$  où  $k$  est le nombre de classes
while  $k > 1$  do
  for  $j = 1$  to  $k$  do
    for all  $j' > j$  do
      - Calculer la distance entre les classes  $D^j$  et  $D^{j'}$ 
    end for
  end for
  - Réunir les deux classes à distance minimale
  -  $k \leftarrow k - 1$ 
end while

```

L'algorithme décrit un chemin qui s'inscrit dans un graphe orienté où les arcs correspondent à des transferts simples générateurs de $\mathbb{T}^{sg}(\mathbb{D})$ tel que sur la figure fig-2.3. En particulier, un dendogramme correspond à un chemin du graphe orienté partant de la partition triviale à N classes et "remontant" vers la partition triviale à 1 classe. A chaque étape, l'algorithme consiste à déterminer le transfert simple générateur de $\mathbb{T}^{sg}(\mathbb{D})$ qui permet de fusionner les deux classes à distance minimale.

2.2.2.3 Les méthodes inspirées de l'AR et de la similarité régularisée

L'heuristique présentée est initialement dédiée aux données qualitatives et tente d'approcher la partition optimale modélisée au sens de l'AR par le modèle de programmation linéaire défini par (1.11), page 7. Nous rappelons ci-dessous ce modèle dans le cas de données complètes²¹ et étendu à une version paramétrée par un seuil de similarité α :

$$\max_X \text{Condorcet}(\mathbf{C}, X) = \sum_{i=1}^N \sum_{i'=1}^N \left(\mathbf{C}_{ii'} - \alpha \frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) X_{ii'} \quad (2.2.3)$$

$$\begin{aligned} X_{ii'} &\in \{0, 1\} && \text{(binarité)} \\ X_{ii} &= 1 && \forall i = 1, \dots, N \quad \text{(réflexivité)} \\ X_{i'i} &= X_{ii'} && \forall i, i' = 1, \dots, N \quad \text{(symétrie)} \\ X_{ii'} + X_{i'i''} - X_{ii''} &\leq 1 && \forall i, i', i'' = 1, \dots, N \quad \text{(transitivité)} \end{aligned}$$

où M est le nombre de variables qualitatives et α est un seuil de similarité compris entre 0 et 1 qui est fixé par défaut à 0.5 ce qui, dans le cadre de l'AR, représente un seuil théorique²² lié à la notion de majorité²³ et propre au critère de Condorcet.

Cette heuristique ne nécessite pas de fixation d'un nombre de classes mais il est cependant requis un nombre de classes maximal au delà duquel, la création de nouvelles classes n'est plus possible.

A chaque itération, l'algorithme parcourt chaque objet. On calcule la contribution de l'objet courant avec toutes les classes constituées. Soit un objet O^i et une classe \mathcal{R}_k , la contribution de O^i avec \mathcal{R}_k vaut :

$$\sum_{O^{i'} \in \mathcal{R}_k} \mathbf{C}_{ii'} = \sum_{O^{i'} \in \mathcal{R}_k} \left(\mathbf{C}_{ii'} - \alpha \frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) \quad (2.2.4)$$

où $\mathbf{C}_{ii'} = \left(\mathbf{C}_{ii'} - \alpha \frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right)$ est la contribution atomique de la mise en relation entre les objets O^i et $O^{i'}$.

21. ie sans données manquantes
22. appelé aussi borne de Solomon et Fortier [SF66]
23. cf page 7

Parmi l'ensemble des contributions de l'objet courant O^i avec l'ensemble des classes existantes, on retient celle qui est la plus forte ainsi que l'identité de la classe correspondante. Si cette contribution est négative et si le nombre de classes constituées est inférieur au nombre maximal de classes autorisé, on crée une nouvelle classe dont l'objet courant est le premier élément. *A contrario*, si la contribution est positive, ou si le nombre de classes constituées est égal au nombre maximal de classes autorisé, on affecte l'objet courant à la classe avec laquelle il partage cette contribution maximale. Le critère que l'on cherche à maximiser est celui de Condorcet mais celui-ci peut être associé par ailleurs, à des similarités régularisées telles que définies dans [BM98].

L'opération de base de cet algorithme est donc la mesure donnée par (2.2.4) de la contribution d'un objet O^i avec une classe \mathcal{R}_k . Celle-ci nécessite *a priori* $\#\mathcal{R}_k P$ opérations où P est le nombre total de modalités toute variable qualitative confondue. En effet, il est nécessaire de calculer la contribution atomique de l'objet O^i avec chaque objet $O^{i'}$ appartenant à la classe \mathcal{R}_k . La complexité du calcul d'une contribution atomique $C_{ii'}$, est en $\mathcal{O}(P)$. Ainsi pour évaluer la contribution de O^i avec une classe \mathcal{R}_k il faut bien $\#\mathcal{R}_k P$ opérations.

En fait, en écrivant les mesures de similarités $C_{ii'}$ tels des produits scalaires entre vecteurs binaires \vec{O}^i de terme général $O_j^i = \mathbf{K}_{ij}$, $j = 1, \dots, P$, où \mathbf{K} est le tableau disjonctif complet; nous pouvons employer les propriétés de linéarité du produit scalaire afin de mettre en évidence la réduction de la complexité du calcul de la contribution de O^i avec une classe \mathcal{R}_k de $\mathcal{O}(\#\mathcal{R}_k P)$ à $\mathcal{O}(P)$. Nous avons effectivement le résultat suivant :

$$\begin{aligned} \sum_{O^{i'} \in \mathcal{R}_k} C_{ii'} &= \sum_{O^{i'} \in \mathcal{R}_k} \left(C_{ii'} - \alpha \frac{C_{ii'} + C_{i'i'}}{2} \right) \\ &= \sum_{O^{i'} \in \mathcal{R}_k} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \alpha \frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) \\ &= \left(\langle \vec{O}^i, \sum_{O^{i'} \in \mathcal{R}_k} \vec{O}^{i'} \rangle - \sum_{O^{i'} \in \mathcal{R}_k} \alpha \frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) \\ &= \left(\langle \vec{O}^i, \vec{G}^k \rangle - \alpha \frac{(\#\mathcal{R}_k) \langle \vec{O}^i, \vec{O}^i \rangle + S(\vec{G}^k)}{2} \right) \end{aligned} \quad (2.2.5)$$

où $\vec{G}^k = \sum_{O^i \in \mathcal{R}_k} \vec{O}^i$, représentant de la classe \mathcal{R}_k , est à une constante $1/\#\mathcal{R}_k$ près, le barycentre de la classe \mathcal{R}_k ; et où $S(\vec{G}^k) = \sum_j G_j^k = \sum_{O^i \in \mathcal{R}_k} \langle \vec{O}^i, \vec{O}^i \rangle$ en raison de la binarité des vecteurs.

La somme sur les objets $O^{i'} \in \mathcal{R}_k$ disparaît et il ne reste que l'évaluation de quantités calculables à partir de P opérations élémentaires.

Suite à ces propriétés, nous pouvons dire que l'heuristique AR1 a globalement une complexité en $\mathcal{O}(N)$ à condition que le nombre de classes maximal et le nombre de modalités soient inférieurs au nombre d'individus et que le nombre d'itérations soit faible.

Chaque opération effectuée par l'algorithme consiste à se déplacer pas à pas sur le graphe des partitions en appliquant successivement des transferts simples générateurs de $\mathbb{T}^{sg}(\mathbb{O})$. L'ordre dans lequel sont triés les objets au départ influe sur le résultat car, à chaque itération, nous gardons toujours ce même ordre de traitement. Ainsi, peu importe la classe à laquelle appartient l'objet O^i , à chaque itération, le i -ème transfert concernera celui de cet objet. Comme le transfert d'un objet change la structure d'une partition, nous voyons qu'un ordre différent à de fortes chances de déboucher sur un résultat différent.

L'originalité de cette approche est la possibilité de créer des classes. Cependant, en pratique, le nombre de classes évolue très peu après la première itération. Dans la majorité des cas, la première partition constituée est fondamentale puisqu'elle constitue une partition initiale qui ne changera que peu dans la suite du processus. Ainsi, les itérations suivantes consistent en des raffinements de la partition initiale par re-transferts d'objets dans des classes "déjà" constituées.

Nous pouvons donc considérer que la méthode emploie des transferts simples générateurs mais parcourt, contrairement aux centres mobiles, pas à pas le graphe des partitions. Nous pourrions toutefois, de la même manière que pour les méthodes de type centres mo-

Algorithm 3 Heuristique AR1

Require: $kmax$ = Nombre de classes maximal, $seuil$ = Seuil de similarité et itr = nombre d'itérations

- Prendre le premier objet comme élément de la première classe constituée
- $k = 1$ où k est le nombre de classes

for $l = 1$ to itr **do**

for $i = 1$ to N **do**

for $j = 1$ to m **do**

- Calculer la contribution entre l'objet O^i et la classe \mathcal{R}_k

end for

- $k^* \leftarrow$ l'indice de la classe avec laquelle O^i a la plus forte contribution
- $s^* \leftarrow$ la contribution entre O^i et la classe avec laquelle il a la plus forte contribution

if $s^* < seuil \wedge m < kmax$ **then**

- Créer une nouvelle classe dont O^i est le premier élément
- $m \leftarrow m + 1$

else

- Affecter O^i à la classe avec laquelle il a la plus forte contribution

end if

end for

end for

biles, enrichir cette approche en autorisant des opérations de transferts doubles [MM81] mais pour des petits sous-ensembles d'objets²⁴.

Nous faisons de plus remarquer, que **cette méthode peut utiliser, sous certaines conditions, d'autres critères d'association maximale**²⁵ Δ **que celui de Condorcet**. Notons de manière générique le critère $\Delta(C, X)$ de la manière suivante :

$$\Delta(C, X) = \sum_{i=1}^N \sum_{i'=1}^N C_{ii'} X_{ii'} \quad (2.2.6)$$

où $C_{ii'}$ représente la contribution de la mise en relation des objets O^i et $O^{i'}$ dans une même classe. Il est alors nécessaire, en pratique, que le calcul de base de la mesure de la contribution d'un objet avec une classe puisse être calculée avec une complexité de $O(P)$.

Enfin, nous rapprochons cette sous-section du paragraphe 3.2.1.2 du chapitre 3 page 89 où nous proposons une **extension des critères d'association maximal à des données non binaires**. De cette manière, l'heuristique décrite ici dans le cadre de variables qualitatives, pourrait être étendue, par l'intermédiaire d'interprétations géométriques²⁶, à des variables quantitatives. Nous montrerons par ailleurs, le bien fondé des extensions en montrant notamment les liens étroits entre le critère de Condorcet et le critère de la différence inertielle en sous-section 3.2.2.

2.2.3 Les nouvelles heuristiques de type centres mobiles et l'opération de "cassage" des classes

Nous nous intéressons ici plus particulièrement aux méthodes de type centres mobiles. Nous remarquons deux caractéristiques de cette méthode de classification automatique :

- la fixation à k du nombre de classes implique que les chemins effectués par l'algorithme sont locaux aux partitions au nombre de classes constant k

24. cf remarque concernant les centres mobiles

25. que nous étudions dans le chapitre 3 section 3.2

26. ie en termes de produit scalaire

- la méthode décrit des chemins constitués de transferts simples. Or ces derniers permettent de diminuer le nombre de classes mais pas de l'augmenter

Nous cherchons à mettre en évidence le caractère peu "naturel" de l'opération de "cassage" de classes qui permettrait d'augmenter le nombre de classes d'une partition. Au regard de l'interprétation de ces méthodes en terme de graphe orienté de partitions, cette remarque correspond au fait que les transferts simples sont non bijectifs. Quant aux transferts doubles, bien qu'étant bijectifs, ils relient des partitions ayant la même distribution. En conséquence, les chemins que décrivent les méthodes de type centres mobiles dans le graphe de partitions ne peuvent pas aller vers une partition avec un nombre de classes plus grand.

Il apparaît alors qu'un moyen d'améliorer les performances en qualité de ces heuristiques, est de définir des opérations de "cassage" de classes permettant de se libérer de la fixation rigide du nombre de classes.

Récemment, de nombreux travaux ont cherché à définir ces opérations de "cassage". Nous en étudions succinctement 2 d'entre elles qui sont respectivement issues des articles [PM00] et [AE03].

Ces deux algorithmes reposent sur le même processus que nous décrivons dans l'algorithme 4.

Algorithm 4 Algorithmes des centres mobiles avec opération de "cassage"

Require: $kmin$ = Nombre de classes minimal, $kmax$ = Nombre de classes maximal

- $k \leftarrow kmin$

while $k \leq kmax$ **do**

- Classifier les objets avec un algorithme classique des centres mobiles pour lequel le nombre de classes vaut k

for $l = 1$ to k **do**

- Classifier localement les objets de la classe D^l avec un algorithme classique des centres mobiles pour lequel le nombre de classes vaut 2

- **Tester** si le modèle à 2 classes est meilleur que celui à 1 classe

if le modèle à 2 classes est meilleur que celui à 1 classe **then**

- Mettre à jour la partition

- $k \leftarrow k + 1$

end if

end for

end while

La première extension est proposée par D. Pelleg et A. Moore. Leur algorithme se nomme "X-means". L'étape de cassage définie par ces auteurs consiste à calculer un score afin de déterminer le meilleur modèle à retenir. La méthode retenue par ces auteurs est le "BIC scoring" ("Bayesian Information Criterion") basé sur la vraisemblance sous l'hypothèse de gaussianité. Ce score est utilisé afin de déterminer localement si une classe constituée doit être séparée en deux ou non. Il est également employé afin de déterminer la meilleure partition parmi toutes celles déterminées à chaque itération du processus.

La deuxième approche est due à G. Amerly et C. Elkan et se nomme "G-means". L'étape de "cassage" repose sur l'hypothèse que les objets appartenant à une même classe doivent être distribués autour de leur centre de gravité selon une loi gaussienne multidimensionnelle. Pour chaque classe, on calcule une statistique de test de Anderson-Darling. Afin de réduire les temps de calcul de ces statistiques, G. Amerly et C. Elkan proposent de projeter le sous-nuage de points constitué des objets de la classe étudiée sur un sous-espace à une dimension. Ce sous-espace est la droite reliant les deux centres de gravités G^1 et G^2 établis par l'application de l'algorithme classique de centres mobiles sur les objets de la classe courante.

Après avoir choisi un niveau de test, on identifie si la statistique de test basée sur les données projetées se trouve dans la région critique ou non. Si l'hypothèse qu'une loi gaussienne est rejetée, la classe est symboliquement "cassée" en remplaçant le centre de gravité initial par les deux centres de gravité G^1 et G^2 .

Nous voyons qu'il n'existe pas une unique façon de "casser" des classes constituées alors que l'opération de transfert semble simple et unique. A chaque étape en effet, nous cherchons

le transfert optimal au sens des critères inertiels. *A contrario*, lorsqu'il s'agit de casser une classe, on peut utiliser plusieurs critères qui repose sur différentes hypothèses puisque les critères inertiels favorisent systématiquement les partitions dont le nombre de classe est grand.

L'avantage de ces nouvelles méthodes est qu'elles permettent de relaxer la contrainte de la fixation du nombre de classes. L'inconvénient par contre, est d'ordre théorique. Il réside dans le fait **qu'on utilise un critère de cassage particulier si bien qu'on ne sait plus exactement quel critère initial la partition obtenue tente d'optimiser.**

Forts de ces constats, nous définissons dans la sous-section suivante un nouvel algorithme qui ne nécessite pas la fixation d'un nombre de classes et qui ne possèdent pas les différents inconvénients exposés.

2.2.4 Une nouvelle heuristique de l'application de l'AR en classification automatique

Nous présentons ici, un nouvel algorithme de classification automatique pour données qualitatives que nous nommerons AR2 et qui est issu d'un travail en commun avec J. Lemoine [APLB05], [LBAP06]. Cet algorithme cherche de la même manière que l'heuristique AR1, à approcher au mieux la solution donnée par le modèle suivant :

$$\max_X \text{Condorcet}(\mathbf{C}, X) = \sum_{i=1}^N \sum_{i'=1}^N \left(\mathbf{C}_{ii'} - \alpha \frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) X_{ii'} \quad (2.2.7)$$

$$\begin{array}{lll} X_{ii'} \in \{0, 1\} & & \text{(binarité)} \\ X_{ii} = 1 & \forall i = 1, \dots, N & \text{(réflexivité)} \\ X_{i'i} = X_{ii'} & \forall i, i' = 1, \dots, N & \text{(symétrie)} \\ X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 & \forall i, i', i'' = 1, \dots, N & \text{(transitivité)} \end{array}$$

où M est le nombre de variables qualitatives et α est un seuil de similarité compris entre 0 et 1 qui est fixé par défaut à 0.5.

L'approche de cette heuristique est complètement différente des méthodes précédentes étant donné qu'elle repose sur une modélisation distincte. En effet, nous avons introduit le graphe orienté des partitions et nous avons montré que les trois heuristiques présentées ci-dessus utilisaient des opérations de transferts simples. En particulier, nous avons cherché à mettre en évidence les limites théoriques de ce genre d'opérations de base : les transferts simples ne permettent pas de casser des classes ce qui limite l'exploration des noeuds du graphe et donc les chances d'atteindre la partition optimale.

L'algorithme que nous exposons est basé sur une **opération d'intersection / union** de classes que nous noterons opération I/U. Celle-ci consiste à déterminer, lorsque deux classes sont en intersection pour un sous-ensemble d'objets, la configuration se ramenant à une sous-partition²⁷ contribuant le plus au critère.

Nous introduisons en détail cette opération dans le cadre du critère de Condorcet. Puis nous donnons l'algorithme dans sa généralité.

Nous avons représenté dans la figure fig-2.4 deux classes, notées D^k et D^j , qui sont en intersection pour un ensemble d'objets noté $D^c = D^k \cap D^j$.

A partir de ces classes en intersection, il existe 5 différentes possibilités de se ramener à une partition. Pour cela, introduisons les classes atomiques disjointes suivantes :

^{27.} ie un sous-ensemble de classes disjointes de ce sous-ensemble d'objets

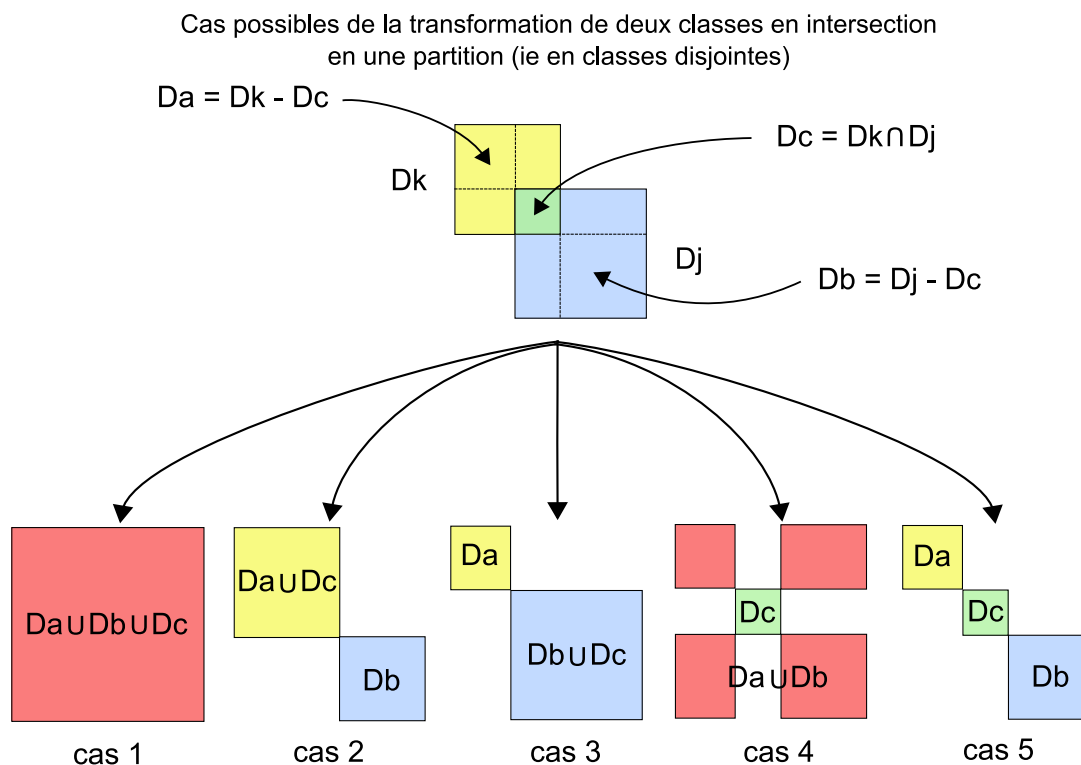


FIGURE 2.4 – Les 5 possibilités de partitions à partir de l'intersection de 2 classes

- $D^c = D^k \cap D^j$
- $D^a = D^k - D^c$
- $D^b = D^j - D^c$

où $D^j - D^{j'} = D^j \cap \overline{D^{j'}}$, est l'ensemble des objets appartenant à D^j mais n'appartenant pas à $D^{j'}$.

Au vu de ces notations, nous pouvons avoir les 5 configurations suivantes qui sont par ailleurs illustrées dans la figure fig-2.4 :

1. 1 partition en une classe : $\{D^a \cup D^b \cup D^c\}$
- 3 partitions en 2 classes disjointes :
 2. $\{D^a \cup D^c\}, \{D^b\}$
 3. $\{D^a\}, \{D^b \cup D^c\}$
 4. $\{D^a \cup D^b\}, \{D^c\}$
5. 1 partition en 3 classes disjointes : $\{D^a\}, \{D^b\}, \{D^c\}$

Considérons alors la matrice C des contributions dont le terme général relatif au critère de

Condorcet étendu par l'ajout d'un paramètre $\alpha \in [0, 1]$, est donné par, $\forall i, i' = 1, \dots, N$:

$$C_{ii'} = \left(C_{ii'} - \alpha \frac{C_{ii} + C_{i'i'}}{2} \right) \quad (2.2.8)$$

Supposons que les classes en intersection concernent un sous-ensemble d'objets de \mathbb{O} que nous noterons $\mathbb{O}_{D^k \cup D^j}$. Nous cherchons à savoir, sans remettre en cause les classes atomiques qui sont disjointes, quelle configuration parmi les 5 possibilités énumérées précédemment, maximise au mieux le critère sur le sous-ensemble $\mathbb{O}_{D^k \cup D^j}$. Pour cela introduisons les quantités suivantes :

- L_{ab} le lien entre les classes D^a et D^b
- L_{ac} le lien entre les classes D^a et D^c
- L_{bc} le lien entre les classes D^b et D^c

Le lien entre deux classes disjointes D^k et $D^{k'}$ est la mesure de la contribution supplémentaire apportée par la réunion des classes D^k et $D^{k'}$. Nous avons donc l'équation suivante :

$$L_{kk'} = \sum_{O^i \in D^k} \sum_{O^{i'} \in D^{k'}} C_{ii'} \quad (2.2.9)$$

$$= \sum_{O^i \in D^k} \sum_{O^{i'} \in D^{k'}} \left(C_{ii'} - \alpha \frac{C_{ii} + C_{i'i'}}{2} \right) \quad (2.2.10)$$

Représentations des liens entre les classes atomiques $D_a = D_k - D_c$; $D_c = D_k \cap D_j$; $D_b = D_j - D_c$

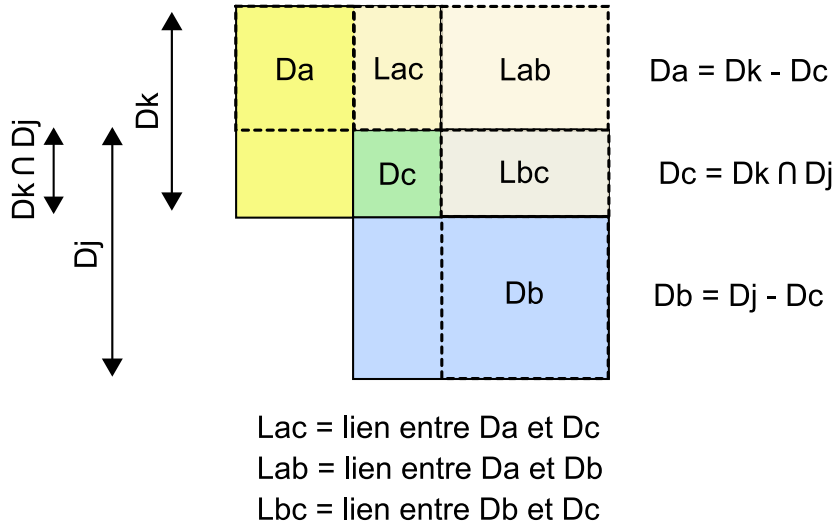


FIGURE 2.5 – Liens entre les classes atomiques issus de l'intersection de 2 classes

Sur la figure fig-2.5 supposons que la matrice représentée soit celle de la restriction de la matrice des contributions C aux objets appartenant à $\mathbb{O}_{D^k \cup D^j}$. Nous voyons alors que les liens entre classes atomiques sont représentés par des rectangles. Prenons, sans perte de généralité, le cas de D^a et D^b . Nous voyons qu'il est nécessaire d'évaluer $\#D^a \#D^c$ termes $C_{ii'}$ avec $(O^i, O^{i'}) \in D^a \times D^c$ et de les sommer afin d'obtenir la valeur de L_{ac} . Si l'évaluation d'une seule contribution $C_{ii'}$ coûte P opérations élémentaires où P est le nombre total de modalités alors la complexité de l'évaluation de L_{ac} est de l'ordre de $\mathcal{O}(\#D^a \#D^c P)$ ce qui se rapproche d'une complexité en $\mathcal{O}(N^2 P)$ plutôt que $\mathcal{O}(P)$.

En fait, il est possible de réduire la complexité du calcul du lien $L_{kk'}$ entre deux classes disjointes D^k et $D^{k'}$ en considérant des représentants de classe. Pour cela reprenons les notations employées pour la description de l'algorithme AR1. Posons alors, \vec{O}^i le vecteur de composante O_1^i, \dots, O_P^i avec $O_j^i = \mathbf{K}_{ij}, \forall i = 1, \dots, N$ et $\forall j = 1, \dots, P$. Dans ce cas, $\mathbf{C}_{ii'} = \langle \vec{O}^i, \vec{O}^{i'} \rangle$. Nous pouvons alors étendre la simplification donnée en 2.2.5 précédemment lors de l'étude de l'algorithme AR1 au cas plus général suivant :

$$\begin{aligned} L_{kk'} &= \sum_{O^i \in D^k} \sum_{O^{i'} \in D^{k'}} \left(\mathbf{C}_{ii'} - \alpha \frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) \\ &= \sum_{O^i \in D^k} \sum_{O^{i'} \in D^{k'}} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \alpha \frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) \\ &= \left(\langle \sum_{O^i \in D^k} \vec{O}^i, \sum_{O^{i'} \in D^{k'}} \vec{O}^{i'} \rangle - \alpha \frac{\#D^{k'} \sum_{O^i \in D^k} \langle \vec{O}^i, \vec{O}^i \rangle + \#D^k \sum_{O^{i'} \in D^{k'}} \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) \quad (2.2.11) \\ &= \left(\langle \vec{G}^k, \vec{G}^{k'} \rangle - \alpha \frac{\#D^{k'} S(\vec{G}^k) + \#D^k S(\vec{G}^{k'})}{2} \right) \end{aligned}$$

où $G^k = \sum_{O^i \in D^k} \vec{O}^i$ et $G^{k'} = \sum_{O^{i'} \in D^{k'}} \vec{O}^{i'}$ sont respectivement les représentants des classes D^k et $D^{k'}$ et où $S(\vec{G}^k) = \sum_{O^i \in D^k} \langle \vec{O}^i, \vec{O}^i \rangle = \sum_j G_j^k$ en raison de la binarité des vecteurs.

Cette "astuce" nous permet donc de réduire la complexité des opérations de base de l'algorithme AR2 à $\mathcal{O}(P)$.

Nous nous apercevons alors que la meilleure configuration de partition est :

1. $\{D^a \cup D^b \cup D^c\}$; si $\max(L_{ac} + L_{bc} + L_{ab}, L_{ac}, L_{bc}, L_{ab}, 0) = L_{ab} + L_{ac} + L_{bc}$
2. $\{D^a \cup D^c\}, \{D^b\}$; si $\max(L_{ac} + L_{bc} + L_{ab}, L_{ac}, L_{bc}, L_{ab}, 0) = L_{ac}$
3. $\{D^a\}, \{D^b \cup D^c\}$; si $\max(L_{ac} + L_{bc} + L_{ab}, L_{ac}, L_{bc}, L_{ab}, 0) = L_{bc}$
4. $\{D^a \cup D^b\}, \{D^c\}$; si $\max(L_{ac} + L_{bc} + L_{ab}, L_{ac}, L_{bc}, L_{ab}, 0) = L_{ab}$
5. $\{D^a\}, \{D^b\}, \{D^c\}$; si $\max(L_{ac} + L_{bc} + L_{ab}, L_{ac}, L_{bc}, L_{ab}, 0) = 0$

Nous venons de présenter l'opération de base de notre algorithme qui, contrairement aux transferts simples, permet soit de diminuer le nombre de classes (cas 1.), soit de conserver le même nombre de classes (cas 2., 3., 4.), soit encore d'augmenter le nombre de classes (cas 5.).

Il reste à préciser le mécanisme général : pourquoi y aurait-il des classes en intersection ?

Dans les notations de cette sous-section, nous avons notamment noté par D^j une classe plutôt que par \mathcal{R}_j , contrairement à AR1. Ceci n'est pas anodin puisqu'en fait, les classes D^j que nous utilisons afin de construire une partition approchant la solution optimale du modèle en programmation linéaire sont les classes engendrées par les descripteurs (ou modalités) $D^j, j = 1, \dots, P$; issus des variables qualitatives. L'intuition est la suivante : si dans la solution optimale les objets O^i et $O^{i'}$ sont dans la même classe alors il doit exister au moins un descripteur D^j que partage ces deux objets.

Ainsi, ce sont les descripteurs et plus précisément, les classes d'objets qu'ils engendrent, qui permettent de construire la partition solution. Or des descripteurs provenant de différentes variables engendrent des classes en intersection²⁸. De cette manière, une partition courante $\mathcal{R} = \{D^k\}_{k=1, \dots, p}$ où p est le nombre de classes, est constamment remise en cause par l'arrivée d'un nouveau descripteur D^j . Grâce à l'opération I/U, nous déterminons à partir de la partition courante en intersection avec ce nouveau descripteur, la partition nouvelle maximisant au mieux le critère de Condorcet.

Il est nécessaire à présent de préciser deux points : l'ordre de prise en compte des descrip-

²⁸. par exemple la modalité "avoir les cheveux blonds" peut être en intersection avec la modalité "avoir les yeux bleus"

teurs et la récursivité due au fait qu'un descripteur peut être en intersection avec plusieurs classes de la partition courante.

Il existe en effet une dépendance eu égard à l'ordre de prise en compte des descripteurs : nous voyons que changer cet ordre conduirait à une solution finale différente. Nous nous basons alors sur la notion de **qualité d'un descripteur** D^j noté Q_j , qui est en fait le lien de D^j avec lui même ce qui représente la contribution au critère de la classe engendrée par D^j :

$$\begin{aligned}
 Q_j &= L_{jj} \\
 &= \sum_{O^i \in D^j} \sum_{O^{i'} \in D^j} C_{ii'} \\
 &= \sum_{O^i \in D^j} \sum_{O^{i'} \in D^j} \left(C_{ii'} - \alpha \frac{C_{ii} + C_{i'i'}}{2} \right) \\
 &= \left(\langle \vec{G}^j, \vec{G}^j \rangle - \alpha \#D^j S(\vec{G}^j) \right) \tag{2.2.12}
 \end{aligned}$$

Nous trions alors en amont les descripteurs selon leur qualité ce qui nous donne un ordre de prise en compte de ces derniers. En pratique, opter pour l'ordre décroissant ou l'ordre croissant influe peu sur le résultat final. Cependant, préférer l'ordre décroissant permet d'avoir une partition stable plus rapidement ce qui conduit à des temps de traitement plus courts que pour l'ordre croissant.

Lorsqu'un descripteur est en intersection avec plusieurs classes de la partition courante, il existe un deuxième ordre de prise en compte à étudier : quelle intersection l'opération I/U doit-elle traiter en premier ?

Il nous a alors semblé naturel de trier les intersections selon leur cardinal et d'effectuer ainsi en premier les opérations I/U correspondant aux intersections de cardinaux les plus grands.

Par ailleurs, lorsque nous effectuons la première opération I/U, nous changeons localement à la fois la classe de la partition concernée et le descripteur. De ce fait, lorsque nous appliquons l'opération de base à une deuxième classe en intersection de la partition courante, le descripteur a changé. Nous donnons ainsi le descripteur modifié qui est retourné à la suite d'une opération I/U selon les différents cas. Pour fixer les idées, nous noterons par un indice k les classes appartenant à la partition et par un indice j le descripteur en intersection avec les classes de la partition. De ce fait, D^a, D^b, D^c , correspondent aux classes atomiques de l'intersection de D^k avec D^j . D'un point de vue ensembliste voici le devenir des éléments D^j après l'opération I/U :

1. $D^j \leftarrow D^j \cup D^a$
2. $D^j \leftarrow D^j - D^c$
3. $D^j \leftarrow D^j$
4. $D^j \leftarrow (D^j - D^c) \cup D^a$
5. $D^j \leftarrow D^j - D^c$

Voici maintenant, le devenir de la classe D^k de la partition suite à l'opération I/U :

1. $D^k \leftarrow \emptyset$
2. $D^k \leftarrow D^k$
3. $D^k \leftarrow D^k - D^c$

4. $D^k \leftarrow D^k - D^a$

5. $D^k \leftarrow D^k - D^c$ et création d'une nouvelle classe $D^{k'} = D^c$

Nous avons décrit, les étapes effectuées lorsque nous calculons la partition \mathcal{R}' issue de l'intersection $\mathcal{R} \cap D^j$ d'une partition courante et d'un descripteur. Dans le cadre du formalisme relationnel, cette partition, est telle que $\text{Condorcet}(\mathbf{C}, X') \geq \text{Condorcet}(\mathbf{C}, X)$ où X' est la matrice relationnelle représentative de \mathcal{R}' . Nous itérons ces différentes étapes pour l'ensemble des descripteurs et nous aboutissons à une partition finale qui en pratique est très proche de l'optimal si celui-ci n'est pas atteint.

Nous donnons ci-dessous un "sketch" de l'algorithme général de l'heuristique AR2. Nous ne donnerons pas en effet les détails de certains calculs telle que la détermination du cardinal de l'intersection de D^j avec les classes D^k de \mathcal{R} . Nous précisons néanmoins, le fait que l'algorithme utilise des tris de complexité²⁹ linéaire étant donné que nous sommes dans le contexte particulier pour lequel cette réduction de complexité est permise [CLR02].

Algorithm 5 Heuristique AR2

```

- Calculer la qualité  $Q_j$  de chaque descripteur  $D^j$ 
- Trier ces qualités selon l'ordre décroissant
for  $j = 1$  to  $P$  do
  - Calculer le cardinal des intersections de la classe  $D^j$  avec les classes  $D^k$  de la partition  $\mathcal{R}$ 
  - Trier ces cardinaux selon l'ordre décroissant
  while  $\bigcup(D^k \cap D^j) \neq \emptyset$  do
    - Effectuer l'opération I/U entre  $D^k$  et  $D^j$ 
    - Effectuer les mises à jour concernant les représentants  $\vec{G}^k$  et  $\vec{G}^j$ 
    - Effectuer les mises à jour concernant la partition  $\mathcal{R} = \{D^k\}$ 
  end while
end for

```

Nous donnons dans la figure fig-2.6, une représentation graphique de l'algorithme proposé.

Nous supposons comme pour les heuristiques présentées précédemment, que le nombre P de descripteurs est très inférieur au nombre N d'objets. Nous faisons de plus l'hypothèse que le nombre d'intersections entre une classe D^j et une partition \mathcal{R} est aussi très inférieur à N . Dans ce cas, nous pouvons dire que la complexité de l'algorithme AR2 est en $\mathcal{O}(N)$.

Nous pouvons également faire la même remarque que pour AR1 : l'algorithme peut utiliser d'autres critères $\Delta(\mathbf{C}, X)$ à condition que les calculs de liens issus des doubles sommes sur des rectangles de la matrice des contributions puissent être effectués avec une complexité en $\mathcal{O}(P)$.

Notons enfin que l'algorithme nécessite des classes d'objets génératrices afin de calculer la partition solution. Si cet algorithme est clairement adapté à des données binaires telles que les variables qualitatives, il n'en va pas de même pour les variables quantitatives. Cependant, nous pouvons créer des classes d'objets en discrétisant les variables quantitatives. Par ailleurs, contrairement aux approches consistant à se ramener de manière exclusive à des données binaires, nous pouvons ici garder les données initiales³⁰ pour la mesure des contributions et utiliser les classes engendrées par la discrétisation uniquement comme descripteurs "alimentant" l'algorithme. Nous serons toutefois, dépendant de la méthode de discrétisation utilisée.

Nous avons présenté les principes et fonctionnements d'un nouvel algorithme de classification automatique. Celui-ci cherche à dépasser les limites propres aux méthodes classiques que nous avons présentés précédemment. Nous donnons en guise de conclusion, les principaux

29. les algorithmes de tris ont en général une complexité en $\mathcal{O}(N \log(N))$ où N est la taille du vecteur d'éléments à trier

30. ie quantitatives

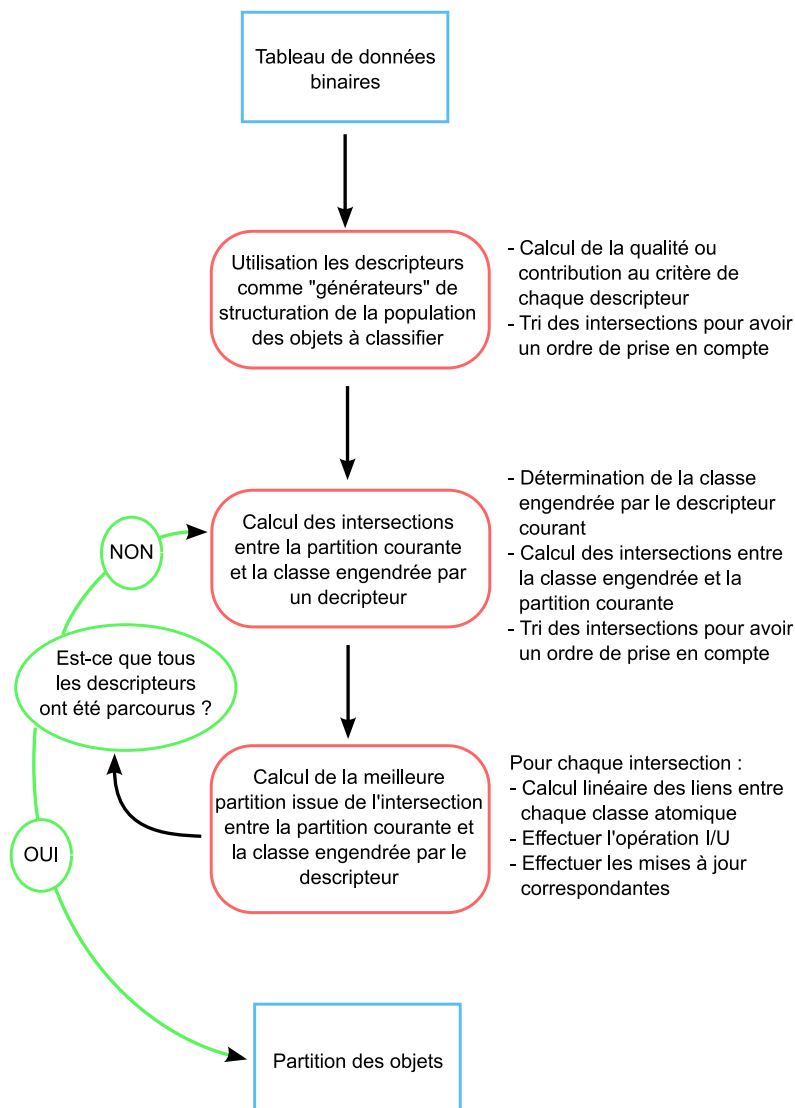


FIGURE 2.6 – Représentation graphique de l’algorithme AR2

avantages d’AR2 :

- utilisation d’une opération de base I/U qui permet d’explorer le graphe des partitions de manière plus efficace et d’approcher ainsi de manière plus précise l’optimal
- fixation d’aucun paramètre et en particulier, pas de fixation du nombre de classes
- non dépendance vis à vis de l’ordre des objets contrairement à AR1
- stabilité vis à vis de la duplication³¹ d’objets

Après avoir étudié certaines heuristiques, nous nous intéressons dans le chapitre suivant, aux indices de similarité entre objets et aux critères de partitionnement qui sont deux autres notions fondamentales du problème de la classification automatique.

31. si on duplique tous les objets de la base, on obtient la "même" partition avec dans chaque classe les objets et leurs répliques

CHAPITRE 3

Présentation unifiée des mesures de similarité et des critères de partitionnement

Sommaire

3.1	Extension des indices de similarité entre vecteurs binaires à des vecteurs non binaires	69
3.1.1	Indices de similarité entre vecteurs binaires et leur formalisme relationnel	70
3.1.1.1	Rappels sur les indices de similarité entre vecteurs binaires . . .	70
3.1.1.2	Les indices de similarité dans le formalisme relationnel	72
3.1.1.3	Critère de Condorcet étendu à des indices de similarité	74
3.1.2	Extension des indices de similarité entre vecteurs binaires à des vecteurs quelconques	76
3.1.2.1	Interprétation géométrique des expressions relationnelles	76
3.1.2.2	Interprétation des indices de similarité en fonction de paramètres géométriques	76
3.1.2.3	Etude des caractéristiques des indices de similarité dans le cadre de vecteurs quelconques	78
3.1.2.4	Définition d'une nouvelle classe d'indices de similarité pour des vecteurs quelconques	80
3.1.2.5	Critère de Condorcet et indices de similarité d'ordre t entre vecteurs quelconques	86
3.2	Extension des critères d'association maximale à des variables quantitatives et la notion d'écart à une tendance centrale	88
3.2.1	Critères d'association maximale et écart à une tendance centrale	88
3.2.1.1	Principe d'écart à une tendance centrale	88
3.2.1.2	Interprétations géométriques des critères d'association maximale, extension à des données quelconques	89
3.2.2	Liens entre les critères inertiels et le critère de Condorcet	91
3.2.2.1	Expression relationnelle de l'inertie totale	93
3.2.2.2	Expression relationnelle de l'inertie intra-classes	94
3.2.2.3	Expression relationnelle de l'inertie inter-classes	95

3.2.2.4	Lien entre le critère de la différence inertielle et le critère de Condorcet	95
3.2.3	Métriques relationnelles et régularisation de l'espace de description . .	97
3.3	Les notions d'indépendance et d'indétermination dans le cas de la mesure d'association entre variables qualitatives	99
3.3.1	Les différents types d'indépendance définis dans le cadre de l'AR	99
3.3.1.1	L'indépendance statistique	99
3.3.1.1.1	L'écart à l'indépendance statistique	99
3.3.1.1.2	Le rapport des chances ou "odds-ratio"	100
3.3.1.2	L'indétermination ou l'approche logique de l'indépendance entre relations d'équivalence	101
3.3.1.3	L'indépendance géométrique et indices de corrélation entre relations d'équivalence	103
3.3.2	Indice d'écart à l'indétermination pondérée entre relations d'équivalence	108

Dans ce chapitre, nous exploitons le codage relationnel présenté dans le chapitre 2 pour la mesure de similarité entre objets et la mesure d'association entre partitions.

L'objectif que nous visons dans un premier temps est de contribuer à l'"unification" des mesures statistiques de similarité et d'association pour des données numériques. Ainsi, nous étendons des mesures définies pour des données qualitatives à des données quantitatives. Au final, nous donnons des indices de similarité et des critères de partitionnement qui sont valables à la fois pour des données qualitatives et quantitatives.

Dans un second temps, nous revenons exclusivement vers l'étude des variables qualitatives et nous rappelons l'existence de plusieurs formes d'indépendances : statistique, logique et géométrique. Nous montrons que celles-ci sont duales selon le type de codage¹ dans lequel nous nous plaçons. Par ailleurs, nous nous attarderons de manière plus approfondie sur l'indépendance logique, ou encore l'indétermination, que nous généralisons en introduisant des pondérations.

Enfin, nous cherchons à étendre la notion de similarité régularisée. Toujours dans le souci de rapprocher les concepts logiques définis en AR des concepts géométriques de l'analyse de données en général, nous interprétons la régularisation dans un contexte géométrique. Ainsi, après avoir constaté que les similarités régularisées définies dans [BM98] correspondaient à des métriques diagonales, nous proposons d'étendre la régularisation à des métriques non diagonales : les produits scalaires qui fondent les mesures de similarité sont issues alors de métriques qui considèrent les liens entre variables.

Remarque 3.0.1 Notations

Le tableau de données de référence à partir duquel nous étudions les variables qualitatives est le tableau disjonctif complet² \mathbf{K} . Celui-ci est composé de P colonnes représentant les modalités D^j , $j = 1, \dots, P$, toutes variables V^k , $k = 1, \dots, M$, confondues. Ces modalités forment des descripteurs binaires qui peuvent être considérées comme des variables de présence / absence à part entière. Ainsi par souci de cohérence vis à vis des notations, nous noterons par D^j une variable quantitative et non par V^k , lorsque nous étendrons les définitions fondées sur le tableau \mathbf{K} au cas quantitatif.

3.1 Extension des indices de similarité entre vecteurs binaires à des vecteurs non binaires

Nous considérons dans le cadre de cette section, un ensemble d'objets \mathbb{O} de cardinal N , qui sont décrits par un ensemble de descripteurs \mathbb{D} de cardinal P . La description des objets consiste

1. contingentiel ou relationnel
2. présenté en introduction page 8

à renseigner pour ces derniers la présence ou l'absence de chaque descripteur. Nous avons de manière générale le tableau de données binaires \mathbf{K} de dimension $(N \times P)$, dont le terme général vaut, $\forall i = 1, \dots, N; \forall j = 1, \dots, P$:

$$\mathbf{K}_{ij} = \begin{cases} 1 & \text{si } O^i \text{ possède la modalité } D^j \in \mathbb{D} \\ 0 & \text{sinon} \end{cases} \quad (3.1.1)$$

Ce type de tableau que nous nommerons tableau disjonctif, se rencontre dans des contextes différents. Nous pouvons citer par exemple les données textuelles où les objets sont des documents et les descripteurs des lemmes ou tout autre "descripteur linguistique" tels que les co-occurrences [LS94]. Dans ce cas, $\mathbf{K}_{ij} = 1$ si le document O^i possède le lemme D^j ; 0 sinon.

En analyse de données, nous pouvons citer principalement les variables qualitatives. En effet, supposons un tableau de données \mathcal{T} constitué de N lignes et de M colonnes. Les lignes correspondent aux objets $O^i, i = 1, \dots, N$ et les colonnes à des variables qualitatives notées $V^k, k = 1, \dots, M$. Chaque variable possède p_k modalités et si nous les réunissons toutes, nous obtenons $\mathbb{D} = \{D^j; j = 1, \dots, P\}$ où $P = \sum_k p_k$, qui est l'ensemble des modalités ou descripteurs³. A partir de \mathcal{T} , nous pouvons ainsi déterminer le tableau disjonctif en explicitant les modalités de chaque variable et en codant leur présence ou leur absence au sein de chaque objet. Dans ce cas, $\mathbf{K}_{ij} = 1$ si l'objet O^i possède la modalité $D^j \in \mathbb{D}$; 0 sinon. Dans le contexte de variables qualitatives, nous parlerons de tableau disjonctif complet lorsque il n'y a aucune donnée manquante. Ceci se traduit formellement par la relation suivante :

$$\sum_{j=1}^P \mathbf{K}_{ij} = \mathbf{K}_i = M \quad \forall i = 1, \dots, N \quad (3.1.2)$$

Qu'il s'agisse de documents ou d'objets quelconques, nous avons :

- les lignes de \mathbf{K} qui codent des vecteurs binaires de dimension P que nous notons par $\{\vec{O}^i; i = 1, \dots, N\}$ et qui représentent les objets. Nous avons alors $O_j^i = \mathbf{K}_{ij}, \forall i = 1, \dots, N; \forall j = 1, \dots, P$
- les colonnes de \mathbf{K} qui codent des vecteurs binaires de dimension N que nous notons par $\{\vec{D}^j; j = 1, \dots, P\}$ et qui représentent les descripteurs. Nous avons alors $D_i^j = \mathbf{K}_{ij}, \forall i = 1, \dots, N; \forall j = 1, \dots, P$

Il existe alors de nombreux indices qui mesurent la similarité entre objets représentés par des vecteurs binaires.

Nous rappelons dans un premier temps ces indices et notamment leurs formulations en termes de matrices relationnelles collectives \mathbf{C} .

Nous interprétons ensuite ceux-ci dans un **contexte géométrique** et nous donnons alors une extension à des vecteurs non binaires.

Enfin, en appliquant ces indices à la mesure de similarité entre descripteurs $\vec{D}^j, j = 1, \dots, P$; nous définissons des métriques dites relationnelles qui permettent, dans un certain sens, de "régulariser" l'espace de description.

3.1.1 Indices de similarité entre vecteurs binaires et leur formalisme relationnel

3.1.1.1 Rappels sur les indices de similarité entre vecteurs binaires

Les indices de similarité entre vecteurs binaires sont définis selon une approche logique, à partir des quantités suivantes :

3. toutes variables confondues

Quantités logiques	Interprétation	Dénomination
$11_{ii'}$	Nombre de 1 en commun aux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$	Accord positif entre O^i et $O^{i'}$
$00_{ii'}$	Nombre de 0 en commun aux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$	Accord négatif entre O^i et $O^{i'}$
$10_{ii'}$	Nombre de couples (1,0) issus du croisement $(\vec{O}^i, \vec{O}^{i'})$	Désaccord relatif de O^i avec $O^{i'}$
$01_{ii'}$	Nombre de couples (0,1) issus du croisement $(\vec{O}^i, \vec{O}^{i'})$	Désaccord relatif de $O^{i'}$ avec O^i

TABLE 3.1 – Définition des quantités $11_{ii'}$, $00_{ii'}$, $10_{ii'}$ et $01_{ii'}$ entre deux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$

Nous avons aussi la présentation sous forme du tableau à double entrée tab-3.2.

	$O_j^{i'} = 1$	$O_j^{i'} = 0$	Marge
$O_j^i = 1$	$11_{ii'} = \sum_j \mathbf{K}_{ij} \mathbf{K}_{i'j}$	$10_{ii'} = \sum_j \mathbf{K}_{ij} (1 - \mathbf{K}_{i'j})$	$\mathbf{K}_i.$
$O_j^i = 0$	$01_{ii'} = \sum_j (1 - \mathbf{K}_{ij}) \mathbf{K}_{i'j}$	$00_{ii'} = \sum_j (1 - \mathbf{K}_{ij}) (1 - \mathbf{K}_{i'j})$	$P - \mathbf{K}_i.$
Marge	$\mathbf{K}_{i'.$	$P - \mathbf{K}_{i'.$	P

TABLE 3.2 – Définition des quantités dont sont fonctions les indices de similarité entre vecteurs binaires

Nous pouvons maintenant donner la définition de quelques uns de ces indices.

Définition 3.1.1 *Indice de Dice*

$$\text{Dice}(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + \frac{1}{2}(10_{ii'} + 01_{ii'})} \quad (3.1.3)$$

Définition 3.1.2 *Indice de Jaccard*

$$\text{Jaccard}(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + 10_{ii'} + 01_{ii'}} \quad (3.1.4)$$

Définition 3.1.3 *Indice de Sokal, Sneath, Anderberg (SSA)*

$$\text{SSA}(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + 2(10_{ii'} + 01_{ii'})} \quad (3.1.5)$$

Définition 3.1.4 *Indice de Sorensen*

$$\text{Sorensen}(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + \frac{1}{4}(10_{ii'} + 01_{ii'})} \quad (3.1.6)$$

Définition 3.1.5 *Indice symétrique d'Anderberg (AS)*

$$\text{AS}(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + \frac{1}{8}(10_{ii'} + 01_{ii'})} \quad (3.1.7)$$

Définition 3.1.6 *Indice d'Ochiaï*

$$\text{Ochiaï}(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{\sqrt{(11_{ii'} + 10_{ii'})(11_{ii'} + 01_{ii'})}} \quad (3.1.8)$$

Définition 3.1.7 *Indice de Kulczynski*

$$\text{Kulczynski}(\vec{O}^i, \vec{O}^{i'}) = \frac{1}{2} \left(\frac{11_{ii'}}{11_{ii'} + 10_{ii'}} + \frac{11_{ii'}}{11_{ii'} + 01_{ii'}} \right) \quad (3.1.9)$$

3.1.1.2 Les indices de similarité dans le formalisme relationnel

Considérons le tableau **K**. Les colonnes de ce dernier sont des descripteurs binaires de présence / absence qui définissent des relations d'équivalence partielle sur les objets. Notons alors par C^j la matrice relationnelle relative au descripteur D^j qui code en fait, la classe des objets qui possèdent ce descripteur : $C_{ii'}^j = 1$ si O^i et $O^{i'}$ possède le descripteur D^j ; 0 sinon.

Dans le cas particulier des variables qualitatives, nous avons la relation suivante :

$$\sum_{\{D^j \in V^k\}} C^j = C^k \quad (3.1.10)$$

où C^k est la matrice relationnelle codant la partition⁴ engendrée par V^k .

Notons alors par **C**, la matrice relationnelle collective constituée de la somme de chaque matrice C^j :

$$\mathbf{C} = \sum_{j=1}^P C^j \quad (3.1.11)$$

Nous avons alors dans tab-3.3 l'écriture des termes $11_{ii'}$, $00_{ii'}$, $10_{ii'}$ et $01_{ii'}$ en fonction de la matrice **C**.

Ces différentes grandeurs possèdent les propriétés définies ci-dessous dans le cas où il s'agit de données binaires issues de variables qualitatives et lorsque nous raisonnons sur des vecteurs d'objets.

4. ie l'ensemble des classes disjointes engendrées par chaque modalité D^j de V^k cf chapitre 2, paragraphe 2.1.1.1.3

Accords/ désaccords	Définition à partir du tableau disjonctif complet	Définition à partir de la matrice relationnelle collective
$11_{ii'}$	$= \sum_j \mathbf{K}_{ij} \mathbf{K}_{i'j}$	$= \mathbf{C}_{ii'}$
$00_{ii'}$	$= P - \left(\sum_j \mathbf{K}_{ij} + \sum_j \mathbf{K}_{i'j} - \sum_j \mathbf{K}_{ij} \mathbf{K}_{i'j} \right)$	$= P - (\mathbf{C}_{ii} + \mathbf{C}_{i'i'} - \mathbf{C}_{ii'})$
$10_{ii'}$	$= \sum_j \mathbf{K}_{ij} - \sum_j \mathbf{K}_{ij} \mathbf{K}_{i'j}$	$= \mathbf{C}_{ii} - \mathbf{C}_{ii'}$
$01_{ii'}$	$= \sum_j \mathbf{K}_{i'j} - \sum_j \mathbf{K}_{ij} \mathbf{K}_{i'j}$	$= \mathbf{C}_{i'i'} - \mathbf{C}_{ii'}$

TABLE 3.3 – Expressions relationnelles des quantités $11_{ii'}$, $00_{ii'}$, $10_{ii'}$, $01_{ii'}$

Propriété 3.1.1 *Relations entre les quantités* $11_{ii'}$, $00_{ii'}$, $10_{ii'}$, $01_{ii'}$

- $11_{ii'} + 00_{ii'} + 10_{ii'} + 01_{ii'} = P \quad \forall i, i' = 1, \dots, N$
- le nombre total de désaccords vaut : $10_{ii'} + 01_{ii'} = \mathbf{C}_{ii} + \mathbf{C}_{i'i'} - 2\mathbf{C}_{ii'}$
- le nombre total d'accords vaut : $11_{ii'} + 00_{ii'} = P - (\mathbf{C}_{ii} + \mathbf{C}_{i'i'} - 2\mathbf{C}_{ii'})$

Dans le cas de variables qualitatives et de données complètes⁵, nous avons de plus :

- $11_{ii'} + \frac{1}{2}(10_{ii'} + 01_{ii'}) = M \quad \forall i, i' = 1, \dots, N$
- $11_{ii} = \mathbf{C}_{ii} = M \quad \forall i = 1, \dots, N$

Nous donnons ci-dessous l'expression relationnelle des indices de similarité entre vecteurs binaires en utilisant les formules de passages de tab-3.3.

Propriété 3.1.2 *Indice de Dice en termes relationnels*

$$\text{Dice}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{\frac{1}{2}(\mathbf{C}_{ii} + \mathbf{C}_{i'i'})} \quad (3.1.12)$$

Propriété 3.1.3 *Indice de Jaccard en termes relationnels*

$$\text{Jaccard}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{\mathbf{C}_{ii} + \mathbf{C}_{i'i'} - \mathbf{C}_{ii'}} \quad (3.1.13)$$

Propriété 3.1.4 *Indice de Sokal, Sneath, Anderberg (SSA) en termes relationnels*

$$\text{SSA}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{2(\mathbf{C}_{ii} + \mathbf{C}_{i'i'}) - 3\mathbf{C}_{ii'}} \quad (3.1.14)$$

5. ie $\sum_j \mathbf{K}_{ij} = M \quad \forall i = 1, \dots, N$ où M est le nombre de variables qualitatives

Propriété 3.1.5 *Indice de Sorensen en termes relationnels*

$$\text{Sorensen}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{\frac{1}{2}((\mathbf{C}_{ii} + \mathbf{C}_{i'i'})/2 + \mathbf{C}_{ii'})} \quad (3.1.15)$$

Propriété 3.1.6 *Indice symétrique d'Anderberg (AS) en termes relationnels*

$$\text{AS}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{\frac{1}{4}((\mathbf{C}_{ii} + \mathbf{C}_{i'i'})/2 + 3\mathbf{C}_{ii'})} \quad (3.1.16)$$

Propriété 3.1.7 *Indice d'Ochiaï en termes relationnels*

$$\text{Ochiaï}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{\sqrt{\mathbf{C}_{ii}\mathbf{C}_{i'i'}}} \quad (3.1.17)$$

Propriété 3.1.8 *Indice de Kulczynski en termes relationnels*

$$\text{Kulczynski}(\vec{O}^i, \vec{O}^{i'}) = \frac{\mathbf{C}_{ii'}}{(2\mathbf{C}_{ii}\mathbf{C}_{i'i'})/(\mathbf{C}_{ii} + \mathbf{C}_{i'i'})} \quad (3.1.18)$$

3.1.1.3 Critère de Condorcet étendu à des indices de similarité

Notons de manière générique par $\mathbf{S}(\vec{O}^i, \vec{O}^{i'}) = \mathbf{S}_{ii'}$, un indice de similarité parmi ceux présentés ci-dessus. Tous ces indices sont compris entre 0 et 1 : la similarité est maximale lorsque $\vec{O}^i = \vec{O}^{i'}$; elle est nulle lorsque $11_{ii'} = 0$. Par ailleurs, il s'agit d'indices symétriques : $\mathbf{S}_{ii'} = \mathbf{S}_{i'i}$, $\forall i, i' = 1, \dots, N$.

Nous pouvons définir un indice de dissimilarité à partir de $\mathbf{S}_{ii'}$ que nous notons $\bar{\mathbf{S}}_{ii'}$:

Définition 3.1.8 *Indice de similarité et de dissimilarité*

Soit $\mathbf{S}_{ii'}$ un indice de similarité mesuré sur deux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$. Supposons que cet indice varie dans l'intervalle $[a, b]$ avec $a < b$ et $b > 0$. Il doit alors vérifier les propriétés suivantes :

- $\mathbf{S}_{ii} = b \quad \forall i = 1, \dots, N$: auto-similarité maximale
- $\mathbf{S}_{ii'} = \mathbf{S}_{i'i} \quad \forall i, i' = 1, \dots, N$: symétrie

Nous pouvons associer à $\mathbf{S}_{ii'}$ un indice de dissimilarité noté $\bar{\mathbf{S}}_{ii'}$ donné par :

$$\bar{\mathbf{S}}_{ii'} = \mathcal{M}^t(\mathbf{S}_{ii}, \mathbf{S}_{i'i'}) - \mathbf{S}_{ii'} \quad (3.1.19)$$

où $\mathcal{M}^t(A, B)$ est une moyenne généralisée d'ordre t avec $-\infty < t < +\infty$ définie par (1.8) page 6 ou plus généralement par déf-3.1.10 page 81. Par défaut, nous prendrons $t = 1$ ce qui nous donne :

$$\bar{\mathbf{S}}_{ii'} = \frac{\mathbf{S}_{ii} + \mathbf{S}_{i'i'}}{2} - \mathbf{S}_{ii'} \quad (3.1.20)$$

Nous dirons enfin que des indices sont normalisés, lorsque $[a, b] = [-1, 1]$ ou $[a, b] = [0, 1]$. Dans ce cas, nous avons :

$$\bar{\mathbf{S}}_{ii'} = 1 - \mathbf{S}_{ii'} \quad (3.1.21)$$

Nous pouvons alors paramétrer le critère de Condorcet présenté en introduction page 5 par un indice de similarité particulier et son indice de dissimilarité associé.

Définition 3.1.9 Critère de Condorcet fonction d'un indice de similarité

Soit $S_{ii'}$ un indice de similarité vérifiant les hypothèses énoncées dans déf-3.1.8 et variant entre 0 et 1. Soit alors, $\bar{S}_{ii'} = 1 - S_{ii'}$, l'indice de dissimilarité associé. Nous pouvons étendre le critère de Condorcet à d'autres indices de similarité $S_{ii'}$ que $C_{ii'}$ de la manière suivante :

$$\text{Condorcet}(\mathbf{S}, X) = \sum_{i=1}^N \sum_{i'=1}^N S_{ii'} X_{ii'} + \sum_{i=1}^N \sum_{i'=1}^N \bar{S}_{ii'} \bar{X}_{ii'} \quad (3.1.22)$$

Dans ce cas nous avons les interprétations suivantes du critère de Condorcet :

$$\begin{aligned} \max \text{Condorcet}(\mathbf{S}, X) &\Leftrightarrow \max \sum_{i,i'} S_{ii'} X_{ii'} + \sum_{i,i'} \bar{S}_{ii'} \bar{X}_{ii'} && \text{“maximisation des accords 11 et 00”} \\ &\Leftrightarrow \max \sum_{i,i'} (S_{ii'} - \bar{S}_{ii'}) X_{ii'} && \text{“maximisation des contributions”} \\ &\Leftrightarrow \max \sum_{i,i'} (S_{ii'} - \frac{1}{2}) X_{ii'} && \text{“règle majoritaire”} \end{aligned} \quad (3.1.23)$$

où X est une matrice relationnelle codant une relation d'équivalence.

Dans [MM81] et [Mar03], J.F. Marcotorchino montre que bon nombre des indices cités ci-dessus sont fonctions de l'indice de Dice. De plus, ces indices peuvent être interprétés en terme de règle majoritaire : deux objets sont *a priori* dans une même classe si l'ensemble des variables supporte, au sens d'une règle de vote, cette relation. Nous rappelons ces résultats dans le tableau tab-3.4 dans le cas de données complètes⁶.

Indice	Expression en fonction de Dice	Règle majoritaire correspondante
Dice($\vec{O}^i, \vec{O}^{i'}$)		Dice($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{1}{2}M$
Jaccard($\vec{O}^i, \vec{O}^{i'}$)	= $\frac{\text{Dice}(\vec{O}^i, \vec{O}^{i'})}{2 - \text{Dice}(\vec{O}^i, \vec{O}^{i'})}$	Jaccard($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{2}{3}M$
SSA($\vec{O}^i, \vec{O}^{i'}$)	= $\frac{\text{Dice}(\vec{O}^i, \vec{O}^{i'})}{4 - 3\text{Dice}(\vec{O}^i, \vec{O}^{i'})}$	SSA($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{4}{5}M$
Sorensen($\vec{O}^i, \vec{O}^{i'}$)	= $\frac{2\text{Dice}(\vec{O}^i, \vec{O}^{i'})}{\text{Dice}(\vec{O}^i, \vec{O}^{i'}) + 1}$	Sorensen($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{1}{3}M$
AS($\vec{O}^i, \vec{O}^{i'}$)	= $\frac{4\text{Dice}(\vec{O}^i, \vec{O}^{i'})}{3\text{Dice}(\vec{O}^i, \vec{O}^{i'}) + 1}$	AS($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{1}{5}M$
Ochiai($\vec{O}^i, \vec{O}^{i'}$)	= Dice($\vec{O}^i, \vec{O}^{i'}$)	Ochiai($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{1}{2}M$
Kulczynski($\vec{O}^i, \vec{O}^{i'}$)	= Dice($\vec{O}^i, \vec{O}^{i'}$)	Kulczynski($\vec{O}^i, \vec{O}^{i'}$) > $\frac{1}{2} \Rightarrow C_{ii'} > \frac{1}{2}M$

TABLE 3.4 – Expressions des indices de similarité en fonction de Dice et règles majoritaires correspondantes

La traduction en termes de règles majoritaires permet de mieux comprendre les distinctions entre les différents indices de similarité. Par exemple, si nous prenons l'indice de Jaccard comme

6. ie $C_{ii} = M \quad \forall i = 1, \dots, M$

mesure de similarité, le critère de Condorcet associé indique que deux objets O^i et $O^{i'}$ seront potentiellement dans la même classe si l'indice de Jaccard est supérieur ou égal à $1/2$. Ceci correspond en fait à prendre comme mesure de similarité le nombre, $C_{ii'}$, de variables supportant " O^i et $O^{i'}$ sont dans la même classe" et de le comparer à une borne, ou encore une majorité qui est de $2/3$. Le critère de Condorcet associé à l'indice de Jaccard correspond ainsi à une règle majoritaire à $2/3$ des voix. Il est donc plus sévère que les critères de Condorcet associés aux indices de Dice, Ochiaï ou Kulczynski qui correspondent à une règle majoritaire à $1/2$ des voix.

3.1.2 Extension des indices de similarité entre vecteurs binaires à des vecteurs quelconques

Nous donnons des interprétations géométriques aux quantités $11_{ii'}$, $10_{ii'}$ et $01_{ii'}$. A partir de ces dernières, nous redéfinissons les indices de similarité en fonction de deux paramètres que sont le cosinus formé par l'angle de deux vecteurs et le rapport de leurs normes.

Nous aboutissons à une approche géométrique des indices de similarité entre vecteurs binaires qui peuvent de ce fait, être étendus naturellement à des vecteurs quantitatifs en considérant l'extension continue des paramètres géométriques.

Il existe des liens étroits entre les concepts logiques propres à l'AR et des concepts géométriques plus répandus en analyse de données.

Nous exploitons ces relations dans une perspective d'unification des mesures définies à la fois pour des variables qualitatives et des variables quantitatives.

3.1.2.1 Interprétation géométrique des expressions relationnelles

Les interprétations en termes géométriques des quantités $11_{ii'}$, $10_{ii'}$ et $01_{ii'}$, ont été plus aisément mises en évidence à partir des expressions relationnelles de ces grandeurs. Nous résumons ces interprétations dans le tableau tab-3.5.

Suite à ces considérations, nous présentons dans le paragraphe suivant des interprétations géométriques des différents indices de similarité.

3.1.2.2 Interprétation des indices de similarité en fonction de paramètres géométriques

Soit $S_{ii'}$ un indice de similarité mesuré entre les vecteurs \vec{O}^i et $\vec{O}^{i'}$ parmi ceux présentés précédemment. Ce dernier peut être défini en fonction de deux paramètres :

- $\theta_{ii'} = \cos(\widehat{\vec{O}^i \vec{O}^{i'}})$: **le cosinus de l'angle formé par les deux vecteurs binaires.** En raison de la binarité des vecteurs, l'équivalence entre l'expression logique et l'expression géométrique des indices implique qu'ils sont définis pour $\theta_{ii'} \in [0, 1]$
- $\gamma_{ii'} = \frac{\max(\|\vec{O}^i\|, \|\vec{O}^{i'}\|)}{\min(\|\vec{O}^i\|, \|\vec{O}^{i'}\|)}$: **le rapport de la plus grande norme des deux sur la plus petite.**
Nous avons donc $\gamma_{ii'} \geq 1$

Notons que ces paramètres sont des rationnels et non des réels étant donné qu'ils sont définis sur des vecteurs dont les valeurs sont soit 0 soit 1. Par conséquent, ils s'expriment sous la forme de fractions entre nombres entiers.

Le cosinus de l'angle formé par deux vecteurs est une mesure de similarité déjà largement employée en statistique appliquée. En revanche, **le rapport des normes n'est pas, à notre connaissance, un paramètre utilisé pour mesurer la similarité entre deux vecteurs.** Il nous semble intéressant alors de l'exploiter. Tout d'abord, parce qu'à cosinus constant il

Accords/ désaccords	Définition à partir de \mathbf{C}	Produit scalaire	Interprétation géométrique
$11_{ii'}$	$= \mathbf{C}_{ii'}$	$= \langle \vec{O}^i, \vec{O}^{i'} \rangle$	Produit scalaire euclidien
11_{ii}	$= \mathbf{C}_{ii}$	$= \ \vec{O}^i\ ^2$	Norme au carré
$10_{ii'}$	$= \mathbf{C}_{ii} - \mathbf{C}_{ii'}$	$= \langle \vec{O}^i, \vec{O}^i - \vec{O}^{i'} \rangle$	Distance relative de \vec{O}^i à $\vec{O}^{i'}$
$01_{ii'}$	$= \mathbf{C}_{i'i'} - \mathbf{C}_{ii'}$	$= \langle \vec{O}^{i'}, \vec{O}^{i'} - \vec{O}^i \rangle$	Distance relative de $\vec{O}^{i'}$ à \vec{O}^i
$10_{ii'} + 01_{ii'}$	$= \mathbf{C}_{ii} + \mathbf{C}_{i'i'} - 2\mathbf{C}_{ii'} = 2\bar{\mathbf{C}}_{ii'}$	$= \ \vec{O}^i - \vec{O}^{i'}\ ^2$	Distance euclidienne entre $\vec{O}^{i'}$ et \vec{O}^i
$10_{ii'} + 11_{ii'}$	$= \mathbf{C}_{ii}$	$= \ \vec{O}^i\ ^2$	Norme au carré de \vec{O}^i
$11_{ii'} + 01_{ii'}$	$= \mathbf{C}_{i'i'}$	$= \ \vec{O}^{i'}\ ^2$	Norme au carré de $\vec{O}^{i'}$

TABLE 3.5 – Interprétations géométriques des quantités $11_{ii'}$, $10_{ii'}$, $01_{ii'}$

permet d'affiner la mesure de similarité entre les vecteurs. Par exemple, deux vecteurs colinéaires positivement qui ne sont pas de même norme ne sont pas par définition identiques. Ainsi, leur mesure de similarité devrait être distincte de 1. Ensuite, notre approche est justifiée dans la mesure où elle vise une extension d'indices de similarité déjà existants.

Nous donnons ci-dessous la définition des indices de similarité présentés précédemment en fonction des deux paramètres géométriques.

Propriété 3.1.9 *Interprétation géométrique de l'indice de Dice*

$$\text{Dice}(\theta_{ii'}, \gamma_{ii'}) = \frac{\theta_{ii'}}{\left(\frac{1+\gamma_{ii'}}{2\gamma_{ii'}}\right)} \quad (3.1.24)$$

Propriété 3.1.10 *Interprétation géométrique de l'indice de Jaccard*

$$\text{Jaccard}(\theta_{ii'}, \gamma_{ii'}) = \frac{\theta_{ii'}}{2\left(\frac{1+\gamma_{ii'}}{2\gamma_{ii'}}\right) - \theta_{ii'}} \quad (3.1.25)$$

Propriété 3.1.11 *Interprétation géométrique de l'indice de Sokal, Sneath, Anderberg (SSA)*

$$\text{SSA}(\theta_{ii'}, \gamma_{ii'}) = \frac{\theta_{ii'}}{4\left(\frac{1+\gamma_{ii'}}{2\gamma_{ii'}}\right) - 3\theta_{ii'}} \quad (3.1.26)$$

Propriété 3.1.12 *Interprétation géométrique de l'indice de Sorensen*

$$\text{Sorensen}(\theta_{ii'}, \gamma_{ii'}) = \frac{\theta_{ii'}}{\frac{1}{2} \left(\frac{1+\gamma_{ii'}}{2\gamma_{ii'}} \right) + \frac{1}{2}\theta_{ii'}} \quad (3.1.27)$$

Propriété 3.1.13 *Interprétation géométrique de l'indice symétrique d'Anderberg (AS)*

$$\text{AS}(\theta_{ii'}, \gamma_{ii'}) = \frac{\theta_{ii'}}{\frac{1}{4} \left(\frac{1+\gamma_{ii'}}{2\gamma_{ii'}} \right) + \frac{3}{4}\theta_{ii'}} \quad (3.1.28)$$

Propriété 3.1.14 *Interprétation géométrique de l'indice d'Ochiaï*

$$\text{Ochiaï}(\theta_{ii'}, \gamma_{ii'}) = \theta_{ii'} \quad (3.1.29)$$

Propriété 3.1.15 *Interprétation géométrique de l'indice de Kulczynski*

$$\text{Kulczynski}(\theta_{ii'}, \gamma_{ii'}) = \frac{\theta_{ii'}}{\left(\frac{2\gamma_{ii'}}{1+\gamma_{ii'}} \right)} \quad (3.1.30)$$

Démonstration

L'ensemble de ces résultats sont obtenus simplement à partir de l'expression relationnelle des indices. Dans un premier temps, nous remplaçons $\mathbf{C}_{ii'}$ par $\langle \vec{O}^i, \vec{O}^{i'} \rangle$. Par ailleurs, nous avons $\langle \vec{O}^i, \vec{O}^{i'} \rangle = \theta_{ii'} \|\vec{O}^i\| \|\vec{O}^{i'}\|$.

Dans un second temps, nous posons $\frac{\|\vec{O}^{i'}\|}{\|\vec{O}^i\|} = \gamma_{ii'}$. Nous supposons, sans perte de généralité, que $\gamma_{ii'} \geq 1$ ce qui implique que $\|\vec{O}^{i'}\| \geq \|\vec{O}^i\|$. En effet, prendre $\vec{O}^{i'}$ comme étant le vecteur de plus grande norme importe peu puisque les indices font jouer un rôle symétrique aux deux vecteurs.

En effectuant ces différents remplacements et en simplifiant nous obtenons pour chaque indice son expression en fonction des paramètres $\theta_{ii'}$ et $\gamma_{ii'}$.

□

3.1.2.3 Etude des caractéristiques des indices de similarité dans le cadre de vecteurs quelconques

Nous étudions les caractéristiques des indices présentés sur le domaine $[-1, 1] \times [1, +\infty[$ en supposant leur **extension continue**. L'objectif est notamment de déterminer si le comportement des indices sur ce domaine en fait de "bons candidats" pour la mesure de similarité entre vecteurs non binaires.

Dans la suite, nous aurons donc :

- $-1 \leq \theta_{ii'} \leq 1 \quad \forall i, i' = 1, \dots, N$
- $1 \leq \gamma_{ii'} \quad \forall i, i' = 1, \dots, N$

A la différence des vecteurs binaires, les indices de similarité étendus à des données quelconques peuvent prendre des valeurs négatives. Nous parlerons alors de vecteurs ou d'objets

“antisimilaires” ce qui indique des objets dont les profils sont opposés.

Lorsque l’indice entre deux objets est positif, il s’agit *a contrario* des objets dont les profils se ressemblent.

Lorsque l’indice est nul, ceci indique une indépendance géométrique entre deux vecteurs puisque nous avons dans ce cas $\theta_{ii'} = 0$. Il s’agit alors d’objets dont les profils sont ni ressemblants, ni opposés.

Les caractéristiques que nous nous proposons d’étudier sont les suivantes :

- Monotonie vis à vis de chacun des deux paramètres
- Valeurs en des points caractéristiques
- Intervalle des valeurs prises par les indices

Nous résumons ces différents points dans les tableaux tab-3.6 et tab-3.7.

$S_{ii'}$	$\frac{\partial S_{ii'}}{\partial \theta_{ii'}}$	$\frac{\partial S_{ii'}}{\partial \gamma_{ii'}}$	Monotonie selon $\theta_{ii'}$	Monotonie selon $\gamma_{ii'}$
Dice($\theta_{ii'}, \gamma_{ii'}$)	$\frac{2\gamma_{ii'}}{1+\gamma_{ii'}^2}$	$2\theta_{ii'} \frac{1-\gamma_{ii'}^2}{(1+\gamma_{ii'}^2)^2}$	↗	↘ si $\theta_{ii'} > 0$ ↗ si $\theta_{ii'} < 0$
Jaccard($\theta_{ii'}, \gamma_{ii'}$)	$\frac{(1+\gamma_{ii'}^2)/\gamma_{ii'}}{((1+\gamma_{ii'}^2)/\gamma_{ii'}-\theta_{ii'})^2}$	$\theta_{ii'} \frac{(1-\gamma_{ii'}^2)/\gamma_{ii'}^2}{((1+\gamma_{ii'}^2)/\gamma_{ii'}-\theta_{ii'})^2}$	↗	↘ si $\theta_{ii'} > 0$ ↗ si $\theta_{ii'} < 0$
SSA($\theta_{ii'}, \gamma_{ii'}$)	$\frac{2(1+\gamma_{ii'}^2)/\gamma_{ii'}}{(2(1+\gamma_{ii'}^2)/\gamma_{ii'}-3\theta_{ii'})^2}$	$2\theta_{ii'} \frac{(1-\gamma_{ii'}^2)/\gamma_{ii'}^2}{(2(1+\gamma_{ii'}^2)/\gamma_{ii'}-\theta_{ii'})^2}$	↗	↘ si $\theta_{ii'} > 0$ ↗ si $\theta_{ii'} < 0$
Sorensen($\theta_{ii'}, \gamma_{ii'}$)	$\frac{(1+\gamma_{ii'}^2)/4\gamma_{ii'}}{((1+\gamma_{ii'}^2)/4\gamma_{ii'}+\theta_{ii'}/2)^2}$	$\theta_{ii'} \frac{(1-\gamma_{ii'}^2)/\gamma_{ii'}^2}{4((1+\gamma_{ii'}^2)/4\gamma_{ii'}+\theta_{ii'}/2)^2}$	↗	↘ si $\theta_{ii'} > 0$ ↗ si $\theta_{ii'} < 0$
AS($\theta_{ii'}, \gamma_{ii'}$)	$\frac{(1+\gamma_{ii'}^2)/8\gamma_{ii'}}{((1+\gamma_{ii'}^2)/8\gamma_{ii'}+3\theta_{ii'}/4)^2}$	$\theta_{ii'} \frac{(1-\gamma_{ii'}^2)/\gamma_{ii'}^2}{8((1+\gamma_{ii'}^2)/8\gamma_{ii'}+3\theta_{ii'}/4)^2}$	↗	↘ si $\theta_{ii'} > 0$ ↗ si $\theta_{ii'} < 0$
Kulczynski($\theta_{ii'}, \gamma_{ii'}$)	$\frac{1+\gamma_{ii'}^2}{2\gamma_{ii'}}$	$-2\theta_{ii'} \frac{1-\gamma_{ii'}^2}{(2\gamma_{ii'}^2)^2}$	↗	↗ si $\theta_{ii'} > 0$ ↘ si $\theta_{ii'} < 0$
Ochiai($\theta_{ii'}$)	1	-	↗	-

TABLE 3.6 – Etude analytique des indices de similarité

Il apparaît à la suite de cette rapide étude des caractéristiques des indices en fonction des

$S_{ii'}$	Indice quand $\gamma_{ii'} = 1$	$(\theta_{ii'}, \gamma_{ii'}) = (1, 1)$	$(\theta_{ii'}, \gamma_{ii'}) = (-1, 1)$	Indice quand $\theta_{ii'} = 1$	Limite pour $\gamma_{ii'} \rightarrow +\infty$	Intervalle des valeurs prises
Dice($\theta_{ii'}, \gamma_{ii'}$)	$\theta_{ii'}$	1	-1	$\frac{2\gamma_{ii'}}{1+\gamma_{ii'}^2}$	0	$[-1, 1]$
Jaccard($\theta_{ii'}, \gamma_{ii'}$)	$\frac{\theta_{ii'}}{2-\theta_{ii'}}$	1	-1/3	$\frac{\gamma_{ii'}}{1+\gamma_{ii'}^2-\gamma_{ii'}}$	0	$[-\frac{1}{3}, 1]$
SSA($\theta_{ii'}, \gamma_{ii'}$)	$\frac{\theta_{ii'}}{4-3\theta_{ii'}}$	1	-1/7	$\frac{\gamma_{ii'}}{2+2\gamma_{ii'}^2-3\gamma_{ii'}}$	0	$[-\frac{1}{7}, 1]$
Sorensen($\theta_{ii'}, \gamma_{ii'}$)	$\frac{\theta_{ii'}}{\frac{1}{2}(1+\theta_{ii'})}$	1	$-\infty$	$\frac{4\gamma_{ii'}}{(1+\gamma_{ii'})^2}$	0	$[-\infty, 1]$
Ochiai($\theta_{ii'}, \gamma_{ii'}$)	$\theta_{ii'}$	1	-1	1	-	$[-1, 1]$
Kulczynski($\theta_{ii'}, \gamma_{ii'}$)	$\theta_{ii'}$	1	-1	$\frac{1+\gamma_{ii'}^2}{2\gamma_{ii'}}$	$+\infty$	$[-1, +\infty]$

TABLE 3.7 – Etude des caractéristiques des indices de similarité

paramètres $\theta_{ii'}$ et $\gamma_{ii'}$, qu'ils ne sont pas tous de "bons candidats" pour la mesure de similarité entre vecteurs non binaires. En effet, nombreux de ces indices ne sont pas bornés et possèdent des points en lesquels il ne sont pas définis.

Cependant, parmi les indices étudiés, il y en a qui pourraient être des mesures de similarité adaptées pour des vecteurs de données non binaires. Il s'agit des indices de Dice et d'Ochiaï dont le comportement et l'interprétation est en accord avec ce que nous attendons d'une mesure de similarité :

- ils sont croissants en $\theta_{ii'}$ et décroissant en $\gamma_{ii'}$
- ils sont compris entre -1 et 1

Dans le paragraphe suivant, nous généralisons ces indices de similarité.

3.1.2.4 Définition d'une nouvelle classe d'indices de similarité pour des vecteurs quelconques

Dans ce paragraphe, nous définissons une famille d'indices de similarité qui pourrait être employée pour la mesure de similarité entre vecteurs quelconques⁷ et qui généralise les indices tels que Dice et Ochiaï. Dans la mesure où ces indices sont fondés sur des paramètres géométriques et non plus logiques, ils sont aussi et surtout valables pour la mesure de similarité d'objets décrits par des variables quantitatives.

Cette famille d'indice de similarité fait intervenir la notion de moyenne généralisée que nous

7. binaires ou non binaires

rappelons ci-dessous.

Définition 3.1.10 Moyenne généralisée d'un ensemble de nombres

Soit $\mathbb{A} = \{A_1, \dots, A_N\}$ un ensemble de nombre. Nous appelons moyenne d'ordre t de l'ensemble des nombres de \mathbb{A} , avec $-\infty < t < +\infty$ le nombre noté $\mathcal{M}^t(\mathbb{A})$ défini par :

$$\mathcal{M}^t(A_1, \dots, A_N) = \left[\frac{1}{N} \sum_{i=1}^N A_i^t \right]^{\frac{1}{t}} \quad (3.1.31)$$

Nous avons alors les correspondances bien connues suivantes :

- $t = 1 \quad \leftrightarrow \quad \mathcal{M}^1(\mathbb{A}) = \text{Moyenne arithmétique des éléments de } \mathbb{A}$
- $t \rightarrow 0 \quad \leftrightarrow \quad \mathcal{M}^t(\mathbb{A}) \rightarrow \text{Moyenne géométrique des éléments de } \mathbb{A}$
- $t = -1 \quad \leftrightarrow \quad \mathcal{M}^{-1}(\mathbb{A}) = \text{Moyenne harmonique des éléments de } \mathbb{A}$
- $t \rightarrow +\infty \quad \leftrightarrow \quad \mathcal{M}^t(\mathbb{A}) = \max(\mathbb{A})$
- $t \rightarrow -\infty \quad \leftrightarrow \quad \mathcal{M}^t(\mathbb{A}) = \min(\mathbb{A})$

Deux moyennes généralisées d'ordres différents, d'un même ensemble de nombres, possèdent la propriété suivante :

Propriété 3.1.16 Inégalités entre moyennes d'ordre t d'un ensemble de nombres

Soient s et t deux entiers tel que $-\infty < s < t < +\infty$. Nous avons alors la relation suivante :

$$\mathcal{M}^s(\mathbb{A}) \leq \mathcal{M}^t(\mathbb{A}) \quad (3.1.32)$$

Nous obtenons l'égalité si $A_i = A \quad \forall i = 1, \dots, N$.

Si nous reprenons les expressions relationnelles des indices de Dice, d'Ochiaï et de Kulczynski données page 72, nous obtenons par ailleurs, la propriété suivante :

Propriété 3.1.17 Indices de Dice, d'Ochiaï et de Kulczynski en fonction des moyennes

L'expression relationnelle des indices de Dice, d'Ochiaï et de Kulczynski correspondent à un même type d'indice de similarité. La différence entre ces indices correspond à l'utilisation d'une moyenne d'ordre distinct.

- $\text{Dice}(\vec{O}^i, \vec{O}^{i'}) = \mathbf{C}_{ii'} / \mathcal{M}^1(\mathbf{C}_{ii}, \mathbf{C}_{i'i'}) = \langle \vec{O}^i, \vec{O}^{i'} \rangle / \mathcal{M}^1(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)$
- $\text{Ochiaï}(\vec{O}^i, \vec{O}^{i'}) = \mathbf{C}_{ii'} / \mathcal{M}^{t \rightarrow 0}(\mathbf{C}_{ii}, \mathbf{C}_{i'i'}) = \langle \vec{O}^i, \vec{O}^{i'} \rangle / \mathcal{M}^{t \rightarrow 0}(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)$
- $\text{Kulczynski}(\vec{O}^i, \vec{O}^{i'}) = \mathbf{C}_{ii'} / \mathcal{M}^{-1}(\mathbf{C}_{ii}, \mathbf{C}_{i'i'}) = \langle \vec{O}^i, \vec{O}^{i'} \rangle / \mathcal{M}^{-1}(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)$

Nous faisons remarquer, que dans le cas de données quantitatives, le coefficient d'Ochiaï défini ci-dessus se ramène au coefficient de corrélation de Bravais-Pearson $\rho(\vec{O}^i, \vec{O}^{i'}) = \theta_{ii'}$.

Remarque 3.1.1 Notations

Pour alléger les notations, nous notons par $\mathcal{M}^{t \rightarrow 0}(\mathbb{A})$ le nombre limite donné par $\lim_{t \rightarrow 0} \mathcal{M}^t(\mathbb{A})$.

Nous définissons une nouvelle classe d'indices qui englobe ceux de Dice et d'Ochiaï.

Définition 3.1.11 Indices de similarité d'ordre t

Soient deux vecteurs \vec{O}^i et $\vec{O}^{i'}$ représentant les profils de deux objets dans un espace de description à M dimensions. L'indice de similarité d'ordre $t > 0$ entre ces deux vecteurs est donné par la formule suivante :

$$\mathbf{S}_{ii'}^t = \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\mathcal{M}^t(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)} \quad (3.1.33)$$

L'indice de Kulczynski ne fait pas partie des indices d'ordre $t > 0$. Comme nous l'avons constaté dans le paragraphe précédent, il ne possède pas des caractéristiques qui répondent à nos attentes en terme de mesure de similarité. En particulier, plus le rapport des normes est distant de 1 et plus l'indice de Kulczynski est fort ce qui est contraire à l'intuition.

L'ordre correspondant à l'indice de Kulczynski est $t = -1$. Lorsque l'ordre est inférieur à 0, l'indice défini fait jouer un rôle au rapport des normes qui est contraire à ce que nous recherchons.

En fait, de manière générale, nous avons les relations suivantes :

- $t < 0 \Rightarrow \mathbf{S}_{ii'}^t$, est d'autant plus fort que le rapport des normes s'éloigne de 1
- $t \rightarrow 0 \Rightarrow \mathbf{S}_{ii'}^t$, ne fait jouer aucun rôle au rapport des normes : l'indice d'Ochiaï représente le cosinus de l'angle et est indépendant du rapport des normes
- $t > 0 \Rightarrow \mathbf{S}_{ii'}^t$, est d'autant plus fort que le rapport des normes est proche de 1

Autrement dit, la valeur de t indique le "rôle" et l' "intensité du rôle" que l'indice défini fait jouer au rapport des normes. Dans le paragraphe suivant, nous étudions plus concrètement l'impact de l'ordre t d'une similarité dans le cadre du critère de Condorcet.

Nous donnons à présent quelques propriétés des indices de similarité que nous venons de définir.

Propriété 3.1.18 Les indices de similarité d'ordre t sont bornés et sont symétriques

- $|\mathbf{S}_{ii'}^t| \leq 1 \quad \forall t > 0.$
- $\mathbf{S}_{ii'}^t = \mathbf{S}_{i'i}^t \quad \forall i, i' = 1, \dots, N.$

Propriété 3.1.19 Inégalités entre indices de similarité d'ordre t

Soient \vec{O}^i et $\vec{O}^{i'}$ deux vecteurs représentatifs de deux objets dans un espace de description à M variables. Soient s et t deux entiers tel que $-\infty < s < t < +\infty$. En conséquence de la propriété prop-3.1.16, nous avons la relation suivante entre $\mathbf{S}_{ii'}^s$ et $\mathbf{S}_{ii'}^t$, deux indices de similarité d'ordres respectifs s et t entre ces deux vecteurs :

$$\begin{cases} \mathbf{S}_{ii'}^s \geq \mathbf{S}_{ii'}^t & \text{si } \theta_{ii'} > 0 \\ \mathbf{S}_{ii'}^s \leq \mathbf{S}_{ii'}^t & \text{si } \theta_{ii'} < 0 \end{cases} \quad (3.1.34)$$

Nous obtenons l'égalité si $\vec{O}^i = \vec{O}^{i'}$.

Démonstration

Nous donnons la démonstration des deux propriétés précédentes. Selon l'inégalité de Cauchy-Schwartz, nous avons :

$$|\langle \vec{O}^i, \vec{O}^{i'} \rangle| \leq \sqrt{\langle \vec{O}^i, \vec{O}^i \rangle \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle} \Rightarrow \left| \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\sqrt{\langle \vec{O}^i, \vec{O}^i \rangle \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}} \right| \leq 1.$$

Or nous savons que :

$$\lim_{t \rightarrow 0} \mathcal{M}^t(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle) = \sqrt{\langle \vec{O}^i, \vec{O}^i \rangle \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}.$$

De plus, en raison de la propriété prop-3.1.16, nous avons :

$$0 < \mathcal{M}^s(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle) \leq \mathcal{M}^t(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle) \quad \forall s < t.$$

Nous en déduisons la relation suivante :

$$\frac{1}{\mathcal{M}^s(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)} \geq \frac{1}{\mathcal{M}^t(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)} > 0 \quad \forall s < t.$$

Nous avons finalement :

$$1 \geq \left| \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\mathcal{M}^s(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)} \right| \geq \left| \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\mathcal{M}^t(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)} \right| \geq 0 \quad \forall s < t.$$

ou encore :

$$0 \leq |\mathbf{S}_{ii'}^t| \leq |\mathbf{S}_{ii'}^s| \leq 1 \quad \forall s < t$$

Dans la mesure où $\mathbf{S}_{ii'}^t$ et $\mathbf{S}_{ii'}^s$ sont de même signe, celui de $\theta_{ii'}$, nous avons bien :

$$0 \leq \mathbf{S}_{ii'}^t \leq \mathbf{S}_{ii'}^s \text{ ou } 0 \geq \mathbf{S}_{ii'}^t \geq \mathbf{S}_{ii'}^s \quad \forall s < t.$$

□

Propriété 3.1.20 Expression des indices de similarité d'ordre t en fonction des paramètres géométriques

Soient :

- $\theta_{ii'} = \cos(\widehat{\vec{O}^i \vec{O}^{i'}})$ tel que $-1 \leq \theta_{ii'} \leq 1$
- $\gamma_{ii'} = \frac{\max(\|\vec{O}^i\|, \|\vec{O}^{i'}\|)}{\min(\|\vec{O}^i\|, \|\vec{O}^{i'}\|)}$ tel que $1 \leq \gamma_{ii'}$
- $\gamma_i^{i'} = \frac{\|\vec{O}^{i'}\|}{\|\vec{O}^i\|}$ tel que $-\infty < \gamma_i^{i'} < +\infty$

Les indices de similarité d'ordre t entre \vec{O}^i et $\vec{O}^{i'}$ peuvent être exprimés selon les paramètres précédents de la manière suivante :

$$\boxed{\mathbf{S}_{ii'}^t = \frac{\theta_{ii'}}{\mathcal{M}^t(\gamma_i^{i'}, \gamma_{i'}^i)}} \quad (3.1.35)$$

ou encore :

$$\boxed{\mathbf{S}_{ii'}^t = \theta_{ii'} \left(\frac{2^{1/t} \gamma_{ii'}}{(1 + \gamma_{ii'}^{2t})^{1/t}} \right)} \quad (3.1.36)$$

Démonstration

$$\begin{aligned} \mathbf{S}_{ii'}^t &= \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\mathcal{M}^t(\langle \vec{O}^i, \vec{O}^i \rangle, \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)} \\ &= \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\left[\frac{1}{2} (\langle \vec{O}^i, \vec{O}^i \rangle)^t + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle^t \right]^{1/t}} \\ &= \frac{\theta_{ii'} \|\vec{O}^i\| \|\vec{O}^{i'}\|}{\left[\frac{1}{2} (\|\vec{O}^i\|^{2t} + \|\vec{O}^{i'}\|^{2t}) \right]^{1/t}} \\ &= \frac{\theta_{ii'}}{\left[\frac{1}{2} \left(\frac{\|\vec{O}^i\|^{2t} + \|\vec{O}^{i'}\|^{2t}}{\|\vec{O}^i\|^t \|\vec{O}^{i'}\|^t} \right) \right]^{1/t}} \\ &= \frac{\theta_{ii'}}{\left[\frac{1}{2} \left(\left(\frac{\|\vec{O}^i\|}{\|\vec{O}^{i'}\|} \right)^t + \left(\frac{\|\vec{O}^{i'}\|}{\|\vec{O}^i\|} \right)^t \right) \right]^{1/t}} \\ &= \frac{\theta_{ii'}}{\mathcal{M}^t(\gamma_i^{i'}, \gamma_{i'}^i)} \end{aligned}$$

Puis en supposant, sans perte de généralité, que $\|\vec{O}^{i'}\| \geq \|\vec{O}^i\|$, nous avons $\gamma_{ii'} = \gamma_i^{i'}$ et donc $\gamma_i^{i'} = 1/\gamma_{ii'}$. Nous avons alors :

$$\begin{aligned}
 S_{ii'}^t &= \frac{\theta_{ii'}}{\mathcal{M}^t(\gamma_{ii'}, 1/\gamma_{ii'})} \\
 &= \frac{\theta_{ii'}}{[\frac{1}{2}((\gamma_{ii'})^t + (1/\gamma_{ii'})^t)]^{1/t}} \\
 &= \frac{\theta_{ii'}}{[\frac{1}{2}((1+(\gamma_{ii'})^{2t})/(\gamma_{ii'})^t)]^{1/t}} \\
 &= \theta_{ii'} \left(\frac{2^{1/t} \gamma_{ii'}}{(1+\gamma_{ii'}^{2t})^{1/t}} \right)
 \end{aligned}$$

□

Afin de justifier le bien fondé des indices de similarité d'ordre t , nous donnons ci-dessous une deuxième interprétation géométrique de ces derniers.

Considérons deux vecteurs représentés dans un espace à M dimensions. En statistiques appliquées, nous savons que la projection orthogonale du vecteur \vec{O}^i sur le vecteur $\vec{O}^{i'}$ constitue, au sens d'une distance euclidienne, la meilleure approximation de \vec{O}^i en fonction de $\vec{O}^{i'}$. Cette approche géométrique est couramment utilisée dans les techniques d'analyse de données telles que la régression linéaire ou l'analyse canonique notamment dans le cas où les vecteurs sont représentatifs de variables quantitatives [Sap06].

Nous avons la relation suivante, qui est illustrée sur la figure fig-3.1 :

$$\vec{O}^i = \underbrace{\left(\frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle} \right) \vec{O}^{i'}}_{\text{Vecteur issu de la projection de } \vec{O}^i \text{ sur } \vec{O}^{i'}} + \underbrace{\left[\vec{O}^i - \left(\frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle} \right) \vec{O}^{i'} \right]}_{\text{Vecteur de l'erreur entre } \vec{O}^i \text{ et son vecteur issu de la projection sur } \vec{O}^{i'}} \quad (3.1.37)$$

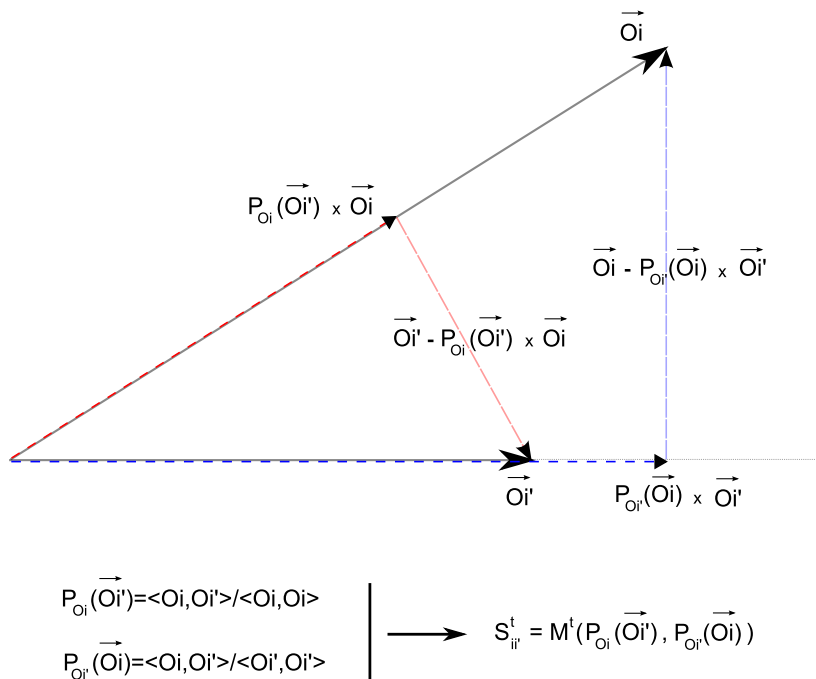


FIGURE 3.1 – Interprétation géométrique de la similarité d'ordre t entre deux vecteurs

Notons par $P_{\vec{O}^{i'}}(\vec{O}^i)$ le coefficient de colinéarité du vecteur issu de la projection de \vec{O}^i sur

$\vec{O}^{i'}$:

$$P_{\vec{O}^{i'}}(\vec{O}^i) = \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle} \quad (3.1.38)$$

Ce dernier exprime le lien entre $\vec{O}^{i'}$ et la meilleure approximation de \vec{O}^i en fonction de $\vec{O}^{i'}$. Ainsi, plus le vecteur \vec{O}^i est semblable au vecteur $\vec{O}^{i'}$ en terme de direction et de longueur, plus $P_{\vec{O}^{i'}}(\vec{O}^i)$ est proche de 1.

Le coefficient de colinéarité donne une mesure de la ressemblance entre le vecteur projeté et le vecteur sur lequel on le projette. Dans la mesure où il s'agit d'une mesure non symétrique, pour mesurer la similarité entre deux vecteurs en se basant sur cette intuition géométrique il faut considérer à la fois $P_{\vec{O}^{i'}}(\vec{O}^i)$ et $P_{\vec{O}^i}(\vec{O}^{i'})$. Une manière naturelle de combiner ces deux quantités, est de prendre une moyenne de ces dernières. En fait, nous avons la propriété suivante :

Propriété 3.1.21 *Expression de $S_{ii'}^t$ en termes de moyennes de coefficients de colinéarité*

$$\begin{aligned} S_{ii'}^t &= \mathcal{M}^{-t}(P_{\vec{O}^{i'}}(\vec{O}^i), P_{\vec{O}^i}(\vec{O}^{i'})) \quad \forall t > 0 \\ &= \mathcal{M}^{-t} \left(\frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}, \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^i, \vec{O}^i \rangle} \right) \quad \forall t > 0 \end{aligned} \quad (3.1.39)$$

Démonstration

$$\begin{aligned} \mathcal{M}^{-t}(P_{\vec{O}^{i'}}(\vec{O}^i), P_{\vec{O}^i}(\vec{O}^{i'})) &= \left[\frac{1}{2} \left((P_{\vec{O}^{i'}}(\vec{O}^i))^{-t} + (P_{\vec{O}^i}(\vec{O}^{i'}))^{-t} \right) \right]^{-1/t} \\ &= \left[\frac{1}{2} \left(\left(\frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^i, \vec{O}^i \rangle} \right)^{-t} + \left(\frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle} \right)^{-t} \right) \right]^{-1/t} \\ &= \left[\frac{1}{2} \left(\left(\frac{\langle \vec{O}^i, \vec{O}^i \rangle}{\langle \vec{O}^i, \vec{O}^{i'} \rangle} \right)^t + \left(\frac{\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{\langle \vec{O}^i, \vec{O}^{i'} \rangle} \right)^t \right) \right]^{-1/t} \\ &= \left[\frac{(\langle \vec{O}^i, \vec{O}^i \rangle)^t + (\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle)^t}{2(\langle \vec{O}^i, \vec{O}^{i'} \rangle)^t} \right]^{-1/t} \\ &= \left[\frac{(\langle \vec{O}^i, \vec{O}^{i'} \rangle)^t}{\frac{1}{2}(\langle \vec{O}^i, \vec{O}^i \rangle^t + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle^t)} \right]^{1/t} \\ &= \frac{\langle \vec{O}^i, \vec{O}^{i'} \rangle}{\left[\frac{1}{2}(\langle \vec{O}^i, \vec{O}^i \rangle^t + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle^t) \right]^{1/t}} \\ &= S_{ii'}^t \end{aligned}$$

□

Remarque 3.1.2 *Lien entre les coefficients de colinéarité et les indices de Rappel et de Précision*

Dans le cas de données binaires et plus particulièrement dans le cadre de la "recherche d'information"⁸, deux autres indicateurs sont couramment utilisés pour l'évaluation d'un moteur de recherche : le rappel et la précision. Considérons par \vec{O}^i un vecteur binaire caractérisant un ensemble de "vraies" réponses correspondant à une requête et par $\vec{O}^{i'}$ un deuxième vecteur binaire correspondant cette fois-ci à un ensemble de réponses restitué par un moteur de recherche vis à vis de la même requête. Alors les coefficients rappel et précision entre ces deux vecteurs notés respectivement $R(\vec{O}^i, \vec{O}^{i'})$ et $P(\vec{O}^i, \vec{O}^{i'})$ sont définis par :

$$R(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + 10_{ii'}} \quad (3.1.40)$$

$$P(\vec{O}^i, \vec{O}^{i'}) = \frac{11_{ii'}}{11_{ii'} + 01_{ii'}} \quad (3.1.41)$$

8. sur internet par exemple

Nous avons en fait la propriété que :

$$R(\vec{O}^i, \vec{O}^{i'}) = P_{\vec{O}^i}(\vec{O}^{i'}) \quad (3.1.42)$$

$$P(\vec{O}^i, \vec{O}^{i'}) = P_{\vec{O}^{i'}}(\vec{O}^i) \quad (3.1.43)$$

Dans le cadre des données binaires, l'expression des indices de Dice, d'Ochiaï et de Kulczynski en fonction d'une moyenne des coefficients de rappel et de précision a été donnée par J.F. Marcotorchino dans [Mar03].

3.1.2.5 Critère de Condorcet et indices de similarité d'ordre t entre vecteurs quelconques

Dans le paragraphe 3.1.1.3, nous avons proposé une extension du critère de Condorcet à des indices de similarité dans le cadre des vecteurs binaires. **Nous proposons ici d'étudier succinctement, des critères de Condorcet qui sont fonctions des indices de similarité d'ordre t définis au paragraphe précédent. Dans la mesure où cette approche s'inscrit dans le cadre de vecteurs quelconques, elle peut être considérée comme étant une extension du critère de Condorcet à des vecteurs quelconques.**

En raison de l'équation (3.1.23), nous cherchons à maximiser le critère de Condorcet suivant :

$$\max_X \text{Condorcet}(\mathbf{S}^t, X) = \sum_{i,i'} (\mathbf{S}_{ii'}^t - \frac{1}{2}) X_{ii'} \quad (3.1.44)$$

où $t > 0$ et X est une matrice relationnelle vérifiant les contraintes relationnelles d'une partition⁹.

Nous nous intéressons aux domaines relatifs à $\theta_{ii'}$ et $\gamma_{ii'}$ pour lesquels la contribution $(\mathbf{S}_{ii'}^t - \frac{1}{2})$ est positive.

Explicitons (3.1.44), en remplaçant $\mathbf{S}_{ii'}^t$ par son expression en fonction des paramètres géométriques. Nous obtenons le critère ci-dessous :

$$\max_X \sum_{i,i'} \left(\theta_{ii'} \left(\frac{2^{1/t} \gamma_{ii'}}{(1 + \gamma_{ii'}^{2t})^{1/t}} \right) - \frac{1}{2} \right) X_{ii'} \quad (3.1.45)$$

Nous avons alors :

$$\begin{aligned} \left(\theta_{ii'} \left(\frac{2^{1/t} \gamma_{ii'}}{(1 + \gamma_{ii'}^{2t})^{1/t}} \right) - \frac{1}{2} \right) \geq 0 &\Leftrightarrow \theta_{ii'} \left(\frac{2^{1/t} \gamma_{ii'}}{(1 + \gamma_{ii'}^{2t})^{1/t}} \right) \geq \frac{1}{2} \\ &\Leftrightarrow \theta_{ii'} \geq \frac{(1 + \gamma_{ii'}^{2t})^{1/t}}{2^{1+1/t} \gamma_{ii'}} \end{aligned}$$

Nous nous intéressons donc à la courbe suivante qui nous donne la frontière délimitant la région que nous étudions :

$$\theta_{ii'} = f_t(\gamma_{ii'}) = \frac{(1 + \gamma_{ii'}^{2t})^{1/t}}{2^{1+1/t} \gamma_{ii'}} \quad (3.1.46)$$

Dans la figure fig-3.2, nous avons tracé la frontière $f_t(\gamma_{ii'})$ pour $t \rightarrow 0$ et pour $t \in \{0.5, 1.0, 1.5, 2.0, 2.5, 3.0\}$.

A chaque $\gamma_{ii'}$ en abscisse, pour que $\mathbf{S}_{ii'}^t$ soit supérieur à $1/2$, il faut que $\theta_{ii'}$, en ordonnée, soit de valeur supérieure à la courbe considérée. Par exemple, lorsque $t = 0.5$, et lorsque $\gamma_{ii'} = 3$, il faut que $\theta_{ii'}$ soit approximativement supérieur à 0.65 pour que la contribution soit positive. Par ailleurs, les courbes indiquent, la valeur de $\gamma_{ii'}$ à partir de laquelle la contribution ne peut plus être positive quelque soit la valeur de $\theta_{ii'}$. En effet, nous voyons par exemple, que pour $t = 1$, si $\gamma_{ii'} > 3.7$, alors il faudrait un cosinus supérieur à 1 pour que la contribution soit positive ce qui n'est pas possible.

9. cf page 36

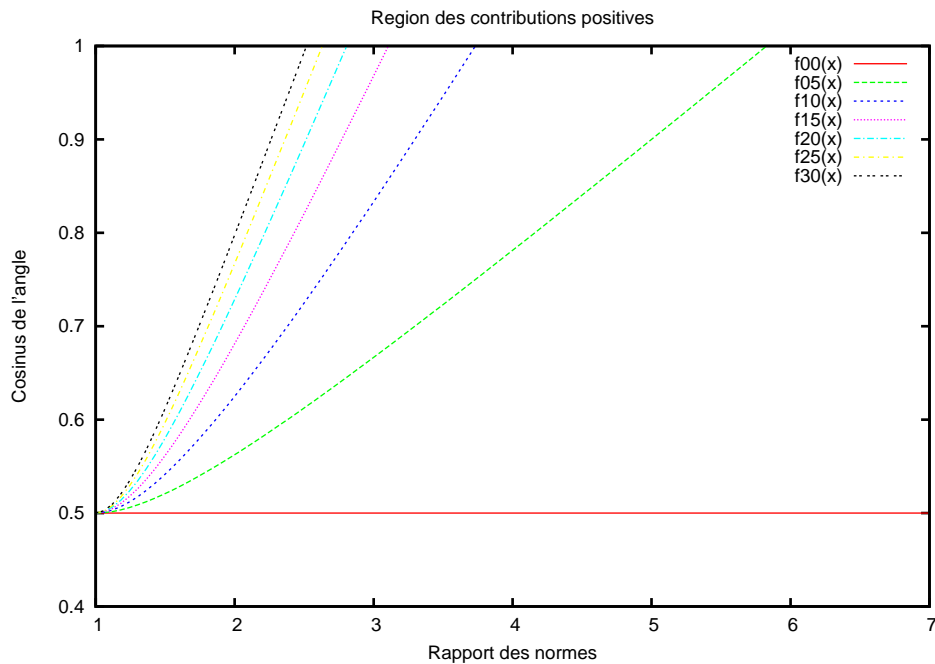


FIGURE 3.2 – Domaines des contributions positives des indices de similarité d'ordre t

La borne $1/2$ dite borne de Solomon et Fortier [SF66], qui est issue de l'interprétation du critère de Condorcet, peut paraître sévère. En pratique il est utile de pouvoir :

1. soit permettre à un utilisateur de fixer lui même ce paramètre sachant qu'il doit être compris entre 0 et 1
2. soit donner à ce paramètre une valeur dépendante des données étudiées telle qu'une tendance centrale

Dans le premier cas, une borne égale à 0 peut paraître satisfaisante dans la mesure où les indices de similarité peuvent prendre des valeurs négatives. Ce sera en particulier le cas pour des données qui sont centrées puisque le signe de l'indice dépend de l'angle formé par les deux vecteurs centrés. Nous rapprochons cette remarque de la formule (3.2.28) obtenue au paragraphe 3.2.2.4 page 97 concernant le lien entre le critère de Condorcet et les critères inertiels.

Dans le deuxième cas, nous proposons d'utiliser une tendance centrale, telle qu'une moyenne ou la médiane qui est calculée sur l'ensemble des paires d'objets. Cela revient à dire que deux objets O^i et $O^{i'}$ sont potentiellement dans la même classe si leur indice de similarité est supérieur à la moyenne des indices de similarité calculés sur l'ensemble des paires d'objets. Cette approche est relative au principe d'écart à une tendance centrale que nous mettons en évidence ultérieurement au paragraphe 3.2.1.1.

En conclusion de ce paragraphe, une première extension de l'approche relationnelle appliquée à des variables quantitatives ou non binaires consiste donc à résoudre le problème suivant :

$$\begin{array}{l}
 \max_X \sum_{i,i'} (\mathbf{S}_{ii'}^t - \mathcal{M}(\{\mathbf{S}_{ii'}^t\})) X_{ii'} \\
 \text{sous les contraintes} \\
 X_{ii'} \in \{0, 1\} \quad \text{(binarité)} \\
 X_{ii} = 1 \quad \forall i = 1, \dots, N \quad \text{(réflexivité)} \\
 X_{i'i} = X_{ii'} \quad \forall i, i' = 1, \dots, N \quad \text{(symétrie)} \\
 X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 \quad \forall i, i', i'' = 1, \dots, N \quad \text{(transitivité)}
 \end{array} \tag{3.1.47}$$

où $\mathcal{M}(\{\mathbf{S}_{ii'}^t\})$ est une tendance centrale telle une moyenne généralisée d'ordre t' , $\mathcal{M}^{t'}(\{\mathbf{S}_{ii'}^t\})$, ou telle la médiane, qui est calculée sur l'ensemble des similarités $\{\mathbf{S}_{ii'}^t\}_{i,i'=1,\dots,N}$.

Notons toutefois, qu'il est nécessaire, dans un premier temps, de réduire les données afin de ne pas avoir d'effets indésirables d'unités de mesure hétérogènes donnant aux variables des ordres de grandeurs parfois très différentes. D'ailleurs, comme le rapport des normes entre vecteur devient un paramètre contribuant à la mesure de similarité, il est effectivement, important d'être indépendant des unités de mesures.

En revanche, le centrage de données est plus discutable dans le sens où le centre de gravité du nuage de points devient la tendance centrale de référence à partir de laquelle notamment, nous pouvons dire si des objets sont similaires¹⁰, antisimilaires¹¹ ou encore indépendants¹².

Nous verrons dans les sous-sections 3.2.1 et 3.2.2, que le principe d'écart à une tendance centrale représentée par une moyenne, est vérifié pour de nombreux critères. Ce principe ne se résume pas uniquement au centrage de données avant traitement, mais concerne de manière plus générale, le centrage des mesures de similarité par une moyenne de ces dernières.

3.2 Extension des critères d'association maximale à des variables quantitatives et la notion d'écart à une tendance centrale

Après avoir étudié la notion de similarité entre objets, nous nous intéressons à présent aux critères de partitionnement qui fondent avec les algorithmes de classification¹³ et les mesures de similarité les trois notions fondamentales en classification automatique.

L'objectif de cette section est de définir de nouveaux critères de partitionnement qui sont valides à la fois pour des objets décrits par des variables qualitatives et des variables quantitatives. Dans un premier temps, notre approche est la même que précédemment : nous étendons des critères initialement définis pour des données qualitatives à des données quantitatives en considérant des interprétations géométriques.

Dans un second temps, nous considérons l'approche duale qui consiste à montrer que les critères initialement définis pour des données quantitatives peuvent être interprétés en termes relationnels et aboutir à des critères d'association maximale. Nous montrons en particulier que les critères inertiels sont proches du critère de Condorcet et lorsque les données sont centrées ils sont identiques à de nombreux autres critères d'association maximale.

3.2.1 Critères d'association maximale et écart à une tendance centrale

3.2.1.1 Principe d'écart à une tendance centrale

Nous commençons par donner des caractéristiques générales des critères de partitionnement. En effet, nous montrons que les critères issus du tableau tab-1.5 vérifient le principe d'écart à une tendance centrale exposé ci-dessous.

Les critères de partitionnement peuvent être vus comme utilisant un indice de similarité particulier que l'on compare à une tendance centrale, le plus souvent une moyenne de ces

10. ie $\theta_{ii'} > 0$

11. ie $\theta_{ii'} < 0$

12. ie $\theta_{ii'} = 0$

13. cf chapitre 2

mesures de similarité.

De cette manière, les objets O^i et $O^{i'}$ sont potentiellement dans une même classe lorsque leur mesure de similarité est supérieure à cette tendance centrale des mesures de similarité.

Nous résumons ces résultats, qui sont faciles à vérifier, dans le tableau tab-3.8.

Critère d'association maximale	Indice de similarité associé	Tendance centrale associée
Belson(\mathbf{C}, X)	$\mathbf{S}_{ii'}^B = \mathbf{C}_{ii'} - \frac{\mathbf{C}_{i.} + \mathbf{C}_{.i'}}{N} + \frac{\mathbf{C}_{..}}{N^2}$	$\mathcal{M}^1(\{\mathbf{S}_{ii'}^B\}) = 0$
Rand(\mathbf{C}, X)	$\mathbf{S}_{ii'}^R = \mathbf{C}_{ii'}$	$\frac{1}{2} \mathcal{M}^1(\mathbf{S}_{ii}^R, \mathbf{S}_{i'i'}^R)$ (locale)
Jordan(\mathbf{C}, X)	$\mathbf{S}_{ii'}^J = \mathbf{C}_{ii'} - \frac{\mathbf{C}_{i.} + \mathbf{C}_{.i'}}{2N}$ ou $\mathbf{S}_{ii'}^{J'} = \mathbf{C}_{ii'}$	$\mathcal{M}^1(\{\mathbf{S}_{ii'}^J\}) = 0$ ou $\mathcal{M}^1(\frac{\mathbf{S}_{i.}^J}{N}, \frac{\mathbf{S}_{.i'}^J}{N})$
LM(\mathbf{C}, X)	$\mathbf{S}_{ii'}^{LM} = \hat{\mathbf{C}}_{ii'}$	$\mathcal{M}^1(\{\mathbf{S}_{ii'}^{LM}\}) = \frac{M}{N}$
E(\mathbf{C}, X)	$\mathbf{S}_{ii'}^E = \mathbf{C}_{ii'}$	$\mathcal{M}^1(\{\mathbf{S}_{ii'}^E\}) = \frac{1}{N^2} \sum_{i,i'} \mathbf{C}_{ii'}$

TABLE 3.8 – Principe d'écart à une tendance centrale vérifié par les critères de partitionnement issus de l'association maximale

3.2.1.2 Interprétations géométriques des critères d'association maximale, extension à des données quelconques

Nous donnons à présent des interprétations géométriques à ces critères d'association maximale ce qui permet leur extension à des données quantitatives.

Nous nous appuyons sur les interprétations géométriques données dans le tableau tab-3.5 page 77. Nous obtenons alors les expressions en termes de produits scalaires des critères de partitionnement issus de l'agrégation maximale. Nous présentons ces résultats dans le tableau tab-3.9 où \vec{G} est le centre de gravité défini par :

$$\vec{G} = \frac{1}{N} \sum_i \vec{O}^i \tag{3.2.1}$$

Les critères de partitionnement ainsi définis, peuvent donc être appliqués à des données quantitatives. Dans ce contexte \mathbf{C} représente la matrice de comparaisons par paires des produits scalaires et X est une matrice relationnelle représentative de la partition recherchée.

Dans le cas du critère de Light et Margolin, LM, nous avons interprété l'expression relationnelle $\hat{\mathbf{C}}_{ii'} = \sum_j \frac{\mathbf{K}_{ij} \mathbf{K}_{i'j}}{\mathbf{K}_{.j}}$ comme étant le produit scalaire des objets dont les données ont été réduites par la norme des variables.

Critère d'association maximale	↔	Expression en termes de produits scalaires
Belson(\mathbf{C}, X)	=	$\sum_{i,i'} (\langle \vec{O}^i - \vec{G}, \vec{O}^{i'} - \vec{G} \rangle) X_{ii'}$
Rand(\mathbf{C}, X)	=	$\frac{1}{N^2} \left[2 \sum_{i,i'} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \frac{1}{2} \frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) X_{ii'} - \sum_{i,i'} \left(\frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} - \langle \vec{O}^i, \vec{O}^{i'} \rangle \right) \right]$
Jordan(\mathbf{C}, X)	=	$\frac{1}{N} \sum_{i,i'} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \frac{\langle \vec{O}^i, \vec{G} \rangle + \langle \vec{G}, \vec{O}^{i'} \rangle}{2} \right) X_{ii'}$
LM(\mathbf{C}, X)	=	$\sum_{i,i'} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \frac{1}{N^2} \sum_{i,i'} \langle \vec{O}^i, \vec{O}^{i'} \rangle \right) X_{ii'}$
E(\mathbf{C}, X)	=	$\sum_{i,i'} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \langle \vec{G}, \vec{G} \rangle \right) X_{ii'}$

TABLE 3.9 – Critères de partitionnement issus de l'association maximale en termes de produits scalaires

Posons :

$$\vec{O}^i = (\hat{O}_1^i, \dots, \hat{O}_j^i, \dots, \hat{O}_P^i) \text{ où } \hat{O}_j^i = \frac{O_j^i}{\sqrt{\sum_i (O_j^i)^2}} \quad (3.2.2)$$

Nous avons alors :

$$\hat{\mathbf{C}}_{ii'} = \langle \vec{O}^i, \vec{O}^{i'} \rangle \quad (3.2.3)$$

Lorsque les données sont relatives à des variables qualitatives nous avons bien l'égalité ci-dessus étant donné que $\sqrt{\mathbf{K}_{\cdot j}} = \sqrt{\sum_j \mathbf{K}_{ij}^2}$ en raison de la binarité des termes \mathbf{K}_{ij} , $i = 1, \dots, N$, $j = 1, \dots, P$.

Lorsque les données sont relatives à des variables quantitatives, nous interprétons donc la similarité de rareté comme étant un produit scalaire de métrique diagonale de terme général $\mathbf{M}_{jj}^{LM} = 1/\|\vec{D}^j\|^2$, $j = 1, \dots, P$.

Nous savons par ailleurs, qu'à toute métrique euclidienne quelconque, correspond une mé-

trique euclidienne canonique $M = I_P$ associée à des données transformées [Sap06] :

$$\begin{aligned} {}^t\vec{O}^i \cdot I_P \cdot \vec{O}^{i'} &= \sum_j \frac{O_j^i}{\|\vec{D}^j\|} \frac{O_j^{i'}}{\|\vec{D}^j\|} \\ &= \sum_j \frac{O_j^i O_j^{i'}}{\|\vec{D}^j\|^2} \\ &= {}^t\vec{O}^i \cdot M^{LM} \cdot \vec{O}^{i'} \end{aligned}$$

Pour conclure ce paragraphe, nous dirons qu'une deuxième extension de l'AR dans la classification automatique d'objets décrits par des variables quantitatives consiste à résoudre le programme linéaire suivant :

$\max_X \Delta(C, X)$			
sous les contraintes			
$X_{ii'} \in \{0, 1\}$		(binarité)	(3.2.4)
$X_{ii} = 1$	$\forall i = 1, \dots, N$	(réflexivité)	
$X_{i'i} = X_{ii'}$	$\forall i, i' = 1, \dots, N$	(symétrie)	
$X_{ii'} + X_{i'i''} - X_{ii''} \leq 1$	$\forall i, i', i'' = 1, \dots, N$	(transitivité)	

où $\Delta(C, X)$ est un critère d'association maximal et où C est la matrice carrée des produits scalaires entre objets.

Nous avons montré que les critères de partitionnement $\Delta(C, X)$, dépendaient d'un indice de similarité particulier et qu'ils vérifiaient le principe d'écart à une tendance centrale. Nous pouvons donc considérer que **les modèles par programmation linéaire (3.1.47) et (3.2.4) relèvent du même modèle général.**

En résumé, l'ensemble des modèles proposés forme une extension à des variables quantitatives de l'application de l'AR en classification automatique.

3.2.2 Liens entre les critères inertiels et le critère de Condorcet

Pour aller plus loin dans l'unification des critères de partitionnement, nous étudions ici les critères inertiels qui sont en fait, les critères de partitionnement les plus couramment employés dans le traitement de données quantitatives.

Nous cherchons à exprimer ces derniers en termes relationnels et nous aboutissons à un formalisme du critère de la différence inertielle proche de celui du critère de Condorcet. Nous reprenons les travaux de J.F. Marcotorchino qui dans [Mar91a], a montré les liens entre le critère de la différence inertielle et le critère de Condorcet pondéré¹⁴ dans le cadre des variables qualitatives.

Il justifie ainsi l'emploi de la classification des objets issue de la maximisation du critère de Condorcet pondéré en complément des projections des objets sur des axes orthogonaux à inertie maximale donnés par une Analyse Factorielle des Correspondances Multiples (AFCM).

En reprenant les articles de I.C. Lerman [Ler79] et de N. Howard, [How69], J.F. Marcotorchino aboutit à la conclusion suivante :

"[...] si nous avons un nuage de points éclaté en paquets au niveau des projections sur les premiers axes d'une AFCM, c'est à dire une "bonne" situation d'interprétabilité, il y a de

14. ie critère de Condorcet pour lequel la matrice des similarités est donné par \hat{C} la matrice des similarités de rareté

grandes chances pour qu'une partition de la population en classes optimales vis à vis du critère d'inertie totale d'inter-classes du fait de sa possibilité de capter l'information en profondeur (c'est à dire non restreinte aux premiers axes) nous donne une interprétation meilleure de la structure des données."

"[...] si nous étions capables de fournir une partition "optimale" de la population vis à vis du critère de l'inertie inter-classes, nous aurions accompli une tâche fondamentale dans l'aide à l'interprétation d'une AFCM".

Ainsi, l'objet de cette sous-section est de généraliser cette démarche au cas du traitement des données quantitatives par une Analyse en Composantes Principales (ACP). **Nous montrons qu'il est utile pour une meilleure interprétation, d'accompagner la représentation des données dans un espace réduit et optimal donné par l'ACP de données quantitatives, de la partition solution de (3.2.4) où $\Delta(\mathbf{C}, X) = \text{Condorcet}(\mathbf{C}, X)$ ou $\text{Rand}(\mathbf{C}, X)$.**

Nous rappelons ci-dessous les critères inertiels dans leur formalisme classique. Ces critères sont ceux sur lesquels est fondée l'ACP.

Soit donc un ensemble de N objets \mathbb{O} décrits par un ensemble de P descripteurs. Ces objets définissent dans l'espace de description des vecteurs notés $\vec{O}^i, i = 1, \dots, N$ pondérés par des poids¹⁵ μ_i tels que $\sum_i \mu_i = 1$.

Considérons, une partition de ces objets notée \mathcal{R} constituée de p classes notées $\mathcal{R}_j, j = 1, \dots, p$. Notons par $p_j = \#\mathcal{R}_j$ le cardinal de la classe \mathcal{R}_j . Nous avons donc $\sum_j p_j = N$. Notons de plus $\nu_j = \sum_{O^i \in \mathcal{R}_j} \mu_i$, le poids de la classe \mathcal{R}_j . Nous avons $\sum_j \nu_j = 1$

Définissons alors par \vec{G}^j le centre de gravité relatif à la classe \mathcal{R}_j :

$$\vec{G}^j = \frac{1}{\nu_j} \sum_{O^i \in \mathcal{R}_j} \mu_i \vec{O}^i.$$

Nous avons alors \vec{G} le centre de gravité de l'ensemble du nuage de points qui vaut :

$$\vec{G} = \sum_i \mu_i \vec{O}^i = \sum_j \nu_j \vec{G}^j.$$

A partir de ces différentes définitions nous pouvons présenter les différents types d'inerties.

Définition 3.2.1 Inertie totale

Nous noterons par \mathbb{IT} , l'inertie totale d'un nuage de points. Celle-ci est définie selon l'équation suivante :

$$\mathbb{IT} = \sum_{i=1}^N \mu_i \|\vec{O}^i - \vec{G}\|^2 \quad (3.2.5)$$

Nous avons aussi la définition équivalente suivante :

$$\mathbb{IT} = \frac{1}{2} \sum_{i=1}^N \sum_{i'=1}^N \mu_i \mu_{i'} \|\vec{O}^i - \vec{O}^{i'}\|^2 \quad (3.2.6)$$

Définition 3.2.2 Inertie intra-classe de la classe \mathcal{R}_j

Nous noterons par \mathbb{IW}_j , l'inertie intra-classe de la classe \mathcal{R}_j . Celle-ci est définie selon l'équation suivante :

$$\mathbb{IW}_j = \sum_{O^i \in \mathcal{R}_j} \mu_i \|\vec{O}^i - \vec{G}^j\|^2 \quad (3.2.7)$$

Définition 3.2.3 Inertie intra-classe d'une partition \mathcal{R}

Nous noterons par $\mathbb{IW}(\mathcal{R})$, l'inertie intra-classe d'une partition \mathcal{R} . Celle-ci est la somme des inerties intra-classes de chaque classe $\mathcal{R}_j, j = 1, \dots, p$ de la partition :

$$\mathbb{IW}(\mathcal{R}) = \sum_{j=1}^p \mathbb{IW}_j \quad (3.2.8)$$

15. le cas classique consiste à prendre $\mu_i = 1/N \quad \forall i = 1, \dots, N$

Définition 3.2.4 Inertie inter-classe d'une partition \mathcal{R}

Nous noterons par $\mathbb{I}\mathbb{B}(\mathcal{R})$, l'inertie inter-classe d'une partition \mathcal{R} . Celle-ci est définie selon l'équation suivante :

$$\mathbb{I}\mathbb{B}(\mathcal{R}) = \sum_{j=1}^p \nu_j \|\vec{G}^j - \vec{G}\|^2 \quad (3.2.9)$$

Théorème 3.2.1 Théorème de Huyghens

Peu importe la partition \mathcal{R} considérée, nous avons toujours la relation suivante :

$$\mathbb{I}\mathbb{T} = \mathbb{I}\mathbb{W}(\mathcal{R}) + \mathbb{I}\mathbb{B}(\mathcal{R}) \quad (3.2.10)$$

Dans ce qui suit, nous donnons les expressions relationnelles des différents inerties présentées ci-dessus. Partant des expressions appelées ci-dessus, la démarche consiste :

1. à développer les distances euclidiennes en termes de produits scalaires
2. puis à introduire la matrice relationnelle X représentative de la partition \mathcal{R} , ce qui permet de faire disparaître les sommations sur j
3. enfin à utiliser les correspondances entre produits scalaires et termes relationnels données dans la table [tab-3.5](#)

3.2.2.1 Expression relationnelle de l'inertie totale

Propriété 3.2.1 Expression relationnelle de l'inertie totale d'un nuage de points

Nous avons les différentes expressions équivalentes suivantes de $\mathbb{I}\mathbb{T}$:

$$\mathbb{I}\mathbb{T} = \sum_i \mu_i \mathbf{C}_{ii} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \quad (3.2.11)$$

$$\mathbb{I}\mathbb{T} = \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii'} \quad (3.2.12)$$

Dans le cas classique où $\mu_i = \frac{1}{N} \quad \forall i = 1, \dots, N$, nous avons :

$$\mathbb{I}\mathbb{T} = \frac{1}{N} \sum_i \mathbf{C}_{ii} - \frac{1}{N^2} \sum_{i,i'} \mathbf{C}_{ii'} \quad (3.2.13)$$

$$\mathbb{I}\mathbb{T} = \frac{1}{N^2} \sum_{i,i'} \bar{\mathbf{C}}_{ii'} \quad (3.2.14)$$

Démonstration

$$\begin{aligned} \mathbb{I}\mathbb{T} &= \frac{1}{2} \sum_{i,i'} \mu_i \mu_{i'} \|\vec{O}^i - \vec{O}^{i'}\|^2 \\ &= \frac{1}{2} \sum_{i,i'} \mu_i \mu_{i'} \langle \vec{O}^i - \vec{O}^{i'}, \vec{O}^i - \vec{O}^{i'} \rangle \\ &= \frac{1}{2} \sum_{i,i'} \mu_i \mu_{i'} \left(\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle - 2 \langle \vec{O}^i, \vec{O}^{i'} \rangle \right) \\ &= \sum_{i,i'} \mu_i \mu_{i'} \left(\frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} - \langle \vec{O}^i, \vec{O}^{i'} \rangle \right) \\ &= \sum_{i,i'} \mu_i \mu_{i'} \left(\frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} - \mathbf{C}_{ii'} \right) \\ &= \sum_i \mu_i \mathbf{C}_{ii} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \end{aligned}$$

Par ailleurs, nous avons $\frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} - \mathbf{C}_{ii'} = \bar{\mathbf{C}}_{ii}$ et donc :

$$\begin{aligned} \mathbb{IT} &= \sum_{i,i'} \mu_i \mu_{i'} \left(\frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} - \mathbf{C}_{ii'} \right) \\ &= \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii} \end{aligned}$$

□

3.2.2.2 Expression relationnelle de l'inertie intra-classes

Propriété 3.2.2 Expressions relationnelles de l'inertie intra-classe d'une partition \mathcal{R}

Nous avons les différentes expressions équivalentes suivantes de $\mathbb{IW}(\mathcal{R})$ où X est la matrice relationnelle représentative de \mathcal{R} :

$$\mathbb{IW}(X) = \sum_i \mu_i \mathbf{C}_{ii} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \quad (3.2.15)$$

$$\mathbb{IW}(X) = \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \quad (3.2.16)$$

Dans le cas classique où $\mu_i = \frac{1}{N} \quad \forall i = 1, \dots, N$, nous avons :

$$\mathbb{IW}(X) = \frac{1}{N} \sum_i \mathbf{C}_{ii} - \frac{1}{N} \sum_{i,i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{X_i} \right) \quad (3.2.17)$$

$$\mathbb{IW}(X) = \frac{1}{N} \sum_{i,i'} \bar{\mathbf{C}}_{ii'} \left(\frac{X_{ii'}}{X_i} \right) \quad (3.2.18)$$

Démonstration

Nous commençons par déterminer l'expression relationnelle de \mathbb{IW}_j :

$$\begin{aligned} \mathbb{IW}_j &= \sum_{O^i \in \mathcal{R}_j} \mu_i \|\bar{\mathbf{O}}^i - \bar{\mathbf{G}}^j\|^2 \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \langle \bar{\mathbf{O}}^i - \bar{\mathbf{G}}^j, \bar{\mathbf{O}}^i - \bar{\mathbf{G}}^j \rangle \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \left(\langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle + \langle \bar{\mathbf{G}}^j, \bar{\mathbf{G}}^j \rangle - 2 \langle \bar{\mathbf{O}}^i, \bar{\mathbf{G}}^j \rangle \right) \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \left(\langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle + \left\langle \frac{1}{\nu_j} \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'}, \frac{1}{\nu_j} \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \right\rangle \right. \\ &\quad \left. - 2 \left\langle \bar{\mathbf{O}}^i, \frac{1}{\nu_j} \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \right\rangle \right) \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \left(\langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle + \frac{1}{\nu_j^2} \langle \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'}, \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \rangle \right. \\ &\quad \left. - \frac{2}{\nu_j} \langle \bar{\mathbf{O}}^i, \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \rangle \right) \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle \\ &\quad + \frac{1}{\nu_j} \langle \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'}, \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \rangle \quad \text{car } \sum_{O^i \in \mathcal{R}_j} \mu_i = \nu_j \\ &\quad - \frac{2}{\nu_j} \langle \sum_{O^i \in \mathcal{R}_j} \mu_i \bar{\mathbf{O}}^i, \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \rangle \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle - \frac{1}{\nu_j} \langle \sum_{O^i \in \mathcal{R}_j} \mu_i \bar{\mathbf{O}}^i, \sum_{O^{i'} \in \mathcal{R}_j} \mu_{i'} \bar{\mathbf{O}}^{i'} \rangle \\ &= \sum_{O^i \in \mathcal{R}_j} \mu_i \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle - \frac{1}{\nu_j} \sum_{O^i \in \mathcal{R}_j} \sum_{O^{i'} \in \mathcal{R}_j} \mu_i \mu_{i'} \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^{i'} \rangle \end{aligned}$$

Nous déterminons $\mathbb{IW}(\mathcal{R})$ en sommant sur l'ensemble des classes \mathcal{R}_j :

$$\begin{aligned} \mathbb{IW}(\mathcal{R}) &= \sum_j \mathbb{IW}_j \\ &= \sum_i \mu_i \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle - \sum_j \frac{1}{\nu_j} \sum_{O^i \in \mathcal{R}_j} \sum_{O^{i'} \in \mathcal{R}_j} \mu_i \mu_{i'} \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^{i'} \rangle \end{aligned}$$

Nous introduisons les notations relationnelles en "remplaçant" les indices sous les sommes qui restreignent les sommations aux objets appartenant à une même classe par une matrice relationnelle X représentative de la partition \mathcal{R} et qui indique justement ce type d'appartenance.

$$\mathbb{IW}(X) = \sum_i \mu_i \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^i \rangle - \sum_{i,i'} \frac{1}{\nu_{j,i}} \mu_i \mu_{i'} \langle \bar{\mathbf{O}}^i, \bar{\mathbf{O}}^{i'} \rangle X_{ii'} \quad \text{où } \nu_{j,i} \text{ est le poids de la classe } \mathcal{R}_j \text{ à}$$

laquelle appartient l'objet O^i .

Nous remarquons de plus que le poids $\nu_{j,i}$ de la classe \mathcal{R}_j de O^i est donné par $\sum_{i''} \mu_{i''} X_{ii''}$. Nous avons alors :

$$\mathbb{I}W(X) = \sum_i \mu_i \langle \vec{O}^i, \vec{O}^i \rangle - \sum_{i,i'} \mu_i \mu_{i'} \langle \vec{O}^i, \vec{O}^{i'} \rangle \frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}}$$

En prenant $\mathbf{C}_{ii'} = \langle \vec{O}^i, \vec{O}^{i'} \rangle$, nous aboutissons à la formulation donnée par (3.2.15). Pour déterminer

(3.2.16), il suffit de voir que $\sum_i \mu_i \langle \vec{O}^i, \vec{O}^i \rangle = \sum_{i,i'} \mu_i \mu_{i'} \left(\frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) \frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}}$, de factoriser par $\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}}$, et d'identifier $\left(\frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} - \langle \vec{O}^i, \vec{O}^{i'} \rangle \right)$ à $\bar{\mathbf{C}}_{ii'}$.

□

3.2.2.3 Expression relationnelle de l'inertie inter-classes

Propriété 3.2.3 Expressions relationnelles de l'inertie inter-classe d'une partition \mathcal{R}

Nous avons les différentes expressions équivalentes suivantes de $\mathbb{I}B(\mathcal{R})$ où X est la matrice relationnelle représentative de \mathcal{R} :

$$\mathbb{I}B(X) = \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \quad (3.2.19)$$

$$\mathbb{I}B(X) = \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \text{ où, } \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) = 1 - \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \quad (3.2.20)$$

Dans le cas classique où $\mu_i = \frac{1}{N} \quad \forall i = 1, \dots, N$, nous avons :

$$\mathbb{I}B(X) = \frac{1}{N} \sum_{i,i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{X_{i.}} \right) - \frac{1}{N^2} \sum_{i,i'} \mathbf{C}_{ii'} \quad (3.2.21)$$

$$\mathbb{I}B(X) = \frac{1}{N^2} \sum_{i,i'} \bar{\mathbf{C}}_{ii'} - \frac{1}{N} \sum_{i,i'} \bar{\mathbf{C}}_{ii'} \left(\frac{X_{ii'}}{X_{i.}} \right) \quad (3.2.22)$$

Démonstration

Selon le théorème de Huyghens, nous avons : $\mathbb{I}B(X) = \mathbb{I}T - \mathbb{I}W(X)$. En utilisant les équations (3.2.11) et (3.2.15), nous avons :

$$\begin{aligned} \mathbb{I}B(X) &= \left(\sum_i \mu_i \mathbf{C}_{ii} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \right) - \left(\sum_i \mu_i \mathbf{C}_{ii} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \right) \\ &= - \left(\sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \right) \\ &= \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \end{aligned}$$

De manière symétrique, en prenant les équations (3.2.12) et (3.2.16), nous obtenons :

$$\begin{aligned} \mathbb{I}B(X) &= \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii'} - \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii'} \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \\ &= \sum_{i,i'} \mu_i \mu_{i'} \bar{\mathbf{C}}_{ii'} \left(1 - \left(\frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \right) \right) \end{aligned}$$

□

3.2.2.4 Lien entre le critère de la différence inertielle et le critère de Condorcet

Le critère de partitionnement donné par la différence inertielle $\mathbb{I}B(\mathcal{R}) - \mathbb{I}W(\mathcal{R})$ que l'on cherche à maximiser, est souvent utilisé en analyse de données. Son expression relationnelle permet de rapprocher ce critère du critère de Condorcet. Nous aboutissons à la propriété suivante en combinant les équations (3.2.19) et (3.2.16).

Propriété 3.2.4 Expression relationnelle du critère de la différence inertielle

Le critère de la différence inertielle pour une partition donnée \mathcal{R} représentée par sa matrice relationnelle X , est exprimée en termes relationnels par l'équation ci-dessous :

$$\mathbb{I}\mathbb{B}(X) - \mathbb{I}\mathbb{W}(X) = \frac{1}{N} \sum_{i,i'} (\mathbf{C}_{ii'} - \bar{\mathbf{C}}_{ii'}) \left(\frac{X_{ii'}}{X_i} \right) - \frac{1}{N^2} \sum_{i,i'} \mathbf{C}_{ii'} \quad (3.2.23)$$

En explicitant $\bar{\mathbf{C}}_{ii'}$, nous avons la formulation équivalente ci-dessous :

$$\mathbb{I}\mathbb{B}(X) - \mathbb{I}\mathbb{W}(X) = \frac{2}{N} \sum_{i,i'} \left(\mathbf{C}_{ii'} - \frac{1}{2} \left(\frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) \right) \left(\frac{X_{ii'}}{X_i} \right) - \frac{1}{N^2} \sum_{i,i'} \mathbf{C}_{ii'} \quad (3.2.24)$$

La deuxième partie étant indépendante de X , nous constatons que la maximisation du critère de la différence inertielle par programmation linéaire rappelle celle du critère de Condorcet donné par les équations (3.1.23) pour lequel les similarités $\mathbf{S}_{ii'}$ sont les produits scalaires $\mathbf{C}_{ii'}$.

La différence essentielle entre ces deux critères est que celui de la différence inertielle pondère les contributions par le poids de la classe de chaque objet qui vaut dans le cas simple X_i . (dans l'exemple de l'objet O^i).

Nous prenons le parti du critère de Condorcet pour lequel cette pondération n'est pas. En effet, il est bien connu que, sans la fixation d'un nombre de classes de la partition \mathcal{R} , la maximisation du critère de la différence inertielle conduit irrémédiablement à la partition triviale à N singletons. Or, ceci est justement dû à la pondération du poids de la classe.

Nous étendons les résultats de [Mar91a] et pensons que le critère de Condorcet appliqué à des données quantitatives où les termes $\mathbf{C}_{ii'}$ sont les produits scalaires entre vecteurs d'objets peut être pris comme critère de partitionnement "non trivial" associé à une ACP.

Propriété 3.2.5 Critère de Condorcet comme critère de partitionnement non trivial pouvant être associé à une ACP

La partition obtenue à partir de la maximisation du critère de Condorcet¹⁶ ci-dessous est complémentaire à une analyse des projections issues d'une ACP :

$$\begin{aligned} \max_X \text{Condorcet}(\mathbf{C}, X) &\Leftrightarrow \max_X \sum_{i,i'} \left(\mathbf{C}_{ii'} - \frac{1}{2} \left(\frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) \right) X_{ii'} \\ &\Leftrightarrow \max_X \sum_{i,i'} \left(\langle \vec{O}^i, \vec{O}^{i'} \rangle - \frac{1}{2} \left(\frac{\langle \vec{O}^i, \vec{O}^i \rangle + \langle \vec{O}^{i'}, \vec{O}^{i'} \rangle}{2} \right) \right) X_{ii'} \end{aligned} \quad (3.2.25)$$

où X est une matrice relationnelle vérifiant les propriétés d'une relation d'équivalence.

Il existe une deuxième interprétation du critère de la différence inertielle dans le cadre de l'AR. Pour cela, développons (3.2.24) :

$$\begin{aligned} \mathbb{I}\mathbb{B}(X) - \mathbb{I}\mathbb{W}(X) &= 2 \sum_{i,i'} \mu_i \mu_{i'} \left(\mathbf{C}_{ii'} - \frac{1}{2} \frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) \frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \\ &= 2 \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} - \sum_{i,i'} \mu_i \mu_{i'} \left(\frac{\mathbf{C}_{ii} + \mathbf{C}_{i'i'}}{2} \right) \frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} \\ &\quad - \sum_{i,i'} \mu_i \mu_{i'} \mathbf{C}_{ii'} \end{aligned}$$

Le deuxième terme se simplifie puisque :

$$\begin{aligned} \sum_{i,i'} \mu_i \mu_{i'} \frac{\mathbf{C}_{ii}}{2} \frac{X_{ii'}}{\sum_{i''} \mu_{i''} X_{ii''}} &= \frac{\sum_i \mu_i \mathbf{C}_{ii} \sum_{i'} \mu_{i'} X_{ii'}}{2 \sum_{i''} \mu_{i''} X_{ii''}} \\ &= \frac{1}{2} \sum_i \mu_i \mathbf{C}_{ii} \end{aligned}$$

Nous avons alors, en remplaçant les μ_i par $1/N$:

$$\mathbb{I}\mathbb{B}(X) - \mathbb{I}\mathbb{W}(X) = \frac{2}{N} \sum_{i,i'} \mathbf{C}_{ii'} \frac{X_{ii'}}{X_i} - \frac{1}{N} \sum_i \mathbf{C}_{ii} \quad (3.2.26)$$

16. ou de Rand cf tab-3.9

Remarque 3.2.1 *Equivalence entre IW , IB et $\text{IB} - \text{IW}$ comme critères de partitionnement*

En considérant les équations (3.2.15), (3.2.19) et (3.2.26), nous retrouvons le résultat bien connu suivant :

$$\min_X \text{IW}(X) \Leftrightarrow \max_X \text{IB}(X) \Leftrightarrow \max_X (\text{IB}(X) - \text{IW}(X)) \quad (3.2.27)$$

La deuxième interprétation du critère de la différence inertielle dans le cadre de l'AR montre que la maximisation du critère suivant, pourrait aussi être un critère de partitionnement adapté à l'analyse des projections d'une ACP :

$$\max_X \sum_{i,i'} \mathbf{C}_{ii'} X_{ii'} \quad (3.2.28)$$

Ce dernier est plus laxiste que le critère de Condorcet étant donné que les contributions sont ici plus "facilement" positives. Ce critère est particulièrement adapté à des données centrées puisque le signe de $\mathbf{C}_{ii'} = \langle \vec{O}^i, \vec{O}^{i'} \rangle$ est donné par l'angle formé par les deux vecteurs. Ainsi, centrer les données permet de comparer les profils d'objets relativement à un profil moyen donné par le centre de gravité.

Il existe donc des produits scalaires positifs propres à des objets similaires qui auront tendance à se regrouper dans une même classe ; et des produits scalaires négatifs conduisant à des objets antisimilaires qui auront tendance à se repousser.

Nous constatons aussi que de nombreux critères d'association maximale étendus à des données quantitatives peuvent se réduire, dans le cas de données centrées, au critère de la différence inertielle non pondérée par le poids de la classe. Nous avons la propriété suivante :

Propriété 3.2.6 *Liens entre le critère de la différence inertielle et les autres critères d'association maximale*

Lorsque les données sont centrées les critères de Belson, de Jordan, de Light et Margolin et d'écart carré à l'indépendance, donnés dans tab-3.9 sont similaires au critère de partitionnement de la différence inertielle donnée par (3.2.26).

3.2.3 Métriques relationnelles et régularisation de l'espace de description

Nous nous intéressons dans cette sous-section à la régularisation des produits scalaires qui définissent l'ensemble des mesures de similarité que nous avons étudiées. La similarité régularisée fut initialement proposée par H. Benhadda et J.F. Marcotorchino dans [Ben98] et [BM98]. Nous avons rappelé, dans le cas des variables qualitatives, les différents types de ces similarités en introduction page 9.

Si nous nous plaçons dans un cadre géométrique, la similarité régularisée revient à utiliser une métrique particulière dans le calcul du produit scalaire. Notons par \mathbf{M}^α cette métrique. Il s'agit d'une matrice carrée de dimension P qui est diagonale :

$$\mathbf{M}_{jj'}^\alpha = \begin{cases} \alpha_j & \text{si } j = j' \\ 0 & \text{sinon} \end{cases} \quad (3.2.29)$$

Nous avons alors l'interprétation géométrique de la similarité régularisée suivante :

$$\mathbf{S}_{ii'}^\alpha = \langle \vec{O}^i, \vec{O}^{i'} \rangle^\alpha = {}^t \vec{O}^i \cdot \mathbf{M}^\alpha \cdot \vec{O}^{i'} \quad (3.2.30)$$

L'idée poursuivie dans ce paragraphe est d'étendre cette métrique à des matrices non diagonales. Là encore, cette approche s'inscrit dans un cadre géométrique et permet, de ce fait, de traiter naturellement des vecteurs de profils d'objets de nature non binaire.

Nous proposons de définir des métriques comme étant des matrices de comparaisons par paires des indices de similarité d'ordre t mesurés pour chaque couple de descripteurs.

Définition 3.2.5 Métriques relationnelles

Soit un ensemble \mathbb{O} constitué de N objets, décrit par un ensemble de P descripteurs binaires ou non-binaires \mathbb{D} . Nous appellerons métrique relationnelle d'ordre t une matrice carrée d'ordre P , notée \mathbf{M}^t dont le terme général $\mathbf{M}_{jj'}^t, \forall j, j' = 1, \dots, P$; est donné par la relation ci-dessous :

$$\mathbf{M}_{jj'}^t = \mathbf{S}_{jj'}^t = \frac{\langle \vec{D}^j, \vec{D}^{j'} \rangle}{\mathcal{M}^t(\langle \vec{D}^j, \vec{D}^j \rangle, \langle \vec{D}^{j'}, \vec{D}^{j'} \rangle)} \quad (3.2.31)$$

Nous noterons par $\langle \vec{O}^i, \vec{O}^{i'} \rangle_{\mathbf{M}^t}$ le produit scalaire entre les vecteurs représentatifs des objets O^i et $O^{i'}$ induit par la métrique relationnelle d'ordre t . Nous avons alors :

$$\langle \vec{O}^i, \vec{O}^{i'} \rangle_{\mathbf{M}^t} = {}^t\vec{O}^i \cdot \mathbf{M}^t \cdot \vec{O}^{i'} = \sum_{j=1}^P \sum_{j'=1}^P \mathbf{S}_{jj'}^t O_j^i O_{j'}^{i'} \quad (3.2.32)$$

où $\vec{O}^i = (O_1^i, O_2^i, \dots, O_P^i)$ et O_j^i est la valeur prise par l'objet O^i pour le descripteur D^j .

Considérons l'équation (3.2.32) et pour une meilleure compréhension considérons, sans perte de généralité, le cas de variables qualitatives.

Si l'objet O^i possède la modalité D^j et l'objet $O^{i'}$ le descripteur $D^{j'}$, alors ce "matching" comptera d'autant plus dans la mesure de similarité si les descripteurs D^j et $D^{j'}$ sont fortement similaires.

Prenons un exemple afin d'illustrer ces propos. Soit $D^j =$ "avoir les cheveux blonds", $D^{j'} =$ "avoir les yeux bleus" et $D^{j''} =$ "pratiquer le foot-ball". D^j et $D^{j'}$ sont des descripteurs *a priori* plus fortement "similaires" que D^j et $D^{j''}$. De ce fait, dans la mesure de similarité entre deux objets O^i et $O^{i'}$, le "matching" entre D^j et $D^{j'}$ doit peser davantage que le "matching" entre D^j et $D^{j''}$.

Le terme de régularisation d'espace de description est employé ici dans le sens où les vecteurs de description ne sont plus considérés comme étant indépendants comme pour les métriques diagonales. Les métriques relationnelles permettent de tenir compte des similarités entre descripteurs dans le produit scalaire entre vecteurs d'objets. De cette façon, les descripteurs positivement similaires contribuent positivement au produit scalaire et inversement pour des descripteurs antisimilaires ou opposés. Autrement dit, nous prenons en compte les dépendances entre descripteurs dans les mesures de similarité entre objets.

De manière symétrique, P.C. Mahalanobis utilise comme métrique, l'inverse de la matrice des variances-covariances ou des corrélations entre descripteurs [Sap06]. Cette métrique régularise l'espace de description d'une manière duale à ce qui vient d'être énoncé. En effet, elle cherche à corriger les dépendances de sorte à ce que des descripteurs corrélés deviennent indépendants.

Dans un cadre classificatoire, nous estimons que le problème de colinéarité entre descripteurs qui influe de manière redondante dans le calcul de similarité ou de distance entre objets et que cherche à corriger la métrique de Mahalanobis, doit être repositionné dans une étape en amont du processus de classification. Cette phase consiste en une **sélection des variables les plus pertinentes** qui devront participer à la phase de classification. Dès lors que des variables sont sélectionnées, nous pensons que les liens, au sens d'une métrique relationnelle, qu'elles entretiennent entre elles, devraient être pris en compte dans la mesure de similarité entre objets.

Notons cependant, que ce produit scalaire n'est pas calculable rapidement étant donné qu'il nécessite $\mathcal{O}(P^3)$ opérations¹⁷. Il n'est donc pas *a priori*, adapté au traitement de grandes bases de données.

3.3 Les notions d'indépendance et d'indétermination dans le cas de la mesure d'association entre variables qualitatives

Nous revenons vers l'étude exclusive des variables qualitatives. Nous étudions différentes notions d'indépendance sur lesquelles sont basés les critères d'association.

Nous rappelons qu'il existe plusieurs types d'indépendance dans l'étude de l'association entre deux variables qualitatives : statistiques, logiques, géométriques. Nous montrons en fait qu'il existe des correspondances entre ces indépendances selon le codage utilisé pour représenter les variables qualitatives : l'indépendance statistique définie en codage contingentiel correspond, dans une certaine mesure, à l'indépendance géométrique en codage relationnel et *vice versa*. Nous enrichissons les travaux d'A. Najah Idrissi à ce sujet [NI00] et nous mettons en évidence des formulations en terme de covariance de différents autres critères. L'intérêt est de montrer que la plupart des critères d'association sont dans leurs codages relationnels fondés sur des concepts géométriques.

Dans un deuxième temps, nous nous arrêtons plus particulièrement sur la notion d'indétermination qui est une approche logique de l'indépendance, initialement proposée par J.F. Marcotorchino [Mar85]. Nous clarifions certaines ambiguïtés et étendons la notion d'indétermination à l'indétermination pondérée. Un des intérêts de ce type d'indépendance est qu'il est défini pour d'autres types de relation que celui des partitions : nous verrons au chapitre 6 que pour les relations d'ordre, les critères d'association sont majoritairement fondés sur ce concept.

Nous emploierons des notations correspondantes au codage contingentiel et relationnel. Le codage contingentiel a été présenté en introduction page 9 et les formules de passage entre ces deux codages sont données par le tableau tab-1.4 page 12.

Les résultats de cette section, ont donné lieu à un article [APM07], qui a été accepté pour la conférence "Applied stochastic models and data analysis" 2007.

3.3.1 Les différents types d'indépendance définis dans le cadre de l'AR

3.3.1.1 L'indépendance statistique

3.3.1.1.1 L'écart à l'indépendance statistique

Considérons deux variables V^k et V^l avec p_k et p_l modalités que nous noterons respectivement $\{D^u; u = 1, \dots, p_k\}$ et $\{D^v; v = 1, \dots, p_l\}$. Soit \mathbf{n}^{kl} le tableau de contingence croisant les modalités de ces deux variables. Les variables sont indépendantes si pour l'ensemble des paires (D^u, D^v) la probabilité jointe d'avoir à la fois la modalité D^u de V^k et D^v de V^l est égale au produit des probabilités marginales.

Définition 3.3.1 *Indépendance statistique entre deux variables qualitatives*

17. complexité d'un produit matriciel

$$V^k \perp_S V^l \Leftrightarrow \frac{\mathbf{n}_{uv}^{kl}}{N} = \frac{\mathbf{n}_u^{kl}}{N} \frac{\mathbf{n}_v^{kl}}{N} \quad \forall (D^u, D^v) : D^u \in V^k, D^v \in V^l \quad (3.3.1)$$

Dés lors, plusieurs critères intègrent cette notion et nous retrouvons très souvent des termes fonction de l'écart à l'indépendance " $f\left(\mathbf{n}_{uv}^{kl} - \frac{\mathbf{n}_u^{kl}\mathbf{n}_v^{kl}}{N}\right)$ " dans leur définition. C'est le cas des critères de Belson, du Chi-deux de Tchuprow, et de Jordan.

Si les autres critères ne sont pas des fonctions explicites de l'écart à l'indépendance, ils sont néanmoins tous nuls lorsque les variables sont effectivement indépendantes au sens statistique à l'exception du critère de Rand ou de Condorcet et de Janson et Vegelius.

3.3.1.1.2 Le rapport des chances ou "odds-ratio"

Pour des variables binaires, il existe une autre mesure d'association bien connue qui est celle du rapport des chances ou encore "odds-ratio" [DLS05]. Considérons deux variables qualitatives V^k et V^l . Une fois codées dans le formalisme relationnel, elles se transforment en variables binaires sous la forme d'une matrice carrée structurée [Mar81]. Nous pouvons alors considérer le tableau de contingence 2×2 donné par tab-3.10.

	C^l	\bar{C}^l	Marges
C^k	$11_{kl} = \sum_{i,i'} C_{ii'}^k C_{ii'}^l$	$10_{kl} = \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l$	$\sum_{i,i'} C_{ii'}^k$
\bar{C}^k	$01_{kl} = \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l$	$00_{kl} = \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l$	$\sum_{i,i'} \bar{C}_{ii'}^k$
Marges	$\sum_{i,i'} C_{ii'}^l$	$\sum_{i,i'} \bar{C}_{ii'}^l$	N^2

TABLE 3.10 – Définition des accords et des désaccords entre deux matrices relationnelles de partition

Pour ce type de tableau, nous définissons le rapport des chances comme étant le ratio suivant :

$$\text{OR}(V^k, V^l) = \frac{11_{kl}00_{kl}}{10_{kl}01_{kl}} \quad (3.3.2)$$

En cas d'indépendance statistique ce ratio vaut 1. Une deuxième approche de la mesure d'indépendance statistique entre les deux variables considérées consiste alors à prendre le critère suivant :

$$V^k \perp_S V^l \Leftrightarrow 11_{kl}00_{kl} - 10_{kl}01_{kl} = 0 \quad (3.3.3)$$

En développant l'expression ci-dessus, nous aboutissons au résultat suivant :

Propriété 3.3.1 Lien entre le rapport des chances et le critère de Lerman

Le critère relationnel mesurant l'écart à l'indépendance au sens de l'équation (3.3.3) entre deux matrices

relationnelles C^k et C^l est à une constante multiplicative près, celui du critère d'ecart carré à l'indépendance, $E(C^k, C^l)$, qui est lui même lié au numérateur du critère modifié de Lerman : $Lerman'(C^k, C^l)$.

$$\begin{aligned} 11_{kl}00_{kl} - 10_{kl}01_{kl} = 0 &\Leftrightarrow E(C^k, C^l) = 0 \\ &\Leftrightarrow \sum_{i,i'} \left(C_{ii'}^k - \sum \frac{C_{ii'}^k}{N^2} \right) \left(C_{ii'}^l - \sum \frac{C_{ii'}^l}{N^2} \right) = 0 \end{aligned} \quad (3.3.4)$$

Autrement dit nous avons les propriétés suivantes :

$$OR(V^k, V^l) = 1 \Leftrightarrow E(C^k, C^l) = Lerman'(C^k, C^l) = 0 \quad (3.3.5)$$

Démonstration

$$\begin{aligned} 11_{kl}00_{kl} - 10_{kl}01_{kl} &= \sum_{i,i'} C_{ii'}^k C_{ii'}^l - \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l - \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l \\ &= \sum_{i,i'} C_{ii'}^k C_{ii'}^l \left[\sum (1 - C_{ii'}^k - C_{ii'}^l + C_{ii'}^k C_{ii'}^l) \right] \\ &\quad - \left[\left(\sum C_{ii'}^k - \sum C_{ii'}^l \right) \left(\sum C_{ii'}^l - \sum C_{ii'}^k \right) \right] \\ &= \left[N^2 \sum C_{ii'}^k C_{ii'}^l - \sum C_{ii'}^k \sum C_{ii'}^l - \sum C_{ii'}^l \sum C_{ii'}^k + \left(\sum C_{ii'}^k C_{ii'}^l \right)^2 \right] \\ &\quad - \left[\sum C_{ii'}^k \sum C_{ii'}^l - \sum C_{ii'}^k \sum C_{ii'}^l - \sum C_{ii'}^l \sum C_{ii'}^k + \left(\sum C_{ii'}^k C_{ii'}^l \right)^2 \right] \\ &= N^2 \sum_{i,i'} C_{ii'}^k C_{ii'}^l - \sum_{i,i'} C_{ii'}^k \sum_{i,i'} C_{ii'}^l \\ &= N^2 \sum_{i,i'} \left(C_{ii'}^k - \sum \frac{C_{ii'}^k}{N^2} \right) \left(C_{ii'}^l - \sum \frac{C_{ii'}^l}{N^2} \right) \end{aligned}$$

□

Par ailleurs, nous rappelons que nous pouvons exprimer le critère modifié Lerman' tel un coefficient tétrachorique mesuré sur des matrices relationnelles¹⁸. En effet, ce coefficient est défini par l'expression suivante :

$$Tetra(C^k, C^l) = \frac{(11_{kl}00_{kl} - 10_{kl}01_{kl})}{\sqrt{(11_{kl} + 10_{kl})(01_{kl} + 00_{kl})(11_{kl} + 01_{kl})(10_{kl} + 00_{kl})}} \quad (3.3.6)$$

Nous avons alors la propriété suivante :

Propriété 3.3.2 Identité entre le critère de Lerman modifié et le coefficient tétrachorique

$$Lerman'(C^k, C^l) = Tetra(C^k, C^l) \quad (3.3.7)$$

Nous faisons remarquer que cette égalité n'est pas valable dans le cas de la formulation initiale de I.C. Lerman en raison de la non prise en compte des termes unitaires de la diagonale des matrices relationnelles. Ceci justifie d'autant plus la formulation modifiée donnée par A. Najah Idrissi et J.F. Marcotorchino [NI00].

3.3.1.2 L'indétermination ou l'approche logique de l'indépendance entre relations d'équivalence

L'expression des critères d'association dans le codage relationnel a permis de mieux appréhender un autre concept de la mesure d'une relation entre deux variables qualitatives qui est sous-jacent à certains critères d'association : l'indétermination. Selon la définition de J.F. Marcotorchino [Mar85], l'indétermination est vérifiée par le critère de Rand, le critère de Janson et Vegelius et le critère de Jordan.

A la base, deux variables qualitatives V^k et V^l , représentées par leurs matrices relationnelles C^k et C^l , sont indéterminées si le nombre d'accords entre ces variables est égale au nombre de

18. initialement, dans le cas discret, le coefficient tétrachorique est mesuré sur des vecteurs binaires cf tab-3.2

désaccords.

Le nombre d'accords entre deux variables est donné par $\sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l$. On distingue accords positifs, $\sum_{i,i'} C_{ii'}^k C_{ii'}^l$, et accords négatifs $\sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l$.

Le nombre de désaccords est quant à lui donné par le désaccord relatif de C^k avec C^l , $\sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l$, additionné du désaccord relatif de C^l avec C^k , $\sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l$.

Nous pouvons définir les accords et les désaccords entre deux matrices relationnelles selon le même formalisme présenté pour les vecteurs binaires en page 71 et que nous avons rappelé dans le paragraphe précédent.

Deux variables sont alors indéterminées si nous avons :

$$V^k \perp_L V^l \Leftrightarrow 11_{kl} + 00_{kl} - 10_{kl} - 01_{kl} = 0 \quad (3.3.8)$$

$$\Leftrightarrow \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l = 0 \quad (3.3.9)$$

En raison du caractère logique de l'indétermination, nous parlerons aussi d'"indépendance logique". Le critère $2\text{Condorcet}(C^k, C^l) - 1$ ou $2\text{Rand}(C^k, C^l) - 1$ est nul lorsqu'il y a indétermination au sens de l'équation ci-dessus.

Dans [Mar85], est donné l'équivalent de l'indétermination définie précédemment dans le codage contingential. Cependant, cette équivalence implique nécessairement que les deux variables V^k et V^l étudiées aient exactement deux modalités chacune. Nous avons alors la relation suivante :

$$\begin{aligned} \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l &= \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l \\ \Leftrightarrow \mathbf{n}_{uv}^{kl} &= \frac{\mathbf{n}_{u\cdot}^{kl}}{2} + \frac{\mathbf{n}_{\cdot v}^{kl}}{2} - \frac{N}{4} \quad \forall (D^u, D^v) \in \{V_1^k, V_2^k\} \times \{V_1^l, V_2^l\} \end{aligned} \quad (3.3.10)$$

En remarquant cette relation, J.F. Marcotorchino étend et définit finalement le concept d'indétermination au travers de la relation donnée ci-dessous :

Définition 3.3.2 *Indétermination entre deux variables qualitatives selon Marcotorchino*

$$V^k \perp V^l \Leftrightarrow \mathbf{n}_{uv}^{kl} = \frac{\mathbf{n}_{u\cdot}^{kl}}{p_l} + \frac{\mathbf{n}_{\cdot v}^{kl}}{p_k} - \frac{N}{p_l p_k} \quad \forall (D^u, D^v) \in V^k \times V^l \quad (3.3.11)$$

Cette approche est issue d'un raisonnement effectué sur des sous-tableaux 2×2 du tableau de contingence \mathbf{n}_{uv}^{kl} et non pas sur les matrices relationnelles. En particulier, nous avons la relation suivante :

$$\sum_{u,v} \left(\mathbf{n}_{uv}^{kl} - \frac{\mathbf{n}_{u\cdot}^{kl}}{p_l} - \frac{\mathbf{n}_{\cdot v}^{kl}}{p_k} + \frac{N}{p_k p_l} \right)^2 = \frac{1}{4p_k p_l} \sum_{u,v} \sum_{u',v'} ((\mathbf{n}_{uv}^{kl} + \mathbf{n}_{u'v'}^{kl}) - (\mathbf{n}_{uv'}^{kl} + \mathbf{n}_{u'v}^{kl}))^2 \quad (3.3.12)$$

Dans le cadre de la définition donnée par (3.3.11), les critères de Janson et Vegelius, et de Jordan sont alors effectivement nuls en cas d'indétermination.

Toutefois, si nous définissons l'indétermination par (3.3.11), nous n'avons plus l'équivalence avec l'équation (3.3.8). En particulier, nous pouvons avoir $\mathbf{n}_{uv}^{kl} = \frac{\mathbf{n}_{u\cdot}^{kl}}{p_l} + \frac{\mathbf{n}_{\cdot v}^{kl}}{p_k} - \frac{N}{p_l p_k} \quad \forall (D^u, D^v) \in V^k \times V^l$ et $11_{kl} + 00_{kl} \neq 10_{kl} + 01_{kl}$.

Pour cette raison, nous proposons dans un premier temps, une autre interprétation de l'indétermination définie par (3.3.11) qui est géométrique et qui fait appel aux transformations

de **W.S. Torgerson [Tor52]**. Nous présentons cette approche dans le paragraphe suivant dans lequel nous exposons l'indépendance géométrique. Plus particulièrement, la relation (3.3.11) sera étudiée dans le cadre du critère de Janson et Vegelius. Nous montrerons par ailleurs les liens que ce dernier entretient avec le critère de Belson.

Nous conservons donc finalement la notion d'indétermination donnée par l'équation (3.3.8). Nous dirons donc que V^k et V^l sont indéterminées si :

$$V^k \perp_L V^l \Leftrightarrow 11_{kl} + 00_{kl} - 10_{kl} - 01_{kl} = 0 \quad (3.3.13)$$

$$\Leftrightarrow \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l = 0 \quad (3.3.14)$$

Par conséquent, seul le critère $2\text{Rand}(C^k, C^l) - 1$ s'annule en cas d'indétermination entre les variables V^k et V^l . Nous constatons par ailleurs, que l'indétermination donnée par (3.3.13) est fondée sur un **modèle additif** contrairement à l'indépendance statistique qui au vu de la relation (3.3.3) est basée sur un **modèle multiplicatif**. Par conséquent, il existe une véritable distinction dans l'interprétation de ces deux approches de la mesure de la relation entre deux variables qualitatives.

Nous revenons alors dans un deuxième temps et de manière plus approfondie vers le concept d'indétermination et nous l'étendons à la notion d'indétermination pondérée dans la sous-section 3.3.2 afin de définir davantage de critères basés sur ce type d'approche. Ce faisant, le concept d'indétermination donné par (3.3.11), s'inscrit alors dans le cadre général de l'indétermination pondérée pour laquelle elle représente un cas particulier. **Nous verrons par ailleurs, lors de l'étude de l'association entre relations d'ordre dans le chapitre 6 que de nombreux critères d'association entre ces types de relation sont basés sur ce même concept d'indétermination.**

3.3.1.3 L'indépendance géométrique et indices de corrélation entre relations d'équivalence

A la suite des travaux de A. Najah Idrissi [NI00], il apparaît que, dans la représentation relationnelle des variables qualitatives, l'ensemble des types d'indépendance précédemment présentés pouvait se ramener à des interprétations géométriques : nombreux de ces critères s'expriment ou se rapprochent d'une expression sous la forme d'un coefficient de corrélation ou d'une covariance. Nous avons rappelé ces résultats en introduction page 11.

Nous enrichissons les résultats de A. Najah Idrissi en rappelant les critères de Belson, de Jordan et d'écart carré à l'indépendance écrits sous leur forme de covariance. L'intérêt de ce paragraphe réside finalement dans le tableau tab-3.11 où nous présentons de manière synthétique l'ensemble de ces résultats.

Propriété 3.3.3 Expression relationnelle en terme de covariance du critère de Belson

$$\text{Belson}(C^k, C^l) = \sum_{i=1}^N \sum_{i'=1}^N \left(C_{ii'}^k - \frac{C_{i.}^k + C_{.i'}^k}{N} + \frac{C_{..}^k}{N^2} \right) \left(C_{ii'}^l - \frac{C_{i.}^l + C_{.i'}^l}{N} + \frac{C_{..}^l}{N^2} \right) \quad (3.3.15)$$

Démonstration

$$\begin{aligned} \text{Nous montrons que } & \sum \left(C_{ii'}^k - \frac{C_{i.}^k + C_{.i'}^k}{N} + \frac{C_{..}^k}{N^2} \right) \left(-\frac{C_{i.}^l + C_{.i'}^l}{N} + \frac{C_{..}^l}{N^2} \right) = 0 \\ & \sum \left(C_{ii'}^k - \frac{C_{i.}^k + C_{.i'}^k}{N} + \frac{C_{..}^k}{N^2} \right) \left(-\frac{C_{i.}^l + C_{.i'}^l}{N} + \frac{C_{..}^l}{N^2} \right) = -\sum \left(C_{ii'}^k \left(\frac{C_{i.}^l + C_{.i'}^l}{N} \right) \right) + \sum \left(\frac{C_{i.}^k + C_{.i'}^k}{N} \right) \left(\frac{C_{i.}^l + C_{.i'}^l}{N} \right) \\ & \quad - \frac{C_{..}^k}{N^2} \sum \left(\frac{C_{i.}^l + C_{.i'}^l}{N} \right) + \frac{C_{..}^l}{N^2} \sum \left(C_{ii'}^k - \frac{C_{i.}^k + C_{.i'}^k}{N} + \frac{C_{..}^k}{N^2} \right) \\ & = -2 \sum_i \frac{C_{i.}^k C_{i.}^l}{N} + 2 \sum_i \frac{C_{i.}^k C_{..}^l}{N^2} + 2 \sum_i \frac{C_{i.}^k C_{.i}^l}{N} \\ & \quad - 2 \sum_i \frac{C_{i.}^k C_{..}^l}{N} + \frac{C_{..}^l}{N^2} (C_{..}^k - 2C_{..}^k + C_{..}^k) \\ & = 0 \end{aligned}$$

□

Propriété 3.3.4 *Expression relationnelle en terme de covariance du critère de Jordan*

$$\text{Jordan}(C^k, C^l) = \frac{1}{N} \sum_{i,i'} \left(C_{ii'}^k - \frac{C_{i.}^k}{N} \right) \left(C_{ii'}^l - \frac{C_{i.}^l}{N} \right) \quad (3.3.16)$$

Démonstration

La démonstration est similaire à celle du critère de Belson

□

Propriété 3.3.5 *Expression relationnelle en terme de covariance du critère d'écart carré à l'indépendance*

$$\text{E}(C^k, C^l) = \sum_{i,i'} \left(C_{ii'}^k - \frac{C_{i.}^k}{N} \right) \left(C_{ii'}^l - \frac{C_{i.}^l}{N} \right) \quad (3.3.17)$$

Démonstration

La démonstration est similaire à celle du critère de Belson

□

Nous pouvons synthétiser l'ensemble des expressions relationnelles des critères d'association qui s'exprime sous la forme d'un coefficient de corrélation. Ce qui distingue finalement un critère d'un autre sont d'une part, les transformations faites aux matrices relationnelles et d'autre part, les moyennes utilisées dans le calcul du coefficient de corrélation. De manière générale, nous avons la forme générique ci-dessous qui dépend de la transformation f et des tendances centrales μ^k et μ^l :

$$\Delta(C^k, C^l, f, \mu^k, \mu^l) = \frac{\sum_{i,i'} (f(C_{ii'}^k) - \mu^k) (f(C_{ii'}^l) - \mu^l)}{\sqrt{\sum_{i,i'} (f(C_{ii'}^k) - \mu^k)^2 \sum_{i,i'} (f(C_{ii'}^l) - \mu^l)^2}} \quad (3.3.18)$$

Nous donnons dans le tableau tab-3.11, la correspondance entre les paramètres (f, μ^k, μ^l) et les critères d'association.

Les matrices relationnelles sont des matrices binaires qui peuvent être considérées comme étant des vecteurs dans un espace tensoriel de base $\{e_i \otimes e_{i'}; i, i' = 1, \dots, N\}$ ¹⁹. Exprimés sous la forme d'un coefficient de corrélation, nous voyons que ces derniers s'annulent lorsque le cosinus de l'angle formé par les deux vecteurs C^k et C^l vaut 0.

Nous dirons donc que deux variables qualitatives sont géométriquement indépendantes, dans le cadre de l'AR, si nous avons la relation suivante :

$$V^k \perp_G V^l \Leftrightarrow \Delta(C^k, C^l, f, \mu^k, \mu^l) = 0 \quad (3.3.19)$$

$$\Leftrightarrow \sum_{i,i'} (f(C_{ii'}^k) - \mu^k) (f(C_{ii'}^l) - \mu^l) = 0 \quad (3.3.20)$$

Remarque 3.3.1 *A propos du critère de Light et Margolin*

Parmi l'ensemble des critères proposés, celui de Light et Margolin issu du critère τ_b de Goodman et Kruskal ne correspond, à aucun des types d'indépendances précédemment cités. Il est asymétrique contrairement aux autres critères. Il est néanmoins nul dans le cas de l'indépendance statistique. Nous n'irons donc pas plus loin quant aux interprétations de ce critère.

19. où les e_i sont les vecteurs de la base canonique et \otimes le produit de Kronecker

$f(C_{ii'})$	μ^k	μ^l	Critère apparenté	Relation avec le critère apparenté
$f(C_{ii'}) = C_{ii'} - \frac{C_{i.} + C_{.i'}}{N} + \frac{C_{..}}{N^2}$ (Transformation de Torgerson)	0	0	Critère de Belson	Le critère de Belson correspond au numérateur du coefficient
$f(C_{ii'}) = C_{ii'}$	1/2	1/2	Critère de Rand	$2\text{Rand}(C^k, C^l) - 1$
$f(C_{ii'}) = C_{ii'}$	$1/p_k$	$1/p_l$	Critère de Janson et Vegelius	$JV(C^k, C^l)$
$f(C_{ii'}) = C_{ii'}$	$C_{..}^k/N^2$	$C_{..}^l/N^2$	Critère de Lerman, Critère d'écart carré à l'indépendance, et coefficient tétrachorique	$\text{Tetra}(C^k, C^l)$
$f(C_{ii'}) = C_{ii'}/C_{i.}$	1/N	1/N	Critère du Chi-deux Tchuprow	$\chi_T^2(C^k, C^l)$
$f(C_{ii'}) = C_{ii'}$	$C_{i.}^k/N$	$C_{i.}^l/N$	Critère de Jordan	Le critère de Jordan correspond au numérateur du coefficient à un facteur près

TABLE 3.11 – Tableau de correspondance des critères d'association en termes relationnels sous forme de coefficient de corrélation

Nous mettons en avant, le critère de Belson dont l'interprétation est résolument géométrique. En effet, dans le codage relationnel nous constatons que celui-ci correspond à un coefficient de corrélation sur des matrices relationnelles dont les données ont subi la transformation de Torgerson [Sap06]. Cette transformation est bien connue du domaine du "Multidimensional Scaling" en analyse de données. Partant d'un tableau de comparaisons par paires des produits scalaires entre objets, telle que la matrice \mathbf{C} , ou des distances obtenues à partir d'une métrique euclidienne, la transformation de Torgerson opérant sur ce tableau permet d'obtenir les produits scalaires entre objets centrés :

$$\mathbf{C}_{ii'} - \frac{\mathbf{C}_{i.}}{N} - \frac{\mathbf{C}_{.i'}}{N} + \frac{\mathbf{C}_{..}}{N^2} = \langle \vec{O}^i - \vec{G}, \vec{O}^{i'} - \vec{G} \rangle \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} \quad (3.3.21)$$

où $\vec{G} = \frac{1}{N} \sum_i \vec{O}^i$.

La diagonalisation du tableau ayant subi la transformation de Torgerson permet par la suite de représenter le nuage de points dans un espace à dimension réduite optimal. Cette projection conserve au mieux l'inertie initiale du nuage de points. Dans le cas de métriques euclidiennes, cette méthode est donc semblable à une ACP ou une AFCM selon la nature des variables.

Dans le cas du critère de Belson, les matrices considérées sont C^k et C^l . Celles-ci sont des matrices relationnelles binaires qui peuvent être vues comme étant des matrices de produits scalaires calculés en considérant les restrictions respectives du tableau disjonctif complet \mathbf{K} aux variables V^k et V^l . Autrement dit, $C_{ii'}^k$ est la valeur binaire du produit scalaire entre les vecteurs issus de la matrice disjonctive²⁰ K^k relative à V^k , représentatifs des objets O^i et $O^{i'}$.

Nous nous consacrons maintenant aux liens qu'entretiennent le numérateur du critère de Janson et Vegelius et le critère de Belson. Nous remarquons en effet une certaine dualité entre ces deux critères relative à la dualité contingence / relationnelle.

Propriété 3.3.6 Dualité entre critère de Belson et le (numérateur du) critère de Janson et Vegelius

Il existe une dualité de formalisme entre le critère de Belson²¹ et le numérateur du critère de Janson et Vegelius²² qui est relative à l'expression de ces derniers en codage contingentiel et relationnel.

Le critère de Belson en contingence repose sur l'écart à l'indépendance statistique ce qui s'exprime en relationnel par une indépendance géométrique associée à la transformation de Torgerson.

Le numérateur du critère de Janson et Vegelius est dual au critère de Belson dans le sens où en contingence, il repose sur une indépendance géométrique relative à la transformation de Torgerson ce qui s'exprime en relationnel par un certain type d'écart à l'indépendance statistique.

Nous mettons en évidence le lien entre ces deux critères dans le tableau tab-3.12 et dans ce qui suit.

	Ecart à l'indépendance statistique	Indépendance géométrique et Transformation de Torgerson
Belson	$\sum_{u,v} \left(\mathbf{n}_{uv}^{kl} - \frac{\mathbf{n}_u^{kl} \mathbf{n}_{\cdot v}^{kl}}{N} \right)^2$	$\sum_{i,i'} \left(C_{ii'}^k - \left[\frac{C_{i\cdot}^k}{N} + \frac{C_{\cdot i'}^k}{N} - \frac{C_{\cdot\cdot}^k}{N^2} \right] \right) \left(C_{ii'}^l - \left[\frac{C_{i\cdot}^l}{N} + \frac{C_{\cdot i'}^l}{N} - \frac{C_{\cdot\cdot}^l}{N^2} \right] \right)$
Janson et Vegelius (numérateur)	$\sum_{i,i'} \left(C_{ii'}^k - \frac{1}{p_k} \right) \left(C_{ii'}^l - \frac{1}{p_l} \right)$	$\sum_{u,v} \left(\mathbf{n}_{uv}^{kl} - \left[\frac{\mathbf{n}_u^{kl}}{p_l} + \frac{\mathbf{n}_{\cdot v}^{kl}}{p_k} - \frac{\mathbf{n}_{\cdot\cdot}^{kl}}{p_k p_l} \right] \right)^2$

TABLE 3.12 – Dualité dans le formalisme contingence / relationnel entre le critère de Belson et le critère de Janson et Vegelius correspondant à une dualité indépendance statistique / indépendance géométrique lié à la décomposition de Torgerson

Les termes \mathbf{n}_{uv}^{kl} peuvent être exprimés en fonction de produits scalaires entre vecteurs binaires représentatifs des modalités appartenant à deux variables qualitatives V^k et V^l distinctes :

$$\mathbf{n}_{uv}^{kl} = \langle \vec{D}^u, \vec{D}^v \rangle \quad \forall (D^u, D^v) \in V^k \times V^l \quad (3.3.22)$$

D'ailleurs en analyse de données, le tableau de comparaisons par paires des produits scalaires entre l'ensemble des descripteurs binaires constituant un tableau disjonctif complet est

20. cf page 8 ou matrice d'affectation à la page 34

21. cf (1.20) page 10 pour les expressions contingentielles du critère de Belson

22. cf (1.22) page 10 pour les expressions contingentielles du critère de Janson et Vegelius

appelé tableau de Burt et est noté \mathbf{B} . Celui-ci est le dual du tableau de Condorcet collectif \mathbf{C} au sens où, si \mathbf{K} est le tableau disjonctif complet, nous avons :

$$\mathbf{C} = \mathbf{K} \cdot {}^t\mathbf{K} \quad \text{et} \quad \mathbf{B} = {}^t\mathbf{K} \cdot \mathbf{K} \quad (3.3.23)$$

Le tableau de contingence \mathbf{n}^{kl} est une sous-matrice du tableau de Burt. De ce fait, nous pouvons lui associer des interprétations géométriques. Nous avons ainsi la relation suivante, qui donne finalement une interprétation géométrique à la relation (3.3.11) :

$$\mathbf{n}_{uv}^{kl} - \frac{\mathbf{n}_{u.}^{kl}}{p_l} - \frac{\mathbf{n}_{.v}^{kl}}{p_k} + \frac{\mathbf{n}_{..}^{kl}}{p_k p_l} = \langle \vec{D}^u - \vec{G}^k, \vec{D}^v - \vec{G}^l \rangle \quad \forall (D^u, D^v) \in V^k \times V^l \quad (3.3.24)$$

où $\vec{G}^k = \frac{1}{p_k} \sum_{D^u \in V^k} \vec{D}^u$ et $\vec{G}^l = \frac{1}{p_l} \sum_{D^v \in V^l} \vec{D}^v$.

A présent, notons symboliquement par $P(\mathbf{n}_{uv}^{kl})$ la probabilité qu'un objet ait à la fois la modalité D^u de V^k et D^v de V^l . Sous l'hypothèse d'indépendance, nous avons :

$$\begin{aligned} P(\mathbf{n}_{uv}^{kl}) &= P("O^i \text{ a la modalité } D^u \text{ de } V^k \text{ et la modalité } D^v \text{ de } V^l") \\ &= P("O^i \text{ a la modalité } D^u \text{ de } V^k")P("O^i \text{ a la modalité } D^v \text{ de } V^l") \\ &= \frac{\mathbf{n}_{u.}^{kl}}{N} \frac{\mathbf{n}_{.v}^{kl}}{N} \end{aligned}$$

Nous venons de rappeler l'interprétation probabiliste de l'indépendance statistique sous-jacente au critère de Belson en termes contingentiels. De la même manière, nous pouvons associer au numérateur du critère de Janson et Vegelius exprimé en termes relationnels des interprétations probabilistes. Notons symboliquement par $P(C_{ii'}^k)$ la probabilité que les objets O^i et $O^{i'}$ aient la même modalité de la variable V^k . Sous l'hypothèse d'équiprobabilité des modalités d'une variable et sous l'hypothèse d'indépendance, nous avons alors :

$$\begin{aligned} P(C_{ii'}^k) &= \sum_{u=1}^{p_k} P("O^i \text{ a la modalité } D^u \text{ de } V^k \text{ et } O^{i'} \text{ a la modalité } D^u \text{ de } V^k") \\ &= \sum_{u=1}^{p_k} P("O^i \text{ a la modalité } D^u \text{ de } V^k")P("O^{i'} \text{ a la modalité } D^u \text{ de } V^k") \\ &= \sum_{u=1}^{p_k} \frac{1}{p_k} \frac{1}{p_k} \\ &= \frac{1}{p_k} \end{aligned}$$

Nous pouvons donc d'une part interpréter l'expression contingentielle du numérateur du critère de Janson et Vegelius comme étant une transformation de Torgerson et d'autre part interpréter son expression relationnelle comme étant une mesure de l'association entre deux variables sous l'hypothèse de l'équiprobabilité des modalités et de l'indépendance statistique. **Si nous rapprochons le numérateur du critère de Janson et Vegelius du critère de Belson nous voyons que l'indépendance statistique en contingentiel se traduit par une indépendance géométrique (au sens de la transformation de Torgerson) en relationnel et vice versa.**

Enfin, pour aller plus loin, nous pouvons mettre en liaison les résultats de ce paragraphe et l'étude des indices de similarité d'ordre t que nous avons exposée en paragraphe 3.1.2.4. **Les critères d'association mis sous la forme de coefficients de corrélation peuvent être interprétés comme étant des indices d'Ochiaï entre matrices relationnelles transformées et centrées**²³ :

$$\Delta(C^k, C^l, f, \mu^k, \mu^l) = \text{Ochiaï}(f(C^k) - \mu^k U_N, f(C^l) - \mu^l U_N) \quad (3.3.25)$$

Nous proposons donc d'étendre les critères Δ à des indices de similarité d'ordre $t > 0$ et faire jouer un rôle aux normes de Frobenius des matrices relationnelles transformées et centrées.

23. au sens où nous retranchons une moyenne μ à tous les termes de la matrice

Nous rappelons que la norme de Frobenius au carré d'une matrice C , que nous notons par $\|C\|_F^2$ est donnée par la relation suivante :

$$\|C\|_F^2 = \sum_{i,i'} C_{ii'}^2 \quad (3.3.26)$$

Nous pouvons associer à la norme de Frobenius, la métrique de Frobenius dont différentes expressions sont les suivantes :

$$\langle C^k, C^l \rangle_F = \sum_{i,i'} C_{ii'}^k C_{ii'}^l = \text{Trace}({}^t C^k \cdot C^l) \quad (3.3.27)$$

Nous définissons alors la **famille des indices de corrélation d'ordre t** entre variables qualitatives suivante :

Définition 3.3.3 *Indices de corrélation d'ordre t comme critères de mesure d'association entre deux matrices relationnelles*

Soient deux variables qualitatives V^k et V^l représentées par leurs matrices relationnelles respectives C^k et C^l . Nous appellerons indice de corrélation d'ordre $t > 0$ entre les matrices relationnelles transformées et centrées de paramètres (f, μ^k, μ^l) l'indice défini de la manière suivante :

$$\begin{aligned} S_C^t(C^k, C^l, f, \mu^k, \mu^l) &= \frac{\langle C^k, C^l \rangle_F}{M^t(\langle C^k, C^k \rangle_F, \langle C^l, C^l \rangle_F)} \\ &= \frac{\sum_{i,i'} C_{ii'}^k C_{ii'}^l}{M^t(\sum_{i,i'} C_{ii'}^k, \sum_{i,i'} C_{ii'}^l)} \end{aligned} \quad (3.3.28)$$

où

$$\begin{aligned} C^k &= f(C^k) - \mu^k U_N \\ C^l &= f(C^l) - \mu^l U_N \end{aligned} \quad (3.3.29)$$

En résumé, nous avons vu que l'indépendance statistique et l'indétermination pouvaient se ramener à une indépendance géométrique mesurée entre matrices relationnelles qui doivent être vues comme des vecteurs dans un espace vectoriel approprié.

Par ailleurs, nous avons proposé d'étendre les critères d'association à des indices de corrélation basés sur les indices de similarité d'ordre t . Ces derniers permettent de tenir compte de la différence entre les densités de chacune de matrices relationnelles étudiées. Cette densité représente la norme de Frobenius de la matrice relationnelle. Plus les densités sont distinctes, moins les indices de corrélations d'ordre t seront forts.

3.3.2 Indice d'écart à l'indétermination pondérée entre relations d'équivalence

Nous revenons vers la notion d'indétermination qui est définie par la relation suivante :

$$\begin{aligned} V^k \perp_L V^l &\Leftrightarrow 11_{kl} + 00_{kl} - 10_{kl} - 01_{kl} = 0 \\ &\Leftrightarrow \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l = 0 \end{aligned}$$

Deux variables qualitatives sont indéterminées si pour l'ensemble des paires d'objets que chacune rassemble (ou non), il y a autant d'accords (positifs et négatifs) que de désaccords.

Nous pouvons exprimer cette relation de différentes façons. Sachant que $\bar{C}_{ii'}^k = 1 - C_{ii'}^k$, nous avons les trois formulations suivantes qui sont équivalentes :

$$11_{kl} + 00_{kl} - 10_{kl} - 01_{kl} = \begin{cases} \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l - \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l \\ \text{ou} \\ \sum_{i,i'} (C_{ii'}^k - \bar{C}_{ii'}^k) (C_{ii'}^l - \bar{C}_{ii'}^l) \\ \text{ou} \\ 4 \sum_{i,i'} (C_{ii'}^k - \frac{1}{2}) (C_{ii'}^l - \frac{1}{2}) \end{cases} \quad (3.30)$$

Nous cherchons à étendre le concept d'indétermination à celui d'indétermination pondérée en associant des poids à chacun des termes 11_{kl} , 00_{kl} , 10_{kl} et 01_{kl} . Nous noterons ces poids respectivement μ_{11}^{kl} , μ_{00}^{kl} , μ_{10}^{kl} et μ_{01}^{kl} .

Nous proposons la définition suivante : **deux variables qualitatives sont indéterminées si pour l'ensemble des paires d'objets que chacune rassemble (ou non), le poids des accords (positifs et négatifs) est égal au poids des désaccords**. Plus formellement, nous avons la définition suivante :

Définition 3.3.4 Critère d'écart à l'indétermination pondérée

Soient V^k et V^l deux variables qualitatives représentées par leurs matrices relationnelles C^k et C^l . Soient μ_{11}^{kl} , μ_{00}^{kl} , μ_{10}^{kl} et μ_{01}^{kl} des poids accordés aux accords et désaccords définis précédemment. Nous définissons alors le critère d'écart à l'indétermination pondérée $\Lambda(C^k, C^l, \mu_0^k, \mu_0^l)$ entre ces deux variables par :

$$\begin{aligned} \Lambda(C^k, C^l, \mu_{11}^{kl}, \mu_{00}^{kl}, \mu_{10}^{kl}, \mu_{01}^{kl}) &= \mu_{11}^{kl} 11_{kl} + \mu_{00}^{kl} 00_{kl} - \mu_{10}^{kl} 10_{kl} - \mu_{01}^{kl} 01_{kl} \\ &= \mu_{11}^{kl} \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \mu_{00}^{kl} \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l \\ &\quad - \mu_{10}^{kl} \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l - \mu_{01}^{kl} \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l \end{aligned} \quad (3.31)$$

Nous nous intéressons plus particulièrement au cas où **les poids sont séparables**, c'est à dire des poids pour lesquels nous avons :

- $\mu_{11}^{kl} = \mu_1^k \mu_1^l$
- $\mu_{00}^{kl} = \mu_0^k \mu_0^l$
- $\mu_{10}^{kl} = \mu_1^k \mu_0^l$
- $\mu_{01}^{kl} = \mu_0^k \mu_1^l$

où :

- μ_1^k le poids d'un terme $C_{ii'}^k$ qui vaut 1 $\forall i, i' = 1, \dots, N$
- μ_0^k le poids d'un terme $\bar{C}_{ii'}^k$ qui vaut 1 ou encore le poids d'un couple $C_{ii'}^k$ qui vaut 0 $\forall i, i' = 1, \dots, N$
- μ_1^l le poids d'un terme $C_{ii'}^l$ qui vaut 1 $\forall i, i' = 1, \dots, N$
- μ_0^l le poids d'un terme $\bar{C}_{ii'}^l$ qui vaut 1 ou encore le poids d'un terme $C_{ii'}^l$ qui vaut 0 $\forall i, i' = 1, \dots, N$

Nous définissons de plus ce que nous entendons par système de poids normalisés. Il s'agit de poids vérifiant la propriété suivante :

$$\mu_0^k = 1 - \mu_1^k \quad (3.3.32)$$

$$\mu_0^l = 1 - \mu_1^l \quad (3.3.33)$$

Dans le cas de **ponds normalisés**, nous pouvons exprimer tous les poids en fonction de μ_1^k et μ_1^l .

- $\mu_{11}^{kl} = \mu_1^k \mu_1^l$
- $\mu_{00}^{kl} = (1 - \mu_1^k)(1 - \mu_1^l)$
- $\mu_{10}^{kl} = \mu_1^k(1 - \mu_1^l)$
- $\mu_{01}^{kl} = \mu_1^l(1 - \mu_1^k)$

Nous pouvons dès lors aboutir à la propriété ci-dessous qui permet notamment, d'exprimer l'indétermination pondérée sous la forme d'une covariance.

Définition 3.3.5 Critère d'écart à l'indétermination pondérée dans le cas de poids séparables

Soient V^k et V^l deux variables qualitatives représentées par leurs matrices relationnelles C^k et C^l . Soient $\mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l$ un système de poids des termes unitaires et nuls des matrices relationnelles C^k et C^l . Le critère d'écart à l'indétermination pondérée $\Lambda(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l)$ peut alors s'écrire de la manière suivante :

$$\Lambda(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) = \sum_{i,i'} \left(\mu_1^k C_{ii'}^k - \mu_0^k \bar{C}_{ii'}^k \right) \left(\mu_1^l C_{ii'}^l - \mu_0^l \bar{C}_{ii'}^l \right) \quad (3.3.34)$$

Remarque 3.3.2 Lien avec la définition de l'indétermination selon Marcotorchino

Dans le cas où nous avons $\mu_1^k = p_k - 1, \mu_0^k = 1, \mu_1^l = p_l - 1, \mu_0^l = 1$, nous aboutissons à la définition de l'indétermination selon J.F. Marcotorchino donnée par (3.3.11) et définie dans [Mar85]. Nous avons dans ce cas particulier, la relation suivante :

$$\begin{aligned} \Lambda(C^k, C^l, (p_k - 1), 1, (p_l - 1), 1) = 0 &\Leftrightarrow (p_k - 1)(p_l - 1) \sum_{i,i'} C_{ii'}^k C_{ii'}^l + \sum_{i,i'} \bar{C}_{ii'}^k \bar{C}_{ii'}^l \\ &\quad - (p_k - 1) \sum_{i,i'} C_{ii'}^k \bar{C}_{ii'}^l - (p_l - 1) \sum_{i,i'} \bar{C}_{ii'}^k C_{ii'}^l = 0 \\ &\Leftrightarrow p_k p_l \sum_{u,v} \left(\mathbf{n}_{uv}^{kl} - \left[\frac{\mathbf{n}_{u.}^{kl}}{p_l} + \frac{\mathbf{n}_{.v}^{kl}}{p_k} - \frac{\mathbf{n}_{.}^{kl}}{p_k p_l} \right] \right)^2 = 0 \end{aligned}$$

Nous avons par ailleurs la propriété ci-dessous lorsque les poids sont de plus normalisés :

Propriété 3.3.7 Critère d'écart à l'indétermination pondérée dans le cas de poids séparables et normalisés

Si en plus d'être séparable, le système de poids $\mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l$ est normalisé, nous avons l'écriture suivante du critère d'écart à l'indétermination qui rappelle celle d'une covariance :

$$\Lambda(C^k, C^l, \mu_1^k, (1 - \mu_1^k), \mu_1^l, (1 - \mu_1^l)) = \sum_{i,i'} (C_{ii'}^k - (1 - \mu_1^k)) (C_{ii'}^l - (1 - \mu_1^l)) \quad (3.3.35)$$

Démonstration

Nous avons dans un premier temps pour (3.3.34) :

$$\begin{aligned}
 \sum_{i,i'} \begin{pmatrix} \mu_1^k C_{ii'}^k - (1 - \mu_1^k) \bar{C}_{ii'}^k \\ \mu_1^l C_{ii'}^l - (1 - \mu_1^l) \bar{C}_{ii'}^l \end{pmatrix} &= \sum_{i,i'} \left[\mu_1^k \mu_1^l C_{ii'}^k C_{ii'}^l + (1 - \mu_1^k)(1 - \mu_1^l) \bar{C}_{ii'}^k \bar{C}_{ii'}^l \right. \\
 &\quad \left. - (1 - \mu_1^l) \mu_1^k C_{ii'}^k \bar{C}_{ii'}^l - (1 - \mu_1^k) \mu_1^l \bar{C}_{ii'}^k C_{ii'}^l \right] \\
 &= \mu_1^k \mu_1^l \sum C_{ii'}^k C_{ii'}^l + (1 - \mu_1^k)(1 - \mu_1^l) \sum \bar{C}_{ii'}^k \bar{C}_{ii'}^l \\
 &\quad - (1 - \mu_1^l) \mu_1^k \sum C_{ii'}^k \bar{C}_{ii'}^l - (1 - \mu_1^k) \mu_1^l \sum \bar{C}_{ii'}^k C_{ii'}^l \\
 &= \mu_{11}^{kl} 11_{kl} + \mu_{00}^{kl} 00_{kl} - \mu_{10}^{kl} 10_{kl} - \mu_{01}^{kl} 01_{kl}
 \end{aligned}$$

Dans un deuxième temps, nous avons pour (3.3.35) :

$$\begin{aligned}
 \mu_1^k \mu_1^l \sum C_{ii'}^k C_{ii'}^l &+ (1 - \mu_1^k)(1 - \mu_1^l) \sum \bar{C}_{ii'}^k \bar{C}_{ii'}^l \\
 - (1 - \mu_1^l) \mu_1^k \sum C_{ii'}^k \bar{C}_{ii'}^l &- (1 - \mu_1^k) \mu_1^l \sum \bar{C}_{ii'}^k C_{ii'}^l \\
 &= \sum C_{ii'}^k C_{ii'}^l \left[\mu_1^k \mu_1^l + (1 - \mu_1^k)(1 - \mu_1^l) + (1 - \mu_1^l) \mu_1^k + (1 - \mu_1^k) \mu_1^l \right] \\
 &\quad - \sum C_{ii'}^k \left[(1 - \mu_1^k)(1 - \mu_1^l) + (1 - \mu_1^l) \mu_1^k \right] \\
 &\quad - \sum C_{ii'}^l \left[(1 - \mu_1^k)(1 - \mu_1^l) + (1 - \mu_1^k) \mu_1^l \right] \\
 &\quad + N^2 (1 - \mu_1^k)(1 - \mu_1^l) \\
 &= \sum C_{ii'}^k C_{ii'}^l - (1 - \mu_1^l) \sum C_{ii'}^k - (1 - \mu_1^k) \sum C_{ii'}^l + N^2 (1 - \mu_1^k)(1 - \mu_1^l) \\
 &= \sum (C_{ii'}^k C_{ii'}^l - (1 - \mu_1^l) C_{ii'}^k - (1 - \mu_1^k) C_{ii'}^l + (1 - \mu_1^k)(1 - \mu_1^l)) \\
 &= \sum_{i,i'} (C_{ii'}^k - (1 - \mu_1^k)) (C_{ii'}^l - (1 - \mu_1^l))
 \end{aligned}$$

□

Nous proposons une version normalisée du critère d'indétermination pondérée qui est similaire aux indices de corrélation d'ordre t entre deux variables qualitatives et qui nous conduit finalement à la définition d'indices d'écart à l'indétermination pondérée.

Définition 3.3.6 *Indice d'ordre t d'écart à l'indétermination pondérée entre deux matrices relationnelles de partitions*

Soient deux variables V^k et V^l représentées par leurs matrices relationnelles respectives C^k et C^l . Nous appellerons indice d'écart à l'indétermination pondéré d'ordre $t > 0$ entre les matrices relationnelles, munies de poids séparables $(\mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l)$, l'indice défini de la manière suivante :

$$\begin{aligned}
 S_t^t(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) &= \frac{\langle C^k, C^l \rangle_F}{\mathcal{M}^t(\langle C^k, C^k \rangle_F, \langle C^l, C^l \rangle_F)} \\
 &= \frac{\sum_{i,i'} C_{ii'}^k C_{ii'}^l}{\mathcal{M}^t(\sum_{i,i'} C_{ii'}^k C_{ii'}^k, \sum_{i,i'} C_{ii'}^l C_{ii'}^l)}
 \end{aligned} \tag{3.3.36}$$

où

C^k, C^l	Approche logique de l'indétermination	Approche géométrique de l'indétermination lorsque de plus les poids sont normalisés
$C^k =$	$\mu_1^k C^k - \mu_0^k \bar{C}^k$	$C^k - (1 - \mu_1^k) U_N$
$C^l =$	$\mu_1^l C^l - \mu_0^l \bar{C}^l$	$C^l - (1 - \mu_1^l) U_N$

et où U_N est la matrice carrée de dimension N remplie de 1.

Nous concluons cette sous-section en discutant des liens et distinctions entre indices de corrélation et indices d'écart à l'indétermination pondérée que nous avons définis.

Remarque 3.3.3 *Liens entre indices de corrélation et indices d'écart à l'indétermination entre deux variables qualitatives*

Indices de corrélation et indices d'écart à l'indétermination pondérée sont des mesures de liaisons entre variables qualitatives, représentées par leurs matrices relationnelles, qui sont tous deux définis par des indices de similarité d'ordre t que nous avons introduits en déf-3.1.11 page 82. Cependant, leurs interprétations sont bien distinctes. Pour fixer les idées plaçons nous dans le cas des poids séparables et normalisés suivants : $\mu_1^k = \frac{1}{N^2} \sum_{i,i'} C_{ii'}^k$ et $\mu_0^k = \frac{1}{N^2} \sum_{i,i'} \bar{C}_{ii'}^k = 1 - \mu_1^k$.

- En ce qui concerne les indices de corrélation :
 - les données sont d'abord transformées selon une fonction f , puis centrées selon une **moyenne des "données unitaires"** telle que $\mu^k = \mu_1^k$
 - l'indice s'annule en cas d'indépendance géométrique entre les matrices relationnelles
 - l'indice vaut 1 lorsque les matrices relationnelles sont identiques
 - l'ordre t fait jouer un rôle plus ou moins important au rapport des normes de Frobenius des matrices relationnelles transformées et centrées. Nous précisons que plus la différence de densité entre deux matrices relationnelles est grande, plus l'indice de corrélation est proche de 0.
- En ce qui concerne les indices d'écart à l'indétermination pondérée :
 - les données ne sont pas transformées mais la simplification sous la forme d'une covariance montre que les données sont "centrées" selon une **moyenne des "données nulles"** telle que $(1 - \mu^k) = \mu_0^k = \frac{1}{N^2} \sum_{i,i'} \bar{C}_{ii'}^k = 1 - \frac{1}{N^2} \sum_{i,i'} C_{ii'}^k$. Son interprétation est davantage logique que géométrique. Toutefois, sa simplification permet de mettre en évidence une formulation proche de ceux des indices de corrélation :

$$\boxed{S_I^t(C^k, C^l, \mu^k, (1 - \mu^k), \mu^l, (1 - \mu^l)) = S_C^t(C^k, C^l, Id, (1 - \mu^k), (1 - \mu^l))} \quad (3.3.38)$$

De manière plus générale, c'est à dire pour des poids non normalisés, nous avons :

$$\boxed{S_I^t(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) = S_C^t\left(C^k, C^l, Id, \frac{\mu_0^k}{\mu_1^k + \mu_0^k}, \frac{\mu_0^l}{\mu_1^l + \mu_0^l}\right)} \quad (3.3.39)$$

Par ailleurs, lorsque $\mu_1^k = \mu_0^k = \mu_1^l = \mu_0^l$, nous voyons que nous aboutissons au critère $2\text{Rand}(C^k, C^l) - 1 = S_C^t(C^k, C^l, Id, 1/2, 1/2)$

- l'indice s'annule en cas d'indétermination pondérée entre les matrices relationnelles
- l'indice vaut 1 lorsque les matrices relationnelles sont identiques
- l'ordre t fait jouer un rôle plus ou moins important aux "normes" de Frobenius des matrices relationnelles "centrées". Indices de corrélation et d'écart à l'indétermination pondérée font jouer le même rôle aux normes. Ici aussi, plus la différence entre les densités respectives des deux matrices est grande, plus l'indice d'écart à l'indétermination est proche de 0.

CHAPITRE 4

Les enveloppes de Condorcet et le calcul consensuel

Sommaire

4.1 Structures algébriques booléennes des matrices relationnelles	114
4.1.1 Algèbre des relations binaires	115
4.1.1.1 Algèbre des relations binaires non nécessairement structurées	115
4.1.1.2 Stabilité des propriétés relationnelles sous l'action des intersections et des unions relationnelles	117
4.1.2 Sur les fondements logiques des contraintes relationnelles des matrices relationnelles	118
4.2 Les enveloppes de Condorcet	124
4.2.1 Définitions et propriétés de l'enveloppe de Condorcet	125
4.2.1.1 Complémentaires d'une enveloppe de Condorcet	127
4.2.1.2 Décomposition d'une enveloppe de Condorcet selon la majorité m	128
4.2.2 Calcul de l'enveloppe de Condorcet par une approche combinatoire	131
4.2.2.1 Notations	131
4.2.2.2 Calcul de l'enveloppe de Condorcet par inversion du triangle de Pascal	132
4.2.2.3 Propriétés des coefficients $\alpha_k^{m,M}$	137
4.2.2.3.1 Les coefficients $\alpha_k^{m,M}$ et les coefficients binomiaux	137
4.2.2.3.2 Les coefficients $\alpha_k^{m,M}$ et les nombres de Lah	138
4.2.2.3.3 Les coefficients $\alpha_k^{m,M}$ et les coefficients binomiaux négatifs	139
4.3 Le calcul consensuel ou l'application du calcul de l'enveloppe de Condorcet à des données non binaires	142
4.3.1 Le calcul consensuel comme méthode d'agrégation et son application dans le cadre de relations d'équivalence	143
4.3.2 Modélisation de la mesure consensuelle de la similarité entre objets	146
4.3.2.1 Cas des variables qualitatives	146
4.3.2.1.1 Matrices relationnelles individuelles de similarités	146
4.3.2.1.2 Matrices d'intersections relationnelles de similarités	146
4.3.2.1.3 Matrices d'intersections relationnelles d'indices de similarité	147

4.3.2.2	Cas des variables quantitatives	147
4.3.2.2.1	Matrices relationnelles individuelles de similarités	147
4.3.2.2.2	Matrices d'intersections relationnelles d'indices de similarité	148
4.3.2.3	Cas des variables qualitatives et quantitatives	149

Nous abordons dans ce chapitre des aspects logiques et combinatoires de l'AR.

Tout d'abord, nous définissons des structures algébriques booléennes sur les matrices relationnelles et rappelons succinctement les opérations de l'algèbre relationnelle qui est l'algèbre des relations binaires. Ces dernières interviendront régulièrement dans la suite de ce mémoire notamment au chapitre 6.

Nous nous intéressons ensuite aux fondements logiques des contraintes linéaires issues de l'AR en montrant que celles-ci sont obtenues par les opérateurs de conjonction et de disjonction de Lukasiewicz. Nous proposons une méthode qui permettrait, dans de nombreux cas, de modéliser sous forme de contraintes linéaires des propriétés relationnelles qui s'expriment initialement sous la forme de propositions logiques.

Nous développons enfin une méthode d'agrégation de nature combinatoire que nous nommons enveloppe de Condorcet et calcul consensuel. Celle-ci repose en fait sur les formules combinatoires de Ch. Jordan qui généralisent celle de H. Poincaré. Elle permet de donner au problème d'agrégation une approche logique puisque nous calculons la mesure qu'il y ait une majorité m de variables ou juges supportant la relation $O^i \mathcal{R} O^{i'}$ entre deux objets. Nous proposons cette méthode comme étant une troisième extension de l'AR au traitement de données de nature non nécessairement binaire. En particulier, les données étudiées sont de type floues ou probabilistes. Nous montrons aussi en quoi notre approche pourrait constituer une méthode originale de la classification automatique de données numériques hétérogènes.

Nous nous appuyons dans ce chapitre, sur les structures algébriques booléennes des matrices relationnelles. Compte tenu du caractère binaire de ces dernières, l'étude de leurs aspects algébriques se place en effet, dans le cadre de l'algèbre de Boole. Cette dernière est la structure algébrique sous-jacente à la fois à la théorie des ensembles et à la logique bivalente.

Par conséquent, tout au long de ce chapitre, nous interpréterons les matrices relationnelles selon deux manières, qu'il est utile de préciser en préambule. Nous aurons effectivement la double interprétation suivante :

- **ensembliste**, puisque la matrice relationnelle, d'un point de vue global, code l'ensemble des paires d'objets appartenant au graphe de la relation \mathcal{R} étudiée
- **logique**, étant donné que localement à une paire d'objets $(O^i, O^{i'})$, nous codons la valeur de vérité de la proposition $O^i \mathcal{R} O^{i'}$: " O^i est en relation avec $O^{i'}$ selon la relation \mathcal{R} "

Par conséquent, nous utiliserons dans la suite de ce mémoire de thèse de manière indifférente les termes intersection et conjonction ou les termes union et disjonction.

4.1 Structures algébriques booléennes des matrices relationnelles

Nous appliquons dans un premier temps, les opérations d'union et d'intersection définies en algèbre booléenne, aux matrices relationnelles binaires. Nous introduisons ensuite le calcul et l'algèbre des relations proposés initialement par A. De Morgan et C.S. Peirce puis développées notamment par A. Tarski [Tar41].

Dans un second temps, nous nous intéressons aux fondements logiques de l'AR en investiguant l'écriture sous forme de contraintes linéaires des propriétés relationnelles. Nous verrons comment cette approche permet *a priori*, de généraliser certains aspects de l'AR à tout type de relations binaires structurées.

Enfin, nous étudions la stabilité de quelques propriétés relationnelles sous l'action des opérations booléennes afin de dégager des structures algébriques à des types particuliers de relations binaires.

4.1.1 Algèbre des relations binaires

4.1.1.1 Algèbre des relations binaires non nécessairement structurées

Nous considérons l'ensemble des matrices carrées binaires de taille N que nous notons par \mathbb{C} . Les éléments de cet ensemble seront appelés matrices relationnelles. L'ensemble des matrices relationnelles des relations binaires structurées¹ est donc inclus dans \mathbb{C} .

Nous pouvons alors munir l'ensemble \mathbb{C} des opérations d'intersection, d'union et de complémentaire suivants :

Définition 4.1.1 Opérations d'intersection, d'union et de complémentaires relationnelles

Soit C^k et C^l deux matrices appartenant à l'ensemble \mathbb{C} .

L'intersection (ou la conjonction ou la multiplication) relationnelle entre les matrices C^k et C^l est notée $C^k \otimes C^l$ et son terme général est donné par la relation suivante, $\forall i, i' = 1, \dots, N$:

$$[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = \begin{cases} 1 & \text{si } C_{ii'}^k = 1 \text{ et } C_{ii'}^l = 1 \\ 0 & \text{sinon} \end{cases} \quad (4.1.1)$$

L'union (ou la disjonction ou l'addition) relationnelle entre ces deux matrices est notée $C^k \oplus C^l$ et son terme général est donné par la relation suivante, $\forall i, i' = 1, \dots, N$:

$$[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = \begin{cases} 1 & \text{si } C_{ii'}^k = 1 \text{ ou } C_{ii'}^l = 1 \\ 0 & \text{sinon} \end{cases} \quad (4.1.2)$$

Le complémentaire d'une matrice C^k , noté \overline{C}^k , est donné par la relation suivante :

$$\overline{C}^k = U_N - C^k \quad (4.1.3)$$

où U_N est la matrice carrée de taille N remplie de 1.

Remarque 4.1.1 A propos des symboles \oplus et \otimes

Nous avons choisi d'utiliser pour l'union et l'intersection relationnelles, des symboles distincts des notations classiques \cup et \cap utilisée en théorie des ensembles. En effet, tout d'abord, comme le mode de représentation des ensembles que définissent les graphes des relations étudiées est, dans le cadre de l'AR, une matrice², nous avons préféré avoir des notations rappelant les symboles arithmétiques $+$ et \times . Puis, nous étendrons l'algèbre de Boole au cas d'algèbres multivalentes qui utilisent des valeurs d'appartenance ou de vérité qui ne sont plus nécessairement 0 ou 1 mais un réel compris entre 0 et 1. Dans ce cas, les unions et les intersections relationnelles peuvent prendre plusieurs formes selon l'intensité de la conjonction³ entre deux propositions. Ainsi, en raison de la représentation sous forme de matrices des ensembles définis par les relations, et de notre volonté d'être générique en laissant la possibilité de choisir un type de

1. nous dirons qu'une relation binaire est structurée si elle possède des propriétés relationnelles cf tab-1.1 et tab-1.2

2. une matrice relationnelle représente un sous-ensemble du produit cartésien $\mathbb{O} \times \mathbb{O}$ de la manière suivante : $C_{ii'}^k = v^k(O^i \mathcal{R}^k O^{i'})$ où v^k est la fonction caractéristique de \mathcal{R}^k qui renvoie 1 si la paire $(O^i, O^{i'})$ appartient au graphe de la relation ; 0 sinon

3. ces conjonctions sont des fonctions particulières appelées "triangular norms" et notée "t-norms" que nous présenterons ultérieurement cf déf-4.3.1 page 142

conjonction particulier, nous avons décidé d'adopter des notations particulières. Dans un souci de clarté, nous ne manquerons pas de préciser la définition de ces symboles selon le contexte dans lequel nous nous plaçons.

Suite à la remarque précédente, nous précisons ci-dessous, notre interprétation des différentes notations que nous utiliserons dans la section 4.1 toute entière.

Remarque 4.1.2 *Signification des termes $C_{ii'}^k$, et des symboles \oplus et \otimes dans le cadre de la section 4.1*

Dans le cadre de la section 4.1, nous privilégierons l'interprétation logique (bivalente) des termes de la matrice relationnelle. Nous avons alors :

- $C_{ii'}^k = v(O^i \mathcal{R}^k O^{i'}) =$ valeur de vérité binaire de la proposition “ O^i est en relation avec $O^{i'}$ selon la relation \mathcal{R}^k ”
- $[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = v(O^i \mathcal{R}^k O^{i'} \vee O^i \mathcal{R}^l O^{i'}) =$ valeur de vérité binaire de la disjonction “ O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R}^k ou la relation \mathcal{R}^l ”
- $[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = v(O^i \mathcal{R}^k O^{i'} \wedge O^i \mathcal{R}^l O^{i'}) =$ valeur de vérité binaire de la conjonction “ O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R}^k et la relation \mathcal{R}^l ”

Propriété 4.1.1 *Algèbre de Boole $\mathcal{B}(\mathbb{C})$*

L'ensemble \mathbb{C} des matrices carrées binaires de taille N , muni de l'opération unaire de complémentarité et des opérations binaires d'intersection et d'union relationnelles dont les éléments neutres sont respectivement U_N et \bar{U}_N forme une algèbre de Boole que nous notons $\mathcal{B}(\mathbb{C})$.

$$\mathcal{B}(\mathbb{C}) = \{\mathbb{C}, -, \otimes, \oplus, U_N, \bar{U}_N\} \quad (4.1.4)$$

où U_N est la matrice carrée de taille N remplie de 1 et $\bar{U}_N = U_N - U_N$, est la matrice carrée de taille N remplie de 0.

En pratique, la conjonction et la disjonction de deux matrices carrées binaires peuvent être données par différentes opérations que nous empruntons à la logique multivalente [KMP00] et dont nous citerons quelques unes dans la sous-section 4.3. Dans le cadre de la logique bivalente dans lequel nous nous plaçons pour l'instant et pour fixer les idées, nous pourrions prendre les expressions particulières suivantes :

$$[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = \min(C_{ii'}^k, C_{ii'}^l) \quad (4.1.5)$$

$$[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = \max(C_{ii'}^k, C_{ii'}^l) \quad (4.1.6)$$

Les structures algébriques peuvent par ailleurs, être enrichies de diverses autres opérations unaires et binaires. C'est l'objet de l'algèbre des relations binaires qui cherchent à étendre l'algèbre de Boole à des structures relationnelles qui sont plus riches [L'A58] [Mad91] [Pra92]. Dans ce contexte, les opérations suivantes dites peirciennes ont été rajoutées.

Définition 4.1.2 *Opérations peirciennes de multiplication relative, d'addition relative et de converse*

Soient C^k et C^l des matrices appartenant à l'ensemble \mathbb{C} .

La multiplication relative entre les matrices C^k et C^l sera notée $C^k \odot C^l$ étant donné qu'il s'agit d'une multiplication matricielle particulière. Son terme général est donné par la relation suivante, $\forall i, i' = 1, \dots, N$:

$$[C^k \odot C^l]_{ii'} = \bigoplus_{i''=1}^N C_{ii''}^k \otimes C_{i''i'}^l \quad (4.1.7)$$

$$= \begin{cases} 1 & \text{si } \exists i'' \in \{1, \dots, N\} : C_{ii''}^k \otimes C_{i''i'}^l = 1 \\ 0 & \text{sinon} \end{cases} \quad (4.1.8)$$

L'addition relative entre ces deux matrices sera $C^k \uplus C^l$ et son terme général est donné par la relation suivante, $\forall i, i' = 1, \dots, N$:

$$[C^k \uplus C^l]_{ii'} = \bigotimes_{i''=1}^N C_{ii''}^k \oplus C_{i''i'}^l \quad (4.1.9)$$

$$= \begin{cases} 1 & \text{si } \forall i'' \in \{1, \dots, N\} : C_{ii''}^k \oplus C_{i''i'}^l = 1 \\ 0 & \text{sinon} \end{cases} \quad (4.1.10)$$

La converse (ou la transposée) d'une matrice C^k , noté \check{C}^k , est donné par la relation suivante :

$$\check{C}^k = {}^t C^k \quad (4.1.11)$$

Nous avons introduit les opérations peirciennes notamment pour les raisons suivantes :

- la transitivité est une propriété très structurante dans le cas des relations binaires et la multiplication relative aussi appelée multiplication matricielle booléenne permet d'obtenir des fermetures transitives des relations
- la converse ou transposée d'une relation d'ordre total donne la relation duale naturelle de ce type de relation et non son complémentaire contrairement aux relations d'équivalence. Nous emploierons d'ailleurs cette opération unaire lors de l'étude des critères d'association entre relations d'ordre total dans la section 6.1 du chapitre 6 afin de mettre en évidence la notion d'écart à l'indétermination pour ce type de relation
- les opérations peirciennes enfin, trouvent des applications en théorie du choix social. Nous montrons en sous-section 6.2.2 du chapitre 6 qu'elles sont liées aux propriétés relationnelles dans un contexte d'agrégation de relations selon une règle majoritaire

4.1.1.2 Stabilité des propriétés relationnelles sous l'action des intersections et des unions relationnelles

Nous donnons succinctement des résultats sur les structures algébriques de certaines classes de relations binaires structurées. En effet, certaines propriétés sont stables sous l'action de différentes opérations parmi la complémentarité, la disjonction et la conjonction comme nous le mentionnons dans la propriété ci-dessous.

Propriété 4.1.2 Stabilité des propriétés relationnelles

Certaines des propriétés relationnelles sont stables selon l'action unaire ou binaire de la complémentarité, de l'intersection relationnelle ou de l'union relationnelle. Nous résumons ces résultats à l'aide du tableau tab-4.1.

Propriétés relationnelles	Complémentarité	Intersection relationnelle	Union relationnelle
Réflexivité		✓	✓
Symétrie	✓	✓	✓
Asymétrie		✓	
Transitivité		✓	
Totalité			✓

TABLE 4.1 – Stabilité des propriétés relationnelles sous l'action des opérations booléennes

Démonstration

Nous donnons les indications de la démonstration. En ce qui concerne la réflexivité, la symétrie et la totalité, le résultat est immédiat dans le cas de l’affirmation. Dans le cas de l’infirmation des autres propriétés, on trouve aisément des contre-exemples.

L’asymétrie est stable pour l’intersection relationnelle. Prenons deux matrices relationnelles C^k et C^l qui sont asymétriques. Par conséquent, ces dernières vérifient les configurations possibles suivantes :

$C_{ii'}^k$	1	0	0
$C_{i'i}^k$	0	1	0

Compte tenu des configurations possibles, il suffit alors d’énumérer tous les cas de croisement $C_{ii'}^k, C_{ii'}^l$, qui représente le terme général de $C^k \otimes C^l$, et de vérifier que $C_{ii'}^k C_{ii'}^l + C_{i'i}^k C_{i'i}^l \leq 1$. Nous avons alors le résultat :

$C_{ii'}^k$	1	0	0	1	0	0	1	0	0
$C_{i'i}^k$	0	1	0	0	1	0	0	1	0
$C_{ii'}^l$	1	1	1	0	0	0	0	0	0
$C_{i'i}^l$	0	0	0	1	1	1	0	0	0
$C_{ii'}^k C_{ii'}^l + C_{i'i}^k C_{i'i}^l$	1	0	0	0	1	0	0	0	0

Une démarche similaire nous donne le résultat pour la stabilité de la transitivité sous l’action de l’intersection relationnelle.

□

Un des résultats intéressant que nous pouvons rappeler suite à cette étude est le suivant.

Propriété 4.1.3 Monoïde des partitions

Soit l’ensemble des matrices relationnelles des relations d’équivalence définies sur un ensemble d’objets \mathbb{O} que nous notons $\mathbb{RE}(\mathbb{O})$. L’ensemble de ces matrices relationnelles sont des éléments du sous-ensemble de \mathbb{C} qui vérifient simultanément les propriétés de réflexivité, de symétrie et de transitivité. Ce sous-ensemble muni de l’opération d’intersection relationnelle forme un monoïde.

En d’autres termes l’intersection relationnelle de deux matrices relationnelles de partitions est une matrice relationnelle de partition.

4.1.2 Sur les fondements logiques des contraintes relationnelles des matrices relationnelles

L’étude des différentes opérations de la logique multivalente a permis de donner des fondements logiques aux résultats de l’AR qui permettent, dans le cadre des matrices relationnelles, de coder les propriétés relationnelles par des contraintes linéaires d’égalités ou d’inégalités.

Soient A et B deux propositions. Dans le contexte de la logique classique issue de l’algèbre booléenne, ces deux propositions peuvent prendre deux valeurs : 1 si la proposition est vraie, 0 si elle est fausse.

La logique de Lukasiewicz cherche à étendre la logique classique, vers une logique trivalente pour laquelle, une proposition peut de plus posséder la valeur $1/2$: c’est le cas des propositions incertaines.

Pour permettre le calcul propositionnel dans le cadre de sa logique trivalente, J. Lukasiewicz définit deux nouvelles opérations donnant la valeur logique de la conjonction et de la disjonction de deux propositions. Ces opérations sont respectivement :

$$v(A \wedge B) = W(A, B) = \max(v(A) + v(B) - 1, 0) \tag{4.1.12}$$

$$v(A \vee B) = W^*(A, B) = \min(v(A) + v(B), 1) \tag{4.1.13}$$

où $v(A) \in \{0, 1/2, 1\}$ représente la valeur logique de la proposition A .

Nous avons alors la table de vérité donnée par tab-4.2.

$v(A)$	$v(B)$	$v(A \wedge B)$	$v(A \vee B)$	$v(A \Rightarrow B) = v(\bar{A} \vee B)$
0	0	0	0	1
0	1	0	1	1
1	0	0	1	0
1	1	1	1	1
0	1/2	0	1/2	1
1	1/2	1/2	1	1/2
1/2	0	0	1/2	1/2
1/2	1	1/2	1	1
1/2	1/2	0	1	1

TABLE 4.2 – Tables de vérité de la logique trivalente de Lukasiewicz

Nous ne nous intéressons pas au traitement des propositions incertaines. Par contre, la première partie du tableau nous montre que les opérations de Lukasiewicz permettent de retrouver les résultats classique du calcul propositionnel issu de la logique bivalente. **Ce que nous cherchons à montrer c'est en quoi les opérations de conjonction et de disjonction de Lukasiewicz justifient et étendent, dans le cadre de l'AR, la modélisation des propriétés relationnelles sous la forme de contraintes linéaires d'égalités ou d'inégalités.**

Au préalable, nous remarquons à l'aide de cette table les propriétés suivantes.

Propriété 4.1.4 *Contraintes linéaires issues de l'implication et de l'équivalence entre propositions*

Nous avons les propriétés suivantes qui retranscrivent les relations d'implication et d'équivalence entre propositions sous la forme de contraintes d'inégalités et d'égalités.

$$\{v(A \Rightarrow B) = 1\} \Leftrightarrow \{v(A) \leq v(B)\} \quad (4.1.14)$$

$$\{v(A \Leftrightarrow B) = 1\} \Leftrightarrow \{v(A) = v(B)\} \quad (4.1.15)$$

Par ailleurs, il est utile de rappeler pour la suite, les expressions algébriques des fonctions min et max de deux nombres a et b :

$$\min(a, b) = \frac{1}{2}(a + b) - \frac{1}{2} |a - b| \quad (4.1.16)$$

$$\max(a, b) = \frac{1}{2}(a + b) + \frac{1}{2} |a - b| \quad (4.1.17)$$

Nous rappelons aussi, les expressions suivantes :

$$|a| \leq b \Leftrightarrow -b \leq a \leq b \quad (4.1.18)$$

$$|a| \geq b \Leftrightarrow a \leq -b \vee b \leq a \quad (4.1.19)$$

A partir de ces rappels, nous démontrons comment les propriétés de symétrie, d'antisymétrie forte, de totalité et de transitivité qui s'énoncent sous la forme de conjonctions, de disjonctions et d'implications entre propositions, peuvent s'écrire sous forme d'inégalités linéaires en utilisant les opérations W et W^* .

Démonstration *Symétrie*

$\{O^i \mathcal{R} O^{i'} \Rightarrow O^{i'} \mathcal{R} O^i\}$ et de la même manière, $\{O^{i'} \mathcal{R} O^i \Rightarrow O^i \mathcal{R} O^{i'}\}$

Par conséquent, une relation binaire est symétrique si et seulement si $\{O^i \mathcal{R} O^{i'} \Leftrightarrow O^{i'} \mathcal{R} O^i\}$

Ainsi, $v(O^i \mathcal{R} O^{i'}) = v(O^{i'} \mathcal{R} O^i)$. Donc nous avons :

$$C_{ii'} = C_{i'i}$$

□

Démonstration Antisymétrie forte

$$\begin{aligned} \{O^i \mathcal{R} O^{i'} \Rightarrow O^{i'} \overline{\mathcal{R}} O^i\} &\Leftrightarrow v(O^i \mathcal{R} O^{i'}) \leq v(O^{i'} \overline{\mathcal{R}} O^i) \\ &\Leftrightarrow C_{ii'} \leq \overline{C}_{i'i} \\ &\Leftrightarrow C_{ii'} + C_{i'i} \leq 1 \end{aligned}$$

□

Démonstration Transitivité

$$\begin{aligned} \{O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''}\} \\ &\Leftrightarrow v(O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''}) \leq v(O^i \mathcal{R} O^{i''}) \\ &\Leftrightarrow \max(C_{ii'} + C_{i'i''} - 1, 0) \leq C_{ii''} \\ &\Leftrightarrow C_{ii'} + C_{i'i''} - 1 \leq C_{ii''} \text{ puisque } C_{ii''} \geq 0 \\ &\Leftrightarrow C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 \end{aligned}$$

□

Démonstration Totalité

$$\begin{aligned} \{O^i \mathcal{R} O^{i'} \vee O^{i'} \overline{\mathcal{R}} O^i\} \\ &\Leftrightarrow v(O^i \mathcal{R} O^{i'} \vee O^{i'} \overline{\mathcal{R}} O^i) = 1 \\ &\Leftrightarrow \min(C_{ii'} + C_{i'i}, 1) = 1 \\ &\Leftrightarrow C_{ii'} + C_{i'i} \geq 1 \end{aligned}$$

□

Pour robustifier nos résultats, nous ajoutons aux précédentes propriétés d'une part, celles données par P. Michaud dans [MM79] et qui sont les propriétés d'intermédiarité et de cohérence; d'autres part, celles d'antisymétrie faible, d'acyclicité, de Ferrers, de semi-transitivité et de quasi-transitivité :

- Intermédiarité : Si O^i et $O^{i'}$ sont en relation et, $O^{i''}$ et $O^{i'''}$ sont non comparables, et $O^{i''}$ et $O^{i'''}$ sont aussi en relation alors nous avons O^i en relation avec $O^{i'''}$. Cette propriété est traduite sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge (O^{i'} \overline{\mathcal{R}} O^{i''} \wedge O^{i''} \overline{\mathcal{R}} O^{i'''}) \wedge O^{i''} \mathcal{R} O^{i'''} \Rightarrow O^i \mathcal{R} O^{i'''} \quad \forall (O^i, O^{i'}, O^{i''}, O^{i'''}) \in \mathbb{O}^4$$

- Cohérence : Si O^i et $O^{i'}$ sont en relation et, $O^{i'}$ et $O^{i''}$ sont aussi en relation alors les deux propositions suivantes ne peuvent être simultanément vérifiées : O^i et $O^{i''}$ ne sont pas comparables et $O^{i'}$ et $O^{i''}$ ne sont pas comparables. Cette propriété s'exprime sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow \neg \left((O^i \overline{\mathcal{R}} O^{i''} \wedge O^{i''} \overline{\mathcal{R}} O^i) \wedge (O^{i'} \overline{\mathcal{R}} O^{i''} \wedge O^{i''} \overline{\mathcal{R}} O^{i'}) \right) \quad \forall (O^i, O^{i'}, O^{i''}, O^{i'''}) \in \mathbb{O}^4$$

- Antisymétrie faible : Si O^i et $O^{i'}$ sont en relation et, $O^{i'}$ et O^i sont aussi en relation alors $O^i = O^{i'}$. Cette propriété s'exprime sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^i \Rightarrow O^i = O^{i'}$$

- Acyclicité : Si O^i et $O^{i'}$ sont en relation et, $O^{i'}$ et $O^{i''}$ sont aussi en relation alors $O^{i''}$ n'est pas en relation avec O^i . Cette propriété s'exprime sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^{i''} \overline{\mathcal{R}} O^i$$

- Propriété de de Ferrers : Si O^i et $O^{i'}$ sont en relation et, $O^{i''}$ et $O^{i''}$ sont aussi en relation alors O^i est en relation avec $O^{i''}$ ou $O^{i''}$ est en relation avec $O^{i'}$. Cette propriété s'exprime sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge O^{i''} \mathcal{R} O^{i'''} \Rightarrow O^i \mathcal{R} O^{i'''} \vee O^{i''} \mathcal{R} O^{i'}$$

- Propriété de semi-transitivité : Si O^i et $O^{i'}$ sont en relation et, $O^{i'}$ et $O^{i''}$ sont aussi en relation alors O^i est en relation avec $O^{i''}$ ou $O^{i''}$ est en relation avec $O^{i'}$. Cette propriété s'exprime sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''} \vee O^{i''} \mathcal{R} O^{i'}$$

- Propriété de quasi-transitivité : Si O^i et $O^{i'}$ sont en relation et, $O^{i'}$ et O^i ne sont pas en relation et, $O^{i'}$ et $O^{i''}$ sont en relation et, $O^{i''}$ et O^i ne sont pas en relation, alors O^i est en relation avec $O^{i''}$ et $O^{i''}$ n'est pas en relation avec O^i . Cette propriété s'exprime sous la forme propositionnelle suivante :

$$O^i \mathcal{R} O^{i'} \wedge O^{i'} \bar{\mathcal{R}} O^i \wedge O^{i'} \mathcal{R} O^{i''} \wedge O^{i''} \bar{\mathcal{R}} O^i \Rightarrow O^i \mathcal{R} O^{i''} \wedge O^{i''} \bar{\mathcal{R}} O^i$$

Ces propriétés permettent entre autre de définir⁴ les relations suivantes :

- relation d'ordre large : relation binaire réflexive, faiblement antisymétrique et transitive
- relation d'ordre d'intervalles : relation binaire transitive, asymétrique et ayant la propriété d'intermédiarité
- relation de quasi-ordre : relation binaire transitive, asymétrique et ayant de plus les propriétés d'intermédiarité et de cohérence

Par ailleurs, nous utiliserons la propriété d'acyclicité dans le cadre de la théorie du choix social au chapitre 6, section 6.2.

Les contraintes linéaires associées à ces propriétés relationnelles sont données ci-dessous. Soit une relation binaire \mathcal{R} représentée par sa matrice relationnelle C , nous avons :

- \mathcal{R} possède la propriété d'intermédiarité si :

$$C_{ii'} - (C_{i'i''} + C_{i''i'}) + C_{i''i'''} - C_{ii'''} \leq 1 \quad \forall i, i', i'', i''' = 1, \dots, N \quad (4.1.20)$$

- \mathcal{R} possède la propriété de cohérence si :

$$C_{ii'''} + C_{i''i'''} + C_{i''i'''} + C_{i''i'''} \geq C_{ii'} + C_{i'i''} - 1 \quad \forall i, i', i'', i''' = 1, \dots, N \quad (4.1.21)$$

- \mathcal{R} est faiblement antisymétrique si :

$$\boxed{\begin{cases} C_{ii'} + C_{i'i} + C_{i''i''} - C_{i'i''} \leq 2 & \forall i, i', i'' = 1, \dots, N \\ C_{ii'} + C_{i'i} - C_{i''i''} + C_{i'i''} \leq 2 & \forall i, i', i'' = 1, \dots, N \end{cases}} \quad (4.1.22)$$

- \mathcal{R} est acyclique si :

$$\boxed{C_{ii'} + C_{i'i''} + C_{i''i} \leq 2 \quad \forall i, i', i'', i''' = 1, \dots, N} \quad (4.1.23)$$

- \mathcal{R} est de Ferrers si :

$$\boxed{C_{ii'} + C_{i''i'''} - C_{ii'''} - C_{i''i'} \leq 1 \quad \forall i, i', i'', i''' = 1, \dots, N} \quad (4.1.24)$$

- \mathcal{R} est semi-transitive si :

$$\boxed{C_{ii'} + C_{i'i''} - C_{ii''} - C_{i''i'''} \leq 1 \quad \forall i, i', i'', i''' = 1, \dots, N} \quad (4.1.25)$$

- \mathcal{R} est quasi-transitive si :

$$\boxed{\begin{cases} C_{ii'} - C_{i'i} + C_{i''i''} - C_{i''i'} - C_{ii''} \leq 1 & \forall i, i', i'' = 1, \dots, N \\ C_{ii'} - C_{i'i} + C_{i''i''} - C_{i''i'} + C_{i''i} \leq 2 & \forall i, i', i'' = 1, \dots, N \end{cases}} \quad (4.1.26)$$

4. nous reprenons notamment les définitions de ces relations données par [MM79]

Nous montrons là aussi, que les contraintes linéaires sont obtenues à partir de la conjonction et de la disjonction de Lukasiewicz.

Démonstration Intermédierité

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge (O^{i'} \overline{\mathcal{R}}O^{i''} \wedge O^{i''} \overline{\mathcal{R}}O^i) \wedge O^{i''} \mathcal{R}O^{i'''} \Rightarrow O^i \mathcal{R}O^{i'''}\} \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i'} \overline{\mathcal{R}}O^{i''} \wedge O^{i''} \overline{\mathcal{R}}O^i \wedge O^{i'''} \mathcal{R}O^{i''''}) \leq v(O^i \mathcal{R}O^{i''''}) \\
 & \Leftrightarrow \max(\max(\max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) + \overline{C}_{i''i'} - 1, 0) + C_{i''i'''} - 1, 0) \leq C_{ii''''} \\
 & \Leftrightarrow \max(\max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) + \overline{C}_{i''i'} - 1, 0) + C_{i''i'''} - 1 \leq C_{ii''''} \text{ puisque } C_{ii''''} \geq 0 \\
 & \Leftrightarrow \max(\max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) + \overline{C}_{i''i'} - 1, 0) \leq 1 + C_{ii''''} - C_{i''i'''} \\
 & \Leftrightarrow \max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) + \overline{C}_{i''i'} - 1 \leq 1 + C_{ii''''} - C_{i''i'''} \text{ puisque } 1 + C_{ii''''} - C_{i''i'''} \geq 0 \\
 & \Leftrightarrow \max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) \leq 2 + C_{ii''''} - C_{i''i'''} - \overline{C}_{i''i'} \\
 & \Leftrightarrow \max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) \leq 2 + C_{ii''''} - C_{i''i'''} - 1 + C_{i''i'} \\
 & \Leftrightarrow \max(C_{ii'} + \overline{C}_{i'i''} - 1, 0) \leq 1 + C_{ii''''} + C_{i''i'} - C_{i''i'''} \\
 & \Leftrightarrow C_{ii'} + \overline{C}_{i'i''} - 1 \leq 1 + C_{ii''''} + C_{i''i'} - C_{i''i'''} \text{ puisque } 1 + C_{ii''''} + C_{i''i'} - C_{i''i'''} \geq 0 \\
 & \Leftrightarrow C_{ii'} + 1 - C_{i'i''} - 1 \leq 1 + C_{ii''''} + C_{i''i'} - C_{i''i'''} \\
 & \Leftrightarrow C_{ii'} - (C_{i'i''} + C_{i''i'}) + C_{i''i'''} - C_{ii''''} \leq 1
 \end{aligned}$$

□

Démonstration Cohérence

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^{i''} \Rightarrow \neg((O^i \overline{\mathcal{R}}O^{i'''} \wedge O^{i'''} \overline{\mathcal{R}}O^i) \wedge (O^{i''} \overline{\mathcal{R}}O^{i''''} \wedge O^{i''''} \overline{\mathcal{R}}O^{i''}))\} \\
 & \Leftrightarrow \{O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^{i''} \Rightarrow \neg(O^i \overline{\mathcal{R}}O^{i'''} \wedge O^{i'''} \overline{\mathcal{R}}O^i \wedge O^{i''} \overline{\mathcal{R}}O^{i''''} \wedge O^{i''''} \overline{\mathcal{R}}O^{i''})\} \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^{i''}) \leq 1 - v(O^i \overline{\mathcal{R}}O^{i'''} \wedge O^{i'''} \overline{\mathcal{R}}O^i \wedge O^{i''} \overline{\mathcal{R}}O^{i''''} \wedge O^{i''''} \overline{\mathcal{R}}O^{i''}) \\
 & \Leftrightarrow \max(C_{ii'} + C_{i'i''} - 1, 0) \leq 1 - \max(\overline{C}_{ii'''} + \overline{C}_{i''i} + \overline{C}_{i''i''} + \overline{C}_{i''i'''} - 3, 0) \\
 & \Leftrightarrow C_{ii'} + C_{i'i''} - 1 \leq 1 - \max(\overline{C}_{ii'''} + \overline{C}_{i''i} + \overline{C}_{i''i''} + \overline{C}_{i''i'''} - 3, 0) \\
 & \text{puisque } 1 - \max(\overline{C}_{ii'''} + \overline{C}_{i''i} + \overline{C}_{i''i''} + \overline{C}_{i''i'''} - 3, 0) \geq 0 \\
 & \Leftrightarrow \max(\overline{C}_{ii'''} + \overline{C}_{i''i} + \overline{C}_{i''i''} + \overline{C}_{i''i'''} - 3, 0) \leq 2 - C_{ii'} - C_{i'i''} \\
 & \Leftrightarrow \overline{C}_{ii'''} + \overline{C}_{i''i} + \overline{C}_{i''i''} + \overline{C}_{i''i'''} - 3 \leq 2 - C_{ii'} - C_{i'i''} \text{ puisque } 2 - C_{ii'} - C_{i'i''} \geq 0 \\
 & \Leftrightarrow 4 - (C_{ii'''} + C_{i''i} + C_{i''i''} + C_{i''i'''}) - 3 \leq 2 - C_{ii'} - C_{i'i''} \\
 & \Leftrightarrow -(C_{ii'''} + C_{i''i} + C_{i''i''} + C_{i''i'''}) \leq 1 - C_{ii'} - C_{i'i''} \\
 & \Leftrightarrow C_{ii'''} + C_{i''i} + C_{i''i''} + C_{i''i'''} \geq C_{ii'} + C_{i'i''} - 1
 \end{aligned}$$

□

Démonstration Antisymétrie faible

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^i \Rightarrow O^i = O^{i'}\} \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^i) \leq v(\{v(O^i \mathcal{R}O^{i''}) \Leftrightarrow v(O^{i'} \mathcal{R}O^{i''})\}) \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^i) \leq v(\{v(O^i \mathcal{R}O^{i''}) \Rightarrow v(O^{i'} \mathcal{R}O^{i''}) \wedge v(O^{i'} \mathcal{R}O^{i''}) \Leftrightarrow v(O^i \mathcal{R}O^{i''})\}) \\
 & \Leftrightarrow \max(C_{ii'} + C_{i'i} - 1, 0) \leq \max(\min(\overline{C}_{ii''} + C_{i'i''}, 1) + \min(\overline{C}_{i'i''} + C_{ii''}, 1) - 1, 0) \\
 & \Leftrightarrow C_{ii'} + C_{i'i} - 1 \leq \min(\overline{C}_{ii''} + C_{i'i''}, 1) + \min(\overline{C}_{i'i''} + C_{ii''}, 1) - 1 \\
 & \Leftrightarrow C_{ii'} + C_{i'i} \leq \min(\overline{C}_{ii''} + C_{i'i''}, 1) + \min(\overline{C}_{i'i''} + C_{ii''}, 1)
 \end{aligned}$$

Nous utilisons alors l'expression algébrique de la fonction min. Nous avons donc :

$$\begin{aligned}
 & C_{ii'} + C_{i'i} \leq \min(\overline{C}_{ii''} + C_{i'i''}, 1) + \min(\overline{C}_{i'i''} + C_{ii''}, 1) \\
 & \Leftrightarrow C_{ii'} + C_{i'i} \leq \frac{1}{2}(\overline{C}_{ii''} + C_{i'i''} + 1) - \frac{1}{2} |\overline{C}_{ii''} + C_{i'i''} - 1| + \frac{1}{2}(\overline{C}_{i'i''} + C_{ii''} + 1) - \frac{1}{2} |\overline{C}_{i'i''} + C_{ii''} - 1| \\
 & \Leftrightarrow C_{ii'} + C_{i'i} \leq \frac{1}{2}(1 + 1 + 2) - \frac{1}{2} |C_{ii''} - C_{i'i''}| - \frac{1}{2} |C_{i'i''} - C_{ii''}| \\
 & \text{comme } \overline{C}_{ii''} + C_{i'i''} = 1 \text{ par définition} \\
 & \Leftrightarrow C_{ii'} + C_{i'i} \leq 2 - |C_{ii''} - C_{i'i''}| \\
 & \Leftrightarrow \begin{cases} C_{ii'} + C_{i'i} + C_{ii''} - C_{i'i''} \leq 2 \\ C_{ii'} + C_{i'i} - C_{ii''} + C_{i'i''} \leq 2 \end{cases} \text{ comme } |C_{ii''} - C_{i'i''}| \leq A \Leftrightarrow \{C_{ii''} - C_{i'i''} \leq A\} \wedge \{-C_{ii''} + C_{i'i''} \leq A\}
 \end{aligned}$$

□

Démonstration Acyclicité

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge O^{i'} \mathcal{R}O^{i''} \Rightarrow O^{i''} \overline{\mathcal{R}}O^i\} \\
 & \Leftrightarrow \max(C_{ii'} + C_{i'i''} - 1, 0) \leq \overline{C}_{i''i} \\
 & \Leftrightarrow C_{ii'} + C_{i'i''} - 1 \leq \overline{C}_{i''i} \text{ puisque } \overline{C}_{i''i} \geq 0 \\
 & \Leftrightarrow C_{ii'} + C_{i'i''} - 1 \leq 1 - C_{i''i} \text{ puisque } \overline{C}_{i''i} = 1 - C_{i''i} \\
 & \Leftrightarrow C_{ii'} + C_{i'i''} + C_{i''i} \leq 2
 \end{aligned}$$

□

Démonstration Propriété de de Ferrers

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge O^{i''} \mathcal{R}O^{i'''} \Rightarrow O^i \mathcal{R}O^{i'''} \vee O^{i''} \mathcal{R}O^{i'}\} \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i''} \mathcal{R}O^{i'''}) \leq v(O^i \mathcal{R}O^{i'''} \vee O^{i''} \mathcal{R}O^{i'}) \\
 & \Leftrightarrow \max(C_{ii'} + C_{i''i'''} - 1, 0) \leq \min(C_{ii'''} + C_{i''i'}, 1) \\
 & \Leftrightarrow C_{ii'} + C_{i''i'''} - 1 \leq \min(C_{ii'''} + C_{i''i'}, 1) \text{ puisque } \min(C_{ii'''} + C_{i''i'}, 1) \geq 0 \\
 & \Leftrightarrow C_{ii'} + C_{i''i'''} - 1 \leq \frac{1}{2}(C_{ii'''} + C_{i''i'} + 1) - \frac{1}{2} | C_{ii'''} + C_{i''i'} - 1 | \\
 & \Leftrightarrow | C_{ii'''} + C_{i''i'} - 1 | \leq -2(C_{ii'} + C_{i''i'''}) + C_{ii'''} + C_{i''i'} + 3 \\
 & \Leftrightarrow \begin{cases} C_{ii'''} + C_{i''i'} - 1 \leq -2(C_{ii'} + C_{i''i'''}) + C_{ii'''} + C_{i''i'} + 3 \\ 2(C_{ii'} + C_{i''i'''}) - C_{ii'''} - C_{i''i'} - 3 \leq C_{ii'''} + C_{i''i'} - 1 \end{cases} \\
 & \Leftrightarrow \begin{cases} C_{ii'} + C_{i''i'''} \leq 2 \\ C_{ii'} + C_{i''i'''} - C_{ii'''} - C_{i''i'} \leq 1 \end{cases}
 \end{aligned}$$

Comme la première équation est toujours vraie, nous retenons uniquement la seconde qui donne bien le résultat annoncé.

□

Démonstration Semi-transitivité

Cette démonstration est similaire à celle de la propriété de Ferrers.

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge O^{i''} \mathcal{R}O^{i'''} \Rightarrow O^i \mathcal{R}O^{i'''} \vee O^{i''} \mathcal{R}O^{i'}\} \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i''} \mathcal{R}O^{i'''}) \leq v(O^i \mathcal{R}O^{i'''} \vee O^{i''} \mathcal{R}O^{i'}) \\
 & \Leftrightarrow \max(C_{ii'} + C_{i''i'''} - 1, 0) \leq \min(C_{ii'''} + C_{i''i'''}, 1) \\
 & \Leftrightarrow C_{ii'} + C_{i''i'''} - 1 \leq \min(C_{ii'''} + C_{i''i'''}, 1) \text{ puisque } \min(C_{ii'''} + C_{i''i'''}, 1) \geq 0 \\
 & \Leftrightarrow C_{ii'} + C_{i''i'''} - 1 \leq \frac{1}{2}(C_{ii'''} + C_{i''i'''} + 1) - \frac{1}{2} | C_{ii'''} + C_{i''i'''} - 1 | \\
 & \Leftrightarrow | C_{ii'''} + C_{i''i'''} - 1 | \leq -2(C_{ii'} + C_{i''i'''}) + C_{ii'''} + C_{i''i'''} + 3 \\
 & \Leftrightarrow \begin{cases} C_{ii'''} + C_{i''i'''} - 1 \leq -2(C_{ii'} + C_{i''i'''}) + C_{ii'''} + C_{i''i'''} + 3 \\ 2(C_{ii'} + C_{i''i'''}) - C_{ii'''} - C_{i''i'''} - 3 \leq C_{ii'''} + C_{i''i'''} - 1 \end{cases} \\
 & \Leftrightarrow \begin{cases} C_{ii'} + C_{i''i'''} \leq 2 \\ C_{ii'} + C_{i''i'''} - C_{ii'''} - C_{i''i'''} \leq 1 \end{cases}
 \end{aligned}$$

Comme la première équation est toujours vraie, nous retenons uniquement la seconde qui donne bien le résultat annoncé.

□

Démonstration Quasi-transitivité

$$\begin{aligned}
 & \{O^i \mathcal{R}O^{i'} \wedge O^{i''} \overline{\mathcal{R}}O^{i'''} \wedge O^{i'''} \mathcal{R}O^{i''} \wedge O^{i''''} \overline{\mathcal{R}}O^{i''} \Rightarrow O^i \mathcal{R}O^{i''} \wedge O^{i'''} \overline{\mathcal{R}}O^{i''}\} \\
 & \Leftrightarrow v(O^i \mathcal{R}O^{i'} \wedge O^{i''} \overline{\mathcal{R}}O^{i'''} \wedge O^{i'''} \mathcal{R}O^{i''} \wedge O^{i''''} \overline{\mathcal{R}}O^{i''}) \leq v(O^i \mathcal{R}O^{i''} \wedge O^{i'''} \overline{\mathcal{R}}O^{i''}) \\
 & \Leftrightarrow \max(C_{ii'} + (1 - C_{i''i'''}) + C_{i''i'''} + (1 - C_{i''i'''}) - 3, 0) \leq \min(C_{ii''}, 1 - C_{i''i''}) \\
 & \Leftrightarrow C_{ii'} + 1 - C_{i''i''} + C_{i''i'''} + 1 - C_{i''i'''} - 3 \leq \min(C_{ii''}, 1 - C_{i''i''}) \text{ puisque } \min(C_{ii''}, 1 - C_{i''i''}) \geq 0 \\
 & \Leftrightarrow C_{ii'} - C_{i''i''} + C_{i''i'''} - C_{i''i'''} - 1 \leq \frac{1}{2}(C_{ii''} - C_{i''i''} + 1) - \frac{1}{2} | C_{ii''} + C_{i''i''} - 1 | \\
 & \Leftrightarrow | C_{ii''} + C_{i''i''} - 1 | \leq C_{ii''} - C_{i''i''} + 1 - 2(C_{ii'} - C_{i''i''} + C_{i''i'''} - C_{i''i'''} - 1) \\
 & \Leftrightarrow \begin{cases} C_{ii''} + C_{i''i''} - 1 \leq C_{ii''} - C_{i''i''} + 1 - 2(C_{ii'} - C_{i''i''} + C_{i''i'''} - C_{i''i'''} - 1) \\ 2(C_{ii'} - C_{i''i''} + C_{i''i'''} - C_{i''i'''} - 1) - (C_{ii''} - C_{i''i''} + 1) \leq C_{ii''} + C_{i''i''} - 1 \end{cases} \\
 & \Leftrightarrow \begin{cases} C_{ii'} - C_{i''i''} + C_{i''i'''} - C_{i''i'''} + C_{i''i''} \leq 2 \quad \forall i, i', i'' = 1, \dots, N \\ C_{ii'} - C_{i''i''} + C_{i''i'''} - C_{i''i'''} - C_{ii''} \leq 1 \quad \forall i, i', i'' = 1, \dots, N \end{cases}
 \end{aligned}$$

□

Nous pensons qu'il est possible de définir de nombreuses propriétés relationnelles simples sous la forme de contraintes linéaires qui sont fonctions de variables bivalentes en appliquant la méthode suivante :

1. Formuler les propriétés des relations sous la forme de propositions combinant des propositions atomiques de type $O^i \mathcal{R}O^{i'}$ ou $O^i \overline{\mathcal{R}}O^{i'}$
2. Remplacer chaque proposition $O^i \mathcal{R}O^{i'}$ par un terme relationnel $C_{ii'}$ et $O^i \overline{\mathcal{R}}O^{i'}$ par $\overline{C}_{ii'} = 1 - C_{ii'}$

3. Remplacer les implications entre propositions atomiques par des inégalités et les équivalences par des égalités comme nous l'avons fait remarquer en (4.1.14) et (4.1.15)
4. Remplacer les conjonctions \wedge et disjonctions \vee entre propositions par les opérations de Lukasiewicz (4.1.12) et (4.1.13) ou les opérations min et max de manière à privilégier les simplifications menant à l'expression (4.1.18)
5. Simplifier les expressions :
 - en remarquant les quantités strictement positives ou strictement inférieures à 1
 - ou en remplaçant les fonctions min et max par leurs expressions algébriques (4.1.16) et (4.1.17) et en explicitant les valeurs absolues selon (4.1.18)

De la sorte, nous pouvons, à l'instar du problème de la recherche d'un consensus entre plusieurs relations binaires structurées, linéariser les contraintes d'un problème qui s'expriment par des propositions logiques. Si de plus le problème fait apparaître un critère linéaire en l'inconnu, nous pouvons alors résoudre celui-ci par programmation linéaire. Nous rappelons à ce propos les travaux de P. Michaud issus de [Mic90] dans lesquels ce dernier montre que certaines formes propositionnelles se mettent sous la forme de contraintes linéaires. Il fait ainsi le lien entre l'analyse logique par des systèmes experts et l'analyse numérique par l'utilisation de programmes linéaires en nombres bivalents.

Nous avons introduit l'algèbre booléenne des matrices relationnelles. Les différentes opérations de cette section nous permettent de définir une méthode d'agrégation basée sur des processus majoritaires. Nous exposons cette approche dans la section suivante.

4.2 Les enveloppes de Condorcet

Les enveloppes de Condorcet que nous présentons dans cette section entrent dans le cadre du problème de l'agrégation de plusieurs relations binaires. Il s'agit d'une approche qui mêle combinatoire et règle de vote avec la notion d'une majorité m donnée.

Cette problématique qui initialement concernait des données binaires fut relevée dans un premier temps par J.F. Marcotorchino dans [Mar85] puis, étudiée dans un second temps par S. Ghashghaie dans [Gha90]. Bien que le principe ait été bien défini, aucune formulation explicite n'a été cependant donnée jusqu'à présent. Nous montrons que le processus majoritaire que définit l'enveloppe de Condorcet aboutit en fait aux formules combinatoires de Ch. Jordan [Com70].

Nous commençons par définir plus formellement ce que nous entendons par l'enveloppe de Condorcet d'un ensemble de M relations à une majorité m donnée.

Puis nous montrons comment nous pouvons déterminer cette enveloppe par une approche combinatoire.

Enfin, nous étendons ces définitions à des données non binaires. **Nous appliquons cette méthode à la modélisation du problème de classification de données hétérogènes dans la mesure où l'approche proposée permet de combiner des informations de nature distincte tout en permettant de garder les spécificités de chaque type d'information.**

Dans le cadre de cette section, nous nous plaçons toujours dans un contexte où les matrices relationnelles sont binaires et pour lesquelles les termes sont interprétés d'un point de vue logique. Nous rappelons alors ci-dessous la signification des notations utilisées :

Remarque 4.2.1 *Signification des termes $C_{ii'}^k$ et des symboles \oplus et \otimes dans le cadre de la section 4.2*

Dans le cadre de cette section, nous aurons de la même manière que pour la section 4.1, les définitions

suivantes :

- $C_{ii'}^k = v(O^i \mathcal{R}^k O^{i'}) =$ valeur de vérité binaire de la proposition “ O^i est en relation avec $O^{i'}$ selon la relation \mathcal{R}^k ”
- $[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = v(O^i \mathcal{R}^k O^{i'} \vee O^i \mathcal{R}^l O^{i'}) =$ valeur de vérité binaire de la disjonction “ O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R}^k ou la relation \mathcal{R}^l ”
- $[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = v(O^i \mathcal{R}^k O^{i'} \wedge O^i \mathcal{R}^l O^{i'}) =$ valeur de vérité binaire de la conjonction “ O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R}^k et la relation \mathcal{R}^l ”

Pour fixer les idées, nous pourrions prendre comme précédemment les expressions particulières suivantes pour l'intersection \otimes et la disjonction \oplus :

$$[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = \min(C_{ii'}^k, C_{ii'}^l) \quad (4.2.1)$$

$$[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = \max(C_{ii'}^k, C_{ii'}^l) \quad (4.2.2)$$

4.2.1 Définitions et propriétés de l'enveloppe de Condorcet

Soient M relations binaires structurées \mathcal{R}^k , $k = 1, \dots, M$ sur un ensemble de N objets. Soit m un entier compris entre 0 et M qui représente la notion de majorité. Dans ce qui suit, nous noterons C^k la matrice relationnelle individuelle, représentative de la relation binaire \mathcal{R}^k et par $\mathbb{C} = \sum_{k=1}^M C^k$, la matrice relationnelle collective. Nous définissons l'enveloppe de Condorcet de la manière suivante :

Définition 4.2.1 *Enveloppe de Condorcet d'un ensemble de M relations binaires structurées relative à une majorité m donnée*

L'enveloppe de Condorcet de M relations binaires relative à une majorité m (nous dirons aussi d'ordre m) est une matrice binaire et carrée de taille N , notée E_m^M , dont le terme général $[E_m^M]_{ii'}$, $\forall i, i' = 1, \dots, N$ est tel que :

$$[E_m^M]_{ii'} = \begin{cases} 1 & \text{s'il existe au moins } m \text{ relations qui indiquent que } O^i \text{ est} \\ & \text{en relation avec } O^{i'} \text{ ou encore si } C_{ii'} \geq m \\ 0 & \text{sinon} \end{cases} \quad (4.2.3)$$

L'enveloppe de Condorcet est une matrice relationnelle qui n'hérite pas obligatoirement des propriétés des matrices relationnelles qu'elle agrège. Ainsi, l'enveloppe de Condorcet d'un ensemble de partitions n'est pas en général une partition. La définition logique ci-dessous nous permet dans un premier temps d'éclairer ces propos puis, dans un second temps, d'établir certaines propriétés.

Propriété 4.2.1 *Définition logique de l'enveloppe de Condorcet*

Soit l'algèbre de Boole $\mathcal{B}(\mathbb{C}) = \{\mathbb{C}, -, \otimes, \oplus, U_N, \bar{U}_N\}$, introduite en prop-4.1.1. L'enveloppe de Condorcet de M matrices relationnelles d'ordre m peut être définie de la manière suivante :

$$E_m^M = \bigoplus_{1 \leq i_1 < i_2 < \dots < i_m \leq M} (C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_m}) \quad (4.2.4)$$

Par convention, dans le cas où $m < 1$, nous prendrons $E_m^M = U_N$ et dans le cas où $m > M$, nous prendrons $E_m^M = \bar{U}_N$.

Démonstration

Il suffit de voir que $[\bigoplus_{1 \leq i_1 < i_2 < \dots < i_m \leq M} C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_m}]_{ii'}$ vaut 1 s'il y a au moins un m -uplet (i_1, i_2, \dots, i_m) tel que l'intersection $[C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_m}]_{ii'}$ vaut 1. Or ceci indique précisément qu'il y a au moins m matrices relationnelles dont les termes (i, i') valent 1.

□

La définition logique de l'enveloppe de Condorcet fait intervenir des disjonctions. Or nous avons vu dans le paragraphe 4.1.1.2 que la propriété de transitivité n'était pas stable sous l'action de cette opération. Par conséquent si les éléments C^k sont des matrices relationnelles transitives, telles que les relations d'équivalence ou d'ordre, alors l'enveloppe ne sera pas, en général, transitive.

Il existe une méthode simple permettant de calculer l'enveloppe de Condorcet d'ordre m à partir de M relations binaires. Celle-ci procède de la manière suivante :

1. Calcul des matrices relationnelles C^k associées à chaque relation binaire \mathcal{R}^k .
2. Calcul de la matrice relationnelle collective $\mathbf{C} = \sum_{k=1}^M C^k$
3. **Ecrétage** de la matrice relationnelle collective au niveau m donné

Nous donnons ci-dessous un exemple afin d'illustrer la problématique. Considérons alors les 4 matrices relationnelles de partitions suivantes :

$$C^1 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & ; & C^2 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad (4.2.5)$$

$$C^3 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & ; & C^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (4.2.6)$$

La matrice relationnelle collective $\mathbf{C} = C^1 + C^2 + C^3 + C^4$ vaut alors :

$$\mathbf{C} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 4 & 3 & 1 & 1 \\ 3 & 4 & 2 & 0 \\ 1 & 2 & 4 & 1 \\ 1 & 0 & 1 & 4 \end{pmatrix} \end{matrix} \quad (4.2.7)$$

A partir de \mathbf{C} , nous pouvons ainsi déterminer par écrétage les enveloppes de Condorcet suivantes :

$$E_4^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & ; & E_3^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (4.2.8)$$

$$E_2^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & ; & E_1^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad (4.2.9)$$

Malgré cette approche simple, nous donnons dans la sous-section 4.2.2, une méthode combinatoire afin de déterminer l'enveloppe de Condorcet. En effet, mis à part l'aspect théorique de cette approche, nous chercherons à lui donner des applications dans la sous-section 4.3 où nous étendrons notamment les définitions des enveloppes de Condorcet à des données non binaires. Les formules combinatoires alors établies permettent de définir une nouvelle méthode d'agrégation de données.

Pour l'heure, nous donnons dans le paragraphe suivant, des propriétés de l'enveloppe de Condorcet où nous continuons pour l'instant, à étudier des matrices relationnelles binaires.

4.2.1.1 Complémentaires d'une enveloppe de Condorcet

Compte tenu des propriétés de l'algèbre de Boole, nous avons plusieurs écritures de l'enveloppe de Condorcet. Celles-ci font intervenir deux types de complémentaires que nous introduisons ci-dessous.

Définition 4.2.2 Complémentaire de type 1 de l'enveloppe de Condorcet

Le complémentaire de type 1 d'une enveloppe de Condorcet E_m^M , est noté \overline{E}_m^M et est donné par la relation suivante :

$$\overline{E}_m^M = U_N - E_m^M \quad (4.2.10)$$

En d'autres termes, nous avons $\forall i, i' = 1, \dots, N$:

$$\boxed{[\overline{E}_m^M]_{ii'} = \begin{cases} 1 & \text{s'il existe au plus } m-1 \text{ relations qui indiquent que } O^i \text{ est en} \\ & \text{relation avec } O^{i'} \text{ ou encore si } C_{ii'} \leq m-1 \\ 0 & \text{sinon} \end{cases}} \quad (4.2.11)$$

Nous pouvons aussi définir le complémentaire de type 1 de la manière suivante :

$$\overline{E}_m^M = \bigotimes_{1 \leq i_1 < i_2 < \dots < i_m \leq M} (C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_m}) \quad (4.2.12)$$

Définition 4.2.3 Complémentaire de type 2 de l'enveloppe de Condorcet

Le complémentaire de type 2 d'une enveloppe de Condorcet E_m^M , sera noté $\overline{\overline{E}}_m^M$ et sera donné par la relation suivante :

$$\overline{\overline{E}}_m^M = \bigoplus_{1 \leq i_1 < i_2 < \dots < i_m \leq M} (\overline{C}^{i_1} \otimes \overline{C}^{i_2} \otimes \dots \otimes \overline{C}^{i_m}) \quad (4.2.13)$$

En d'autres termes, nous avons $\forall i, i' = 1, \dots, N$:

$$\boxed{[\overline{\overline{E}}_m^M]_{ii'} = \begin{cases} 1 & \text{s'il existe au moins } m \text{ relations qui indiquent que } O^i \text{ n'est pas} \\ & \text{en relation avec } O^{i'} \text{ ou encore si } \overline{C}_{ii'} \geq m \\ 0 & \text{sinon} \end{cases}} \quad (4.2.14)$$

Suite à ces définitions, nous avons les propriétés suivantes.

Propriété 4.2.2 Relations entre complémentaire de type 1 et complémentaire de type 2

Nous avons la relation suivante entre les deux types de complémentarité d'une enveloppe :

$$\boxed{\overline{\overline{E}}_m^M = \overline{E}_{M-m+1}^M} \quad (4.2.15)$$

Démonstration

$$\begin{aligned}
 [\overline{E}_m^M]_{ii'} = 1 &\Leftrightarrow \text{il existe au moins } m \text{ relations qui indiquent que } O^i \text{ n'est pas en relation avec } O^{i'} \\
 &\Leftrightarrow \mathbf{C}_{ii'} \geq m \\
 &\Leftrightarrow M - \mathbf{C}_{ii'} \geq m \\
 &\Leftrightarrow \mathbf{C}_{ii'} \leq M - m \\
 &\Leftrightarrow [\overline{E}_{M-m+1}^M]_{ii'} = 1
 \end{aligned}$$

□

Propriété 4.2.3 Formulation de E_m^M en fonction des deux types de complémentaires

Une conséquence de la propriété prop-4.2.2 est la définition de l'enveloppe de Condorcet suivante :

$$\boxed{E_m^M = \overline{\overline{E}_{M-m+1}^M}} \quad (4.2.16)$$

Démonstration

 Selon (4.2.15), nous avons, en remplaçant m par $M - m + 1$: $\overline{E}_{M-m+1}^M = \overline{\overline{E}_m^M}$. Nous en déduisons donc que $\overline{\overline{E}_{M-m+1}^M} = E_m^M$.

□

4.2.1.2 Décomposition d'une enveloppe de Condorcet selon la majorité m

Les enveloppes de Condorcet définies en termes des opérations de l'algèbre de Boole peuvent être interprétées comme étant des ensembles dont les éléments sont des paires d'objets. La conjonction \otimes et la disjonction \oplus sont respectivement les fonctions caractéristiques de l'intersection et de l'union des deux ensembles C^k et C^l . Nous avons alors les propriétés suivantes :

Propriété 4.2.4 Relations d'inclusions entre des enveloppes d'ordre distincts

 Soit $\{E_m^M; m = 0, \dots, M\}$ l'ensemble des enveloppes de Condorcet définies sur un même ensemble de M relations binaires. Nous avons les propriétés suivantes :

$$E_m^M \subseteq E_{m'}^M \quad \forall m' \leq m \quad (4.2.17)$$

ou encore

$$[E_m^M]_{ii'} \leq [E_{m'}^M]_{ii'} \quad \forall m' \leq m \text{ et } \forall i, i' = 1, \dots, N \quad (4.2.18)$$

Démonstration

$$\begin{aligned}
 [E_m^M]_{ii'} = 1 &\Leftrightarrow \mathbf{C}_{ii'} \geq m \\
 &\Rightarrow \mathbf{C}_{ii'} \geq m' \quad \forall m' \leq m \\
 &\Rightarrow [E_{m'}^M]_{ii'} = 1 \quad \forall m' \leq m
 \end{aligned}$$

□

Nous pouvons définir par ailleurs, les matrices relationnelles suivantes :

Définition 4.2.4 Enveloppe de Condorcet partielle d'ordre m

 Soient $\{E_m^M; m = 0, \dots, M\}$ l'ensemble des enveloppes de Condorcet d'ordre m d'un ensemble de M relations binaires. Nous définissons H_m^M , $m = 0, \dots, M$ l'enveloppe partielle d'ordre m par :

$$H_m^M = E_m^M - E_{m+1}^M \quad \forall m = 0, \dots, M \quad (4.2.19)$$

Le terme général de H_m^M noté $[H_m^M]_{ii'}$, $\forall i, i' = 1, \dots, N$; a pour signification :

$$[H_m^M]_{ii'} = \begin{cases} 1 & \text{s'il y a exactement } m \text{ relations qui indiquent que} \\ & O^i \text{ et } O^{i'} \text{ sont en relation} \\ 0 & \text{sinon} \end{cases} \quad (4.2.20)$$

Ces enveloppes partielles vérifient certaines propriétés :

Propriété 4.2.5 Propriétés des enveloppes partielles

- $H_m^M \otimes H_{m'}^M = \bar{U}_N \quad \forall m, m' = 0, \dots, M : m \neq m'$
- $E_m^M = \sum_{k=m}^M H_k^M$

Démonstration

La première propriété exprime le fait qu'une paire $(O^i, O^{i'})$ a un unique nombre de relations qui supportent la relation entre ces deux objets. Ainsi, en termes ensemblistes, les enveloppes partielles d'ordres distincts sont d'intersections nulles.

La deuxième propriété s'obtient facilement en développant d'une part la somme en prenant $E_k^M - E_{k+1}^M$ pour H_k^M et en observant d'autre part que les termes s'annulent consécutivement deux à deux pour ne laisser que E_m^M .

□

Voici pour illustration, une enveloppe partielle concernant l'exemple présenté page 126.

$$\begin{aligned} H_2^4 = E_2^4 - E_3^4 &= \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \\ O^2 & \begin{pmatrix} 1 & 1 & 1 & 0 \end{pmatrix} \\ O^3 & \begin{pmatrix} 0 & 1 & 1 & 0 \end{pmatrix} \\ O^4 & \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} - \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \\ O^2 & \begin{pmatrix} 1 & 1 & 0 & 0 \end{pmatrix} \\ O^3 & \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\ O^4 & \begin{pmatrix} 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \\ &= \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \\ O^2 & \begin{pmatrix} 0 & 0 & 1 & 0 \end{pmatrix} \\ O^3 & \begin{pmatrix} 0 & 1 & 0 & 0 \end{pmatrix} \\ O^4 & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \end{aligned} \quad (4.2.21)$$

Suite à ces définitions, nous donnons des propriétés de décomposition de l'enveloppe de Condorcet selon la majorité m lorsque nous distinguons deux groupes disjoints de relations binaires.

Propriété 4.2.6 Décomposition booléenne d'une enveloppe de Condorcet selon la majorité

Soit un ensemble de M relations binaires que nous séparons en deux groupes distincts l'un composé de p_1 relations et l'autre composé des p_2 relations complémentaires avec $p_2 = M - p_1$. Nous pouvons alors décomposer l'enveloppe E_m^M selon la formule suivante :

$$E_m^M = \bigoplus_{q=0}^{p_2} (E_{m-q}^{p_1} \otimes E_q^{p_2}) \quad (4.2.22)$$

Nous rappelons que par convention, nous avons $E_m^M = U_N$, $\forall m < 1$ et $E_m^M = \bar{U}_N$, $\forall m > M$.

Démonstration

$[E_{m-q}^{p_1} \otimes E_q^{p_2}]_{ii'} = 1$ signifie que, pour le cas particulier où dans le premier groupe nous retenons uniquement une majorité partielle de $m - q$ relations et pour le deuxième groupe une majorité partielle de q , il y a au moins m relations sur l'ensemble des $p_1 + p_2 = M$, qui supportent " O^i est en relation avec $O^{i'}$ ". L'intersection permet ainsi de compléter mutuellement les deux ensembles de relations afin d'aboutir pour l'ensemble à une majorité globale égale à m .

A une valeur q donnée correspond une configuration particulière de la décomposition de m selon $m - q$ d'une part et q d'autre part. De ce fait, la somme booléenne sur l'ensemble des q permet de parcourir l'ensemble des configurations possibles entre les deux groupes conduisant à une majorité globale égale à m . Par conséquent, $[\bigoplus_{q=0}^{p_2} E_{m-q}^{p_1} \otimes E_q^{p_2}]_{ii'} = 1$ signifie bien qu'il y a au moins m relations qui indiquent que " O^i est en relation avec $O^{i'}$ ".

□

Nous pouvons dans (4.2.22) restreindre la sommation à $0 \leq q \leq \min(m, p_2)$. Toutefois, il est utile de garder la sommation jusqu'à p_2 notamment pour démontrer le résultat suivant où nous exprimons la propriété de décomposition avec une sommation arithmétique⁵ et faisant intervenir uniquement les conjonctions \otimes .

Propriété 4.2.7 Décomposition arithmétique d'une enveloppe de Condorcet selon les enveloppes partielles

Nous nous plaçons dans les mêmes conditions que pour la prop-4.2.6. Nous avons alors la décomposition arithmétique de l'enveloppe E_m^M suivante :

$$E_m^M = \sum_{q=0}^{p_2} (E_{m-q}^{p_1} \otimes H_q^{p_2}) \quad (4.2.23)$$

Démonstration

En posant $E_{m-q}^{p_1} = \sum_{k_1=m-q}^{p_1} H_{k_1}^{p_1}$ et $E_q^{p_2} = \sum_{k_2=q}^{p_2} H_{k_2}^{p_2}$ dans (4.2.22), nous avons :

$$\begin{aligned} E_m^M &= \bigoplus_{q=0}^{p_2} \left(\sum_{k_1=m-q}^{p_1} H_{k_1}^{p_1} \otimes \sum_{k_2=q}^{p_2} H_{k_2}^{p_2} \right) \\ &= \bigoplus_{q=0}^{p_2} \left(\sum_{k_1=m-q}^{p_1} \sum_{k_2=q}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2}) \right) \end{aligned}$$

Enumérons les sommations pour les différents q en commençant par $q = p_2$ et en remontant jusqu'à $q = 0$. Nous avons :

q	$\sum_{k_1=m-q}^{p_1} \sum_{k_2=q}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$
p_2	$\sum_{k_1=m-p_2}^{p_1} \sum_{k_2=p_2}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$
$p_2 - 1$	$\sum_{k_1=m-p_2+1}^{p_1} \sum_{k_2=p_2-1}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$
$p_2 - 2$	$\sum_{k_1=m-p_2+2}^{p_1} \sum_{k_2=p_2-2}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$
\vdots	\vdots
2	$\sum_{k_1=m-2}^{p_1} \sum_{k_2=2}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$
1	$\sum_{k_1=m-1}^{p_1} \sum_{k_2=1}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$
0	$\sum_{k_1=m}^{p_1} \sum_{k_2=0}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$

Il y a des redondances dans les sommations pour différents q ce qui explique la nécessité d'une somme booléenne afin de maintenir la binarité. Par exemple la sommation pour $q = p_2 - 1$ est en partie incluse dans la sommation pour $q = p_2$: c'est le cas pour $k_1 \geq m - p_2 + 1 \wedge k_2 \geq p_2$; ainsi seuls les éléments correspondants aux régions d'indices suivantes ne sont pas inclus, $k_1 \geq m - p_2 + 1 \wedge k_2 = p_2 - 1$. Ces répétitions sont dues au fait que les sommations sur k_2 sont incluses les unes dans les autres selon que la valeur de q décroît. Si nous évitons ces redondances nous pouvons supprimer l'addition booléenne dans la décomposition pour la remplacer par une addition arithmétique.

Nous obtenons alors le tableau suivant :

5. ie non booléenne

q	$\sum_{k_1=m-q}^{p_1} \sum_{k_2=q}^{p_2} (H_{k_1}^{p_1} \otimes H_{k_2}^{p_2})$	En fonction des E_m^M
p_2	$\sum_{k_1=m-p_2}^{p_1} H_{k_1}^{p_1} \otimes H_{p_2}^{p_2}$	$E_{m-p_2}^{p_1} \otimes H_{p_2}^{p_2}$
$p_2 - 1$	$\sum_{k_1=m-p_2+1}^{p_1} H_{k_1}^{p_1} \otimes H_{p_2-1}^{p_2}$	$E_{m-p_2+1}^{p_1} \otimes H_{p_2-1}^{p_2}$
$p_2 - 2$	$\sum_{k_1=m-p_2+2}^{p_1} H_{k_1}^{p_1} \otimes H_{p_2-2}^{p_2}$	$E_{m-p_2+2}^{p_1} \otimes H_{p_2-2}^{p_2}$
\vdots	\vdots	\vdots
2	$\sum_{k_1=m-2}^{p_1} H_{k_1}^{p_1} \otimes H_2^{p_2}$	$E_{m-2}^{p_1} \otimes H_2^{p_2}$
1	$\sum_{k_1=m-1}^{p_1} H_{k_1}^{p_1} \otimes H_1^{p_2}$	$E_{m-1}^{p_1} \otimes H_1^{p_2}$
0	$\sum_{k_1=m}^{p_1} H_{k_1}^{p_1} \otimes H_0^{p_2}$	$E_m^{p_1} \otimes H_0^{p_2}$

Le résultat se simplifiant selon la troisième colonne, nous avons bien $E_m^M = \sum_{q=0}^{p_2} E_{m-q}^{p_1} \otimes H_q^{p_2}$.

□

Nous illustrons cette propriété en reprenant notre exemple donné page 126. Nous prenons $p_1 = 1$ et $p_2 = 3$ avec comme sous-ensembles correspondants $\{C^1\}$ et $\{C^2, C^3, C^4\}$. Nous allons montrer que nous avons bien : $E_2^4 = E_2^1 \otimes H_0^3 + E_1^1 \otimes H_1^3 + E_0^1 \otimes H_2^3 + E_{-1}^1 \otimes H_3^3$. Compte tenu des conventions, nous avons $E_2^1 = \bar{U}_N$ et $E_0^1 = E_{-1}^1 = U_N$. Nous pouvons donc d'ores et déjà simplifier l'équation en $E_2^4 = E_1^1 \otimes H_1^3 + H_2^3 + H_3^3$. Nous donnons ci-dessous les valeurs de H_2^3 et H_3^3 sachant que E_1^1 se ramène à C^1 donné page 126 :

$$H_1^3 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix} ; \quad H_2^3 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix} \quad (4.2.24)$$

$$H_3^3 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

Nous avons alors :

$$E_1^1 \otimes H_1^3 + H_2^3 + H_3^3 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} = E_2^4 \quad (4.2.25)$$

Nous faisons remarquer que le résultat peut être étendu à un nombre de groupes disjoints quelconque puisque l'enveloppe $E_{m-q}^{p_1}$ peut être à nouveau décomposé en deux sous-groupes et ainsi de suite.

Nous verrons ultérieurement plus en détail l'utilité de mettre en évidence la propriété de décomposition d'une enveloppe de Condorcet. En effet, celle-ci peut être vue comme étant une fonction d'agrégation de données floues et dans ce cas, la propriété de décomposition correspond à une propriété de **quasi-associativité** qui est nécessaire pour son utilisation pratique.

4.2.2 Calcul de l'enveloppe de Condorcet par une approche combinatoire

4.2.2.1 Notations

Nous introduisons des notations qui permettront de simplifier l'écriture des enveloppes. Celles-ci sont reprises de L. Comtet [Com70].

Définition 4.2.5 Somme d'intersections relationnelles d'ordre k

Soient C^k , $k = 1, \dots, M$, des matrices relationnelles munies de la loi interne d'intersection relationnelle \otimes introduite en déf-4.1.1.

Pour tout k tel que $1 \leq k \leq M$, nous notons par S_k^M , la somme des matrices relationnelles issues de toutes les intersections relationnelles entre les $\binom{M}{k}$ combinaisons possibles de k matrices relationnelles parmi M au sens du produit terme à terme.

$$S_k^M = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq M} (C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_k}) \quad (4.2.26)$$

S_k^M est une matrice carrée de taille N non nécessairement binaire de terme général $[S_k^M]_{ii'}$, $\forall i, i' = 1, \dots, N$.

L'ensemble des k -uplets (i_1, i_2, \dots, i_k) tel que $1 \leq i_1 < i_2 < \dots < i_k \leq M$ est de cardinal $\binom{M}{k}$. Ses éléments sont naturellement ordonnés. Nous noterons par $[S_k^M]^j$, la matrice relationnelle dont le k -uplet associé est le j -ème élément ordonné de l'ensemble des k -uplets.

$[S_k^M]^j$ est une matrice relationnelle d'une partition dont le terme général est noté $[S_k^M]_{ii'}^j$, $\forall i, i' = 1, \dots, N$.

Par convention, dans le cas où $k < 1$, nous prendrons $S_k^M = U_N$ et dans le cas où $k > M$, nous prendrons $S_k^M = \bar{U}_N$.

Pour fixer les idées, nous donnons un exemple constitué de 4 matrices relationnelles C^1, C^2, C^3, C^4 . Nous avons dans ce cas, S_2^4 qui vaut $C^1 \otimes C^2 + C^1 \otimes C^3 + C^1 \otimes C^4 + C^2 \otimes C^3 + C^2 \otimes C^4 + C^3 \otimes C^4$. S_2^4 est bien la somme de $\binom{4}{2} = 6$ termes et nous avons par exemple $[S_2^4]^3 = C^1 \otimes C^4$ et $[S_2^4]^6 = C^3 \otimes C^4$.

Si nous reprenons l'exemple présenté page 126, nous avons alors :

$$S_2^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 6 & 3 & 0 & 0 \\ 3 & 6 & 1 & 0 \\ 0 & 1 & 6 & 0 \\ 0 & 0 & 0 & 6 \end{pmatrix} & ; & S_3^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 4 & 1 & 0 & 0 \\ 1 & 4 & 0 & 0 \\ 0 & 0 & 4 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix} \end{matrix} \quad (4.2.27)$$

$$S_4^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

4.2.2.2 Calcul de l'enveloppe de Condorcet par inversion du triangle de Pascal

Nous cherchons à présent à déterminer l'enveloppe de Condorcet par une approche combinatoire. Nous montrons que celle-ci consiste finalement à inverser le triangle de Pascal.

Afin de comprendre la démarche suivie, nous présentons un exemple introductif.

Supposons que nous voulions calculer E_2^4 , l'enveloppe de Condorcet calculée sur un ensemble de $M = 4$ relations à une majorité $m = 2$. Selon déf-4.2.1, E_2^4 est une matrice carrée binaire de taille N qui identifie l'ensemble des paires d'objets $(O^i, O^{i'})$ pour lesquels il y a au moins 2 relations sur 4 qui indique que O^i est en relation avec $O^{i'}$.

Si pour la paire $(O^i, O^{i'})$, il y a exactement 2 relations pour lesquelles ces deux objets sont en relation alors il existe une unique intersection relationnelle $C_{ii'}^k C_{ii'}^{k'}$ qui vaut 1, avec $k, k' = 1, \dots, M$ et $k < k'$. Considérons l'équation suivante :

$$S_2^4 = C^1 \otimes C^2 + C^1 \otimes C^3 + C^1 \otimes C^4 + C^2 \otimes C^3 + C^2 \otimes C^4 + C^3 \otimes C^4$$

Nous avons les différents cas suivants :

- $[S_2^4]_{ii'} = 0$ alors il y a moins de 2 relations qui indiquent que O^i est en relation avec $O^{i'}$
- $[S_2^4]_{ii'} = 1$ alors il y a exactement 2 relations qui indiquent que O^i est en relation avec $O^{i'}$
- $[S_2^4]_{ii'} > 1$ alors il y a plus de 2 relations qui indiquent que O^i est en relation avec $O^{i'}$. Plus précisément, nous avons :
 - $[S_2^4]_{ii'} = \binom{3}{2} = 3$, s'il y a exactement 3 relations qui indiquent que O^i est en relation avec $O^{i'}$
 - $[S_2^4]_{ii'} = \binom{4}{2} = 6$, s'il y a exactement 4 relations qui indiquent que O^i est en relation avec $O^{i'}$

De la même manière, si pour la paire $(O^i, O^{i'})$, il y a exactement 3 relations pour lesquelles ces deux objets sont en relation alors il existe une unique intersection relationnelle $C_{ii'}^k, C_{ii'}^{k'}, C_{ii'}^{k''}$ qui vaut 1, avec $k, k', k'' = 1, \dots, M$ et $k < k' < k''$. Dans ce cas, nous considérons l'équation suivante :

$$S_3^4 = C^1 \otimes C^2 \otimes C^3 + C^1 \otimes C^2 \otimes C^4 + C^1 \otimes C^3 \otimes C^4 + C^2 \otimes C^3 \otimes C^4$$

Nous avons les différents cas suivants :

- $[S_3^4]_{ii'} = 0$ alors il y a moins de 3 relations qui indiquent que O^i est en relation avec $O^{i'}$
- $[S_3^4]_{ii'} = 1$ alors il y a exactement 3 relations qui indiquent que O^i est en relation avec $O^{i'}$
- $[S_3^4]_{ii'} > 1$ alors il y a plus de 3 relations qui indiquent que O^i est en relation avec $O^{i'}$. Plus précisément, nous avons $[S_3^4]_{ii'} = \binom{4}{3} = 4$, s'il y a exactement 4 relations qui indiquent que O^i est en relation avec $O^{i'}$

Enfin, si pour la paire $(O^i, O^{i'})$, il y a exactement 4 relations pour lesquelles ces deux objets sont en relation alors nous avons $[S_4^4]_{ii'} = C_{ii'}^1 \otimes C_{ii'}^2 \otimes C_{ii'}^3 \otimes C_{ii'}^4 = 1$. Dans le cas où il y a moins de 4 relations, nous avons $[S_4^4]_{ii'} = 0$.

L'enveloppe de Condorcet est une matrice relationnelle dans les termes sont soit 1, soit 0. Or, dans l'ensemble des sommes S_k^M que nous venons d'expliquer dans le cas $M = 4$ et $m = 2$, il nous faut combiner ces termes afin d'arriver à des valeurs binaires.

1. s'il y a exactement 2 relations qui indiquent que O^i est en relation avec $O^{i'}$, les paires (i, i') qui vérifient cette condition sont identifiées par l'unique somme S_2^4 . De plus, il s'agit justement des termes unitaires de cette matrice somme : S_2^4 permet déjà d'affecter 1 aux termes (i, i') pour lesquels il y a 2 relations qui supportent " O^i est en relation avec $O^{i'}$ "
2. les paires $(O^i, O^{i'})$ pour lesquels il y a exactement 3 relations qui indiquent que O^i est en relation avec $O^{i'}$, sont identifiées par :
 - les termes de S_2^4 qui valent $\binom{3}{2} = 3$
 - les termes de S_3^4 qui valent $\binom{3}{3} = 1$
 A ce stade, nous voyons que la combinaison linéaire suivante $S_2^4 - 2S_3^4$ permet d'affecter 1 aux termes (i, i') pour lesquels soit il y a 2 ou 3 relations qui indiquent que O^i est en relation avec $O^{i'}$. En effet, pour les paires supportés par 3 relations, nous avons $[S_2^4]_{ii'} - 2[S_3^4]_{ii'} = 3 - 2 \times 1 = 1$.
3. enfin, les paires $(O^i, O^{i'})$ pour lesquels il y a exactement 4 relations qui indiquent que O^i est en relation avec $O^{i'}$, sont identifiées par :
 - les termes de S_2^4 qui valent $\binom{4}{2} = 6$
 - les termes de S_3^4 qui valent $\binom{4}{3} = 4$
 - les termes de S_4^4 qui valent $\binom{4}{4} = 1$
 Nous avons à l'étape précédente considéré la combinaison linéaire $S_2^4 - 2S_3^4$. Pour cette matrice, les termes (i, i') supportés par 4 relations ont pour valeur $[S_2^4]_{ii'} - 2[S_3^4]_{ii'} =$

$6 - 2 \times 4 = -2$. En complétant par la dernière somme S_4^4 , nous voyons finalement que la combinaison linéaire $S_2^4 - 2S_3^4 + 3S_4^4$ permet à nouveau de ramener à 1 la valeur des termes correspondants aux paires supportées par 4 relations. Au final, les termes (i, i') de $S_2^4 - 2S_3^4 + 3S_4^4$ qui valent 1 sont les paires pour lesquels, il y a au moins 2 relations qui indiquent que $(O^i, O^{i'})$ sont en relation. Par ailleurs les termes nuls sont les paires supportées par strictement moins de 2 relations. Nous avons donc une expression arithmétique de E_2^4 qui est la suivante :

$$E_2^4 = S_2^4 - 2S_3^4 + 3S_4^4$$

Notons par $\alpha_k^{m,M}$ avec $m \leq k \leq M$ le coefficient de la combinaison linéaire associé à la somme S_k^M . Nous avons dans l'exemple ci-dessus $\alpha_2^{2,4} = 1$, $\alpha_3^{2,4} = -2$ et $\alpha_4^{2,4} = 3$. Formellement, il est facile de voir que la détermination de ces coefficients est obtenue à partir de la relation suivante :

$$\begin{pmatrix} \binom{2}{2} & 0 & 0 \\ \binom{3}{3} & \binom{3}{3} & 0 \\ \binom{4}{2} & \binom{4}{3} & \binom{4}{4} \end{pmatrix} \cdot \begin{pmatrix} \alpha_2^{2,4} \\ \alpha_3^{2,4} \\ \alpha_4^{2,4} \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \quad (4.2.28)$$

Pour être plus concret, reprenons l'exemple constitué de quatre matrices relationnelles C^1, C^2, C^3 et C^4 données page 126.

Nous laissons alors le soin au lecteur de vérifier que sur l'exemple donné page 126 et page 132, nous avons bien :

$$S_2^4 - 2S_3^4 + 3S_4^4 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} & = & E_2^4 \end{matrix} \quad (4.2.29)$$

De manière générale, nous remarquons que la valeur d'un terme (i, i') de S_k^M correspondant à une paire d'objets supportée par exactement $k + j$ relations, avec $j \geq 0$, est $\binom{k+j}{k}$. En considérant la démarche décrite dans l'exemple précédent, nous aboutissons à la formalisation du problème général qui consiste en la résolution du système suivant :

$$\boxed{\mathbf{P}^{m,M} \cdot \mathbf{A}^{m,M} = U_{M-m+1,1}} \quad (4.2.30)$$

où

$$\mathbf{P}^{m,M} = \begin{pmatrix} \binom{m}{m} & 0 & 0 & 0 & 0 & \dots & 0 \\ \binom{m+1}{m} & \binom{m+1}{m+1} & 0 & 0 & 0 & \dots & 0 \\ \binom{m+2}{m} & \binom{m+2}{m+1} & \binom{m+2}{m+2} & 0 & 0 & \dots & 0 \\ \binom{m+3}{m} & \binom{m+3}{m+1} & \binom{m+3}{m+2} & \binom{m+3}{m+3} & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \binom{M}{m} & \binom{M}{m+1} & \binom{M}{m+2} & \binom{M}{m+3} & \binom{M}{m+4} & \dots & \binom{M}{M} \end{pmatrix} \quad (4.2.31)$$

$$\mathbf{A}^{m,M} = \begin{pmatrix} \alpha_{m,m}^{m,M} \\ \alpha_{m+1,m}^{m,M} \\ \alpha_{m+2,m}^{m,M} \\ \alpha_{m+3,m}^{m,M} \\ \vdots \\ \alpha_{M,m}^{m,M} \end{pmatrix} \quad (4.2.32)$$

$$U_{M-m+1,1} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \quad (4.2.33)$$

- $\mathbf{P}^{m,M}$ est une matrice carré de taille $M - m + 1$ qui est extraite du triangle de Pascal des coefficients binomiaux
- $\mathbf{A}^{m,M}$ est le vecteur de taille $M - m + 1$ des coefficients de la combinaison linéaire que nous cherchons à déterminer explicitement
- $U_{M-m+1,1}$ est le vecteur de taille $M - m + 1$ rempli de 1

L'ensemble des équations linéaires que décrit le système (4.2.30) est donné par l'équivalence ci-dessous :

$$\mathbf{P}^{m,M} \cdot \mathbf{A}^{m,M} = U_{M-m+1,1} \Leftrightarrow \left\{ \sum_{k'=m}^k \alpha_{k'}^{m,M} \binom{k}{k'} = 1 \right\}_{k=m,\dots,M} \quad (4.2.34)$$

Nous voyons que le vecteur des coefficients $\alpha_k^{m,M}$, $k = m, \dots, M$; est obtenue en inversant $\mathbf{P}^{m,M}$:

$$\boxed{\mathbf{A}^{m,M} = (\mathbf{P}^{m,M})^{-1} \cdot U_{M-m+1,1}} \quad (4.2.35)$$

Nous rappelons la formule d'inversion du triangle de Pascal qui nous permet alors de donner une formulation explicite des $\alpha_k^{m,M}$ en fonction des coefficients binomiaux. La démonstration de cette formule d'inversion pourra être trouvée dans [Com70].

Propriété 4.2.8 Formule d'inversion du triangle de Pascal

Soient, $\{a_k; k = 0, \dots, n\}$ et $\{b_k; k = 0, \dots, n\}$ deux familles d'éléments tels que :

$$\sum_{k'=0}^k a_{k'} \binom{k}{k'} = b_k \quad (4.2.36)$$

alors :

$$a_k = \sum_{k'=0}^k (-1)^{k-k'} \binom{k}{k'} b_{k'} \quad (4.2.37)$$

Propriété 4.2.9 Expression des coefficients $\alpha_k^{m,M}$ en fonction des coefficients binomiaux

En appliquant la formule d'inversion de Pascal aux familles suivantes, $a_k = \alpha_k^{m,M}$ et $b_k = 1, \forall k = m, \dots, M$, nous obtenons l'expression suivante des coefficients de l'enveloppe de Condorcet :

$$\alpha_k^{m,M} = \sum_{k'=m}^k (-1)^{k-k'} \binom{k}{k'} \quad \forall k = m, \dots, M \quad (4.2.38)$$

Par convention, dans le cas où $k < m$ ou $k > M$, nous prendrons $\alpha_k^{m,M} = 0$.

Démonstration

L'application de la formule d'inversion est valable dans notre cas étant donné que \mathbf{P} est une sous-matrice triangulaire du triangle de Pascal dont le premier terme \mathbf{P}_{11} correspond à un terme de la diagonale $\binom{m}{m} = 1$ et dont les termes au-dessus de la diagonale sont nuls.

□

Nous avons déterminé une formulation explicite des coefficients $\alpha_k^{m,M}$. Nous avons alors exprimé l'enveloppe de Condorcet et les enveloppes de Condorcet partielles par une combinaison linéaire des S_k^M .

Propriété 4.2.10 Expression arithmétique de l'enveloppe de Condorcet

L'enveloppe de Condorcet de M relations binaires relative à une majorité m est donnée par :

$$E_m^M = \sum_{k=m}^M \alpha_k^{m,M} S_k^M \quad (4.2.39)$$

où les coefficients $\alpha_k^{m,M}, \forall k = m, \dots, M$ sont donnés par (4.2.38).

Propriété 4.2.11 Expression arithmétique d'une enveloppe de Condorcet partielle d'ordre m

Nous avons défini en déf-4.2.4 les enveloppes de Condorcet partielles d'ordre k par $H_m^M = E_m^M - E_{m+1}^M \quad \forall 0 \leq m \leq M$. Nous avons alors l'expression arithmétique suivante de H_m^M :

$$H_m^M = \sum_{k=m}^M (\alpha_k^{m,M} - \alpha_k^{m+1,M}) S_k^M \quad (4.2.40)$$

où les coefficients $\alpha_k^{m,M}, \forall k = m, \dots, M$ sont donnés par (4.2.38).

Démonstration

$$\begin{aligned} H_m^M &= E_m^M - E_{m+1}^M \\ &= \sum_{k=m}^M \alpha_k^{m,M} S_k^M - \sum_{k=m+1}^M \alpha_k^{m+1,M} S_k^M \\ &= \sum_{k=m}^M \alpha_k^{m,M} S_k^M - \sum_{k=m}^M \alpha_k^{m+1,M} S_k^M \quad \text{car, par convention, } \alpha_{m+1}^{m+1,M} = 0 \\ &= \sum_{k=m}^M (\alpha_k^{m,M} - \alpha_k^{m+1,M}) S_k^M \end{aligned}$$

□

Nous donnons dans tab-4.3 les coefficients $\alpha_k^{m,M}$ dans le cas particulier où la majorité m vaut $\frac{M}{2}$ si M est pair et $\lfloor \frac{M}{2} \rfloor + 1$ si M est impair où $\lfloor a \rfloor$ est le plus grand entier relatif inférieur ou égal à a .

$M \setminus k$	1	2	3	4	5	6	7	8	9	10
1	1									
2	1	-1								
3	0	1	-2							
4	0	1	-2	3						
5	0	0	1	-3	6					
6	0	0	1	-3	6	-10				
7	0	0	0	1	-4	10	-20			
8	0	0	0	1	-4	10	-20	35		
9	0	0	0	0	1	-5	15	-35	70	
10	0	0	0	0	1	-5	15	-35	70	-126

TABLE 4.3 – Tabulation des coefficients $\alpha_k^{m,M}$ dans le cas particulier où m représente la “majorité large” de M

4.2.2.3 Propriétés des coefficients $\alpha_k^{m,M}$

4.2.2.3.1 Les coefficients $\alpha_k^{m,M}$ et les coefficients binomiaux

Les coefficients $\alpha_k^{m,M}$ peuvent s’exprimer de manière plus simple à partir du triangle de Pascal en utilisant les propriétés de ce dernier. Parmi les nombreuses formules de récurrence des coefficients binomiaux, nous avons la propriété dite de “hockey stick” en raison de la forme particulière qu’elle décrit sur le triangle de Pascal :

Propriété 4.2.12 Propriété du “Hockey stick”

$$\sum_{q=1}^{n-p} (-1)^{q+1} \binom{n}{p+q} = \binom{n-1}{p} \quad (4.2.41)$$

De cette propriété, nous obtenons l’expression simplifiée des $\alpha_k^{m,M}$ suivante :

Propriété 4.2.13 Expression simplifiée des coefficients $\alpha_k^{m,M}$ en fonction des coefficients binomiaux
 Nous avons l’expression suivante des coefficients $\alpha_k^{m,M}$:

$$\alpha_k^{m,M} = (-1)^{k-m} \binom{k-1}{m-1} \quad \forall k = m, \dots, M \quad (4.2.42)$$

Démonstration

En posant $n = k$, $p = m - 1$, $q = k'$ et en faisant finalement varier k' de m à k , nous aboutissons à une réécriture de la première partie de (4.2.41) qui est à un facteur $(-1)^{-m}$ près, l’expression des $\alpha_k^{m,M}$ donnée par (4.2.38) d’où le résultat.

□

Propriété 4.2.14 Expression simplifiée des coefficients $\alpha_k^{m,M} - \alpha_k^{m+1,M}$ d’une enveloppe de Condorcet partielle d’ordre m en fonction des coefficients binomiaux

Nous avons l’expression suivante des coefficients $\alpha_k^{m,M} - \alpha_k^{m+1,M}$:

$$\alpha_k^{m,M} - \alpha_k^{m+1,M} = (-1)^{k-m} \binom{k}{m} \quad \forall k = m, \dots, M \quad (4.2.43)$$

Démonstration

$$\begin{aligned}
 \alpha_k^{m,M} - \alpha_k^{m+1,M} &= (-1)^{k-m} \binom{k-1}{m-1} - (-1)^{k-m-1} \binom{k-1}{m} \\
 &= (-1)^{k-m} \binom{k-1}{m-1} - (-1)^{k-m+1} \binom{k-1}{m} \\
 &= (-1)^{k-m} \left(\binom{k-1}{m-1} + \binom{k-1}{m} \right) \\
 &= (-1)^{k-m} \binom{k}{m} \qquad \qquad \qquad \text{1ère identité de Pascal}
 \end{aligned}$$

□

Nous avons en fait retrouvé les formules combinatoires de Ch. Jordan qui généralisent celles de H. Poincaré, dite formule de crible, qui permet d’obtenir le cardinal de l’union de plusieurs ensembles.

Ch. Jordan étend les résultats de H. Poincaré dans un cadre probabiliste avec des épreuves répétées et introduit la notion de m cas favorables [Jor39], [Com70].

Notre approche couple donc ces aspects combinatoires et les matrices relationnelles qui sont les ensembles que nous manipulons. Par ailleurs, nous nous plaçons dans le cadre d’une problématique d’agrégation de données avec la notion de majorité.

Dans le tableau tab-4.4, nous donnons l’expression des enveloppes de Condorcet selon les formules de Jordan avant de donner d’autres interprétations de celles-ci. En effet, nous étudions ci-après, les liens des coefficients $\alpha_k^{m,M}$ avec les nombres de Lah et les coefficients binomiaux négatifs.

Enveloppes de Condorcet	Expression arithmétique de Jordan
E_m^M	$\sum_{k=m}^M (-1)^{k-m} \binom{k-1}{m-1} S_k^M$
H_m^M	$\sum_{k=m}^M (-1)^{k-m} \binom{k}{m} S_k^M$

TABLE 4.4 – Enveloppes de Condorcet selon les coefficients combinatoires de Jordan

4.2.2.3.2 Les coefficients $\alpha_k^{m,M}$ et les nombres de Lah

Nous pouvons exprimer les $\alpha_k^{m,M}$, en fonction des nombres de Lah. Ces derniers sont de 2 types : les nombres de Lah non signés que nous noterons par $Lah(k, m)$ et les nombres de Lah signés que nous noterons par $Lah'(k, m)$. Nous avons respectivement :

$$Lah(k, m) = \binom{k-1}{m-1} \frac{k!}{m!} \tag{4.2.44}$$

$$Lah'(k, m) = (-1)^k \binom{k-1}{m-1} \frac{k!}{m!} \tag{4.2.45}$$

$Lah(k, m)$ représente le nombre de partitions possibles en m classes d’un ensemble de k éléments ordonnés tels qu’à l’intérieur de chaque classe les éléments soient ordonnées linéairement⁶.

$Lah'(k, m)$ sont définis en combinatoire comme étant les coefficients issus de l’inversion entre les factorielles croissantes et les factorielles décroissantes. Ils vérifient des identités qui

6. ordre large et total cf section 6.1

sont en complémentarité avec celles de J. Stirling [PP]. Ce sont ces derniers que nous prenons en compte tout particulièrement.

Dans le cadre de notre problématique, nous voyons que nous avons les relations suivantes :

Propriété 4.2.15 *Expression des coefficients $\alpha_k^{m,M}$ en fonction des nombres de Lah signés*

Nous avons l'expression suivante des coefficients $\alpha_k^{m,M}$:

$$\alpha_k^{m,M} = (-1)^m \frac{m!}{k!} \text{Lah}'(k, m) \quad \forall k = m, \dots, M \quad (4.2.46)$$

De la même manière, nous avons les propriétés suivantes :

Propriété 4.2.16 *Expression simplifiée des coefficients $\alpha_k^{m,M} - \alpha_k^{m+1,M}$ d'une enveloppe de Condorcet partielle d'ordre m en fonction des nombres de Lah signés*

Nous avons l'expression suivante des coefficients $\alpha_k^{m,M} - \alpha_k^{m+1,M}$:

$$\alpha_k^{m,M} - \alpha_k^{m+1,M} = (-1)^m \frac{(m+1)!}{(k+1)!} \text{Lah}'(k+1, m+1) \quad \forall k = m, \dots, M \quad (4.2.47)$$

Dans le tableau tab-4.5, nous donnons l'expression des enveloppes de Condorcet en mettant en évidence les nombres de Lah signés.

Enveloppes de Condorcet	Expression arithmétique en fonction des $\text{Lah}'(k, m)$
E_m^M	$((-1)^m m!) \sum_{k=m}^M \frac{\text{Lah}'(k, m)}{k!} S_k^M$
H_m^M	$((-1)^{m+1} (m+1)!) \sum_{k=m}^M \frac{\text{Lah}'(k+1, m+1)}{(k+1)!} S_k^M$

TABLE 4.5 – Enveloppes de Condorcet selon les nombres de Lah

4.2.2.3.3 Les coefficients $\alpha_k^{m,M}$ et les coefficients binomiaux négatifs

Les coefficients $\alpha_k^{m,M}$ sont enfin intimement liés aux coefficients binomiaux négatifs définis par la formule suivante :

$$\binom{-n}{p} = \frac{(-n)(-n-1)(-n-2)\dots(-n-p+1)}{p!} \quad (4.2.48)$$

$$= (-1)^p \binom{n+p-1}{p} \quad (4.2.49)$$

Nous avons en effet la propriété suivante :

Propriété 4.2.17 *Expression des coefficients $\alpha_k^{m,M}$ en fonction des coefficients binomiaux négatifs*
 Nous avons l'expression suivante des coefficients $\alpha_k^{m,M}$:

$$\alpha_k^{m,M} = \binom{-m}{k-m} \quad \forall k = m, \dots, M \quad (4.2.50)$$

Démonstration

$$\begin{aligned} \binom{-m}{k-m} &= (-1)^{k-m} \binom{m+k-m-1}{k-m} && \text{selon (4.2.49)} \\ &= (-1)^{k-m} \binom{k-1}{(k-1)-(k-m)} && \text{car } \binom{n}{p} = \binom{n}{n-p} : 4\text{\`eme identit e de Pascal} \\ &= (-1)^{k-m} \binom{k-1}{m-1} \\ &= \alpha_k^{m,M} && \text{selon (4.2.42)} \end{aligned}$$

□

Propri t  4.2.18 *Propri t  de r currence des coefficients binomiaux n gatifs*
 Les coefficients binomiaux n gatifs respectent les propri t s suivantes :

$$\binom{-n}{p} = (-1)^p \sum_{k=0}^p \left| \binom{-n+1}{k} \right| \quad (4.2.51)$$

$$\binom{-n}{p} = \binom{-n+1}{p} - \binom{-n}{p-1} \quad (4.2.52)$$

Nous d duisons de ces propri t s l'expression des coefficients des enveloppes H_m^M en fonction des coefficients binomiaux n gatifs ainsi que des relations de r currences entre les coefficients $\alpha_k^{m,M}$.

Propri t  4.2.19 *Expression des coefficients $\alpha_k^{m,M} - \alpha_k^{m+1,M}$ d'une enveloppe de Condorcet partielle d'ordre m en fonction des coefficients binomiaux n gatifs*
 Nous avons l'expression suivante des coefficients $\alpha_k^{m,M} - \alpha_k^{m+1,M}$:

$$\alpha_k^{m,M} - \alpha_k^{m+1,M} = \binom{-m-1}{k-m} \quad \forall k = m, \dots, M \quad (4.2.53)$$

D monstration

Le r sultat est obtenu en exprimant $\alpha_k^{m,M}$ et $\alpha_k^{m+1,M}$ en fonction des coefficients binomiaux n gatifs et en appliquant (4.2.52) avec $n = m + 1$ et $p = k - m$.

□

Propri t  4.2.20 *Relation de r currence entre les $\alpha_k^{m,M}$*

$$\alpha_k^{m,M} = \alpha_{k-1}^{m-1,M'} - \alpha_{k-1}^{m,M''} \quad \text{o  } 0 \leq m < k \leq \min(M, M' + 1, M'' + 1) \quad (4.2.54)$$

D monstration

Il suffit de v rifier que $\alpha_k^{m,M} = \alpha_{k-1}^{m-1,M'} - \alpha_{k-1}^{m,M''}$ est  quivalent   $\binom{-m}{k-m} = \binom{-m+1}{k-m} - \binom{-m}{k-m-1}$ qui est  quivalent   (4.2.52) en posant $n = m$ et $p = k - m$.

Par ailleurs, nous prenons $k \leq \min(M, M' + 1, M'' + 1)$ car dans le cas contraire au moins un terme s'annule en raison du fait que $\alpha_k^{m,M} = 0$ pour $k > M$, ce qui rend fausse l' galit .

□

L'expression arithmétique des enveloppes de Condorcet en fonction des coefficients binomiaux négatifs rappellent quelque peu celle du développement selon la formule du binôme de Newton dans le cas négatif : $(a + b)^{-m} = \sum_{k=0}^{\infty} \binom{-m}{k} a^k b^{-(m+k)}$. En particulier, lorsque $a = 1$, nous avons $(1 + b)^{-m} = \sum_{k=0}^{\infty} \binom{-m}{k} b^{-(m+k)}$.

Dans le tableau tab-4.6, nous donnons l'expression des enveloppes de Condorcet selon les coefficients binomiaux négatifs et nous montrons le rapprochement avec la formule du binôme dans le cas négatif.

Enveloppes de Condorcet	Expression arithmétique	Expression rappelant la formule du binôme dans le cas négatif
E_m^M	$\sum_{k=m}^M \binom{-m}{k-m} S_k^M$	$\sum_{k=0}^{M-m} \binom{-m}{k} S_{m+k}^M$
H_m^M	$\sum_{k=m}^M \binom{-m-1}{k-m} S_k^M$	$\sum_{k=0}^{M-m} \binom{-m-1}{k} S_{m+k}^M$

TABLE 4.6 – Enveloppes de Condorcet selon les coefficients binomiaux négatifs

Nous donnons dans tab-4.7, le triangle de Pascal ainsi que les coefficients binomiaux négatifs à partir desquels, nous pouvons déterminer les coefficients $\alpha_k^{m,M}$.

$n \setminus p$	0	1	2	3	4	5	6	7	8	9	10
-10	1	-10	55	-220	715	-2002	5005	-11440	24310	-48620	92378
-9	1	-9	45	-165	495	-1287	3003	-6435	12870	-24310	43758
-8	1	-8	36	-120	330	-792	1716	-3432	6435	-11440	19448
-7	1	-7	28	-84	210	-462	924	-1716	3003	-5005	8008
-6	1	-6	21	-56	126	-252	462	-792	1287	-2002	3003
-5	1	-5	15	-35	70	-126	210	-330	495	-715	1001
-4	1	-4	10	-20	35	-56	84	-120	165	-220	286
-3	1	-3	6	-10	15	-21	28	-36	45	-55	66
-2	1	-2	3	-4	5	-6	7	-8	9	-10	11
-1	1	-1	1	-1	1	-1	1	-1	1	-1	1
0	1										
1	1	1									
2	1	2	1								
3	1	3	3	1							
4	1	4	6	4	1						
5	1	5	10	10	5	1					
6	1	6	15	20	15	6	1				
7	1	7	21	35	35	21	7	1			
8	1	8	28	56	70	56	28	8	1		
9	1	9	36	84	126	126	84	36	9	1	
10	1	10	45	120	210	252	210	120	45	10	1

TABLE 4.7 – Triangle de Pascal et coefficients binomiaux négatifs

4.3 Le calcul consensuel ou l'application du calcul de l'enveloppe de Condorcet à des données non binaires

Nous avons défini les enveloppes selon une approche logique à l'aide d'opérations booléennes. Nous avons fait remarquer, que l'interprétation pouvait aussi être en termes ensemblistes. Par conséquent, les définitions et propriétés que nous avons données précédemment peuvent être étendues à des ensembles mesurés, c'est à dire des données non binaires appartenant à l'intervalle $[0, 1]$.

L'enveloppe de Condorcet, que nous appellerons aussi calcul consensuel, nous donne alors une mesure consensuelle associée à une majorité m de la relation entre deux objets :

$$[E_m^M]_{ii'} = \text{Mesure consensuelle qu'il y ait au moins } m \text{ relations qui indiquent que } O^i \text{ est en relation avec } O^{i'} \quad (4.3.1)$$

Lorsque nous passons de la logique bivalente à la logique multivalente, il existe plusieurs façons de mesurer la valeur de vérité floue⁷ de la conjonction ou de la disjonction de deux propositions floues. Ces conjonctions et disjonctions particulières sont appelées "t-norms" et "t-conorms". Nous rappelons leurs définitions ci-dessous avant de donner quelques exemples.

Définition 4.3.1 "Triangular norms" et "triangular conorms"

Une "t-norm" est une opération binaire \otimes sur l'intervalle $[0, 1]$, c'est à dire une fonction de $[0, 1]^2$ dans $[0, 1]$ tel que pour $x, y, z \in [0, 1]$, les axiomes suivants sont respectés :

- $x \otimes y = y \otimes x$
- $x \otimes (y \otimes z) = (x \otimes y) \otimes z$
- $x \otimes y \leq x \otimes z \quad \forall y \leq z$
- $x \otimes 1 = x$

Les "t-conorms", \oplus , sont définis à partir des "t-norms" selon l'équation suivante : $x \oplus y = 1 - (1 - x) \otimes (1 - y)$.

Remarque 4.3.1 "t-norms" et copules

Il existe, en théorie des probabilités, des fonctions permettant de modéliser la nature de la dépendance entre deux ou plusieurs variables aléatoires réelles. Il s'agit des copules qui sont des fonctions permettant d'obtenir la distribution jointe d'un ensemble de variables aléatoires réelles à partir de leurs distributions marginales [Nel99], [SS83], [Men42]. Les copules sont des fonctions très proches des "t-norms" [KMP00]. Toutefois, dans le cas de variables aléatoires discrètes, leurs définitions et leurs études sont moins immédiates. Pour cette raison, nous aborderons le problème dans le cadre de la logique multivalente dont l'utilisation est inscrite dans un cadre formel moins "strict" que les copules.

Soient C^k, C^l deux éléments de \mathbb{C} dont les termes sont des réels appartenant à $[0, 1]$. Nous présentons ci-dessous des exemples classiques de "t-norms".

$$[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = \begin{cases} M(C_{ii'}^k, C_{ii'}^l) = \min(C_{ii'}^k, C_{ii'}^l) \\ \text{ou} \\ \Pi(C_{ii'}^k, C_{ii'}^l) = C_{ii'}^k C_{ii'}^l \\ \text{ou} \\ W(C_{ii'}^k, C_{ii'}^l) = \max(C_{ii'}^k + C_{ii'}^l - 1, 0) \end{cases} \quad (4.3.2)$$

7. le terme floue fait référence à une valeur de vérité ou d'appartenance, dans le cas ensembliste ou "fuzzy sets", appartenant à $[0, 1]$

De la même manière, nous présentons les “t-conorms” correspondants.

$$[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = \begin{cases} M^*(C_{ii'}^k, C_{ii'}^l) = \max(C_{ii'}^k, C_{ii'}^l) \\ \text{ou} \\ \Pi^*(C_{ii'}^k, C_{ii'}^l) = C_{ii'}^k + C_{ii'}^l - C_{ii'}^k C_{ii'}^l \\ \text{ou} \\ W^*(C_{ii'}^k, C_{ii'}^l) = \min(C_{ii'}^k + C_{ii'}^l, 1) \end{cases} \quad (4.3.3)$$

Les cas particuliers M et W exposés sont appelés bornes de Fréchet-Hoeffding et représentent respectivement des bornes supérieure et inférieure de l'ensemble des “t-norms”. Nous retrouvons d'ailleurs, la conjonction et la disjonction de Lukasiewicz que nous avons exploité en sous-section 4.1.2. Nous avons par ailleurs, la propriété suivante concernant les relations entre “t-norms” et “t-conorms”.

Propriété 4.3.1 Relations de De Morgan

Les opérations de conjonctions \otimes et de disjonctions \oplus précédemment citées sont telles que :

$$C^k \otimes C^l = \overline{\overline{C^k} \oplus \overline{C^l}} = U_N - \overline{C^k} \oplus \overline{C^l} \quad (4.3.4)$$

$$C^k \oplus C^l = \overline{\overline{C^k} \otimes \overline{C^l}} = U_N - \overline{C^k} \otimes \overline{C^l} \quad (4.3.5)$$

Suite à ces rappels, nous précisons finalement la signification des termes employés dans cette section qui représente donc une extension des résultats énoncés précédemment aux cas de données non binaires.

Remarque 4.3.2 Signification des termes $C_{ii'}^k$, et des symboles \oplus et \otimes dans le cadre de la sous-section 4.3

Dans le cadre de cette section, nous aurons :

- $C_{ii'}^k = \mu(O^i \mathcal{R}^k O^{i'}) =$ valeur de vérité comprise entre 0 et 1, de la proposition “ O^i est en relation avec $O^{i'}$ selon la relation \mathcal{R}^k ”
- $[C^k \oplus C^l]_{ii'} = C_{ii'}^k \oplus C_{ii'}^l = \mu(O^i \mathcal{R}^k O^{i'} \vee O^i \mathcal{R}^l O^{i'}) =$ valeur de vérité comprise entre 0 et 1, de la disjonction “ O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R}^k ou la relation \mathcal{R}^l ”
- $[C^k \otimes C^l]_{ii'} = C_{ii'}^k \otimes C_{ii'}^l = \mu(O^i \mathcal{R}^k O^{i'} \wedge O^i \mathcal{R}^l O^{i'}) =$ valeur de vérité comprise entre 0 et 1, de la conjonction “ O^i est en relation avec $O^{i'}$ pour la relation \mathcal{R}^k et la relation \mathcal{R}^l ”

4.3.1 Le calcul consensuel comme méthode d'agrégation et son application dans le cadre de relations d'équivalence

L'enveloppe de Condorcet peut être vue comme une fonction d'agrégation de relations binaires floues. Nous rappelons la définition formelle d'un tel opérateur et renvoyons aux références suivantes pour de plus amples détails [Det00], [KMP00].

Définition 4.3.2 Opérateur d'agrégation

Un opérateur d'agrégation est une fonction Agreg de $\bigcup_{k=1}^M [0, 1]^k$ dans $[0, 1]$ qui satisfait les conditions suivantes :

- $\text{Agreg}(a) = a \quad \forall a \in [0, 1]$

- $\text{Agreg}(0, \dots, 0) = 0$ et $\text{Agreg}(1, \dots, 1) = 1$
- $\text{Agreg}(a_1, \dots, a_M) \leq \text{Agreg}(b_1, \dots, b_M)$ si $a_k \leq b_k \quad \forall k = 1, \dots, M$

Parmi les propriétés supplémentaires que peuvent vérifier les opérateurs d'agrégation, l'associativité est, en pratique, importante. Elle indique en particulier, qu'il est possible de mettre à jour progressivement des données supplémentaires sans avoir besoin à re-calculer du début l'agrégation :

$$\text{Agreg}(\text{Agreg}(a, b), c) = \text{Agreg}(a, \text{Agreg}(b, c)) = \text{Agreg}(a, b, c) \quad (4.3.6)$$

En raison de la propriété prop-4.2.7 de décomposition, **les enveloppes de Condorcet sont quasi-associatives** au sens suivant.

Supposons, sans perte de généralité, que la stratégie de choix de la majorité est de prendre $m = \frac{M}{2}$ si M est pair et $\lfloor \frac{M}{2} \rfloor + 1$ si M est impair. Ceci correspond à prendre la "majorité large". Supposons tout d'abord que M est pair. Si une nouvelle variable est disponible, $M + 1$ est impair et la majorité m , au sens de la stratégie choisie au préalable, augmente aussi de 1. Dans ce cas, en appliquant la propriété de décomposition de l'enveloppe, nous pouvons mettre à jour l'information apportée par une nouvelle variable :

$$E_{m+1}^{M+1} = E_{m+1}^M \otimes H_0^1 + E_m^M \otimes H_1^1 \quad (4.3.7)$$

Supposons maintenant que M est impair. Dans ce cas, $M + 1$ est pair mais la majorité m , au sens de la stratégie choisie ne change pas. La mise à jour de la nouvelle variable s'effectue de la manière suivante :

$$E_m^{M+1} = E_m^M \otimes H_0^1 + E_{m-1}^M \otimes H_1^1 \quad (4.3.8)$$

Ainsi, l'enveloppe de Condorcet est quasi-associative dans la mesure où il est possible de mettre à jour progressivement l'information apportée par une nouvelle variable en gardant en mémoire, pour un même ensemble de M variables, 3 enveloppes de Condorcet uniquement : E_{m-1}^M , E_m^M et E_{m+1}^M .

Nous nous plaçons désormais, dans le contexte des relations d'équivalence. Nous appliquons donc les enveloppes de Condorcet comme méthode d'agrégation de données non nécessairement binaires et nous cherchons à aboutir à un modèle relationnel de classification automatique de données hétérogènes. L'approche consiste à donner une modélisation nouvelle de la mesure consensuelle de similarité entre objets décrits soient par des variables qualitatives, soient par des variables quantitatives, soient par les deux. Nous poursuivons donc notre extension de l'AR à la classification de données autres que binaires. Dans le cadre de ce chapitre, nous proposons le modèle suivant :

$\max_X \sum_{i, i'} ([E_m^M]_{ii'} - \mathcal{M}(\{[E_m^M]_{ii'}\})) X_{ii'}$		
sous les contraintes		
$X_{ii'} \in \{0, 1\}$		(binarité)
$X_{ii} = 1$	$\forall i = 1, \dots, N$	(réflexivité)
$X_{i'i} = X_{ii'}$	$\forall i, i' = 1, \dots, N$	(symétrie)
$X_{ii'} + X_{i'i''} - X_{ii''} \leq 1$	$\forall i, i', i'' = 1, \dots, N$	(transitivité)

où $\mathcal{M}(\{[E_m^M]_{ii'}\})$ est une tendance centrale telle une moyenne généralisée d'ordre t' , $\mathcal{M}^{t'}(\{[E_m^M]_{ii'}\})$, ou telle la médiane, qui est calculée sur l'ensemble des valeurs $\{[E_m^M]_{ii'}\}_{i, i'=1, \dots, N}$.

Autrement dit, O^i et $O^{i'}$ seront susceptibles d'être dans une même classe si leur mesure consensuelle de similarité $[E_m^M]_{ii'}$ est supérieure à une moyenne de l'ensemble des mesures consensuelles de similarité $\mathcal{M}^{t'}(\{[E_m^M]_{ii'}\})$.

La mesure consensuelle de similarité, selon l'expression de l'enveloppe de Condorcet en fonction des coefficients binomiaux négatifs, est donnée par :

$$[E_m^M]_{ii'} = \sum_{k=0}^{M-m} \binom{-m}{k} [S_{m+k}^M]_{ii'} \quad \forall i, i' = 1, \dots, N \quad (4.3.10)$$

où

$$[S_{m+k}^M]_{ii'} = \sum_{1 \leq i_1 < i_2 < \dots < i_k \leq M} [C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_k}]_{ii'} \quad (4.3.11)$$

A la différence des extensions du modèle relationnel données au chapitre 3 par (3.1.47) et (3.2.4), le modèle suggéré ici est davantage de nature logique et combinatoire que géométrique. L'intérêt est en particulier de proposer des nuances dans les mesures de similarité en intégrant la notion de majorité. Pour illustrer l'intérêt de cette démarche, nous donnons ci-dessous un exemple.

Supposons que les mesures non binaires dont nous disposons soient interprétées dans un contexte probabiliste : $C_{ii'}^k \in [0, 1]$ donne la probabilité pour la variable V^k que les objets O^i et $O^{i'}$ soient dans la même classe. Dans ce cas, si $m = M$ et si nous émettons l'hypothèse d'indépendance entre variables, nous pouvons interpréter E_M^M comme une matrice relationnelle de "vraisemblances" puisque dans ce cas :

$$[E_M^M]_{ii'} = \prod_k C_{ii'}^k = \text{"Vraisemblance" que } O^i \text{ et } O^{i'} \text{ soient dans la même classe} \quad (4.3.12)$$

Le cas d'une "vraisemblance" classique correspond au cas $m = M$. Or, celui-ci peut, dans le cas de la classification, engendrer des inconvénients. Prenons l'exemple numérique suivant constitué de 4 variables donnant chacune la probabilité que O^i et $O^{i'}$ soient dans la même classe :

Variables	$C_{ii'}^1$	$C_{ii'}^2$	$C_{ii'}^3$	$C_{ii'}^4$
Probabilité	0.9	0.8	0.7	0

Calculons alors les mesures consensuelles pour les différentes majorités à partir de (4.3.10) et pour $\otimes = \Pi$. Nous obtenons alors les probabilités suivantes :

Majorité m	4	3	2	1
$[E_m^M]_{ii'}$	0.000	0.504	0.902	0.994

Dans cet exemple simple, 3 variables sur 4 donnent une forte probabilité à l'évènement " O^i et $O^{i'}$ sont dans la même classe". Selon la vraisemblance classique, ce qui correspond à une majorité unanime $m = 4$, nous obtenons $C_{ii'}^4 = 0$. Ce résultat est peu satisfaisant. Nous voyons alors que les probabilités consensuelles sont moins sensibles au fait qu'une variable sur quatre uniquement affirme avec certitude que les objets ne sont pas dans la même classe. De cette façon, nous illustrons le fait que le calcul consensuel permet d'être moins dépendant des valeurs extrêmes ou aberrantes. Il permet de donner une mesure de similarité qui tient compte de la majorité des variables qui participent à la mesure globale de similarité, ce qui, eu égard à l'exemple précédent, est plus robuste.

4.3.2 Modélisation de la mesure consensuelle de la similarité entre objets

Nous exposons dans ce qui suit, plusieurs approches de la modélisation de la mesure consensuelle de similarité entre objets en attribuant différentes interprétations aux éléments $[S_{m+k}^M]_{ii'}$, $\forall k = m, \dots, M$.

4.3.2.1 Cas des variables qualitatives

Dans le cadre de ce paragraphe, nous considérons des variables qualitatives V^k , $k = 1, \dots, M$ ayant chacune p_k modalités. Chaque variable V^k est représentée par sa matrice relationnelle binaire C^k .

4.3.2.1.1 Matrices relationnelles individuelles de similarités

Nous proposons tout d'abord de pondérer la similarité logique $C_{ii'}^k = 1$ de la relation entre O^i et $O^{i'}$ par la probabilité empirique que deux objets pris au hasard puissent avoir la même modalité de la variable V^k considérée. Cette démarche est proche de la similarité régularisée que nous avons rappelée en introduction page 9. Chaque variable qualitative V^k est donc représentée par une matrice relationnelle non binaire C^k , dont le terme général est le suivant :

$$C_{ii'}^k = \frac{C_{ii'}^k}{N^2} \quad \forall i, i' = 1, \dots, N \quad (4.3.13)$$

Par la suite, la conjonction $C^k \otimes C^l$, entre deux matrices relationnelles, pourra prendre plusieurs formes parmi les différentes "t-norms"⁸ existantes et selon différentes interprétations [KMP00]. Dans le cas d'indépendance entre variables par exemple, nous pourrions prendre : $C^k \otimes C^l = \Pi(C^k, C^l) = C^k C^l$ où $C^k C^l$ est le produit terme à terme⁹ entre les deux matrices relationnelles. Une fois la conjonction \otimes choisie ou déterminée, elle permet de calculer la valeur des termes de chaque élément S_{m+k}^M :

$$[S_{m+k}^M]_{ii'} = \sum_{1 \leq i_1 < i_2 < \dots < i_{m+k} \leq M} [C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_{m+k}}]_{ii'} \quad (4.3.14)$$

Finalement la valeur de l'enveloppe de Condorcet est obtenue par combinaison linéaire selon l'équation (4.3.10).

4.3.2.1.2 Matrices d'intersections relationnelles de similarités

Nous proposons ensuite de donner des poids directement aux intersections relationnelles sans passer par l'attribution de poids à chaque matrice relationnelle individuelle. Nous restons cependant dans le même esprit que précédemment : le poids affecté à une intersection relationnelle $[C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_{m+k}}]$ est la probabilité jointe empirique que deux objets pris au hasard aient les mêmes modalités de chacune des variables appartenant à l'intersection :

$$[S_{m+k}^M]_{ii'} = \sum_{1 \leq i_1 < i_2 < \dots < i_{m+k} \leq M} \frac{[C^{i_1} C^{i_2} \dots C^{i_{m+k}}]_{..}}{N^2} [C^{i_1} \otimes C^{i_2} \otimes \dots \otimes C^{i_{m+k}}]_{ii'} \quad (4.3.15)$$

où

$$[C^{i_1} C^{i_2} \dots C^{i_{m+k}}]_{..} = \sum_{i, i'} [C^{i_1} C^{i_2} \dots C^{i_{m+k}}]_{ii'} \quad (4.3.16)$$

8. dans le cas d'une interprétation logique de l'opération \otimes

9. aussi appelé dans un contexte plus général produit de Schur ou de Hadamard

4.3.2.1.3 Matrices d'intersections relationnelles d'indices de similarité

Dans cette dernière approche, concernant les variables qualitatives, nous proposons d'utiliser les indices de similarité que nous avons étudiés en paragraphe 3.1.2.4 page 80 comme mesure floue de la relation entre deux objets.

A chaque élément S_{m+k}^M , correspond un sous-ensemble de $m+k$ variables. Nous prenons alors comme mesure de la relation entre deux objets relative à cette somme, l'indice de similarité d'ordre t entre objets, calculé en tenant uniquement compte des variables appartenant au sous-ensemble¹⁰.

Notons alors par $S_{i_1, \dots, i_{m+k}}^t(\vec{O}^i, \vec{O}^{i'})$ l'indice de similarité d'ordre t défini par l'équation (3.1.33) page 82 mais pour lequel, seules les variables $V^{i_1}, \dots, V^{i_{m+k}}$ interviennent pour le calcul des produits scalaires $\langle \vec{O}^i, \vec{O}^{i'} \rangle$, $\langle \vec{O}^i, \vec{O}^i \rangle$ et $\langle \vec{O}^{i'}, \vec{O}^{i'} \rangle$. Nous avons alors :

$$[S_{m+k}^M]_{ii'} = \sum_{1 \leq i_1 < i_2 < \dots < i_{m+k} \leq M} S_{i_1, \dots, i_{m+k}}^t(\vec{O}^i, \vec{O}^{i'}) \quad (4.3.17)$$

4.3.2.2 Cas des variables quantitatives

Dans ce paragraphe, nous considérons des variables quantitatives V^k , $k = 1, \dots, M$, à partir desquelles, nous déterminons des mesures non binaires de la similarité entre objets.

4.3.2.2.1 Matrices relationnelles individuelles de similarités

De manière parallèle au paragraphe 4.3.2.1.1, nous proposons ici de représenter chaque variable V^k par une matrice relationnelle C^k dont les termes donnent une mesure de la similarité entre objets. Nous proposons dans un premier temps, ci-dessous, la mesure de similarité normalisée engendrée par une variable quantitative qui est initialement proposée par Gower [Gow71] :

$$C_{ii'}^k = 1 - \frac{|V_i^k - V_{i'}^k|}{\max_{i''}(V_{i''}^k) - \min_{i''}(V_{i''}^k)} \quad (4.3.18)$$

Etant donné que cette mesure peut être sensible aux données extrêmes, nous pouvons utiliser les rangs de chaque objet de la variable triée V^{*k} à la place des valeurs numériques. Notons par $\text{Rang}(V^k, O^i)$ le rang de l'objet O^i dans la variable triée dans l'ordre croissant V^{*k} . Nous pouvons prendre comme mesure de similarité plus robuste celle donnée par la formule suivante :

$$C_{ii'}^k = 1 - \frac{|\text{Rang}(V^k, O^i) - \text{Rang}(V^k, O^{i'})|}{N - 1} \quad (4.3.19)$$

Cependant, dans le cas de données irrégulières, les mesures basées sur les rangs lissent trop les mesures de similarité. Nous proposons une troisième méthode à mi-chemin entre les similarités données par (4.3.18) et (4.3.19). Celle-ci repose sur les triordonnances qui ont été étudiées par I.C. Lerman puis par S. Chah [Ler81], [Cha85].

Nous introduisons pour cela les différentes quantités suivantes qui sont illustrées par la figure fig-4.1 :

- $\mathbb{A}^k(a, b) = \#\{O^i : \min(a, b) \leq V_i^k \leq \max(a, b)\}$ est le nombre d'objets dont les valeurs, pour la variable quantitative V^k , sont comprises entre $\min(a, b)$ et $\max(a, b)$
- $t_-^k = \min_i(\{V_i^k\}) - \epsilon$ est la valeur minimale de la variable V^k à laquelle nous retranchons $\epsilon > 0$

10. au sens d'une restriction de l'ensemble des variables à ce sous-ensemble de $m+k$ variables en intersection

- $t_+^k = \max_i(\{V_i^k\}) + \epsilon$ est la valeur maximale de la variable V^k à laquelle nous ajoutons $\epsilon > 0$
- $t_{ii'}^{k-} = \min(V_i^k, V_{i'}^k) - |V_i^k - V_{i'}^k|$ est la borne supérieure de la valeur des objets $O^{i'}$ tel que $V_{i'}^k \leq \min(V_i^k, V_{i'}^k)$ et $|V_{i'}^k - \min(V_i^k, V_{i'}^k)| \geq |V_i^k - V_{i'}^k|$
- $t_{ii'}^{k+} = \max(V_i^k, V_{i'}^k) + |V_i^k - V_{i'}^k|$ est la borne inférieure de la valeur des objets $O^{i'}$ tel que $V_{i'}^k \geq \max(V_i^k, V_{i'}^k)$ et $|V_{i'}^k - \max(V_i^k, V_{i'}^k)| \geq |V_i^k - V_{i'}^k|$

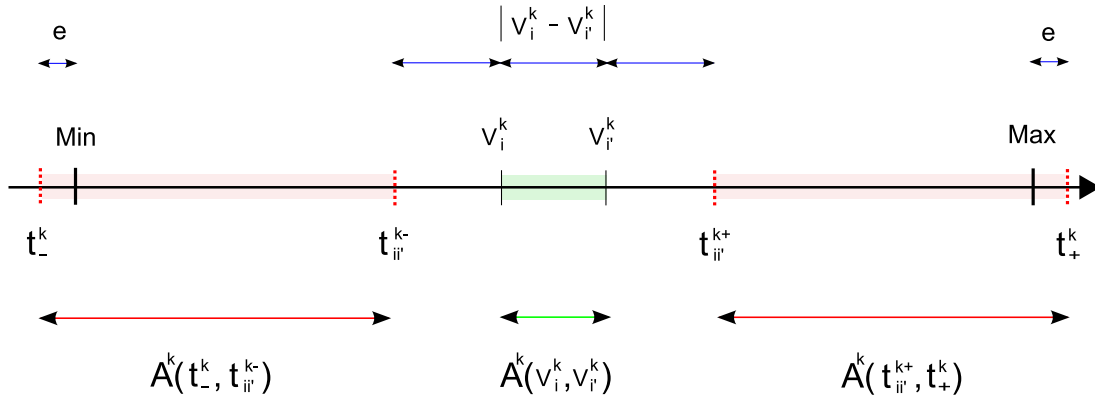


FIGURE 4.1 – Représentations des quantités intervenant dans le calcul de similarités entre objets en se basant sur les triordonnances

Nous proposons alors la mesure de similarité suivante :

$$C_{ii'}^k = \frac{1}{2} \left(1 - \frac{\mathbb{A}^k(V_i^k, V_{i'}^k) - \mathbb{A}^k(t_-^k, t_{ii'}^{k-}) - \mathbb{A}^k(t_{ii'}^{k+}, t_+^k)}{N} \right) \quad (4.3.20)$$

Nous pouvons aussi prendre des mesures de similarité régularisées pour des variables quantitatives telles qu'elles ont été proposées par H. Benhadda dans [Ben98].

A partir de ces matrices relationnelles individuelles, nous déterminerons les différentes conjonctions en utilisant une "t-norm" particulière. Nous pouvons dans un premier temps déterminer la valeur des éléments sommes (4.3.14) puis dans un deuxième temps l'enveloppe de Condorcet par combinaison linéaire en utilisant (4.3.10).

4.3.2.2.2 Matrices d'intersections relationnelles d'indices de similarité

L'idée poursuivie dans ce paragraphe est identique à celle du paragraphe 4.3.2.1.3. Nous calculons l'indice de similarité d'ordre $t > 0$ pour chaque élément de la somme S_{m+k}^M en restreignant les produits scalaires aux seules variables appartenant au sous-ensemble de l'intersection relationnelle. Nous notons de la même manière $\mathbf{S}_{i_1, \dots, i_{m+k}}^t(\vec{O}^i, \vec{O}^{i'})$, l'indice de similarité d'ordre t

entre les objets O^i et $O^{i'}$ restreint aux variables $V^{i_1}, \dots, V^{i_{m+k}}$. Nous proposons donc :

$$[S_{m+k}^M]_{ii'} = \sum_{1 \leq i_1 < i_2 < \dots < i_{m+k} \leq M} \mathbf{S}_{i_1, \dots, i_{m+k}}^t(\vec{O}^i, \vec{O}^{i'}) \quad (4.3.21)$$

4.3.2.3 Cas des variables qualitatives et quantitatives

Le traitement simultané de variables de types hétérogènes est souvent délicat étant donné les distinctions de leurs caractéristiques propres : les variables qualitatives sont de nature binaire et infèrent donc une structure rigide sur les données, tandis que les variables quantitatives montrent un *continuum* entre les similarités entre données et par conséquent une structure "floue".

Si nous souhaitons pouvoir classifier des objets décrits à la fois par des variables qualitatives et des variables quantitatives de manière simultanée, il peut être utile de disposer d'une méthode permettant d'une part de tenir compte des spécificités de chacun des types de variable, d'autre part, de combiner de façon cohérente l'information provenant de ces deux sources particulières.

Nous proposons d'employer la propriété de décomposition des enveloppes de Condorcet afin de constituer une méthode qui réponde aux propriétés citées précédemment.

Considérons un ensemble d'objets décrits par un groupe de p_1 variables qualitatives et un groupe de p_2 variables quantitatives. Dans le cadre du problème de données mixtes, nous proposons de résoudre le problème de classification posé par le programme (4.3.9) en remplaçant E_m^M par sa décomposition suivante :

$$E_m^M = \sum_{q=0}^{p_2} E_{m-q}^{p_1} \otimes H_q^{p_2} \quad (4.3.22)$$

Nous pouvons dès lors spécifier plusieurs types de conjonctions :

- un premier type de "t-norm" relatif aux variables qualitatives (resp. quantitatives) qui fonde les éléments $E_{m-q}^{p_1}$
- un deuxième type de "t-norm" relatif aux variables quantitatives (resp. qualitatives) qui fonde les éléments $H_q^{p_2}$
- un troisième type de "t-norm" relatif à la conjonction des deux groupes de variables de natures distinctes.

Deuxième partie

Aspects algébriques et combinatoires de l'AR des relations d'ordre et applications en théorie du choix social et en théorie des tresses

CHAPITRE 5

Etudes des codages de l'ensemble des classements possibles d'un ensemble d'objets

Sommaire

5.1 Définitions et rappels	152
5.1.1 Les différents formalismes matriciels des classements et les formules de passage correspondantes	153
5.1.1.1 Le codage algébrique	153
5.1.1.2 Le codage par affectation	154
5.1.1.3 Le codage relationnel	154
5.1.1.4 Lien entre le codage par affectation et le codage relationnel	157
5.1.1.5 Le codage sous forme de matrices de paires d'objets nommés	158
5.1.2 Rappels sur le groupe des permutations	159
5.1.2.1 Quelques propriétés du groupe des permutations	159
5.1.2.2 Le treillis des permutations et le graphe de Cayley des permutations	165
5.2 Action des transpositions sur les classements	170
5.2.1 Action des transpositions génératrices sur les matrices d'affectation et les matrices relationnelles	171
5.2.1.1 Action des transpositions sur les matrices d'affectation	171
5.2.1.2 Action des transpositions sur les matrices relationnelles	172
5.2.2 Action des transpositions sur les matrices de paires d'objets nommés	174

Dans ce chapitre, nous étudions l'ensemble des relations d'ordre strict et total sur un ensemble d'objets \mathcal{O} . Une relation d'ordre strict et total correspond à un classement sur les objets tel qu'il n'y ait pas d'*ex-aequo* ni de données manquantes. Ainsi, dans la suite, nous ne ferons pas de distinction entre les termes ordre strict et total et classement.

Ce chapitre est, dans l'esprit, très proche du chapitre 2 qui est consacré à l'étude de l'ensemble des relations d'équivalence sur un ensemble d'objets.

Dans un premier temps, nous rappelons plusieurs formalismes des relations d'ordre strict et total. Les classements étant totaux, nous montrons en particulier, qu'il existe une bijection entre le codage par affectation et le codage relationnel contrairement aux relations d'équivalence. Cette propriété nous permet de définir un nouveau type de codage nommé "codage sous forme de matrices de paires d'objets nommés" qui nous sera utile pour la modélisation des tresses¹ au chapitre 7.

Nous rappelons dans un second temps certaines propriétés du groupe des permutations qui entretient des liens très étroits avec l'ensemble des classements².

Nous étudierons enfin l'action des permutations sur une relation d'ordre strict et total. Selon que les permutations concernent les objets ou les rangs, nous aurons deux types d'action distincts³.

En somme, les objectifs de ce chapitre sont, d'une part, de donner un formalisme commun à l'étude des relations d'équivalence et celle des relations d'ordre par l'intermédiaire de l'AR. D'autre part, nous exploitons les liens entre classements et permutations ainsi que les actions des permutations sur des classements afin de définir un nouveau codage que nous chercherons à appliquer dans le cadre de la modélisation des tresses.

Remarque 5.0.1 Signification des éléments de l'ensemble des descripteurs \mathbb{D}

Dans le chapitre 2, nous avons étudié les relations d'équivalence pour lesquelles les descripteurs $D^j \in V^k$ pouvaient être des modalités non ordonnées⁴ d'une variable qualitative V^k ⁵.

Nous cherchons à employer dans le présent chapitre les mêmes notations que celles utilisées au chapitre 2 afin de donner une présentation unifiée et cohérente de l'étude des différents types de relations binaires dans le cadre de ce mémoire de thèse.

Dans ce chapitre, les "descripteurs", $\{D^1, \dots, D^N\}$, que nous emploierons seront en fait strictement équivalents à l'ensemble des N rangs donnés par l'ensemble $\mathbb{N}_N = \{1, \dots, N\}$. Nous mettons en bijection chaque élément D^j avec chaque rang $j : D^j \leftrightarrow j, \forall j = 1, \dots, N$. Nous avons par conséquent l'ordre suivant induit sur les "descripteurs" : $D^1 < D^2 < \dots < D^N$. Lorsque nous ferons référence à l'entier j plutôt qu'au descripteur ordonné D^j correspondant, nous parlerons d'indice de rang j ou de position j .

5.1 Définitions et rappels

Nous rappelons différents types de codage des relations d'ordre strict et total. Il s'agit des mêmes types de formalismes que ceux décrits en section 2.1 : le codage algébrique, le codage par affectation et le codage relationnel.

Nous introduisons de plus, un nouveau codage des classements appelé matrice des paires d'objets nommés. Celui-ci possède des propriétés intéressantes qui peuvent être exploitées pour la modélisation des tresses comme nous le verrons au chapitre 7.

Dans ce qui suit, nous noterons par \mathbb{O} l'ensemble des objets sur lequel est définie la relation d'ordre et par \mathbb{D} , l'ensemble de cardinal N des rangs possibles auxquels sont affectés les objets. Nous rappelons quelques propriétés du groupe des permutations et nous mettons ensuite en évidence la structure sous forme de graphe de Cayley que vérifie ce groupe.

1. les tresses sont dans un certain sens une extension des ordres stricts et totaux
2. en effet, nous pouvons considérer qu'il existe une bijection entre l'ensemble des permutations d'un ensemble d'objets et l'ensemble des classements de ces derniers cf [MM78]
3. cette double possibilité constitue une propriété fondamentale dans notre modélisation des tresses du chapitre 7
4. les modalités peuvent être ordonnées mais nous ne tenons pas compte de cette information
5. cf rq-2.1.1

5.1.1 Les différents formalismes matriciels des classements et les formules de passage correspondantes

Définition 5.1.1 Relation d'ordre strict et total

Une relation d'ordre strict et total, \mathcal{R} , d'un ensemble \mathbb{O} de N objets, $\{O^i, i = 1, \dots, N\}$, est une relation binaire sur cet ensemble d'éléments qui respecte les propriétés suivantes :

- Irréflexivité : $O^i \overline{\mathcal{R}} O^i \quad \forall O^i \in \mathbb{O}$
- Antisymétrie forte ou asymétrie : $O^i \mathcal{R} O^{i'} \Rightarrow O^{i'} \overline{\mathcal{R}} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} : O^i \neq O^{i'}$
- Transitivité : $O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''} \quad \forall (O^i, O^{i'}, O^{i''}) \in \mathbb{O} \times \mathbb{O} \times \mathbb{O}$
- Totalité : $O^i \mathcal{R} O^{i'} \vee O^{i'} \mathcal{R} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} : O^i \neq O^{i'}$

De même que pour les relations d'équivalence, il existe plusieurs manières de coder une relation d'ordre strict et total qui sont les suivantes :

- Codage algébrique
- Codage par affectation
- Codage relationnel
- Codage sous forme de matrices de paires d'objets nommés

Nous prendrons un exemple constitué de 4 objets que nous noterons par \mathcal{R}' pour illustrer ces différents formalismes. Il s'agira du classement suivant : $O^3 < O^1 < O^4 < O^2$ où $O^i < O^{i'}$ signifie que le rang de O^i est strictement inférieur au rang de $O^{i'}$.

5.1.1.1 Le codage algébrique

Ce codage est celui qui est classiquement rencontré en théorie des groupes. Il est donné par la relation ci-dessous.

Définition 5.1.2 Codage algébrique

Soit une relation d'ordre strict et total \mathcal{R} de N objets qui sont classés selon les rangs suivants : $D^1 < D^2 < \dots < D^N$. Le codage algébrique représentant \mathcal{R} est noté κ et est défini selon le formalisme suivant :

$$\kappa = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ \kappa(O^1) & \kappa(O^2) & \dots & \kappa(O^N) \end{pmatrix} \quad (5.1.1)$$

où $\kappa(O^i) \in \{D^1, \dots, D^N\}$, représente le rang auquel est affecté l'objet O^i selon \mathcal{R} . κ est une **bijection de \mathbb{O} dans \mathbb{D}** . Nous dirons que l'objet O^i est classé avant l'objet $O^{i'}$ selon κ , ce que nous noterons par abus de langage par $O^i <_{\kappa} O^{i'}$ ou $O^i < O^{i'}$ s'il n'y a pas d'ambiguïté, si nous avons $\kappa(O^i) < \kappa(O^{i'})$.

Propriété 5.1.1 Propriété structurelle des classements

κ étant notamment une injection, nous avons la propriété suivante :

$$\kappa(O^i) \neq \kappa(O^{i'}) \quad \forall i, i' = 1, \dots, N : i \neq i' \quad (5.1.2)$$

Autrement dit, il n'y a pas d'ex-aequo possible.

La représentation de l'ordre strict et total \mathcal{R}' selon le codage algébrique est la suivante :

$$\kappa' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^2 & D^4 & D^1 & D^3 \end{pmatrix}$$

5.1.1.2 Le codage par affectation

Définition 5.1.3 Codage par affectation

Soit un classement \mathcal{R} de N objets classés selon les rangs suivants $D^1 < D^2 < \dots < D^N$. Le codage par affectation de cette relation d'ordre strict et total consiste en une matrice carrée notée K de dimension N dont le terme général K_{ij} , $i, j = 1, \dots, N$; est donnée par la relation suivante :

$$K_{ij} = \begin{cases} 1 & \text{si } O^i \text{ est de rang } D^j \\ 0 & \text{sinon} \end{cases} \quad (5.1.3)$$

K sera appelée matrice d'affectation.

A l'instar des partitions, le codage par affectation peut être vue comme une représentation matricielle du codage algébrique.

Propriété 5.1.2 Propriétés de la matrice K

- $\sum_j K_{ij} = K_{i.} = 1 \quad \forall i = 1, \dots, N$
- $\sum_i K_{.j} = K_{.j} = 1 \quad \forall j = 1, \dots, N$
- $\sum_{i,j} K_{ij} = K_{..} = N$

La matrice d'affectation représentative de \mathcal{R}' est la suivante :

$$K' = \begin{matrix} & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

La structure des matrices d'affectation que nous introduisons ici, est celle des matrices de permutation que nous avons introduites page 35, sont semblables. Il existe toutefois, une nuance entre ces deux types de matrice : les matrices d'affectations croisent deux ensembles distincts tandis que les matrices de permutations croisent un seul et même ensemble. Nous distinguerons par conséquent ces deux types de matrices.

5.1.1.3 Le codage relationnel

Définition 5.1.4 Codage relationnel

Soit un classement \mathcal{R} de N objets classés selon les rangs ordonnés suivants $D^1 < D^2 < \dots < D^N$. Le codage relationnel de cette relation d'ordre strict et total consiste en une matrice carrée notée C de dimension N dont le terme général $C_{ii'}$ où $i, i' = 1, \dots, N$; est donné par la relation suivante :

$$C_{ii'} = \begin{cases} 1 & \text{si } O^i \text{ est classé avant } O^{i'} \text{ selon } \mathcal{R} \\ 0 & \text{sinon} \end{cases} \quad (5.1.4)$$

Nous donnons les propriétés caractéristiques du type de la relation qu'une matrice relationnelle de classement vérifie.

Propriété 5.1.3 Propriétés relationnelles de la matrice C

La matrice relationnelle vérifie les contraintes suivantes qui retranscrivent les propriétés relationnelles d'une relation d'ordre strict et total :

- Irréflexivité : $C_{ii} = 0 \quad \forall i = 1, \dots, N$
- Asymétrie : $C_{ii'} + C_{i'i} \leq 1 \quad \forall i, i' = 1, \dots, N : i \neq i'$
- Transitivité : $C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 \quad \forall i, i', i'' = 1, \dots, N$
- Totalité : $C_{ii'} + C_{i'i} \geq 1 \quad \forall i, i' = 1, \dots, N : i \neq i'$

Remarque 5.1.1 Asymétrie et totalité

Les contraintes d'asymétrie et de totalité peuvent être combinées pour donner un seul type de contraintes :

$$C_{ii'} + C_{i'i} = 1 \quad \forall i, i' = 1, \dots, N : i \neq i' \quad (5.1.5)$$

Nous donnons de plus, des propriétés relatives au classement codé que permet d'exprimer le codage relationnel.

Propriété 5.1.4 Propriétés structurelles de la matrice C

- $\sum_i C_{ii'} + 1 = C_{.i'} + 1 = \text{rang de l'objet } O^{i'} \text{ ou encore, } \kappa(O^{i'}) = D^{C_{.i'}+1} \quad \forall i' = 1, \dots, N$
- $N - \sum_{i'} C_{ii'} = N - C_{i.} = \text{rang de l'objet } O^i \text{ ou encore, } \kappa(O^i) = D^{N-C_{i.}} \quad \forall i = 1, \dots, N$
- $\sum_{i, i'} C_{ii'} = \frac{N(N-1)}{2}$

Démonstration

$C_{.i'}$ donne le nombre d'objets dont le classement est avant l'objet $O^{i'}$. Soit O^{i_1} l'objet classé en première position, nous avons alors $C_{.i_1} = 0$. De la même manière, soit O^{i_2} l'objet classé en deuxième position, nous avons alors $C_{.i_2} = 1$ et ainsi de suite. Nous voyons bien que $C_{.i'} + 1$ donne le rang de l'objet $O^{i'}$. De manière symétrique, $C_{i.}$ donne le nombre d'objets dont le classement est après l'objet O^i . Avec un même raisonnement, nous arrivons au résultat complémentaire du précédent.

□

A l'instar des partitions, le complémentaire d'un ordre strict et total est une relation binaire structurée⁶ : une relation d'ordre large, totale et sans *ex-aequo*⁷.

Propriété 5.1.5 Complémentaire d'une relation d'ordre strict et total

Soit un classement \mathcal{R} représenté par sa matrice relationnelle C . Soit \bar{C} la matrice relationnelle du complémentaire de \mathcal{R} défini par :

$$\bar{C} = U_N - C \quad (5.1.6)$$

où U_N est la matrice carrée de dimension N remplie de 1.

Alors, \bar{C} possède les propriétés d'une relation d'ordre large, total et sans *ex-aequo* :

6. ce qui n'est pas le cas en général et notamment pour les préordres
7. qui peut être aussi vue comme un ordre réflexif, asymétrique et total. En effet, la propriété d'asymétrie empêche les relations symétriques $C_{ii'} = C_{i'i} = 1$ nécessaires aux *ex-aequos*

- *Réflexivité* : $\bar{C}_{ii} = 1 \quad \forall i = 1, \dots, N$
- *Antisymétrie faible* : $\{\bar{C}_{ii'} + \bar{C}_{i'i} + \bar{C}_{ii''} - \bar{C}_{i'i''} \leq 2\} \wedge \{\bar{C}_{ii'} + \bar{C}_{i'i} - \bar{C}_{ii''} + \bar{C}_{i'i''} \leq 2\} \quad \forall i, i' = 1, \dots, N$
- *Totalité* : $\bar{C}_{ii'} + \bar{C}_{i'i} \geq 1 \quad \forall i, i' = 1, \dots, N : i \neq i'$
- *Pas d'ex-aequo* : $\nexists(i, i') : \{i \neq i'\} \wedge \{\bar{C}_{ii'} = \bar{C}_{i'i} = 1\}$

Si κ est le codage algébrique de \mathcal{R} alors nous noterons $\bar{\kappa}$ le codage algébrique de $\bar{\mathcal{R}}$:

$$\bar{\kappa} = \begin{pmatrix} O^1 & \cdots & O^i & \cdots & O^N \\ D^{C_{1,1}+1} & \cdots & D^{C_{i,1}+1} & \cdots & D^{C_{N,1}+1} \end{pmatrix} \quad (5.1.7)$$

Démonstration

Il s'agit d'un résultat bien connu mais que nous démontrons toutefois dans le contexte de l'AR.

Le résultat sur la réflexivité est immédiat.

La propriété de totalité provient de la propriété d'antisymétrie forte de C sur les paires $(O^i, O^{i'})$ tel que $i \neq i'$:

$$\begin{aligned} C_{ii'} + C_{i'i} \leq 1 &\Leftrightarrow -1 \leq -(C_{ii'} + C_{i'i}) \\ &\Leftrightarrow 1 \leq 2 - (C_{ii'} + C_{i'i}) \\ &\Leftrightarrow \bar{C}_{ii'} + \bar{C}_{i'i} \geq 1 \end{aligned}$$

De manière symétrique, \bar{C} possède la propriété d'antisymétrie forte du fait que C soit totale. Toutefois, nous savons que l'antisymétrie forte n'est pas compatible avec la réflexivité, c'est la raison pour laquelle elle n'est pas définie pour les couples $O^i = O^{i'}$. Nous préférons la remplacer par l'antisymétrie faible qui est complètement compatible avec à la fois l'antisymétrie forte et la réflexivité.

En ce qui concerne la propriété de n'avoir aucun ex-aequo, celle-ci provient du fait que $C_{ii'} = 1 - C_{i'i}$ avec $i \neq i'$. En effet, nous avons alors $\bar{C}_{ii'} = 1 - \bar{C}_{i'i}$ pour $i \neq i'$. Nous ne pouvons donc, avoir à la fois $\bar{C}_{ii'} = 1$ et $\bar{C}_{i'i} = 1$. Ainsi, au vu de la propriété d'antisymétrie faible, nous ne pouvons avoir d'ex-aequo.

□

Propriété 5.1.6 Propriétés structurelles de la matrice \bar{C}

- $\sum_i \bar{C}_{ii'} = \bar{C}_{.i'} = \text{rang de l'objet } O^{i'} \text{ pour l'ordre large, total et sans ex-aequo engendré par } \bar{C}$
 $\forall i' = 1, \dots, N$
- $N - \sum_{i'} \bar{C}_{ii'} + 1 = N - \bar{C}_{i.} + 1 = \text{rang de l'objet } O^i \text{ pour l'ordre large, total et sans ex-aequo engendré par } \bar{C} \quad \forall i = 1, \dots, N$
- $\sum_{i,i'} \bar{C}_{ii'} = \frac{N(N+1)}{2}$

De manière plus intuitive, le complémentaire de $\mathcal{R}' = O^3 < O^1 < O^4 < O^2$ est l'ordre large, total et sans ex-aequo $\bar{\mathcal{R}}' = O^3 \geq O^1 \geq O^4 \geq O^2$.

Propriété 5.1.7 Converse ou "transposée" d'une relation d'ordre strict et total

Soit un classement \mathcal{R} représenté par sa matrice relationnelle C . Notons par $\check{C} = {}^t C$ la matrice relationnelle transposée de C qui, dans le cadre de l'algèbre relationnelle, est aussi appelée la converse de C . Alors, \check{C} possède les propriétés d'une relation d'ordre strict et total.

Démonstration

L'irréflexivité et l'asymétrie sont immédiates. Pour ce qui est de la transitivité, nous avons $\forall i, i', i'' = 1, \dots, N$:

$$\begin{aligned} C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 &\Leftrightarrow C_{i''i'} + C_{i'i} - C_{i''i} \leq 1 \quad \text{en remplaçant } i \text{ par } i'' \text{ et vice-versa} \\ &\Leftrightarrow {}^t C_{i'i''} + {}^t C_{ii'} - {}^t C_{ii''} \leq 1 \end{aligned}$$

De manière générale, il est facile de voir que l'opération de transposition permet de conserver les propriétés relationnelles ce qui implique que la nature de la relation transposée ne change pas.

□

Propriété 5.1.8 Classement dual engendré par la converse ou "transposée" d'une matrice relationnelle

Soit \mathcal{R} une relation d'ordre strict et total représentée par κ son codage algébrique et C sa matrice relationnelle. Alors \check{C} est la matrice relationnelle représentative de l'ordre strict et total ci-dessous qui est symboliquement noté $\check{\kappa}$ ou $\check{\mathcal{R}}$:

$$\check{\kappa} = \begin{pmatrix} O^1 & \cdots & O^i & \cdots & O^N \\ D^{C_{1.+1}} & \cdots & D^{C_{i.+1}} & \cdots & D^{C_{N.+1}} \end{pmatrix} \quad (5.1.8)$$

Nous pouvons représenter l'ordre dual obtenu pour \mathcal{R}' . Nous avons, $\check{\mathcal{R}}' = O^3 > O^1 > O^4 > O^2$.

Remarque 5.1.2 Distinction entre $\bar{\kappa}$ et $\check{\kappa}$

Nous avons en terme algébrique $\bar{\kappa} = \check{\kappa}$. Cependant, la première est réflexive et faiblement antisymétrique tandis que la seconde est irréflexive et asymétrique. En terme relationnel, la nature des ordres engendrés n'est pas la même ce qui implique que les matrices relationnelles respectives sont différentes⁸.

Nous verrons dans le chapitre 6, dans lequel nous étudions les relations d'ordre large, qu'il est utile de distinguer ordre complémentaire et ordre transposé ou dual. Les deux types d'opérations unaires correspondants, représentent deux types d'ordres duaux dont les interprétations sont distinctes. A bien des égards, nous verrons que l'ordre donné par l'opération \check{C} est davantage le dual naturel d'une relation d'ordre que l'ordre complémentaire donné par \bar{C} .

Nous donnons l'exemple de la matrice représentative de la relation \mathcal{R}' dans son codage relationnel ainsi que son complémentaire et sa transposée :

$$C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix} ; \quad \bar{C}' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \end{pmatrix} \end{matrix} ;$$

$$\check{C}' = {}^t C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

Dans le cas particulier des ordres stricts et totaux, complémentaire et converse sont liées : seuls les termes de la diagonale semble les différencier. Nous verrons en chapitre 6 dans lequel nous étudions les ordres larges avec *ex-aequos* qu'il est dans ce cas important de distinguer ces deux opérations unaires étant donné que le complémentaire d'un ordre large n'est pas en général un ordre ce qui n'est pas le cas de sa converse.

5.1.1.4 Lien entre le codage par affectation et le codage relationnel

Similairement aux relations d'équivalence, il est possible de passer du codage d'affectation d'une relation d'ordre strict et total, au codage relationnel de cette dernière par l'intermédiaire

8. il s'agit d'une propriété générale de l'AR dont le formalisme, qui permet de représenter les informations atomiques par paires d'objets, permet de modéliser tout type de relation ce qui n'est pas forcément le cas des autres types de codage

d'une matrice de passage R_N . Cette dernière est une matrice relationnelle codant l'ordre sur les rangs permettant de classer les objets. Cet aspect fut mis en évidence par S. Ghashghaie dans [Gha90].

Propriété 5.1.9 Du codage par affectation au codage relationnel

Soit un classement \mathcal{R} et soit K sa matrice d'affectation associée. A partir de K , nous obtenons la matrice relationnelle C de \mathcal{R} , en appliquant la relation suivante :

$$C = K \cdot R_N \cdot {}^t K \quad (5.1.9)$$

où R_N est une matrice relationnelle de passage qui code la relation d'ordre sur les éléments de l'ensemble \mathbb{D} . Dans le cas où les éléments de \mathbb{D} sont ordonnés de la manière suivante, $D^1 < D^2 < \dots < D^N$. Le terme général de R_N est ainsi défini par, $\forall j, j' = 1, \dots, N$:

$$[R_N]_{jj'} = \begin{cases} 1 & \text{si } D^j < D^{j'} \text{ c'est à dire si } j < j' \\ 0 & \text{sinon} \end{cases} \quad (5.1.10)$$

R_N représente la matrice de passage standard dans le cas des ordres stricts pour lesquels les rangs sont ordonnés de la manière suivante $D^1 < D^2 < \dots < D^N$.

Contrairement aux partitions, où plusieurs matrices d'affectation peuvent représenter la même matrice relationnelle, les classements n'ont pas de classe d'équivalence. **A chaque matrice d'affectation correspond un unique ordre strict et total. De ce fait, nous pouvons déterminer à partir d'une matrice relationnelle codant un classement sa matrice d'affectation associée.**

Remarque 5.1.3 Du codage par affectation d'un ordre dual au codage relationnel

Soit un classement \mathcal{R} et soit K sa matrice d'affectation associée. A partir de K , nous obtenons la matrice relationnelle \check{C} de $\check{\mathcal{R}}$, en appliquant la relation suivante :

$$\check{C} = K \cdot {}^t R_N \cdot {}^t K \quad (5.1.11)$$

Propriété 5.1.10 Du codage relationnel au codage par affectation

Soit une relation d'ordre strict et total \mathcal{R} et soit C sa matrice relationnelle associée. A partir de C , nous pouvons déterminer la matrice d'affectation K de \mathcal{R} , en appliquant la relation suivante :

$$K_{ij} = \frac{\prod_{i' \neq i} (C_{i'} + 1 - j)}{\prod_{j' \neq j} (j' - j)} \quad \forall i, j = 1, \dots, N \quad (5.1.12)$$

où $C_{.i'} = \sum_i C_{ii'}$

Démonstration

Nous avons $C_{.i'} + 1 = \text{rang de l'objet } O^{i'}$ selon prop-5.1.6.

Comme chaque rang est affecté à un et un seul objet, si l'indice de rang de O^i est j alors $\prod_{i' \neq i} (C_{.i'} + 1 - j) \neq 0$ et vaut exactement $\prod_{j' \neq j} (j' - j)$. Dans ce cas, nous avons bien $K_{ij} = 1$.

Si l'indice de rang de O^i n'est pas j , alors il existe un unique objet $O^{i^*} \neq O^i$ pour lequel l'indice de rang j a été affecté. Dans ce cas, nous avons bien $\prod_{i' \neq i} (C_{.i'} + 1 - j) = K_{ij} = 0$

□

5.1.1.5 Le codage sous forme de matrices de paires d'objets nommés

Nous introduisons ici un nouveau codage qui sera à la base de la modélisation des tresses que nous verrons au chapitre 7.

Ce codage consiste en une matrice présentant une structure triangulaire dont les éléments sont des "paires d'objets nommés". La disposition de ces paires à l'intérieur de la matrice possède des propriétés de correspondance avec les rangs des objets de chaque paire.

Définition 5.1.5 *Le codage sous forme de matrice de paires d'objets nommés*

Soit un classement \mathcal{R} de N objets classés selon les rangs suivants $D^1 < D^2 < \dots < D^N$ et représentés par son codage algébrique :

$$\kappa = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ \kappa(O^1) & \kappa(O^2) & \dots & \kappa(O^N) \end{pmatrix}$$

où $\kappa(O^i) \in \{D^1, \dots, D^N\}, \forall i = 1, \dots, N$.

Le codage sous forme de matrice de paires d'objets nommés de cette relation d'ordre strict et total consiste en une matrice carrée notée L de dimension $N - 1$ dont le terme général $L_{jj'}$ où $j = 1, \dots, N - 1$ et $j' = j, \dots, N - 1$, est donné par la relation suivante

$$L_{jj'} = \left[(\kappa)^{-1}(D^j), (\kappa)^{-1}(D^{j'+1}) \right] \quad \forall j = 1, \dots, N - 1; \forall j' = j, \dots, N - 1 \quad (5.1.13)$$

où $(\kappa)^{-1}(D^j) \in \{O^1, \dots, O^N\}, \forall j = 1, \dots, N$.

Nous emploierons les notations suivantes :

- $[L_{jj'}]_1 = (\kappa)^{-1}(D^j) =$ premier terme de la paire $L_{jj'}$
- $[L_{jj'}]_2 = (\kappa)^{-1}(D^{j'+1}) =$ deuxième terme de la paire $L_{jj'}$

Propriété 5.1.11 *Propriétés structurelles de la matrice L*

- $[L_{jj'}]_1 =$ objet rangé en j -ème position $\forall j = 1, \dots, N - 1; \forall j' = j, \dots, N - 1$
- $[L_{jj'}]_2 =$ objet rangé en $(j' + 1)$ -ème position $\forall j = 1, \dots, N - 1; \forall j' = j, \dots, N - 1$
- l'ensemble des paires $\{L_{jj'}; j = 1, \dots, N - 1; j' = j, \dots, N\}$ représente les arcs du graphe orienté représentatif de la relation d'ordre strict et total $G(\mathcal{R})$. Autrement dit, si C est la matrice relationnelle de \mathcal{R} alors nous avons la relation suivante :

$$C_{ii'} = 1 \Leftrightarrow \exists!(j, j'), j = 1, \dots, N - 1, j' = j, \dots, N - 1 : L_{jj'} = [O^i, O^{i'}] \quad (5.1.14)$$

où $j = C_{.i} + 1$ et $j' = C_{.i'}$

A titre illustratif, nous donnons la matrice de paires d'objets nommés représentative de la relation \mathcal{R}' :

$$L' = \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ & [O^1, O^4] & [O^1, O^2] \\ & & [O^4, O^2] \end{pmatrix}$$

5.1.2 Rappels sur le groupe des permutations

5.1.2.1 Quelques propriétés du groupe des permutations

Nous rappelons dans ce paragraphe quelques définitions et propriétés du groupe des permutations d'un ensemble d'éléments $\mathbb{O} = \{O^1, \dots, O^N\}$. En effet, il existe une correspondance biunivoque entre une permutation et un ordre strict et total comme nous le précisons par la suite et en déf-5.1.7. Il est donc utile de rappeler les propriétés du groupe des permutations pour mieux appréhender les structures algébriques des ordres stricts et totaux.

Définition 5.1.6 Groupe des permutations de \mathbb{O}

L'ensemble des permutations des éléments d'un ensemble \mathbb{O} , est l'ensemble des bijections de \mathbb{O} dans lui-même. Nous noterons une permutation π de la manière suivante :

$$\pi = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ \pi(O^1) & \pi(O^2) & \dots & \pi(O^N) \end{pmatrix} \quad (5.1.15)$$

où $\pi(O^i) \in \{O^1, \dots, O^N\}$.

L'ensemble des permutations muni de la composition d'applications \circ forme un groupe. Dans le cadre du groupe des permutations, **la composition de π' par π'' se fait par la droite.**

$$(\pi' \circ \pi'')(O^i) = \pi'(\pi''(O^i)) \quad \forall i = 1, \dots, N \quad (5.1.16)$$

L'élément neutre du groupe noté π^e est l'application identité pour laquelle : $\pi^e(O^i) = O^i, \forall i = 1, \dots, N$. Par ailleurs, le groupe des permutations de l'ensemble \mathbb{O} sera noté $\mathbb{S}(\mathbb{O})$.

Remarque 5.1.4 Permutations et transferts doubles

Le groupe des permutations d'un ensemble d'éléments est identique au groupe des opérations de transferts doubles définis sur cet ensemble d'éléments et que nous avons présentés lors de l'étude des actions sur les partitions au paragraphe 2.1.2.3 page 43.

Les permutations étant des bijections d'un ensemble dans lui-même, elle peuvent hériter des représentations que nous avons étudiées pour les ordres stricts et totaux que nous rappelons être des bijections d'un ensemble d'objets \mathbb{O} dans l'ensemble des rangs \mathbb{D} . Dans ce cas, nous remplaçons l'ensemble des rangs par l'ensemble des éléments de \mathbb{O} pour lesquels, nous considérons en plus l'ordre suivant $O^1 < O^2 < \dots < O^N$ qui est l'ordre initial pris par défaut dans l'ensemble de cette sous-section. Dans le cas d'un codage par affectation d'une permutation, nous noterons par P la matrice associée à π que nous appellerons aussi matrice de permutation. Nous avons alors $\forall i, i' = 1, \dots, N$:

$$P_{ii'} = \begin{cases} 1 & \text{si } \pi(O^i) = O^{i'} \\ 0 & \text{sinon} \end{cases} \quad (5.1.17)$$

Dans le cas d'une matrice relationnelle C représentative de la permutation π , il existe une double interprétation⁹ qui est d'une part celle d'une liste, d'autre part, celle d'un classement. Afin d'illustrer ces propos nous prenons l'exemple de la permutation suivante :

$$\pi' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^4 & O^1 & O^3 \end{pmatrix} \quad (5.1.18)$$

Si nous considérons l'interprétation classique d'une permutation en terme de liste, π' représente l'élément O^2, O^4, O^1, O^3 . Par contre, si nous considérons, comme dans le paragraphe précédent, l'interprétation en terme de classement pour lequel nous supposons l'ordre initial $O^1 < O^2 < O^3 < O^4$, π' représente alors l'élément $O^3 < O^1 < O^4 < O^2$. La distinction entre ces deux aspects est considérée dans la convention suivante qui montre la correspondance bi-univoque entre une permutation et le classement qui lui est associé lorsque nous considérons la représentation relationnelle.

Définition 5.1.7 Permutation en tant que classement et en tant que liste dans le cadre des matrices relationnelles

Soit π une permutation de $\mathbb{S}(\mathbb{O})$. Nous noterons par C la matrice relationnelle représentant le **classement** engendré par π pour lequel nous considérons l'ordre initial $O^1 < \dots < O^N$. Nous avons alors :

$$C = P \cdot R_N \cdot {}^t P \quad (5.1.19)$$

9. selon que l'on considère un ordre initial sur les éléments de \mathbb{O} ou non

où R_N est la matrice relationnelle¹⁰ codant l'ordre $O^1 < \dots < O^N$.

Soit π^{-1} la permutation inverse de π . Nous noterons par C^- la matrice relationnelle représentant le classement engendré par π^{-1} et qui correspond à l'ordre engendré par la liste représentée par π . Nous avons alors :

$$C^- = {}^tP \cdot R_N \cdot P \quad (5.1.20)$$

Pour illustrer la définition précédente, nous reprenons l'exemple ci-dessus pour lequel nous avons :

$$C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \end{matrix} ; \quad C'^- = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

Par ailleurs, nous pouvons aussi vérifier sur cet exemple la relation suivante issue des propriétés prop-5.1.6 permettant, à l'inverse, de passer de l'écriture relationnelle C' à l'écriture algébrique π' :

$$\pi' = \left(\begin{matrix} O^1 & \dots & O^i & \dots & O^N \\ O^{C'.1+1} & \dots & O^{C'.i+1} & \dots & O^{C'.N+1} \end{matrix} \right) \quad (5.1.21)$$

Remarque 5.1.5 Groupe symétrique

Lorsque l'ensemble ordonné étudié est l'ensemble des N premiers nombres entiers naturels, $\mathbb{N}_N = \{1, 2, \dots, N\}$, le groupe des permutations est aussi appelé groupe symétrique. Celui-ci est noté \mathbb{S}_N .

Dans le contexte dans lequel nous nous plaçons, il peut donc y avoir deux interprétations des permutations. En ce qui concerne le formalisme relationnel, nous distinguons le classement de la liste selon que nous utilisons respectivement les notations C et C^- . Il est en fait, utile d'étudier les codages relationnels d'une permutation interprétée telle un classement ou une liste étant donné que ceux-ci permettent de caractériser certaines des propriétés du groupe des permutations. En particulier, nous verrons que les paires d'inversions d'une permutation ainsi que la distance de Kendall entre deux listes, peuvent s'exprimer par l'intermédiaire des codages relationnels.

Nous donnons à présent quelques définitions et propriétés du groupe des permutations que nous emploierons par la suite. Nous donnons donc, lorsque cela sera possible, l'interprétation des propriétés des permutations dans le codage relationnel.

Propriété 5.1.12 Inverse d'une permutation

Soit π une permutation de $\mathbb{S}(\mathbb{O})$ représentée par sa matrice de permutation P soit par sa matrice relationnelle C . Les représentations par affectation et relationnelle de la permutation inverse π^{-1} de π sont respectivement notées P^{-1} et C^- et sont obtenues de la manière suivante :

- $P^{-1} = {}^tP$
- $C^- = ({}^tP)^2 \cdot C \cdot (P)^2$

Démonstration

Le résultat relatif au codage par affectation est connu.

En ce qui concerne le codage relationnel, nous avons :

10. cf (5.1.10) mais en considérant l'ensemble ordonné \mathbb{O}

$$\begin{aligned}
 {}^tP \cdot {}^tP \cdot C \cdot P \cdot P &= {}^tP \cdot {}^tP \cdot (P \cdot R_N \cdot {}^tP) \cdot P \cdot P \\
 &= {}^tP \cdot R_N \cdot P \quad \text{car } P \cdot {}^tP = I_N \\
 &= C^- \quad \text{car } {}^tP \text{ est la matrice de permutation associée à } \pi^{-1}
 \end{aligned}$$

□

Remarque 5.1.6 *Ordre dual et permutation inverse*

Nous pouvons représenter les permutations par des matrices relationnelles d'une manière qui est identique à la représentation relationnelle d'un ordre strict et total. Nous précisons néanmoins, que les opérateurs unaires de complémentarité, $\bar{\pi}$, et de converse, $\check{\pi}$, que nous avons définies dans la section précédente et qui donnent des ordres duaux à un ordre π , ne doivent pas être confondues avec la permutation inverse, π^{-1} , d'une permutation π . Nous avons le tableau récapitulatif suivant où π est une permutation dont l'ordre initial sur les objets est $O^1 < \dots < O^N$:

Codage algébrique	Codage par affectation	Codage relationnel (π en terme de classement)	(5.1.22)
π	P	$C = P \cdot R_N \cdot {}^tP$	
π^{-1}	$P^{-1} = {}^tP$	$C^- = {}^tP \cdot R_N \cdot P$	
$\check{\pi}$	$\check{P} \neq {}^tP$	$\check{C} = {}^tC = P \cdot {}^tR_N \cdot {}^tP$	
$\bar{\pi}$	$\bar{P} = {}^tP$	$\bar{C} = U_N - C = {}^tC + I_N$	

où R_N est donnée par (5.1.10) dans le cas où les éléments sont les $O^i, i = 1, \dots, N$. Nous faisons remarquer que \check{C} et \bar{C} sont des opérations unaires définies pour des matrices relationnelles. Elles ne définissent aucunement une quelconque opération sur des codages autres que ces dernières. Elles sont néanmoins utilisées symboliquement pour le codage algébrique et par affectation.

Définition 5.1.8 *Point fixe et point mobile d'une permutation*

Un point fixe ou un invariant de π est un élément O^i tel que : $\pi(O^i) = O^i$. A contrario, un point mobile de π est un élément O^i pour lequel : $\pi(O^i) \neq O^i$.

Si nous considérons les représentations par affectation et relationnelle, nous voyons respectivement que O^i est un invariant de π si :

- $P_{ii} = 1$
- $\sum_{i'} C_{i'i} + 1 = C_{.i} + 1 = i$

Nous définissons maintenant, la notion de cycle dans une permutation d'éléments. L'ensemble des propriétés rappelées interviendront notamment dans le chapitre 6 lors de la recherche de conditions nécessaires et suffisantes de l'acyclicité d'une relation d'ordre collective.

Définition 5.1.9 *k-cycle*

Un k -cycle γ de \mathbb{O} avec $2 \leq k \leq N$, est une permutation dont les éléments mobiles, en nombre k , sont de la forme :

$$O^i ; \gamma(O^i) ; (\gamma(O^i))^2 ; \dots ; (\gamma(O^i))^{k-1} \tag{5.1.23}$$

et

$$O^i = (\gamma(O^i))^k \quad (5.1.24)$$

où O^i est un élément mobile de γ et où la puissance est au sens de la composition d'applications \circ .
Si nous avons $O^i = O^{i_1}; \gamma(O^{i_1}) = O^{i_2}; \dots; (\gamma(O^{i_1}))^{k-1} = (\gamma(O^{i_{k-1}})) = O^{i_k}$ alors on note habituellement un tel cycle par :

$$\gamma = (O^{i_1}, O^{i_2}, \dots, O^{i_k}) \quad (5.1.25)$$

Définition 5.1.10 Transposition

Une transposition de \mathbb{O} est un 2-cycle de cet ensemble. Nous noterons par $\pi^{ii'}$ le 2-cycle $(O^i, O^{i'})$.

Supposons que les éléments sont ordonnés et rangés de la manière suivante : $O^1 < O^2 < \dots < O^N$. Parmi l'ensemble des 2-cycles, nous distinguons alors un sous-ensemble particulier constitué de $N - 1$ éléments qui est l'ensemble générateur de $\mathbb{S}(\mathbb{O})$.

$$\pi^i = \pi^{i,i+1} = (O^i, O^{i+1}) \quad \forall i = 1, \dots, N - 1 \quad (5.1.26)$$

Ce sous-ensemble sera noté $\mathbb{S}^g(\mathbb{O})$ et constitue l'ensemble des générateurs du groupe $\mathbb{S}(\mathbb{O})$. Ces générateurs seront aussi nommés transpositions d'éléments juxtaposés ou transpositions génératrices ou encore transpositions d'ordre 1¹¹.

Nous donnons un exemple afin d'illustrer l'emploi du terme "transpositions d'éléments juxtaposés" :

$$\pi' \circ \pi^2 = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^4 & O^1 & O^3 \end{pmatrix} \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^1 & O^3 & O^2 & O^4 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^1 & O^4 & O^3 \end{pmatrix}$$

Nous voyons dans cet exemple que composer π' par π^2 consiste à permuter les éléments $\pi'(O^2)$ et $\pi'(O^3)$: la liste que code π' qui est O^2, O^4, O^1, O^3 , devient, sous l'action de π^2 la liste O^2, O^1, O^4, O^3 . Autrement dit, la composition de π par π^2 permet de transposer les éléments de la liste codée par π qui sont en position 2 et en position 3.

Propriété 5.1.13 Décomposition d'une permutation en produit de cycles

- Toute permutation $\pi \in \mathbb{S}(\mathbb{O})$ différente de l'identité peut s'écrire sous la forme d'un produit de cycles dont les supports sont disjoints, cette décomposition étant unique à l'ordre près
- Tout cycle est un produit de transpositions
- En conséquence, toute permutation est un produit de transpositions

Démonstration

La démonstration de ces propriétés peut être trouvée dans le livre de L. Comtet [Com70] qui développe largement ce domaine.

□

Propriété 5.1.14 Décomposition d'une transposition $\pi^{ii'} \in \mathbb{S}(\mathbb{O})$ en un produit de générateurs $\pi^i \in \mathbb{S}^g(\mathbb{O})$

Soit $\pi^{ii'} \in \mathbb{S}(\mathbb{O})$ un 2-cycle et soit $\pi^i = \pi^{i,i+1}$, $i = 1, \dots, N - 1$, le sous-ensemble générateurs $\mathbb{S}^g(\mathbb{O})$.

11. cf aussi déf-5.2.1 page 171

Nous pouvons exprimer tout élément $\pi^{ii'}$ comme étant le produit de transpositions génératrices de la manière suivante. $\forall i, i' : 1, \dots, N - 1 : i < i'$ nous avons :

$$\pi^{ii'} = \pi^i \circ \pi^{i+1} \circ \dots \circ \pi^{i'-1} \circ \pi^{i'-2} \circ \dots \circ \pi^i \quad (5.1.27)$$

$$\pi^{i'i} = \pi^{i'-1} \circ \pi^{i'-2} \circ \dots \circ \pi^i \circ \pi^{i+1} \circ \dots \circ \pi^{i'-1} \quad (5.1.28)$$

L'ensemble des propriétés de décompositions d'une permutation montre bien que les éléments $\pi^i = \pi^{i,i+1}$ sont générateurs du groupe des permutations puisque toute permutation se ramène à un produit de ces générateurs.

Les générateurs du groupe des permutations possèdent des propriétés appelées relations. En théorie des groupes, de telles structures possèdent un formalisme particulier appelé présentation d'un groupe. La **présentation du groupe des permutations** est donnée par la propriété suivante.

Propriété 5.1.15 Présentation du groupe des permutations

Soit $\mathbb{S}^g(\mathbb{O})$ l'ensemble des générateurs du groupe des permutations $\mathbb{S}(\mathbb{O})$. Nous avons alors la présentation suivante du groupe des permutations :

$$\left\langle \pi^1, \dots, \pi^{N-1} \left| \begin{array}{l} \pi^i \circ \pi^i = \pi^\epsilon \quad \forall i = 1, \dots, N - 1 \\ \pi^i \circ \pi^{i'} = \pi^{i'} \circ \pi^i \quad \forall i, i' = 1, \dots, N - 1 : |i - i'| > 1 \\ \pi^i \circ \pi^{i'} \circ \pi^i = \pi^{i'} \circ \pi^i \circ \pi^{i'} \quad \forall i, i' = 1, \dots, N - 1 : |i - i'| = 1 \end{array} \right. \right\rangle \quad (5.1.29)$$

Ci-dessous, nous supposons de plus que les éléments possèdent l'ordre suivant : $O^1 < O^2 < \dots < O^N$. Nous rappelons la définition de l'ensemble des inversions d'une permutation. Ce sont des ensembles de paires d'objets. Nous montrons qu'ils peuvent être déterminés à l'aide des matrices relationnelles et des opérations de l'algèbre relationnelle ¹².

Définition 5.1.11 Ensemble des inversions d'une permutation

Soit π une permutation de $\mathbb{S}(\mathbb{O})$ dont les éléments sont initialement ordonnés de la manière suivante : $O^1 < O^2 < \dots < O^N$. Un couple d'objets $(O^i, O^{i'})$ est une inversion de π s'il vérifie la propriété suivante :

$$\{O^i < O^{i'}\} \wedge \{\pi(O^{i'}) < \pi(O^i)\} \quad (5.1.30)$$

Propriété 5.1.16 Matrice de comparaisons par paires de l'ensemble des inversions d'une permutation

Soit π une permutation représentée par sa matrice de permutation P et par sa matrice relationnelle C . L'ensemble des inversions de π peut être représentée par une matrice de comparaisons par paires notée I donnée par :

$$I = R_N \otimes \underbrace{(P \cdot {}^t R_N \cdot {}^t P)}_C \quad (5.1.31)$$

où R_N est la matrice relationnelle de passage standard codant l'ordre initial $O^1 < O^2 < \dots < O^N$ qui est définie par (5.1.10) et \otimes est l'intersection relationnelle ¹³. Nous nommerons cette matrice de comparaisons par paires, **matrice des inversions d'une permutation**.

Démonstration

R_N représente l'ensemble des paires telles que $O^i < O^{i'}$.

12. cf page 116
13. cf déf-4.1.1 page 115

$P \cdot R_N \cdot {}^tP$ représente l'action de la permutation π sur une matrice relationnelle codant ici l'ordre $O^1 < O^2 < \dots < O^N$. Nous avons donc les paires $(O^i, O^{i'})$ pour lesquels, $\pi(O^i) < \pi(O^{i'})$. Ainsi, par sa transposée ${}^t(P \cdot R_N \cdot {}^tP)$ qui vaut aussi $P \cdot {}^tR_N \cdot {}^tP$, nous obtenons l'ensemble des paires pour lesquelles $\pi(O^i) > \pi(O^{i'})$ qui est donc l'ordre dual du classement engendré par la permutation ¹⁴. L'intersection relationnelle de ces deux matrices nous donne ainsi l'ensemble des inversions de π .
□

Afin d'illustrer les définitions précédentes, nous étudions rapidement la permutation π'' suivante :

$$\pi'' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^4 & O^3 & O^1 \end{pmatrix} \quad (5.1.32)$$

O^3 est l'unique point fixe de π'' . (O^1, O^2, O^4) est un 3-cycles. Ainsi, π'' peut aussi être écrite selon sa décomposition en cycles : $\pi'' = (O^1, O^2, O^4)(O^3)$.

La matrice des inversions de π'' est alors la suivante :

$$\begin{aligned} I'' &= R_N \otimes (P'' \cdot {}^tR_N \cdot {}^tP'') \\ &= \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \otimes \left(\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix} \right) \\ &= \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{aligned}$$

Nous avons donc $\mathbb{I}(\pi'') = \{(O^1, O^4); (O^2, O^3); (O^2, O^4); (O^3, O^4)\}$.

Remarque 5.1.7 Interprétations des paires d'inversions

Dans [ECH⁺92], W.P. Thurston montre que les paires d'inversions $(O^i, O^{i'})$ d'une permutation π , sont les objets dont les images ¹⁵ par la permutation sont "juxtaposées" et dont les transpositions permettent de passer successivement de π à π^ϵ .

Dans l'exemple ci-dessus, nous pouvons donc vérifier que nous avons : $\pi'' \circ (O^1, O^4) \circ (O^2, O^3) \circ (O^2, O^4) \circ (O^3, O^4) = \pi^\epsilon$. Par ailleurs, les paires d'inversions $(O^i, O^{i'})$ sélectionnées à chaque étape sont telles que les images $\pi''(O^i)$ et $\pi''(O^{i'})$ sont juxtaposées au sens de l'ordre initial $O^1 < \dots < O^N$. Nous avons par exemple :

1. $\pi''(O^1) = O^2$ et $\pi''(O^4) = O^1$ où O^2 et O^1 sont juxtaposés
2. $(\pi \circ (O^1, O^4))(O^2) = O^4$ et $(\pi \circ (O^1, O^4))(O^3) = O^3$ où O^4 et O^3 sont juxtaposés
3. ...

5.1.2.2 Le treillis des permutations et le graphe de Cayley des permutations

Comme pour les relations d'équivalence, il existe un ordre partiel sur les permutations. Celui-ci est appelé ordre de Bruhat.

14. cf déf-5.1.7

15. ie les rangs

Définition 5.1.12 *Ordre partiel de Bruhat sur l'ensemble des permutations*

Nous dirons que la permutation π^k est inférieure ou égale à la permutation π^l au sens de Bruhat, ce que nous noterons par $\pi^k \preceq \pi^l$, si nous pouvons obtenir π^l à partir de π^k en appliquant à cette dernière une succession de transpositions¹⁶ $\pi^{ii'}$ = $(O^i, O^{i'})$ tel que $O^i < O^{i'}$ au sens de l'ordre initial¹⁷.

En termes ensemblistes, nous avons la propriété suivante [ECH⁺92] :

$$\pi^k \preceq \pi^l \Leftrightarrow \mathbb{I}(\pi^k) \subseteq \mathbb{I}(\pi^l) \tag{5.1.33}$$

où $\mathbb{I}(\pi^k)$ et $\mathbb{I}(\pi^l)$ sont les ensembles des inversions de π^k et π^l .

En termes de matrices des inversions cela se traduit par la relation ci-dessous :

$$\pi^k \preceq \pi^l \Leftrightarrow I_{ii'}^k \leq I_{ii'}^l \quad \forall i, i' = 1, \dots, N \tag{5.1.34}$$

où I^k et I^l sont les matrices de comparaisons par paires des inversions de π^k et π^l .

Nous renvoyons le lecteur à [FH69] et au chapitre 9 de [ECH⁺92] pour de plus amples détails sur les définitions et les liens entre l'ordre partiel et les ensembles d'inversions.

Cet ordre permet de conférer au groupe des permutations une structure de treillis.

Propriété 5.1.17 *Structure de treillis*

L'ensemble des permutations d'un ensemble d'objets, muni de la relation d'ordre partiel de Bruhat \preceq , respecte une structure de treillis pour laquelle :

- La borne supérieure est la permutation converse de l'identité :

$$\tilde{\pi}^\epsilon = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ O^N & O^{N-1} & \dots & O^1 \end{pmatrix} \tag{5.1.35}$$

- La borne inférieure est la permutation identité notée π^ϵ :

$$\pi^\epsilon = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ O^1 & O^2 & \dots & O^N \end{pmatrix} \tag{5.1.36}$$

Nous donnons dans la figure fig-5.1 la représentation du treillis des permutations de l'ensemble d'objets $\{O^1, O^2, O^3\}$. Une permutation π^l pointe vers une permutation π^k si $\pi^k \preceq \pi^l$.

Le groupe de permutations d'un ensemble d'objets peut aussi être appréhender sous l'angle de la géométrie combinatoire au travers de son graphe de Cayley dont nous rappelons la définition ci-dessous.

Définition 5.1.13 *Graphe de Cayley du groupe des permutations*

Soit $\mathbb{S}(\mathbb{O})$ le groupe des permutations définies sur l'ensemble \mathbb{O} et soit $\mathbb{S}^g(\mathbb{O})$ le sous-ensemble $\{\pi^1, \dots, \pi^{N-1}\}$ des transpositions génératrices du groupe $\mathbb{S}(\mathbb{O})$. Nous pouvons alors représenter ce dernier par un graphe non orienté appelé graphe de Cayley dont les noeuds sont les éléments de $\mathbb{S}(\mathbb{O})$ et les arcs les éléments de $\mathbb{S}^g(\mathbb{O})$. Deux permutations π^k et π^l seront reliées à condition qu'il existe un générateur $\pi^i = (O^i, O^{i+1})$, $i \in \{1, \dots, N - 1\}$ tel que :

$$\{\pi^k = \pi^l \circ \pi^i\} \vee \{\pi^l = \pi^k \circ \pi^i\} \tag{5.1.37}$$

car $\pi^i = (\pi^i)^{-1}$.

16. agissant sur les objets : cf page 169 et section 5.2

17. ie si $i < i'$

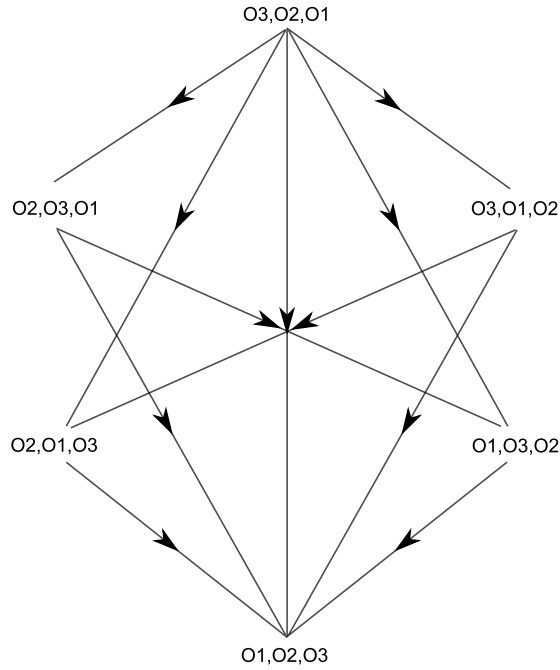


FIGURE 5.1 – Exemple de treillis de permutations

Contrairement à la structure de treillis, le graphe de Cayley du groupe des permutations d'un ensemble d'objets est non orienté et cyclique. Il définit d'ailleurs, un **polytope appelé permutoèdre**.

Dans [ECH⁺92], W.P. Thurston fait le lien entre les ensembles d'inversions et la longueur¹⁸ d'une permutation au sens du graphe de Cayley.

De manière plus générale, la distance entre deux permutations représentées dans le contexte du graphe de Cayley, est donnée par la distance de Kendall et peut être modélisée dans le formalisme relationnel [MM78].

Propriété 5.1.18 Formalisme relationnel de la distance de Kendall sur le graphe de Cayley

La distance de Kendall entre deux permutations π^k et π^l notée $d_K(\pi^k, \pi^l)$ est donnée par le nombre minimal de transpositions d'éléments juxtaposés de $\mathbb{S}^g(\mathbb{O})$ permettant d'aller d'une permutation à une autre¹⁹.

Soient C^{k-} et C^{l-} les matrices relationnelles associées aux permutations $(\pi^k)^{-1}$ et $(\pi^l)^{-1}$ alors nous avons :

$$d_K(\pi^k, \pi^l) = d_K(C^{k-}, C^{l-}) = \frac{1}{2} \sum_{i=1}^N \sum_{i'=1}^N (C_{ii'}^{k-} - C_{ii'}^{l-})^2 \tag{5.1.38}$$

Démonstration

Le rapprochement entre la distance de Kendall entre deux listes et de la longueur de la géodésique entre deux permutations en fonction des matrices relationnelles est montré par J.F. Marcotorchino et P. Michaud dans [MM78]. Nous précisons que notre codage des permutations par des matrices relationnelles implique que nous raisonnons sur les permutations inverses des permutations étudiées. En effet, le codage

18. nombre de transpositions génératrices séparant une permutation de la permutation identité

19. longueur de la géodésique séparant les deux permutations sur le graphe de Cayley

de la permutation π'' par exemple donnée par (5.1.32) code en fait le classement engendré par la permutation qui sous forme de liste ordonnée est $O^4 < O^1 < O^3 < O^2$. Or la permutation π'' , représente la liste O^2, O^4, O^3, O^1 . Celle-ci, dans le codage relationnel correspond donc au "classement" engendré par la liste, $O^2 < O^4 < O^3 < O^1$, et qui est donné par la matrice relationnelle de $(\pi'')^{-1}$. Ceci explique pourquoi, par souci de rigueur, apparaissent les matrices relationnelles des permutations inverses²⁰.
□

Nous donnons en exemple le graphe de Cayley de l'ensemble $\{O^1, O^2, O^3, O^4\}$ dans la figure fig-5.2. Nous représentons les différentes permutations sous la forme d'une liste. Prenons l'exemple de l'élément O^3, O^1, O^4, O^2 , celui correspond à la permutation suivante :

$$\begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^3 & O^1 & O^4 & O^2 \end{pmatrix}$$

Par ailleurs, les générateurs $\{\pi^1, \pi^2, \pi^3\}$ seront notés par $\{P^1, P^2, P^3\}$ et sont représentés dans la figure par des arcs de formes distinctes.

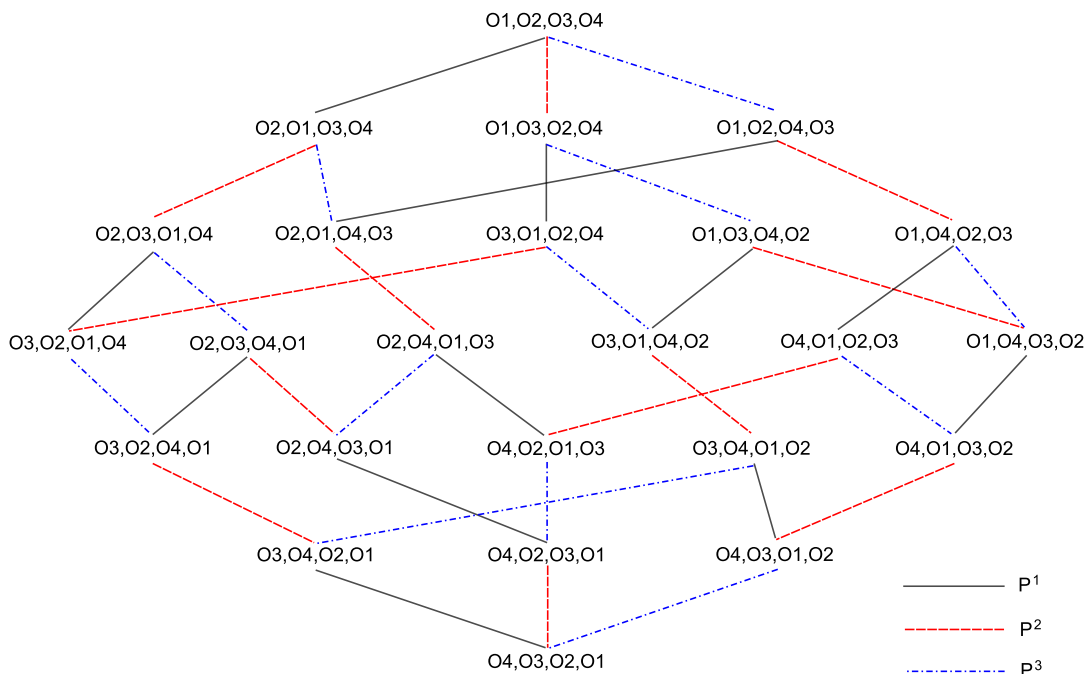


FIGURE 5.2 – Graphe de Cayley des permutations d'un ensemble d'objets

Nous calculons la distance de Kendall entre les permutations π'' et π''' suivantes :

$$\pi'' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^4 & O^3 & O^1 \end{pmatrix} ; \quad \pi''' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^3 & O^2 & O^1 & O^4 \end{pmatrix}$$

Si nous déterminons cette distance à partir du formalisme relationnel, il nous faut d'abord déterminer les permutations inverses. Nous avons :

$$(\pi'')^{-1} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^4 & O^1 & O^3 & O^2 \end{pmatrix} ; \quad (\pi''')^{-1} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^3 & O^2 & O^1 & O^4 \end{pmatrix}$$

20. cf déf-5.1.7

Puis, nous devons calculer les matrices relationnelles correspondantes qui sont les suivantes :

$$C'' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix} ; \quad C''' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}$$

Ensuite, nous avons la matrice de terme général $(C''_{ii'} - C'''_{ii'})^2$ qui vaut :

$$\begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix} \end{matrix}$$

Nous obtenons finalement $d_K(\pi'', \pi''') = \frac{1}{2} \sum_{i=1}^N \sum_{i'=1}^N (C''_{ii'} - C'''_{ii'})^2 = 3$.

Une méthode graphique et beaucoup plus adéquate permettant de calculer la distance de Kendall pour deux permutations est de compter le nombre d'intersections engendrées par les droites rejoignant les objets dans les deux listes. Nous illustrons ces propos dans le cas de notre exemple dans la figure fig-5.3.

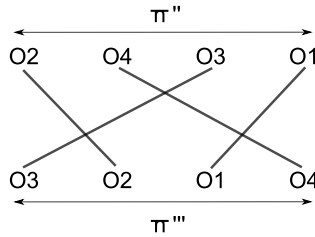


FIGURE 5.3 – Méthode graphique de calcul de la distance de Kendall entre deux permutations

La distance de Kendall, en terme des matrices relationnelles, permet d'identifier les paires qui sont discordantes au sens de la définition donnée page 17 et consiste à compter leur nombre.

Ces paires discordantes donnent, au niveau de la permutation courante, la dénomination des objets qu'il faut transposer afin de passer d'une permutation à une autre. Par ailleurs, ces objets appartenant aux paires discordantes, ont des images réciproques " $(\pi)^{-1}$ ", qui donnent des objets qui sont juxtaposés au sens de l'ordre initial. L'action de ces paires discordantes se traduit en fait par une composition à gauche de la permutation. Ces transpositions sont dans un certain sens génératrices étant donné que leurs équivalents (donnés pas les images réciproques) dans l'espace dual des transposition par composition à droite le sont. Nous illustrons ces propos en reprenant notre exemple précédent. Nous avons comme paires discordantes les paires suivantes $\{(O^1, O^4), (O^2, O^3), (O^3, O^4)\}$:

1. $(O^3, O^4) \circ \pi'' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^1 & O^2 & O^4 & O^3 \end{pmatrix} \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^4 & O^3 & O^1 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^4 & O^1 \end{pmatrix}$
2. $(O^1, O^4) \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^4 & O^1 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^4 & O^2 & O^3 & O^1 \end{pmatrix} \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^4 & O^1 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^1 & O^4 \end{pmatrix}$
3. $(O^2, O^3) \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^1 & O^4 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^1 & O^3 & O^2 & O^4 \end{pmatrix} \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^1 & O^4 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^3 & O^2 & O^1 & O^4 \end{pmatrix} = \pi'''$

Dans cet exemple, nous pouvons voir qu'à toute transposition génératrice $\pi^i, i = 1, \dots, N-1$ agissant à droite d'une permutation, nous pouvons faire correspondre des transpositions d'objets²¹ dont l'action à gauche d'une permutation est identique à la première. C'est le cas de la transposition (O^1, O^4) du deuxième point. Cette paire agissant à gauche de la permutation considérée, correspond à la transposition génératrice $\pi^3 = (O^3, O^4)$ agissant à droite de cette permutation :

$$(O^1, O^4) \circ \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^4 & O^1 \end{pmatrix} = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^2 & O^3 & O^4 & O^1 \end{pmatrix} \circ \pi^3 \quad (5.1.39)$$

En effet, π^3 consiste à permuter les objets d'indices de rang²² 3 et 4 or, dans la permutation courante, l'objet correspondant au rang 3 est O^4 et celui correspondant au rang 4 est O^1 .

Cette correspondance, action à droite, action à gauche, est exploitée dans la section suivante et plus particulièrement dans la sous-section 5.2.2 mais dans le contexte des relations d'ordre total et strict²³. Il s'agit d'une dualité qui constituera l'originalité de la modélisation des tresses que nous proposerons ultérieurement.

5.2 Action des transpositions sur les classements

Nous revenons donc vers l'étude des relations d'affectation d'un ensemble d'objets à un ensemble de rangs. Il s'agit d'une relation binaire sur deux ensembles distincts et qui représente des classements entre les objets. **Contrairement au groupe des permutations d'un ensemble d'objets totalement ordonnés, les classements tels que nous les avons présentés initialement distinguent l'ensemble des objets \mathbb{O} , de l'ensemble des rangs \mathbb{D} .**

Nous étudions l'action du groupe des permutations sur les ordres stricts et totaux. Comme nous l'avons exposée ci-dessus, et aussi lors de l'étude des relations d'équivalence, l'action des permutations est de deux types : action à droite et action à gauche qui correspondent respectivement à action sur les objets et action sur les rangs.

Nous présentons, dans cette section, l'ensemble des ordres stricts et totaux sous la forme d'un automate à états finis tel que nous l'avons présenté en déf-5.2.2.

Notre objectif est de présenter de manière unifiée les structures algébriques des relations d'ordre strict et total et celles des relations d'équivalence dans un contexte algébrique fondé sur les graphes dont les noeuds sont des éléments d'un premier ensemble et les arcs des actions d'éléments d'un deuxième ensemble²⁴. Au delà de cette unification des aspects algébriques des relations binaires classiques, cette section permet d'introduire l'approche que nous développerons dans le chapitre 7 pour la modélisation des tresses : ces dernières étant une extension des permutations ou des ordres stricts et totaux, elles possèdent aussi cette même structure d'automate.

Remarque 5.2.1 Notations

Nous faisons la même remarque à propos des notations que celle effectuée en rq-2.1.4 au chapitre 2 : par souci de clarté, τ fera référence à des permutations²⁵ agissant sur les objets de \mathbb{O} et σ fera référence à des permutations agissant sur les descripteurs qui sont ici des rangs de \mathbb{D} . Les matrices d'affectations représentatives de τ et de σ seront notées T et S .

21. non nécessairement juxtaposés

22. au sens de l'ordre engendré par une liste

23. qui sont en correspondance biunivoque avec le groupe des permutations

24. c'est la raison pour laquelle, nous parlons davantage d'automates que de graphes de Cayley pour lesquels les arcs sont des éléments d'un sous-ensemble générateur de l'ensemble constituant les noeuds du graphe

25. et plus particulièrement de transpositions d'éléments juxtaposés

Nous rappelons²⁶ que l'ensemble des permutations d'un ensemble d'objets est identique à l'ensemble des opérations de transferts doubles que nous avons introduits dans un contexte plus général²⁷ lors de l'étude des partitions au chapitre 2.

5.2.1 Action des transpositions génératrices sur les matrices d'affectation et les matrices relationnelles

Les résultats de cette sous-section sont globalement ceux énoncés au paragraphe 2.1.2.3 page 43.

Dans ce qui suit, nous étudions les actions de transpositions, sur une relation d'ordre strict et total notée \mathcal{R} qui pourra être représentée selon plusieurs codages.

5.2.1.1 Action des transpositions sur les matrices d'affectation

Soit \mathcal{R} une relation d'ordre strict et total représentée par son codage algébrique noté κ , ou par son codage par affectation K .

Soit τ ou σ une permutation représentée par leurs matrices de permutation T ou S .

Il est utile pour la suite de distinguer des sous-ensembles de $\mathbb{S}^g(\mathbb{D})$ que nous appelons transpositions d'ordre m :

Définition 5.2.1 Transposition d'ordre m

Soit $\sigma^{jj'}$ une transposition de $\mathbb{S}(\mathbb{D})$. Nous dirons que $\sigma^{jj'}$ est une transposition d'ordre m si $|j - j'| = m$, où $1 \leq m \leq N - 1$.

L'ensemble des éléments d'ordre m seront notés $\mathbb{S}_m(\mathbb{D})$.

Propriété 5.2.1 Propriétés des sous-ensembles $\mathbb{S}_m(\mathbb{D})$, $1 \leq m \leq N - 1$

- $\{\bigcup_m \mathbb{S}_m(\mathbb{D}) = \mathbb{S}(\mathbb{D})\} \wedge \{\mathbb{S}_m(\mathbb{D}) \cap \mathbb{S}_{m'}(\mathbb{D}) = \emptyset \quad \forall m, m' = 1, \dots, N - 1 : m \neq m'\}$: les sous-ensembles d'ordre m forment une partition de $\mathbb{S}(\mathbb{D})$
- $\#\mathbb{S}_m(\mathbb{D}) = N - m$
- $\mathbb{S}_1(\mathbb{D}) = \mathbb{S}^g(\mathbb{D})$: $\mathbb{S}_1(\mathbb{D})$ correspond à l'ensemble des générateurs du groupe des permutations sur l'ensemble des rangs \mathbb{D} où $D^1 < \dots < D^N$. Nous parlerons donc aussi de transpositions génératrices.

Nous pouvons maintenant définir succinctement les actions des transpositions sur le classement \mathcal{R} dans le tableau tab-5.1. Compte tenu du formalisme commun entre les structures étudiées ici et celles étudiées au chapitre 2, nous renvoyons le lecteur vers ce dernier et plus particulièrement vers la section 2.1.3 page 44 pour de plus amples détails.

Dans le cas de l'action sur les objets, nous définissons l'ensemble des transpositions comme étant l'ensemble $\{\tau^{ii'}; i, i' = 1, \dots, N : i \neq i'\} = \mathbb{S}(\mathbb{O})$ de cardinal $\frac{N(N-1)}{2}$.

Dans le cas de l'action sur les rangs, nous nous restreignons à l'ensemble des transpositions génératrices ou d'ordre 1, éléments de $\mathbb{S}^g(\mathbb{D})$: $\{\sigma^j = \sigma^{j,j+1}; j = 1, \dots, N - 1\}$.

26. cf rq-5.1.4 page 160

27. dans la mesure où les transferts doubles sont un sous-ensemble de l'ensemble des transferts

Type de transposition	Codage algébrique	Codage par affectation
$\tau^{ii'} \in \mathbb{S}(\mathbb{O}) \quad \forall i, i' = 1, \dots, N$	$\tau^{ii'}(\mathcal{R}) = \kappa \circ \tau^{ii'}$	$\tau^{ii'}(\mathcal{R}) = T^{ii'} \cdot K$
$\sigma^{j,j+1} \in \mathbb{S}^g(\mathbb{D}) \quad \forall j = 1, \dots, N-1$	$\sigma^{j,j+1}(\mathcal{R}) = \sigma^{j,j+1} \circ \kappa$	$\sigma^{j,j+1}(\mathcal{R}) = K \cdot S^{j,j+1}$

TABLE 5.1 – Action des transpositions sur les matrices d'affectation

Nous pouvons ainsi définir, à l'instar des relations d'équivalence, le graphe orienté des relations d'ordre strict et total sous la forme d'un automate à états finis²⁸.

Définition 5.2.2 *Graphe orienté de l'ensemble des classements d'un ensemble \mathbb{O}*

Le graphe orienté de l'ensemble des classements de \mathbb{O} est un automate $A(\mathbb{O}) = (\mathbb{E}, \Sigma, \mu, \mathbb{A}, E^\epsilon)$ où :

- $\mathbb{E} = \text{ROST}(\mathbb{O}) =$ Ensemble des relations d'ordre strict et total de l'ensemble \mathbb{O}
- Σ :
 - $\Sigma = \mathbb{S}(\mathbb{O}) =$ ensemble des permutations agissant sur les objets
 - $\Sigma = \mathbb{S}^g(\mathbb{D}) =$ ensemble des permutations agissant sur les rangs
- μ :
 - $\mu = \circ$: composition à droite si les transpositions agissent sur les objets
 - $\mu = \circ$: composition à gauche si les transpositions agissent sur les rangs
- $\mathbb{A} = \emptyset$
- $E^\epsilon = \kappa^\epsilon = O^1 < \dots < O^N$ le classement "identité",

Dans la figure fig-5.4, les éléments constituant les noeuds du graphe sont des classements contrairement au graphe de Cayley de la figure fig-5.2 pour lesquels, les noeuds sont des listes. Nous donnons le graphe des classements dont les arcs sont des transpositions de $\mathbb{S}(\mathbb{O})$ agissant sur les objets.

5.2.1.2 Action des transpositions sur les matrices relationnelles

Dans le cas de l'action des transpositions sur les matrices relationnelles codant des classements, nous avons la même dualité "affectation / relationnel" que nous avons évoquée dans le paragraphe 2.1.3.2.3 page 49 mais pour des matrices de passage \mathbf{R} distinctes de celles des partitions :

- Dans le cas des partitions, $\mathbf{R} = I_N =$ partition triviale sur les classes de \mathbb{D} indiquant par défaut qu'aucune classe n'est associée à une autre
- Dans le cas des classements, $\mathbf{R} = R_N =$ matrice relationnelle sur les rangs éléments de \mathbb{D} qui, code l'ordre strict et total : $D^1 < D^2 < \dots < D^N$

Propriété 5.2.2 *Dualité Affectation / Relationnel*

28. au sens de déf-2.2.1

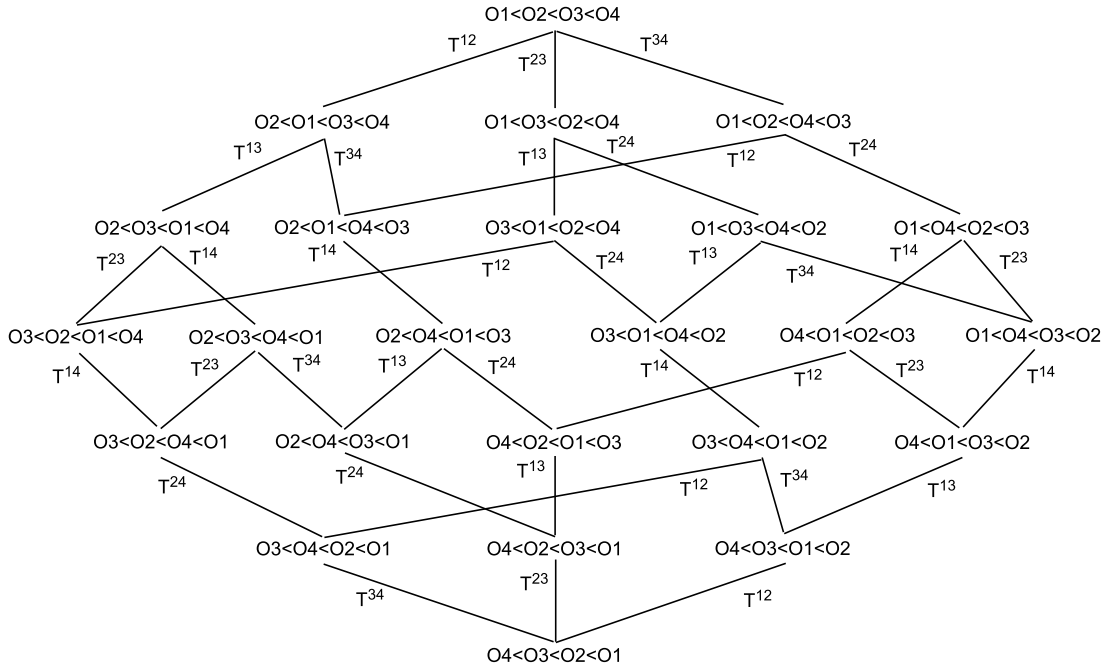


FIGURE 5.4 – Graphe orienté des permutations d'un ensemble d'objets

Soit une relation d'ordre strict et total \mathcal{R} représentée soit par sa matrice d'affectation K , soit par sa matrice relationnelle C .

Soit τ une transposition agissant sur l'ensemble des objets \mathbb{O} et soit σ une transposition génératrice agissant sur l'ensemble des rangs \mathbb{D} qui sont respectivement représentées par leurs matrices d'affectation associées T et S .

Soit $\mathcal{R}' = \tau(\mathcal{R})$ le classement obtenu après l'action de τ ou de σ sur \mathcal{R} et soit C' la matrice relationnelle associée au classement \mathcal{R}' .

Soit \mathbf{R} une matrice relationnelle codant une relation d'ordre strict et total sur les éléments de \mathbb{D} dite matrice de passage. Soit enfin, \mathbf{D} une matrice d'affectation ou de transposition, les deux types de matrices vérifiant les mêmes propriétés structurales.

Les actions de τ sur \mathcal{R} dans le cadre du codage relationnel sont de la forme générale suivante :

$$C' = \mathbf{D} \cdot \mathbf{R} \cdot {}^t\mathbf{D} \quad (5.2.1)$$

que nous exprimons par la formulation suivante :

$$\text{“Relationnel} = \text{Affectation} \cdot \text{Relationnel} \cdot {}^t\text{Affectation”}$$

Nous avons les différents cas suivants que nous illustrons dans le tableau tab-5.2. Dans celui-ci R_N est la matrice relationnelle de passage standard codant l'ordre strict et total suivant : $D^1 < D^2 < \dots < D^N$.

L'action de σ^j sur les rangs de \mathcal{R} représentée par C , consiste à modifier une matrice relationnelle de passage \mathbf{R} qui code une relation d'ordre strict et total. Ces actions sont plus précisément décrites dans la définition déf-2.1.22 et lors de la remarque rq-2.2.3.

Type de transposition	Définition de \mathbf{R}	Définition de \mathbf{D}
$\tau^{ii'} \in \mathbb{S}(\mathbb{O}) \quad \forall i, i' = 1, \dots, N$	$\mathbf{R} = K \cdot R_N \cdot {}^t K = C$	$\mathbf{D} = T^{ii'}$
$\sigma^{j,j+1} \in \mathbb{S}^g(\mathbb{D}) \quad \forall j = 1, \dots, N-1$	$\mathbf{R} = S^{j,j+1} \cdot R_N \cdot {}^t S^{j,j+1}$	$\mathbf{D} = K$

TABLE 5.2 – Dualité affectation/relationnel lors de l'action des transpositions sur des matrices relationnelles codant des classements

5.2.2 Action des transpositions sur les matrices de paires d'objets nommés

Ce codage a la particularité d'identifier les paires d'objets nommés d'un classement donné qui sont relatives aux transpositions génératrices agissant sur les rangs de ce classement. Reprenons l'exemple donné en page 159 de la matrice de paires d'objets nommés représentative du classement $\mathcal{R}' = O^3 < O^1 < O^4 < O^2$:

$$L' = \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ [O^1, O^4] & [O^1, O^2] & [O^4, O^2] \end{pmatrix}$$

Sur cet exemple, appliquer la transposition d'ordre 1, σ^1 , qui permute les objets en positions 1 et 2, consiste à transposer les objets $(\kappa')^{-1}(D^1) = O^3$ et $(\kappa')^{-1}(D^2) = O^1$. Or, cette paire est donnée par L'_{11} .

De la même manière, appliquer la transposition σ^2 ou σ^3 à \mathcal{R}' revient à transposer les objets O^1 et O^4 pour σ^2 ou O^4 et O^2 pour σ^3 .

Suite à cet exemple, nous pouvons énoncer la propriété suivante.

Propriété 5.2.3 Correspondance entre les transpositions génératrices agissant sur les rangs et les transpositions génératrices agissant sur les objets correspondant à un classement donné

Soit \mathcal{R} un classement et soit κ son codage algébrique et L son codage sous forme de matrices de paires d'objets nommés.

Soit, $\{\tau^{ii'}; i, i' = 1, \dots, N\}$, les éléments de $\mathbb{S}(\mathbb{O})$, l'ensemble des opérations de transpositions agissant sur les objets.

Soit de la même manière, $\{\sigma^{jj'}; j = 1, \dots, N-1; j' = 2, \dots, N : j < j'\}$, les éléments de $\mathbb{S}(\mathbb{D})$, l'ensemble des opérations de transpositions agissant sur les rangs. Nous avons alors la correspondance suivante, où $j = 1, \dots, N-1; j' = 2, \dots, N; j < j'$:

$$\sigma^{jj'} \circ \kappa \Leftrightarrow \kappa \circ \tau^{ii'} \text{ où } \tau^{ii'} = \phi(L_{j,j'-1}) = ([L_{j,j'-1}]_1, [L_{j,j'-1}]_2) \quad (5.2.2)$$

où ϕ est l'application qui à une paire d'objets nommés $[O^i, O^{i'}]$ lui associe la transposition $\tau^{ii'} = (O^i, O^{i'})$ de $\mathbb{S}(\mathbb{O})$. Nous avons la correspondance suivante pour les sous-ensemble de $\mathbb{S}_m(\mathbb{D})$ des transpositions d'ordre m :

$$\sigma^{j,j+m} \circ \kappa \Leftrightarrow \kappa \circ \tau^{ii'} \text{ où } \tau^{ii'} = \phi(L_{j,j+m-1}) = ([L_{j,j+m-1}]_1, [L_{j,j+m-1}]_2) \quad (5.2.3)$$

De manière plus particulière, nous avons la correspondance suivante pour l'ensemble des transpositions génératrices, les éléments $\sigma^j = \sigma^{j,j+1}$ de $\mathbb{S}_1(\mathbb{D}) = \mathbb{S}^g(\mathbb{D})$:

$$\sigma^j \circ \kappa \Leftrightarrow \kappa \circ \tau^{ii'} \text{ où } \tau^{ii'} = \phi(L_{jj}) = ([L_{jj}]_1, [L_{jj}]_2) \quad (5.2.4)$$

Nous parlerons alors, dans ce dernier cas, de **paires d'objets nommés et juxtaposés**.

Démonstration

Cette propriété de correspondance est obtenue à partir des propriétés des matrices de paires d'objets nommés données en prop-5.1.11.

□

La transposition des objets rangés en positions j et $j + 1$ donnée par σ^j , correspond à la transposition $(O^i, O^{i'})$ pour lesquels, l'indice du rang de O^i est j ²⁹, et celui de $O^{i'}$ est $j + 1$.

Etant donnée un classement \mathcal{R} , l'application ϕ permet de déterminer au travers de la matrice L , toutes les transpositions de $\mathbb{S}(\mathbb{O})$ qui sont propres à la permutation donnée, et qui correspondent chacune à une transposition de $\mathbb{S}(\mathbb{D})$.

Par abus de langage, nous omettrons l'application ϕ lorsque nous appliquons, dans le cadre du codage sous forme de matrices de paires d'objets nommés, une transposition d'objets à un classement donné.

Nous donnons à proprement dit, l'action d'une transposition sur un classement dans le cadre des matrices de paires d'objets nommés.

Définition 5.2.3 *Produit (à droite) d'une paire d'objets nommés par une transposition de $\mathbb{S}(\mathbb{O})$*

Soit une transposition $\tau^{ii'}$ de $\mathbb{S}(\mathbb{O})$ et soit une paire d'objets nommés $[O^l, O^{l'}]$. Le produit non commutatif de $[O^l, O^{l'}]$ par $\tau^{ii'}$, noté $[O^l, O^{l'}] * (O^i, O^{i'})$ est donné par la relation suivante³⁰ :

$$[O^l, O^{l'}] * \tau = [\tau(O^l), \tau(O^{l'})] \quad (O^l, O^{l'}) \in \mathbb{O} \times \mathbb{O} \quad (5.2.5)$$

Dans le cas particulier où $\tau = \tau^{ii'}$ est une transposition de $\mathbb{S}(\mathbb{O})$, nous avons :

$$[O^l, O^{l'}] * (O^i, O^{i'}) = \begin{cases} [O^i, O^{i'}] & \text{si } \{l = i'\} \wedge \{l' \neq i\} \wedge \{l' \neq i'\} \\ [O^l, O^i] & \text{si } \{l \neq i\} \wedge \{l \neq i'\} \wedge \{l' = i'\} \\ [O^{i'}, O^{l'}] & \text{si } \{l = i\} \wedge \{l' \neq i\} \wedge \{l' \neq i'\} \\ [O^l, O^{i'}] & \text{si } \{l \neq i\} \wedge \{l \neq i'\} \wedge \{l' = i\} \\ [O^i, O^{i'}] & \text{si } \{l = i'\} \wedge \{l' = i\} \\ [O^{i'}, O^i] & \text{si } \{l = i\} \wedge \{l' = i'\} \\ [O^l, O^{l'}] & \text{si } \{l \neq i\} \wedge \{l \neq i'\} \wedge \{l' \neq i\} \wedge \{l' \neq i'\} \end{cases} \quad (5.2.6)$$

Définition 5.2.4 *Produit (à droite) d'une matrice de paires d'objets nommés par une transposition de $\mathbb{S}(\mathbb{O})$*

Soit une transposition $\tau^{ii'}$ de $\mathbb{S}(\mathbb{O})$ et une matrice de paires d'objets nommés L représentative d'un classement \mathcal{R} . Le produit de L par $\tau^{ii'}$, noté $L * \tau^{ii'}$ est une matrice de paires d'objets nommés dont le terme général est donné par la relation suivante :

$$\boxed{[L * \tau^{ii'}]_{jj'} = L_{jj'} * \tau^{ii'}} \quad (5.2.7)$$

29. $j \leftrightarrow D^j \quad \forall j = 1, \dots, N - 1$

30. plus généralement, la composition $*$ est la composition entre une paire d'objets nommés $[O^i, O^{i'}]$ et une permutation quelconque π de $\mathbb{S}(\mathbb{O})$ dont les transpositions $\tau^{ii'}$ sont des 2-cycles

Nous donnons ci-dessous la propriété de correspondance entre l'action sur les rangs et l'action sur les objets.

Définition 5.2.5 *Correspondance entre les actions des $\sigma^j \in \mathbb{S}^g(\mathbb{D})$ et les actions des $\tau^{ii'} \in \mathbb{S}(\mathbb{O})$*
Soit \mathcal{R} un classement représenté par sa matrice de paires d'objets nommés L . Soit $\sigma^j \in \mathbb{S}^g(\mathbb{D})$. La matrice de paires d'objets nommés correspondante au classement $\mathcal{R}' = \sigma^j(\mathcal{R})$ est donnée par :

$$L * \phi(L_{jj})$$

où ϕ est l'application qui associe à une paire d'objets nommés de L , une transposition d'objets.
Par abus de langage, nous écrirons :

$$\boxed{\sigma^j(\mathcal{R}) = L * L_{jj}} \quad (5.2.8)$$

Désormais, lorsque nous parlerons de transpositions, nous sous-entendrons des éléments $\tau^{ii'}$ de $\mathbb{S}(\mathbb{O})$ donnés par les paires d'objets nommés qui sont les termes des matrices L , sauf lorsque cela sera précisé.

L'action d'une transposition génératrice sur un classement représenté par une matrice de paires d'objets nommés nécessite *a priori* l'évaluation de $\frac{N(N-1)}{2}$ compositions. Nous donnons ci-dessous une propriété remarquable permettant de calculer de manière rapide la matrice de paires d'objets nommés $L' = L * (O^i, O^{i'})$ issue de l'action d'une transposition provenant de la diagonale de la matrice L pour laquelle, $\exists j = 1, \dots, N-1 : (O^i, O^{i'}) = \phi(L_{jj})$. Cette propriété ramène le nombre d'opérations à $N-1$ permettant ainsi de "linéariser" le calcul de la détermination du résultat des actions de transpositions sur des matrices de paires d'objets nommés.

Nous introduisons pour cela une matrice générique, qui pourra être pré-calculée et à partir de laquelle, nous pouvons déterminer de manière rapide les seules véritables opérations à effectuer lors de l'action d'une transposition sur une matrice de paires d'objets nommés.

Définition 5.2.6 *Matrice générique de détermination de la matrice de paires d'objets nommés résultantes de l'action d'une transposition*

Notons \mathcal{L} la matrice rectangulaire de taille $(N \times N - 1)$ dont le terme général est donné par :

$$\boxed{\mathcal{L}_{jj'} = \begin{cases} L_{jj'} & \text{si } \{1 \leq j \leq N-1\} \wedge \{j \leq j' \leq N-1\} \\ L_{j',j-1} & \text{si } \{2 \leq j \leq N\} \wedge \{1 \leq j' \leq j-1\} \end{cases}} \quad (5.2.9)$$

Pour $N = 4$, nous avons :

$$\mathcal{L} = \begin{matrix} & \begin{matrix} 1 & 2 & 3 \end{matrix} \\ \begin{matrix} 1 \\ 2 \\ 3 \\ 4 \end{matrix} & \begin{pmatrix} L_{11} & L_{12} & L_{13} \\ L_{11} & L_{22} & L_{23} \\ L_{12} & L_{22} & L_{33} \\ L_{13} & L_{23} & L_{33} \end{pmatrix} \end{matrix} \quad (5.2.10)$$

Théorème 5.2.1 *Règle de détermination d'une matrice L' issue du produit $L * L_{j^*,1}$ à l'aide de la matrice générique \mathcal{L}*

Soit L une matrice de paires d'objets nommés représentative d'un ordre strict et total \mathcal{R} sur un ensemble de N objets. Soit $L' = L * L_{j^*,j^*}$ la matrice de paires d'objets nommés représentative de l'ordre strict et total de $\mathcal{R}' = \sigma^{j^*}(\mathcal{R})$ où $j^* \in 1, \dots, N-1$. Soit \mathcal{L} la matrice générique présentée en déf-5.2.6 de taille $(N \times N - 1)$.

Alors $L' = L * L_{j^*,j^*}$ est obtenue en effectuant $N-1$ opérations élémentaires identifiées à l'aide de la matrice \mathcal{L} et qui sont les suivantes :

- l'élément $L'_{j^*j^*}$ est la paire donnée par $L_{j^*j^*}$ mais dont les éléments ont été permutés : $L'_{j^*j^*} = [[L_{j^*j^*}]_2, [L_{j^*j^*}]_1] = r(L_{j^*j^*})$ où $r([O^i, O^{i'}]) = [O^{i'}, O^i]$ est l'opérateur qui permute les objets d'une paire.
- pour les paires dont les objets sont relatifs aux indices de rang j^* et $j^* + 1$, L' est obtenue à partir de L en interchangeant les termes de L donnés par $\mathcal{L}_{j^*j'}$ avec les termes de L donnés par $\mathcal{L}_{j^*+1,j'}$ sauf pour $j' = j^*$ dont la transformation est donnée par le point précédent
- toutes les paires dont les objets sont de rangs distincts de j^* et $j^* + 1$, ne sont pas affectées par l'action de la transposition, nous avons par conséquent pour ces paires d'objets nommés $L'_{jj'} = L_{jj'}$

Pour démontrer le théorème précédent, nous commencerons par démontrer la propriété suivante :

Lemme 5.2.1 Propriété de la matrice génératrice \mathcal{L}

Soit \mathcal{L} une matrice générique telle que définie par déf-5.2.6. Nous avons les relations suivantes :

$$\boxed{\begin{array}{l} \forall j, j', \{j = 1, \dots, N - 1\} \wedge \{j \leq j' \leq N - 1\} : [\mathcal{L}_{jj'}]_1 = (\kappa)^{-1}(D^j) \\ \forall j, j', \{j = 2, \dots, N\} \wedge \{1 \leq j' \leq j - 1\} : [\mathcal{L}_{jj'}]_2 = (\kappa)^{-1}(D^j) \end{array}} \quad (5.2.11)$$

En d'autres termes, pour chaque ligne $j = 1, \dots, N - 1$, nous avons les informations suivantes qui sont toutes relatives à l'objet dont le rang est D^j :

1. les termes de $\mathcal{L}_{jj'}$ pour lesquels $\{j \leq j' \leq N - 1\}$, sont tels que les premiers objets de la paire, $[\mathcal{L}_{jj'}]_1$, sont d'indice de rang j
2. les termes de $\mathcal{L}_{jj'}$ pour lesquels $\{1 \leq j' \leq j - 1\}$, sont tels que les deuxièmes objets de la paire, $[\mathcal{L}_{jj'}]_2$, sont d'indice de rang j

Démonstration

Pour le premier sous-ensemble, cela est dû à la définition même des matrices de paires d'objets nommés données par déf-5.1.5. Pour le deuxième sous-ensemble, il suffit de voir qu'il correspond à la transposée de L . Par conséquent, sur une même ligne, les indices de rang correspondant au deuxième terme de la paire d'objets sont constants et valent j par construction. En effet, nous avons $\mathcal{L}_{jj'} = L_{j',j-1}$. Or par définition³¹, le rang du deuxième objet correspondant à l'élément $L_{j',j-1}$ vaut $j - 1 + 1 = j$.

□

Nous donnons alors ci-dessous la démonstration du théorème théo-5.2.1 :

Démonstration

Appliquer à un classement κ une transposition d'objets $\tau^{ii'}$ correspondante à une transposition génératrice σ^{j^*} consiste à transposer les objets $O^i = (\kappa)^{-1}(D^{j^*})$ et $O^{i'} = (\kappa)^{-1}(D^{j^*+1})$. Dans le cadre du codage par matrices de paires d'objets nommés, nous voyons que les termes concernés sont ceux donnés par les lignes j^* et $j^* + 1$ de la matrice générique \mathcal{L} en raison du lemme lem-5.2.1. Par construction, la matrice \mathcal{L} est organisée de telle manière qu'il suffit de permuter les termes de L donnés par $\mathcal{L}_{j^*j'}$ avec les termes de L donnés par $\mathcal{L}_{j^*+1,j'}$ pour $j' \neq j^*$. En effet, nous constatons que les opérations élémentaires précédentes de permutations entre paires d'objets nommés, correspondent à celles qui sont relatives aux objets de rangs j^* et $j^* + 1$.

Il reste la permutation concernant les objets O^i et $O^{i'}$. Ces derniers, restent juxtaposés mais $O^{i'}$ passe devant O^i si bien que $L_{j^*j^*}$ devient $L'_{j^*j^*} = r(L_{j^*j^*})$.

□

31. cf déf-5.1.5 page 159

Afin d'illustrer cette propriété, considérons à nouveau la matrice L' représentative du classement $O^3 < O^1 < O^4 < O^2$:

$$L' = \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ & [O^1, O^4] & [O^1, O^2] \\ & & [O^4, O^2] \end{pmatrix}$$

Nous avons par exemple $\sigma^2(\mathcal{R}') = O^3 < O^4 < O^1 < O^2$. En appliquant les règles décrites dans prop-5.2.1, nous obtenons L'' représentative de $O^3 < O^4 < O^1 < O^2$ en effectuant sur L' les changements suivants, sachant que $j^* = 2$:

- $L''_{22} = r(L'_{22}) = r([O^1, O^4]) = [O^4, O^1]$
- $\{\mathcal{L}_{11} \leftrightarrow \mathcal{L}_{12}\} \Rightarrow \{L'_{11} \leftrightarrow L'_{12}\} \Rightarrow \{\{L''_{11} = L'_{12} = [O^3, O^4]\} \wedge \{L''_{12} = L'_{11} = [O^3, O^1]\}\}$
- $\{\mathcal{L}_{23} \leftrightarrow \mathcal{L}_{33}\} \Rightarrow \{L'_{23} \leftrightarrow L'_{33}\} \Rightarrow \{\{L''_{23} = L'_{33} = [O^4, O^2]\} \wedge \{L''_{33} = L'_{23} = [O^1, O^2]\}\}$
- Et pour tous les autres termes : $L''_{jj'} = L'_{jj'}$

Au vu de la matrice générique, nous avons appliqué les opérations représentées ci-dessous :

$$\begin{aligned} \mathcal{L} &= \begin{pmatrix} L'_{11} & L'_{12} & L'_{13} \\ L'_{11} & L'_{22} & L'_{23} \\ \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow \\ L'_{12} & r(L'_{22}) & L'_{33} \\ L'_{13} & L'_{23} & L'_{33} \end{pmatrix} \\ &= \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ [O^3, O^1] & [O^1, O^4] & [O^1, O^2] \\ \uparrow\downarrow & \uparrow\downarrow & \uparrow\downarrow \\ [O^3, O^4] & r([O^1, O^4]) & [O^4, O^2] \\ [O^3, O^2] & [O^1, O^2] & [O^4, O^2] \end{pmatrix} \end{aligned}$$

Nous obtenons bien la matrice de paires d'objets nommés représentant le classement $O^3 < O^4 < O^1 < O^2$:

$$L' = \begin{pmatrix} [O^3, O^4] & [O^3, O^1] & [O^3, O^2] \\ & [O^4, O^1] & [O^4, O^2] \\ & & [O^1, O^2] \end{pmatrix}$$

Nous donnons à présent, en terme de matrices de paires d'objets nommés, l'expression de la propriété de décomposition d'une transposition agissant sur les rangs d'ordre $m > 1$ en fonction de transpositions génératrices σ^j , que nous avons rappelée en prop-5.1.14.

Propriété 5.2.4 *Expression des transpositions d'ordre $m > 1$ en fonction des transpositions génératrices dans le cadre des matrices de paires d'objets nommés*

Soit un classement \mathcal{R} représenté par une matrice de paires d'objets nommés L . Considérons le sous-ensemble des paires $L_{jj'}$ tel que $j' > j$. Ce sous-ensemble représente l'ensemble des paires d'objets nommés dont les rangs des objets sont distants de $m > 1$. Ces paires indiquent les objets qu'il faut permuter lorsque nous appliquons une transposition agissant sur les rangs $\sigma^{jj+m} \in \mathbb{S}_m(\mathbb{D})$ avec $j' = j + m - 1$. Nous pouvons alors exprimer la propriété prop-5.1.14 dans le codage des matrices de paires d'objets nommés de la manière suivante :

$$\begin{array}{l} L * L_{jj'} = L * L_{jj} * \dots * L_{j,j'-1} * L_{jj'} * L_{j'j'} * \dots * L_{j-1,j'} \\ L * L_{j'j'} = L * L_{j'j'} * \dots * L_{j-1,j'} * L_{j'j'} * L_{jj} * \dots * L_{j,j'-1} \end{array} \quad (5.2.12)$$

Dans (5.2.12), pour alléger les notations, nous avons omis le parenthésage. Il faut garder à l'esprit que $*$ traduit l'action d'une transposition³² sur un classement : $*$ n'est aucunement une composition entre transpositions. Les actions se font à droite de la matrice de paires d'objets nommés de manière successive et dans l'ordre énoncé. Nous pouvons cependant, exprimer la relation (5.2.12) en mettant en évidence \circ , la composition de transpositions qui, nous le rappellerons s'effectue par la droite puisque ces transpositions agissent sur les objets. Nous avons donc l'expression équivalente suivante, $\forall j, j' : \{1 \leq j \leq N - 1\} \wedge \{j \leq j' \leq N - 1\}$:

$$\begin{aligned} L * (L_{jj'}) &= L * (L_{jj} \circ \dots \circ L_{j,j'-1} \circ L_{jj'} \circ L_{j',j'} \circ \dots \circ L_{j-1,j'}) \\ L * (L_{jj'}) &= L * (L_{j',j'} \circ \dots \circ L_{j-1,j'} \circ L_{jj'} \circ L_{jj} \circ \dots \circ L_{j,j'-1}) \end{aligned}$$

ou encore :

$$L * L_{jj'} = L * \left(\bigcirc_{j''=0}^{j'-j-1} L_{j,j+j''} \circ L_{jj'} \circ \bigcirc_{j''=0}^{j'-j+1} L_{j'-j'',j'} \right) \quad (5.2.13)$$

$$L * L_{jj'} = L * \left(\bigcirc_{j''=0}^{j'-j+1} L_{j'-j'',j'} \circ L_{jj'} \circ \bigcirc_{j''=0}^{j'-j-1} L_{j,j+j''} \right) \quad (5.2.14)$$

Démonstration

En considérant les expressions équivalentes (5.2.13), nous voyons que la propriété prop-5.2.4 correspond à la "traduction" dans le codage de matrices de paires d'objets nommés de la propriété prop-5.1.14.

□

L'ordre dans lequel, nous appliquons des transpositions d'objets qui sont génératrices dans l'espace dual des transpositions agissant sur des rangs, décrit un chemin particulier sur les matrices de paires d'objets nommés. Ceci correspond à une propriété remarquable du codage que nous présentons. Nous l'illustrons dans l'exemple suivant.

Si nous appliquons à \mathcal{R}' la transposition σ^{14} qui permute les objets de rangs 1 et 4³³, cela revient à permuter les objets qui constituent le terme L'_{13} . Celui-ci correspond à la paire $[O^3, O^2]$ qui donne la transposition $\phi([O^3, O^2]) = (O^3, O^2)$.

Selon la propriété précédente, nous avons alors : $L' * L'_{13} = L * L'_{11} * L'_{12} * L'_{13} * L'_{33} * L'_{23}$. Cette expression est aussi égale à : $L' * L'_{13} = L' * L'_{33} * L'_{23} * L'_{13} * L'_{11} * L'_{12}$. Nous pouvons par exemple vérifier que $L' * (O^3, O^1) * (O^3, O^4) * (O^3, O^2) * (O^4, O^2) * (O^1, O^2)$ donne bien la matrice de paires d'objets nommés relatives à l'ordre $O^2 < O^1 < O^4 < O^3$.

Les deux types de chemins parcourus par cette action au vu de la matrice des paires d'objets nommés sont particuliers et correspondent aux représentations suivantes :

$$\begin{pmatrix} [O^3, O^1](\mathbf{1}) & \rightarrow & [O^3, O^4](\mathbf{2}) & \rightarrow & [O^3, O^2](\mathbf{3}) \\ & & [O^1, O^4] & & [O^1, O^2](\mathbf{5}) \\ & & & & \uparrow \\ & & & & [O^4, O^2](\mathbf{4}) \end{pmatrix}$$

$$\begin{pmatrix} [O^3, O^1](\mathbf{4}) & \rightarrow & [O^3, O^4](\mathbf{5}) & & [O^3, O^2](\mathbf{3}) \\ & & [O^1, O^4] & & \uparrow \\ & & & & [O^1, O^2](\mathbf{2}) \\ & & & & \uparrow \\ & & & & [O^4, O^2](\mathbf{1}) \end{pmatrix}$$

Il s'agit de chemins "triangulaires" dont le "sommet" est la transposition d'ordre m considérée.

32. ie que nous prenons l'image $(O^i, O^{i'})$ de la paire $[O^i, O^{i'}]$ donnée par l'application ϕ

33. les rangs sont distancés de $|1 - 4| = 3$ unités

CHAPITRE 6

Sur l'association, l'agrégation et le consensus de relations d'ordre

Sommaire

6.1	Sur les critères d'association entre deux relations d'ordre	181
6.1.1	Etude des codages des classements avec <i>ex-aequos</i>	182
6.1.1.1	Le codage algébrique	183
6.1.1.2	Le codage par affectation	183
6.1.1.3	Le codage relationnel	184
6.1.1.4	Lien entre le codage par affectation et le codage relationnel	186
6.1.2	Critères d'association et nouvelles formulations relationnelles	186
6.1.3	Critères d'écart à l'indétermination pondérée entre relations d'ordre	190
6.2	L'effet Condorcet, le théorème d'impossibilité d'Arrow et les conditions de restriction sur le domaine des préférences individuelles	192
6.2.1	Rappels en théorie du choix social	193
6.2.1.1	Le théorème d'impossibilité d'Arrow	193
6.2.1.2	Les conditions de restriction sur l'ensemble des préférences individuelles aboutissant à des préférences collectives acycliques par la règle de la majorité simple	194
6.2.2	Caractérisations algébriques de la nature de la relation collective obtenue par différentes règles majoritaires dans un cadre relationnel	198
6.2.2.1	Règles majoritaires	199
6.2.2.2	Caractérisation de la nature d'une relation collective à partir des propriétés algébriques d'une matrice de comparaisons par paires de contributions	200
6.2.2.3	Version contrainte des règles majoritaires et résolution par maximisation des critères de vote associés par programmation linéaire en nombres bivalents	212

Dans le chapitre 5, nous avons étudié les relations d'ordre strict et total. Nous avons dans un premier temps présenté différents formalismes de ces dernières. Nous avons ensuite, étudié des propriétés algébriques des classements et avons abouti à une modélisation sous forme d'automate avec une présentation de la dualité "action à droite / action à gauche" par l'intermédiaire du codage par matrices de paires d'objets nommés.

Dans ce chapitre, nous mettons en avant le codage relationnel que nous étendons à l'étude des relations d'ordre large. Ces relations binaires correspondent à des classements avec possibilité d'*ex-aequos* et de données manquantes. La thématique de ce chapitre rentre donc dans le cadre des préoccupations classiques de l'AR [Gha90].

Parallèlement au chapitre 3, nous étudions les relations d'ordre large dans le contexte des problématiques suivantes qui avaient intéressé les relations d'équivalence :

- la mesure de l'association entre deux relations d'ordre
- l'agrégation et la recherche d'une relation d'ordre consensuelle

Notre premier objectif est, à l'instar des mesures d'association dans le contexte des relations d'équivalence, de contribuer à l'unification des différentes mesures d'association entre relations d'ordre. Nous donnons notamment une interprétation nouvelle à l'ordre dual par l'intermédiaire de l'opération unaire issue de l'algèbre relationnelle : la converse ou encore la "transposée" d'une relation. Celle-ci permet de simplifier de nombreuses notations et de donner de nouvelles formulations davantage symétriques des critères d'association. Enfin, elle permet d'unifier indétermination dans le cadre des relations d'équivalence et indétermination dans le cadre des relations d'ordre.

Notre seconde contribution concerne la théorie du choix social. Nous nous plaçons alors dans un cadre davantage théorique. En effet, nous revenons sur les aspects axiomatiques du problème de la possibilité du calcul d'un ordre collectif transitif à partir de plusieurs ordres individuels. Nous rappelons les contributions majeures apportées par de nombreux économistes sur ce sujet. Nous rappelons ainsi, le théorème d'impossibilité d'Arrow et les travaux de D. Black, K. Inada et A.K. Sen autour de ce théorème concernant les conditions de restriction sur les préférences individuelles conduisant à une préférence collective transitive. Nous proposons une généralisation de leurs résultats en définissant un formalisme original.

Remarque 6.0.1 *Signification des éléments de l'ensemble des descripteurs* \mathbb{D}

Les relations d'ordre large étudiées peuvent se présenter en pratique sous différentes formes. Dans le contexte de l'analyse de données, il peut s'agir de variables qualitatives possédant des modalités ordonnées. Dans le cadre des statistiques non-paramétriques il peut s'agir de variables quantitatives pour lesquelles nous nous intéressons aux rangs induits sur les objets. Dans le contexte de l'agrégation multicritère, il peut s'agir tout simplement de rangs donnés à un ensemble d'objets avec potentiellement des ex-aequos.

Dans tous les cas, nous notons, comme précédemment, les descripteurs d'une relation d'ordre large par D^j avec $j = 1, \dots, N$. Ces descripteurs sont initialement ordonnés selon l'ordre suivant : $D^1 < D^2 < \dots < D^N$.

6.1 Sur les critères d'association entre deux relations d'ordre

Nous rappelons dans un premier temps le codage relationnel des relations d'ordre dans un cadre plus général que celui des relations d'ordre strict et total.

Puis, nous investiguons les critères d'association permettant de mesurer la liaison entre deux relations d'ordre. Nous donnons des formulations relationnelles de ces critères qui sont différentes de celles proposées par S. Ghashghaie dans [Gha90]. Nos formules mettent en évidence une forme symétrique qui rappelle les indices d'écart à l'indétermination que nous avons introduits en sous-section 3.3.2 pour des relations d'équivalences.

Cette formulation nous permet enfin d'étendre les critères d'association étudiés au concept d'indétermination pondérée.

6.1.1 Etude des codages des classements avec *ex-aequos*

Parmi les relations d'ordre, nous étudions ici les ordres larges et totaux ou les préordres totaux. Nous rappelons leurs définitions.

Définition 6.1.1 Relation d'ordre large et total

Une relation d'ordre large et total, \mathcal{R} , d'un ensemble \mathbb{O} de N objets, $\{O^i, i = 1, \dots, N\}$, est une relation binaire sur cet ensemble d'éléments qui respecte les propriétés suivantes :

- *Réflexivité* : $O^i \mathcal{R} O^i \quad \forall O^i \in \mathbb{O}$
- *Antisymétrie faible* : $O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^i \Rightarrow O^i = O^{i'} \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O}$
- *Transitivité* : $O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''} \quad \forall (O^i, O^{i'}, O^{i''}) \in \mathbb{O} \times \mathbb{O} \times \mathbb{O}$
- *Totalité* : $O^i \mathcal{R} O^{i'} \vee O^{i'} \mathcal{R} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} : O^i \neq O^{i'}$

Définition 6.1.2 Relation de préordre total

Une relation de préordre total, \mathcal{R} , d'un ensemble \mathbb{O} de N objets, $\{O^i, i = 1, \dots, N\}$, est une relation binaire sur cet ensemble d'éléments qui respecte les propriétés suivantes :

- *Réflexivité* : $O^i \mathcal{R} O^i \quad \forall O^i \in \mathbb{O}$
- *Transitivité* : $O^i \mathcal{R} O^{i'} \wedge O^{i'} \mathcal{R} O^{i''} \Rightarrow O^i \mathcal{R} O^{i''} \quad \forall (O^i, O^{i'}, O^{i''}) \in \mathbb{O} \times \mathbb{O} \times \mathbb{O}$
- *Totalité* : $O^i \mathcal{R} O^{i'} \vee O^{i'} \mathcal{R} O^i \quad \forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O} : O^i \neq O^{i'}$

Au vu des définitions précédentes, nous constatons que les préordres sont équivalents aux ordres larges. Nous le montrons dans la propriété suivante. Ainsi, dans ce qui suit, nous ne ferons aucune distinction entre ces deux termes et nous notons par $\mathbb{ROLT}(\mathbb{O})$ l'ensemble des relations d'ordre large et total possibles sur l'ensemble des objets \mathbb{O} .

Propriété 6.1.1 Lien entre préordre et ordre large

Toute relation réflexive et transitive est faiblement antisymétrique.

Démonstration

Considérons une relation de préordre total \mathcal{R} et soit C sa matrice relationnelle. Nous rappelons que les matrices relationnelles vérifient des contraintes linéaires relatives aux propriétés relationnelles de la relation binaire qu'elles représentent. Ces contraintes ont été données dans tab-1.2 page 4. En ce qui concerne l'antisymétrie faible, nous avons démontré son écriture sous forme de contraintes linéaires page 122.

Considérons justement la propriété d'antisymétrie faible. Cette dernière correspond aux contraintes suivantes :

$$\begin{cases} C_{ii'} + C_{i'i} + C_{ii''} - C_{i'i''} \leq 2 & \forall i, i', i'' = 1, \dots, N \\ C_{ii'} + C_{i'i} - C_{ii''} + C_{i'i''} \leq 2 & \forall i, i', i'' = 1, \dots, N \end{cases}$$

Plus particulièrement cette propriété sert à vérifier que lorsque deux objets sont symétriquement en relation, $C_{ii'} = C_{i'i} = 1$, alors ils doivent avoir exactement le même profil.

Nous montrons que dans le cas où $C_{ii'} = C_{i'i} = 1$, les inégalités de transitivité de la relation de préordre entraînent celles de l'antisymétrie faible. $\forall (O^i, O^{i'}) \in \mathbb{O} \times \mathbb{O}$, nous avons :

$$\begin{cases} C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 & \forall i'' \\ C_{i'i} + C_{ii''} - C_{i'i''} \leq 1 & \forall i'' \\ C_{ii'} = C_{i'i} = 1 \end{cases} \Rightarrow \begin{cases} C_{ii'} + C_{i'i} + C_{i'i''} - C_{ii''} \leq 1 + C_{i'i} & \forall i'' \\ C_{ii'} + C_{i'i} + C_{ii''} - C_{i'i''} \leq 1 + C_{ii'} & \forall i'' \\ C_{ii'} = C_{i'i} = 1 \end{cases}$$

$$\Rightarrow \begin{cases} C_{ii'} + C_{i'i} + C_{i'i''} - C_{ii''} \leq 2 & \forall i'' \\ C_{ii'} + C_{i'i} + C_{ii''} - C_{i'i''} \leq 2 & \forall i'' \end{cases}$$

L'hypothèse de réflexivité est nécessaire afin que l'implication ci-dessus soit valable quelque soit $i'' =$

$1, \dots, N$, ce qui correspond strictement à la définition de l'antisymétrie faible.

□

Les relations d'ordre que nous étudions désormais sont celles que l'on rencontre le plus souvent en pratique. Elles correspondent par exemple à des classements avec *ex-aequos*. Dans un premier temps, nous supposons que ces classements sont totaux, c'est à dire qui ne présentent pas de données manquantes. Nous prendrons comme exemple illustratif le classement \mathcal{R}' suivant, $O^3 < O^1 \sim O^2 < O^4$, que nous pouvons représenter par le tableau ci-dessous.

$$\mathcal{R}' = \begin{array}{|c|c|} \hline \text{Objets} & \text{Modalités ordonnées} \\ \hline O^1 & D^2 \\ \hline O^2 & D^2 \\ \hline O^3 & D^1 \\ \hline O^4 & D^4 \\ \hline \end{array}$$

où les D^j sont des rangs tels que $D^1 < D^2 < \dots < D^N$ dans le cas où il y a N objets.

Nous présentons rapidement les différents types classiques de codage de ces ordres. Puis nous exposons les liens qui lient un ordre large et total à une relation d'ordre strict et à une relation d'équivalence.

6.1.1.1 Le codage algébrique

Définition 6.1.3 Codage algébrique

Soit une relation d'ordre large et total \mathcal{R} de N objets qui sont classés selon les rangs ordonnés suivants : $D^1 < D^2 < \dots < D^N$. Le codage algébrique représentant \mathcal{R} est noté comme précédemment, κ , et est défini selon le formalisme suivant :

$$\kappa = \begin{pmatrix} O^1 & O^2 & \dots & O^N \\ \kappa(O^1) & \kappa(O^2) & \dots & \kappa(O^N) \end{pmatrix} \quad (6.1.1)$$

où $\kappa(O^i) \in \{D^1, \dots, D^N\}$, représente le rang auquel est affecté l'objet O^i selon \mathcal{R} .

- nous dirons que l'objet O^i est de rang strictement inférieur à celui de l'objet $O^{i'}$ selon κ si nous avons $\kappa(O^i) < \kappa(O^{i'})$
- nous dirons que l'objet O^i est *ex-aequo* avec l'objet $O^{i'}$ selon κ si nous avons $\kappa(O^i) = \kappa(O^{i'})$

La représentation de l'ordre large et total \mathcal{R}' selon le codage algébrique est la suivante :

$$\kappa' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^2 & D^2 & D^1 & D^4 \end{pmatrix}$$

6.1.1.2 Le codage par affectation

Définition 6.1.4 Codage par affectation

Soit une relation d'ordre large et total \mathcal{R} de N objets classés selon les rangs ordonnés suivants $D^1 < D^2 < \dots < D^N$. Le codage par affectation de cette relation d'ordre large et total consiste en une matrice carrée notée K de dimension N dont le terme général K_{ij} , $i, j = 1, \dots, N$; est donné par la relation suivante :

$$K_{ij} = \begin{cases} 1 & \text{si } O^i \text{ est de rang } D^j \\ 0 & \text{sinon} \end{cases} \quad (6.1.2)$$

K sera appelée matrice d'affectation

Le codage par affectation d'une relation d'ordre large et total est structurellement la même que celle d'une partition¹. En effet, chaque objet est affecté à un rang et plusieurs objets peuvent avoir un même rang. Ce que ne code pas de manière explicite la représentation par affectation est l'ordre induit sur les rangs.

Par ailleurs, dans le cas d'une variable qualitative ordonnée, les modalités ont une signification particulière et ne sont pas interchangeables. Plaçons nous dans un contexte différent où par exemple les éléments de \mathbb{D} n'ont pas de signification intrinsèque vis à vis de l'appartenance et indiquent uniquement un ordre de grandeur. Supposons que les descripteurs D^j soient en fait des rangs $j : D^j \leftrightarrow j \quad \forall j = 1, \dots, N$. Dans ce cas, il existe plusieurs types de matrices de permutations qui soient représentatives d'un même ordre large et total dans le cas où la relation possède plusieurs *ex-aequos*. Par exemple, nous donnons l'ordre κ'' suivant qui code le même ordre que celui de κ' :

$$\kappa'' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^2 & D^2 & D^1 & D^3 \end{pmatrix} \quad (6.1.3)$$

Nous ne nous attarderons pas sur la définition d'une relation d'équivalence entre matrices d'affectation de relations d'ordre large et total, comme nous l'avions fait pour le cas des partitions, car elle est dépendante de la nature des descripteurs. Cependant nous établissons la **convention** suivante :

Remarque 6.1.1 Convention sur les rangs pour les ordres avec ex-aequos

Supposons que nous ayons $\#D^j$ objets ayant le même rang D^j . Alors le rang du ou des objets qui sont directement supérieurs à D^j sera le rang $D^{j+\#D^j}$ et non D^{j+1} .

De cette manière et en commençant par $j = 1$, nous voyons, qu'il existe une et une seule représentation pour un ordre large, total et avec ex-aequos.

Si nous prenons l'exemple de κ' et κ'' , nous voyons qu'avec cette convention, c'est κ' l'unique représentant de l'ordre large et total \mathcal{R}' .

Propriété 6.1.2 Propriétés de la matrice K

- $\sum_j K_{ij} = K_{i.} = 1 \quad \forall i = 1, \dots, N$
- $\sum_i K_{.j} = K_{.j} = \#D^j \quad \forall j = 1, \dots, N$, le nombre d'objets ex-aequos de rang D^j
- $\sum_{i,j} K_{ij} = K_{..} = N$, nombre total d'objets

La matrice d'affectation représentative de \mathcal{R}' est la suivante :

$$K' = \begin{matrix} & D^1 & D^2 & D^3 & D^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix}$$

6.1.1.3 Le codage relationnel

Définition 6.1.5 Codage relationnel

Soit une relation d'ordre large et total \mathcal{R} de N objets classés selon les rangs ordonnés suivants $D^1 <$

1. cf page 154

$D^2 < \dots < D^N$. Le codage relationnel de cette relation d'ordre large et total consiste en une matrice carrée notée C de dimension N dont le terme général $C_{ii'}$ où $i, i' = 1, \dots, N$; est donné par la relation suivante :

$$C_{ii'} = \begin{cases} 1 & \text{si } O^i \text{ est classé avant } O^{i'} \text{ ou si } O^i \text{ est de même rang que } O^{i'} \text{ selon } \mathcal{R} \\ 0 & \text{sinon} \end{cases} \quad (6.1.4)$$

Propriété 6.1.3 Propriétés relationnelles de la matrice C

La matrice relationnelle vérifie les contraintes suivantes qui retranscrivent les propriétés relationnelles d'une relation d'ordre large et total :

- *Réflexivité* : $C_{ii} = 1 \quad \forall i = 1, \dots, N$
- *Antisymétrie faible* : $\begin{cases} C_{ii'} + C_{i'i} + C_{i''i''} - C_{ii''} \leq 2 & \forall i, i', i'' = 1, \dots, N \\ C_{ii'} + C_{i'i} + C_{ii''} - C_{i'i''} \leq 2 & \forall i, i', i'' = 1, \dots, N \end{cases}$
- *Transitivité* : $C_{ii'} + C_{i'i''} - C_{ii''} \leq 1 \quad \forall i, i', i'' = 1, \dots, N$
- *Totalité* : $C_{ii'} + C_{i'i} \geq 1 \quad \forall i, i' = 1, \dots, N : i \neq i'$

Le codage relationnel d'une relation d'ordre large et total représente explicitement l'ordre induit sur l'ensemble des objets par les rangs auxquels ces derniers sont affectés. De ce fait, contrairement au codage par affectation, la représentation relationnelle d'un ordre large et total est structurellement distincte de celle d'une partition.

Propriété 6.1.4 Propriété structurelle de la matrice C

$$N + 1 - C_{i.} = \text{rang de l'objet } O^i \quad \forall i = 1, \dots, N \quad (6.1.5)$$

Comme nous l'avons fait remarquer en [rq-5.1.2](#), il existe, dans le cadre des ordres, deux possibilités de détermination d'un ordre dual à un ordre \mathcal{R} représentée par C , sa matrice relationnelle : le complémentaire $\bar{C} = U_N - C$ et la converse ou transposée $\check{C} = {}^t C$. Contrairement aux ordres stricts et totaux et aux précédentes recherches entreprises en AR autour de l'étude des relations d'ordre, nous ne nous intéressons pas au complémentaire \bar{C} . En effet, pour un ordre large et total, celui-ci n'induit pas forcément une relation d'ordre². Par conséquent, **il est davantage vrai de représenter l'ordre dual d'un ordre large et total κ par sa converse ou transposée, $\check{\kappa}$. Cette dernière est en effet, un ordre large et total également étant donné que l'opération de transposition d'une matrice relationnelle permet à \check{C} de conserver les propriétés relationnelles de C . Cette approche nous permettra de simplifier de manière significative les notations des précédents travaux entamés dans les domaines qui nous concernent ici et de mettre en évidence de nouveaux résultats.**

Propriété 6.1.5 Ordre large et total engendré par la transposée d'une matrice relationnelle d'un ordre large et total

Soit \mathcal{R} une relation d'ordre large et total représentée par κ et sa matrice relationnelle C . Alors \check{C} est la matrice relationnelle représentative de l'ordre dual de C qui est de même nature que cette dernière. Elle est symboliquement noté $\check{\kappa}$ ou $\check{\mathcal{R}}$ et est donnée par :

$$\check{\kappa} = \begin{pmatrix} O^1 & \dots & O^i & \dots & O^N \\ D^{N+1-C_{.1}} & \dots & D^{N+1-C_{.i}} & \dots & D^{N+1-C_{.N}} \end{pmatrix} \quad (6.1.6)$$

Nous donnons l'exemple de la matrice représentative de la relation \mathcal{R}' dans son codage

2. le complémentaire de la propriété de transitivité n'est pas la transitivité mais l'inégalité triangulaire

relationnel ainsi que sa transposée :

$$C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} ; \quad \check{C}' = {}^t C' = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 1 \end{pmatrix} \end{matrix}$$

L'ordre dual induit par la transposée de κ' est le suivant :

$$\check{\kappa}' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^2 & D^2 & D^4 & D^1 \end{pmatrix}$$

6.1.1.4 Lien entre le codage par affectation et le codage relationnel

Propriété 6.1.6 Du codage par affectation au codage relationnel

Soit une relation d'ordre large et total \mathcal{R} et soit K sa matrice d'affectation associée. A partir de K , nous obtenons la matrice relationnelle C de \mathcal{R} , en appliquant la relation suivante :

$$C = K \cdot Q_N \cdot {}^t K \quad (6.1.7)$$

où Q_N est une matrice relationnelle de passage qui code la relation d'ordre sur les éléments de l'ensemble \mathbb{D} ainsi que la relation d'équivalence engendrée par la multi-affectation d'un descripteur D^j à plusieurs objets. Dans le cas où les éléments de \mathbb{D} sont ordonnés de la manière suivante, $D^1 < D^2 < \dots < D^N$, nous avons $Q_N = I_N + R_N$ qui est définie par :

$$[Q_N]_{jj'} = \begin{cases} 1 & \text{si } D^j \leq D^{j'} \text{ c'est à dire si } j \leq j' \\ 0 & \text{sinon} \end{cases} \quad (6.1.8)$$

$Q_N = I_N + R_N$ représente la matrice de passage standard dans le cas des ordres larges.

A partir de la définition donnée à la matrice de passage Q_N dans le cadre des ordres larges et totaux, nous mettons en évidence les parties symétriques et antisymétriques de ce type de relation définie dans [Gha90]. En fait, **une relation d'ordre large et total peut être vue comme étant l'addition d'une relation d'équivalence et une relation d'ordre strict.**

Définition 6.1.6 Parties symétrique et antisymétrique d'une relation d'ordre large

Soit \mathcal{R} une relation d'ordre large et total, et soit C sa matrice relationnelle associée. Nous avons la décomposition $C = \check{C} + \dot{C}$ où :

- $\check{C} = K \cdot I_N \cdot {}^t K$: la partie symétrique qui est une matrice relationnelle codant une relation d'équivalence
- $\dot{C} = K \cdot R_N \cdot {}^t K$: la partie asymétrique qui est une matrice relationnelle codant une relation d'ordre strict et non nécessairement total³

6.1.2 Critères d'association et nouvelles formulations relationnelles

Nous avons rappelé en introduction page 17, l'écriture relationnelle donnée dans [Gha90] des critères d'association entre relations d'ordre. Les critères rappelés sont ceux de Goodman et Kruskal, de Somers, de Kendall, de Deuchler, et de Kim.

3. ie partielle

Dans cette sous-section, nous donnons une nouvelle écriture relationnelle de ces critères qui est davantage symétrique. Cette nouvelle formulation permet de mettre en évidence l'équivalence entre le concept de concordance et d'indétermination.

La conséquence de ce résultat est une contribution nouvelle à l'unification entre les différentes mesures utilisées en statistiques. Nous montrons en effet, que certains critères d'association entre relations d'ordre sont des cas particuliers des indices d'écart à l'indétermination pondérée que nous avons présentés en sous-section 3.3.2 page 111 dans le cadre des relations d'équivalence.

Par conséquent, le concept d'écart à l'indétermination pondérée peut être étendu au cas des relations d'ordre. Cet aspect est exposé dans la sous-section suivante.

La nouvelle formulation relationnelle des critères d'association est obtenue à partir des formules de passage entre mesures de concordances⁴ et quantités relationnelles que nous donnons dans le tableau tab-6.1. Celles-ci, ainsi que les critères qu'elles définissent, sont valables pour des ordres stricts ou larges qui sont totaux ou non. Toutefois, il faut préciser, que lorsqu'il y a des données manquantes, la quantité N intervenant dans les formules du tableau tab-6.1 relatives à l'évaluation du nombre de paires en *ex-aequos*, doit être le nombre de données renseignées. Prenons le cas d'une relation d'ordre large relative à une variable V^k représentée par C^k sa matrice relationnelle associée dont le terme général est donné par (6.1.4). Alors le nombre d'objets pour lesquels le classement est renseigné est donné par⁵ $\sum_i C_{ii}^k$.

Concordance	\Leftrightarrow	Relationnel
$\text{Conc}(C^k, C^l) - \text{Disc}(C^k, C^l)$	$=$	$\frac{1}{2} \sum_{i=1}^N \sum_{i'=1}^N (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)$
$\text{Conc}(C^k, C^l) + \text{Disc}(C^k, C^l)$	$=$	$\frac{1}{2} \sum_{i,i'} (\bar{C}_{ii'}^k + \check{C}_{ii'}^k) (\bar{C}_{ii'}^l + \check{C}_{ii'}^l)$
$\text{Tie}(C^k)$	$=$	$\frac{N(N-1)}{2} - \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k)^2$
$\text{Tie}(C^k, C^l)$	$=$	$\frac{N(N-1)}{2} - \frac{1}{2} \sum_{i,i'} (C_{ii'}^k C_{ii'}^l - \check{C}_{ii'}^k \check{C}_{ii'}^l)^2$
où $\bar{C}_{ii'}^k = 1 - C_{ii'}^k$ et $\check{C}_{ii'}^k = C_{i'i}^k$		

TABLE 6.1 – Nouvelles formules de passage “concordance”/relationnel

Démonstration

Nous démontrons les nouvelles formules de correspondances. Pour les deux premiers termes, nous partons des formulations du tableau tab-6.1 et nous montrons que nous aboutissons aux formualtions de S. Ghashghaie données par tab-1.6 et plus particulièrement par (1.55).

4. qui ont été initialement introduites page 17

5. en effet, si l'objet O^i n'a pas de classement pour la variable V^k alors $C_{ii}^k = 0$ et de manière générale $C_{ii'}^k = C_{i'i}^k = 0, \forall i'$

$$\begin{aligned}
 \text{Conc}(C^k, C^l) - \text{Disc}(C^k, C^l) &= \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l) \\
 &= \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k) C_{ii'}^l - \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k) C_{i'i}^l \\
 &= \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k) C_{ii'}^l + \frac{1}{2} \sum_{i,i'} (C_{i'i}^k - C_{ii'}^k) C_{i'i}^l \\
 &\quad \text{en remplaçant dans le deuxième terme } i \text{ par } i' \text{ et vice versa nous avons} \\
 &= \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k) C_{ii'}^l + \frac{1}{2} \sum_{i',i} (C_{ii'}^k - C_{i'i}^k) C_{ii'}^l \\
 &= \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k) C_{ii'}^l \\
 \text{Conc}(C^k, C^l) + \text{Disc}(C^k, C^l) &= \frac{1}{2} \sum_{i,i'} (\bar{C}_{ii'}^k + \bar{C}_{i'i}^k) (\bar{C}_{ii'}^l + \bar{C}_{i'i}^l) \\
 &= \frac{1}{2} \sum_{i,i'} (1 - C_{ii'}^k + 1 - \check{C}_{ii'}^k) (1 - C_{ii'}^l + 1 - \check{C}_{ii'}^l) \\
 &= \frac{1}{2} \sum_{i,i'} (1 - C_{ii'}^k + 1 - C_{i'i}^k) (1 - C_{ii'}^l + 1 - C_{i'i}^l) \\
 &= \frac{1}{2} \sum_{i,i'} (2 - C_{ii'}^k - C_{i'i}^k) (1 - C_{ii'}^l + 1 - C_{i'i}^l) \\
 &= \frac{1}{2} \sum_{i,i'} (2 - C_{ii'}^k - C_{i'i}^k) (1 - C_{ii'}^l) + \frac{1}{2} \sum_{i,i'} (2 - C_{ii'}^k - C_{i'i}^k) (1 - C_{i'i}^l) \\
 &= \sum_{i,i'} (2 - C_{ii'}^k - C_{i'i}^k) (1 - C_{ii'}^l) \\
 &\quad \text{car } \sum_{i,i'} (2 - C_{ii'}^k - C_{i'i}^k) (1 - C_{i'i}^l) = \sum_{i',i} (2 - C_{i'i}^k - C_{ii'}^k) (1 - C_{ii'}^l) \\
 &\quad \text{en remplaçant } i \text{ par } i' \text{ et vice versa}
 \end{aligned}$$

Pour démontrer la nouvelle expression relationnelle de $\text{Tie}(C^k)$, le nombre de paires d'objets étant ex-aequos, nous appliquons le raisonnement suivant :

- $\frac{N(N-1)}{2} = \binom{N}{2}$ est le nombre de paires parmi les N objets (dont les données ont été renseignées dans le cas de données manquantes)
- Nous avons, $(C_{ii'}^k - C_{i'i}^k)^2$ qui vaut 1 si $\{C_{ii'}^k = 1 \wedge C_{i'i}^k = 0\}$ ou $\{C_{ii'}^k = 0 \wedge C_{i'i}^k = 1\}$. Ainsi, $(C_{ii'}^k - C_{i'i}^k)^2 = 1$ si les objets O^i et $O^{i'}$ ne sont pas ex-aequos. En sommant sur tous les paires (i, i') , et en tenant compte de la symétrie, nous constatons que $\frac{1}{2} \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)^2$ est le nombre de paires qui ne sont pas ex-aequos parmi les $\binom{N}{2}$ paires d'objets possibles.

De ce fait, le complémentaire $\frac{N(N-1)}{2} - \frac{1}{2} \sum_{i,i'} (C_{ii'}^k - C_{i'i}^k)^2$ est bien le nombre de paires qui sont ex-aequos parmi les $\binom{N}{2}$ paires d'objets possibles.

Etant donné que $\text{Tie}(C^k, C^l)$ est le nombre d'ex-aequos de la partition engendrée par l'intersection relationnelle $C^k \otimes C^l$, nous avons $\text{Tie}(C^k, C^l) = \text{Tie}(C^k \otimes C^l)$.

□

A l'aide des nouvelles formules de passages, nous donnons ci-dessous les nouvelles expressions relationnelles des différents critères d'association entre relations d'ordre. Mis à part les critères asymétriques de Somers, tous les autres critères d'association entre relations d'ordre considérées sont symétriques en C^k et C^l .

Propriété 6.1.7 Expression relationnelle du critère de Goodman et Kruskal

$$\text{GK}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)}{\sum_{i,i'} (\bar{C}_{ii'}^k + \bar{C}_{i'i}^k) (\bar{C}_{ii'}^l + \bar{C}_{i'i}^l)} \quad (6.1.9)$$

Propriété 6.1.8 Expression relationnelle du critère asymétrique de Somers

$$\text{Somers}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)}{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k)^2} \quad (6.1.10)$$

$$\text{Somers}(C^l, C^k) = \frac{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)}{\sum_{i,i'} (C_{ii'}^l - \check{C}_{ii'}^l)^2} \quad (6.1.11)$$

Propriété 6.1.9 *Expression relationnelle du critère de Kendall*

$$\text{Kendall}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)}{\sqrt{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k)^2 \sum_{i,i'} (C_{ii'}^l - \check{C}_{ii'}^l)^2}} \quad (6.1.12)$$

Propriété 6.1.10 *Expression relationnelle du critère de Deuchler*

$$\text{Deuchler}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)}{\sum_{i,i'} (C_{ii'}^k C_{ii'}^l - \check{C}_{ii'}^k \check{C}_{ii'}^l)^2} \quad (6.1.13)$$

Propriété 6.1.11 *Expression relationnelle du critère de Kim*

$$\text{Kim}(C^k, C^l) = \frac{\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k) (C_{ii'}^l - \check{C}_{ii'}^l)}{\frac{1}{2} \left(\sum_{i,i'} (C_{ii'}^k - \check{C}_{ii'}^k)^2 + \sum_{i,i'} (C_{ii'}^l - \check{C}_{ii'}^l)^2 \right)} \quad (6.1.14)$$

Les nouvelles expressions relationnelles des critères d'association entre relations d'ordre permettent de mettre en évidence le concept d'indétermination sous-jacent à ces derniers. En effet, nous avons discuté de la nature de l'ordre dual d'un ordre large. Nous avons retenu la converse $\check{C}^k = {}^t C^k$ comme étant le dual naturel d'un ordre large et total et non son complémentaire contrairement aux relations d'équivalence : c'est cette interprétation qui nous a permis de mettre en évidence la notion d'écart à l'indétermination.

Soient les quantités d'accords $11_{kl}, 00_{kl}$ et de désaccords $10_{kl}, 01_{kl}$, données par les expressions du tableau tab-6.2.

	$C_{ii'}^l$	$\check{C}_{ii'}^l$
$C_{ii'}^k$	$11_{kl} = \sum_{i,i'} C_{ii'}^k C_{ii'}^l$	$10_{kl} = \sum_{i,i'} C_{ii'}^k \check{C}_{ii'}^l$
$\check{C}_{ii'}^k$	$01_{kl} = \sum_{i,i'} \check{C}_{ii'}^k C_{ii'}^l$	$00_{kl} = \sum_{i,i'} \check{C}_{ii'}^k \check{C}_{ii'}^l$

TABLE 6.2 – Définition des accords et des désaccords entre deux matrices relationnelles d'ordre

Dans le cas des ordres, nous avons les particularités suivantes, $11_{kl} = 00_{kl}$ et $10_{kl} = 01_{kl}$. Par ailleurs, l'indétermination se traduit par :

$$\text{Conc}(C^k, C^l) - \text{Disc}(C^k, C^l) = 0 \Leftrightarrow 11_{kl} + 00_{kl} - 10_{kl} - 01_{kl} = 0 \quad (6.1.15)$$

En effet, nous avons :

$$\begin{aligned} 11_{kl} + 00_{kl} - 10_{kl} - 01_{kl} = 0 &\Leftrightarrow \sum_{i,i'} \left(C_{ii'}^k C_{ii'}^l + \check{C}_{ii'}^k \check{C}_{ii'}^l - C_{ii'}^k \check{C}_{ii'}^l - \check{C}_{ii'}^k C_{ii'}^l \right) = 0 \\ &\Leftrightarrow \sum_{i,i'} \left(C_{ii'}^k - \check{C}_{ii'}^k \right) \left(C_{ii'}^l - \check{C}_{ii'}^l \right) = 0 \end{aligned}$$

6.1.3 Critères d'écart à l'indétermination pondérée entre relations d'ordre

Nous avons montré dans la sous-section précédente que les critères d'association entre relations d'ordre s'annulaient en cas d'indétermination.

Nous mettons plus particulièrement en exergue les critères de Kendall et de Kim dont les nouvelles formulations relationnelles rappellent :

- celles des indices de similarité d'ordre t dans le cas où les matrices relationnelles sont représentées dans un espace vectoriel particulier. Nous avons introduit ces mesures de similarités lors de l'étude de l'extension des indices d'Ochiaï et de Dice au cas des variables quantitatives en déf-3.1.11 page 82
- celles des indices d'indétermination pondérée que nous avons présentés dans le cadre des relations d'équivalence en déf-3.3.6 page 111

Dans le cas des ordres nous prenons \check{C}^k , la converse de C^k , comme étant l'opération unaire donnant le dual de C^k et non \bar{C}^k pour les raisons que nous avons déjà évoquées précédemment. Par ailleurs, nous sommes dans le cas d'un système de pondération uniforme pour lequel $\mu_0^k = \mu_0^l = \mu_1^k = \mu_1^l = 1/2$. Dans le contexte que nous venons de décrire, nous avons alors les propriétés suivantes :

Propriété 6.1.12 *Critères de Kendall et de Kim vus sous l'angle des indices de similarité entre vecteurs et les indices d'écart à l'indétermination*

$$\boxed{\text{Kendall}(C^k, C^l) = \text{Ochiaï}(\vec{C}^k - \check{C}^k, \vec{C}^l - \check{C}^l) = S_I^{t \rightarrow 0}(C^k, C^l, 1/2, 1/2, 1/2, 1/2)} \quad (6.1.16)$$

$$\boxed{\text{Kim}(C^k, C^l) = \text{Dice}(\vec{C}^k - \check{C}^k, \vec{C}^l - \check{C}^l) = S_I^1(C^k, C^l, 1/2, 1/2, 1/2, 1/2)} \quad (6.1.17)$$

Nous rappelons que dans l'interprétation des indices de similarité d'ordre t , le produit scalaire entre deux matrices relationnelles est celui de Frobenius donné⁶ par $\langle \vec{C}^k, \vec{C}^l \rangle_F = \sum_{i,i'} C_{ii'}^k C_{ii'}^l = \text{Trace}({}^t C^k \cdot C^l)$.

De manière générale, nous pouvons étendre les indices d'écart à l'indétermination que nous avons définis pour les relations d'équivalence aux relations d'ordre.

Définition 6.1.7 *Indice d'ordre t d'écart à l'indétermination pondérée entre deux matrices relationnelles d'ordres*

Soient deux variables ordonnées V^k et V^l représentées par leurs matrices relationnelles d'ordres C^k et C^l . Nous appellerons indice d'écart à l'indétermination pondérée d'ordre $t > 0$ entre les matrices

6. cf page 108

relationnelles, muni de poids séparables⁷ de paramètres $(\mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l)$, l'indice défini de la manière suivante :

$$\begin{aligned}
 S_I^t(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) &= \frac{\langle C^k, C^l \rangle_F}{\mathcal{M}^t(\langle C^k, C^k \rangle_F, \langle C^l, C^l \rangle_F)} \\
 &= \frac{\langle C^k, C^l \rangle_F}{\mathcal{M}^t(\sum_{i,i'} C_{ii'}^k, \sum_{i,i'} C_{ii'}^l)}
 \end{aligned}
 \tag{6.1.18}$$

où

$$\begin{aligned}
 C^k &= \mu_1^k C^k - \mu_0^k \check{C}^k \\
 C^l &= \mu_1^l C^l - \mu_0^l \check{C}^l
 \end{aligned}
 \tag{6.1.19}$$

Nous nous arrêtons par ailleurs, sur les critères asymétriques de Somers par aller plus loin dans notre interprétation géométrique des matrices relationnelles d'ordres et leurs mesures d'association. En effet, les critères asymétriques de Somers peuvent être rapprochés des coefficients de projection que nous avons présentés au chapitre 3, en propriété prop-3.1.21 page 85 et qui permettent de définir d'une deuxième façon les indices d'indétermination d'ordre t . Nous avons ainsi la propriété suivante :

Propriété 6.1.13 Expression de S_I^t en termes de moyennes de critères de Somers

Soient deux variables ordonnées V^k et V^l représentées par leurs matrices relationnelles d'ordres C^k et C^l . Considérons les matrices $C^k = C^k - \check{C}^k$ et $C^l = C^l - \check{C}^l$. Nous avons dans un premier temps la propriété suivante :

$$\text{Somers}(C^k, C^l) = P_{C^k}(C^l) = \frac{\langle C^k, C^l \rangle_F}{\langle C^k, C^k \rangle_F}
 \tag{6.1.20}$$

où C^k et C^l sont considérées comme étant des vecteurs dans un espace vectoriel approprié. Conformément à la propriété prop-3.1.21 donnée page 85, nous avons la propriété suivante :

$$S_I^t(C^k, C^l, 1/2, 1/2, 1/2, 1/2) = \mathcal{M}^{-t}(\text{Somers}(C^k, C^l), \text{Somers}(C^l, C^k)) \quad \forall t > 0
 \tag{6.1.21}$$

Ainsi les critères de Kendall et de Kim sont respectivement les moyennes géométriques ($t \rightarrow 0$) et arithmétiques ($t = 1$) des coefficients asymétriques de Somers.

Par ailleurs, la propriété reste vraie pour le cas général de l'indétermination pondérée où les matrices C^k et C^l sont définies par (6.1.19) et où nous définissons le critère asymétrique de Somers pondéré de la manière suivante :

$$\text{Somers}(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) = \frac{\langle C^k, C^l \rangle_F}{\langle C^k, C^k \rangle_F}
 \tag{6.1.22}$$

Nous avons alors $\forall t > 0$:

$$S_I^t(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) = \mathcal{M}^{-t} \left(\begin{array}{c} \text{Somers}(C^l, C^k, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l), \\ \text{Somers}(C^k, C^l, \mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l) \end{array} \right)
 \tag{6.1.23}$$

Nous pouvons dire que, le formalisme relationnel, a permis d'unifier, dans un certain sens, des mesures d'association entre des partitions d'une part et des ordres d'autre part. **L'écart à**

7. cf page 109

L'indétermination pondérée joue dans ce cadre un rôle central puisque c'est la notion logique à partir de laquelle nous pouvons définir des critères d'association communs aux deux types de relations. Nous avons par ailleurs proposé un modèle général de ces critères au travers des indices d'écart à l'indétermination pondérée. Ces derniers permettent de tenir compte de la différence de structure entre deux relations en termes de densités⁸. Plus deux relations sont de densités distinctes, plus leur indice d'indétermination pondérée sera faible :

- dans le cas des partitions, plus la distribution des tailles des classes sont distinctes entre deux relations d'équivalence, plus leur indice d'écart à l'indétermination pondérée sera faible
- dans le cas des ordres totaux, plus le nombre d'*ex-aequos* ou de données manquantes sont distincts entre deux classements, plus leur indice d'écart à l'indétermination pondérée sera faible

6.2 L'effet Condorcet, le théorème d'impossibilité d'Arrow et les conditions de restriction sur le domaine des préférences individuelles

Dans cette section, nous nous intéressons à une problématique issue du domaine des sciences économiques. Il s'agit de la théorie du choix social qui est une thématique qui semble être différente des problèmes que nous avons étudiés précédemment.

Cependant, la théorie du choix social s'intéresse notamment au problème de l'agrégation de relations de préférences. Il s'agit en fait de déterminer une règle d'agrégation, qui permette de déduire d'un ensemble de M relations d'ordre, une relation d'ordre collective. Ce problème entre ainsi dans le cadre de l'AR et est proche du problème de la classification automatique qui est l'agrégation de relations d'équivalence.

Nous nous intéressons aux aspects axiomatiques de ce problème. Nous commençons alors par rappeler un résultat essentiel qui est dû à K.J. Arrow et qui a largement contribué à l'obtention de son prix Nobel en économie en 1972 [Arr74]. Nous rappelons aussi une partie des différents travaux qu'a engendrés ce théorème. Ils consistent à déterminer des conditions que doit vérifier l'ensemble des préférences individuelles pour ne pas générer d'effet Condorcet⁹.

Nous avons cherché à interpréter ces conditions dans le formalisme de l'algèbre relationnelle et l'algèbre matricielle. Nous nous basons alors sur la matrice de Condorcet collective C qui nous permet de définir une matrice de contributions C dont les éléments appartiennent à un dioïde particulier. Nous étudions ces propriétés algébriques pour différentes règles majoritaires dont celle de Condorcet.

Notre objectif est de généraliser de manière significative les résultats obtenus dans ce domaine. Pour ce faire nous nous plaçons dans un formalisme basé sur les matrices de comparaisons par paires dont les termes sont des éléments d'une structure algébrique particulière.

Remarque 6.2.1 Notations

Dans le cadre de cette section, les objets sont des candidats ou alternatives et sont au nombre de N . Les variables sont des votants ou juges et sont au nombre de M . De cette manière, \mathcal{O} et \mathcal{V} sont respectivement l'ensemble des candidats et l'ensemble des votants.

De plus, nous dirons qu'un candidat O^i est au moins aussi préféré à un autre candidat O^j , pour le

8. qui varie selon le nombre d'objets *ex-aequos* ou le nombre de données manquantes

9. cf introduction page 16

votant V^k , si le rang de O^i est plus petit ou égal à celui de $O^{i'}$. Cette proposition correspond à la notation $O^i \mathcal{R}^k O^{i'}$ et à l'écriture symbolique $O^i \leq O^{i'}$: " O^i est classé avant ou au même rang que $O^{i'}$ ". Nous faisons ainsi le lien entre relation de préférence et relation d'ordre.

6.2.1 Rappels en théorie du choix social

6.2.1.1 Le théorème d'impossibilité d'Arrow

Afin de rappeler le théorème d'impossibilité d'Arrow, nous rappelons tout d'abord, les axiomes sur lequel celui-ci repose :

1. axiome d'universalité
2. axiome de non dictature
3. axiome de Pareto par paires
4. axiome d'indépendance

Définition 6.2.1 Axiome d'universalité

L'ensemble des relations de préférences exprimées par les votants sont tout type de classement avec la possibilité d'avoir des ex-aequos. Ces classements engendrent des relations de préordre total ou d'ordre large et total¹⁰. En d'autres termes, les matrices relationnelles représentatives des classements individuels appartiennent à $\text{ROLT}(\mathbb{O})$.

Définition 6.2.2 Axiome de non dictature

Il ne doit pas exister un votant V^k tel que, quels que soient les candidats O^i et $O^{i'}$ et quelle que soit l'opinion des autres votants, la préférence collective entre ces deux candidats soit imposée par celle du votant V^k .

Définition 6.2.3 Axiome de Pareto par paires ou d'unanimité

Quels que soient les candidats O^i et $O^{i'}$, si chaque votant préfère strictement le candidat O^i au candidat $O^{i'}$ alors il doit en être de même pour le choix collectif.

Définition 6.2.4 Axiome d'indépendance (vis à vis des alternatives non pertinentes)

Soient $\kappa^1, \dots, \kappa^M$ et $\kappa'^1, \dots, \kappa'^M$ deux ensembles de classements individuels. Considérons une paire de candidats $(O^i, O^{i'})$. Si les restrictions à O^i et $O^{i'}$ des préférences individuelles de chaque votant sont identiques dans les deux ensembles, alors les deux préférences collectives concernant ces candidats doivent aussi être les mêmes pour ces deux ensembles.

Nous avons dès lors le théorème d'impossibilité d'Arrow suivant :

Théorème 6.2.1 Théorème d'impossibilité d'Arrow

Si $M \geq 2$ et $N \geq 3$ alors il n'existe pas de règle d'agrégation dont le résultat soit un préordre total et qui vérifie simultanément les quatres axiomes présentés précédemment.

10. nous avons montré que ces types de relations sont équivalents en prop-182

Ce théorème a eu un impact considérable en sciences économiques et politiques et a engendré une vaste littérature tant son impact dans le cadre de la notion de démocratie est grande.

K.J. Arrow proposa différents axes de recherches relatifs à son théorème afin d'enrichir sa théorie :

1. permettre au choix collectif d'être moins restrictif qu'un préordre total
2. **restreindre l'ensemble possible des préférences individuelles à des domaines particuliers de sorte que la règle d'agrégation aboutisse à un préordre total**
3. travailler autour des axiomes et notamment de celui de l'indépendance qu'il suggère de rendre moins restrictif voire d'abandonner

Parmi ces extensions, nous nous intéressons plus particulièrement au deuxième point. Nous rappelons dans le paragraphe suivant un certain nombre de résultats importants concernant celui-ci.

6.2.1.2 Les conditions de restriction sur l'ensemble des préférences individuelles aboutissant à des préférences collectives acycliques par la règle de la majorité simple

Dans le cadre de ce paragraphe, nous adopterons la règle d'agrégation dite règle de la majorité simple par paire.

Définition 6.2.5 Règle de la majorité simple (par paire) et représentation relationnelle

Un candidat O^i sera collectivement préféré ou indifférent à un candidat $O^{i'}$ si le nombre total de voix supportant " O^i préféré ou indifférent à $O^{i'}$ " est supérieur ou égal au nombre de voix supportant " $O^{i'}$ préféré ou indifférent à O^i ".

Dans le formalisme relationnel, notons par C^k les matrices relationnelles individuelles représentatives des préférences exprimées par chaque votant V^k . Nous avons $\mathbf{C} = \sum_k C^k$ qui est la matrice relationnelle collective aussi appelée matrice de Condorcet collective.

Soit X la matrice relationnelle représentative de la relation obtenue en appliquant la règle de la majorité simple. Nous avons alors :

$$X_{ii'} = \begin{cases} 1 & \text{si } \mathbf{C}_{ii'} \geq \check{\mathbf{C}}_{ii'} \text{ ou encore si } \mathbf{C}_{ii'} \geq \mathbf{C}_{i'i} \\ 0 & \text{sinon} \end{cases} \quad (6.2.1)$$

La règle de la majorité simple attribuée à K.O. May¹¹ [May52] est une règle d'agrégation qui vérifie les axiomes de non dictature, de Pareto par paires et d'indépendance. De plus, si nous n'imposons aucune contrainte quant aux préordres totaux que peuvent exprimer les votants, nous sommes dans le contexte du théorème d'Arrow présenté dans le paragraphe précédent. Malheureusement, cette règle d'agrégation ne conduit pas à une relation collective qui soit un préordre total, du fait de la non transitivité des préférences collectives obtenues.

En effet, il est facile de vérifier que la matrice relationnelle X est réflexive et totale. En revanche, il arrive très souvent que X ne soit pas transitive. Ce phénomène est l'**effet Condorcet**¹² : sur un triplet de candidats $(O^i, O^{i'}, O^{i''})$, nous pouvons avoir $X_{ii'} = X_{i'i''} = X_{i''i} = 1$ ce qui implique le cycle¹³ $O^i \mathcal{R}^X O^{i'}, O^{i'} \mathcal{R}^X O^{i''}, O^{i''} \mathcal{R}^X O^i$ et par conséquent une incohérence dans les préférences collectives.

11. mais initialement proposée sous forme littérale par M.J.A Condorcet. D'ailleurs, dans le cas de relations d'ordre strict et total, la règle de la majorité simple est similaire à la règle majoritaire de Condorcet défini par le critère $\text{Condorcet}(\mathbf{C}, X)$

12. cf page 16

13. ie dans ce cas précis une relation non transitive

De nombreuses recherches ont été entreprises visant à déterminer des conditions sur les préférences individuelles qui permettent d'obtenir une relation collective X qui soit transitive. Nous les rappelons ci-dessous.

- conditions d'unimodalité supérieure de D. Black [Bla48]
- conditions d'unimodalité inférieure de K. Inada [Ina64]
- conditions de séparabilité en deux groupes de K. Inada [Ina64]
- conditions de restriction sur des triplets de A.K. Sen [Sen66] qui généralisent les précédentes conditions

Nous nous basons sur l'article récent de C. Elsholtz et C. List [EL05]. En particulier, ces derniers donnent une formulation des conditions de Sen plus simple à appréhender.

L'ensemble des résultats que nous allons rappeler se situe dans le contexte de la règle de la majorité simple et raisonne donc sur des triplets de candidats en s'appuyant sur la propriété suivante :

Propriété 6.2.1 Cycle sur des triplets

S'il existe un cycle¹⁴ sur k candidats, pour $3 < k \leq N$, dans la relation collective X alors ceci implique qu'il existe au moins un cycle sur 3 candidats dans X .

Démonstration

La démonstration de ce résultat pourra être trouvée dans [Sen66] ou [EL05].

□

En considérant la contraposée de la propriété précédente, nous pouvons donc raisonner sur trois candidats au lieu de N : **si pour tout triplet il n'existe pas de cycle, alors la relation collective est acyclique.**

Les définitions que nous allons présenter sont des conditions suffisantes sur l'ensemble des préférences individuelles. Si tout triplet respecte un type de conditions énoncé, alors la relation collective obtenue par la règle de la majorité simple, est acyclique. Dans les cas considérés, la relation collective est même transitive. Les conditions de Black et d'Inada concernent des relations d'ordre strict et total et supposent que le nombre de votants est impair. Les conditions de Sen sont davantage générales : elles concernent les préordres totaux et elles supposent uniquement que le nombre de votants qui ne sont pas indifférents¹⁵ entre les 3 candidats du triplet, c'est à dire pour lesquels nous n'avons pas $O^i \sim O^{i'} \sim O^{i''}$, est impair.

Soit un ensemble de relations d'ordre strict et total \mathcal{R}^k , $k = 1, \dots, M$, relatives à des préférences exprimées par un ensemble de M votants \mathbb{V} sur un ensemble de N alternatives \mathbb{O} .

Propriété 6.2.2 Conditions suffisantes d'unimodalité supérieure ou de "single-peakedness" par rapport à une permutation de référence, de D. Black

L'ensemble de relations \mathcal{R}^k , $k = 1, \dots, M$, respecte les conditions d'unimodalité supérieure de D. Black s'il existe une bijection $\pi : \mathbb{O} \rightarrow \mathbb{N}_N$ où $\mathbb{N}_N = \{1, \dots, N\}$ telle que pour tout triplet $(O^i, O^{i'}, O^{i''})$ et

14. ce que nous entendons ici par cycle peut être rapproché de la définition des cycles données dans le cas du groupe des permutations par déf-5.1.9 page 162. Toutefois, il faut préciser que dans cette section " O^i préféré ou indifférent à $O^{i'}$ " se note $O^i \mathcal{R} O^{i'}$ et sera représenté symboliquement par $O^i \leq O^{i'}$. Le lien entre relations d'ordre et relations de préférences doit être vu de la manière suivante : $O^i \leq O^{i'}$ signifie le classement donné au candidat O^i est inférieur ou égal à celui donné à $O^{i'}$ ce qui implique que le candidat O^i est préféré ou indifférent à $O^{i'}$

15. "concerned judges" en anglais

pour tout votant V^k , $k = 1, \dots, M$, nous avons la propriété suivante :

$$\left\{ \left\{ \pi(O^i) < \pi(O^{i'}) < \pi(O^{i''}) \right\} \vee \left\{ \pi(O^{i''}) < \pi(O^{i'}) < \pi(O^i) \right\} \right\} \Rightarrow \left\{ O^i \mathcal{R}^k O^{i'} \Rightarrow O^{i''} \mathcal{R}^k O^{i''} \right\} \quad (6.2.2)$$

Propriété 6.2.3 Conditions suffisantes d'unimodalité inférieure ou de "single-cavedness" par rapport à une permutation de référence, de K. Inada

L'ensemble de relations \mathcal{R}^k , $k = 1, \dots, M$, respecte les conditions d'unimodalité inférieure de K. Inada s'il existe une bijection $\pi : \mathbb{O} \rightarrow \mathbb{N}_N$ où $\mathbb{N}_N = \{1, \dots, N\}$ telle que pour tout triplet $(O^i, O^{i'}, O^{i''})$ et pour tout votant V^k , $k = 1, \dots, M$, nous avons la propriété suivante :

$$\left\{ \left\{ \pi(O^i) < \pi(O^{i'}) < \pi(O^{i''}) \right\} \vee \left\{ \pi(O^{i''}) < \pi(O^{i'}) < \pi(O^i) \right\} \right\} \Rightarrow \left\{ O^{i'} \mathcal{R}^k O^i \Rightarrow O^{i''} \mathcal{R}^k O^{i'} \right\} \quad (6.2.3)$$

Il existe des versions plus "faibles" des deux conditions présentées précédemment. Celles-ci autorisent différentes bijections π pour chaque triplet d'alternatives. Ces conditions suffisent pour éviter les phénomènes de cycles.

Les conditions précédentes engendrent une représentation graphique particulière sur les classements de chaque votant qui précise mieux la dénomination donnée à ces propriétés. Nous donnons deux exemples sur la figure fig-6.1.

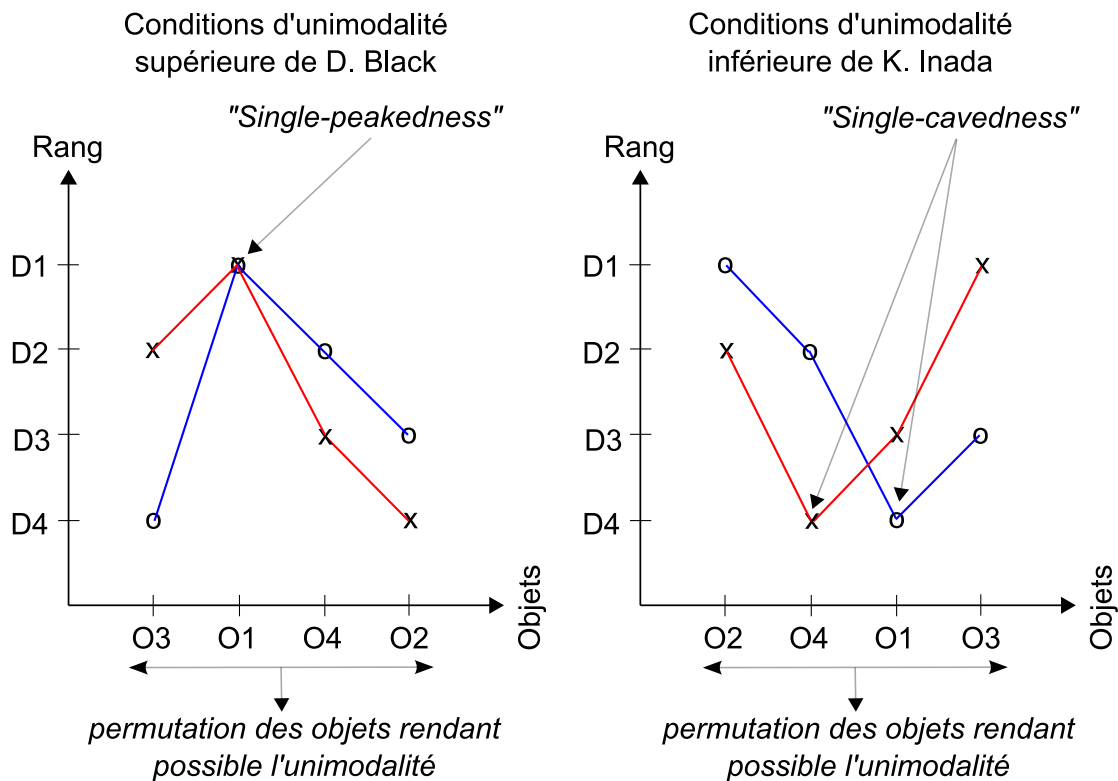


FIGURE 6.1 – Illustration des conditions d'unimodalité de D. Black et de K. Inada

Propriété 6.2.4 Conditions suffisantes de séparabilité en deux groupes de K. Inada

L'ensemble de relations \mathcal{R}^k , $k = 1, \dots, M$, respecte les conditions de séparabilité (en deux groupes) si celui-ci peut être séparé en deux groupes disjoints et non vides D^j et $D^{j'}$ tel que pour chaque votant V^k la propriété suivante est vérifiée :

$$\left\{ \forall O^i \in D^j \text{ et } \forall O^{i'} \in D^{j'} : O^i \mathcal{R}^k O^{i'} \right\} \vee \left\{ \forall O^i \in D^j \text{ et } \forall O^{i'} \in D^{j'} : O^{i'} \mathcal{R}^k O^i \right\} \quad (6.2.4)$$

Avant de rappeler les conditions de Sen, il est nécessaire de rappeler la définition des "valeurs" que peut prendre chaque candidat d'un triplet.

Définition 6.2.6 Valeur des candidats d'un triplet

Soit $\{O^i, O^{i'}, O^{i''}\}$ un triplet d'alternatives. Nous pouvons associer à chaque candidat $O \in \{O^i, O^{i'}, O^{i''}\}$ une valeur $v \in \{1, 2, 3\}$ de la manière suivante :

- 1 si le rang de O est plus petit que les deux autres ("best")
- 2 si le rang de O est compris entre le rang des deux autres ("medium")
- 3 si le rang de O est plus grand que les deux autres ("worst")

Propriété 6.2.5 Conditions suffisantes de restriction sur des triplets de candidats de A.K. Sen

L'ensemble de relations \mathcal{R}^k , $k = 1, \dots, M$, respecte les conditions de restriction¹⁶ de A.K. Sen, si pour chaque triplet de candidats $(O^i, O^{i'}, O^{i''})$, il existe $O \in \{O^i, O^{i'}, O^{i''}\}$ et une valeur $v \in \{1, 2, 3\}$ tel qu'aucun votant n'attribue à O la valeur v pour ce triplet.

Les conditions de Sen sont plus générales que celles de de Black et d'Inada. En effet, elles s'appliquent à des relations de préordres totaux qui sont plus larges que les ordres stricts et totaux. De plus, elles contiennent ces dernières dans la mesure où tout triplet vérifiant les conditions de Sen vérifient au moins une des conditions de Black ou d'Inada [Sen66]. De la même manière les hypothèses concernant le nombre de votants sont plus laxistes : les conditions de Sen supposent que c'est le nombre de votants qui ne sont pas indifférents par rapport aux trois candidats du triplet, qui est impair alors que les autres conditions supposent, de manière plus restrictive, que c'est le nombre **total** de votants qui est impair.

Dans [EL05], C. Elsholtz et C. List donnent une définition des conditions de Sen dans un formalisme différent. Ce dernier est relativement proche des matrices relationnelles. Leur approche consiste à énumérer pour chaque triplet, tous les préordres totaux possibles et de ne considérer que la partie asymétrique¹⁷ des préordres totaux. Nous tenons alors compte pour chaque triplet $(O^i, O^{i'}, O^{i''})$, des différents cas suivants :

- $O^i \leq O^{i'} \leq O^{i''}$
- $O^i \leq O^{i''} \leq O^{i'}$
- $O^{i'} \leq O^i \leq O^{i''}$
- $O^{i'} \leq O^{i''} \leq O^i$
- $O^{i''} \leq O^i \leq O^{i'}$
- $O^{i''} \leq O^{i'} \leq O^i$
- $O^i \sim O^{i'} \leq O^{i''}$
- $O^i \sim O^{i''} \leq O^{i'}$
- $O^i \leq O^{i'} \sim O^{i''}$
- $O^{i'} \sim O^{i''} \leq O^i$
- $O^{i'} \leq O^i \sim O^{i''}$
- $O^{i''} \leq O^i \sim O^{i'}$

Pour chacune de ces configurations nous les précédons du symbole # pour désigner le cardinal de l'ensemble des votants dont la restriction des préférences au triplet étudié correspond à la configuration considérée. Par exemple $\#(O^i \leq O^{i'} \leq O^{i''})$ donne le nombre de votants dont le préordre total restreint aux candidats $(O^i, O^{i'}, O^{i''})$ donne l'ordre $O^i \leq O^{i'} \leq O^{i''}$. Les conditions de Sen peuvent alors être exprimées de la manière suivante :

Propriété 6.2.6 Conditions suffisantes de restriction sur des triplets de candidats données par Elsholtz et List

L'ensemble de relations \mathcal{R}^k , $k = 1, \dots, M$, respecte les conditions de Sen, si pour chaque triplet de candidats $(O^i, O^{i'}, O^{i''})$, la proposition suivante est vérifiée :

16. conditions de "triplewise value restriction" en anglais

17. au sens de la définition donnée par déf-6.1.6 page 186

$$\left. \begin{array}{l}
 \left\{ \begin{array}{l}
 \#(O^i \leq O^{i'} \leq O^{i''}) + \#(O^i \sim O^{i'} \leq O^{i''}) + \#(O^i \leq O^{i'} \sim O^{i''}) = 0 \\
 \#(O^{i''} \leq O^i \leq O^{i'}) + \#(O^i \sim O^{i''} \leq O^{i'}) + \#(O^{i''} \leq O^i \sim O^{i'}) = 0 \\
 \#(O^{i'} \leq O^{i''} \leq O^i) + \#(O^{i'} \sim O^{i''} \leq O^i) + \#(O^{i'} \leq O^i \sim O^{i''}) = 0
 \end{array} \right. \\
 \wedge \\
 \left\{ \begin{array}{l}
 \#(O^{i''} \leq O^{i'} \leq O^i) + \#(O^{i'} \sim O^{i''} \leq O^i) + \#(O^{i''} \leq O^i \sim O^{i'}) = 0 \\
 \#(O^{i'} \leq O^i \leq O^{i''}) + \#(O^i \sim O^{i'} \leq O^{i''}) + \#(O^{i'} \leq O^i \sim O^{i''}) = 0 \\
 \#(O^i \leq O^{i''} \leq O^{i'}) + \#(O^i \sim O^{i''} \leq O^{i'}) + \#(O^i \leq O^{i'} \sim O^{i''}) = 0
 \end{array} \right.
 \end{array} \right\} \quad (6.2.5)$$

Démonstration

Ce résultat est démontré dans [EL05] et repose sur l'explicitation logique des conditions de Sen et de l'équation (6.2.5) : si on distribue cette dernière condition selon \wedge , nous explicitons chacun des 9 cas possibles où chaque candidat du triplet n'est pas associé à une des trois valeurs $v \in \{1, 2, 3\}$. Par exemple la conjonction du premier élément de la première parenthèse et du premier élément de la seconde parenthèse donne la proposition atomique suivante : $\{\#(O^i \leq O^{i'} \leq O^{i''}) + \#(O^i \sim O^{i'} \leq O^{i''}) + \#(O^i \leq O^{i'} \sim O^{i''}) = 0\}$ et $\{\#(O^{i''} \leq O^i \leq O^{i'}) + \#(O^{i'} \sim O^{i''} \leq O^{i'}) + \#(O^{i''} \leq O^i \sim O^{i'}) = 0\}$. Celle-ci exprime le fait que le candidat $O^{i'}$ n'est jamais en position intermédiaire "medium" vis à vis des candidats O^i et $O^{i''}$.

Nous donnons un deuxième exemple en considérant le premier élément de la première parenthèse et le troisième élément de la seconde parenthèse. Nous obtenons la proposition atomique suivante : $\{\#(O^i \leq O^{i'} \leq O^{i''}) + \#(O^i \sim O^{i'} \leq O^{i''}) + \#(O^i \leq O^{i'} \sim O^{i''}) = 0\}$ et $\{\#(O^i \leq O^{i''} \leq O^{i'}) + \#(O^i \sim O^{i''} \leq O^{i'}) + \#(O^i \leq O^{i'} \sim O^{i''}) = 0\}$. Celle-ci exprime alors le fait que le candidat O^i n'est jamais en position "best" vis à vis des candidats $O^{i'}$ et $O^{i''}$.

□

Nous venons de rappeler quelques résultats majeurs concernant l'axiomatique du problème de l'agrégation des préférences individuelles donnant des préférences collectives transitives. Dans la sous-section suivante, nous formalisons le problème dans un cadre relationnel et exploitons les comparaisons par paires afin de donner des conditions sur les propriétés des relations consensuelles obtenues par différentes règles majoritaires. Les conditions que nous allons donner, ont la particularité d'utiliser des outils de l'algèbre relationnelle et de l'algèbre matricielle et sont par ailleurs plus générales.

6.2.2 Caractérisations algébriques de la nature de la relation collective obtenue par différentes règles majoritaires dans un cadre relationnel

Dans cette sous-section, nous utiliserons comme algèbre de base le dioïde $\{\mathbb{Z}_M, \oplus = \max, \otimes = \min\}$ où \mathbb{Z}_M est l'ensemble des entiers relatifs suivant $\{-M, -M+1, \dots, -1, 0, 1, \dots, M-1, M\}$. Nous rappelons au préalable, la définition d'un dioïde.

Définition 6.2.7 Dioïde $(\mathbb{A}, \oplus, \otimes)$

$(\mathbb{A}, \oplus, \otimes)$ est un dioïde si \mathbb{A} est un ensemble d'éléments muni de deux lois de composition internes \oplus et \otimes qui vérifie les propriétés suivantes :

- (\mathbb{A}, \oplus) est un monoïde¹⁸ commutatif avec élément neutre ϵ

18. nous avons rappeler brièvement la définition d'un monoïde page 41

- (\mathbb{A}, \otimes) est un monoïde avec élément neutre e
- la relation de préordre canonique relativement à \oplus définie par $a \leq b \Leftrightarrow \exists c : b = a \oplus c$, est une relation d'ordre vérifiant la propriété d'antisymétrie faible
- ϵ est absorbant pour \otimes , c'est à dire, $\forall a \in \mathbb{A} : a \otimes \epsilon = \epsilon \otimes a = \epsilon$
- \otimes est distributive à droite et à gauche relativement à \oplus

Il est facile de vérifier que l'ensemble \mathbb{Z}_M associé, muni des lois de compositions internes $\oplus = \max$ et $\otimes = \min$ est un dioïde dont les éléments neutres sont $\epsilon = -M$ et $e = M$. Les dioïdes sont des structures algébriques moins riches que les anneaux. Cependant, elles permettent de modéliser de nombreux problèmes combinatoires dans la théorie des graphes notamment, comme il est montré dans l'ouvrage de M. Gondran et M. Minoux [GM01].

L'objectif de cette sous-section est de généraliser de manière significative les résultats rappelés précédemment. Notre approche est originale dans la mesure où elle associe plusieurs formalismes :

- nous utilisons les matrices relationnelles de comparaisons par paires afin de tirer les avantages d'une représentation de l'information à un niveau fin de granularité qui est propre à l'AR
- nous définissons des matrices relationnelles de contributions liées à différentes règles majoritaires. Les termes des matrices relationnelles de contributions obtenues sont interprétées comme étant des éléments du dioïde présenté précédemment
- nous replaçons ce formalisme dans le cadre de l'algèbre des relations binaires qui sont donc à présent multivaluées. Dans ce contexte, cette dernière s'étend finalement à l'algèbre matricielle sur le dioïde $\{\mathbb{Z}_M, \oplus = \max, \otimes = \min\}$. Nous utilisons alors des opérateurs tel que le permanent d'une matrice, afin de caractériser les propriétés relationnelles du choix collectif obtenu selon plusieurs règles de vote

Dans le premier paragraphe, nous présentons les différentes règles majoritaires que nous allons étudier. Puis, nous donnerons des propriétés algébriques sur des matrices de comparaisons par paires de contributions permettant de caractériser la nature des relations collectives obtenues. Dans le troisième paragraphe, nous donnons aux différentes règles majoritaires présentées une version contrainte proposée initialement par M.J.A. Condorcet.

6.2.2.1 Règles majoritaires

Nous nous plaçons dans un cadre relationnel : les préférences des votants $V^k, k = 1 \dots, M$, sont codées par des matrices relationnelles C^k et nous notons par \mathbf{C} la matrice relationnelle collective agrégeant par addition les matrices relationnelles individuelles.

Nous étudions 3 types de règles, qui correspondent à 3 critères de vote distincts :

1. la règle de la majorité simple que nous avons définie en déf-6.2.5 et qui est largement utilisée dans le cadre de la théorie du choix social
2. la règle majoritaire à une majorité m où $1 \leq m \leq M$
3. la règle majoritaire avec un écart de p voix entre majorité et minorité où $1 \leq p \leq M$

A chaque type de règle majoritaire que nous étudions, nous lui associons une matrice de contributions notée \mathbf{C} qui est calculée à partir de la matrice \mathbf{C} . Nous cherchons à déterminer

des propriétés sur les matrices C afin de caractériser la nature de la relation collective obtenue à partir des différentes règles majoritaires.

Nous noterons par X^1 la matrice relationnelle représentative de la relation collective obtenue par la règle de la majorité simple. Nous rappelons la définition des règles majoritaires 2 et 3 ci-dessous. Nous noterons par X_m^2 et X_p^3 les relations collectives de ces deux règles qui dépendent respectivement des paramètres m et p . Nous utiliserons la notation suivante pour faire référence à l'une des trois règles considérées : X^r , $r = 1, 2, 3$. Lorsque $r = 2$ ou $r = 3$, nous précisons alors les paramètres respectifs m et p .

Définition 6.2.8 Règle majoritaire (par paires) à une majorité m donnée et représentation relationnelle

Un candidat O^i sera collectivement préféré ou indifférent à un candidat $O^{i'}$ si le nombre total de voix supportant " O^i préféré ou indifférent à $O^{i'}$ " est supérieur ou égal à une majorité m où $1 \leq m \leq M$. Dans le formalisme de l'AR, nous voyons que la matrice relationnelle représentative de la relation collective est en fait l'enveloppe de Condorcet que nous avons étudiée au chapitre 4. Par souci de cohérence vis à vis des notations utilisées dans ce paragraphe, nous notons par X_m^2 (et non E_2^M) la matrice relationnelle représentative de la relation obtenue en appliquant la règle majoritaire à une majorité m donnée. Nous avons alors :

$$[X_m^2]_{ii'} = \begin{cases} 1 & \text{si } C_{ii'} \geq m \\ 0 & \text{sinon} \end{cases} \quad (6.2.6)$$

Le cas particulier $m = \frac{M}{2}$, correspond à la règle majoritaire de Condorcet que nous avons appelée en introduction page 16.

Définition 6.2.9 Règle majoritaire (par paires) avec un écart de p voix entre majorité et minorité et représentation relationnelle

Un candidat O^i sera collectivement préféré ou indifférent à un candidat $O^{i'}$ si la différence entre le nombre total de voix majoritaires supportant " O^i préféré ou indifférent à $O^{i'}$ " et le nombre de voix minoritaires supportant " O^i préféré ou indifférent à $O^{i'}$ " est au moins égale à p où $1 \leq p \leq M$. Dans le formalisme de l'AR, nous constatons que le terme général de la matrice relationnelle représentative de la relation collective notée X_p^3 obtenue en appliquant cette règle est donnée par :

$$[X_p^3]_{ii'} = \begin{cases} 1 & \text{si } C_{ii'} - C_{i'i} \geq p \\ 0 & \text{sinon} \end{cases} \quad (6.2.7)$$

6.2.2.2 Caractérisation de la nature d'une relation collective à partir des propriétés algébriques d'une matrice de comparaisons par paires de contributions

Le contexte dans lequel nous nous positionnons est davantage général que celui habituellement proposé en théorie du choix social. Les préférences que peuvent exprimer les M votants sont des relations d'ordre large non nécessairement totales. Il s'agit donc de relations réflexives, faiblement antisymétriques, transitives qui peuvent posséder à la fois des candidats *ex-aequos* et des données manquantes. De plus, nous ne faisons aucune hypothèse sur la parité du nombre de votants.

Chaque préférence individuelle \mathcal{R}^k est représentée par une matrice relationnelle C^k et nous considérons la matrice relationnelle collective $C = \sum_k C^k$.

A chaque règle majoritaire, nous pouvons lui associer une matrice de comparaisons par paires dite de "contributions" qui sera notée C^r , $r = 1, 2, 3$. Les termes positifs de cette matrice permettent de déterminer la relation collective. Nous aurons :

$$[X^r]_{ii'} = \begin{cases} 1 & \text{si } [C^r]_{ii'} \geq 0 \\ 0 & \text{sinon} \end{cases} \quad (6.2.8)$$

Nous avons en fait la correspondance suivante.

Propriété 6.2.7 *Correspondance entre règles majoritaires et matrices de contributions*

Nous pouvons associer à chaque règle majoritaire 1, 2, 3, des matrices de contributions C^r , $r = 1, 2, 3$, qui sont définies à partir de la matrice relationnelle collective C .

Règle majoritaire	Terme général de C^r
1	$[C^1]_{ii'} = \begin{cases} C_{ii} = M & \forall (i, i') : i = i' \\ C_{ii'} - \check{C}_{ii'} & \forall (i, i') : i \neq i' \end{cases}$
2	$[C_m^2]_{ii'} = \begin{cases} C_{ii} = M & \forall (i, i') : i = i' \\ C_{ii'} - m & \forall (i, i') : i \neq i' \end{cases}$
3	$[C_p^3]_{ii'} = \begin{cases} C_{ii} = M & \forall (i, i') : i = i' \\ \max(C_{ii'} - \check{C}_{ii'} - p, -M) & \forall (i, i') : i \neq i' \end{cases}$

(6.2.9)

Les termes de la matrice C sont alors interprétés comme étant des éléments du dioïde $\{\mathbb{Z}_M, \max, \min\}$.

Il importe peu en pratique de savoir si la relation collective est réflexive ou non. Nous décidons cependant de considérer que les préférences individuelles sont toutes réflexives¹⁹. Nous considérons alors que $[C^r]_{ii} = C_{ii} = M$, $\forall i = 1, \dots, N$. Nous verrons par la suite que cette convention nous permet d'émettre des propriétés algébriques simples en nous appuyant sur le fait que M représentent l'élément neutre de la loi interne $\otimes = \min$.

Nous rappelons à présent des outils algébriques de l'algèbre matricielle dont les propriétés permettront de caractériser les préférences collectives. Ces outils correspondent en fait aux opérations peirciennes que nous avons rappelées en déf-4.1.2 page 116 de la section 4.1.1, mais pour lesquelles nous apportons quelques modifications. Ces opérations sont à la base issues de l'algèbre des relations binaires dans le cas booléen. Nous les appliquons dans un contexte différent puisque les éléments sur lesquels nous raisonnons sont des entiers relatifs compris entre $-M$ et M . Nous précisons que nous ne cherchons pas pour l'instant à interpréter les opérations présentées en termes de vote. Ainsi, C est une matrice carrée quelconque dont les termes sont des éléments de $\{\mathbb{Z}_M, \max, \min\}$.

Définition 6.2.10 *Opérations peirciennes modifiées de multiplication relative, d'addition relative et de converse dans le dioïde $\{\mathbb{Z}_M, \oplus = \max, \otimes = \min\}$*

Soit C et C' deux matrices carrée de taille N dont les termes appartiennent à l'ensemble $\{\mathbb{Z}_M, \max, \min\}$. La multiplication relative modifiée entre les matrices C et C' est notée $C \odot^* C'$ et son terme général est donné par la relation suivante :

$$[C \odot^* C']_{ii'} = \bigoplus_{i'' : i'' \neq i \vee i'' \neq i'} C_{ii''} \otimes C'_{i''i'} \quad (6.2.10)$$

$$= \max_{i'' : i'' \neq i \vee i'' \neq i'} (\min(C_{ii''}, C'_{i''i'})) \quad (6.2.11)$$

¹⁹. dans le cas d'ordres stricts nous considérons, sans conséquence sur le résultat, l'extension réflexive de ces derniers; dans le cas de données manquantes nous prendrons aussi par convention $C_{ii}^k = 1$, si le juge V^k ne renseigne pas le classement relatif au candidat O^i

L'addition relative modifiée entre ces deux matrices est notée $C \uplus^* C'$ et son terme général est donné par la relation suivante :

$$[C \uplus^* C']_{ii'} = \bigotimes_{i'' : i'' \neq i \vee i'' \neq i'} C_{ii''} \oplus C'_{i''i'} \quad (6.2.12)$$

$$= \min_{i'' : i'' \neq i \vee i'' \neq i'} (\max(C_{ii''}, C'_{i''i'})) \quad (6.2.13)$$

La converse d'une matrice C , noté \check{C} , est sa transposée :

$$\check{C} = {}^t C \quad (6.2.14)$$

ou encore

$$\check{C}_{i'i} = C_{i'i} \quad \forall i, i' = 1, \dots, N \quad (6.2.15)$$

Les versions modifiées de la multiplication relative et de l'addition relative consistent à ne pas tenir compte des éléments de la diagonale lors des calculs de type "multiplication matricielle".

A titre illustratif, nous donnons ci-dessous la matrice issue de la multiplication relative modifiée de la matrice C avec C' dans le cas où $N = 3$.

$$\begin{aligned} C \odot^* C' &= \begin{matrix} O^1 \\ O^2 \\ O^3 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 \\ C_{11} & C_{12} & C_{13} \\ C_{21} & C_{22} & C_{23} \\ C_{31} & C_{32} & C_{33} \end{pmatrix} \odot^* \begin{matrix} O^1 \\ O^2 \\ O^3 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 \\ C'_{11} & C'_{12} & C'_{13} \\ C'_{21} & C'_{22} & C'_{23} \\ C'_{31} & C'_{32} & C'_{33} \end{pmatrix} \\ &= \begin{matrix} O^1 \\ O^2 \\ O^3 \end{matrix} \begin{pmatrix} O^1 & O^2 & O^3 \\ \max \begin{pmatrix} \min(C_{12}, C'_{21}) \\ \min(C_{13}, C'_{31}) \end{pmatrix} & \min(C_{13}, C'_{32}) & \min(C_{12}, C'_{23}) \\ \min(C_{23}, C'_{31}) & \max \begin{pmatrix} \min(C_{21}, C'_{12}) \\ \min(C_{23}, C'_{32}) \end{pmatrix} & \min(C_{21}, C'_{13}) \\ \min(C_{32}, C'_{21}) & \min(C_{31}, C'_{12}) & \max \begin{pmatrix} \min(C_{31}, C'_{13}) \\ \min(C_{32}, C'_{23}) \end{pmatrix} \end{pmatrix} \end{matrix}$$

En plus des opérations peirciennes, nous rappelons la définition de la trace et du permanent d'une matrice carrée dans la structure algébrique qui nous concerne. Nous donnons une définition de la trace qui est fonction d'une loi interne. Nous donnons aussi une deuxième définition du permanent appelé permanent modifié qui nous sera utile par la suite.

Définition 6.2.11 Trace d'une matrice carrée dans le dioïde $\{\mathbb{Z}_M, \oplus = \max, \otimes = \min\}$

La trace d'une matrice carrée C de dimension N associée soit à la loi interne \oplus , soit à la loi interne \otimes , sera notée respectivement $\text{Trace}_{\oplus}(C)$ et $\text{Trace}_{\otimes}(C)$ et définie par :

$$\text{Trace}_{\oplus}(C) = \bigoplus_{i=1}^N C_{ii} = \max_{i=1}^N C_{ii} \quad (6.2.16)$$

$$\text{Trace}_{\otimes}(C) = \bigotimes_{i=1}^N C_{ii} = \min_{i=1}^N C_{ii} \quad (6.2.17)$$

Définition 6.2.12 Permanent Perm et permanent modifié Perm^* d'une matrice carrée dans le dioïde $\{\mathbb{Z}_M, \oplus = \max, \otimes = \min\}$

Le permanent²⁰ d'une matrice carrée C de dimension N , noté $\text{Perm}(C)$, est donné par :

$$\text{Perm}(C) = \bigoplus_{\pi \in \mathbb{S}_N} \bigotimes_{i=1}^N C_{i, \pi(i)} \quad (6.2.18)$$

$$= \max_{\pi \in \mathbb{S}_N} \left(\min_{i=1}^N C_{i, \pi(i)} \right) \quad (6.2.19)$$

où \mathbb{S}_N est le groupe symétrique sur les N premiers entiers naturels que nous avons rappelé en [rq-5.1.5](#) au chapitre 5 page 161.

Nous utiliserons une deuxième version du permanent que nous appellerons permanent modifié. Nous le noterons Perm^* . Il est défini par :

$$\text{Perm}^*(C) = \bigoplus_{\pi \in \mathbb{S}_N: \pi \neq \pi^\epsilon} \bigotimes_{i=1}^N C_{i, \pi(i)} \quad (6.2.20)$$

$$= \max_{\pi \in \mathbb{S}_N: \pi \neq \pi^\epsilon} \left(\min_{i=1}^N C_{i, \pi(i)} \right) \quad (6.2.21)$$

$\text{Perm}^*(C)$ se distingue de $\text{Perm}(C)$ en ce qu'il ne tient pas compte de la permutation identité π^ϵ qui est donnée par la diagonale de C .

Nous revenons à la problématique que nous avons étudié dans le cadre du choix social. Nous montrons à présent que les opérations peirciennes modifiées, la trace associée à une loi interne et le permanent modifié permettent de caractériser, à partir de la matrice des contributions C^r , les propriétés relationnelles de la relation collective obtenue en appliquant la règle majoritaire r où $r = 1, 2, 3$.

Propriété 6.2.8 Propriété algébrique de C^r et propriété d'asymétrie de X^r

Soit C^r la matrice des contributions relatives à la règle majoritaire $r = 1, 2, 3$ et soit X^r la matrice relationnelle de la relation collective obtenue selon cette règle. Alors :

$$\begin{aligned} \text{Trace}_{\oplus}(C^r \odot^* C^r) < 0 &\Leftrightarrow X^r \text{ est asymétrique} \\ &\Leftrightarrow \nexists (i, i'), i \neq i' : X_{ii'}^r = X_{i'i}^r = 1 \end{aligned} \quad (6.2.22)$$

Compte tenu de la propriété rappelée par l'équation (3.3.27) page 108, nous avons la formulation suivante équivalente :

$$\begin{aligned} \max_{i, i': i' \neq i} (\min([C^r]_{ii'}, [C^r]_{i'i})) < 0 &\Leftrightarrow X^r \text{ est asymétrique} \\ &\Leftrightarrow \nexists (i, i'), i \neq i' : X_{ii'}^r = X_{i'i}^r = 1 \end{aligned} \quad (6.2.23)$$

Démonstration

$$\begin{aligned} \text{Trace}_{\oplus}(C^r \odot^* C^r) < 0 &\Leftrightarrow \max_{i=1}^N ([C^r \odot^* C^r]_{ii}) < 0 \\ &\Leftrightarrow \max_{i=1}^N (\max_{i'': i'' \neq i} (\min([C^r]_{ii''}, [C^r]_{i''i}))) < 0 \\ &\Leftrightarrow \max_{i, i'': i'' \neq i} (\min([C^r]_{ii''}, [C^r]_{i''i})) < 0 \\ &\Leftrightarrow \forall (i, i''), i'' \neq i : \min([C^r]_{ii''}, [C^r]_{i''i}) < 0 \\ &\Leftrightarrow \forall (i, i''), i'' \neq i : \text{au moins un des deux termes } [C^r]_{ii''} \text{ ou } [C^r]_{i''i} \text{ est inférieur à } 0 \\ &\Leftrightarrow \forall (i, i''), i'' \neq i : \neg \{ [X^r]_{ii''} = [X^r]_{i''i} = 1 \} \\ &\Leftrightarrow \nexists (i, i'), i \neq i' : X_{ii'}^r = X_{i'i}^r = 1 \end{aligned}$$

□

20. un permanent d'une matrice carrée est tel un déterminant de cette dernière pour lequel, nous ne tenons pas compte du signe de la signature des permutations sur lesquelles sont fondées ces opérateurs

Propriété 6.2.9 Propriété algébrique de C^r et propriété de totalité de X^r

Soit C^r la matrice des contributions relatives à la règle majoritaire $r = 1, 2, 3$ et soit X^r la matrice relationnelle de la relation collective obtenue selon cette règle. Alors :

$$\begin{aligned} \text{Trace}_{\otimes}(C^r \uplus^* C^r) \geq 0 &\Leftrightarrow X^r \text{ est totale} \\ &\Leftrightarrow \forall(i, i'), i \neq i' : X_{ii'}^r = 1 \vee X_{i'i}^r = 1 \end{aligned} \quad (6.2.24)$$

Nous avons aussi :

$$\begin{aligned} \min_{i, i' : i' \neq i} (\max([C^r]_{ii'}, [C^r]_{i'i})) \geq 0 &\Leftrightarrow X^r \text{ est totale} \\ &\Leftrightarrow \forall(i, i'), i \neq i' : X_{ii'}^r = 1 \vee X_{i'i}^r = 1 \end{aligned} \quad (6.2.25)$$

Démonstration

Nous démontrons la contraposée de la propriété :

$$\begin{aligned} \text{Trace}_{\otimes}(C^r \uplus^* C^r) < 0 &\Leftrightarrow \min_{i=1}^N ([C^r \uplus^* C^r]_{ii}) < 0 \\ &\Leftrightarrow \min_{i=1}^N (\min_{i'' : i'' \neq i} (\max([C^r]_{ii''}, [C^r]_{i''i}))) < 0 \\ &\Leftrightarrow \min_{i, i'' : i'' \neq i} (\max([C^r]_{ii''}, [C^r]_{i''i})) < 0 \\ &\Leftrightarrow \exists(i, i''), i'' \neq i : \text{les deux termes } [C^r]_{ii''} \text{ et } [C^r]_{i''i} \text{ sont inférieurs à } 0 \\ &\Leftrightarrow \exists(i, i''), i'' \neq i : \max([C^r]_{ii''}, [C^r]_{i''i}) < 0 \\ &\Leftrightarrow \exists(i, i''), i'' \neq i : [X^r]_{ii''} = [X^r]_{i''i} = 0 \end{aligned}$$

Or, s'il existe au moins une paire (i, i'') , $i'' \neq i$, tel que $[X^r]_{ii''} = [X^r]_{i''i} = 0$, alors la relation n'est pas totale.

□

Théorème 6.2.2 Propriété algébrique de C^r et propriété d'acyclicité de X^r

Soit C^r la matrice des contributions relatives à la règle majoritaire $r = 1, 2, 3$ et soit X^r la matrice relationnelle de la relation collective obtenue selon cette règle. Alors :

$$\text{Perm}^*(C^r) < 0 \Leftrightarrow X^r \text{ est acyclique} \quad (6.2.26)$$

Démonstration

Considérons la contraposée de la propriété et montrons que $\text{Perm}^*(C^r) \geq 0$ si et seulement s'il existe au moins un k -cycles avec $2 \leq k \leq N$, dans la relations collective X^r .

$$\begin{aligned} \text{Perm}^*(C^r) \geq 0 &\Leftrightarrow \max_{\pi \in \mathbb{S}_N : \pi \neq \text{id}} (\min_{i=1}^N C_{i, \pi(i)}) \geq 0 \\ &\Leftrightarrow \max_{\pi \in \mathbb{S}_N : \pi \neq \text{id}} (\min(C_{1, \pi(1)}, C_{2, \pi(2)}, \dots, C_{N, \pi(N)})) \geq 0 \end{aligned}$$

Nous utilisons la décomposition en produits de cycles²¹ d'une permutation $\pi = \gamma^1 \gamma^2 \dots \gamma^q$. Soit γ^l un p^l -cycle. γ^l sécrit alors $(O^{i_1^l}, O^{i_2^l}, \dots, O^{i_{p^l}^l})$. Ce dernier vérifie la propriété suivante :

$$\gamma^l(O^{i_1^l}) = O^{i_2^l}; \dots; (\gamma^l(O^{i_1^l}))^{p^l-1} = \gamma^l(O^{i_{p^l-1}^l}) = O^{i_{p^l}^l} \text{ et } (\gamma^l(O^{i_1^l}))^{p^l} = \gamma^l(O^{i_{p^l}^l}) = O^{i_1^l}.$$

En termes des matrices de comparaisons par paires de contributions, le p^l -cycle γ^l concerne les éléments suivants :

$$C_{i_1^l, i_2^l}^r, C_{i_2^l, i_3^l}^r, \dots, C_{i_{p^l-1}^l, i_{p^l}^l}^r, C_{i_{p^l}^l, i_1^l}^r.$$

Si l'ensemble des éléments de la matrice C^r relatifs à un p^l -cycle γ^l sont positifs ou nuls, cela signifie, en terme de la matrice X^r que nous avons :

$$X_{i_1^l, i_2^l}^r = X_{i_2^l, i_3^l}^r = \dots = X_{i_{p^l-1}^l, i_{p^l}^l}^r = X_{i_{p^l}^l, i_1^l}^r = 1.$$

Ceci correspond donc à des préférences cycliques conduisant à l'impossibilité d'ordonner le sous-ensemble des objets O^{i^l} , $i^l \in \{i_1^l, i_2^l, \dots, i_{p^l}^l\} \subseteq \{1, \dots, N\}$.

Dans le cadre de notre modèle, nous avons la relation suivante :

$$\begin{aligned} X_{i_1^l, i_2^l}^r = X_{i_2^l, i_3^l}^r = \dots = X_{i_{p^l-1}^l, i_{p^l}^l}^r = X_{i_{p^l}^l, i_1^l}^r = 1 \\ \Leftrightarrow \min \left(C_{i_1^l, i_2^l}^r, C_{i_2^l, i_3^l}^r, \dots, C_{i_{p^l-1}^l, i_{p^l}^l}^r, C_{i_{p^l}^l, i_1^l}^r \right) \geq 0 \end{aligned} \quad (6.2.27)$$

21. cf prop-5.1.13 page 163

Par ailleurs, la décomposition d'une permutation π en un produit de cycles $\pi = \gamma^1 \gamma^2 \dots \gamma^q$ nous permet d'écrire :

$$\begin{aligned} \min(C_{1,\pi(1)}, C_{2,\pi(2)}, \dots, C_{N,\pi(N)}) &\geq 0 \\ &= \min_{l=1}^q \left(\min(C_{i_1^l, \gamma^l(i_1^l)}, \dots, C_{(\gamma^l(i_1^l))^{p^l-1}, i_1^l}) \right) \geq 0 \end{aligned}$$

En particulier, nous parcourons unitairement l'ensemble des permutations π dont la décomposition en cycle est du type suivant : un unique k -cycle, et $N - k$ points fixes. Les points fixes correspondent à la diagonale de C^r dont les termes sont égaux à M qui est l'élément neutre de $\otimes = \min$ du dioïde considéré. Sur ce sous-ensemble de permutations donnant l'ensemble des k -cycles de manière unitaire, il suffit qu'il n'y en ait qu'un seul qui soit vérifié pour que le permanent soit positif car (6.2.27). De ces différentes considérations, nous pouvons dire que $\text{Perm}^*(C^r) \geq 0$ si et seulement si, il existe au moins un k -cycles avec $2 \leq k \leq N$, dans la relations collective X^r . Par conséquent, $\text{Perm}^*(C^r) < 0$ si et seulement si il n'y a aucun cycle dans la relation collective X^r .

□

Nous faisons remarquer que la propriété précédente teste aussi les transpositions c'est à dire les 2-cycles. Autrement dit, s'il existe au moins une paire (i, i') avec $i \neq i'$ pour lequel $[X^r]_{ii'} = [X^r]_{i'i} = 1$ alors le permanent modifié sera positif ou nul. Ainsi, une relation partiellement symétrique sera considérée comme cyclique.

De plus, il est nécessaire de préciser qu'**acyclicité n'est pas équivalent à transitivité**. En fait, nous pouvons affirmer que la propriété d'acyclicité implique la propriété de transitivité uniquement dans le cas où nous savons que la relation est asymétrique et totale²². Il s'agit en particulier du cas de la règle majoritaire simple avec un nombre de votants impair puisqu'elle donne une relation collective asymétrique et totale²³. Nous avons la démonstration suivante qui étaye ces propos où X est une matrice relationnelle :

$$\begin{aligned} \begin{cases} X_{ii'} + X_{i'i} \leq 1 \\ X_{ii'} + X_{i'i} \geq 1 \\ X_{ii'} + X_{i'i''} + X_{i''i} \leq 2 \end{cases} &\Rightarrow \begin{cases} X_{ii'} + X_{i'i} = 1 \\ X_{ii'} + X_{i'i''} + X_{i''i} \leq 2 \end{cases} \\ &\Rightarrow \begin{cases} X_{ii'} + X_{i'i} = 1 \\ X_{ii'} + X_{i'i''} + (1 - X_{ii''}) \leq 2 \end{cases} \\ &\Rightarrow \begin{cases} X_{ii'} + X_{i'i} = 1 \\ X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 \end{cases} \end{aligned}$$

Nous donnons alors ci-dessous des conditions nécessaires et suffisantes de la propriété de transitivité d'une relation collective.

Théorème 6.2.3 *Propriété algébrique de C^r et propriété de transitivité de X^r*

Soit C^r la matrice des contributions relatives à la règle majoritaire $r = 1, 2, 3$ et soit X^r la matrice relationnelle de la relation collective obtenue selon cette règle. Alors :

$$\boxed{(C^r \odot^* C^r) \otimes \bar{U}_N \preceq C^r \oplus (-U_N) \Leftrightarrow X^r \text{ est transitive}} \quad (6.2.28)$$

où \bar{U}_N est la matrice carrée de dimension N remplie de 0, $-U_N$ celle remplie de -1 et où deux matrices C et C' sont telles que $C \preceq C'$ si et seulement si $[C]_{ii'} \leq [C']_{ii'} \quad \forall i, i' = 1, \dots, N$.

Démonstration

Nous avons tout d'abord :

22. ce qui est souvent considéré dans le cadre de la théorie du choix social qui de plus raisonne sur des triplets

23. contexte dans lequel se place les conditions de Black et d'Inada

X^r est transitive

$$\begin{aligned} &\Leftrightarrow \forall i, i' : \{\exists i'', i'' \neq i \vee i'' \neq i' : [C^r]_{ii''} \geq 0 \wedge [C^r]_{i''i'} \geq 0\} \Rightarrow \{[C^r]_{ii'} \geq 0\} \\ &\Leftrightarrow \forall i, i' : \{\max_{i'', i'' \neq i \vee i'' \neq i'} (\min([C^r]_{ii''), [C^r]_{i''i'}) \geq 0\} \Rightarrow \{[C^r]_{ii'} \geq 0\} \\ &\Leftrightarrow \forall i, i' : \{[C^r \odot^* C^r]_{ii'} \geq 0\} \Rightarrow \{[C^r]_{ii'} \geq 0\} \\ &\Leftrightarrow \forall i, i' : \{\min([C^r \odot^* C^r]_{ii'}, 0) \geq 0\} \Rightarrow \{\max([C^r]_{ii'}, -1) \geq 0\} \end{aligned}$$

Nous expliquons à présent le lien entre les deux dernières équivalences que nous venons d'introduire afin de démontrer la propriété. Nous notons par A la proposition $\{[C^r \odot^* C^r]_{ii'} \geq 0\}$ et par B la proposition $\{[C^r]_{ii'} \geq 0\}$. Soit $v(A) \in \{0, 1\}$ la valeur de vérité de la proposition A . Selon la table de vérité, nous avons la correspondance suivante vis à vis de la relation d'implication :

$v(A)$	$v(B)$	$v(A \Rightarrow B)$
1	1	1
1	0	0
0	1	1
0	0	1

Dans un premier temps, si $v(A) = 1$, alors $\min([C^r \odot^* C^r]_{ii'}, 0) = 0$. Si de plus $v(B) = 1$, nous avons $\max([C^r]_{ii'}, -1) \geq 0$. Dans ce cas, nous avons la conjonction $v(A) = 1$ et $v(B) = 1$, qui est équivalente à $\min([C^r \odot^* C^r]_{ii'}, 0) \leq \max([C^r]_{ii'}, -1)$. Si par contre, $v(B) = 0$, alors $\max([C^r]_{ii'}, -1) = -1$, et dans ce cas, la conjonction $v(A) = 1$ et $v(B) = 0$, est équivalente à $\min([C^r \odot^* C^r]_{ii'}, 0) > \max([C^r]_{ii'}, -1)$.

Dans un deuxième temps, considérons $v(A) = 0$. Nous avons alors $\min([C^r \odot^* C^r]_{ii'}, 0) \leq -1$. Si $v(B) = 1$, nous avons alors l'équivalence suivante : $\min([C^r \odot^* C^r]_{ii'}, 0) \leq \max([C^r]_{ii'}, -1)$. Si $v(B) = 0$, alors nous avons de la même manière : $\min([C^r \odot^* C^r]_{ii'}, 0) \leq \max([C^r]_{ii'}, -1)$.

Nous avons énuméré, pour la relation d'implication, toutes les combinaisons possibles données précédemment par la table de vérité. Si nous nous intéressons à la relation d'ordre engendrée par ces différentes combinaisons, sur les deux quantités $\min([C^r \odot^* C^r]_{ii'}, 0)$ et $\max([C^r]_{ii'}, -1)$ qui fondent les propositions A et B , nous nous apercevons que lorsque la relation d'implication est vraie alors nous avons toujours $\min([C^r \odot^* C^r]_{ii'}, 0) \leq \max([C^r]_{ii'}, -1)$ et lorsque qu'elle est fautive, nous avons au contraire $\min([C^r \odot^* C^r]_{ii'}, 0) > \max([C^r]_{ii'}, -1)$. Nous avons donc :

X^r est transitive

$$\begin{aligned} &\Leftrightarrow \forall i, i' : \min([C^r \odot^* C^r]_{ii'}, 0) \leq \max([C^r]_{ii'}, -1) \\ &\Leftrightarrow [(C^r \odot^* C^r) \otimes \bar{U}_N] \subseteq [C^r \oplus (-U_N)] \end{aligned}$$

□

Le contexte dans lequel nous énonçons nos résultats est plus général que celui dans lequel sont énoncées les conditions de Black, d'Inada et de Sen. En effet, coder les relations selon le formalisme relationnel permet de tenir compte de tout type de relations d' "ordre" qu'elles soient totales ou non, avec des *ex-aequos* ou non voire même transitives ou non²⁴. Par ailleurs, nos résultats concernent 3 règles d'agrégation différentes. Notre approche consiste à raisonner directement sur des paires de candidats plutôt que sur des triplets puisqu'ainsi, nous sommes davantage proche de la définition des propriétés relationnelles. De plus, les propriétés algébriques que nous proposons sont des conditions nécessaires et suffisantes qui caractérisent les propriétés relationnelles de totalité, d'asymétrie, d'acyclicité et de transitivité d'une relation collective X^r obtenue à partir des trois règles majoritaires étudiées.

Enfin, un des intérêts de ce paragraphe est de caractériser les propriétés relationnelles, dans le contexte de l'agrégation des préférences par une règle majoritaire, avec une approche qui est à la croisée de l'algèbre relationnelle étendue à un dioïde non booléen et de l'algèbre matricielle classique. A ce propos, nous faisons remarquer que les opérateurs Trace de la multiplication entre deux matrices peut aussi être assimilé à un "produit scalaire" entre matrices²⁵.

D'autres approches originales du problème d'agrégation de relations de préférences ont été entreprises. Nous citerons par exemple, les travaux de G. Chichilnisky qui étudia le problème

24. les préférences peuvent parfois être non rationnelles

25. le produit scalaire ou métrique de Frobenius $\langle {}^t C^r, C^r \rangle_F$ cf page 108. Toutefois, dans le mesure où, notre algèbre de base n'est pas un anneau, encore moins un corps, nous ne pouvons rigoureusement parler d'espace vectoriel et de produit scalaire

dans un cadre topologique [Chi80]. Par ailleurs, nous rappelons, qu'il existe un indice permettant dans le cas particulier d'un ordre strict et total, de vérifier de manière rapide la présence d'effets Condorcet. Il s'agit de l'indice de hiérarchie de Landau que nous rappelons ci-dessous [MM78].

Propriété 6.2.10 Indice de hiérarchie de Landau

Si nous nous plaçons dans le contexte de la règle de la majorité simple $r = 1$ et dans le cas où le nombre de votants est impair, nous savons que la relation collective représentée par la matrice relationnelle binaire X^1 , donnée par (6.2.8), est asymétrique et totale. Alors dans ce cas précis, un moyen rapide de vérifier si la relation X^1 possède des cycles est de calculer l'indice suivant :

$$\text{Landau}(X^1) = \frac{12}{N^3 - N} \sum_{i=1}^N \left([X^1]_i. - \frac{N-1}{2} \right)^2 \tag{6.2.29}$$

où $[X^1]_i. = \sum_{i'=1}^N [X^1]_{ii'}$. Nous savons en effet, que si $\text{Landau}(X^1) < 1$ alors il y a des cycles et donc des effets Condorcet.

Pour conclure ce paragraphe, nous allons illustrer les conditions nécessaires et suffisantes présentées ici sur un exemple pratique. Celui-ci concerne le classement de 4 candidats O^1, O^2, O^3, O^4 donnés par 5 juges V^1, V^2, V^3, V^4, V^5 selon le tableau tab-6.3.

	V^1	V^2	V^3	V^4	V^5
O^1	1	1	1	4	3
O^2	2	4	4	1	2
O^3	3	2	3	3	1
O^4	4	3	2	2	4

TABLE 6.3 – Exemple d'ensemble de préférences individuelles qui ne vérifient pas les conditions de Sen et qui conduisent à une relation collective transitive

Nous avons représenté l'ensemble de ces classements sur la figure fig-6.2 qui montre que nous ne sommes pas dans le cadre des conditions d'unimodalités de Black ou d'Inada.

Par ailleurs, nous pouvons voir de même sur le triplet O^2, O^3, O^4 , que l'ensemble de ces préférences individuelles, ne vérifient pas les conditions de Sen. En effet, si nous restreignons le tableau tab-6.3 à ce triplet particulier, et si nous réécrivons les rangs en termes de valeur v , nous obtenons le tableau tab-6.4.

	V^1	V^2	V^3	V^4	V^5
O^2	1	3	3	1	2
O^3	2	1	2	3	1
O^4	3	2	1	2	3

TABLE 6.4 – Triplet de l'exemple étudié ne vérifiant pas les conditions de Sen

Nous constatons sur ce tableau que chaque candidat, O^2, O^3, O^4 , prend au moins une fois une valeur $v = 1, 2, 3$. Ce triplet ne vérifie donc pas les conditions de Sen.

Nous montrons sur cet exemple, les limites des conditions de Black, d'Inada, et de Sen.

Pour cela, considérons dans un premier temps, l'agrégation des matrices relationnelles de chacune des relations d'ordre total engendrées par ces classements. La matrice relationnelle

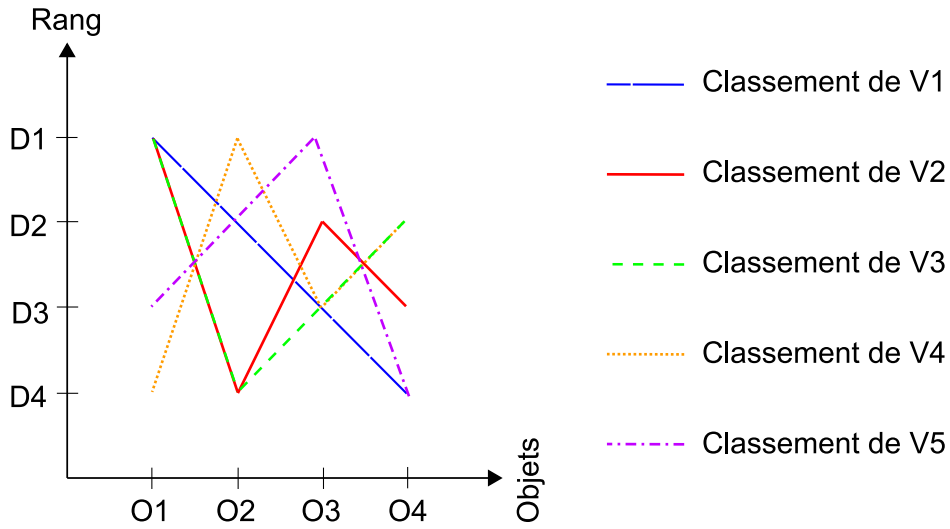


FIGURE 6.2 – Représentation de l'exemple traité mettant en évidence la non vérification des conditions de Black et d'Inada

collective $C = C^1 + C^2 + C^3 + C^4$ obtenue est la suivante :

$$C = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 5 & 3 & 3 & 4 \\ 2 & 5 & 2 & 3 \\ 2 & 3 & 5 & 3 \\ 1 & 2 & 2 & 5 \end{pmatrix} \end{matrix} \quad (6.2.30)$$

Prenons la règle de la majorité simple avec $r = 1$, et calculons la matrice de contributions C^1 correspondante. Nous avons alors selon prop-6.2.7 :

$$C^1 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 5 & 1 & 1 & 3 \\ -1 & 5 & -1 & 1 \\ -1 & 1 & 5 & 1 \\ -3 & -1 & -1 & 5 \end{pmatrix} \end{matrix} \quad (6.2.31)$$

La relation d'ordre collective obtenue par la règle de la majorité simple est alors la suivante :

$$X^1 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (6.2.32)$$

X^1 obtenue est un ordre large, total et sans *ex-aequo* qui est le suivant : $O^1 \leq O^3 \leq O^2 \leq O^4$. La relation collective obtenue selon la règle de la majorité simple est donc transitive. De ce fait,

nous sommes en présence d'un cas, qui ne rentre pas dans le cadre des conditions de Black, d'Inada ou de Sen. **Ainsi et a contrario de ce qui vient d'être mis en évidence, nous montrons ci-dessous que nos conditions sont vérifiées à la fois pour l'asymétrie, la totalité, l'acyclicité et la transitivité. Nous illustrons alors le caractère suffisant de nos conditions.**

Nous avons dans un premier temps :

$$\begin{aligned}
 & \max_{i,i':i' \neq i} (\min([C^1]_{ii'}, [C^1]_{i'i})) \\
 &= \max (\min([C^1]_{12}, [C^1]_{21}), \min([C^1]_{13}, [C^1]_{31}), \dots, \min([C^1]_{34}, [C^1]_{43})) \\
 &= \max (\min(1, -1), \min(1, -1), \min(3, -3) \dots, \min(1, -1)) \\
 &= \max (-1, -1, -3, -1, -1, -1) \\
 &= -1
 \end{aligned} \tag{6.2.33}$$

Nous avons donc $\max_{i,i':i' \neq i} (\min([C^1]_{ii'}, [C^1]_{i'i}))$ qui est strictement inférieur à 0. Nous venons de vérifier que la relation collective obtenue est asymétrique selon la propriété prop-6.2.8.

Nous pouvons vérifier de la même manière que $\min_{i,i':i' \neq i} (\max([C^1]_{ii'}, [C^1]_{i'i})) = 1$, ce qui est supérieur ou égal à 0. Par conséquent, nous vérifions de même que la relation collective est totale selon la propriété prop-6.2.9.

Nous sommes alors dans le cas où acyclicité et transitivité sont identiques comme nous l'avions fait remarquer précédemment. Cependant, afin d'illustrer les deux théorèmes donnant les conditions d'acyclicité et de transitivité, nous montrons qu'elles sont toutes les deux vérifiées pour notre exemple.

Nous calculons donc dans un deuxième temps le permanent modifié de la matrice de contributions C^1 . Le permanent étant un déterminant pour lequel, la signature d'une permutation n'est pas pris en compte, nous pouvons pour le calculer, appliquer la formule de développement selon une colonne de la matrice que nous utilisons couramment pour le calcul d'un déterminant. Nous avons alors :

$$\text{Perm}^*(C^1) = \begin{vmatrix} 5 & 1 & 1 & 3 \\ -1 & 5 & -1 & 1 \\ -1 & 1 & 5 & 1 \\ -3 & -1 & -1 & 5 \end{vmatrix} = \max \left(\begin{array}{l} \min \left(5, \begin{vmatrix} 5 & -1 & 1 \\ 1 & 5 & 1 \\ -1 & -1 & 5 \end{vmatrix} \right), \\ \min \left(-1, \begin{vmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ -1 & -1 & 5 \end{vmatrix} \right), \\ \min \left(-1, \begin{vmatrix} 1 & 1 & 3 \\ 5 & -1 & 1 \\ -1 & -1 & 5 \end{vmatrix} \right), \\ \min \left(-3, \begin{vmatrix} 1 & 1 & 3 \\ 5 & -1 & 1 \\ 1 & 5 & 1 \end{vmatrix} \right) \end{array} \right) \tag{6.2.34}$$

Posons :

$$\begin{aligned}
 A &= \min \left(5, \begin{vmatrix} 5 & -1 & 1 \\ 1 & 5 & 1 \\ -1 & -1 & 5 \end{vmatrix} \right) & B &= \min \left(-1, \begin{vmatrix} 1 & 1 & 3 \\ 1 & 5 & 1 \\ -1 & -1 & 5 \end{vmatrix} \right) \\
 C &= \min \left(-1, \begin{vmatrix} 1 & 1 & 3 \\ 5 & -1 & 1 \\ -1 & -1 & 5 \end{vmatrix} \right) & D &= \min \left(-3, \begin{vmatrix} 1 & 1 & 3 \\ 5 & -1 & 1 \\ 1 & 5 & 1 \end{vmatrix} \right)
 \end{aligned}$$

Nous avons :

$$\begin{aligned}
 A &= \min \left(\begin{array}{c} \left(\min \left(5, \max \left(\begin{array}{c} \min \left(5, \begin{array}{c} 5 \\ -1 \\ 5 \end{array} \end{array} \right) \\ \min \left(1, \begin{array}{c} -1 \\ -1 \\ 5 \end{array} \right) \\ \min \left(-1, \begin{array}{c} -1 \\ 5 \\ 1 \end{array} \right) \end{array} \right) \end{array} \right) \\
 B &= \min \left(\begin{array}{c} \left(\min \left(-1, \max \left(\begin{array}{c} \min \left(1, \begin{array}{c} 5 \\ -1 \\ 5 \end{array} \right) \\ \min \left(1, \begin{array}{c} 1 \\ -1 \\ 5 \end{array} \right) \\ \min \left(-1, \begin{array}{c} 1 \\ 5 \\ 1 \end{array} \right) \end{array} \right) \end{array} \right) \\
 C &= \min \left(\begin{array}{c} \left(\min \left(-1, \max \left(\begin{array}{c} \min \left(1, \begin{array}{c} -1 \\ -1 \\ 5 \end{array} \right) \\ \min \left(5, \begin{array}{c} 1 \\ -1 \\ 5 \end{array} \right) \\ \min \left(-1, \begin{array}{c} 1 \\ -1 \\ 1 \end{array} \right) \end{array} \right) \end{array} \right) \\
 D &= \min \left(\begin{array}{c} \left(\min \left(-3, \max \left(\begin{array}{c} \min \left(1, \begin{array}{c} -1 \\ 5 \\ 1 \end{array} \right) \\ \min \left(5, \begin{array}{c} 1 \\ 5 \\ 1 \end{array} \right) \\ \min \left(1, \begin{array}{c} 1 \\ -1 \\ 1 \end{array} \right) \end{array} \right) \end{array} \right)
 \end{aligned}$$

Puis :

$$\begin{aligned}
 A &= \min \left(\mathbf{5}, \max \left(\begin{array}{c} \min (\mathbf{5}, \max (\min (\mathbf{5}, \mathbf{5}), \min (-1, 1))) \\ \min (1, \max (\min (-1, 5), \min (-1, 1))) \\ \min (-1, \max (\min (-1, 1), \min (5, 1))) \end{array} \right) \right) \\
 B &= \min \left(-1, \max \left(\begin{array}{c} \min (1, \max (\min (5, 5), \min (-1, 1))) \\ \min (1, \max (\min (1, 5), \min (-1, 3))) \\ \min (-1, \max (\min (1, 1), \min (5, 3))) \end{array} \right) \right) \\
 C &= \min \left(-1, \max \left(\begin{array}{c} \min (1, \max (\min (-1, 5), \min (-1, 1))) \\ \min (5, \max (\min (1, 5), \min (-1, 3))) \\ \min (-1, \max (\min (1, 1), \min (-1, 3))) \end{array} \right) \right) \\
 D &= \min \left(-3, \max \left(\begin{array}{c} \min (1, \max (\min (-1, 1), \min (5, 1))) \\ \min (5, \max (\min (1, 1), \min (5, 3))) \\ \min (1, \max (\min (1, 1), \min (-1, 3))) \end{array} \right) \right)
 \end{aligned}$$

Nous précisons que le terme A ci-dessus contient le produit relatif à la permutation triviale π^e . Nous ne tenons donc pas compte de ce terme dont les éléments ont été mis en gras et nous obtenons :

$$\begin{aligned}
 A &= \min (5, \max (\min (5, -1), \min (1, -1), \min (-1, 1))) = \min (5, -1) \\
 B &= \min (-1, \max (\min (1, 5), \min (1, 1), \min (-1, 3))) = \min (-1, 1) \\
 C &= \min (-1, \max (\min (1, -1), \min (5, 1), \min (-1, 1))) = \min (-1, 1) \\
 D &= \min (-3, \max (\min (1, 1), \min (5, 3), \min (1, 1))) = \min (-3, 3)
 \end{aligned}$$

Nous aboutissons finalement à :

$$\text{Perm}^*(C^1) = \max(-1, -1, -1, -3) = -1 \tag{6.2.35}$$

La valeur du permanent modifié étant strictement négative, nous en déduisons selon le théorème théo-6.2.2 que la relation collective est bien acyclique.

Nous avons ensuite, le produit relatif modifié de C^1 avec lui-même qui vaut :

$$C^1 \odot^* C^1 = \begin{array}{c} O^1 \quad O^2 \quad O^3 \quad O^4 \\ \begin{array}{c} O^1 \\ O^2 \\ O^3 \\ O^4 \end{array} \left(\begin{array}{cccc} -1 & 1 & -1 & 1 \\ -1 & -1 & -1 & -1 \\ -1 & -1 & -1 & 1 \\ -1 & -1 & -1 & -1 \end{array} \right)
 \end{array}$$

Le lecteur pourra vérifier que nous avons la relation suivante qui, par le théorème théo-6.2.3, permet d'affirmer que la relation collective issue des préférences individuelles données par le tableau tab-6.3 et obtenue par la règle de la majorité simple, aboutit à un ordre transitif, c'est à dire, sans effet Condorcet :

$$\underbrace{\begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} -1 & 0 & -1 & 0 \end{pmatrix} \\ O^2 & \begin{pmatrix} -1 & -1 & -1 & -1 \end{pmatrix} \\ O^3 & \begin{pmatrix} -1 & -1 & -1 & 0 \end{pmatrix} \\ O^4 & \begin{pmatrix} -1 & -1 & -1 & -1 \end{pmatrix} \end{matrix}}_{(C^1 \odot^* C^1) \otimes \bar{U}_N} \preceq \underbrace{\begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 5 & 1 & 1 & 3 \end{pmatrix} \\ O^2 & \begin{pmatrix} -1 & 5 & -1 & 1 \end{pmatrix} \\ O^3 & \begin{pmatrix} -1 & 1 & 5 & 1 \end{pmatrix} \\ O^4 & \begin{pmatrix} -1 & -1 & -1 & 5 \end{pmatrix} \end{matrix}}_{C^1 \oplus (-U_N)} \quad (6.2.36)$$

Nous faisons remarquer par ailleurs, qu'il n'est pas plus couteux, en terme de complexité, de calculer un produit matriciel que d'énumérer tous les triplets possibles d'un ensemble de candidats. En effet, l'ensemble de ces calculs nécessitent $\mathcal{O}(N^3)$ opérations élémentaires. De ce fait, **les temps de calcul afin de vérifier nos conditions nécessaires et suffisantes²⁶ ne sont pas significativement plus longs que les temps de calculs cherchant à vérifier les conditions de Black, d'Inada ou de Sen²⁷.**

Enfin, nous illustrons le caractère nécessaire de nos conditions ainsi que la prise en compte par celles-ci de divers autres critères de vote. Pour cela, considérons la matrice de contributions liée à la règle majoritaire à une majorité $m = 2$ sur un ensemble de $M = 5$ votants. Notons cette matrice par C_2^2 , nous avons alors :

$$C_2^2 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 5 & 1 & 1 & 2 \end{pmatrix} \\ O^2 & \begin{pmatrix} 0 & 5 & 0 & 1 \end{pmatrix} \\ O^3 & \begin{pmatrix} 0 & 1 & 5 & 1 \end{pmatrix} \\ O^4 & \begin{pmatrix} -1 & 0 & 0 & 5 \end{pmatrix} \end{matrix} \quad (6.2.37)$$

Nous pouvons rapidement vérifier les conditions d'asymétrie et de totalité. Nous avons pour l'un $\max_{i,i':i' \neq i} (\min([C_2^2]_{ii'}, [C_2^2]_{i'i}))$ qui vaut 0, ce qui implique que la relation collective notée X_2^2 obtenue par cette règle majoritaire n'est pas asymétrique. Nous avons pour l'autre, $\min_{i,i':i' \neq i} (\max([C_2^2]_{ii'}, [C_2^2]_{i'i}))$ qui vaut 1, ce qui implique en revanche que la relation collective obtenue est totale.

Nous cherchons à savoir si X_2^2 est transitive ou non. Le lecteur pourra alors vérifier que nous avons la relation suivante qui, selon le théorème théo-6.2.3, permet d'affirmer que la relation collective n'est pas transitive :

$$\underbrace{\begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \\ O^2 & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \\ O^3 & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \\ O^4 & \begin{pmatrix} 0 & 0 & 0 & 0 \end{pmatrix} \end{matrix}}_{(C_2^2 \odot^* C_2^2) \otimes \bar{U}_N} \not\preceq \underbrace{\begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ O^1 & \begin{pmatrix} 5 & 1 & 1 & 2 \end{pmatrix} \\ O^2 & \begin{pmatrix} 0 & 5 & 0 & 1 \end{pmatrix} \\ O^3 & \begin{pmatrix} 0 & 1 & 5 & 1 \end{pmatrix} \\ O^4 & \begin{pmatrix} -1 & 0 & 0 & 5 \end{pmatrix} \end{matrix}}_{C^2 \oplus (-U_N)} \quad (6.2.38)$$

Nous constatons en effet que l'inégalité n'est pas vérifiée pour le terme $(4, 1)$. Or, nous nous apercevons que c'est ce terme précis qui suggère un problème de transitivité puisque la relation collective X_2^2 étant la suivante, nous devrions avoir $\{[X_2^2]_{4,3} = 1 \wedge [X_2^2]_{3,1} = 1 \Rightarrow [X_2^2]_{4,1} = 1\}$,

26. excepté le calcul d'un permanent dont la valeur exacte ne peut être, dans notre cas, calculer en un temps polynomial [Vai79]

27. quoique nous avons donné une démonstration algébrique (relationnelle) de l'extension des conditions de Black, d'Inada et de Sen, l'exemple que nous avons donné permet d'imaginer des conventions "graphiques" de type "double-peakedness" ou "double-cavedness"...

ce qui n'est pas :

$$X_2^2 = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{pmatrix} \end{matrix} \quad (6.2.39)$$

6.2.2.3 Version contrainte des règles majoritaires et résolution par maximisation des critères de vote associés par programmation linéaire en nombres bivalents

Nous avons fait référence en introduction aux travaux de P. Michaud concernant les résultats déjà obtenus par M.J.A. Condorcet au XVIIIème siècle, en théorie du choix social. Tout particulièrement dans [Mic85b], P. Michaud rappelle une seconde règle de Condorcet qui est la règle de la majorité sous contraintes. Le principe est de restreindre l'ensemble des solutions possibles aux seules relations vérifiant les propriétés de la solution que nous souhaitons avoir. Nous avons précédemment rappelé l'effet Condorcet et le théorème d'impossibilité d'Arrow ainsi que certains résultats importants autour de ce théorème. De plus, nous avons donné de nouvelles conditions mais à caractère algébrique. Nous nous plaçons dans ce paragraphe dans un contexte différent et considérons davantage le principe de la règle sous contraintes qui, dans le cas de la règle majoritaire de Condorcet vérifie les 3 premiers axiomes d'Arrow et une version modifiée de l'axiome d'indépendance [Mic85b].

Nous étendons le programme linéaire en nombres bivalents permettant de déterminer les préférences collectives à partir de la règle de la majorité de Condorcet sous contraintes, aux règles majoritaires 1, 2 et 3.

Nous obtenons ainsi une version contrainte de ces règles majoritaires.

Chaque règle majoritaire 1, 2, 3 peut être associée à un critère de vote à l'instar de la règle de la majorité de Condorcet dont le critère Condorcet (C, X) ²⁸ est représentatif. Ces critères de votes sont liés aux termes généraux des matrices de contributions respectives que nous avons donnés en prop-6.2.7.

Propriété 6.2.11 Critères de vote associés aux règles majoritaires 1, 2, 3

1. Critère de vote de la règle de la majorité simple :

$$CV^1(C, X) = \sum_{i,i'} (C_{ii'} - \check{C}_{ii'}) X_{ii'} \quad (6.2.40)$$

2. Critère de vote de la règle majoritaire à une majorité m donnée :

$$CV_m^2(C, X) = \sum_{i,i'} (C_{ii'} - m) X_{ii'} \quad (6.2.41)$$

3. Critère de vote de la règle majoritaire avec un écart de p voix entre majorité et minorité :

$$CV_n^3(C, X) = \sum_{i,i'} (C_{ii'} - \check{C}_{ii'} - p) X_{ii'} \quad (6.2.42)$$

Pour chacun de ces critères, l'ordre consensuel pourra être obtenu à partir des modèles de programmation linéaire en nombres bivalents suivants :

28. cf page 7

Propriété 6.2.12 Détermination de la relation d'ordre collective par une règle majoritaire sous contraintes

La relation d'ordre collective correspondante à la règle majoritaire sous contraintes $r = 1, 2, 3$; est obtenue par le programme linéaire en nombres bivalents suivant :

$$\left. \begin{array}{l} X_{ii'} \in \{0, 1\} \\ X_{ii'} = 0 \\ X_{ii'} + X_{i'i} \leq 1 \\ X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 \end{array} \right\} \begin{array}{l} \forall i, i' \\ \forall i \\ \forall i, i' : i \neq i' \\ \forall i, i', i'' \end{array} \quad \text{si } X \text{ est un ordre strict}$$

ou

$$\max_X CV^r(\mathbf{C}, X) \quad \text{s.t.} \quad \left. \begin{array}{l} X_{ii'} \in \{0, 1\} \\ X_{ii} = 1 \\ X_{ii'} + X_{i'i} + X_{ii''} - X_{i'i''} \leq 2 \\ X_{ii'} + X_{i'i} - X_{ii''} + X_{i'i''} \leq 2 \\ X_{ii'} + X_{i'i''} - X_{ii''} \leq 1 \end{array} \right\} \begin{array}{l} \forall i, i' \\ \forall i \\ \forall i, i', i'' \\ \forall i, i', i'' \\ \forall i, i', i'' \end{array} \quad \text{si } X \text{ est un ordre large} \quad (6.2.43)$$

plus les contraintes suivantes

$$X_{ii'} + X_{i'i} \geq 1 \quad \forall i, i' : i \neq i' \quad \text{si } X \text{ est totale}$$

Nous terminons ce chapitre en faisant la remarque suivante : **les critères de vote que nous avons définis et qui sont les termes des matrices de contributions peuvent, dans un certain sens, être interprétés comme étant des critères d'association maximale appliqués au problème du consensus de relations d'ordre.**

En effet, nous avons étudié à la section 6.1, des critères d'associations entre deux matrices relationnelles de deux relations d'ordre. Si nous considérons le numérateur de ces critères nous avons vu qu'ils étaient tous les mêmes. En particulier, considérons la formulation de S. Ghashghaie de ces critères que nous avons rappelé en introduction page 19. Les numérateurs de ces critères s'écrivent tous selon l'équation suivante :

$$\Delta(C^k, C^l) = \sum_{i, i'} (C_{ii'}^k - \check{C}_{ii'}^k) C_{ii'}^l \quad (6.2.44)$$

Si nous appliquons le modèle d'association maximale au critère ci-dessus nous définissons le critère suivant :

$$\max_X \sum_k \Delta(C^k, X) \quad (6.2.45)$$

Or, étant donnée que (6.2.44) est linéaire en C^k , l'équation (6.2.46) se simplifie en :

$$\max_X \sum_k \Delta(C^k, X) \Leftrightarrow \max_X \Delta(\mathbf{C}, X) \Leftrightarrow \max_X \sum_{i, i'} (C_{ii'} - \check{C}_{ii'}) X_{ii'} \quad (6.2.46)$$

Cette équation est exactement identique à celle du critère de vote associé à la règle de la majorité simple et à une constante p près²⁹ au critère de vote de la règle majoritaire avec un écart entre majorité et minorité. De ce fait, si nous faisons abstraction du dénominateur de ces critères qui donne chacun une interprétation quant aux poids des paires d'*ex-aequos*, nous pouvons considérer que **le critère de vote CV^1 est le critère d'association maximal associé aux critères d'association entre relations d'ordre de Goodman et Kruskal, de Somers, de Kendall, de Deuchler et de Kim, ainsi que celui de la recherche de l'ordre médian selon l'approche de J.G. Kemeny [KS62], [MM78].**

²⁹. nous pouvons aussi définir l'association maximale de critères d'association suivant : $\sum_k \left(\sum_{i, i'} (C_{i'i}^k - C_{i'i}^k - p/M) \right)$ où p/M est la proportion relative à l'écart entre majorité et minorité

Dans le contexte des critères d'association pour partitions, ce modèle a permis de définir des critères de partitionnement que nous avons étudiés dans le chapitre 3. Appliqué à des critères d'association de relations d'ordre, nous constatons qu'il permet alors de définir des critères de vote. Nous faisons ainsi l'extension de l'association maximale dans le cadre des classements.

CHAPITRE 7

Modélisation des tresses et des noeuds par approche relationnelle

Sommaire

7.1 Problèmes autour du groupe des tresses et quelques méthodes de résolution	217
7.1.1 Problème du mot du groupe des tresses ou d'isotopie de mots de tresse	217
7.1.1.1 Présentation du problème	217
7.1.1.2 Mouvements isotopiques de retournement à droite et de retournement à gauche	218
7.1.1.3 L'ordre de Dehornoy sur les tresses et les mouvements isotopiques de réduction de poignées	220
7.1.2 Problèmes du mot de tresse minimal et du mot de tresse "shortlex"	223
7.2 Modélisation des tresses par approche relationnelle	224
7.2.1 Les matrices de paires de brins nommés et les structures de graphe induites	225
7.2.2 Codage du groupe des tresses par matrices de paires de brins nommés	227
7.2.3 Mouvements isotopiques de brins libres	230
7.2.3.1 Identification des paires de brins nommés et juxtaposés consécutives à ajouter lors du déplacement du brin libre	237
7.2.3.2 Détermination de la nature topologique des paires de brins nommés juxtaposés utilisées lors du déplacement d'un brin libre	242
7.3 Les problèmes du mot du groupe des tresses, du mot de tresse minimal et du mot de tresse "shortlex" dans le cadre de la modélisation par paires de brins nommés	253
7.3.1 Modélisation des problèmes algorithmiques autour des tresses sous la forme de problèmes d'optimisation combinatoire	253
7.3.2 Exemples d'application et considérations diverses	263
7.4 Modélisation des noeuds par approche relationnelle	271
7.4.1 Noeuds tressés dans le cadre de la modélisation relationnelle : la clôture d'une tresse et les mouvements de Markov	271
7.4.2 Détermination d'un invariant complet pour les noeuds, exemple d'application	275

Dans ce chapitre, nous nous intéressons au groupe des tresses dont nous avons introduit les concepts de base en page 20. Nous proposons une modélisation originale de ces objets mathématiques. Nous présentons de plus, des procédés qui décrivent des outils qui permettraient de contribuer à la résolution de problèmes autour du groupe des tresses. Ces procédés sont, dans l'esprit, différents des approches proposées à ce jour.

Dans un premier temps, nous rappelons donc les problèmes du mot du groupe des tresses, du mot de tresse minimal et du mot de tresse "shortlex". Parmi ces trois problèmes, nous nous arrêterons tout d'abord plus particulièrement sur le problème du mot du groupe des tresses. Ce problème consiste à déterminer si deux mots de tresse représentent la même tresse géométrique ou non. Parmi les méthodes de résolution les plus efficaces nous rappellerons celles de P. Dehornoy [Deh00a]. Ce dernier définit des procédés, qui consistent en des déplacements isotopiques de brins et qui permettent de résoudre le problème rapidement. Nous introduisons ces procédés étant donné que notre approche consiste, à définir des mouvements isotopiques de brins qui sont plus généraux que ceux adoptés par P. Dehornoy. Ces mouvements se formalisent naturellement dans notre modélisation et seront appelés **déplacements ou mouvements de brins libres**. Nous présentons ensuite le **problème du mot de tresse minimal et de la forme canonique "shortlex"**. Ces problèmes consistent à déterminer une forme réduite et unique d'un mot de tresse représentant une tresse géométrique. Plusieurs chercheurs ont proposé, dans différents contextes, des méthodes afin de déterminer des solutions à ces problèmes. Nous les citons brièvement et nous proposons une approche de modélisation nouvelle de ces problèmes.

Ainsi, dans un second temps, nous présentons de manière approfondie notre modélisation des tresses. Celle-ci est basée sur une extension du codage par matrice de paires d'objets nommés que nous avons défini pour le groupe des permutations qui est un sous-cas particulier du groupe des tresses. Notre méthode permet de raisonner sur les brins d'une tresse tout en conservant l'information classique sur les rangs que manipulent traditionnellement les approches algébriques¹. Ce faisant, nous pouvons déterminer des règles de réécriture d'une tresse dans ce codage qui sont en fait des mouvements isotopiques génériques. Nous présentons une approche différente de la modélisation des problèmes cités précédemment en les appréhendant sous l'angle de l'optimisation combinatoire. Ainsi, de la même manière que pour le problème de classification automatique, ou de la recherche de consensus au sein de relation d'ordres, les problèmes algorithmiques autour du groupe des tresses peuvent être appréhendus comme étant la recherche de chemins dans un graphe polytopique particulier cherchant à optimiser un critère.

Enfin, nous abordons succinctement le problème de la recherche d'un invariant complet pour les noeuds à partir de notre modélisation des tresses. Nous nous basons sur un article de T. Gittings [Git04] dont la détermination de l'invariant défini par ce dernier, se ramène à la recherche d'un mot de tresse "shortlex" particulier. Nous appliquons alors notre méthode à ce problème en montrant comment nous pouvons prendre en compte les mouvements isotopiques de Markov qui lient "tresses d'un noeud" et le noeud lui même. Notre méthode pourrait conduire alors à une méthode de détermination de cet invariant complet.

Remarque 7.0.1 Notations

Dans le cadre de la théorie des groupes, les tresses à N brins forment un groupe noté \mathbb{B}_N dont la présentation est la suivante :

$$\left\langle \begin{array}{l} \sigma^1, \dots, \sigma^{N-1} \\ (\sigma^1)^{-1}, \dots, (\sigma^{N-1})^{-1} \end{array} \middle| \begin{array}{l} \sigma^j \circ (\sigma^j)^{-1} = \sigma^\epsilon \quad \forall j = 1, \dots, N-1 \\ \sigma^j \circ \sigma^{j'} = \sigma^{j'} \circ \sigma^j \quad \forall j, j' = 1, \dots, N-1 : |j - j'| > 1 \\ \sigma^j \circ \sigma^{j'} \circ \sigma^j = \sigma^{j'} \circ \sigma^i \circ \sigma^{j'} \quad \forall j, j' = 1, \dots, N-1 : |j - j'| = 1 \end{array} \right\rangle \quad (7.0.1)$$

où σ^ϵ est le générateur identité qui ne code aucun croisement.

Nous remarquons les liens entre la présentation du groupe des permutations données par prop-5.1.15 page 164 et la présentation du groupe des tresses.

1. ie les générateurs σ^j agissant sur les rangs

Nous noterons par β une tresse géométrique. Par ailleurs, nous noterons par \mathbb{B}_N^+ le monoïde des tresses à N brins généré uniquement par les générateurs d'exposants positifs : $\{\sigma^1, \dots, \sigma^{N-1}\}$. Un élément w de \mathbb{B}_N^+ sera appelé mot de tresse positif.

Nous supposons par la suite que les mots de tresse sont librement réduits, c'est à dire que les motifs $\sigma^j \circ (\sigma^j)^{-1}$ et $(\sigma^j)^{-1} \circ \sigma^j$ sont systématiquement supprimés².

Enfin, nous avons étudié, dans le cadre des permutations ou des ordres stricts et totaux, les transpositions génératrices³ qui sont les éléments générateurs des groupes correspondants. Dans le cas des tresses, nous nommerons ces générateurs "**aires de brins nommés et juxtaposés**" lorsqu'il s'agira de croisements agissant sur les brins. Ces générateurs de croisements de brins ont de plus une signature topologique "au-dessous" et "au-dessus". Nous distinguerons ces deux cas par respectivement une barre au-dessous et au-dessus de la aire de brins nommés.

7.1 Problèmes autour du groupe des tresses et quelques méthodes de résolution

7.1.1 Problème du mot du groupe des tresses ou d'isotopie de mots de tresse

7.1.1.1 Présentation du problème

Un des problèmes d'ordre algébrique qui a été très étudié en théorie des tresses est celui de reconnaître si deux mots de tresse sont topologiquement équivalents, c'est à dire s'ils représentent la même tresse géométrique ou encore s'ils appartiennent à la classe d'équivalence d'une même tresse géométrique. Prenons l'exemple simple des deux mots de tresse suivants $Ababc$ et bac . Nous pouvons passer du premier mot de tresse au second par des mouvements isotopiques issus des relations d'Artin que nous avons rappelées en introduction page 20 :

$$\begin{aligned} \overline{Ababc} &\equiv \overline{Aab\bar{a}c} && \text{application de la relation } R3 \\ &\equiv \overline{Aabac} \\ &\equiv \overline{bac} && \text{application de la relation } R1 \end{aligned}$$

Dans notre exemple, la seule utilisation des relations d'Artin a permis de montrer l'équivalence entre ces deux mots de tresse. Pour des tresses plus compliquées, les relations d'Artin $R2$ et $R3$ ne suffisent pas afin de réduire⁴ les motifs $\sigma^j \circ (\sigma^j)^{-1}$ ou $(\sigma^j)^{-1} \circ \sigma^j$ de manière efficace. De nombreux auteurs ont étudié alors ce problème et proposé des solutions diverses :

- E. Artin dans [Art47]
- F.A. Garside dans [Gar69]
- W. Thurston dans [ECH⁺92]
- P. Dehornoy dans [Deh92] puis [Deh97]
- J. Birman K.H. Ko et S.J. Lee dans [BKL98]

Parmi l'ensemble des méthodes proposées, l'une des plus efficaces d'un point de vue algorithmique, est celle développée par P. Dehornoy dans [Deh97], [Deh00a]. Celle-ci utilise des mouvements de brins simples mais possédant des propriétés particulières. Nous les rappelons dans le paragraphe suivant étant donné que notre approche consiste, dans un certain sens, en une extension de ces mouvements de brins.

2. symboliquement il s'agit du générateur identité σ^e
 3. agissant sur les rangs ou agissant sur les objets
 4. ie réduction libre

7.1.1.2 Mouvements isotopiques de retournement à droite et de retournement à gauche

Une première solution au problème du mot du groupe des tresses proposée par P. Dehornoy dans [Deh92] consiste à déterminer l'expression d'un mot de tresse sous une forme canonique particulière. Celle-ci est inspirée des travaux de F.A. Garside [Gar69] qui indique qu'un mot de tresse w peut être mis sous une forme fractionnaire :

$$w \equiv \mathbf{N}(w) \circ \mathbf{D}(w)^{-1} \quad (7.1.1)$$

où $\mathbf{N}(w)$ et $\mathbf{D}(w)$ sont des tresses positives éléments de \mathbb{B}_N^+ .

Cette forme particulière peut être obtenue à partir des mouvements dits de "retournement" que nous rappelons ci-dessous.

Définition 7.1.1 Mouvement isotopique de retournement à droite

Soit \mathbf{c}_R une fonction dite de "complément" (à droite) définie par :

$$\mathbf{c}_R(\sigma^j, \sigma^{j'}) = \begin{cases} \sigma^\epsilon & \text{si } j = j' \\ \sigma^j & \text{si } |j - j'| > 1 \\ \sigma^j \circ \sigma^{j'} & \text{si } |j - j'| = 1 \end{cases} \quad (7.1.2)$$

Nous avons alors la relation suivante :

$$\sigma^j \circ \mathbf{c}_R(\sigma^{j'}, \sigma^j) \equiv \sigma^{j'} \circ \mathbf{c}_R(\sigma^j, \sigma^{j'}) \Leftrightarrow \begin{cases} \sigma^j \circ \sigma^{j'} \equiv \sigma^{j'} \circ \sigma^j & \forall |j - j'| > 1 \\ \sigma^j \circ \sigma^{j'} \circ \sigma^j \equiv \sigma^{j'} \circ \sigma^j \circ \sigma^{j'} & \forall |j - j'| = 1 \end{cases} \quad (7.1.3)$$

Nous dirons qu'on applique un "retournement à droite" du mot de tresse w si nous appliquons la relation suivante :

$$(\sigma^j)^{-1} \circ \sigma^{j'} \rightarrow \mathbf{c}_R(\sigma^{j'}, \sigma^j) \circ \mathbf{c}_R(\sigma^j, \sigma^{j'})^{-1} \quad (7.1.4)$$

où nous remplaçons un motif $(\sigma^j)^{-1} \circ \sigma^{j'}$ par le motif $\mathbf{c}_R(\sigma^{j'}, \sigma^j) \circ \mathbf{c}_R(\sigma^j, \sigma^{j'})^{-1}$ qui ne change pas la nature de la tresse.

Nous donnons dans la figure fig-7.1 le diagramme correspondant aux deux cas non triviaux d'un mouvement isotopique de retournement à droite.

Définition 7.1.2 Mouvement isotopique de retournement à gauche

Soit \mathbf{c}_L une fonction dite de "complément" (à gauche) définie par :

$$\mathbf{c}_L(\sigma^j, \sigma^{j'}) = \begin{cases} \sigma^\epsilon & \text{si } j = j' \\ \sigma^j & \text{si } |j - j'| > 1 \\ \sigma^{j'} \circ \sigma^j & \text{si } |j - j'| = 1 \end{cases} \quad (7.1.5)$$

Nous avons alors la relation suivante :

$$\mathbf{c}_L(\sigma^{j'}, \sigma^j) \circ \sigma^j \equiv \mathbf{c}_L(\sigma^j, \sigma^{j'}) \circ \sigma^{j'} \Leftrightarrow \begin{cases} \sigma^{j'} \circ \sigma^j \equiv \sigma^j \circ \sigma^{j'} & \forall |j - j'| > 1 \\ \sigma^j \circ \sigma^{j'} \circ \sigma^j \equiv \sigma^{j'} \circ \sigma^j \circ \sigma^{j'} & \forall |j - j'| = 1 \end{cases} \quad (7.1.6)$$

Nous dirons qu'on applique un "retournement à gauche" du mot de tresse w si nous appliquons la relation suivante :

$$\sigma^j \circ (\sigma^{j'})^{-1} \rightarrow \mathbf{c}_L(\sigma^{j'}, \sigma^j)^{-1} \circ \mathbf{c}_L(\sigma^j, \sigma^{j'}) \quad (7.1.7)$$

où nous remplaçons un motif $\sigma^j \circ (\sigma^{j'})^{-1}$ par le motif $\mathbf{c}_L(\sigma^{j'}, \sigma^j)^{-1} \circ \mathbf{c}_L(\sigma^j, \sigma^{j'})$ qui ne change pas la nature de la tresse.

Retournement à droite

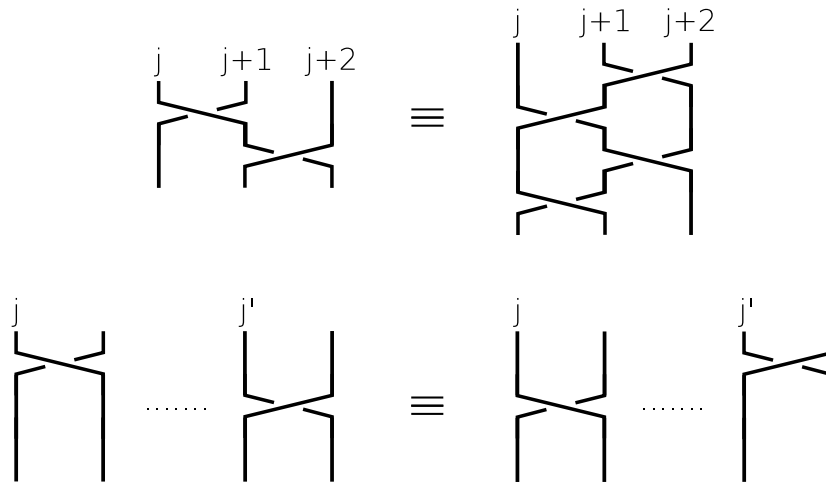


FIGURE 7.1 – Mouvements isotopiques de retournement à droite

Si elles se terminent, les itération successives d'un des deux types de mouvements isotopiques de retournement définis ci-dessus, convergent vers un mot de tresse de forme fractionnaire qui est unique et qui est équivalent au mot de tresse de départ. Cette forme fractionnaire est telle que :

- les générateurs d'exposant négatif se retrouvent en fin de mot si les mouvements sont des retournements à droite
- les générateurs d'exposant négatif se retrouvent en début de mot si les mouvements sont des retournements à gauche

Plus formellement, nous aboutissons aux formes fractionnaires suivantes d'un mot de tresse w :

- dans le cas de retournements à droite successifs, c'est à dire en appliquant itérativement les règles de réécritures données par (7.1.4) :

$$w \equiv \mathbf{N}_R(w) \circ \mathbf{D}_R(w)^{-1} \quad (7.1.8)$$

- dans le cas de retournements à gauche successifs, c'est à dire en appliquant itérativement les règles de réécritures données par (7.1.7) :

$$w \equiv \mathbf{D}_L(w)^{-1} \circ \mathbf{N}_L(w) \quad (7.1.9)$$

où $\mathbf{N}_R(w)$, $\mathbf{D}_R(w)$, $\mathbf{N}_L(w)$, $\mathbf{D}_L(w)$ sont des mots positifs.

P. Dehornoy démontre la proposition et le théorème suivants qui permettent de définir un algorithme de comparaison de deux mots de tresses basé sur ces mouvements [Deh92] :

Proposition 7.1.1 Double retournement et forme fractionnaire

Posons pour tout mot de tresse w :

$$\begin{cases} \mathbf{N}_{LR}(w) &= \mathbf{N}_L(\mathbf{N}_R(w) \circ \mathbf{D}_R(w)^{-1}) \\ \mathbf{D}_{LR}(w) &= \mathbf{D}_L(\mathbf{N}_R(w) \circ \mathbf{D}_R(w)^{-1}) \end{cases} \quad (7.1.10)$$

Nous avons alors les propriétés suivantes :

$$w \equiv \mathbf{D}_{LR}(w)^{-1} \circ \mathbf{N}_{LR}(w) \quad (7.1.11)$$

$$w \equiv w' \Leftrightarrow \{\mathbf{N}_{LR}(w) \equiv \mathbf{N}_{LR}(w')\} \wedge \{\mathbf{D}_{LR}(w) \equiv \mathbf{D}_{LR}(w')\} \quad (7.1.12)$$

Théorème 7.1.1 *Sur les applications de double retournement*

Il existe deux applications \mathbf{D}_{LR} et \mathbf{N}_{LR} de \mathbb{B}_N dans \mathbb{B}_N^+ tel que :

- pour tout mot de tresse w , $\mathbf{D}_{LR}(w)^{-1} \circ \mathbf{N}_{LR}(w)$ représente le même mot de tresse que w
- le mot w représente la tresse triviale β^ϵ si et seulement si les mots $\mathbf{D}_{LR}(w)$ et $\mathbf{N}_{LR}(w)$ sont vides ⁵
- lorsque le nombre de générateurs est borné, la détermination de $\mathbf{D}_{LR}(w)$ et de $\mathbf{N}_{LR}(w)$ pour un mot w de longueur $|w|$ se fait en un nombre d'étapes de l'ordre de $|w|^2$

A l'aide des résultats précédemment cités, on peut définir un algorithme permettant de comparer deux mots de tresse w et w' pour résoudre le problème d'isotopie entre ces deux mots. Il suffit de considérer la tresse $w \circ (w')^{-1}$ et de procéder sur ce mot l'application de double retournement. Si celui-ci aboutit à un mot vide alors on a $w \equiv w'$.

Les procédures que nous venons de décrire ont une complexité quadratique relativement à la longueur du mot de tresse sur lequel nous appliquons le double retournement. Nous rappelons ci-dessous une deuxième procédure, plus élégante encore, toujours due à P. Dehornoy qui d'un point de vue géométrique utilise un mouvement isotopique dit de "réduction de poignée" et qui d'un point de vue algébrique est fondé sur l'existence d'un ordre total sur les tresses à N brins.

7.1.1.3 L'ordre de Dehornoy sur les tresses et les mouvements isotopiques de réduction de poignées

Tout d'abord, nous rappelons l'ordre de Dehornoy sur le groupe des mots de tresse \mathbb{B}_N . Cette propriété est fondamentale dans la compréhension du comportement de l'algorithme de réduction de poignées que nous présentons ensuite et qui donne une résolution efficace du problème du mot du groupe des tresses.

Nous rappelons essentiellement les résultats donnés dans [Deh97] vers lequel nous renvoyons le lecteur pour les démonstrations des différents théorèmes et propositions énoncés.

Définition 7.1.3 *Mot de tresse réduit*

Soit w un mot de tresse de \mathbb{B}_N . Nous dirons que celui-ci est réduit si son indice principal j^- , qui est le plus petit indice de l'ensemble des générateurs composant w , apparaît soit positivement uniquement, soit négativement uniquement.

Définition 7.1.4 *Mot de tresse de type σ^j -positif*

Soit w un mot de tresse de \mathbb{B}_N . Nous dirons que celui-ci est σ^j -positif si le générateur σ^j apparaît avec des puissances positives uniquement. De manière symétrique, nous dirons qu'il est σ^j -négatif si le générateur σ^j apparaît avec des exposants négatifs uniquement.

Proposition 7.1.2 *Mot de tresse réduit et tresse triviale*

Un mot de tresse réduit qui est non vide ne peut représenter la tresse triviale.

Par ailleurs, tout mot de tresse est de manière exclusive soit σ^{j^-} -positif, soit σ^{j^-} -négatif, soit vide.

⁵ ie qui se réduisent elles même à la tresse triviale β^ϵ

Théorème 7.1.2 *Existence d'une décomposition réduite d'un mot de tresse*

Tout mot de tresse admet une décomposition réduite au sens de la définition déf-7.1.4

Corollaire 7.1.1 *Ordre strict et total sur les éléments de \mathbb{B}_N*

Soient β^k et β^l deux tresses à N brins et soient respectivement w^k et w^l , leurs représentations sous formes de mots de tresse éléments de \mathbb{B}_N . Nous dirons que β^k est strictement inférieur à β^l , ce que nous noterons par $\beta^k < \beta^l$, si et seulement si $(w^k)^{-1} \circ w^l$ admet une décomposition réduite qui soit σ^{j^-} -positive où j^- est l'indice principal de $(w^k)^{-1} \circ w^l$.

A partir des résultats précédents, nous voyons que si on dispose d'une méthode permettant de construire une décomposition réduite d'un mot de tresse, alors on dispose par la même d'une méthode permettant de comparer deux mots de tresse. En effet, si w et w' sont deux mots de tresse, alors la décomposition réduite du mot $w \circ (w')^{-1}$ est égale au mot vide si et seulement si $w \equiv w'$.

La méthode décrite ci-dessous permet de construire de manière efficace, en terme de temps de traitements, une décomposition réduite d'un mot de tresse.

Définition 7.1.5 *σ^j -poignée et poignée principale*

Une σ^j -poignée est un mot de tresse de la forme $(\sigma^j)^e \circ v \circ (\sigma^j)^{-e}$ où l'exposant e est pris dans $\{-1, 1\}$ et le mot v ne contient uniquement que des générateurs $\sigma^{j'}$ tel que $j' < j - 1$ ou $j' > j$.

Une poignée principale d'un mot de tresse w est un sous-mot de celui-ci qui soit une σ^{j^-} -poignée où j^- est l'indice principal de w .

Définition 7.1.6 *Mouvement isotopique de réduction de poignée*

Nous dirons qu'on applique une "réduction d'une σ^j -poignée" du mot de tresse w si nous appliquons la relation $\phi_{j,e}$ suivante :

$$\phi_{j,e} : \begin{cases} (\sigma^j)^{\pm 1} & \rightarrow \sigma^e \\ (\sigma^{j+1})^{\pm 1} & \rightarrow (\sigma^{j+1})^{-e} \circ (\sigma^j)^{\pm 1} \circ (\sigma^{j+1})^e \\ (\sigma^{j'})^{\pm 1} & \rightarrow (\sigma^{j'})^{\pm 1} \quad \text{où } j' \neq j \text{ et } j' \neq j+1 \end{cases} \quad (7.1.13)$$

Représentons une σ^j -poignée, $(\sigma^j)^e \circ (v) \circ (\sigma^j)^{-e}$, donnée par déf-7.1.5 de la manière suivante :

$$(\sigma^j)^e \circ [v] \circ (\sigma^j)^{-e} = (\sigma^j)^e \circ [v^0 \circ (\sigma^{j+1})^{d_1} \circ v^1 \circ \dots \circ (\sigma^{j+1})^{d_m} \circ v^m] \circ (\sigma^j)^{-e} \quad (7.1.14)$$

où v^0, \dots, v^m ne contiennent aucun générateur $\sigma^{j'}$ tel que $j - 1 \leq j' \leq j + 1$. Alors l'application $\phi_{j,e}$ consiste à transformer l'expression précédente en l'expression équivalente ci-dessous :

$$[v^0 \circ (\sigma^{j+1})^{-e} \circ (\sigma^j)^{d_1} \circ (\sigma^{j+1})^e \circ v^1 \circ \dots \circ (\sigma^{j+1})^{-e} \circ (\sigma^j)^{d_m} \circ (\sigma^{j+1})^e \circ v^m] \quad (7.1.15)$$

Nous illustrons dans la figure fig-7.2 les mouvements isotopiques de réduction d'une σ^j -poignée quelconque.

Les mouvements isotopiques de réduction de poignées cherchent à simplifier les occurrences des générateurs σ^j qui apparaissent à la fois avec un exposant positif et un exposant négatif⁶. Nous voyons que ce type de mouvement pourrait aider à déterminer la décomposition réduite d'un mot de tresse donnée par déf-7.1.4. Toutefois, ceci n'est pas immédiat. Il est nécessaire de préciser certains points et de démontrer le bien fondé de ces procédures. En particulier, nous faisons remarquer que l'application d'une réduction de poignée peut augmenter le nombre de croisements⁷ et faire apparaître de nouvelles poignées.

6. ie réduction libre

7. ie la longueur d'un mot de tresse

Réduction de poignée

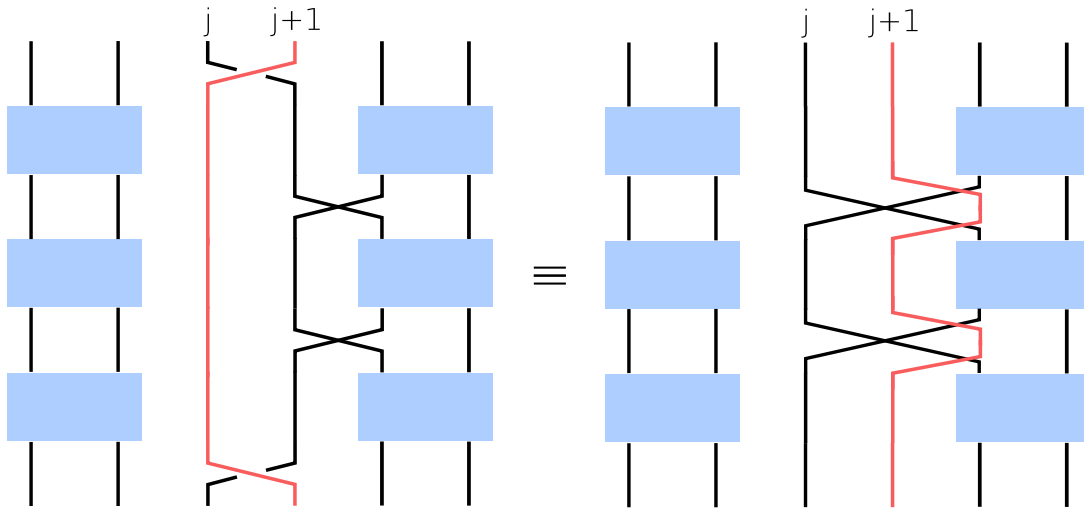


FIGURE 7.2 – Mouvements isotopiques de réduction d’une σ^j -poignée

Définition 7.1.7 Réduction “permise” d’une σ^j -poignée

Une σ^j -poignée, $(\sigma^j)^e \circ v \circ (\sigma^j)^{-e}$, est permise si au moins un des deux générateurs σ^{j+1} et $(\sigma^{j+1})^{-1}$ n’ont pas d’occurrence dans v .

Dans [Deh97], P. Dehornoy montre que les applications successives des réductions permises de σ^j -poignées sur un mot de tresse :

- convergent vers une décomposition réduite de ce mot
- se terminent en un nombre fini d’étapes étant donné que ces mouvements sont acycliques

Ces résultats fondent le théorème et les résultats suivants qui permettent finalement d’affirmer que l’application successive de réduction de poignées permises procure bien une méthode pour comparer deux mots de tresse.

Théorème 7.1.3 La réduction de poignée est noethérienne

Il n’existe pas de séquence infinie de réductions permises de poignée. Plus précisément, si w est un mot de tresse de longueur l et de largeur n , la longueur de toute suite d’application de réductions permises de poignées à partir de w est majorée par $2^{n^4 l}$.

De manière plus générale, nous avons les résultats suivants :

Définition 7.1.8 Mot de tresse complètement réduit

Soit w un mot de tresse de \mathbb{B}_N . Nous dirons que celui-ci est complètement réduit si deux générateurs σ^j et $(\sigma^j)^{-1}$ dans w sont séparés par au moins un générateur $(\sigma^{j-1})^{\pm 1}$.

Lemme 7.1.1 Mot de tresse complètement réduit et existence de poignées

Soit w un mot de tresse de \mathbb{B}_N . Les propositions suivantes sont équivalentes :

- w est complètement réduit
- w ne contient aucune poignée
- on ne peut plus réduire w par réduction de poignée

En pratique P. Dehornoy propose deux algorithmes pour l'application de la réduction de poignées qui diffèrent par la stratégie adoptée dans le choix de la poignée à réduire.

Le premier algorithme nommé "FullHred" consiste, à partir d'un mot de tresse w , à appliquer une réduction permise⁸ à la première poignée rencontrée qu'elle soit principale ou non jusqu'à obtenir une forme complètement réduite du mot.

Le deuxième algorithme nommé "GreedyHred" consiste, à partir d'un mot de tresse w , à appliquer une réduction permise à la première σ^j -poignée rencontrée jusqu'à obtenir une forme réduite du mot.

Dans les deux cas, les algorithmes convergent vers le mot de tresse vide si le mot de tresse initial représente la tresse triviale. Sinon, le mot de tresse obtenu est respectivement soit complètement réduit au sens de déf-7.1.8, soit réduit au sens de déf-7.1.4.

Dans [Deh97], P. Dehornoy émet finalement une remarque concernant l'algorithme "FullHRed". Celui-ci produit un mot de tresse complètement réduit dont la longueur est plus petite que celle du mot initial. Nous faisons ici le lien avec le problème du mot de tresse minimal que nous présenterons dans la sous-section suivante et pour lequel l'algorithme mentionné permet déjà d'obtenir une bonne approximation.

7.1.2 Problèmes du mot de tresse minimal et du mot de tresse "shortlex"

Le problème du mot de tresse minimal présenté ici consiste à déterminer, parmi la classe d'équivalence des mots d'une tresse géométrique, celui qui possède le minimum de croisements ou encore le minimum de générateurs. De manière empirique, on constate que la réduction de poignées présentée précédemment permet d'obtenir un mot de tresse de taille réduite mais cette taille n'est pas forcément optimale. Lorsque le nombre de brins n'est pas fixé, le problème est selon M.S. Paterson et A.A Razborov co- \mathcal{NP} -complet [PA91]. Lorsque le nombre de brins est en revanche fixé, certains travaux soutiennent que le problème peut alors être résolu par un algorithme de complexité polynomiale : il s'agit tout d'abord d'un preprint de K. Tatsuoka [Tat87] complété par les travaux de J.F.P. Hudson dans [Hud96].

Le problème considéré est simple mais ardu. Au vu d'un graphe de Cayley, il s'agit de déterminer une géodésique séparant la tresse triviale de la tresse considérée.

Reprenons l'exemple du mot de tresse suivant, $Ababc$. Nous avons vu page 217 que celui-ci était topologiquement équivalent au mot de tresse bac . Il n'existe ensuite aucune simplification qui permette de diminuer la longueur de ce mot de tresse. Par conséquent nous pouvons dire que ce dernier est un mot de tresse minimal. Nous remarquons qu'il existe plusieurs mots de tresse de longueur minimale. Dans notre exemple, si nous appliquons la relation d'Artin R_2 alors nous obtenons bca qui est ainsi un deuxième mot de tresse minimal.

Si nous donnons initialement un ordre total sur l'ensemble des générateurs tel que $a < A <$

8. ceci implique donc qu'il faille réduire les sous-poignées dans un premier temps

$b < B < c < C \dots$ ⁹, alors il existe un unique mot de tresse minimal que nous appellerons forme canonique “shortlex”. Cette forme unique est donnée par le mot de tresse minimal dont l’ordre¹⁰ qui est engendré par les générateurs le constituant, est le plus petit.

Dans notre exemple, le mot de tresse bac est le mot de tresse “shortlex”. En effet, si nous prenons l’ordre total sur les générateurs donné précédemment, nous voyons que $bac < bca$.

Il existe quelques travaux qui, dans des cas particuliers proposent des méthodes de résolution. Ces derniers concernent essentiellement le groupe des mots de tresse à 3 brins \mathbb{B}_3 . Il s’agit des contributions de R.D. Keever [Kee94], M.A. Berger [Ber94] et H. Simsek et al [SBC03]. Dans un cadre davantage général, nous citons le travail de thèse de P.D. Bangert [Ban02] qui propose un procédé de détermination de la forme “shortlex¹¹” d’un mot de tresse avec une approche qui partant de la forme canonique de Garside, cherche à énumérer exhaustivement les mots de tresse obtenus à l’aide des générateurs d’Artin. Nous voyons que cet algorithme est très couteux en temps de traitement et rappelle l’approche intuitive de l’énumération exhaustive de l’ensemble des partitions possibles d’un ensemble d’objets afin de déterminer la partition optimisant un critère donné.

L’ensemble des algorithmes proposés peuvent s’apparenter à des systèmes de réécriture¹². Ces derniers consistent à identifier des motifs particuliers et de les remplacer systématiquement par d’autres motifs équivalents changeant ainsi le mot de tresse mais non pas la nature de la tresse. Des règles de réécriture sont donc définies en amont de l’algorithme. Il s’agit ensuite de démontrer que le système de réécriture défini est complet, c’est à dire qu’il est convergent et terminal¹³. De plus amples détails sur les systèmes de réécriture, pourront être trouvés dans [DJ90] par exemple.

Dans la section suivante, nous proposons une modélisation nouvelle et originale des tresses. Celle-ci, nous l’espérons, pourrait permettre de contribuer dans un cadre général, c’est à dire pour \mathbb{B}_N avec $N > 2$, à la résolution du problème de la détermination d’un mot de tresse minimal, voire de la forme canonique “shortlex” d’une tresse. Nous proposons pour l’instant des procédés qui, nous le pensons, apporte déjà une solution approchée efficace. Nous n’avons pas démontré que celle-ci définissait une méthode de résolution exacte mais nous espérons progresser dans ce sens prochainement. Quoiqu’il en soit, nous avons tout d’abord cherché à définir une **modélisation originale des tresses et de leurs mouvements**. Celle-ci est fondée sur un codage particulier inspiré des comparaisons par paires. Par ailleurs, dans le cadre de cette modélisation, nous proposons une formulation des problèmes algorithmiques autour du groupe des tresses s’apparentant davantage à des problèmes d’**optimisation combinatoire**.

7.2 Modélisation des tresses par approche relationnelle

Dans cette section nous exposons notre approche de la modélisation des tresses. **Dans la continuité des différents aspects développés dans le cadre de ce mémoire de thèse, nous cherchons à exploiter la représentation de relations à partir des comparaisons par paires des objets sur lesquels portent la relation.** Notre interprétation des tresses, en tant qu’extension du groupe des permutations, est fondée sur une relation particulière qu’entretiennent les brins constituant la tresse. De ce fait, les objets étudiés ici sont les brins et nous représentons l’information sous forme de paires de “brins nommés”.

Nous avons introduit en paragraphe 5.1.1.5 et en sous-section 5.2.2 du chapitre 5 un codage original du groupe des permutations. L’idée poursuivie désormais est d’étendre et d’adapter

9. ie $\sigma^1 < (\sigma^1)^{-1} < \sigma^2 < (\sigma^2)^{-1} < \sigma^3 < (\sigma^3)^{-1} \dots$

10. on parle aussi d’ordre lexicographique

11. appelé aussi ordre lexicographique

12. le terme anglo-saxon est “Term rewriting systems”

13. ie noethérien en algèbre

ce codage au groupe des mots de tresse \mathbb{B}_N . Ce codage, nous le verrons, possède des propriétés qui permettent de modéliser de manière plus intuitive tout type de mouvements isotopiques de brins généralisant de ce fait les mouvements de retournements et de réduction de poignées données dans la section précédente.

Nous verrons ensuite, dans la section 7.3, comment cette nouvelle modélisation permet de reconsidérer dans un formalisme original, les problèmes algorithmiques autour des tresses.

7.2.1 Les matrices de paires de brins nommés et les structures de graphe induites

Nous montrons dans cette sous-section, l’extension aux tresses, du codage par matrices de paires d’objets nommés des permutations. Ce codage met en avant les comparaisons par paires des brins. De ce fait, **contrairement aux approches classiques, l’information codée n’est plus uniquement spécifique aux indices de rang j mais fait intervenir également les brins impliqués lors de ces croisements.**

En effet, les générateurs du groupe des tresses sont traditionnellement les croisements σ^j (resp. $(\sigma^j)^{-1}$) qui consistent à faire passer le brin en position j au-dessous (resp. au-dessus) du brin en position $j+1$. Dans notre approche, nous retenons en fait, l’identité des brins en position j et $j+1$ qui font l’objet du croisement.

Prenons un exemple introductif afin d’illustrer notre démarche et dans un premier temps, ne nous soucions pas de l’information topologique “au-dessous” ou “au-dessus” d’un croisement. Pour cela considérons une tresse géométrique à 4 brins représentée par le mot de tresse suivant : $w = bac$ ou encore $w = \sigma^2 \circ \sigma^1 \circ \sigma^3$. Cet exemple à ainsi la particularité de ne pas faire intervenir de générateurs $(\sigma^j)^{-1}$.

Supposons, sans perte de généralité, que les brins notés O^i , $i = 1, 2, 3, 4$ sont initialement rangés selon l’ordre suivant $O^1 < O^2 < O^3 < O^4$. **La tresse représentée dans son codage par matrices de paires de brins nommés consistera en une succession de $|w| + 1$ couples :**

1. **le premier élément est la matrice de paires de brins nommés L représentative de κ , l’ordre¹⁴ courant des brins. Nous rappelons que nous avons :**

$$L_{jj'} = \left[(\kappa)^{-1}(D^j), (\kappa)^{-1}(D^{j'+1}) \right] \quad \forall j = 1, \dots, N; \forall j' = j, \dots, N-1 \quad (7.2.1)$$

où $(\kappa)^{-1}(D^j) \in \{O^1, \dots, O^N\}$, $\forall j = 1, \dots, N$; $\mathbb{O} = \{O^1, \dots, O^N\}$ est l’ensemble des brins et $\mathbb{D} = \{D^1, \dots, D^N\}$ est l’ensemble des rangs

2. **le second élément consiste en la paire de brins nommés et juxtaposés $(O^i, O^{i'})$, identifiant le croisement relatif à un unique générateur σ^j ¹⁵, qui agit¹⁶ *a posteriori* sur L . Ces paires de brins nommés et juxtaposés correspondent¹⁷ à la diagonale des matrices L , nous avons :**

$$\phi(L_{jj}) = ([L_{jj}]_1, [L_{jj}]_2) = (O^i, O^{i'}) \quad (7.2.2)$$

où ϕ , nous le rappelons, est l’application qui fait correspondre une paire de brins nommés $[O^i, O^{i'}]$ au croisement générateur agissant sur les brins : $(O^i, O^{i'})$

14. dans le sens où une tresse peut être réduite à une permutation qui elle même est en correspondance biunivoque avec une unique relation d’ordre strict et total sur les brins cf page 161

15. agissant sur les rangs

16. ie par la composition *

17. cf prop-5.2.3 page 174

Dans le cas de notre exemple nous avons donc la représentation suivante :

Matrices de paires de brins nommés	Codage classique
$\left[\left(\begin{array}{ccc} [O^1, O^2] & [O^1, O^3] & [O^1, O^4] \\ & [O^2, O^3] & [O^2, O^4] \\ & & [O^3, O^4] \end{array} \right); (O^2, O^3) \right]$	b ou σ^2
$\left[\left(\begin{array}{ccc} [O^1, O^3] & [O^1, O^2] & [O^1, O^4] \\ & [O^3, O^2] & [O^3, O^4] \\ & & [O^2, O^4] \end{array} \right); (O^1, O^3) \right]$	a ou σ^1
$\left[\left(\begin{array}{ccc} [O^3, O^1] & [O^3, O^2] & [O^3, O^4] \\ & [O^1, O^2] & [O^1, O^4] \\ & & [O^2, O^4] \end{array} \right); (O^2, O^4) \right]$	c ou σ^3
$\left[\left(\begin{array}{ccc} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ & [O^1, O^4] & [O^1, O^2] \\ & & [O^4, O^2] \end{array} \right); \epsilon \right]$	

De manière générale, nous avons constaté au paragraphe 5.1.1.5 page 158 qu'à toute relation d'ordre strict et total, il existait une correspondance entre les générateurs σ^j agissant sur les rangs et les objets que ces transpositions font finalement permuer. Cette correspondance est le propre du codage par paires d'objets nommés.

Dans le cadre de ce chapitre, les objets sont les brins et **à tout moment, une tresse dans son "histoire" induit une relation d'ordre strict et total sur ses brins**. Les matrices de paires de brins nommés et les paires de brins nommés et juxtaposés permettent ainsi, dans un premier temps, de modéliser la restriction d'une tresse aux différentes permutations qu'elle engendre sur ses brins.

Peut-on aller plus loin et intégrer l'information topologique des croisements qui est spécifique aux tresses ?

Le codage pressenti s'étend en effet, naturellement au cas des tresses pour lesquelles nous ajoutons les paires de brins nommés duaux qui permettent finalement de coder l'information topologique "au-dessous" ou "au-dessus".

En effet, à toute matrice de paires de brins nommés, nous identifions sur la diagonale les paires de brins nommés et juxtaposés $[O^i, O^{i'}]$. Dans le cas des permutations, l'application ϕ permettait de mettre en bijection un élément $[\cdot, \cdot]$, à une paire d'objets nommés et juxtaposés $\phi([\cdot, \cdot]) = (\cdot, \cdot)$ qui correspond à un croisement générateur σ^j . Il n'y a pas d'ambiguïté dans le cas des permutations, ce qui n'est pas le cas pour les tresses qui pour un même croisement, fait état de deux possibilités : $(\sigma^j)^1$ ou $(\sigma^j)^{-1}$.

Pour les tresses, nous pouvons en fait modifier l'application ϕ en ajoutant la notion de signature topologique qui permet de spécifier si le brin qui est le premier élément de la paire passe "au-dessous" ou "au-dessus" du brin qui est le deuxième élément de la paire. Nous avons alors une relation ϕ qui fait correspondre un élément $[\cdot, \cdot]$ soit à une paire de brins nommés $(\cdot, \cdot)^1$ correspondant alors au générateur $(\sigma^j)^1$, soit à une paire de brins nommés $(\cdot, \cdot)^{-1}$ correspondant

alors au générateur $(\sigma^j)^{-1}$. Nous adoptons finalement les notations suivantes :

$\overline{([L_{jj}]_1, [L_{jj}]_2)}$	indique que le brin $[L_{jj}]_1 = (\kappa)^{-1}(D^j)$ passe au-dessous du brin $[L_{jj}]_2 = (\kappa)^{-1}(D^{j+1})$	(7.2.3)
$([L_{jj}]_1, [L_{jj}]_2)$	indique que le brin $[L_{jj}]_1 = (\kappa)^{-1}(D^j)$ passe au-dessus du brin $[L_{jj}]_2 = (\kappa)^{-1}(D^{j+1})$	

Remarque 7.2.1 A propos des notations : mots de paires de brins nommés topologiquement signés ou non

Lorsque l'information topologique "au-dessous", "au-dessus" est indéterminée, ou n'est pas pertinente (lors de certaines définitions faisant intervenir des paires de brins nommés), nous les noterons alors sans barre. Dans ce cas, nous avons :

$$(O^i, O^{i'}) = \overline{(O^i, O^{i'})} \vee (O^i, O^{i'}) \tag{7.2.4}$$

Dans le cas d'un mot de paires de brins nommés dont l'information topologique est indéterminée, nous parlerons notamment de **mots non topologiquement signés ou non signés**.

Dans le cas où c'est l'information topologique qui n'est pas primordiale, c'est alors l'information d'ordre (ou de permutation) " O^i est de rang inférieur ou non à $O^{i'}$ " qui est significative.

Les différentes considérations énoncées précédemment, nous permettent de définir une structure d'automate particulière sur les tresses à N brins.

Définition 7.2.1 Graphe orienté de l'ensemble des tresses d'un ensemble de brins \mathbb{O}

Le graphe orienté représentant une tresse des brins \mathbb{O} peut être représenté par un automate $A(\mathbb{O}) = (\mathbb{E}, \Sigma, \mu, \mathbb{A}, E^\epsilon)$ pour lequel :

- $\mathbb{E} = \text{ROST}(\mathbb{O}) =$ ensemble des relations d'ordre strict et total κ de l'ensemble \mathbb{O} représenté par des matrices de paires de brins nommés L
- $\Sigma = \{\Sigma_\kappa\} =$ ensemble des croisements de brins juxtaposés dépendants d'une relation d'ordre strict et total $\kappa \in \text{ROST}(\mathbb{O})$. Ces croisements sont donnés par les paires de brins nommés de la diagonale de chaque matrice de paires de brins nommés
- $\mu = * :$ composition à droite de la matrice de paires de brins nommés courantes par un de ses croisements propres¹⁸ et dont l'action est donnée par déf-5.2.3 page 175
- $\mathbb{A} = \epsilon$
- $E^\epsilon = \kappa^\epsilon = O^1 < \dots < O^N$ le classement de brins "identité"

Afin d'illustrer cette structure de graphe, nous représentons celle de l'exemple précédent dans la figure fig-7.3

De manière plus générale, nous présentons la structure d'automate que suggère le groupe \mathbb{B}_3 dans la figure fig-7.4.

7.2.2 Codage du groupe des tresses par matrices de paires de brins nommés

Définition 7.2.2 Mot de tresse dans le formalisme d'automate suggéré par le codage par matrice de paires de brins nommés

18. donnés par une paire de brins nommés de la diagonale

$$\begin{aligned}
 & \begin{pmatrix} [O^1, O^2] & [O^1, O^3] & [O^1, O^4] \\ & [O^2, O^3] & [O^2, O^4] \\ & & [O^3, O^4] \end{pmatrix} \\
 & \quad \downarrow \underline{(O^2, O^3)} \\
 & \begin{pmatrix} [O^1, O^3] & [O^1, O^2] & [O^1, O^4] \\ & [O^3, O^2] & [O^3, O^4] \\ & & [O^2, O^4] \end{pmatrix} \\
 & \quad \downarrow \underline{(O^1, O^3)} \\
 & \begin{pmatrix} [O^3, O^1] & [O^3, O^2] & [O^3, O^4] \\ & [O^1, O^2] & [O^1, O^4] \\ & & [O^2, O^4] \end{pmatrix} \\
 & \quad \downarrow \underline{(O^2, O^4)} \\
 & \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ & [O^1, O^4] & [O^1, O^2] \\ & & [O^4, O^2] \end{pmatrix}
 \end{aligned}$$

FIGURE 7.3 – Chemin dans l’automate de \mathbb{B}_4 représentatif du mot de tresse *bac* partant de la tresse triviale

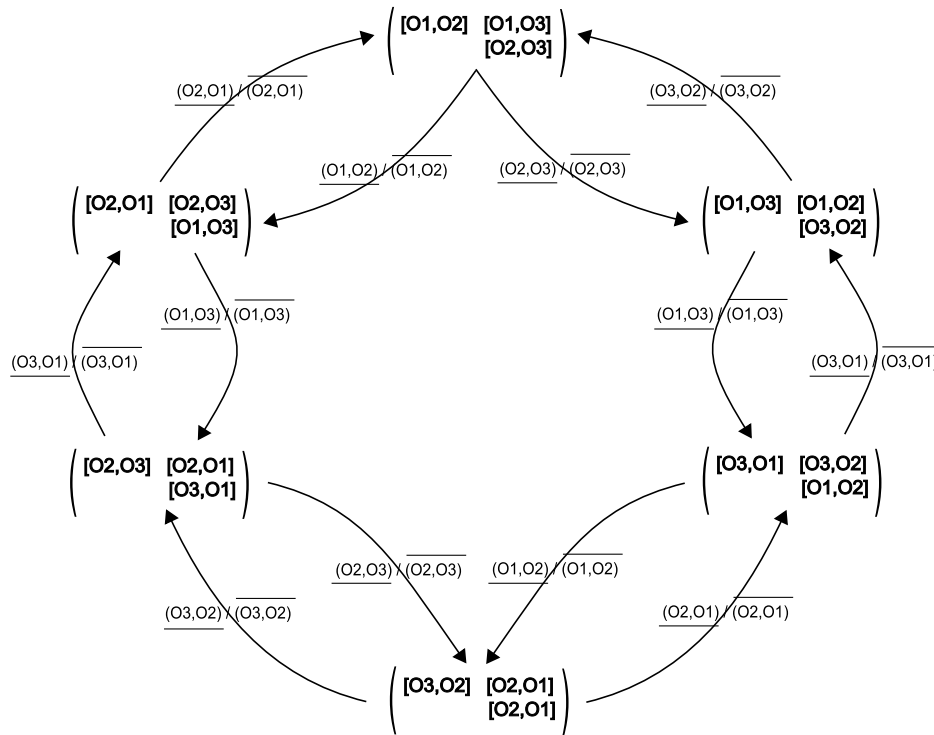


FIGURE 7.4 – Automate du groupe des tresses à 3 brins

Le couple “matrice de paires de brins nommés / paires de brins nommés utilisées comme générateur” donne à un moment donné de l’“histoire” d’une tresse l’ordre dans lequel sont rangés les brins et le

croisement des brins juxtaposés qui s'en suit. Ce couple $(L, (O^i, O^{i'}))$ est tel que :

$$\exists !j : \phi^{-1} \left((O^i, O^{i'}) \right) = [O^i, O^{i'}] = L_{jj} \text{ ce qui correspond de manière biunivoque à } (\sigma^j)^{\pm 1} \quad (7.2.5)$$

Une tresse est donc décrite dans sa totalité en retenant à chaque étape de son histoire les couples $(L, (.,.))$. Cette succession de couples sera notée W et par convention, nous avons :

- l'état de départ est la tresse triviale pour laquelle l'ordre des brins est donné par le codage par matrice de paires de brins nommés L codant $\kappa^\epsilon = O^1 < \dots < O^N$
- l'état final est un couple (L, ϵ) où L code le résultat de l'ordre (ou de la permutation) des brins de la tresse à la suite duquel il n'y a plus aucun croisement

En théorie, il est nécessaire de conserver le couple "matrice de paires de brins nommés / paires de brins nommés et juxtaposés utilisées comme générateur" dans la mesure où ce dernier, qui est un générateur agissant sur des brins juxtaposés en correspondance¹⁹ avec un générateur agissant sur les rangs, est dépendant de l'ordre donné par L dans lequel sont rangés les brins.

Cependant, nous pouvons en pratique conserver uniquement la suite des paires de brins nommés $(.,.)$, c'est à dire les croisements générateurs de brins nommés et juxtaposés. En effet, comme par convention l'état de départ est la tresse triviale et comme les matrices L sont les images obtenues à partir de l'application successive des croisements, nous pouvons alors à tout moment retrouver la matrice L de l'ordre des brins au cours de l'"histoire" d'une tresse. De cette manière, nous noterons par w la succession des paires $(.,.)$, excepté le second membre donné par le couple $(L^{|W|}, \epsilon)$. Ainsi, de manière générale, nous avons $|W| = |w| + 1$.

w décrit l'"histoire" d'une tresse et est en correspondance avec le mot de tresse classique utilisant les générateurs $(\sigma^j)^{\pm 1}$. Par exemple $w = (O^2, O^3)(O^1, O^3)(O^2, O^4)$ associé à l'ordre initial $O^1 < O^2 < O^3 < O^4$, est le mot de tresse dans le codage des paires de brins nommés et juxtaposés correspondant au mot de tresse classique $w = bac$.

Dés lors, les relations d'Artin $R1, R2$ et $R3$ s'interprètent respectivement, dans le codage des paires de brins nommés, de la manière suivante :

Propriété 7.2.1 Relations d'Artin exprimées dans le codage des croisements générateurs de brins juxtaposés

$$\begin{array}{l}
 R1 : \quad \underline{(O^i, O^{i'})} \underline{(O^{i'}, O^i)} \equiv \overline{(O^i, O^{i'})} \overline{(O^{i'}, O^i)} \equiv \epsilon \quad \forall i \neq i' \\
 R2 : \quad (O^i, O^{i'}) (O^{i''}, O^{i'''}) \equiv (O^{i''}, O^{i'''}) (O^i, O^{i'}) \quad \forall i \neq i' \neq i'' \neq i''' \\
 R3 : \quad (O^i, O^{i'}) (O^i, O^{i''}) (O^{i'}, O^{i''}) \equiv (O^{i'}, O^{i''}) (O^i, O^{i''}) (O^i, O^{i'}) \\
 \text{Sauf pour} \quad \underline{(O^i, O^{i'})} \overline{(O^i, O^{i''})} \underline{(O^{i'}, O^{i''})} \neq \underline{(O^{i'}, O^{i''})} \overline{(O^i, O^{i''})} \underline{(O^i, O^{i'})} \\
 \text{les} \\
 \text{cas suivants :} \quad \overline{(O^i, O^{i'})} \underline{(O^i, O^{i''})} \overline{(O^{i'}, O^{i''})} \neq \overline{(O^{i'}, O^{i''})} \underline{(O^i, O^{i''})} \overline{(O^i, O^{i'})}
 \end{array} \quad (7.2.6)$$

où ϵ représente symboliquement, le générateur identité ou trivial qui ne donne aucun croisement.

Nous donnons à présent la définition du support d'un mot w qui nous servira par la suite.

19. cf prop-5.2.3 page 174

Définition 7.2.3 *Support d'un mot de tresse w dans le cas du codage par paires de brins nommés et juxtaposés*

Soit w un mot de tresse représentant les seconds termes d'une succession W de couples $(L, (.,.))$ décrivant une tresse dans le cadre de la modélisation par automate du groupe \mathbb{B}_N donnée par déf-7.2.2. Le support de w noté $\text{Supp}(w)$ est alors l'ensemble des brins O^i qui ont fait l'objet d'au moins un croisement dans w .

Prenons l'exemple du mot suivant $w' = (O^1, O^2)(O^1, O^3)$ issu d'un chemin W' dans l'automate de \mathbb{B}_4 décrivant une tresse. Nous avons alors $\text{Supp}(w') = \{O^1, O^2, O^3\}$.

Au terme de la présentation de cette nouvelle approche de modélisation des tresses, nous retiendrons les points suivants :

- les croisements générateurs qui définissent les tresses sont représentés par des paires de brins nommés et juxtaposés et sont donc dépendants de l'ordre dans lequel sont rangés les brins de la tresse à un moment donné de son histoire
- une tresse est dans notre codage, une succession ou encore un chemin de paires de brins nommés dans l'automate représentatif du groupe \mathbb{B}_N que nous avons, par exemple, présenté pour $N = 3$ dans la figure fig-7.4
- chaque paire de brins nommés $(O^i, O^{i'})$ ou $(\overline{O^i}, \overline{O^{i'}})$ définissant un croisement générateur agissant sur des brins juxtaposés d'une tresse, est en correspondance biunivoque avec un générateur $(\sigma^j)^{\pm 1}$ agissant sur les "rangs" des brins d'une tresse. A tout moment de l'"histoire" d'une tresse, cette correspondance est donnée par la diagonale des matrices L de paires de brins nommés au sens de la relation (7.2.5). Ainsi, il nous est constamment possible de revenir au codage classique, en termes de $(\sigma^j)^{\pm 1}$, d'un mot de tresse. **De ce fait, notre modélisation est au moins équivalente à celle d'E. Artin. Quel est alors l'intérêt des paires de brins nommés ?**

Nous montrons dans la sous-section suivante comment notre modélisation basée sur les paires de brins nommés permet de déterminer plus naturellement tout type de mouvement isotopique des brins d'une tresse.

7.2.3 Mouvements isotopiques de brins libres

Nous cherchons à modéliser les différents déplacements possibles d'un brin d'une tresse qui ne change pas la nature de cette dernière. En d'autres termes, **nous cherchons à déterminer d'autres chemins dans l'automate du groupe \mathbb{B}_N qui soit topologiquement équivalent au chemin décrivant la tresse étudiée.**

Nous nous basons sur la modélisation donnée précédemment. De ce fait, une tresse est une succession de couples $(L, (.,.))$ que nous notons par W . Nous avons vu qu'il était en pratique possible de ne conserver que les seconds membres de ces couples, ce qui nous permet d'avoir un formalisme proche de la modélisation classique d'E. Artin en terme de mots de tresse. Nous notons ainsi, par w la succession des paires de brins nommés générateurs représentant une tresse. Par ailleurs, seules les paires de brins nommés de la diagonale des matrices L nous intéressent *a priori* en raison des relations d'Artin données dans ce codage par (7.2.6). Pourquoi alors, conserver les matrices L dans notre formalisme ?

En fait, ce sont elles, qui en identifiant les paires de brins nommés non nécessairement juxtaposés, vont nous permettre de modéliser tous les déplacements isotopiques de chaque brin comme nous le montrons dans ce qui suit.

Notre démarche consiste à déplacer des brins "libres" définis ci-dessous. Ces brins forment dans des cas particuliers des σ^j -poignées au sens de la définition donnée par P. Dehornoy en

déf-7.1.5.

Définition 7.2.4 Brins libres

Soit w un mot de tresse dans le codage des paires de brins nommés et juxtaposés.

Nous dirons que le brin O^i est libre si nous avons l'un des motifs donnés par les combinaisons suivantes :

$$\left. \begin{array}{l} 1. \overline{(O^i, O^{i'})} \\ 2. \overline{(O^i, O^{i'})} \\ 3. \overline{(O^{i'}, O^i)} \\ 4. \overline{(O^{i'}, O^i)} \end{array} \right\} v \left\{ \begin{array}{l} \overline{(O^{i''}, O^i)} \text{ a.} \\ \overline{(O^{i''}, O^i)} \text{ b.} \\ \overline{(O^i, O^{i''})} \text{ c.} \\ \overline{(O^i, O^{i''})} \text{ d.} \end{array} \right. \quad (7.2.7)$$

où v est un sous-mot non vide ne faisant aucunement intervenir le brin O^i . Chaque couple de l'ensemble cartésien $\{1, 2, 3, 4\} \times \{a, b, c, d\}$ représente un motif définissant un brin libre. Il existe ainsi 16 cas différents. De manière générique, nous noterons par $O^i v O^i$ un brin libre O^i ou encore une configuration de brin libre, qui représentera un des 16 motifs donnés ci-dessus.

Nous illustrons dans la figure fig-7.5 différents cas de brins libres.

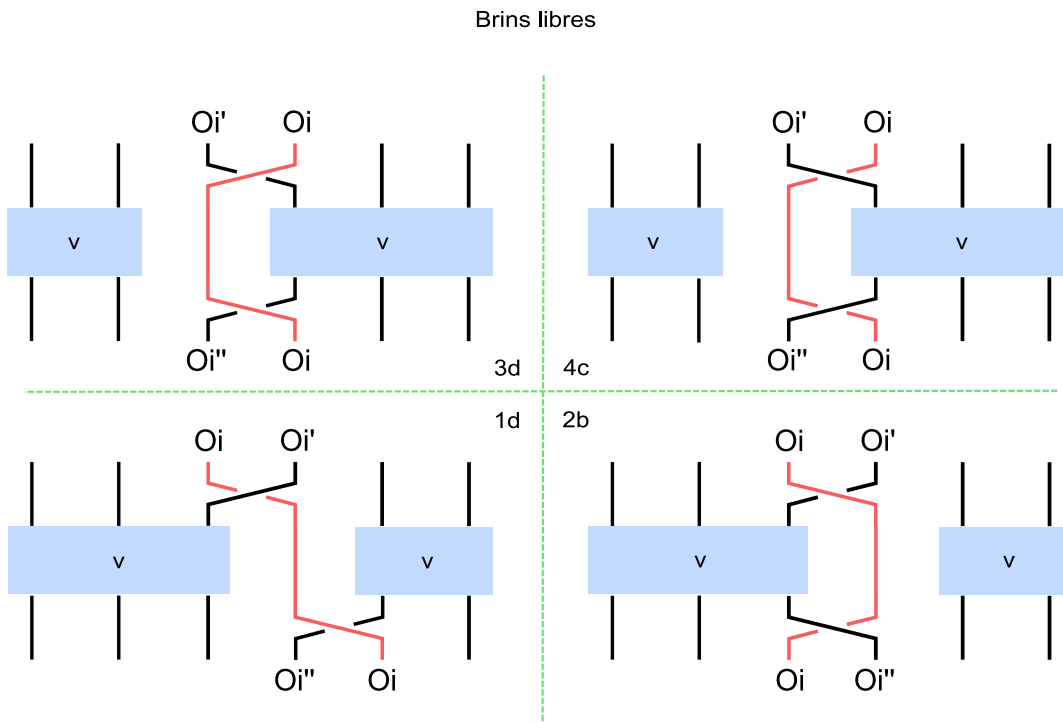


FIGURE 7.5 – Différents cas de brins libres

Les cas 3d et 4c en particulier forment respectivement les motifs $\overline{(O^{i'}, O^i)} v \overline{(O^i, O^{i''})}$ et $\overline{(O^{i'}, O^i)} v \overline{(O^i, O^{i''})}$ et sont de ce fait, des σ^j -poignées telles que définies dans le paragraphe 7.1.1.3.

Définition 7.2.5 Brins libres de type "poignée" et de type "poignée de Dehornoy"

De manière générale, nous dirons que la configuration $O^i v O^i$ vérifie un motif "poignée" si nous avons

un des 8 couples suivants : $\{1, 2\} \times \{a, b\} \cup \{3, 4\} \times \{c, d\}$. Parmi cet ensemble, nous appellerons "poignée de Dehornoy" les configurations suivantes : 1b, 2a, 3d, 4c.

Nous remarquons une propriété qui découle de la définition des brins libres :

Propriété 7.2.2 Rang constant d'un brin libre O^i lors d'une configuration $O^i v O^i$

Pour une configuration de brin libre $O^i v O^i$, le rang du brin libre O^i est par définition toujours le même pour l'ensemble des ordres relatifs au mot v . Nous noterons alors par j^i l'indice du rang D^{j^i} du brin libre O^i .

Nous précisons maintenant ce que nous entendons par mouvement de brins libres.

En préambule, il est utile de rappeler la propriété de décomposition d'une transposition d'objets non juxtaposés en produit de transpositions génératrices d'objets juxtaposés dans le cas du groupe des permutations d'un ensemble \mathbb{O} . Nous avons dans un premier temps rappelé cette propriété prop-5.1.14 en page 163, dans le cadre du codage classique utilisant les générateurs σ^j agissant sur les rangs. Puis, dans un second temps, nous avons exprimé cette propriété dans le formalisme des matrices de paires d'objets nommés en prop-5.2.4 page 178. Nous avons vu dans le cadre de ces dernières, que la décomposition d'une transposition de deux objets non juxtaposés appliqués à un ordre strict et total κ était décrite par un chemin dont les éléments en termes de paires d'objets nommés et juxtaposés étaient donnés par un circuit "triangulaire" au sein de la matrice L représentative de l'ordre κ . Afin d'illustrer ces propos, nous rappelons l'exemple donné page 179.

Considérons l'ordre strict et total $\kappa' = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ D^2 & D^4 & D^1 & D^3 \end{pmatrix}$ dont la matrice de paires d'objets nommés L' est la suivante :

$$L' = \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ & [O^1, O^4] & [O^1, O^2] \\ & & [O^4, O^2] \end{pmatrix}$$

Si nous appliquons la transposition des objets en position 1 et en position 4, cela revient à permuter les objets qui constituent le terme L'_{13} . Celui-ci correspond à la paire $[O^3, O^2]$ qui donne la transposition génératrice agissant sur les objets suivants : $\phi([O^3, O^2]) = (O^3, O^2)$. Selon la propriété prop-5.2.4 donnée page 178, nous avons alors : $L' * L'_{13} = L * L'_{11} * L'_{12} * L'_{13} * L'_{33} * L'_{23}$. Cette expression est aussi égale à : $L' * L'_{13} = L' * L'_{33} * L'_{23} * L'_{13} * L'_{11} * L'_{12}$.

Les deux types de chemins parcourus par cette action au vu de la matrice des paires d'objets nommés décrivent un chemin triangulaire qui est représenté ci-dessous :

$$\left(\begin{array}{ccc} [O^3, O^1](\mathbf{1}) & \rightarrow & [O^3, O^4](\mathbf{2}) & \rightarrow & [O^3, O^2](\mathbf{3}) \\ & & [O^1, O^4] & & [O^1, O^2](\mathbf{5}) \\ & & & & \uparrow \\ & & & & [O^4, O^2](\mathbf{4}) \end{array} \right)$$

Dans le cas des tresses qui nous intéressent ici, nous prenons comme objets les brins. **Les déplacements de brins libres sont issus d'une démarche qui rappelle les propriétés rappelées précédemment.**

Nous donnons dans un premier temps une illustration graphique de l'approche développée. Nous nous restreignons d'abord au cas des tresses pour lesquelles l'information topologique

“au-dessous” ou “au-dessus” est abandonnée²⁰. Nous représentons donc graphiquement d’une part, la propriété précédente qui n’est pas sans rappeler les générateurs de Birman, Ko et Lee [BKL98]; d’autre part, l’extension de cette propriété au cas du déplacement de brins libres.

Nous avons ainsi représenté à gauche de la figure fig-7.6 le déroulement de la transposition de deux brins non juxtaposés qui illustre le chemin triangulaire de l’exemple précédent. Nous faisons remarquer que sur la figure, l’absence d’information topologique est symbolisée par l’absence de coloriage lors d’un croisement de deux brins.

A droite de cette figure, nous avons représenté dans un cas simple, ce que nous entendons par un déplacement à droite d’un brin libre qui est dans notre exemple représenté par le brin O^3 . Nous avons cherché notamment à illustrer le fait que les déplacements de brins libres décrivent aussi des chemins au sein des matrices de paires de brins nommés qui sont proches de ceux décrits par les transpositions de brins non juxtaposés.

Nous retiendrons en particulier le fait qu’il y ait deux phases lors d’un déplacement d’un brin libre. Dans notre exemple qui est un déplacement à droite, nous avons :

1. une **phase ascendante** où le brin libre croise les brins qui le succèdent au sens de l’ordre strict et total dans lequel sont rangés les brins avant le déplacement
2. une **phase descendante** où le brin libre croise à nouveau les mêmes brins mais qui, en raison de la première phase, le précède depuis. Le brin libre revient ainsi à son rang initial.

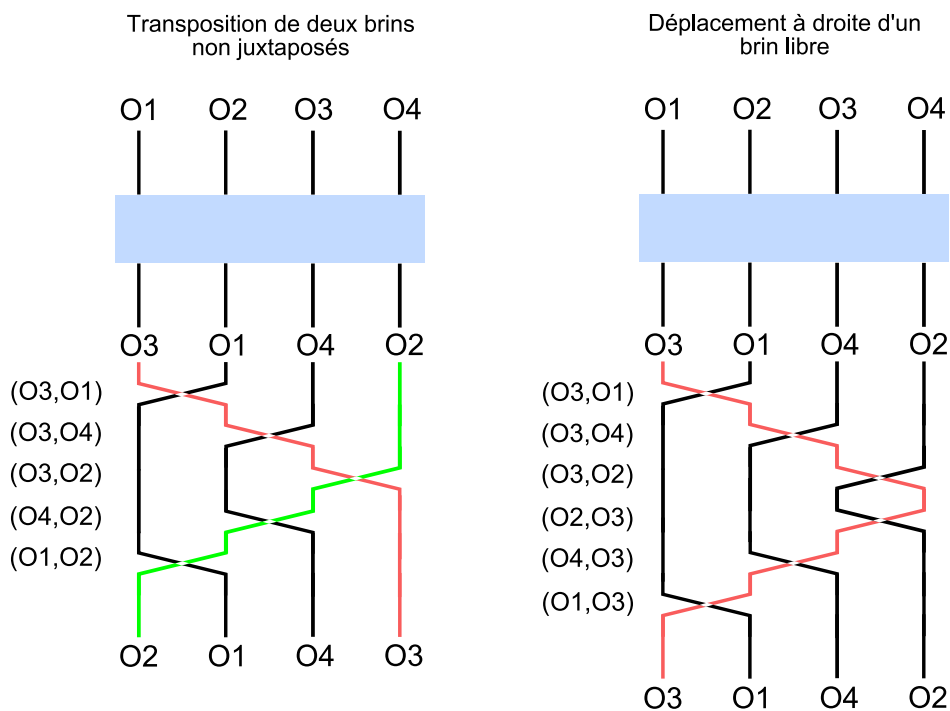


FIGURE 7.6 – Transposition de brins non juxtaposés et déplacement d’un brin libre

Au sens de la matrice de paires de brins nommés codant l’ordre strict et total dans lesquels sont rangés les brins avant la première transposition génératrice engendrant le déplacement, le chemin que ce dernier décrit est, dans le cas de l’exemple précédent, représenté par la succession de paires de brins nommés suivante :

²⁰. ie lorsque nous nous ramenons à des permutations. Dans ce cas, les paires de brins nommés et juxtaposés $(O^i, O^{i'})$ et $(O^i, O^{i'})$ sont indifférentes

- phase ascendante :

$$\left(\begin{array}{l} [O^3, O^1](1) \rightarrow [O^3, O^4](2) \rightarrow [O^3, O^2](3) \\ \\ [O^1, O^4] \qquad [O^1, O^2] \\ \\ [O^4, O^2] \end{array} \right)$$

- phase descendante :

$$\left(\begin{array}{l} r([O^3, O^1])(6) \leftarrow r([O^3, O^4])(5) \leftarrow r([O^3, O^2])(4) \\ \\ [O^1, O^4] \qquad [O^1, O^2] \\ \\ [O^4, O^2] \end{array} \right)$$

où $r([O^i, O^{i'}]) = [O^{i'}, O^i]$ est l'application qui permute les brins d'une paire de brins nommés

Il s'agit d'un chemin de type "va et vient" qui peut être appréhendé à partir des matrices de paires de brins nommés comme nous venons de le voir sur un exemple.

Nous nous sommes placés dans le cas simple qui correspond ici à un brin libre O^3 du type $(O^3, O^1)v(O^1, O^3)$ où $v = \epsilon$. Par ailleurs, nous ne considérons pas encore les informations topologiques des croisements.

La figure fig-7.7 représente un cas plus général qui est celui du déplacement à droite d'un brin libre dans le cas où $v \neq \epsilon$. Cette figure nous permet de visualiser les définitions qui vont suivre.

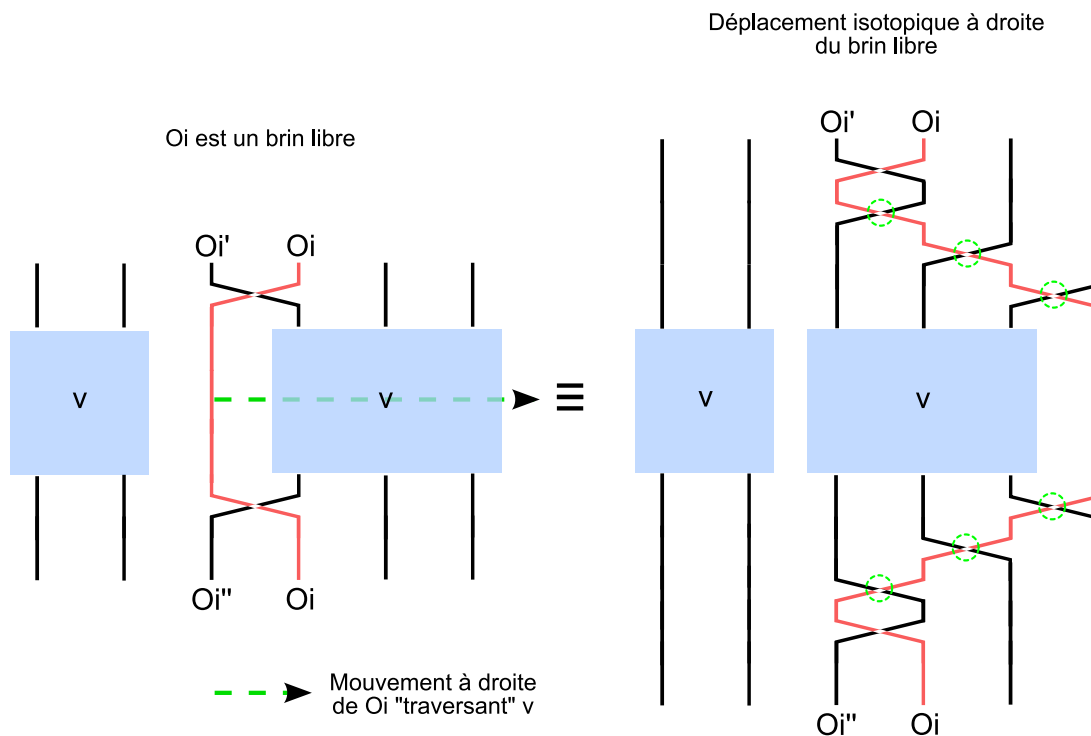


FIGURE 7.7 – Illustration d'un déplacement à droite d'un brin libre sans tenir compte dans un premier temps de la nature topologique des croisements

Définition 7.2.6 Déplacement à droite d'un brin libre

Soit une tresse géométrique β représenté dans le codage de paires de brins nommés par le chemin W dans

l'automate représentatif de \mathbb{B}_N au sens des définitions déf-7.2.1 et déf-7.2.2. Soit un brin libre O^i défini par un des 16 motifs $O^i v O^i$ donnés par déf-7.2.4. Le déplacement à droite de O^i noté $\xrightarrow{O^i}$ consiste en deux phases successives :

1. une **phase ascendante** qui consiste à faire croiser vers la droite le brin libre O^i avec des brins intervenant dans le sous-mot v qui ont un rang supérieur au brin libre. Ceci se traduit par le rajout d'une succession de paires de brins nommés et juxtaposés avant le sous-mot v
2. une **phase descendante** qui consiste à faire croiser vers la gauche le brin libre O^i avec ces mêmes brins intervenant dans le sous-mot v qui, après la phase ascendante, sont de rang inférieur ou égal au rang initial du brin libre. Ceci se traduit par le rajout d'une succession de paires de brins nommés et juxtaposés après le sous-mot v

Définition 7.2.7 Déplacement à gauche d'un brin libre

Soit une tresse β élément de \mathbb{B}_N représenté dans le codage de paires de brins nommés par le chemin W dans l'automate représentatif de \mathbb{B}_N . Soit un brin libre O^i défini par un des 16 motifs $O^i v O^i$. Le déplacement à gauche de O^i noté $\xleftarrow{O^i}$ consiste en deux phases successives :

1. une **phase descendante** qui consiste à faire croiser vers la gauche le brin libre O^i avec des mêmes brins intervenant dans le sous-mot v qui ont un rang inférieur au brin libre. Ceci se traduit par le rajout d'une succession de paires de brins nommés et juxtaposés avant le sous-mot v
2. une **phase ascendante** qui consiste à faire croiser vers la droite le brin libre O^i avec ces mêmes brins intervenant dans le sous-mot v qui, après la phase descendante, sont de rang supérieur ou égal au rang initial du brin libre. Ceci se traduit par le rajout d'une succession de paires de brins nommés et juxtaposés après le sous-mot v

Dans les définitions précédentes, nous faisons référence à des ajouts de paires de brins nommés et juxtaposés avant et après le sous-mot v . Notre objectif désormais est de montrer comment déterminer ces paires de brins nommés qui représentent le mouvement d'un brin libre.

Nous introduisons d'ores et déjà les notations et quelques définitions générales que nous emploierons par la suite.

W est donc le chemin dans l'automate représentatif de \mathbb{B}_N qui code une tresse β à N brins. Soit pour cette tresse, un brin libre O^i défini par un des 16 motifs $O^i v O^i$. Nous notons par U la succession de couples extraits de W représentant la partie de la tresse correspondant au motif $O^i v O^i$ avec $|U| = |v| + 2$.

Nous raisonnerons désormais localement au sous-mot correspondant au brin libre. Ainsi, pour alléger les notations, nous ne ferons pas référence à U dans ce qui suit : le k -ème élément du sous-chemin U est noté U_k ce que nous notons en termes d'éléments de l'automate représentatif de \mathbb{B}_N par $(L^k, (.,.)^k)$.

Pour fixer les idées, considérons l'expression générale suivante :

$$U = (L^1, (.,.)^1) \underbrace{(L^2, (.,.)^2) \dots (L^{|U|-1}, (.,.)^{|U|-1})}_{\text{partie correspondant au sous-mot } v} (L^{|U|}, (.,.)^{|U|}) \tag{7.2.8}$$

où $(.,.)^1$ et $(.,.)^{|U|}$ sont des générateurs faisant intervenir le brin libre O^i correspondant respectivement à un des cas de $\{1, 2, 3, 4\}$ et de $\{a, b, c, d\}$. Nous avons par ailleurs $u = (.,.)^1 \underbrace{(.,.)^2 \dots (.,.)^{|U|-1}}_v (.,.)^{|U|}$,

le mot de paires de brins nommés constitué des seconds membres des couples de U .

Afin d'introduire plus aisément les définitions et les propriétés obtenues, nous les illustrerons à partir d'un exemple particulier.

L'exemple que nous nous proposons de traiter est représenté dans la figure fig-7.8.

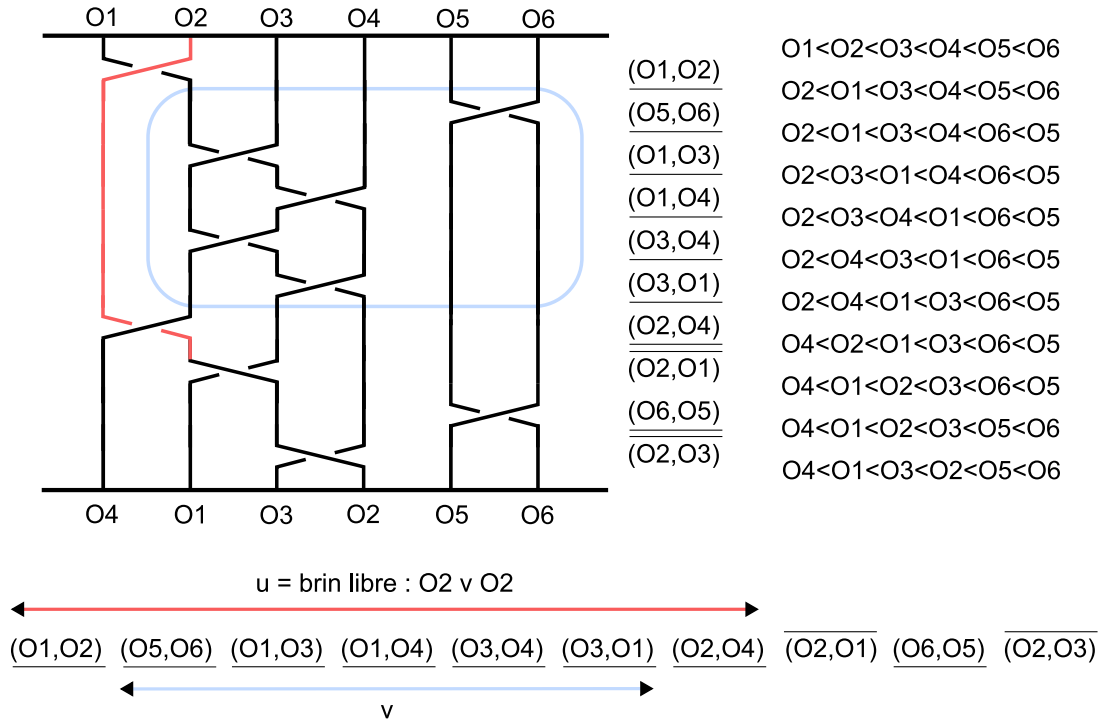


FIGURE 7.8 – Exemple de configuration d'un brin libre O^2vO^2

Sur cette figure, nous avons représenté en rouge un brin libre O^2 . Celui-ci correspond, dans le cadre de notre définition de brins libres en déf-7.2.4, au motif de configuration 3c. Nous avons u la succession des seconds membres des couples de U qui vaut dans ce cas particulier :

$$u = \underbrace{(O^1, O^2)(O^5, O^6)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_v (O^2, O^4) \quad (7.2.9)$$

Si nous ne tenons pas compte du premier et dernier termes de u relatif au brin libre O^2 , nous isolons v qui vaut dans ce cas :

$$v = \underline{(O^5, O^6)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)} \quad (7.2.10)$$

Pour notre exemple, nous avons $|U| = |v| + 2 = 7$. Nous remarquons au passage, que l'indice du rang du brin libre O^2 est bien constant au cours des ordres stricts et totaux représentés par les matrices de paires de brins nommés codant les ordres successifs générés par le mot v et qui vaut ici $j^i = j^2 = 1$.

Nous cherchons à présent, à effectuer un mouvement à droite du brin libre O^2 . Dans un premier temps, nous cherchons à déterminer les paires de brins nommés permettant d'effectuer les phases ascendante puis descendante du mouvement à droite de O^2 . Nous investiguons cette première question dans le paragraphe suivant.

Dans un second temps, après avoir identifié les paires de brins nommés représentant le déplacement d'un brin libre, nous cherchons à déterminer toutes les différentes possibilités de combinaisons de croisements possibles lors de ce déplacement en terme de nature topologique de chaque croisement constituant le mouvement. En effet, il existe plusieurs manières

de déplacer un brin libre. Celui-ci par exemple, peut être au-dessus de tous les brins à croiser ou alors au-dessous. Il existe aussi des mouvements possibles qui consistent à “traverser” les brins à croiser en se “faufilant” entre eux. Pour identifier ces différentes possibilités, nous introduisons les notions de **signature topologique d’un mot de paires de brins nommés (non signé) et de relation d’ordre “topologique” associé à un mot de tresse.**

Dans chacune des deux sous-sections, nous donnerons les propriétés dans un cadre général mais par souci de clarté nous ne manquerons pas d’illustrer ces propriétés sur l’exemple présenté précédemment relatif à la figure fig-7.8.

7.2.3.1 Identification des paires de brins nommés et juxtaposés consécutives à ajouter lors du déplacement du brin libre

Nous introduisons ci-dessous des relations binaires entre brins d’une tresse qui est engendrée par un mot ou un sous-mot de tresse. Celles-ci permettent dans notre contexte, de **définir les brins à croiser lors d’un mouvement de brin libre.**

Définition 7.2.8 Relations binaires “topologiques” engendrées par un mot de paires de brins nommés

Soit w un mot de paires de brins nommés représentant totalement ou partiellement une tresse au sens de déf-7.2.1. La matrice relationnelle, C^w , représentant la relation binaire topologique engendrée sur les brins et relative à w , est donnée par, $\forall i, i' = 1, \dots, N$:

$$C_{ii'}^w = \begin{cases} 1 & \text{si } (O^i, O^{i'}) \in w \text{ ou si } (\overline{O^{i'}}, \overline{O^i}) \in w \text{ ou si } i = i' \\ 0 & \text{sinon} \end{cases} \quad (7.2.11)$$

En d’autres termes, la matrice relationnelle C^w , code la relation \mathcal{R}^w qui est l’ensemble des paires de brins $(O^i, O^{i'})$ pour lesquelles, “ O^i est au-dessous de $O^{i'}$ selon w ”. Au delà de cette définition, si l’on considère sa fermeture symétrique et transitive, cette relation binaire engendre une **partition dite de brins liés**. Cette relation nous intéresse davantage étant donné qu’elle nous aide à identifier les brins qu’il faut croiser lors d’un mouvement de brin libre. Nous la définissons formellement ci-dessous pour un mot quelconque w de paires de brins nommés.

Définition 7.2.9 Relation d’équivalence de liaison entre brins engendrée par un mot de paires de brins nommés

Nous appellerons relation d’équivalence entre brins liés engendrée par le mot w et représentée par une matrice relationnelle de partition $C^{w,e}$, la fermeture transitive de la relation binaire topologique C^w symétrisée qui dans ce cas vaut $C^w \oplus \check{C}^w$. Nous avons alors :

$$C^{w,e} = (C^w \oplus \check{C}^w) \oplus (C^w \oplus \check{C}^w)^2 \oplus \dots \oplus (C^w \oplus \check{C}^w)^{N-1} \quad (7.2.12)$$

où $\check{C}^w = {}^t C^w$ est la converse ou transposée de C^w , \oplus est l’union relationnelle, et où les puissances sont prises au sens des multiplications relatives \odot . Nous rappelons que ces opérations ont été introduites page 115.

Reprenons notre exemple et considérons le sous-mot v qui représente l’ensemble des croisements que le brin libre O^2 devra tenir compte lorsqu’il se déplacera à leur droite”. Nous avons $v = (O^5, O^6)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)$. Nous déterminons les matrices relationnelles binaires et d’équivalence de brins liés engendrées par le sous-mot v . Nous avons dans

notre exemple les résultats suivants :

$$C^v = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.13)$$

$$C^v \oplus \check{C}^v = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 \end{pmatrix} \end{matrix} \quad (7.2.14)$$

La matrice relationnelle $C^v \oplus \check{C}^v$ étant déjà transitive, nous avons donc $C^{v,e}$ qui est identique à celle-ci.

Nous pouvons identifier à partir de la matrice relationnelle représentative de la relation d'équivalence $C^{v,e}$, le sous-ensemble de brins que le brin libre doit "traverser" au cours de son mouvement. En effet, **nous prendrons comme définition de nos mouvements isotopiques de brins libres $O^i v O^i$, le déplacement du brin O^i qui "traverse" la classe de brins liés qui lui est juxtaposée à sa droite²¹, s'il s'agit d'un mouvement à droite; à sa gauche²², s'il s'agit d'un mouvement à gauche.**

Nous illustrons ces propos à l'aide de notre exemple qui représente le cas d'un déplacement à droite. Sur la figure fig-7.8, nous voyons qu'il n'est pas en effet nécessaire pour un mouvement à droite du brin O^2 de croiser les brins O^6 et O^5 . Nous faisons une remarque ci-dessous qui justifie notre définition.

Remarque 7.2.2 "Coarse" et "convex" réduction

Nous précisons que nous pouvons éviter de déterminer cet ensemble minimal et considérer que les mouvements doivent parcourir l'ensemble des brins de rangs supérieurs s'il s'agit d'un déplacement à droite et de rangs inférieurs s'il s'agit d'un déplacement à gauche. Ce cas revient alors à une "coarse" réduction définie par P. Dehornoy [Deh97]. A contrario, identifier l'ensemble minimal des brins à croiser comme ci-dessus correspond à une "convex" réduction selon le même auteur. Nous préférons cette deuxième approche dans la mesure où elle engendre moins de croisements et pour des raisons d'optimisation que nous évoquerons ultérieurement lors de l'étude du problème du mot de tresse minimal.

Nous avons pour l'exemple traité, les classes d'équivalence suivantes $\{O^1, O^3, O^4\}$, $\{O^5, O^6\}$ et $\{O^2\}$. Nous voyons que les brins O^1, O^3, O^4 forment une classe qui est disjointe de la classe formée des brins O^5 et O^6 . Ceci se traduit graphiquement par l'absence de croisement entre les brins de ces deux classes. Par ailleurs, nous faisons remarquer que le brin O^2 étant le brin libre de configuration $O^2 v O^2$, il n'appartient pas, par définition, au support de v . Ceci explique que, relativement au sous-mot v , il ne croise aucun brin et forme ainsi une classe singleton. Ceci est une propriété générale du brin libre au vu de sa configuration $O^i v O^i$.

Nous constatons alors que la relation d'équivalence des brins liés permet d'isoler la classe des brins à traverser puisqu'en effet, le brin libre O^2 aura à croiser les brins de la classe $\{O^1, O^3, O^4\}$ lors de son déplacement à droite.

21. ie la classe dont l'un des brins est d'indice de rang $j^i + 1$

22. ie la classe dont l'un des brins est d'indice de rang $j^i - 1$

Comment de manière générale, connaître la classe de brin qu'il faut "traverser" lors d'un mouvement d'un brin libre ?

Définition 7.2.10 Ensemble des brins à "traverser" lors d'un mouvement de brin libre

Soit une configuration de brin libre $O^i v O^i$ représentée par U la succession de couples $(L^k, (.,.)^k)$; $k = 1, \dots, |U| = |v| + 2$. Soit $C^{w,e}$ la relation d'équivalence de brins liés engendrée par v sur les brins de la tresse. Soit j^i l'indice du rang du brin libre O^i constant au cours des ordres stricts et totaux engendrés par les paires de brins nommés du mot $v : L^2, L^3, \dots, L^{|U|-1}$. Soient alors l'indice de brin i^c défini de la manière suivante :

$$i^c = \begin{cases} \text{indice du brin } [L_{j^i, j^i}^2]_2 & \text{s'il s'agit d'un mouvement à droite} \\ \text{indice du brin } [L_{j^i-1, j^i-1}^2]_1 & \text{s'il s'agit d'un mouvement à gauche} \end{cases} \quad (7.2.15)$$

En d'autres termes i^c est l'indice du brin qui est en position $j^i + 1$ pour l'ordre strict et total représenté par L^2 , s'il s'agit d'un mouvement à droite. i^c est l'indice du brin qui est en position $j^i - 1$ pour l'ordre strict et total représenté par L^2 , s'il s'agit d'un mouvement à gauche.

Nous rappelons une propriété des matrices relationnelles de partition qui nous sera utile :

Définition 7.2.11 Cardinal de la classe d'un brin issue d'une relation d'équivalence de brins liés

Soit $C^{w,e}$ une matrice relationnelle représentative d'une relation d'équivalence de brins liés relative à un mot w . Notons alors par $\mathbb{C}\mathbb{E}(O^i | C^{w,e})$ la classe d'équivalence de O^i , c'est à dire l'ensemble des brins qui sont dans la même classe que O^i selon $C^{w,e}$. Nous avons :

$$\mathbb{C}\mathbb{E}(O^i | C^{w,e}) = \{O^{i'} \in \mathbb{O} : C_{ii'}^{w,e} = 1\} \quad (7.2.16)$$

Nous avons en outre :

$$\#\mathbb{C}\mathbb{E}(O^i | C^{w,e}) = \sum_{i'=1}^N C_{ii'}^{w,e} = C_{i.}^{w,e} \quad (7.2.17)$$

Nous montrons à présent un premier intérêt de la modélisation des tresses par des matrices de paires de brins nommés. En effet, en considérant les définitions précédentes, **nous montrons comment à partir des matrices de paires de brins nommés et de la notion de relation d'équivalence de brins liés engendrée par un mot de paires de brins nommés, nous pouvons identifier les brins qu'il faut "traverser" lors d'un déplacement d'un brin libre ainsi que les paires de brins nommés correspondant qui "matérialisent" les croisements.**

Propriété 7.2.3 Identification des paires de brins nommés représentatives de la phase ascendante et descendante du mouvement d'un brin libre à droite

Soit une configuration de brin libre $O^i v O^i$ représentée par $U = (L^k, (.,.)^k)$, $k = 1, \dots, |U|$ la succession des couples formant la sous-tresse étudiée au sens de l'automate représentatif de \mathbb{B}_N . Soit j^i l'indice du rang de O^i qui est constant tout au long du sous-mot v . Soit i^c l'indice du brin (lié au brin libre) dont la classe, au sens de la relation d'équivalence de brins liés engendrée par v , donne les brins que le brin libre doit croiser lors de son déplacement. Cet indice dépend des configurations distinctes de brin libre, nous avons donné les différentes possibilités précédemment dans déf-7.2.10.

Soit L^2 la matrice de paires de brins nommés en 2-ème position dans U . Les paires de brins nommés consécutives constituant la phase ascendante du déplacement à droite d'un brin libre O^i sont alors données par l'ensemble suivant :

$$\boxed{\{L_{j^i, j^i+j'}^2; j' = 0, \dots, C_{i^c}^{v,e} - 1\}} \quad (7.2.18)$$

Nous considérons que cet ensemble est ordonné selon j' et nous noterons par a pour ascendante le mot de paires de brins nommés **non topologiquement signé** qu'il définit.

Soit enfin $L^{|U|}$ la matrice de paires de brins nommés en $|U|$ -ème position dans U . Les paires de brins nommés consécutives constituant la phase descendante du déplacement à droite d'un brin libre O^i sont alors données par l'ensemble suivant :

$$\boxed{\{r(L_{j^i, j^i + C_{i^c}^{v,e} - 1 - j'}^{|U|}); j' = 0, \dots, C_{i^c}^{v,e} - 1\}} \quad (7.2.19)$$

où r est l'application permutant les premier et deuxième éléments de la paire de brins nommés. Nous considérons là aussi que cet ensemble est ordonné selon j' et nous noterons par d pour descendante le mot de paires de brins nommés **non topologiquement signé** qu'il définit.

Propriété 7.2.4 Identification des paires de brins nommés représentatives de la phase ascendante et descendante du mouvement d'un brin libre à gauche

Soit une configuration de brin libre $O^i v O^i$ représentée par $U = (L^k, (.,.)^k)$, $k = 1, \dots, |U|$ la succession des couples formant la sous-tresse étudiée au sens de l'automate représentatif de \mathbb{B}_N . Soit j^i l'indice du rang de O^i qui est constant tout au long du sous-mot v défini par déf-7.2.2. Soit i^c l'indice du brin (lié au brin libre) dont la classe, au sens de la relation d'équivalence de brins liés engendrée par v , donne les brins que le brin libre doit croiser lors de son déplacement. Cet indice dépend des configurations distinctes de brin libre, nous avons donné les différentes possibilités précédemment dans déf-7.2.10.

Soit L^2 la matrice de paires de brins nommés en 2-ème position dans U . Les paires de brins nommés consécutives constituant la phase descendante²³ du déplacement à droite d'un brin libre O^i sont alors données par l'ensemble suivant :

$$\boxed{\{L_{j^i - 1 - j', j^i - 1}^2; j' = 0, \dots, C_{i^c}^{v,e} - 1\}} \quad (7.2.20)$$

Nous considérons que cet ensemble est ordonné selon j' et nous noterons par d pour descendante le mot de paires de brins nommés **non topologiquement signé** qu'il définit.

Soit enfin $L^{|U|}$ la matrice de paires de brins nommés en $|U|$ -ème position dans U . Les paires de brins nommés consécutives constituant la phase ascendante²⁴ du déplacement à gauche d'un brin libre O^i sont alors données par l'ensemble suivant :

$$\boxed{\{r(L_{j^i - C_{i^c}^{v,e} + j', j^i - 1}^{|U|}); j' = 0, \dots, C_{i^c}^{v,e} - 1\}} \quad (7.2.21)$$

où r est l'application permutant les premier et deuxième éléments d'une paire de brins nommés. Nous considérons là aussi que cet ensemble est ordonné selon j' et nous noterons par a pour ascendante le mot de paires de brins nommés **non topologiquement signé** qu'il définit.

Sur notre exemple qui correspond au cas d'un mouvement à droite et donc à la propriété prop-7.2.3, nous avons $j^i = j^2 = 1$, $i^c = 1$ et $C_{i^c}^{v,e} = C_1^{v,e} = 3$ ce qui donne ensemble :

$$a = L_{11}^2 L_{12}^2 L_{13}^2 = (O^2, O^1)(O^2, O^3)(O^2, O^4) \quad (7.2.22)$$

où nous rappelons que nous sous-entendons l'application ϕ qui à un élément $L_{jj'} = [.,.]$ fait correspondre le croisement générateur $(.,.)$.

23. au sens où le rang du brin commence par descendre lors d'un déplacement à gauche

24. au sens où le rang du brin finit par augmenter lors d'un déplacement à gauche

Sur ce même exemple, nous avons par ailleurs le mot d non signé qui vaut :

$$d = r(L_{13}^{[U]})r(L_{12}^{[U]})r(L_{11}^{[U]}) = (O^3, O^2)(O^1, O^2)(O^4, O^2) \quad (7.2.23)$$

Nous précisons que l'application r est nécessaire lors de la phase descendante, puisque, après l'ajout des paires de brins nommés correspondant à la phase ascendante et définissant le mot a , le brin libre, dans le cas d'un déplacement à droite, est de rang supérieur aux brins qu'il vient de croiser. Lors de la phase descendante, nous utilisons la matrice $L^{[U]}$ qui nous donne les paires de brins nommés correspondantes, mais sans tenir compte de la phase ascendante. De ce fait, nous avons avec $L^{[U]}$, les bonnes paires de brins nommés mais avec un sens inverse et par conséquent, avec une mauvaise permutation entre les brins de chaque paire de brins nommés qu'il faut corriger.

Ces propos illustrent l'image du type "va et vient" que nous attribuons à nos mouvements de brins libres au vu des matrices de paires de brins nommés. Nous constatons par ailleurs, que le **mouvement à droite d'un brin libre est de type "va et vient horizontal" alors que le mouvement à gauche d'un brin libre est de type "va et vient vertical"**. Plus concrètement, nous illustrons dans le cas de notre exemple ces deux phases de "va et vient horizontal".

Pour la phase ascendante et donc en amont de v , il s'agit de rajouter le mot a qui présente une succession de paires de brins nommés et juxtaposés²⁵ suivant données par L^2 de U :

$$L^2 = \left(\begin{array}{cccccc} [O^2, O^1](\mathbf{1}) & \rightarrow & [O^2, O^3](\mathbf{2}) & \rightarrow & [O^2, O^4](\mathbf{3}) & \begin{array}{c} [O^2, O^5] \\ [O^2, O^6] \end{array} \\ & & [O^1, O^3] & & [O^1, O^4] & \begin{array}{c} [O^1, O^5] \\ [O^1, O^6] \end{array} \\ & & & & [O^3, O^4] & \begin{array}{c} [O^3, O^5] \\ [O^3, O^6] \end{array} \\ & & & & & \begin{array}{c} [O^4, O^5] \\ [O^4, O^6] \\ [O^5, O^6] \end{array} \end{array} \right)$$

Pour la phase descendante et donc en aval de v , il s'agit de rajouter le mot d . Ces paires de brins nommés sont données par $L^{[U]}$ de U :

$$L^{[U]} = \left(\begin{array}{cccccc} r([O^2, O^4])(\mathbf{6}) & \leftarrow & r([O^2, O^1])(\mathbf{5}) & \leftarrow & r([O^2, O^3])(\mathbf{4}) & \begin{array}{c} [O^2, O^6] \\ [O^2, O^5] \end{array} \\ & & [O^4, O^1] & & [O^4, O^3] & \begin{array}{c} [O^4, O^6] \\ [O^4, O^5] \end{array} \\ & & & & [O^1, O^3] & \begin{array}{c} [O^1, O^6] \\ [O^1, O^5] \end{array} \\ & & & & & \begin{array}{c} [O^3, O^6] \\ [O^3, O^5] \\ [O^6, O^5] \end{array} \end{array} \right)$$

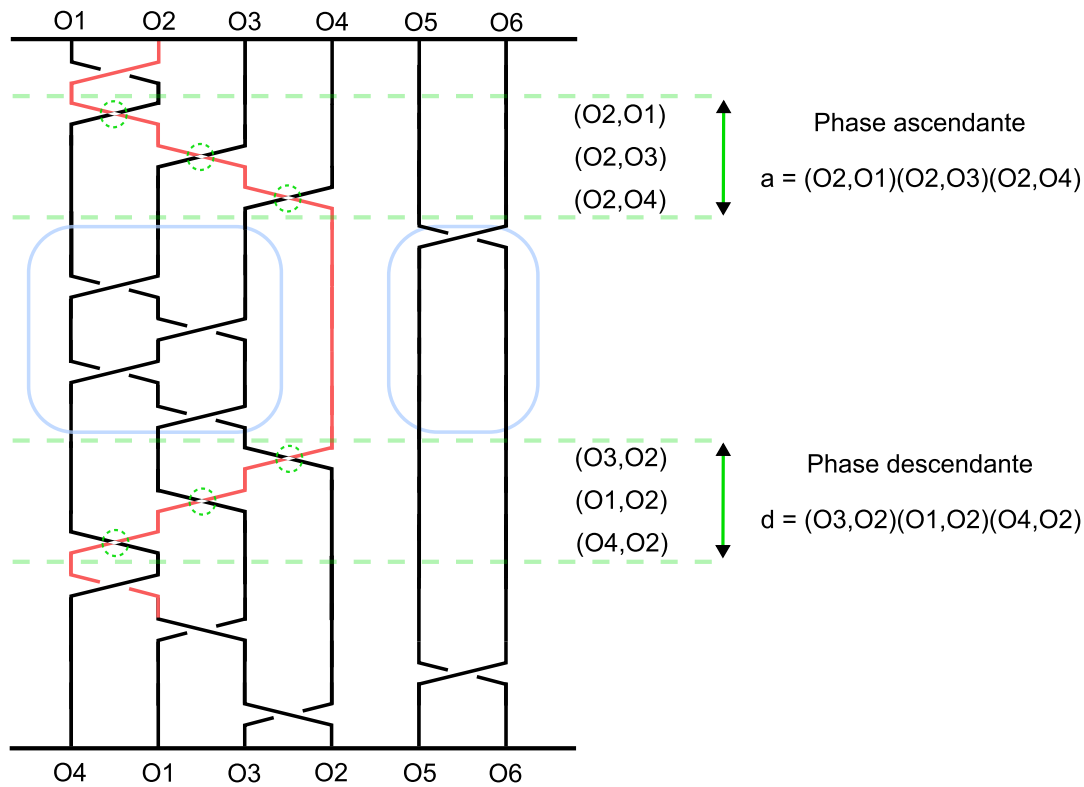
Nous précisons qu'un intérêt de raisonner sur les brins et non sur les rangs est que dans notre cas, les paires de brins nommés et juxtaposés constituant v ne sont pas modifiées lorsque nous effectuons un mouvement. Si nous raisonnions sur les croisements générateurs classiques σ^j , il aurait fallu modéliser les translations des indices des rangs des générateurs.

Nous représentons graphiquement sur la figure fig-7.9, le mouvement à droite du brin libre O^2 de notre exemple qui met donc en évidence, deux phases distinctes : ascendante et descendante.

Nous avons montré ici comment à partir de notre modélisation des tresses par des paires de brins nommés et juxtaposés, nous pouvons définir des relations binaires sur les brins de la tresse étudiée. Ces relations binaires, représentées par des matrices relationnelles, peuvent nous aider à formaliser des informations utiles pour la modélisation de mouvements de brin.

Nous avons en particulier, montré dans ce premier paragraphe, comment la notion de relation d'équivalence de brins liés, engendrée par un mot de paires de brins nommés v issue d'une configuration de brin libre, permettrait de séparer les brins en classes disjointes afin de détecter

25. au vu de la matrice L^2 ces brins ne sont pas juxtaposés mais ils le deviennent progressivement si nous les appliquons consécutivement dans l'ordre énoncé


 FIGURE 7.9 – Exemple de mouvement à droite d'un brin libre O^2vO^2

l'ensemble de brins liés juxtaposés au brin libre qu'il faut "traverser" pour réaliser un mouvement de ce dernier. Ceci se traduit, dans notre modélisation et pour le cas de notre exemple particulier, par un ajout de paires de brins nommés et juxtaposés données pour la partie en amont de v , par la matrice L^2 de U , et par la matrice $L^{|U|}$ pour la partie en aval de v .

Nous n'avons pas en revanche, étudié pour l'instant la nature topologique des paires de brins nommés et juxtaposés utilisées pour modéliser ces deux phases : est-ce que le brin libre qui croise un brin à "traverser" doit être au-dessus ou au-dessous de celui-ci ? Comme nous le verrons dans le paragraphe suivant, il existe en effet, plusieurs façons pour un brin libre de "traverser" un ensemble de brins appartenant au support de v . Cette fois-ci encore, nous définissons une relation binaire sur les brins qui nous permet de formaliser ces possibilités distinctes. Nous utiliserons pour cela, la notion de relation d'ordre "topologique" sur des brins engendrés par un mot de paires de brins nommés.

7.2.3.2 Détermination de la nature topologique des paires de brins nommés juxtaposés utilisées lors du déplacement d'un brin libre

Nous précisons maintenant la nature topologique des paires de brins nommés et juxtaposés de a et de d . Nous continuons notre approche qui consiste à illustrer la méthode en appliquant les différentes définitions et les différentes étapes à notre exemple de la figure fig-7.8.

Nous sommes donc au stade où nous avons identifié les paires de brins nommés permettant de matérialiser le déplacement du brin libre O^2 à droite des brins constituant la classe des brins liés, qui lui est juxtaposée.

Nous avons vérifié que les brins que O^2 croisait lors de sa phase ascendante était naturellement, les mêmes qu'il croisait lors de la phase descendante. Il s'agit dans le cas de notre exemple des brins O^1, O^3, O^4 .

Nous pouvons commencer par constater, que **les croisements ascendant et descendant, d'un brin $O^{i'}$ par le brin libre O^i sont "topologiquement symétriques" au sens suivant : si lors de la phase ascendante, le brin libre O^i est "au-dessous" (resp. "au-dessus") de $O^{i'}$ alors nécessairement, lors de la phase descendante, le brin libre O^i devra être également "au-dessous" (resp. "au-dessus") du brin $O^{i'}$.**

En effet, si nous supposons le contraire, cela conduirait à engendrer un double croisement entre O^i et $O^{i'}$. Or ceci changerait la nature de la tresse et impliquerait que notre mouvement n'est pas isotopique.

Nous avons en fait la propriété générale suivante :

Propriété 7.2.5 Relation topologique symétrique entre les croisements du brin libre O^i avec un brin $O^{i'}$ entre la phase ascendante et descendante lors d'un mouvement isotopique

Soit une configuration de brin libre $O^i \nu O^{i'}$. Soit, $O^{i'}$ appartenant à $\mathbb{C}\mathbb{E}(O^{i'c} | C^{\nu, \epsilon})$ qui est l'ensemble des brins que le brin libre doit "traverser" lors de son mouvement. Nous avons les relations suivantes selon les différentes configurations :

- Cas d'un mouvement à droite :

Phase ascendante (1) \rightarrow Phase descendante (2)		
Si	$\underline{(O^i, O^{i'})}$	alors $\overline{(O^{i'}, O^i)}$
Si	$\overline{(O^i, O^{i'})}$	alors $\underline{(O^{i'}, O^i)}$

(7.2.24)

- Cas d'un mouvement à gauche

Phase descendante (1) \rightarrow Phase ascendante (2)		
Si	$\underline{(O^{i'}, O^i)}$	alors $\overline{(O^i, O^{i'})}$
Si	$\overline{(O^{i'}, O^i)}$	alors $\underline{(O^i, O^{i'})}$

(7.2.25)

Cette propriété met en évidence deux types de "symétrie" concernant les paires de brins nommés ajoutées en amont et en aval de ν pour la modélisation des mouvements isotopiques du brin libre :

- une symétrie de type logique qui concerne l'ordre de permutation engendrée sur les tresses :
 - dans le cas d'un mouvement à droite, le brin libre est avant²⁶ les brins qu'il doit croiser lors de la phase ascendante et il est après²⁷ ces mêmes brins qu'il doit re-croiser lors de la phase descendante. Ainsi si dans a , nous avons $(O^i, O^{i'})$ alors dans d , nous aurons $(O^{i'}, O^i)$.
 - dans le cas d'un mouvement à gauche, le brin libre est après les brins qu'il doit croiser lors de la phase descendante et il est avant ces mêmes brins qu'il doit re-croiser

26. ie de rang inférieur
27. ie de rang supérieur

lors de la phase ascendante. Ainsi si dans d , nous avons $(O^{i'}, O^i)$ alors dans a , nous aurons $(O^i, O^{i'})$.

- une symétrie de type topologique qui est formellement liée à la symétrie précédente : lorsque lors de la première phase, la paire de brins nommés est de type $(O^i, O^{i'})$ alors cela implique que nous ayons $(O^{i'}, O^i)$ lors de la deuxième phase, et *vice versa*

Ainsi par définition d'un mouvement isotopique d'un brin libre, il n'y a pas d'"incohérence" topologique²⁸ entre les paires de brins nommés utilisées en amont et en aval de v . Nous pouvons alors, selon la relation binaire topologique définie par déf-7.2.8, considérer uniquement les paires de brins nommés définies soit par a soit par d . En effet, ces deux mots codent la même information topologique sur les paires de brins impliquées.

Sans perte de généralité, nous raisonnerons alors uniquement sur a , le mot de paires de brins nommés représentant la phase ascendante que le mouvement considéré soit à droite ou à gauche. Nous introduisons maintenant la notion de signature topologique d'un mot de tresse (non signé) qui sera fondamentale pour la suite.

Définition 7.2.12 Mot de paires de brins nommés topologiquement signés

Soit $w = (.,.)^1 \dots (.,.)^{|w|}$ un mot de paires de brins nommés non topologiquement signé de cardinal $|w|$ (tel que a et d présentés précédemment). Soit, α un $|w|$ -uplet de $\{-1, 1\}$. Nous définissons $w(\alpha)$ comme étant le mot de paires de brins nommés topologiquement signé (nous dirons aussi signé) issu de w tel que, $\forall k = 1, \dots, |w|$:

$$w(\alpha)_k = \begin{cases} \underline{(.,.)}^k & \text{si } \alpha_k = 1 \\ \overline{(.,.)}^k & \text{si } \alpha_k = -1 \end{cases} \quad (7.2.26)$$

Si nous notons par $a(\alpha)$ une signature²⁹ du mot a alors, nous noterons par $d[a(\alpha)]$ le mot de paires de brins nommés représentant la phase descendante, qui est obtenu à partir de a en appliquant les relations de symétrie topologique³⁰ exprimées par les tableaux de la propriété prop-7.2.5. Nous nous intéressons à présent de savoir quelle nature topologique, doit-on attribuer à chaque paires de brins nommés constituant a .

Il existe ainsi plusieurs façons de signer un mouvement isotopique d'un brin libre. Si nous reprenons notre exemple, il existe en effet, 3 possibilités du même mouvement à droite $\xrightarrow{O^2}$ du brin libre O^2 qui sont représentées sur les figures fig-7.10, fig-7.11, fig-7.12.

Nous rappelons la valeur initiale du mot u relative à la configuration de brin libre $O^2 v O^2$

$$u = \underbrace{(O^1, O^2)(O^5, O^6)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_v (O^2, O^4) \quad (7.2.27)$$

Nous donnons ci-dessous, pour chacun des trois mouvements isotopiques à droite du brin libre O^2 , le mot signé engendré qui est topologiquement équivalent à u .

Nous avons pour la figure fig-7.10, c'est à dire pour $\alpha = (1, 1, 1)$ l'équivalence suivante localement à u :

$$u \equiv \underbrace{(O^1, O^2)}_{a(1,1,1)} \underbrace{[(O^2, O^1)(O^2, O^3)(O^2, O^4)]}_{v} \underbrace{[(O^3, O^2)(O^1, O^2)(O^4, O^2)]}_{d[a(1,1,1)]} (O^2, O^4) \quad (7.2.28)$$

28. ie si O^i est au-dessus (resp. au-dessous) de $O^{i'}$ pour a alors il doit nécessairement être au-dessus (resp. au-dessous) de $O^{i'}$

29. cf ci-dessous

30. ou leurs contraposées

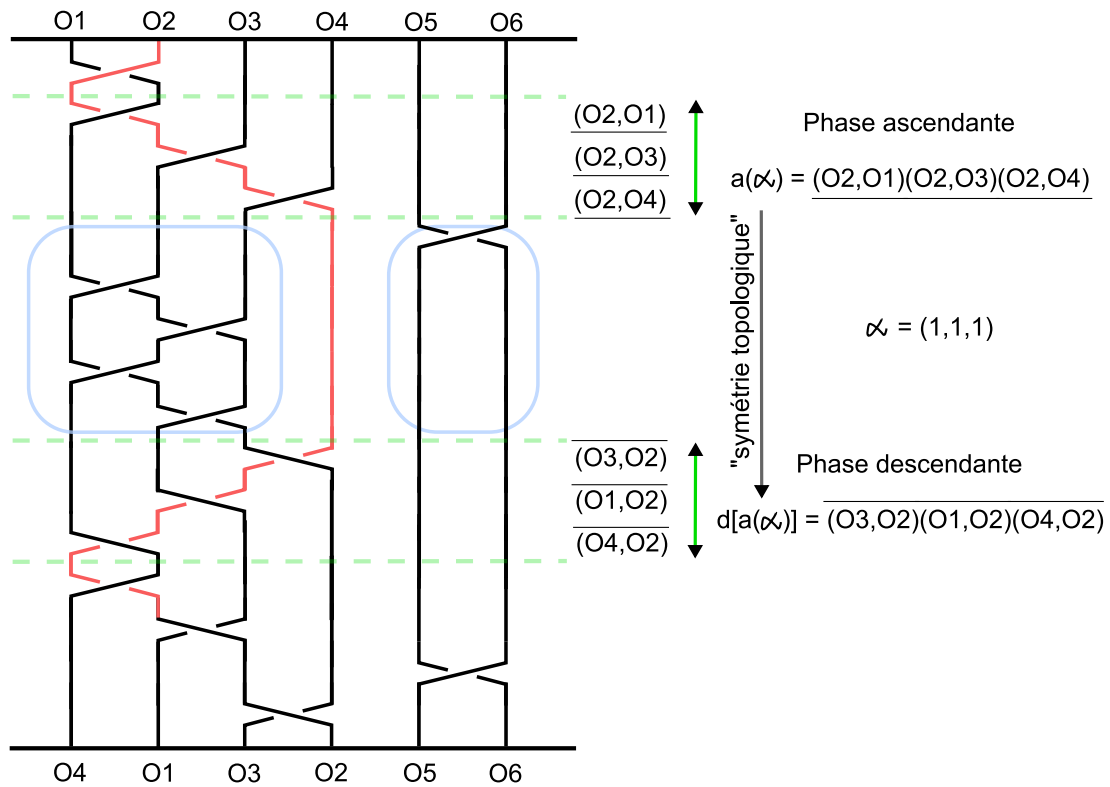


FIGURE 7.10 – Signatures du mouvement isotopique à droite d'un brin libre O^2vO^2 pour $\alpha = (1, 1, 1)$

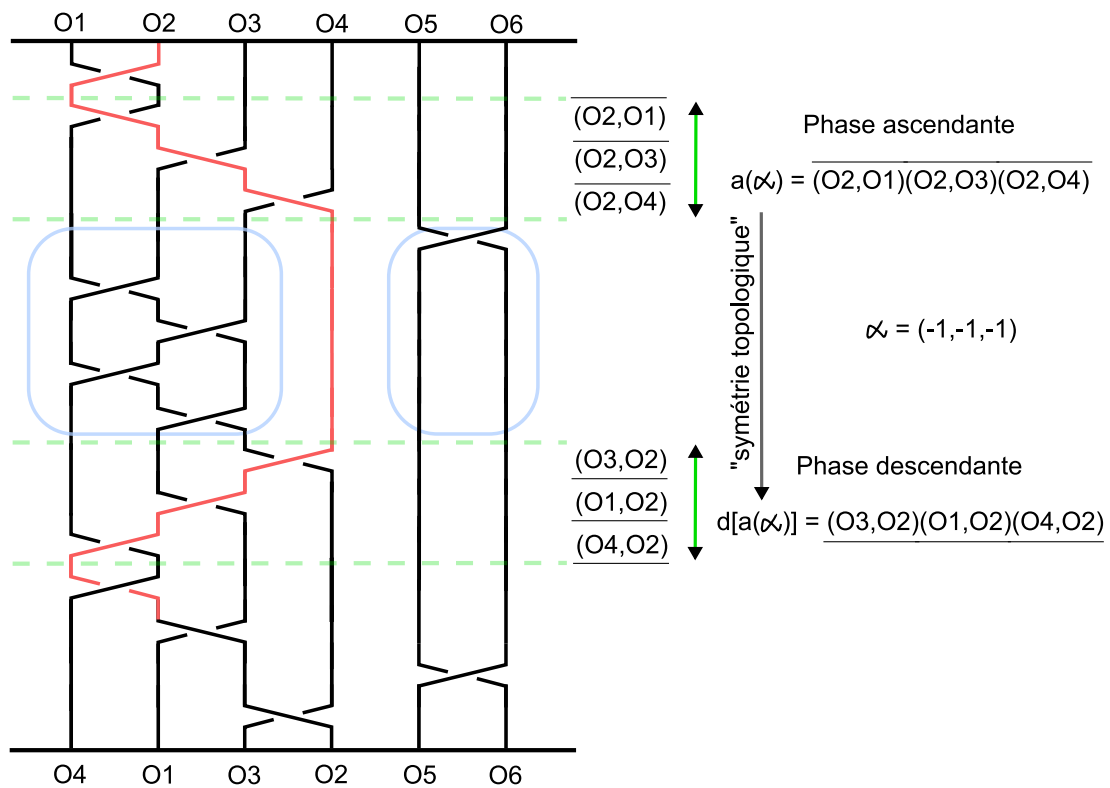


FIGURE 7.11 – Signatures du mouvement isotopique à droite d'un brin libre O^2vO^2 pour $\alpha = (-1, -1, -1)$

Nous avons pour la figure fig-7.11, c'est à dire pour $\alpha = (-1, -1, -1)$ l'équivalence suivante localement à u :

$$u \equiv \underbrace{(O^1, O^2) \left[(O^2, O^1)(O^2, O^3)(O^2, O^4) \right]}_{a(-1, -1, -1)} v \underbrace{\left[(O^3, O^2)(O^1, O^2)(O^4, O^2) \right] (O^2, O^4)}_{d[a(-1, -1, -1)]} \quad (7.2.29)$$

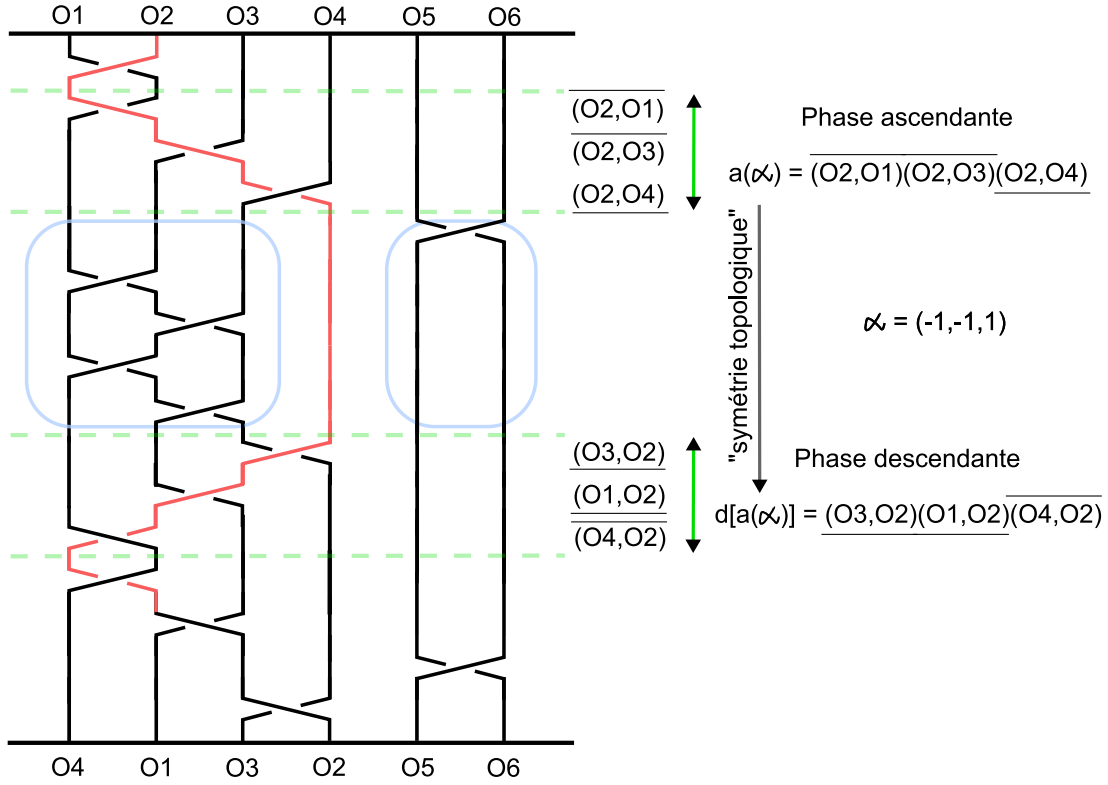


FIGURE 7.12 – Signatures du mouvement isotopique à droite d'un brin libre O^2vO^2 pour $\alpha = (-1, -1, 1)$

Nous avons pour la figure fig-7.12, c'est à dire pour $\alpha = (-1, -1, 1)$ l'équivalence suivante localement à u :

$$u \equiv \underbrace{(O^1, O^2) \left[(O^2, O^1)(O^2, O^3)(O^2, O^4) \right]}_{a(-1, -1, 1)} v \underbrace{\left[(O^3, O^2)(O^1, O^2)(O^4, O^2) \right] (O^2, O^4)}_{d[a(-1, -1, 1)]} \quad (7.2.30)$$

Il n'existe pas d'autre $|a|$ -uplet donnant un mouvement isotopique à droite pour la configuration O^2vO^2 . Nous donnons dans la figure fig-7.13, un quatrième cas de figure correspondant à $\alpha = (-1, 1, -1)$. Le lecteur pourra alors vérifier que la tresse engendrée n'est plus topologiquement équivalente à la tresse d'origine.

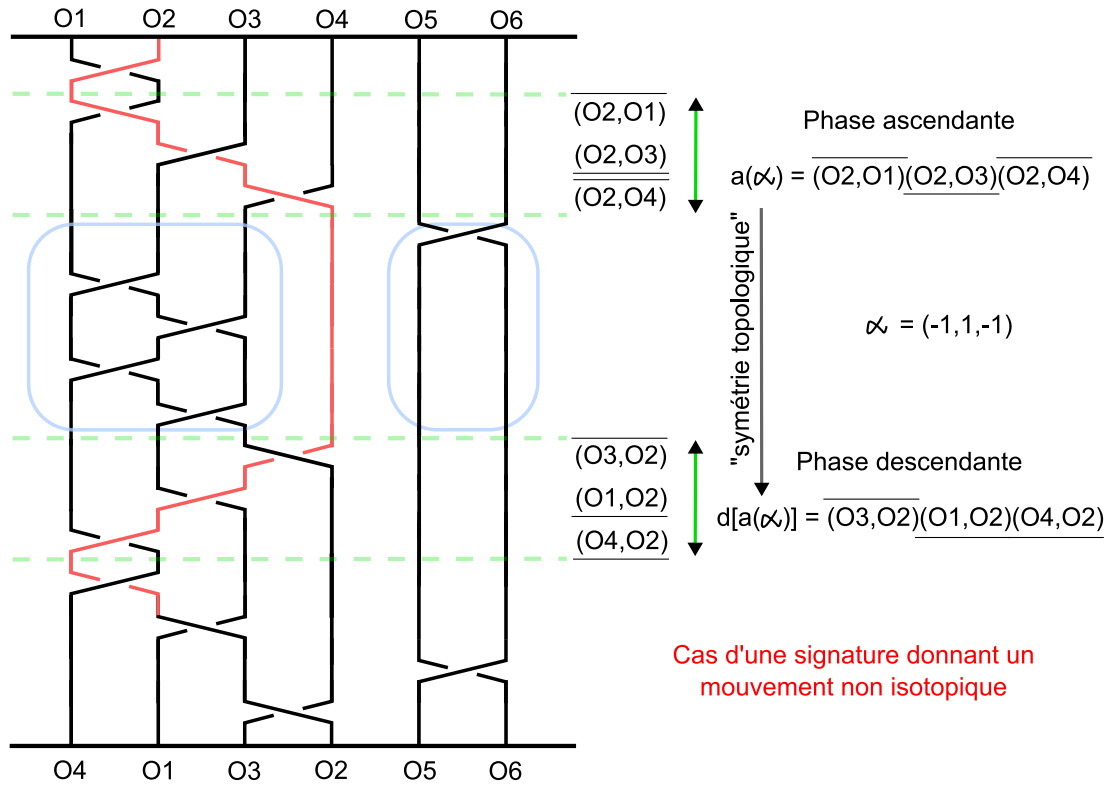


FIGURE 7.13 – Signatures du mouvement à droite d'un brin libre O^2vO^2 pour $\alpha = (-1, 1, -1)$ qui donne un mouvement non isotopique

Nous avons ainsi pour la figure fig-7.13, c'est à dire pour $\alpha = (-1, 1, -1)$ la non équivalence :

$$u \neq \underbrace{(O^1, O^2) \left[(O^2, O^1)(O^2, O^3)(O^2, O^4) \right]}_{a(-1,1,-1)} v \underbrace{\left[(O^3, O^2)(O^1, O^2)(O^4, O^2) \right]}_{d[a(-1,1,-1)]} (O^2, O^4) \quad (7.2.31)$$

Signer un mot de paires de brins nommés a représentant la phase ascendante d'un mouvement de brin libre, revient donc à lui appliquer un $|a|$ -uplet de $\{-1, 1\}$ qui est une combinaison parmi les $2^{|a|}$ possibilités.

Parmi ces différentes combinaisons, **il existe un sous-ensemble d'uplets qui traduit un mouvement isotopique c'est à dire qui ne change pas la nature de la tresse.**

Comment déterminer ce sous-ensemble ?

Pour répondre à cette question, nous définissons à présent, une autre relation binaire déterminée à partir d'un mot de paires de brins nommés. Celle-ci est proche de la définition d'un ordre ce qui nous conduit à l'appeler **ordre "topologique" engendré par un mot sur les brins de la tresse.**

Définition 7.2.13 *Relation d'ordre "topologique" engendrée par un mot de paires de brins nommés*
 Soit w un mot de paires de brins nommés signé. Nous appellerons relation d'ordre topologique engendrée par w , représentée par une matrice relationnelle $C^{w,\circ}$, la fermeture transitive de la relation binaire

topologique représentée par C^w et définie en déf-7.2.8. Nous avons alors :

$$\boxed{C^{w,o} = C^w \oplus (C^w)^2 \oplus \dots \oplus (C^w)^{N-1}} \quad (7.2.32)$$

où \oplus est l'union relationnelle, et où les puissances sont prises au sens des multiplications relatives \odot .

A l'aide de la notion d'ordre topologique, nous pouvons donner les conditions pour qu'une signature d'un mot de paires de brins nommés représentant le déplacement ascendant d'un brin libre représente un mouvement véritablement isotopique.

Théorème 7.2.1 Conditions pour qu'une signature d'un mot de paires de brins nommés définisse un mouvement isotopique

Soit $O^i v O^i$ une configuration de brin libre conduisant à un mouvement de O^i devant "traverser" tout ou partie des brins appartenant au support de v . Soit a le mot de paires de brins nommés non signé représentant la phase ascendante du mouvement considéré. Soit α un $|a|$ -uplet de $\{-1, 1\}$ et soit $a(\alpha)$ le mot issu de a qui est topologiquement signé selon α .

Soient alors $C^{v,o}$ et $C^{a(\alpha),o}$ les matrices relationnelles représentatives des ordres topologiques engendrés respectivement par v et $a(\alpha)$. Alors la signature α définit un mouvement **isotopique** si et seulement s'il n'y a pas de discordance entre les ordres topologiques engendrés par v et $a(\alpha)$. En termes de matrices relationnelles, nous avons la propriété suivante :

$$\boxed{a(\alpha) \text{ définit un mouvement isotopique} \Leftrightarrow \sum_{i=1}^N \sum_{i'=1}^N C_{ii'}^{v,o} \check{C}_{ii'}^{a,o} - N = 0} \quad (7.2.33)$$

où $\check{C}_{ii'} = C_{i'i}$ et où N est le nombre de brins de la tresse qui représente aussi le nombre de discordances non significatives dues à la propriété de réflexivité³¹ qui attribue la valeur 1 à tous les termes de la diagonale. L'expression précédente est équivalente à la suivante :

$$\boxed{a(\alpha) \text{ définit un mouvement isotopique} \Leftrightarrow C^{v,o} \otimes \check{C}^{a,o} = I_N} \quad (7.2.34)$$

où I_N est la matrice identité et où \otimes est l'intersection relationnelle. Nous précisons aussi, que ces formules sont symétriques au sens où on peut remplacer $C^{v,o} \otimes \check{C}^{a,o}$ par $\check{C}^{v,o} \otimes C^{a,o}$.

Démonstration

Nous voyons que le déplacement d'un brin libre $O^i v O^i$ est isotopique, si le brin O^i peut "traverser" les brins de v sans qu'il y ait de "coupures et de re-collages" de ces brins. En d'autres termes, il faut qu'il y ait un "espace" entre les brins de v pour que le brin O^i puisse se "faufiler" entre eux, à droite ou à gauche selon le mouvement considéré. Cette notion d'"espace" est exprimée par la notion d'ordre topologique "au-dessous" existant entre les brins. En effet, nous avons constaté que pour un mot de paires de brins nommés signé topologiquement, il existait un ordre d'"empilement" pour lequel il peut y avoir³² un brin d'indice de rang 1³³ qui est celui qui ne passe "au-dessus" d'aucun autre brin et un brin de rang maximal³⁴ qui est celui qui passe "au-dessus" de tous les autres brins. Par ailleurs, le cas d'un double croisement entre deux brins $O^{i'}$ et $O^{i''}$ tel que $(O^{i'}, O^{i''})(O^{i''}, O^{i'})$, implique qu'il n'y a pas d'espace entre ces deux brins et qu'il faut donc soit passer "au-dessus" de ce double croisement soit "au-dessous". Signer par α un mot a représentant la phase ascendante d'un mouvement, c'est définir la nature des croisements de O^i , le brin libre, avec les autres brins de v . En terme de matrices relationnelles de la

31. que nous avons admise pour les relations binaires topologiques cf déf-7.2.8

32. ce n'est pas toujours le cas notamment lorsqu'il y a des doubles croisements

33. ie le "plus au-dessous"

34. ie le "moins au-dessous"

relation binaire topologique engendrée par $a(\alpha)$, celui-ci revient à remplir de manière asymétrique³⁵ la i -ème ligne et la i -ème colonne de $C^{a(\alpha)}$. La fermeture transitive de $C^{a(\alpha)}$ notée $C^{a(\alpha),o}$ engendre un ordre asymétrique sur les brins de v relativement à O^i . Nous constatons alors que O^i se "faufile" entre les brins de v si et seulement si l'ordre topologique engendré par $a(\alpha)$ sur les brins de v respecte l'ordre topologique engendré par v .

Dit autrement, l'"espace" nécessaire à un brin libre pour pouvoir "traverser" un groupe de brins liés correspond en quelque sorte à l'attribution d'un rang à ce brin libre qui doit s'intégrer dans l'ordre topologique existant entre les brins de v et non pas le remettre en cause.

□

En d'autres termes, le théorème précédent indique, qu'un mouvement isotopique est valide si le mot de brin signé, $a(\alpha)$, engendre un ordre topologique qui ne remet pas en cause celui engendré par v . Ainsi, être en discordance avec l'ordre topologique engendré par v revient à modifier la nature d'une tresse lors du mouvement du brin libre qui n'est alors plus isotopique.

Ce théorème montre finalement le deuxième intérêt de notre modélisation des tresses par des matrices de paires de brins nommés. En effet, il n'existe pas, à notre connaissance, de méthode ou de formalisme permettant de modéliser de manière générique le déplacement isotopique d'un brin.

Nous illustrons ce théorème fondamental dans ce qui suit au travers de notre exemple. Dans celui-ci, nous rappelons que nous avons $v = \overline{(O^5, O^6)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}$. La matrice relationnelle représentative de l'ordre topologique engendré par v est alors la suivante :

$$C^{v,o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.35)$$

Nous sommes dans le cas simple et particulier, où C^v est déjà transitive ce qui implique $C^v = C^{v,o}$

Nous commençons par étudier les ordres topologiques des mouvements isotopiques valides représentées sur les figures fig-7.10 et fig-7.11. Nous montrons que ces mouvements isotopiques sont particuliers puisqu'ils sont en effet toujours valides.

Dans le cas de la figure fig-7.10, nous avons $\alpha = (1, 1, 1)$ ce qui donne $\overline{(O^2, O^1)(O^2, O^3)(O^2, O^4)}$. La matrice relationnelle de $a(1, 1, 1)$ est alors la suivante :

$$C^{a(1,1,1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.36)$$

Dans le cas de la figure fig-7.11, nous avons $\alpha = (-1, -1, -1)$ ce qui donne $\overline{(O^2, O^1)(O^2, O^3)(O^2, O^4)}$.

35. en effet, si $C_{ii'}^{a(\alpha)} = 1$ alors $C_{i'i}^{v,a(\alpha)} = 0$ en raison de la "symétrie" topologique présentée précédemment

La matrice relationnelle de $a(-1, -1, -1)$ est alors la suivante :

$$C^{a(-1,-1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.37)$$

Nous pouvons vérifier que nous avons bien $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(1,1,1),o} - 6 = \sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(-1,-1,-1),o} - 6 = 0$ et que les termes unitaires additionnés correspondent aux éléments diagonaux qui sont au nombre de $N = 6$. Nous avons en fait la propriété suivante caractéristique de ces mouvements isotopiques particuliers qui sont toujours valides quelque soit v .

Propriété 7.2.6 *Mouvements isotopiques “au-dessous” et “au-dessus”*

Le cas particulier des $|a|$ -uplets qui valent soit $(1, 1, \dots, 1)$ soit $(-1, -1, \dots, -1)$ engendrent toujours des mouvements isotopiques valides. Lorsque nous avons $\alpha = (1, 1, \dots, 1)$, nous parlerons d'un mouvement isotopique “au-dessous” étant donné que le brin libre passe au-dessous de tous les brins qu'il doit croiser ce qui ne donne aucune discordance avec l'ordre topologique des brins croisés. De manière symétrique, lorsque nous avons $\alpha = (-1, -1, \dots, -1)$, nous parlerons d'un mouvement isotopique “au-dessus”.

Nous vérifions à présent, que le cas de la figure fig-7.12 pour lequel nous avons $\alpha = (-1, -1, 1)$ respecte bien, lui aussi, le théorème théo-7.2.1. Nous avons dans ce cas $a(-1, -1, 1) = \overline{(O^2, O^1)}(O^2, O^3)(O^2, O^4)$ et l'ordre topologique correspondant :

$$C^{a(-1,-1,1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.38)$$

où nous avons mis en gras, les éléments unitaires issus de la fermeture transitive.

A nouveau, nous pouvons vérifier, que nous avons $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(-1,-1,1),o} - 6 = 0$

Contrairement, aux possibilités précédentes, nous avons présenté le cas de la figure fig-7.13 comme étant un mouvement non isotopique qui modifiait la nature ou la topologie de la tresse étudiée. Nous vérifions que ce cas, ne vérifie pas en effet, le théorème théo-7.2.1. Nous avons $a(-1, 1, -1) = \overline{(O^2, O^1)}(O^2, O^3)\overline{(O^2, O^4)}$ et l'ordre topologique correspondant :

$$C^{a(-1,1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.39)$$

Nous avons alors dans ce cas :

$$C^{v,o} \otimes \check{C}^{a(-1,1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{1} & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.40)$$

où nous avons mis ici en gras, les éléments unitaires en discordances. Nous avons par ailleurs, $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(-1,1,-1),o} - 6 = 2$. Ces éléments s'interprètent de la manière suivante : pour v , nous avons O^3 qui est "au-dessous" de O^1 et O^4 alors que l'ordre topologique engendré par $a(-1,1,-1)$ implique au contraire que O^3 est "au-dessus" de O^1 et O^4 . Ces discordances indiquent donc que le mouvement engendré par $a(-1,1,-1)$ change la nature topologique de la tresse ce qui l'invalidé comme mouvement isotopique.

Pour étayer le théorème théo-7.2.1, nous montrons de la même façon, l'invalidité des 4 autres possibilités restantes en rappelant que nous avons initialement $2^3 = 8$ possibilités.

Dans le cas $\alpha = (-1,1,1)$, nous avons $a(-1,1,1) = \overline{(O^2, O^1)}(O^2, O^3)(O^2, O^4)$ et l'ordre topologique correspondant :

$$C^{a(-1,1,1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 1 & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.41)$$

Nous avons alors dans ce cas l'intersection relationnelle suivante qui donne un nombre de discordances égal à $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(-1,1,1),o} - 6 = 1$:

$$C^{v,o} \otimes \check{C}^{a(-1,1,1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ O^3 & \mathbf{1} & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.42)$$

Dans le cas $\alpha = (1,-1,-1)$, nous avons $a(1,-1,-1) = \overline{(O^2, O^1)}\overline{(O^2, O^3)}\overline{(O^2, O^4)}$ et l'ordre topologique correspondant :

$$C^{a(1,-1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & \mathbf{1} & 0 & 0 & 0 \\ \mathbf{1} & \mathbf{1} & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.43)$$

Nous avons alors dans ce cas l'intersection relationnelle suivante qui donne un nombre de

discordances égal à $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(1,-1,-1),o} - 6 = 2$:

$$C^{v,o} \otimes \check{C}^{a(1,-1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & \mathbf{1} & \mathbf{1} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.44)$$

Dans le cas $\alpha = (1, -1, 1)$, nous avons $a(1, -1, 1) = \underline{(O^2, O^1)} \overline{(O^2, O^3)} \underline{(O^2, O^4)}$ et l'ordre topologique correspondant :

$$C^{a(1,-1,1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ \mathbf{1} & 1 & 1 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.45)$$

Nous avons alors dans ce cas l'intersection relationnelle suivante qui donne un nombre de discordances égal à $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(1,-1,1),o} - 6 = 1$:

$$C^{v,o} \otimes \check{C}^{a(1,-1,1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & \mathbf{1} & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.46)$$

Enfin, dans le cas $\alpha = (1, 1, -1)$, nous avons $a(1, 1, -1) = \underline{(O^2, O^1)} \underline{(O^2, O^3)} \overline{(O^2, O^4)}$ et l'ordre topologique correspondant :

$$C^{a(1,1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ \mathbf{1} & 1 & \mathbf{1} & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.47)$$

Nous avons alors dans ce cas l'intersection relationnelle suivante qui donne un nombre de discordances égal à $\sum_{i,i'} C_{ii'}^{v,o} \check{C}_{ii'}^{a(1,1,-1),o} - 6 = 2$:

$$C^{v,o} \otimes \check{C}^{a(1,1,-1),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 & O^6 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \\ O^6 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & \mathbf{1} & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & \mathbf{1} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.2.48)$$

7.3 Les problèmes du mot du groupe des tresses, du mot de tresse minimal et du mot de tresse “shortlex” dans le cadre de la modélisation par paires de brins nommés

Nous avons donné précédemment une nouvelle modélisation des mots de tresses basée sur un nouveau codage qui exploite les paires de brins nommés $(.,.)$ et leurs matrices L associées représentant l'ordre strict et total (ou la permutation) des brins.

Cette nouvelle modélisation, nous permet de définir des relations binaires d'équivalence dites de brins liés et d'ordre topologique qui permettent, de modéliser d'une manière originale le mouvement isotopique d'un brin.

L'objectif de cette section est de proposer des approches permettant d'aborder les problèmes algorithmiques autour des tresses mais qui soient de nature différente des approches entreprises jusqu'à présent.

Après avoir présenté de manière générale ces méthodes, nous les illustrerons sur des exemples pratiques.

7.3.1 Modélisation des problèmes algorithmiques autour des tresses sous la forme de problèmes d'optimisation combinatoire

Les problèmes autour des tresses, sont très couramment abordés sous l'angle des systèmes de réécriture. Il s'agit de définir des règles de type $v \rightarrow v'$ pour lesquelles, un motif v est systématiquement remplacé par un autre motif v' . Ces règles, ne changent pas la nature de la tresse et définissent ainsi des mouvements isotopiques particuliers de la même manière que les réductions de poignées de Dehornoy ou les relations d'Artin. Tout système de règles défini afin de résoudre un type de problème tel que le problème du mot du groupe des tresses, doit être complet. En d'autres termes, il faut montrer que l'application du système de règles converge vers la solution recherchée et qu'il est acyclique de sorte qu'il n'y ait pas de boucle possible.

Notre approche est distincte de ce qui vient d'être décrit dans l'esprit. En effet, **nous souhaitons aborder les problèmes sous l'angle de l'optimisation combinatoire**. Par ailleurs, **les trois problèmes cités peuvent être vus comme étant des sous-problèmes imbriqués les uns dans les autres dont le problème du mot du groupe des tresses serait le sous-problème le “plus simple” à résoudre et le problème du mot de tresse “shortlex” comme étant celui englobant les deux autres.**

En effet, considérons le cas du problème du mot du groupe des tresses qui consiste à savoir si deux mots de tresse v et v' par exemple représentent ou non la même tresse. Nous avons vu que celui-ci pouvait se ramener au problème qui consiste à savoir si $v \circ (v')^{-1}$ est trivial.

Prenons désormais le cas général d'un mot de tresse w , qui pourra donc être égal dans le cas du problème du mot du groupe des tresses à $v \circ (v')^{-1}$.

Dans notre approche très intuitive, nous poursuivons l'idée que si nous arrivons à définir un procédé permettant de démêler une tresse, au sens de la détection de réductions libres, alors nous pouvons chercher à résoudre les problèmes cités précédemment.

Selon ce procédé un “mot de tresse démêlé” devra se réduire au mot vide si la tresse correspondante est triviale. Si la tresse n'est pas triviale, le procédé devra alors donner une tresse “optimalement simplifiée” au sens où le mot de tresse obtenu devra être de taille minimale. Dans l'hypothèse où le procédé arrive effectivement à obtenir un mot de tresse minimal alors,

le problème de la détermination de la forme canonique “shortlex” consistera à déplacer les brins³⁶ de sorte à avoir le plus petit ordre lexicographique du mot qui est relatif à l’ordre initial sur les générateurs³⁷.

Autrement dit, ce que nous cherchons dans l’ensemble de ces trois problèmes, c’est tout d’abord d’optimiser la taille du mot de tresse. Nous cherchons plus particulièrement à minimiser celle-ci et nous proposons d’utiliser la modélisation précédemment exposée pour y arriver. Plus précisément, l’idée générale du procédé que nous proposons est d’effectuer à chaque étape un mouvement de brin libre permettant de simplifier la tresse en mettant en évidence des réductions libres. L’approche est nuancée vis à vis des systèmes de réécriture. En effet, dans ces derniers, la taille du mot peut augmenter, il faut alors démontrer que le système est *in fine* complet³⁸. **Dans une vision “optimisation” du problème, la taille du mot étant le critère à minimiser, nous n’effectuons aucun déplacement de brin libre qui engendrerait une dégradation stricte du critère. Il faut *a contrario* démontrer formellement que le procédé que nous présentons atteint l’“optimal global”.**

Nous précisons ci-dessous, certains points vis à vis de ces propos.

Tout d’abord il est nécessaire de favoriser l’application de mouvements de brins libres permettant de réduire la taille d’un mot de tresse. En effet, nous avons montré comment nous pouvons détecter et effectuer de manière générique des mouvements isotopiques de brins. Parmi ces derniers, tous ne permettent pas d’effectuer des réductions libres. Dans un premier temps, il est facile de voir que **les mouvements de brin permettant de réduire la taille d’un mot se ramènent à des déplacement de brins libres de type “poignées”³⁹**. Dans notre procédé, nous favoriserons ainsi la recherche de mouvements de brins libres de ce type ci.

Nous introduisons de plus, ci-dessous la notion d’“épuration” d’un sous-mot u intervenant dans la définition d’un brin libre $O^i v O^i$ qui permet soit de mettre en évidence plus aisément des réductions libres soit de limiter la génération de croisements “superflus” lors d’un déplacement isotopique d’un brin libre.

Définition 7.3.1 “Épuration” d’un sous-mot u issu d’une configuration de brin libre $O^i v O^i$ et/ou de son sous-mot associé v

Soit une configuration de brin libre $O^i v O^i$ représenté par un mot de paires de brins nommés u . Soit $V(u)$ le voisinage en amont et en aval de u , c’est à dire l’ensemble des paires de brins nommés qui sont avant et après u . Lorsque nous déplaçons un brin libre O^i dans le contexte de la recherche de réductions libres, il est très utile d’effectuer les opérations suivantes avant d’évaluer le nombre de réductions libres ou le nombre de croisements “superflus” ajoutés par ce déplacement et susceptible d’augmenter de manière inutile la taille du mot de tresse après déplacement :

1. pour favoriser la mise en évidence de réductions libres, il est nécessaire de rapprocher de u les paires de brins nommés appartenant à $V(u)$ qui font intervenir le brin libre O^i en utilisant les relations d’Artin R_2 et R_3 . En effet, comme nous déplaçons le brin libre O^i , il y a potentiellement en amont et en aval de u des réductions libres entre O^i et un autre brin quelconque que ce déplacement peut engendrer
2. pour éviter les croisements “superflus”, il est nécessaire de réduire au maximum la taille du sous-mot v . Pour ce faire, on utilise les mouvements d’Artin R_2 et R_3 afin de faire “sortir” de v des paires de brins nommés “superflus”. Ces paires de brins nommés font intervenir des brins qui ne sont pas liés (au sens de la partition des brins liés) à la classe des brins qui est juxtaposée au brin libre. Vis à vis de la remarque [rq-7.2.2 page 238](#), cela revient à privilégier des mouvements “convex” par rapport à des mouvement “coarse”. Pour ce deuxième cas, nous parlerons plutôt de

36. il faut toutefois démontrer que la classe des mots de tresse minimales d’une tresse β peut être parcourue sans complexifier la tresse, c’est à dire sans augmenter la taille du mot de tresse

37. rappelons que nous pouvons pour tout couple $(L, (.,.))$ retrouver le codage classique en terme des générateurs σ^j en utilisant la relation (7.2.5)

38. ie convergeant et terminal

39. cf déf-7.2.5

l'épuration du sous-mot v plutôt que de l'épuration de u qui correspondra davantage au premier cas

L'un ou l'autre ou les deux types de pré-traitements pourront être étudiés afin de déterminer l'"épuration" optimisant au mieux le critère.

Dans le cadre d'un procédé de démêlage de tresse, lorsque nous détectons un brin libre $O^i v O^i$, nous chercherons donc avant tout, à épurer le mot u ou son sous-mot associé v . Ceci permet de faire apparaître plus facilement des mouvements de brins libres présentant des réduction libres et de ne pas générer de croisements "superflus" dégradant le critère. Si un mot u est notamment épuré au sens du deuxième point donné précédemment, alors nous avons la propriété suivante qui évite de calculer la matrice relationnelle des brins liés lors d'un mouvement isotopique d'un brin libre comme nous l'avions avancé dans la sous-section précédente :

Propriété 7.3.1 Identification des paires de brins nommés représentatives de la phase ascendante et descendante d'un mouvement d'un brin libre dans le cas de mots $u = O^i v O^i$ et v épurés
 Soit une configuration de brin libre $O^i v O^i$ représentée par $U = (L^k, (\cdot, \cdot)^k)$, $k = 1, \dots, |U|$ où v est un mot épuré au sens de déf-7.3.1. Soit $\#\text{Supp}(v)$ le cardinal du support du mot v . Soit j^i l'indice du rang de O^i qui est constant tout au long du sous-mot v . Soit i^c l'indice du brin défini par déf-7.2.10 et qui dépend du type de mouvement (à droite ou à gauche) que nous souhaitons effectuer.

Soit L^2 et $L^{|U|}$ les matrices de paires de brins nommés en 2-ème et $|U|$ -ème position dans U .

Si le mouvement considéré est un **mouvement à droite** nous avons les relations suivantes :

1. Les paires de brins nommés consécutives constituant le mot a représentatif de la **phase ascendante** du déplacement à droite d'un brin libre O^i sont alors données par :

$$a = L_{j^i, j^i}^2 L_{j^i, j^i+1}^2 \cdots L_{j^i, j^i+\#\text{Supp}(v)-1}^2 \quad (7.3.1)$$

2. Les paires de brins nommés consécutives constituant le mot d représentatif la **phase descendante** du déplacement à droite d'un brin libre O^i sont alors données par :

$$d = r(L_{j^i, j^i+\#\text{Supp}(v)-1}^{|U|}) r(L_{j^i, j^i+\#\text{Supp}(v)-2}^{|U|}) \cdots r(L_{j^i, j^i}^{|U|}) \quad (7.3.2)$$

Si le mouvement considéré est un **mouvement à gauche** nous avons les relations suivantes :

1. Les paires de brins nommés consécutives constituant le mot d représentatif de la **phase descendante** du déplacement à gauche d'un brin libre O^i sont alors données par :

$$d = L_{j^i-1, j^i-1}^2 L_{j^i-2, j^i-1}^2 \cdots L_{j^i-\#\text{Supp}(v), j^i-1}^2 \quad (7.3.3)$$

2. Les paires de brins nommés consécutives constituant le mot a représentatif la **phase ascendante** du déplacement à gauche d'un brin libre O^i sont alors données par :

$$a = r(L_{j^i-\#\text{Supp}(v), j^i-1}^{|U|}) r(L_{j^i-\#\text{Supp}(v)+1, j^i-1}^{|U|}) \cdots r(L_{j^i-1, j^i-1}^{|U|}) \quad (7.3.4)$$

Pour illustrer ces propos ainsi que les principes d'"épuration" de la définition déf-7.3.1, prenons l'exemple donné par la figure fig-7.8 page 236. Nous avons au départ :

$$O^2 v O^2 = u = \underbrace{(O^1, O^2)(O^5, O^6)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_v (O^2, O^4) \quad (7.3.5)$$

Nous pouvons considérer en premier lieu le voisinage de u . Nous remarquons qu'en aval de ce dernier, il y a des paires de brins nommés faisant intervenir le brin libre O^2 . Nous avons en effet :

$$w = u \overline{(O^2, O^1)} \overline{(O^6, O^5)} \overline{(O^2, O^3)} \quad (7.3.6)$$

La dernière paire peut être rapprochée de u en utilisant la relation d'Artin $R2$. Conformément au premier principe d'épuration, nous considérons alors le mot suivant :

$$w \equiv u \overline{(O^2, O^1)} \overline{(O^2, O^3)} \overline{(O^6, O^5)} \quad (7.3.7)$$

Nous remarquons en deuxième lieu que le sous-mot v de u n'est pas épuré au sens du deuxième point de déf-7.3.1. En effet, dans cet exemple nous voyons que (O^5, O^6) peut être déplacé et "sortir" de v en utilisant la relation d'Artin $R2$. Nous avons à cet effet :

$$u \equiv \underbrace{(O^5, O^6)(O^1, O^2)}_{v \text{ épuré}} \underbrace{(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_{v \text{ épuré}} (O^2, O^4) \quad (7.3.8)$$

Considérons désormais le mot épuré u suivant et son voisinage obtenu après l'application des deux principes d'épuration :

$$u \overline{(O^2, O^1)} \overline{(O^2, O^3)} \equiv \underbrace{(O^1, O^2)(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_{v \text{ épuré}} (O^2, O^4) \overline{(O^2, O^1)} \overline{(O^2, O^3)}$$

Nous avons alors par (7.2.30) l'équivalence topologique suivante :

$$\begin{aligned} u \overline{(O^2, O^1)} \overline{(O^2, O^3)} &\equiv \underbrace{(O^1, O^2) \left[\overline{(O^2, O^1)} \overline{(O^2, O^3)} \overline{(O^2, O^4)} \right]}_{a(-1, -1, 1)} \underbrace{(O^1, O^3)(O^1, O^4)(O^3, O^4)}_{d[a(-1, -1, 1)]} \\ &\quad \underbrace{(O^3, O^1) \left[\overline{(O^3, O^2)} \overline{(O^1, O^2)} \overline{(O^4, O^2)} \right]}_{d[a(-1, -1, 1)]} \underbrace{(O^2, O^4) \overline{(O^2, O^1)} \overline{(O^2, O^3)}}_{d[a(-1, -1, 1)]} \end{aligned}$$

Nous voyons alors que le mouvement isotopique $\xrightarrow{O^2}$ signé topologiquement par $(-1, -1, 1)$ engendre 4 réductions libres puisque nous avons :

$$u \overline{(O^2, O^1)} \overline{(O^2, O^3)} \equiv \left[\overline{(O^2, O^3)} \overline{(O^2, O^4)} \right] \underbrace{(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_{d[a(-1, -1, 1)]} \quad (7.3.9)$$

Ceci nous donne *in fine* le mot w équivalent ci-dessous :

$$w \equiv \underbrace{(O^5, O^6)}_{d[a(-1, -1, 1)]} \left[\overline{(O^2, O^3)} \overline{(O^2, O^4)} \right] \underbrace{(O^1, O^3)(O^1, O^4)(O^3, O^4)(O^3, O^1)}_{d[a(-1, -1, 1)]} \overline{(O^6, O^5)} \quad (7.3.10)$$

Le mouvement isotopique étudié a permis de diminuer la taille du mot de tresse de 10 à 8 croisements. **Nous venons de décrire, ce que nous entendons par un mouvement isotopique de brin libre permettant de minimiser le critère de la taille d'un mot de tresse.**

Malgré ces différentes considérations, il arrive qu'il n'existe pas, à certaines étapes d'un procédé de démêlage, de mouvements isotopiques de brin libre qui puisse réduire un mot de tresse. Dans ce cas, nous avons possiblement atteint l'"**optimum global**", et la taille du mot de tresse ne peut être davantage réduite. Mais, il arrive souvent que nous soyons en fait dans le cas d'un "**optimum local**". Dans ce contexte, d'autres réductions libres peuvent être effectuées mais

pas de manière immédiate. Il peut exister⁴⁰ alors certains mouvements de brins libres donnant un mot de tresse de même taille mais qui peuvent permettre de faire apparaître *a posteriori* des mouvements de brins libres qui conduiraient alors à nouveau à une réduction de la taille du mot de tresse.

Ainsi au delà du critère de la taille du mot de tresse que nous cherchons à minimiser, nous proposons les sous-critères ci-dessous qui permette de sortir des situations d'“optimum local” lorsque celles-ci se présentent.

Nous avons en effet les propriétés suivantes que nous illustrons à l'aide de la figure fig-7.14 et fig-7.15. La première met en valeur le sous-critère lié aux “poignées de Dehornoy simple” que nous définissons ci-après. La deuxième, met en évidence l'importance des “poignées de Dehornoy” dans la simplification d'une tresse : les faire apparaître permet de conduire à une réduction libre.

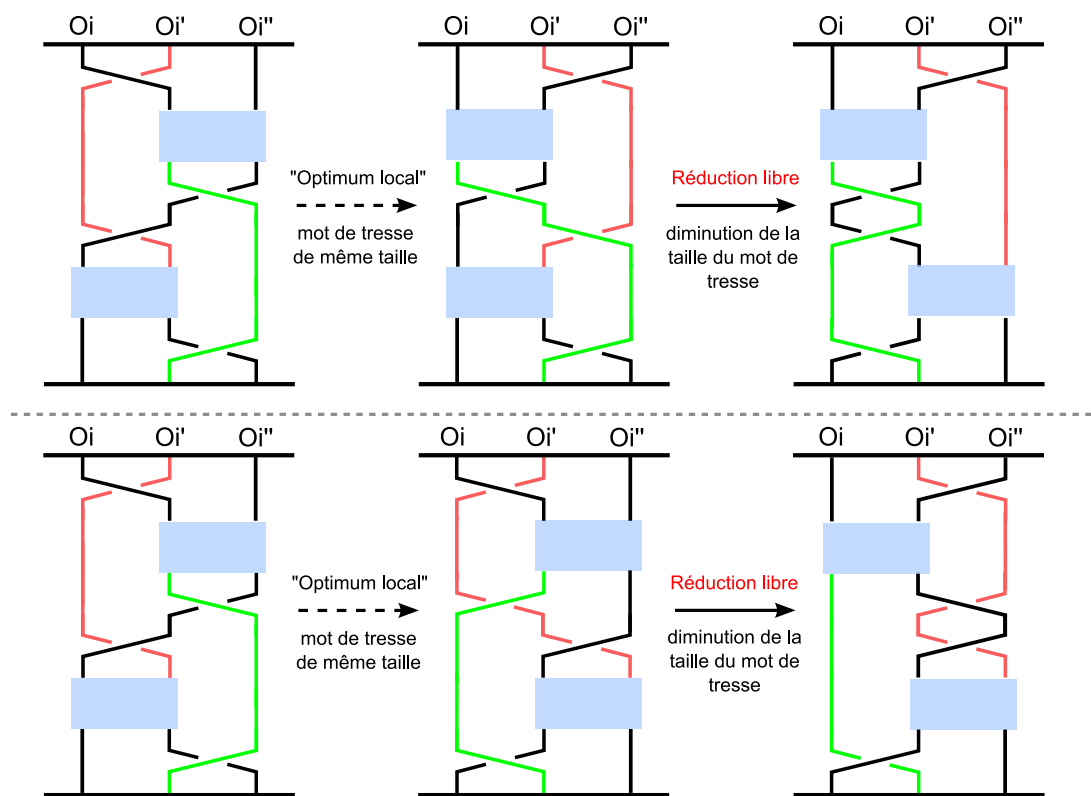


FIGURE 7.14 – Mouvement de brin libre faisant apparaître un brin libre de type “poignée de Dehornoy simple” permettant de “sortir d’un optimum local”

Définition 7.3.2 Brins libres de type “poignée de Dehornoy simple”

Soit w un mot de tresse dans le codage des paires de brins nommés et juxtaposés. Soit une configuration de brin libre $O^i v O^i$ de type “poignée de Dehornoy”, c’est à dire un des 4 cas suivants : 1b, 2a, 3d, 4c. Nous dirons de plus que O^i vérifie sur un sous-mot de w incluant $O^i v O^i$ un motif de brin libre de type “poignée de Dehornoy simple” si nous avons une des séquences suivantes :

- $\underbrace{(O^i, O^{i''})}_{1b} \underbrace{(O^i, O^{i'}) v (O^{i'}, O^i)}_{1b} \underbrace{(O^{i''}, O^i)}$
- $\underbrace{(O^i, O^{i''})}_{2a} \underbrace{(O^i, O^{i'}) v (O^{i'}, O^i)}_{2a} \underbrace{(O^{i''}, O^i)}$

40. principe des “plateaux” en optimisation générale

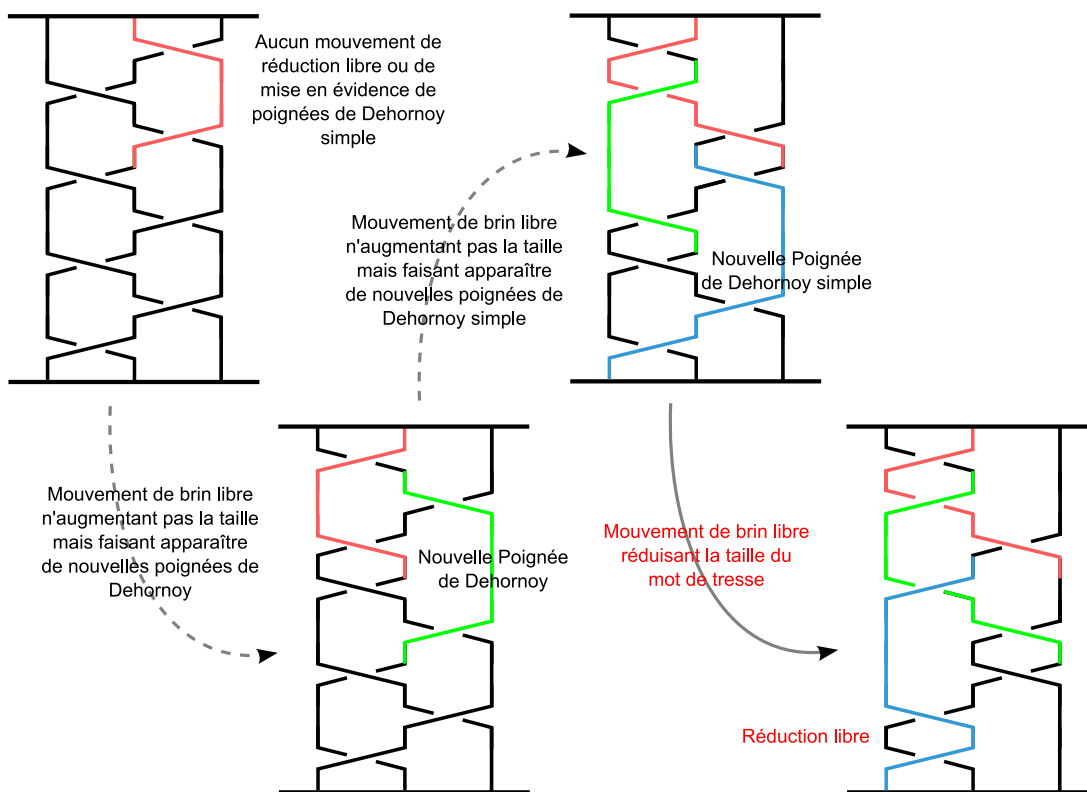


FIGURE 7.15 – Mouvement de brin libre faisant apparaître un brin libre de type “poignée de Dehornoy” permettant ensuite de “sortir d’un optimum local”

- $\overbrace{(O^{i''}, O^i)} \underbrace{(O^{i'}, O^i)v(O^i, O^{i'})}_{3d} \text{ ou } \underbrace{(O^{i'}, O^i)v(O^i, O^{i'})}_{3d} \overbrace{(O^i, O^{i''})}$
- $\overbrace{(O^{i''}, O^i)} \underbrace{(O^{i'}, O^i)v(O^i, O^{i'})}_{4c} \text{ ou } \underbrace{(O^{i'}, O^i)v(O^i, O^{i'})}_{4c} \overbrace{(O^i, O^{i''})}$

Propriété 7.3.2 Sous-critères permettant de “sortir d’un optimum local”

Soit w un mot de tresse dans le codage des paires de brins nommés. Plaçons nous dans le contexte du démêlage d’une tresse par l’application de mouvements isotopiques de brins libres. Dans le cas où il n’existe, à une étape du procédé, aucun mouvement de brin libre permettant de diminuer la taille du mot de tresse, alors nous chercherons à effectuer en premier lieu un mouvement de brin libre gardant la taille du mot de tresse constante et permettant de mettre en évidence des brins libres de type “poignée de Dehornoy simple” comme cela est illustré dans la figure fig-7.14. Si celui-ci n’existe pas, nous chercherons alors dans un deuxième lieu, à effectuer un mouvement de brin libre gardant la taille du mot de tresse constante et permettant de mettre en évidence de nouveaux brins libres de type “poignée de Dehornoy” comme cela est illustré par exemple, dans la figure fig-7.15.

Nous proposons à l’issue de cette sous-section des “sketchs” d’algorithmes qui mettent en valeur les différentes propriétés et définitions étudiées tout au long de cette section et qui permettent de “démêler une tresse”. Nous n’avons pas démontré pour l’instant qu’il s’agissait de procédés de détermination de l’“optimum global”. Il s’agit donc *a priori* de procédés heuristiques qui donnent une approximation de la solution recherchée. En particulier, il faudrait montrer que les deux sous-critères proposés sont suffisants pour sortir de tout type de situation d’“optimum local”.

Le premier procédé est de type “descente” :

Algorithm 6 Démêlage d’une tresse

Require: liste

while La condition d’arrêt n’est pas vérifiée **do**

- Rechercher un mouvement isotopique de brin libre de type “poignée” permettant de diminuer la taille du mot de tresse
- Si un tel mouvement existe, ajouter le mot de tresse obtenu dans la liste et mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

if S’il n’y a pas de mouvement isotopique diminuant la taille du mot de tresse **then**

- Rechercher un mouvement isotopique de brin libre de type “poignée” n’augmentant pas la taille du mot de tresse et permettant de faire apparaître des mouvements de brins libres de type “poignée de Dehornoy simple”
- Si un tel mouvement existe, ajouter le mot de tresse obtenu dans la liste et mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

end if

if S’il n’y a pas de mouvement isotopique diminuant la taille du mot de tresse et ne permettant pas de faire apparaître des mouvements de brins libres de type “poignée de Dehornoy simple” **then**

- Rechercher un mouvement isotopique de brin libre de type “poignée” n’augmentant pas la taille du mot de tresse et permettant de faire apparaître de nouveaux mouvements de brins libres de type “poignée de Dehornoy”
- Si un tel mouvement existe, ajouter le mot de tresse obtenu dans la liste et mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

end if

end while

- Prendre comme mot de tresse celui de la liste qui est la plus optimale au sens de l’ordre “shortlex”

while La condition d’arrêt n’est pas vérifiée **do**

- Rechercher un mouvement isotopique de brin libre de type “poignée” n’augmentant pas la taille du mot de tresse ou alors une relation d’Artin, permettant d’améliorer la forme canonique “shortlex”

end while

Les conditions d’arrêt de l’algorithme sont la non existence à une étape donnée, de mouvement isotopique optimisant au moins un des critères respectifs.

Le procédé proposé vise à résoudre successivement les problèmes du mot de tresse⁴¹, du mot de tresse minimal et du mot de tresse “shortlex”. Si la taille du mot de tresse diminue jusqu’à être nulle, alors la solution aux 3 problèmes est évidente. Si la première condition d’arrêt est atteinte et que le mot de tresse obtenu est de taille non nulle, alors celui-ci correspond à un mot de tresse minimal (ou une approximation). Le procédé cherche alors à déterminer des mouvements isotopiques optimisant la forme canonique “shortlex”.

Nous illustrons graphiquement ce premier procédé dans la figure fig-7.16.

41. dans le cas où $w = v' \circ (v')^{-1}$

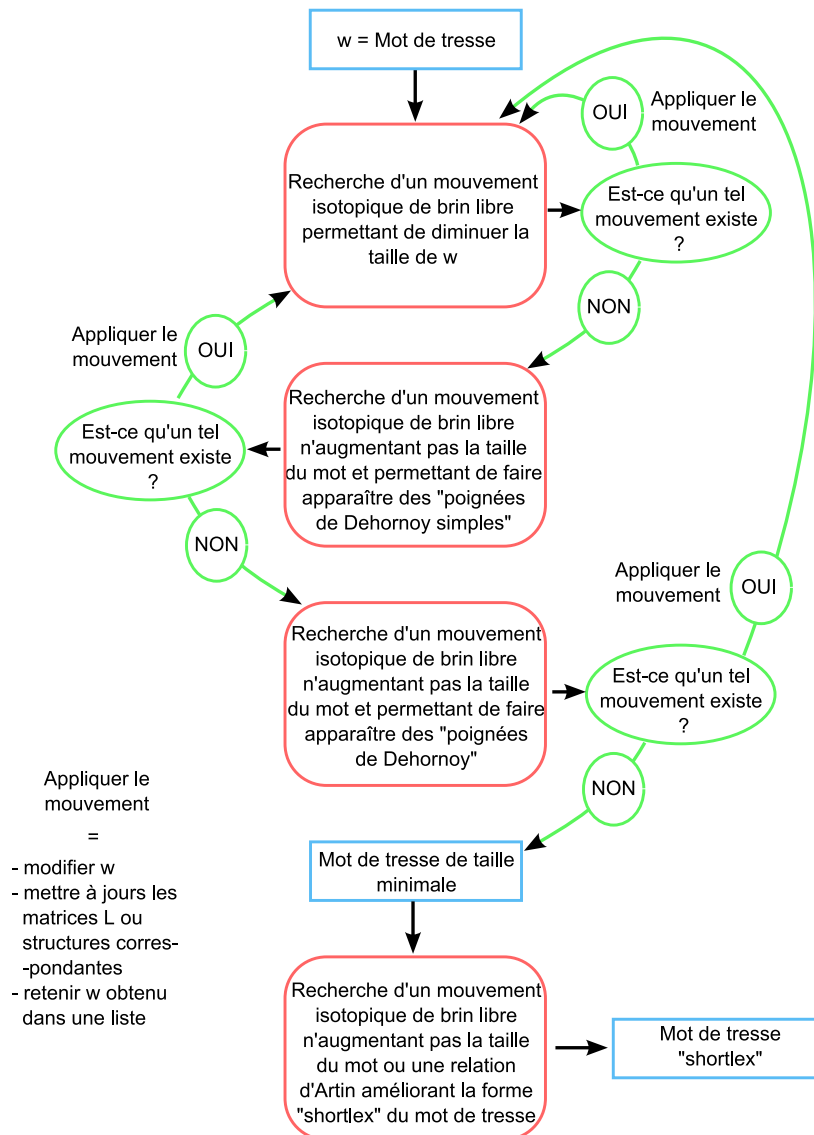


FIGURE 7.16 – Procédé 1 permettant de calculer dans le meilleur des cas la forme canonique “shortlex” d’un mot de tresse

Nous proposons un deuxième procédé qui est inspiré de la méthode de recherche avec liste “tabou”. Celui-ci est donné par l’algorithme donné ci-dessous.

La condition d’arrêt dans ce cas est qu’il n’y ait plus de mouvement isotopique diminuant la taille de la tresse, ni de mouvements isotopiques, n’augmentant pas la taille du mot de tresse, qui conduise à un mot de tresse qui ne soit pas déjà dans la liste des “optima locaux”.

Cette approche est plus coûteuse en temps de traitement mais dans la mesure où elle contient les cas énumérés par le procédé précédent, elle est davantage exhaustive eu égard aux opérations de mouvement de brin libre proposées pour améliorer la solution. La solution obtenue sera alors au moins aussi bonne que celle donnée par le procédé précédent.

Nous illustrons de la même manière ce deuxième procédé dans la figure fig-7.17.

Nous terminerons cette sous-section en faisant une remarque concernant les procédés présentés ici et leurs similarités avec les algorithmes de classification automatique exposés au cha-

Algorithm 7 Démêlage d'une tresse par recherche avec liste "tabou"**Require:** liste**while** La condition d'arrêt n'est pas vérifiée **do**

- Rechercher un mouvement isotopique de brin libre de type "poignée" permettant de diminuer la taille du mot de tresse

- Si un tel mouvement existe, mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

if S'il n'y a pas de mouvement isotopique diminuant la taille du mot de tresse **then**

- Rechercher un mouvement isotopique de brin libre de type "poignée" n'augmentant pas la taille du mot de tresse ou une relation d'Artin, dont le mot de tresse obtenu soit distinct des mots de tresse de la liste

- Si un tel mouvement existe, ajouter à la liste le mot de tresse de taille constante et mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

end if**end while**

- Retourner le mot de tresse de la liste "tabou" dont la forme canonique "shortlex" est la plus petite après avoir testé les relations d'Artin

pitre 2. En effet pour chacun de ces algorithmes, nous cherchons à chaque étape à améliorer le critère. Pour cela nous effectuons des actions ou des mouvements sur les structures étudiées.

Dans le cadre de la classification automatique, nous avons présenté un graphe où les partitions représentent les noeuds et les arcs sont des actions sur une partition permettant de changer sa structure. Nous cherchons alors des chemins permettant de passer d'une partition à une autre de sorte à ce que la valeur du critère de la partition d'arrivée soit meilleure.

Dans le cadre des tresses, nous avons défini un graphe représentatif de \mathbb{B}_N où une tresse est un chemin de ce graphe. Nous cherchons alors à déterminer un autre chemin qui soit équivalent au précédent mais dont la longueur est plus petite. Pour cela, nous effectuons des mouvements qui correspondent localement à des changements de sous-chemins conduisant à des réductions libres.

Par ailleurs, nous savons que, dans le cas d'un mouvement isotopique de brins libres de type "poignée" $O^i v O^i$, il faut *a priori* évaluer $2^{|\#\text{Supp}(v)|}$ signatures différentes et le nombre de réductions libres qu'elles engendrent chacune. Dans le pire des cas, il semble clair que le procédé est exponentiel et que notre approche aboutirait à dire que les problèmes du mot de tresse minimal et "shortlex" seraient \mathcal{NP} -complet. Toutefois, nous constatons qu'il est possible de ne pas effectuer tous les tests de signatures en s'aidant des matrices relationnelles d'ordre topologique. En effet, nous avons mentionné précédemment qu'il était utile de ramener en amont et en aval de u représentatif de $O^i v O^i$, les paires de brins nommés faisant intervenir le brin libre O^i afin d'évaluer le nombre de réductions libres possibles. Dans l'objectif de réduire librement un mot de tresse, l'information apportée par ces paires de brins nommés que nous cherchons à supprimer, nous guide dans la recherche du mouvement isotopique.

Par ailleurs, il est clair que plus le mot v d'un brin libre $O^i v O^i$ engendre une matrice relationnelle d'ordre topologique dense, moins il y a de possibilités de mouvements isotopiques de brin libre.

Ainsi, nous voyons que la complexité peut être réduite en pratique en considérant différents types d'information.

Ainsi, nous n'apportons pas pour l'instant de contributions immédiates aux travaux menés autour des problèmes algorithmiques du groupe des tresses. Toutefois, nous souhaitons mettre l'accent sur le caractère original de la modélisation proposée qui fait notamment intervenir :

- des mouvements isotopiques de brins qui sont génériques
- la notion d'ordre topologique engendrée par un mot de paires de brins nommés

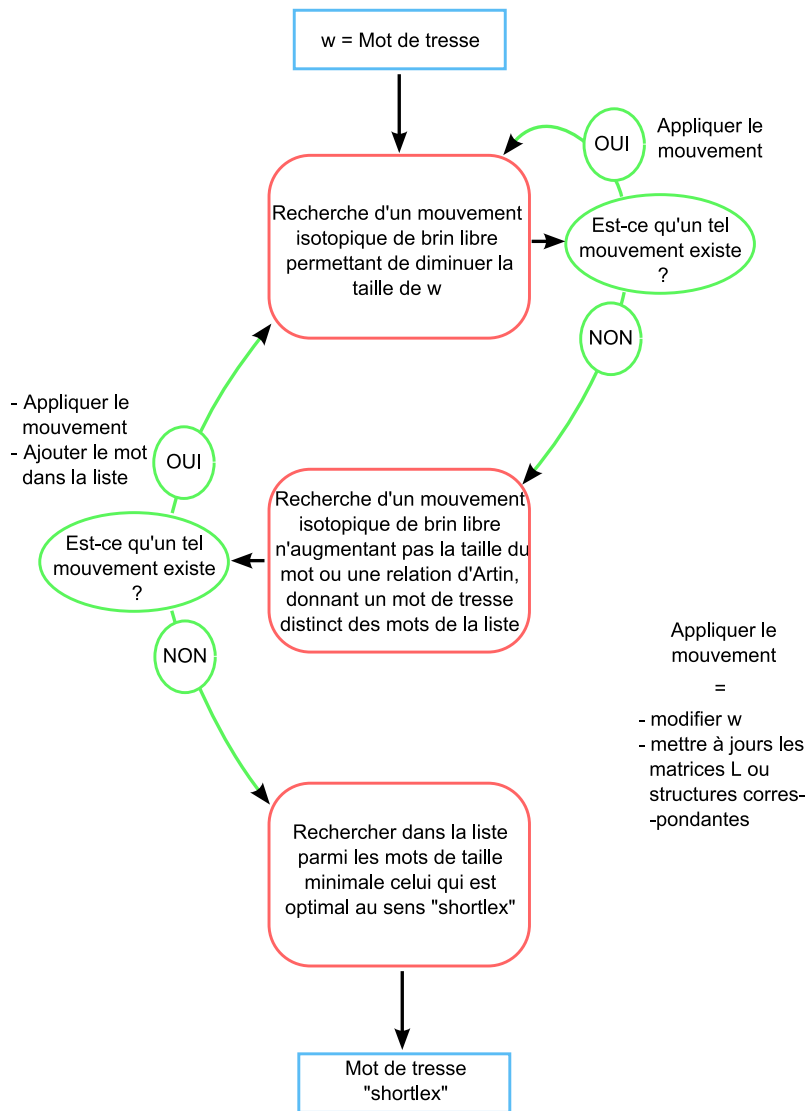


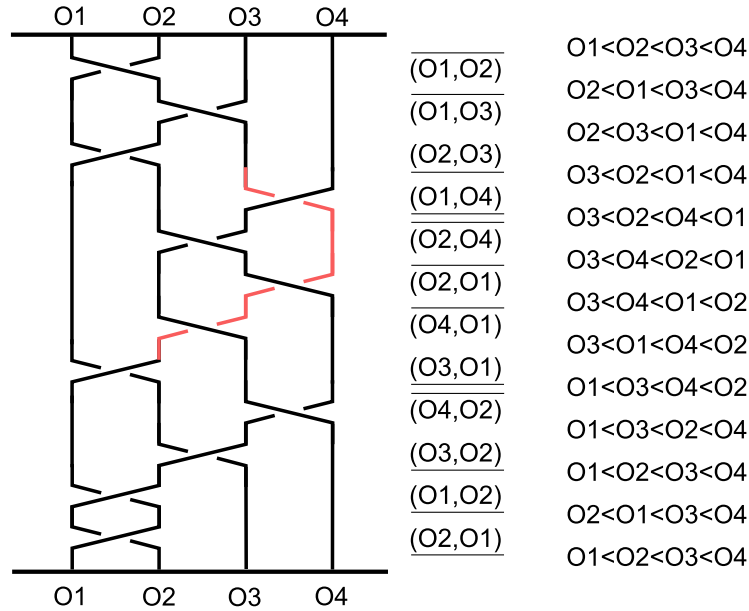
FIGURE 7.17 – Procédé 2 permettant de calculer dans le meilleur des cas la forme canonique “shortlex” d’un mot de tresse

- la vision davantage “optimisation” de ces problèmes

Développer cette approche pourrait, nous le pensons, aider à mettre en évidence de nouvelles propriétés des tresses.

7.3.2 Exemples d'application et considérations diverses

Nous illustrons les différentes étapes du premier procédé proposé pour démêler une tresse sur un exemple tiré de [Deh97]. Le mot traité est le suivant : $w = ABacBCBaCbaa$. Nous représentons cette tresse dans la figure fig-7.18.



$$\overline{(O1, O2)} \overline{(O1, O3)} \overline{(O2, O3)} \overline{(O1, O4)} \overline{(O2, O4)} \overline{(O2, O1)} \overline{(O4, O1)} \overline{(O3, O1)} \overline{(O4, O2)} \overline{(O3, O2)} \overline{(O1, O2)} \overline{(O2, O1)}$$

A B a c B C B a C b a a

FIGURE 7.18 – Exemple d'application : identification du brin libre de type "poignée de Dehornoy simple" conduisant à une réduction de la taille du mot de tresse

En termes de paires de brins nommés, nous avons :

$$w = \overline{(O^1, O^2)} \overline{(O^1, O^3)} \overline{(O^2, O^3)} \overline{(O^1, O^4)} \overline{(O^2, O^4)} \overline{(O^2, O^1)} \overline{(O^4, O^1)} \overline{(O^3, O^1)} \overline{(O^4, O^2)} \overline{(O^3, O^2)} \overline{(O^1, O^2)} \overline{(O^2, O^1)} \quad (7.3.11)$$

En parcourant, le mot de tresse, nous identifions un premier brin libre $O^2 \nu O^2 = \overline{(O^1, O^2)} \overline{(O^1, O^3)} \overline{(O^2, O^3)}$. Celui-ci représente un motif de type $4c$. Ce type de "poignée de Dehornoy" est souvent sujet à un mouvement isotopique à droite "au-dessous" donné par prop-7.2.6. C'est le cas ici, nous pouvons voir que tout autre mouvement isotopique augmenterait la taille du mot de tresse.

De manière générale, pour ce genre de motif, il est d'office inutile de tester les signatures α qui feraient avoir un premier croisement appartenant à $a(\alpha)$ ou un dernier croisement appartenant à $d[a(\alpha)]$, qui impliquerait que le brin libre est au-dessus du brin juxtaposé. En effet, nous constatons que ce type de mouvement isotopique augmenterait systématiquement la taille du mot de tresse. Dans le cadre d'un procédé de diminution de la taille d'un mot de tresse, nous testons uniquement les brins libres de type "poignées". Nous voyons qu'en général il est inutile de tester certaines signatures qui sont celles qui créent des doubles croisements en amont et en aval du déplacement. Ceci ramène donc le nombre de signatures à tester au plus égale à $2^{\#\text{Supp}(v)-2}$, v étant un mot épuré. Comme nous l'avons fait remarquer, en fin de la sous-section précédente, il existe certainement d'autres propriétés permettant de diminuer le nombre de tests de signatures d'un mouvement isotopique de brin libre. Les recherches permettraient d'avoir des éclaircissements sur la complexité du problème.

Revenons à notre exemple d'application, nous voyons que le déplacement isotopique "au-dessous" du brin O^2 ne permet pas de diminuer la taille de la tresse.

Nous poursuivons alors la lecture du mot de tresse et nous détectons un nouveau brin libre de type "poignée de Dehornoy"⁴² : $O^1 v O^1 = \overline{(O^1, O^4)} \overline{(O^2, O^4)} \overline{(O^2, O^1)}$, qui est représentée par un brin rouge sur la figure fig-7.18. Nous remarquons par ailleurs que le croisement suivant ce brin libre, c'est à dire en aval de u , est $\overline{(O^4, O^1)}$ qui fait donc intervenir le brin libre O^1 . Nous cherchons alors à savoir si un mouvement isotopique $\xleftarrow{O^1}$ permet de réduire librement ce croisement. Pour ce faire, nous cherchons une signature α pour laquelle, on aurait pour $a(\alpha)$ la paire de brins nommés $\overline{(O^1, O^4)}$.

Dans ce cas particulier, nous avons :

$$u \overline{(O^4, O^1)} = \overline{(O^1, O^4)} \underbrace{\overline{(O^2, O^4)} \overline{(O^2, O^1)}}_v \overline{(O^4, O^1)} \quad (7.3.12)$$

Nous avons donc $v = \overline{(O^2, O^4)}$, $i^c = 4$ et $j^i = j^1 = 4$. Nous avons par ailleurs, les matrices de paires de brins nommés L^2 et $L^{|U|} = L^3$ représentatives respectivement des ordres $O^3 < O^2 < O^4 < O^1$ et $O^3 < O^4 < O^2 < O^1$ qui valent :

$$L^2 = \begin{pmatrix} [O^3, O^2] & [O^3, O^4] & [O^3, O^1] \\ [O^2, O^4] & [O^2, O^1] & [O^4, O^1] \end{pmatrix} ; \quad L^3 = \begin{pmatrix} [O^3, O^4] & [O^3, O^2] & [O^3, O^1] \\ [O^4, O^2] & [O^4, O^1] & [O^2, O^1] \end{pmatrix} \quad (7.3.13)$$

Il s'agit d'un déplacement à gauche et nous avons alors des chemins de type "va et vient verticaux". v étant épuré, nous avons $\text{Supp}(v) = \{O^2, O^4\}$ et donc $|d| = |a| = \#\text{Supp}(v) = 2$. Nous avons alors :

$$d = \overline{(O^4, O^1)} \overline{(O^2, O^1)} \text{ et } a = r(\overline{(O^4, O^1)}) r(\overline{(O^2, O^1)}) = \overline{(O^1, O^4)} \overline{(O^1, O^2)}.$$

Considérons la matrice relationnelle d'ordre topologique engendrée par $v = \overline{(O^2, O^4)}$:

$$C^{v,o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.3.14)$$

Le brin libre est de type "poignée de Dehornoy". Nous avons fait remarquer qu'il faut avant tout, pour ne pas créer de nouveaux croisements supplémentaires, réduire librement les croisements en amont et en aval de v c'est à dire $\overline{(O^1, O^4)}$ et $\overline{(O^2, O^1)}$. Le premier terme de d et le dernier terme de a doivent être pour cela $\overline{(O^4, O^1)}$ et $\overline{(O^1, O^2)}$. Ceci implique une matrice relationnelle d'ordre topologique de $a(\alpha)$ qui vaut :

$$C^{a(\alpha),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.3.15)$$

Nous avons alors $\alpha = (1, 1)$ ce qui implique :

$$a(1, 1) = \overline{(O^1, O^4)} \overline{(O^1, O^2)} \quad (7.3.16)$$

$$d[(a(1, 1))] = \overline{(O^4, O^1)} \overline{(O^2, O^1)} \quad (7.3.17)$$

42. qui est en fait une "poignée de Dehornoy simple"

Formellement, le mouvement de brin étudié permet les réductions libres suivantes que nous présentons séquentiellement :

$$\begin{aligned}
 \overline{u(O^4, O^1)} &= \underbrace{(O^1, O^4)(O^2, O^4)(O^2, O^1)}_u \overline{(O^4, O^1)} \\
 &\equiv \underbrace{(O^1, O^4)}_u \underbrace{[(O^4, O^1)(O^2, O^1)]}_{d[a(1,1)]} \underbrace{(O^2, O^4)}_v \underbrace{[(O^1, O^4)(O^1, O^2)]}_{a(1,1)} \overline{(O^2, O^1)(O^4, O^1)} \\
 &\equiv \overline{[(O^2, O^1)]} \underbrace{(O^2, O^4)}_v \overline{[(O^1, O^4)]} \overline{(O^4, O^1)} \\
 &\equiv \overline{[(O^2, O^1)]} \underbrace{(O^2, O^4)}_v
 \end{aligned}
 \tag{7.3.18}$$

Nous sommes ainsi passés localement d'un mot $\overline{u(O^4, O^1)}$ de taille 4 à un mot équivalent de taille 2. D'un point de vu global, nous avons donc amélioré le critère puisque w passe d'une taille de 12 croisements à un mot équivalent d'une taille de 10 que nous représentons sur la figure fig-7.19.

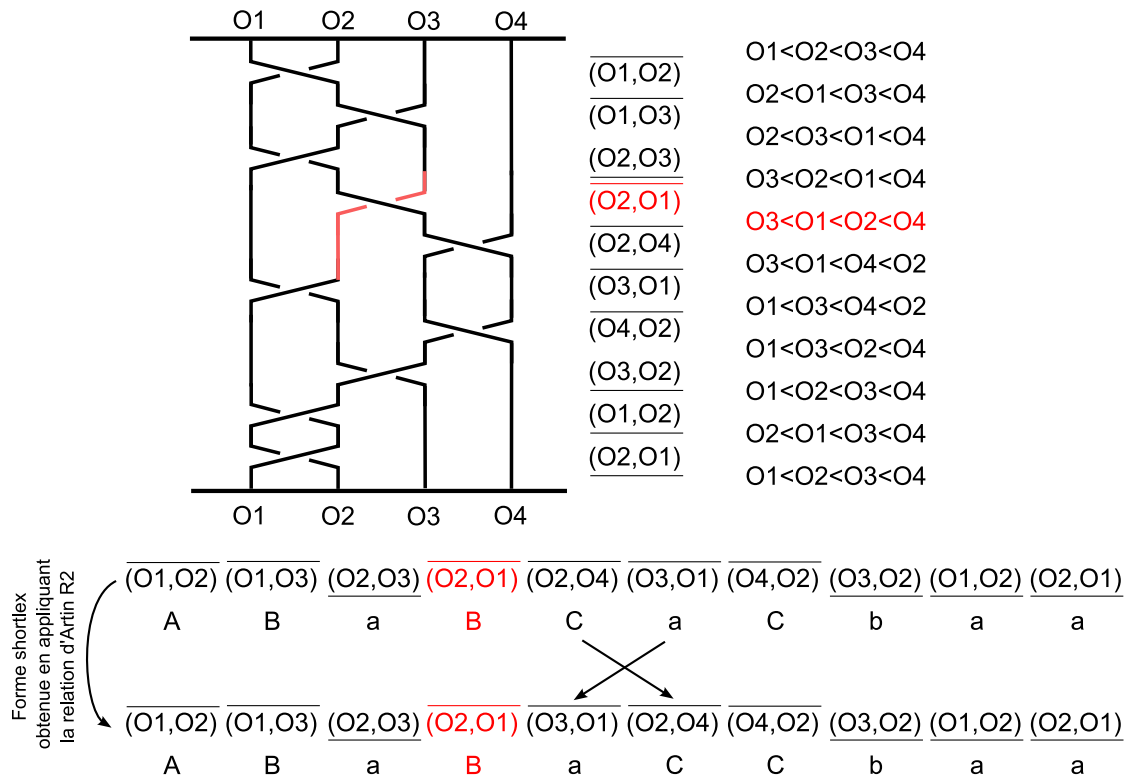


FIGURE 7.19 – Exemple d’application : détermination d’une forme “shortlex” à l’aide d’un mouvement isotopique de brin libre issue d’une poignée qui n’est pas de type de Dehornoy

Suite à ce mouvement, il n’existe pas d’autres mouvements de brins libres permettant soit de diminuer la taille du mot, soit de mettre en évidence des “poignées de Dehornoy simple” sans augmenter le nombre de croisements. Toutefois, il existe des mouvements isotopiques permettant de faire apparaître de nouvelles “poignées de Dehornoy” sans détériorer le critère de la taille. Ces derniers ne débouchent pas finalement sur un mot de tresse de taille plus petite ni même sur un ordre “shortlex” meilleur. Nous n’explicitons pas ici les calculs décrivant ces différentes étapes mais nous donnons cependant dans la figure fig-7.20 une illustration graphique de ce que calculerait à chaque étape le procédé 1 proposé.

Nous “sautons” donc ces étapes qui n’améliorent pas au sens de la taille ni au sens de l’ordre “shortlex” le mot de tresse suivant obtenu après l’application du premier mouvement de brin

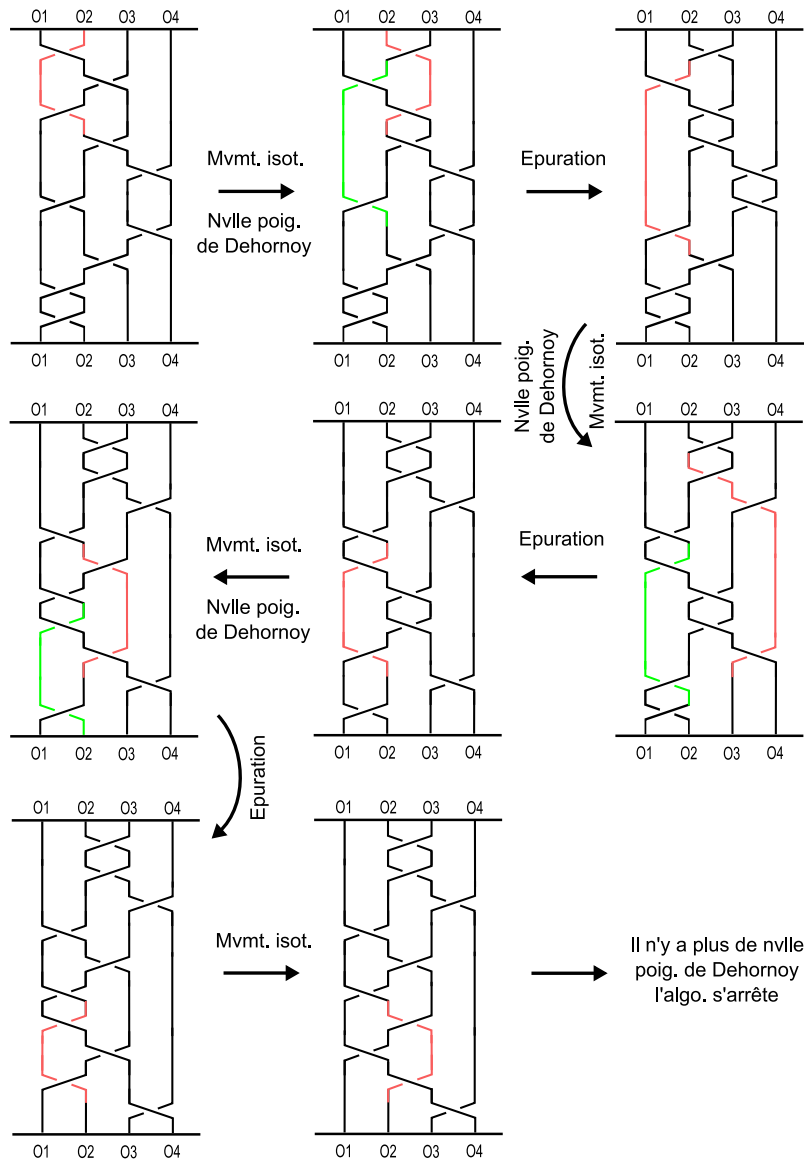


FIGURE 7.20 – Illustration de la terminaison du procédé 1 sur l'exemple traité

libre :

$$w = \overline{(O^1, O^2)} \overline{(O^1, O^3)} \overline{(O^2, O^3)} \overline{(O^2, O^1)} \overline{(O^2, O^4)} \overline{(O^3, O^1)} \overline{(O^4, O^2)} \overline{(O^3, O^2)} \overline{(O^1, O^2)} \overline{(O^2, O^1)}$$

Dans le codage classique, celui-ci est donné par :

$$w = ABaBCaCbaa$$

Nous constatons alors que nous sommes proches de l'“optimum global” du mot de tresse “shortlex” qui est en fait obtenu à l'aide de la relation $R2$ d'Artin qui permet de commuter les paires $\overline{(O^2, O^4)}$ et $\overline{(O^3, O^1)}$.

En effet si nous avons l'ordre strict et total suivant : $a < A < b < B < \dots$, nous voyons que la forme shortlex est obtenue pour le mot suivant :

$$w^* = \overline{(O^1, O^2)} \overline{(O^1, O^3)} \overline{(O^2, O^3)} \overline{(O^2, O^1)} \overline{(O^3, O^1)} \overline{(O^2, O^4)} \overline{(O^4, O^2)} \overline{(O^3, O^2)} \overline{(O^1, O^2)} \overline{(O^2, O^1)}$$

En effet, ce mot correspond, en termes de générateurs classiques, au mot suivant qui est davantage "shortlex" :

$$w^* = ABaBaCCbaa \tag{7.3.19}$$

Nous exposons maintenant un deuxième exemple, qui montre l'efficacité des mouvements de brins libres de type "poignée" qui sont distincts⁴³ des types "poignées de Dehornoy" notamment dans la recherche du mot de tresse "shortlex".

L'exemple étudié est représenté dans la figure fig-7.21 et concerne le mot $w = CBdaacbd c$. En termes de paires de brins nommés nous avons :

$$w = \overline{(O^3, O^4)} \overline{(O^2, O^4)} \overline{(O^3, O^5)} \overline{(O^1, O^4)} \overline{(O^4, O^1)} \overline{(O^2, O^5)} \overline{(O^4, O^5)} \overline{(O^2, O^3)} \overline{(O^4, O^3)} \tag{7.3.20}$$

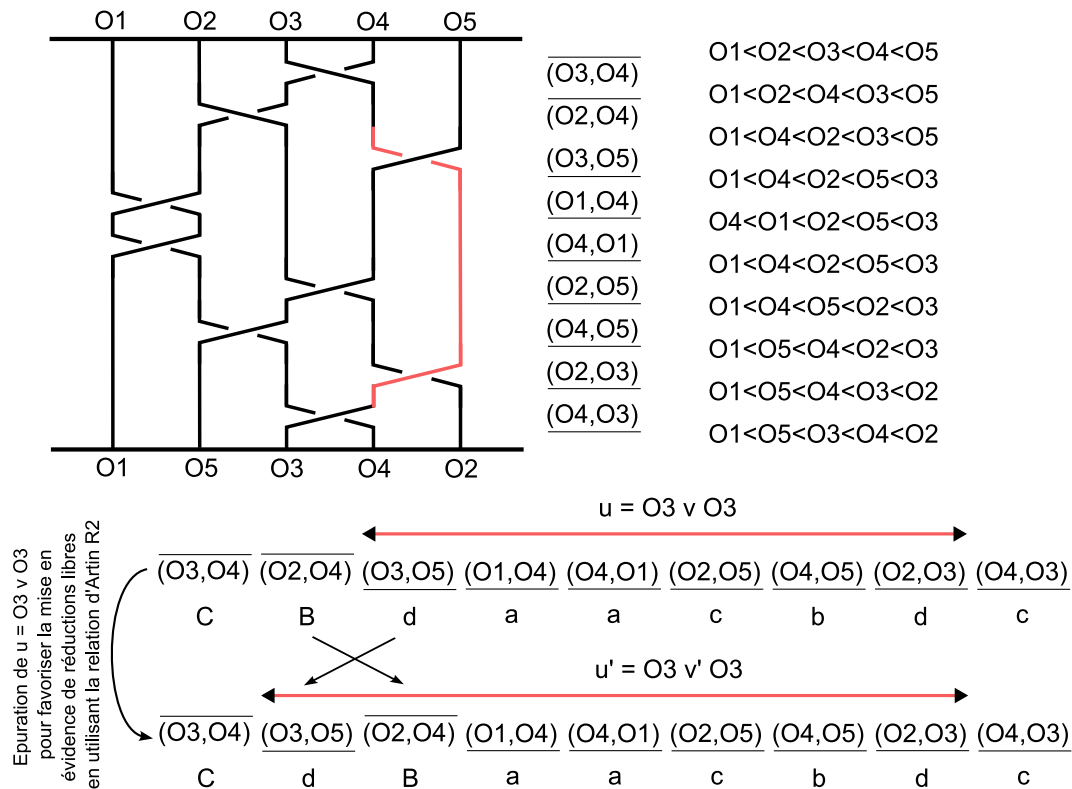


FIGURE 7.21 – Exemple d'application : mot de tresse obtenue après le mouvement isotopique de brin libre et détermination de la forme canonique "shortlex"

Nous voyons qu'il n'existe pas de manière apparente, de brin libre de type "poignée de Dehornoy" pour cet exemple. Il existe en revanche une poignée $O^3 v O^3$ de type $1a$, représentée en rouge sur la figure et pour laquelle :

$$u = \overline{(O^3, O^5)} \underbrace{\overline{(O^1, O^4)} \overline{(O^4, O^1)} \overline{(O^2, O^5)} \overline{(O^4, O^5)}}_v \overline{(O^2, O^3)} \tag{7.3.21}$$

Si on s'intéresse au voisinage de u , on s'aperçoit notamment qu'en amont il y a une paire de brins nommés faisant intervenir le brin libre O^3 , il s'agit de $\overline{(O^3, O^4)}$. Suivant le pré-traitement énoncé dans le cadre de l'"épuration" de u , nous rapprochons ce croisement générateur de u . Nous pouvons effectuer ce pré-traitement en appliquant la relation d'Artin R_2 . Nous obtenons

43. cf déf-7.2.5 page 231

le mot u épuré suivant :

$$u = \underbrace{(O^3, O^5) \overline{(O^2, O^4)} (O^1, O^4) (O^4, O^1) (O^2, O^5) (O^4, O^5) (O^2, O^3)}_v$$

Cette opération va s'avérer très utile dans la mesure où elle conduit à un mot de tresse de taille constante mais davantage "shortlex" que le mot de départ.

Nous avons $i^c = 5$ et $j^i = j^3 = 5$. Les matrices de paires de brins nommés L^2 et $L^{|U|} = L^7$ représentatives respectivement des ordres $O^1 < O^2 < O^4 < O^5 < O^3$ et $O^1 < O^5 < O^4 < O^2 < O^3$, valent alors :

$$L^2 = \begin{pmatrix} [O^1, O^2] & [O^1, O^4] & [O^1, O^5] & [O^1, O^3] \\ & [O^2, O^4] & [O^2, O^5] & [O^2, O^3] \\ & & [O^4, O^5] & [O^4, O^3] \\ & & & [O^5, O^3] \end{pmatrix} \quad (7.3.22)$$

$$L^7 = \begin{pmatrix} [O^1, O^5] & [O^1, O^4] & [O^1, O^2] & [O^1, O^3] \\ & [O^5, O^4] & [O^5, O^2] & [O^5, O^3] \\ & & [O^4, O^2] & [O^4, O^3] \\ & & & [O^2, O^3] \end{pmatrix} \quad (7.3.23)$$

Il s'agit d'un déplacement à gauche et comme précédemment, nous avons des chemins de type "va et vient verticaux". Nous avons $\text{Supp}(v) = \{O^1, O^2, O^4, O^5\}$ et donc $|d| = |a| = \#\text{Supp}(v) = 4$. Nous avons alors :

$d = (O^5, O^3)(O^4, O^3)(O^2, O^3)(O^1, O^3)$ et

$a = r((O^1, O^3))r((O^5, O^3))r((O^4, O^3))r((O^2, O^3)) = (O^3, O^1)(O^3, O^5)(O^3, O^4)(O^3, O^2)$.

Au vu des mouvements a et d , nous pouvons potentiellement augmenter la taille de la tresse de 8 nouveaux croisements. Pour que ce mouvement n'engendre pas de dégradation du critère de la taille, il faut donc avoir au moins 4 réductions libres ce qui correspond à la suppression de 8 croisements.

Considérons à présent, la matrice relationnelle d'ordre topologique engendrée par :

$v = \overline{(O^2, O^4)}(O^1, O^4)(O^4, O^1)(O^2, O^5)(O^4, O^5)$:

$$C^{v,o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \end{matrix} & \begin{pmatrix} 1 & 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.3.24)$$

Pour ne pas créer de nouveaux croisements supplémentaires, il faut systématiquement tenter de réduire librement les croisements en amont et en aval⁴⁴ de v c'est à dire $\overline{(O^3, O^5)}$ et $\overline{(O^2, O^3)}$ dans un premier temps. Pour ce faire, il faudrait que le premier terme de d et le dernier terme de a soient donc $\overline{(O^5, O^3)}$ et $\overline{(O^3, O^2)}$. Ceci implique à ce premier stade une matrice relationnelle d'ordre topologique de $a(\alpha)$ qui vaut :

$$C^{a(\alpha),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \end{matrix} & \begin{pmatrix} 1 & \square & & \times & \square \\ \times & 1 & 1 & \times & 1 \\ & 0 & 1 & & 1 \\ \times & \square & & 1 & \square \\ \times & \times & 0 & \times & 1 \end{pmatrix} \end{matrix} \quad (7.3.25)$$

44. ie les premiers et derniers termes de u

Dans la matrice ci-dessus que nous étudions pour définir un mouvement de brin libre qui soit isotopique, nous avons :

- affecté un 1 aux paires pour lesquelles nous décidons la signification topologique du croisement. Nous savons par ailleurs que si nous avons $C_{ii'}^{a(\alpha),o} = 1$ alors cela implique que $C_{i'i}^{a(\alpha),o} = 0$ en raison de la propriété de symétrie
- laissé un blanc pour les termes relatifs aux éléments constituant a ou d et auxquels nous devons affecter soit 0 soit 1. Ces termes sont en fait les croisements du brin libre O^i avec les éléments $O^{i'}$ du support de v et concernent donc la i -ème ligne et la i -ème colonne de $C^{a(\alpha),o}$
- affecté un \square pour les paires qui seront complétées par transitivité et qui peuvent prendre comme valeur 0 ou 1
- affecté un \times pour les paires qui ne peuvent pas être affectées d'un 1 en raison de la discordance avec la matrice $\check{C}^{v,o}$. Par conséquent, la fermeture transitive ne doit pas engendrer l'affectation d'un 1 à ces termes

Nous cherchons donc à poursuivre la réduction libre tout en respectant les conditions conduisant à un mouvement isotopique. Ainsi, nous cherchons à "remplir" la matrice relationnelle représentative de la signature topologique attribuée aux mouvements a et d en tenant compte des interdictions données par $\check{C}^{v,o}$.

Nous considérons à nouveau le voisinage de u et tentons dans un second temps, de réduire la paire de brins (O^3, O^4) qui est en amont de u et (O^4, O^3) qui est en aval. Pour ce faire, nous regardons si le deuxième élément de d peut être signé topologiquement de la manière suivante (O^4, O^3) . Ceci implique d'ailleurs, par symétrie topologique des mouvements isotopiques, que le troisième terme de a est signé de la manière suivante (O^3, O^4) , ce qui donnerait avec les premières réductions libres un total de 4 réductions libres. Nous posons alors $C_{43}^{a(\alpha),o} = 1$ et nous obtenons la fermeture transitive suivante :

$$C^{a(\alpha),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \end{matrix} & \begin{pmatrix} 1 & \square & & \times & \square \\ \times & 1 & 1 & \times & 1 \\ & 0 & 1 & 0 & 1 \\ \times & \square & 1 & 1 & 1 \\ \times & \times & 0 & \times & 1 \end{pmatrix} \end{matrix} \quad (7.3.26)$$

La fermeture transitive ne donne pas de discordance avec l'ordre topologique engendré par v . A condition que nous puissions poursuivre le remplissage de $C^{a(\alpha),o}$, nous savons déjà que le mouvement engendré n'augmentera pas la taille du mot de tresse.

Il nous reste désormais à choisir les valeurs de $C_{13}^{a(\alpha),o}$ et de $C_{31}^{a(\alpha),o}$. Si nous posons $C_{31}^{a(\alpha),o} = 1$ alors nous constatons que cela aboutit à une interdiction et donc à une discordance puisque nous aurons $C_{23}^{a(\alpha),o} = 1 \wedge C_{31}^{a(\alpha),o} = 1$ implique $C_{21}^{a(\alpha),o} = 1$. Or ce dernier est déjà affecté d'un \times signe d'une interdiction.

Il nous reste alors à tester $C_{13}^{a(\alpha),o} = 1$. Nous voyons finalement que cette dernière affectation définit une signature topologique α qui n'engendre pas de discordances avec l'ordre topolo-

gique engendré par v . En effet, la fermeture transitive de $C^{a(\alpha),o}$ donne :

$$C^{a(\alpha),o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 & O^5 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \\ O^5 \end{matrix} & \begin{pmatrix} 1 & \square & 1 & \times & 1 \\ \times & 1 & 1 & \times & 1 \\ 0 & 0 & 1 & 0 & 1 \\ \times & \square & 1 & 1 & 1 \\ \times & \times & 0 & \times & 1 \end{pmatrix} \end{matrix} \quad (7.3.27)$$

Cette relation d'ordre topologique donne finalement la signature $\alpha = (-1, 1, -1, -1)$ qui relativement à $a = (O^3, O^1)(O^3, O^5)(O^3, O^4)(O^3, O^2)$ donne le mouvement isotopique suivant :

$$a = \overline{(O^3, O^1)}(O^3, O^5)\overline{(O^3, O^4)}(O^3, O^2) \quad (7.3.28)$$

$$d = \overline{(O^5, O^3)}(O^4, O^3)(O^2, O^3)(O^1, O^3) \quad (7.3.29)$$

En appliquant le mouvement décrit par a et d donné précédemment, et après simplification des réductions libres, nous avons finalement le mot de tresse équivalent suivant, que nous avons représenté dans la figure fig-7.22 et qui est davantage "shortlex" que le mot précédent :

$$w \equiv \underline{(O^2, O^3)}(O^1, O^3)\overline{(O^2, O^4)}(O^1, O^4)(O^4, O^1)(O^2, O^5)(O^4, O^5)\overline{(O^3, O^1)}(O^3, O^5)$$

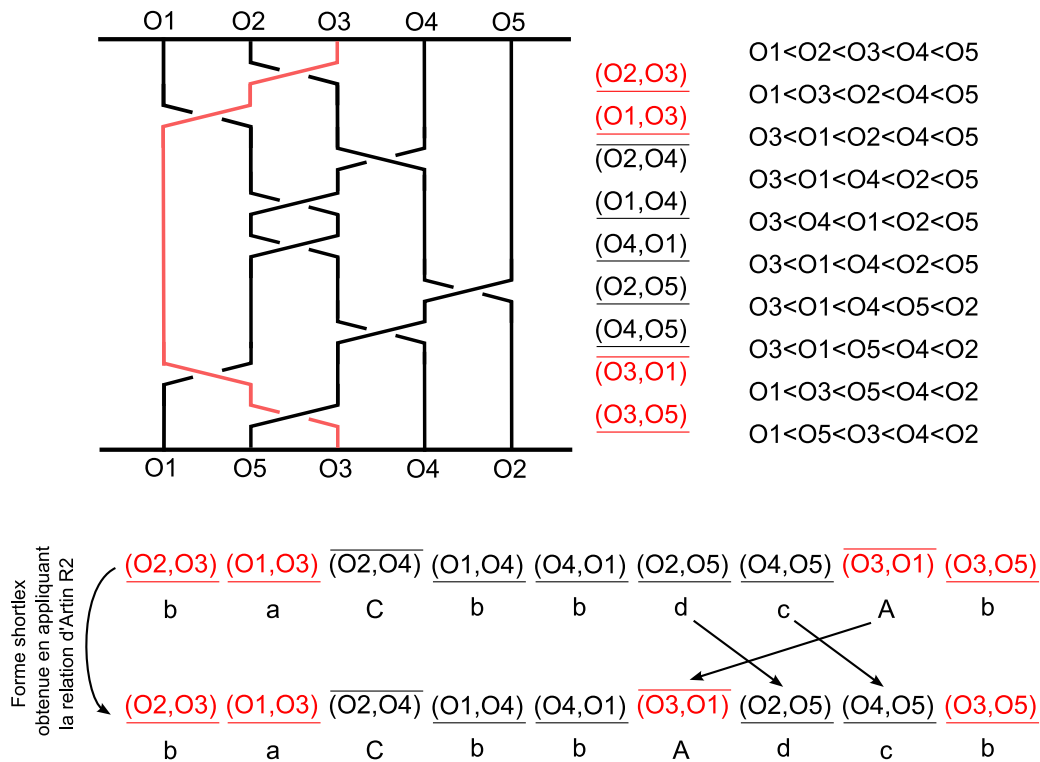


FIGURE 7.22 – Exemple d'application : mot de tresse obtenu après le mouvement isotopique de brin libre et détermination de la forme canonique "shortlex"

Il n'existe plus ensuite de mouvements de brins qui permettent de réduire la taille du mot de tresse ou de faire apparaître des "poignées de Dehornoy simples" ni des nouvelles "poignées de Dehornoy" sans faire augmenter la taille du mot de tresse.

En appliquant les relations d'Artin, nous pouvons alors améliorer la forme "shortlex" du mot de tresse obtenu précédemment afin d'aboutir à l'"optimal global" suivant :

$$w^* = \underline{(O^2, O^3)}(O^1, O^3)\overline{(O^2, O^4)}(O^1, O^4)(O^4, O^1)\overline{(O^3, O^1)}(O^2, O^5)(O^4, O^5)(O^3, O^5) \quad (7.3.30)$$

Traduit en termes de générateurs classiques, nous constatons que l'ordre "shortlex" engendré est effectivement plus petit que l'ordre initial :

$$w^* = baCbbAdcb \tag{7.3.31}$$

7.4 Modélisation des noeuds par approche relationnelle

Nous avons rappelé en introduction page 23, les liens étroits existant entre les tresses et les noeuds. En effet, il a été montré par J.W. Alexander que tout noeud pouvait être obtenu par la clôture d'une tresse [Ale28]. Cette opération consiste à relier le brin en position j au début d'une tresse, au brin en position j à la fin de cette tresse pour $j = 1, \dots, N$ où N est le nombre de brins de la tresse. Nous avons illustré la notion de clôture en introduction dans la figure fig-1.4 pour le cas du noeud de trèfle dont le mot de tresse associé est aaa . Nous notons β le noeud obtenu par la clôture d'une tresse géométrique β .

Dans cette section, nous montrons dans un premier temps comment les différentes propriétés permettant de passer des tresses aux noeuds, que sont les mouvements de Markov, se traduisent dans le cadre de notre modélisation par paires de brins nommés.

Dans un deuxième temps, nous nous intéressons de manière succincte au problème du calcul d'un invariant complet d'un noeud. Nous rappelons pour cela les travaux de T. Gittings [Git04], dont la définition d'un invariant complet se ramène à la détermination d'une forme canonique "shortlex" de la tresse représentant le noeud avec prise en compte des mouvements de Markov.

7.4.1 Noeuds tressés dans le cadre de la modélisation relationnelle : la clôture d'une tresse et les mouvements de Markov

L'objet de cette sous-section, consiste à montrer comment, dans le cadre de notre modélisation, nous pouvons prendre en compte les mouvements de Markov qui permettent d'étudier un noeud à partir de la clôture de sa tresse correspondante.

Dans la continuité de la section précédente, nous cherchons donc à modéliser les noeuds tressés par des paires de brins nommés.

Un noeud correspond à la clôture d'une tresse : il faut donc être en mesure de coder cette information. Nous introduisons pour cela la permutation engendrée par les croisements d'un mot de tresse sur les brins de la tresse considérée.

Définition 7.4.1 Permutation engendrée par une tresse sur ses brins

Soit W une succession de couples $(L, (.,.))$ décrivant une tresse dans le cadre de la modélisation par automate du groupe \mathbb{B}_N donnée par déf-7.2.2 et soit w le mot de tresse des seconds membres de W de taille⁴⁵ $|w| = |W| - 1$. Soit $L^{|W|}$ la dernière matrice de paires de brins nommés donnant l'ordre strict et total dans lesquels sont rangés les brins à l'issu des croisements. Nous noterons par π^w la permutation représentative de la liste des brins ordonnés selon $L^{|W|} = L^{|w|} * (.,.)^{|w|}$. Formellement, nous avons la relation suivante :

$$\pi^w = \left(\begin{array}{cccccc} O^1 & O^2 & \dots & O^i & \dots & O^{N-1} & O^N \\ \left[L_{11}^{|W|} \right]_1 & \left[L_{22}^{|W|} \right]_1 & \dots & \left[L_{ii}^{|W|} \right]_1 & \dots & \left[L_{N-1, N-1}^{|W|} \right]_1 & \left[L_{N-1, N-1}^{|W|} \right]_2 \end{array} \right) \tag{7.4.1}$$

45. nous rappelons que W possède un élément de plus que w et qui représente l'ordre strict et total final engendré par $|w|$ paires de brins nommés

Nous noterons par $\left\{ \left(\begin{smallmatrix} O^i \\ \pi^w(O^i) \end{smallmatrix} \right); i = 1 \dots, N \right\}$ les paires de brins nommés reliés par la clôture de la tresse représentée par w . Lorsque nous considérons un mot de tresse w clôturé qui donne ainsi la définition d'un noeud particulier, nous parlerons de mot de noeud tressé.

Le résultat de permutation sur les brins engendrés par le mot de tresse w nous permet d'obtenir l'identité des brins reliés par la clôture d'une tresse. En effet, le brin $\pi^w(O^i)$ en fin de mot de tresse est celui qui est relié au brin O^i en début de mot de tresse. Par ailleurs, nous constatons qu'à l'évidence, le nombre de cycles⁴⁶ de π^w donne le nombre d'entrelacs. Par conséquent, nous voyons qu'une condition nécessaire et suffisante pour que la clôture d'une tresse donne un noeud et non un entrelac est que la permutation π^w soit cyclique, c'est à dire qui ne possède qu'un seul cycle et aucun point fixe [MK99].

Ainsi, l'ensemble des paires de brins nommés et reliés donné par π^w , nous apporte des informations quant à la clôture d'une tresse. Cet ensemble, associé aux paires de brins nommés du mot de tresse w , nous permet d'aboutir à des propriétés qui expriment la prise en compte des mouvements de Markov dans le cadre de notre modélisation. Ces propriétés sont émises dans l'esprit de la simplification d'un noeud tressé. Ainsi, nous montrons comment, nous pouvons simplifier la taille d'un mot de noeud tressé par $M1$ ou encore comment supprimer un brin "superflu" par $M2$ à partir des paires de brins nommés⁴⁷.

Propriété 7.4.1 Détection et réduction libre issu d'un mouvement de Markov $M1$ dans le cadre de la modélisation relationnelle des noeuds

Soit w un mot de noeud tressé en termes de paires de brins nommés et notons par $|w|$ la taille de ce mot.

Soit π^w la permutation engendrée par le noeud tressé et soit $\left\{ \left(\begin{smallmatrix} O^i \\ \pi^w(O^i) \end{smallmatrix} \right); i = 1, \dots, N \right\}$ les paires de brins nommés reliés par la clôture de la tresse. Nous pouvons, conformément au mouvement de Markov $M1$, réduire librement les paires $(\cdot, \cdot)^1$ et $(\cdot, \cdot)^{|w|}$ si une des conditions suivantes est vérifiée :

$$(\cdot, \cdot)^1 = \overline{(O^i, O^{i'})} \quad \text{et} \quad (\cdot, \cdot)^{|w|} = \overline{(\pi^w(O^{i'}), \pi^w(O^i))}$$

ou

(7.4.2)

$$(\cdot, \cdot)^1 = \overline{(O^i, O^{i'})} \quad \text{et} \quad (\cdot, \cdot)^{|w|} = \overline{(\pi^w(O^{i'}), \pi^w(O^i))}$$

La réduction libre selon un mouvement de Markov $M1$ s'effectue en pratique de la manière suivante :

1. $(L^k, (\cdot, \cdot)^k) \leftarrow (L^{k+1}, (\cdot, \cdot)^{k+1}) \quad \forall k = 1, \dots, |w| - 1$
2. $|W| \leftarrow |W| - 2$
3. $(\cdot, \cdot)^{|W|} \leftarrow \epsilon$

Remarque 7.4.1 A propos des états initiaux des mots de noeuds tressés dans le codage des paires de brins nommés

Nous rappelons qu'une tresse est représentée dans le codage de paires de brins nommés par un chemin W dans l'automate représentatif de \mathbb{B}_N au sens des définitions déf-7.2.1 et déf-7.2.2. Nous avons émis la convention que l'état initial L^1 est la matrice de paires de brins nommés représentative de l'ordre $O^1 < O^2 < \dots < O^N$. Cet état initial est figé pour tout mot de tresse.

Dans le cas d'une clôture d'un mot de tresse définissant un noeud tressé, cet état initial peut être modifié en raison de mouvements de Markov $M1$ ou en raison de passage d'une paire de brins nommés en début de mot vers la fin de ce dernier comme nous le verrons ultérieurement.

46. cf déf-5.1.9 page 162

47. où $M1$ et $M2$ sont les deux types de mouvements de Markov possibles et qui ont été introduits page 24

Nous illustrons graphiquement les conditions de la propriété prop-7.4.1 dans la figure fig-7.23.

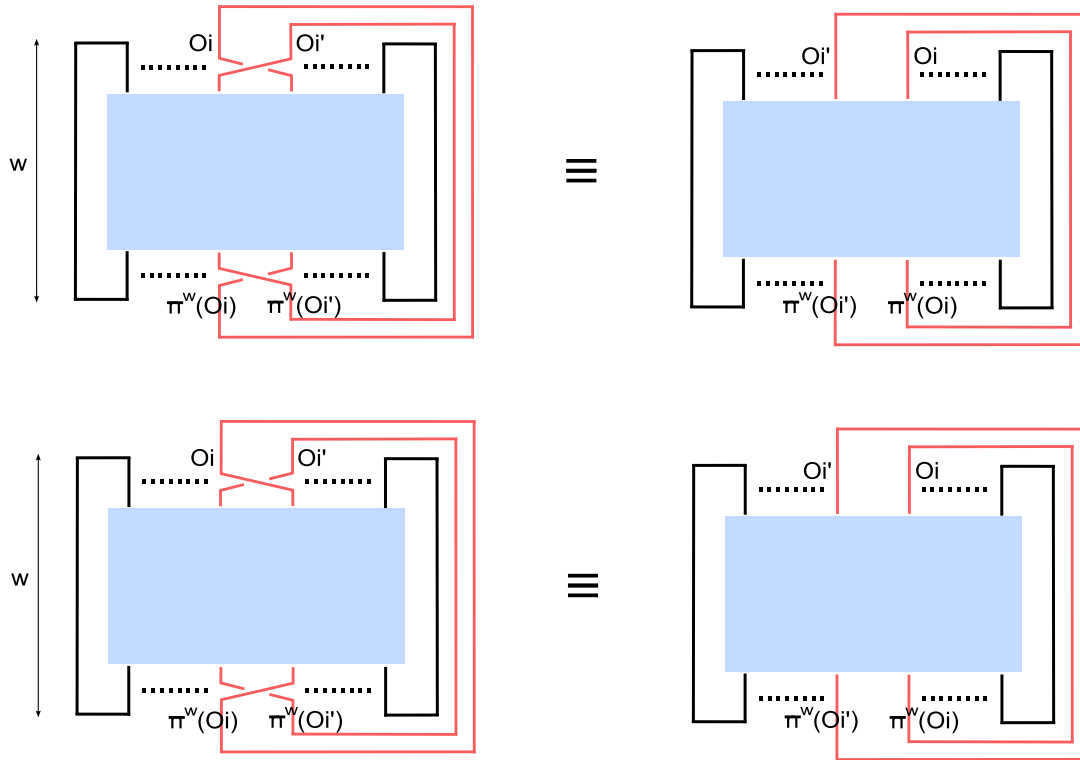


FIGURE 7.23 – Mouvement de Markov $M1$ dans le cadre des paires de brins nommés

Propriété 7.4.2 Détection d'un brin "superflu" selon un mouvement de Markov $M2$ dans le cadre de la modélisation relationnelle des noeuds

Soit w un mot de noeud tressé en termes de paires de brins nommés et notons par $|w|$ la taille de ce mot.

Soit π^w la permutation engendrée par le noeud tressé et soit $\left\{ \begin{pmatrix} O^i \\ \pi^w(O^i) \end{pmatrix}; i = 1, \dots, N \right\}$ les paires de brins nommés reliés par la clôture de la tresse.

Supposons de plus dans un premier temps, que O^i est le brin rangé en première position pour l'ordre strict et total L^1 . Nous avons donc l'hypothèse que $O^i = [L_{11}^1]_1$. O^i est alors un brin "superflu" au sens du mouvement de Markov $M2$ (d'un brin en première position) si nous avons les conditions suivantes qui sont séquentiellement vérifiées :

1. il existe dans w une unique paire de brins nommés et juxtaposés $(O^i, \pi^w(O^i))$
2. supposons alors que la paire de brins $(O^i, \pi^w(O^i))$ corresponde au k -ème élément de w alors il faut vérifier de plus que :
 - O^i n'appartient pas au support du sous-mot $(.,.)^1 \dots (.,.)^{k-1}$
 - $\pi^w(O^i)$ n'appartient pas au support du sous-mot $(.,.)^{k+1} \dots (.,.)^{|w|}$

Supposons dans un deuxième temps, que O^i est le brin rangé en dernière position pour l'ordre strict et total $L^{|W|}$. Nous avons donc l'hypothèse que $O^i = [L_{N-1, N-1}^{|W|}]_2$. O^i est alors un brin "superflu" au sens du mouvement de Markov $M2$ (d'un brin en dernière position) si nous avons les conditions suivantes qui sont séquentiellement vérifiées :

1. il existe dans w une unique paire de brins nommés et juxtaposés $(\pi^w(O^i), O^i)$

2. supposons alors que la paire de brins $(\pi^w(O^i), O^i)$ corresponde au k -ème élément de w alors il faut vérifier de plus que :
- O^i n'appartient pas au support du sous-mot $(.,.)^1 \dots (.,.)^{k-1}$
 - $\pi^w(O^i)$ n'appartient pas au support du sous-mot $(.,.)^{k+1} \dots (.,.)^{|w|}$

Nous illustrons là aussi, dans la figure fig-7.24, les conditions données par prop-7.4.2.

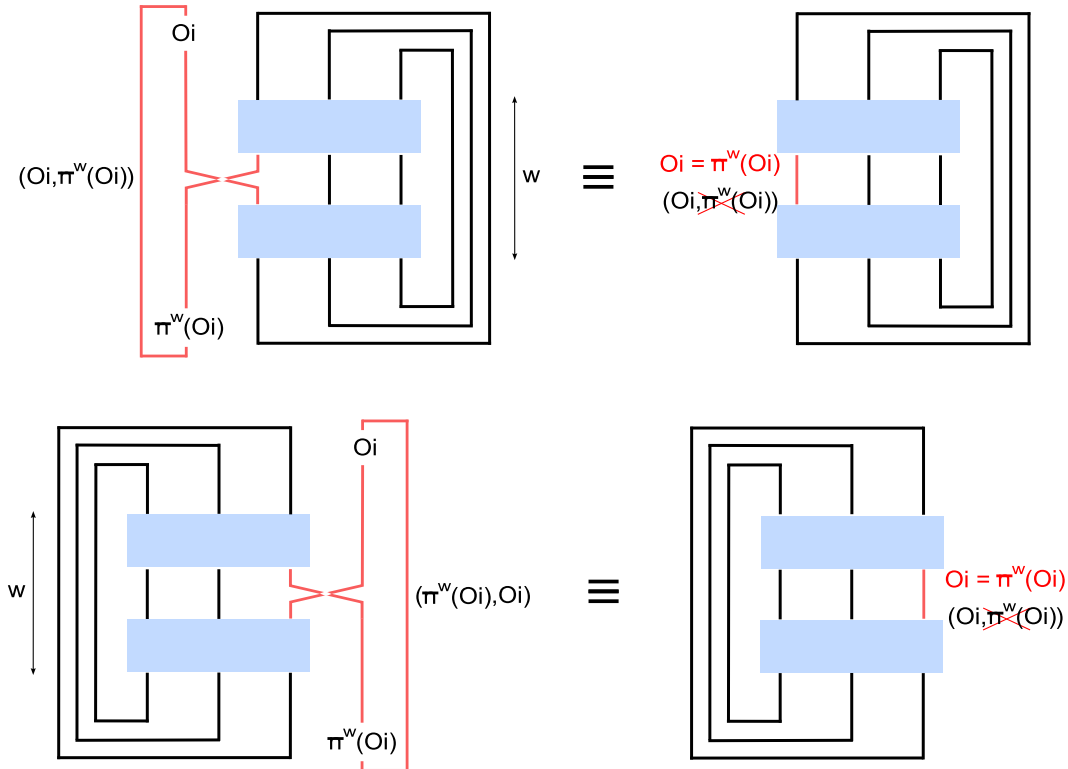


FIGURE 7.24 – Mouvement de Markov $M2$ dans le cadre des paires de brins nommés

Les mouvements de Markov $M1$ permettent des réductions libres. Ainsi, pour prendre en compte cette réduction, il suffit d'enlever les composantes $(L^1, (.,.)^1)$ et $(L^{|w|}, (.,.)^{|w|})$ de W dont la taille diminue ainsi de 2 éléments. Le premier élément devient alors le couple $(L^2, (.,.)^2)$ et l'élément final le couple $(L^{|w|-1} * (.,.)^{|w|-1}, \epsilon) = (L^{|w|}, \epsilon)$. En ce qui concerne, les mouvements de Markov $M2$, ils permettent de supprimer une paire de brins nommés dans w mais aussi, ce qui est nouveau dans le cadre de notre modélisation, la suppression d'un brin "superflu" au sens où nous passons d'une représentation dans \mathbb{B}_N à \mathbb{B}_{N-1} . Nous montrons ci-dessous comment nous mettons à jour la réduction libre et la suppression d'un brin "superflu" lors d'un mouvement $M2$.

Propriété 7.4.3 *Suppression d'un brin "superflu" selon un mouvement de Markov $M2$ dans le cadre de la modélisation relationnelle des noeuds*

Plaçons nous dans le contexte décrit dans prop-7.4.2 et supposons de plus que les conditions pour la détection d'un brin "superflu" en première position (resp. dernière position) soient remplies. Il faut alors effectuer les opérations suivantes afin de prendre en compte dans le cadre de notre codage, la suppression du brin "superflu" en première position (resp. dernière position) :

1. supprimer dans W le k -ème élément correspondant à une diminution d'une unité de la taille du mot de noeud tressé
2. dans chaque matrice couple $(L, (.,.))$ de W remplacer O^i par $\pi^w(O^i)$. Ceci correspond au fait que O^i et $\pi^w(O^i)$ sont un seul et même brin du noeud tressé. Par ailleurs, cette opération matérialise

la disparition su brin O^i

3. pour chaque matrice L de paires de brins nommés de W , supprimer la première ligne (resp. la dernière colonne). Ceci matérialise le passage d'une représentation de \mathbb{B}_N à \mathbb{B}_{N-1}

Nous venons de montrer comment il était possible, dans la perspective de la simplification d'un noeud tressé, de prendre en compte les mouvements de Markov. Nous appliquons les différentes propriétés exposées ici dans la sous-section suivante dans laquelle, nous proposons une méthode qui permettrait dans les meilleurs cas, de déterminer un invariant complet d'un noeud.

7.4.2 Détermination d'un invariant complet pour les noeuds, exemple d'application

Nous montrons comment nous pouvons à partir de notre modélisation par paires de brins nommés, aider au calcul d'un invariant complet des noeuds. Cet invariant est basé sur les tresses comme de nombreux invariants actuels tels que les polynômes de Jones. Toutefois, celui auquel nous cherchons à contribuer n'est pas un polynôme mais une forme canonique du mot de noeud tressé. Il a été défini par T. Gittings [Git04] :

Définition 7.4.2 Tresse minimale d'un noeud comme invariant complet

Parmi l'ensemble des mots de tresse représentant un noeud ou un entrelac, la tresse minimale d'un noeud est celle dont le mot représentatif w^* vérifie, dans l'ordre d'importance énoncé, les 4 propriétés suivantes :

1. w^* a le nombre de croisements minimal
2. w^* a un support de taille minimale, c'est à dire faisant intervenir le minimum de brins
3. l'univers de w^* est minimal
4. le code binaire de la nature topologique des croisements des éléments de w^* est minimal

Les critères permettant de définir la tresse minimale d'un noeud font intervenir les notions d'univers et de code binaire de la nature topologique des croisements d'un mot de tresse.

L'univers d'un mot de tresse fut initialement introduit par L.H. Kauffman [Kau83] et consiste à attribuer un entier à un générateur $(\sigma^j)^{\pm 1}$ indépendamment de la nature topologique de ce dernier. Nous prendrons la correspondance $(\sigma^j)^{\pm 1} \leftrightarrow j$ ce qui indique que le mot de tresse favorisant les générateurs d'indice "petit" en début de mot sont davantage minimaux que les mots de tresse favorisant les générateurs d'indice "grand" en début de mot.

De plus, le code binaire qui permet de prendre en compte l'information topologique, indique qu'il faut de même favoriser les générateurs d'un certain type en début de mot. Si nous prenons, $\forall j = 1 \dots, N, (\sigma^j)^{+1} \leftrightarrow 0$ et $(\sigma^j)^{-1} \leftrightarrow 1$, nous voyons que les croisements du brin en position j "au-dessous" du brin en position $j + 1$ sont favorisés.

Si nous considérons un mot de tresse et que nous cherchons à le rendre minimal par rapport aux critères 3. et 4. uniquement, nous voyons qu'il existe un unique mot de tresse équivalent qui minimise ces deux critères. Par exemple $w = \sigma^1 \circ \sigma^2 \circ (\sigma^1)^{-1}$ à un univers égal à 121 et un code binaire de 001. Nous savons en raison des relations d'Artin que w est équivalent à $(\sigma^2)^{-1} \circ \sigma^1 \circ \sigma^2$. Or ce dernier a un univers qui vaut 212 et un code binaire 100. Selon ces deux critères, nous avons $\sigma^1 \circ \sigma^2 \circ (\sigma^1)^{-1} < (\sigma^2)^{-1} \circ \sigma^1 \circ \sigma^2$.

Nous constatons, que ces deux critères qui tiennent compte d'une part de l'indice des générateurs, d'autre part, de leurs natures topologiques, peuvent être rapprochés de la donnée d'un ordre strict sur l'ensemble des générateurs $(\sigma^j)^{\pm 1}$. En effet, les critères 3. et 4. sont équivalents à la donnée d'un ordre strict et total permettant de comparer deux mots de tresse appartenant à une même classe d'équivalence. Nous pouvons donner un ordre distinct de ce dernier. Si nous posons $\sigma^1 < (\sigma^1)^{-1} < \sigma^2 < \dots < \sigma^{N-1} < (\sigma^{N-1})^{-1}$ alors nous disposons de la même manière d'un unique critère permettant de déterminer dans la classe d'équivalence d'un mot de tresse un élément unique qui est l'élément minimal. Nous constatons donc, à la suite des propos précédents, que la combinaison des propriétés 1., 3. et 4. est équivalente à la recherche d'une forme canonique "shortlex" d'un mot de tresse. Ainsi, même si l'unique mot de tresse "shortlex" obtenu est différent du mot de tresse optimisant l'ordre engendré par 3. et 4., les approches sont similaires.

Nous venons de considérer le rapprochement de la recherche de la tresse minimale d'un noeud, de la détermination d'une forme canonique "shortlex".

Il nous reste à discuter de la propriété 2. de la définition déf-7.4.2. En effet, celle-ci est spécifique aux noeuds et correspond au fait, que la représentation tressée d'un noeud peut aboutir à des brins "superflus", comme nous l'avions représentée en introduction sur la figure fig-1.5. Ce sont en fait les mouvements de Markov M_2 qui nous permettent de supprimer ces brins et nous avons montré précédemment comment les prendre en compte dans le cadre de notre modélisation.

En somme, le calcul de l'invariant de la tresse minimale d'un noeud engendre une extension du problème de la détermination de la forme canonique "shortlex" d'une tresse dans laquelle, les mouvements de Markov sont des mouvements isotopiques supplémentaires aux mouvements de brins libres qu'il faut intégrer. Nous parlerons alors de forme canonique "shortlex" d'un noeud tressé.

Nous pensons que le problème du mot "shortlex" du noeud tressé, est dans une certaine mesure, au plus aussi difficile que le problème du mot "shortlex" d'une tresse. Ceci paraît paradoxal mais en effet, la possibilité pour les noeuds de déplacer une paire de brins nommés en début de mot vers la fin de mot peut aider à débloquer certaines situations pour lesquelles nous sommes sur un "optimum local". Nous précisons ces propos dans la définition suivante :

Définition 7.4.3 *Passage d'une paire de brins nommés du début vers la fin d'un mot de noeud tressé*

Soit w un mot de paires de brins nommés représentant un noeud tressé. Soit π^w la permutation engendrée par w sur les brins du noeud tressé.

En raison de la clôture d'une tresse, un mot de noeud tressé n'a "ni de début, ni de fin". En particulier, nous pouvons déplacer une paire de brins nommés du début d'un mot à la fin de celui-ci.

Supposons que $(\cdot, \cdot)^1 = (O^i, O^{i'})$ et gardons en mémoire cette paire de brins nommés particulière, alors nous pouvons déplacer ce croisement en fin de mot en effectuant les étapes suivantes :

1. $(L^k, (\cdot, \cdot)^k) \leftarrow (L^{k+1}, (\cdot, \cdot)^{k+1}) \quad \forall k = 1, \dots, |w| - 1$
2. $L^{|w|} \leftarrow L^{|W|}$ où nous rappelons que $L^{|W|} = L^{|w|+1}$ est la matrice de paires de brins nommés représentative de l'ordre strict et total ou de la permutation⁴⁸, engendrée par les croisements de w
3. $(\cdot, \cdot)^{|w|} = (\pi^w(O^i), \pi^w(O^{i'}))$ qui traduit le passage de la paire de brins nommés du début vers la fin de mot de noeud tressé
4. $L^{|W|} \leftarrow L^{|w|} * (\pi^w(O^i), \pi^w(O^{i'}))$ est le nouveau résultat d'ordre strict et total ou de permutation engendrée par le "nouveau" mot de noeud tressé qui est équivalent au mot précédent

48. correspondance biunivoque entre un ordre strict et total et une permutation donnée sous forme de liste cf déf-5.1.7 page 160

Nous proposons le procédé suivant de calcul du mot “shortlex” d’un noeud tressé qui est donc en correspondance biunivoque avec l’invariant défini par T. Gittings. Nous illustrons ensuite ce procédé sur un exemple d’application.

Algorithm 8 Démêlage d’une tresse par recherche avec liste “tabou”

Require: liste

- Fin=Faux

while Fin=Faux **do**

while La condition d’arrêt n’est pas vérifiée **do**

 - Rechercher un mouvement de Markov $M1$ ou $M2$ permettant de diminuer la taille du mot de noeud tressé ou de supprimer un brin “superflu”

 - Si un tel mouvement existe, mettre à jour les matrices de paires de brins nommés qui changent après application d’un mouvement de Markov

 - Rechercher un mouvement isotopique de brin libre de type “poignée” permettant de diminuer la taille du mot de noeud tressé au sens de la réduction libre ou au sens de la réduction selon un mouvement de Markov $M1$

 - Si un tel mouvement existe, mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

if S’il n’y a pas de mouvement isotopique ou de Markov diminuant la taille du mot de tresse **then**

 - Rechercher un mouvement isotopique de brin libre de type “poignée” n’augmentant pas la taille du mot de tresse ou une relation d’Artin, dont le mot de tresse obtenu soit distinct des mots de tresse de la liste

 - Si un tel mouvement existe, ajouter à la liste le mot de tresse de taille constante obtenu et mettre à jour les matrices de paires de brins nommés qui changent après application du mouvement de brin libre

end if

end while

 - Prendre comme mot de tresse celui de la liste qui est la plus optimale au sens de l’ordre “shortlex”

while Tant que nous n’obtenons pas de mot distinct des mots de la liste et que Fin=Faux **do**

 - Passer la paire de brins nommés en début de mot de noeud tressé en fin de mot

 - Vérifier si on a fait un cycle complet si oui alors Fin←Vrai

end while

 - Si Fin=Faux alors ajouter le mot obtenu dans la liste

end while

- Retourner le mot de noeud tressé de la liste “tabou” dont la forme canonique “shortlex” est la plus petite

Nous donnons dans fig-7.25 une représentation graphique du procédé que nous venons de décrire.

Nous illustrons à présent ce procédé sur un exemple d’application. Celui-ci est représenté dans la figure fig-7.26 et correspond au mot de tresse suivant :

$$w = bABBCBca \quad (7.4.3)$$

En termes de paires de brins nommés nous avons la définition suivante :

$$w = \underline{(O^2, O^3)} \overline{(O^1, O^3)} \overline{(O^1, O^2)} \overline{(O^2, O^1)} \underline{(O^2, O^4)} \overline{(O^1, O^4)} \underline{(O^1, O^2)} \underline{(O^3, O^4)} \quad (7.4.4)$$

Il n’existe pas de brin libre dont le mouvement isotopique permettrait de diminuer le nombre de croisements de w au sens de la réduction libre des tresses. Toutefois, il existe un mouvement de brin libre isotopique conduisant à une réduction libre au sens d’un mouvement de Markov $M1$.

Ce mouvement concerne le déplacement à gauche du brin libre O^2 , $\xleftarrow{O^2}$. Nous avons plus



FIGURE 7.25 – Procédé permettant de calculer dans le meilleur des cas la forme canonique “shortlex” d’un mot de noeud tressé

précisément la définition suivante de ce brin libre :

$$O^2vO^2 = \underbrace{(O^2, O^3)}_v \overline{(O^1, O^3)} \overline{(O^1, O^2)} \quad (7.4.5)$$

Ce cas est le cas le plus simple d’un mouvement de brin libre. En effet, il se réduit au cas particulier d’une relation d’Artin $R3$ que nous avons présenté en prop-7.2.6 page 229 dans le cadre des paires de brins nommés. Nous l’étudions néanmoins dans le contexte d’un déplacement de brin libre afin d’illustrer à nouveau le mécanisme des mouvements isotopiques des brins libres.

Nous avons, $u = \overline{(O^2, O^3)}\overline{(O^1, O^3)}\overline{(O^1, O^2)}$, $v = \overline{(O^1, O^3)}$, $i^c = 3$ et $j^i = j^2 = 3$. Nous avons par ailleurs, les matrices de paires de brins nommés L^2 et $L^{|U|} = L^3$ représentatives respectivement des ordres $O^1 < O^3 < O^2 < O^4$ et $O^3 < O^1 < O^2 < O^4$ qui valent :

$$L^2 = \begin{pmatrix} [O^1, O^3] & [O^1, O^2] & [O^1, O^4] \\ & [O^3, O^2] & [O^3, O^4] \\ & & [O^2, O^4] \end{pmatrix} ; \quad L^3 = \begin{pmatrix} [O^3, O^1] & [O^3, O^2] & [O^3, O^4] \\ & [O^1, O^2] & [O^1, O^4] \\ & & [O^2, O^4] \end{pmatrix} \quad (7.4.6)$$

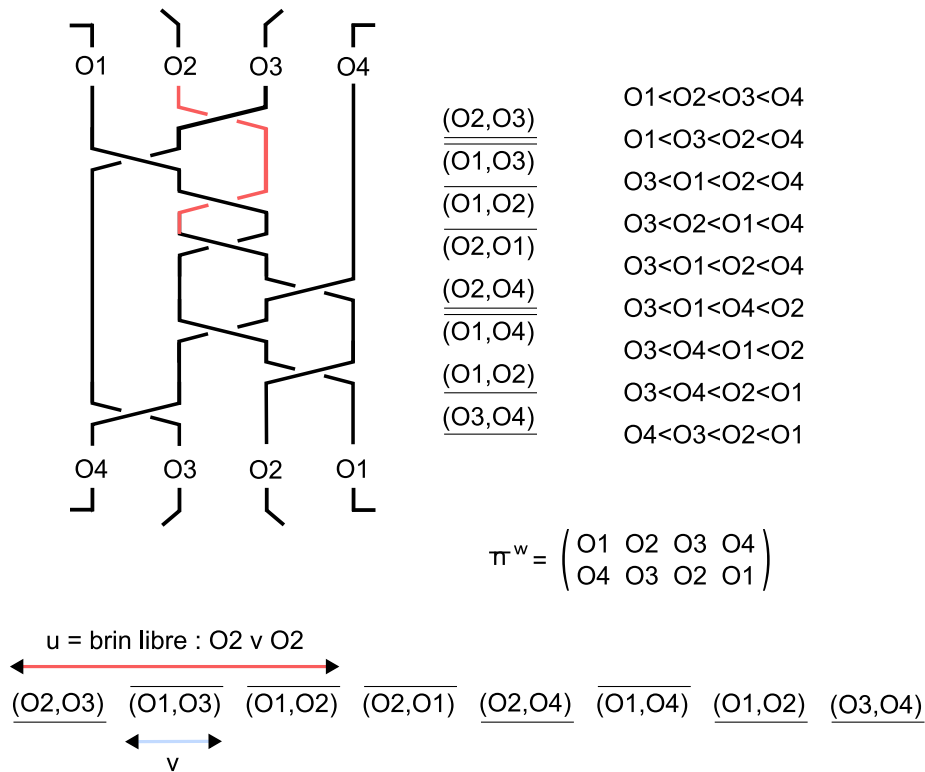


FIGURE 7.26 – Exemple d’application : calcul de la forme canonique “shortlex” d’un mot de noeud tressé

Il s’agit d’un déplacement à gauche et nous avons alors des chemins de type “va et vient verticaux”. v étant épuré, nous avons $\text{Supp}(v) = \{O^1, O^3\}$ et donc $|d| = |a| = \#\text{Supp}(v) = 2$. Nous avons alors : $d = (O^3, O^2)(O^1, O^2)$ et $a = r((O^3, O^2))r((O^1, O^2)) = (O^2, O^3)(O^2, O^1)$.

Considérons la matrice relationnelle d’ordre topologique engendrée par $v = \overline{(O^1, O^3)}$:

$$C^{v,o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \end{matrix} \quad (7.4.7)$$

Pour ne pas créer de nouveaux croisements supplémentaires, il faut réduire librement les croisements en amont et en aval de v c’est à dire $\overline{(O^2, O^3)}$ et $\overline{(O^1, O^2)}$. Ceci implique que le premier terme de d et le dernier terme de a soient $\overline{(O^3, O^2)}$ et $\overline{(O^2, O^1)}$. Ceci a pour conséquence une signature $\alpha = (1, 1)$ dans la mesure où, en raison de la symétrie topologique, nous avons $a(\alpha) = \overline{(O^2, O^3)}(O^2, O^1)$ et $d[a(\alpha)] = \overline{(O^3, O^2)}(O^1, O^2)$.

Il est facile de voir que la signature $\alpha = (1, 1)$ engendre, un mouvement isotopique “au-dessous” qui n’est donc aucunement en discordance avec $C^{v,o}$. En effet, nous avons :

$$C^{\alpha,o} = \begin{matrix} & O^1 & O^2 & O^3 & O^4 \\ \begin{matrix} O^1 \\ O^2 \\ O^3 \\ O^4 \end{matrix} & \begin{pmatrix} 1 & 0 & \square & \square \\ 1 & 1 & 1 & \square \\ \times & 0 & 1 & \square \\ \square & \square & \square & 1 \end{pmatrix} \end{matrix} \quad (7.4.8)$$

Nous avons alors l'coalement, le mot u équivalent suivant dont la taille reste constante :

$$\begin{aligned} u &= \overline{(O^2, O^3)} \overline{(O^1, O^3)} \overline{(O^1, O^2)} \\ &\equiv \overline{(O^2, O^3)} [\overline{(O^3, O^2)} \overline{(O^1, O^2)}] \overline{(O^1, O^3)} [\overline{(O^2, O^3)} \overline{(O^2, O^1)}] \overline{(O^1, O^2)} \\ &\equiv \overline{[(O^1, O^2)] \overline{(O^1, O^3)} \overline{(O^2, O^3)}} \end{aligned} \quad (7.4.9)$$

(7.4.10)

Nous avons à la suite de ce mouvement, le mot de tresse équivalent ci-dessous :

$$w \equiv \overline{[(O^1, O^2)] \overline{(O^1, O^3)} \overline{(O^2, O^3)} \overline{(O^2, O^1)} \overline{(O^2, O^4)} \overline{(O^1, O^4)} \overline{(O^1, O^2)} \overline{(O^3, O^4)}} \quad (7.4.11)$$

L'intérêt de ce mouvement réside dans le fait que ce dernier permet de faire apparaître une réduction libre au sens d'un mouvement de Markov $M1$.

En effet, considérons de plus, π^w , le résultat de permutation engendrée par w . Nous avons :

$$\pi^w = \begin{pmatrix} O^1 & O^2 & O^3 & O^4 \\ O^4 & O^3 & O^2 & O^1 \end{pmatrix} \quad (7.4.12)$$

L'ensemble des paires de brins nommés reliés est alors le suivant :

$$\left\{ \begin{pmatrix} O^1 \\ O^4 \end{pmatrix}, \begin{pmatrix} O^2 \\ O^3 \end{pmatrix}, \begin{pmatrix} O^3 \\ O^2 \end{pmatrix}, \begin{pmatrix} O^4 \\ O^1 \end{pmatrix} \right\} \quad (7.4.13)$$

Nous faisons remarquer, que nous savons d'ores et déjà qu'il s'agit d'un entrelac composé de deux noeuds étant donné que la permutation se décompose en deux cycles⁴⁹ : $\pi^w = (O^1, O^4)(O^2, O^3)$ qui n'est pas à confondre avec des paires de brins nommés.

Nous avons alors, la relation suivante⁵⁰ qui permet de détecter un mouvement de Markov $M1$:

$$(\cdot, \cdot)^1 = \overline{(O^1, O^2)} \quad \text{et} \quad (\cdot, \cdot)^{|w|} = \overline{(\pi^w(O^2), \pi^w(O^1))} = \overline{(O^3, O^4)} \quad (7.4.14)$$

L'application de la réduction libre issue du mouvement de Markov $M1$ nous permet d'avoir le mot de noeud tressé équivalent ci-dessous pour lequel nous précisons que l'état initial représenté par les matrices de paires de brins nommés est désormais L^2 :

$$w \equiv \overline{(O^1, O^3)} \overline{[(O^2, O^3)] \overline{(O^2, O^1)} \overline{(O^2, O^4)} \overline{(O^1, O^4)} \overline{(O^1, O^2)}} \quad (7.4.15)$$

Nous représentons le résultat de cette simplification dans la figure fig-7.27.

Nous avons réussi à réduire la taille du mot de noeud tressé et nous faisons remarquer que les états initiaux et finaux ont changé mais pas l'ensemble des paires de brins nommés reliés.

Nous recommençons alors une nouvelle boucle afin de tester à nouveau des mouvements permettant de réduire la taille du mot. Nous nous apercevons alors, que nous pouvons supprimer un brin "superflu". En effet, O^2 est le brin rangé en première position pour L^1 et nous avons les conditions suivantes qui sont vérifiées :

1. il existe dans w une unique paire de brins nommés et juxtaposés $(O^2, \pi^w(O^2)) = (O^2, O^3)$

49. cf chapitre 5, déf-5.1.9 page 162

50. cf prop-7.4.1

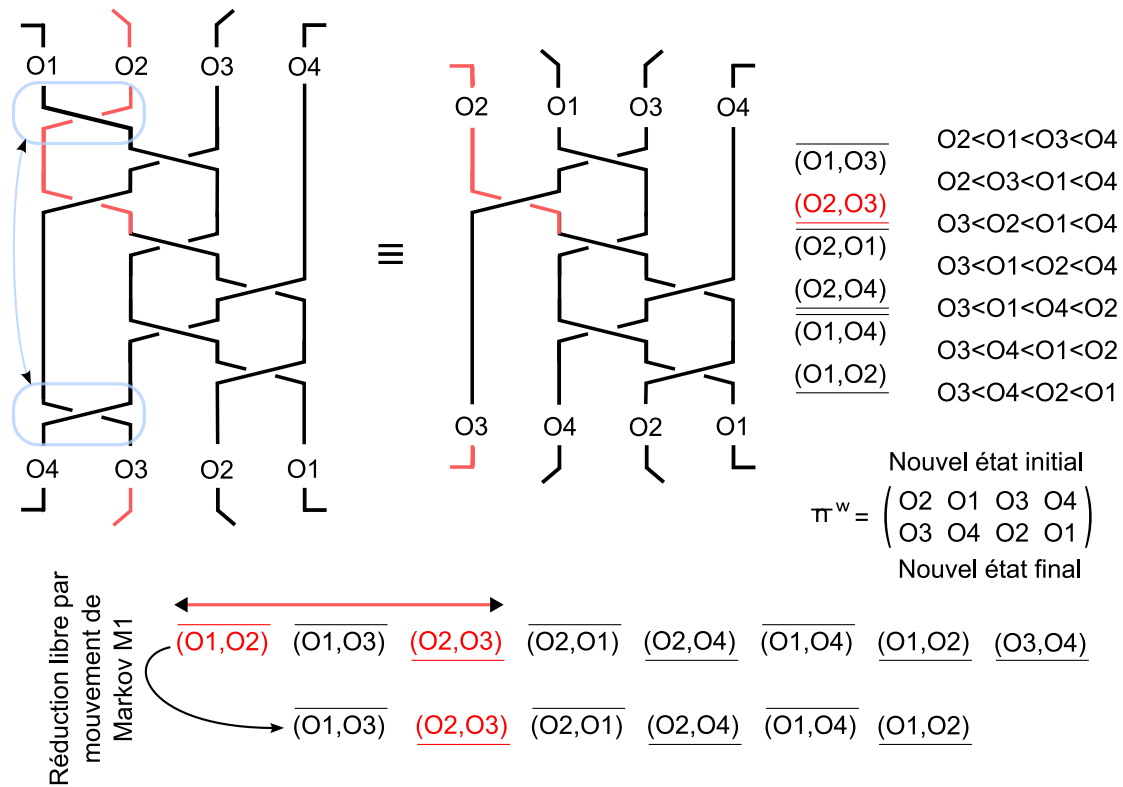


FIGURE 7.27 – Exemple d’application : résultat de la simplification par un mouvement de brin libre débouchant sur la réduction libre issue d’un mouvement de Markov M1

2. cette paire de brins nommés (O^2, O^3) correspond au 2-ème élément de w par ailleurs :
 - O^2 n’appartient pas au support du sous-mot $(\cdot, \cdot)^1 \dots (\cdot, \cdot)^{k-1} = \overline{(O^1, O^3)}$
 - $\pi^w(O^2) = O^3$ n’appartient pas au support du sous-mot $(\cdot, \cdot)^{k+1} \dots (\cdot, \cdot)^{|w|} = \overline{(O^2, O^1)} \underline{\underline{(O^2, O^4)}} \overline{(O^1, O^4)} \underline{\underline{(O^1, O^2)}}$

De ces différentes conditions, nous pouvons selon prop-7.4.2, dire que O^2 est un brin “superflu”. Nous explicitons alors quelques points de la prop-7.4.2 qui indique les étapes de la prise en compte de la suppression de ce brin “superflu” :

1. supprimer dans W le 2-ème élément correspondant à une diminution d’une unité de la taille du mot de noeud tressé. Du point de vue des seconds membres de W , nous avons : ${}^w w \equiv \overline{(O^1, O^3)} \underline{\underline{(O^2, O^1)}} \underline{\underline{(O^2, O^4)}} \overline{(O^1, O^4)} \underline{\underline{(O^1, O^2)}}$
2. pour chaque paire de brins nommés des couples $(L, (\cdot, \cdot))$ de W remplacer O^2 par $\pi^w(O^2) = O^3$. Nous avons les différentes images suivantes des matrices de paires de brins :

$$\begin{aligned}
 L^1 &= \begin{pmatrix} [O^2, O^1] & [O^2, O^3] & [O^2, O^4] \\ & [O^1, O^3] & [O^1, O^4] \\ & & [O^3, O^4] \end{pmatrix} \text{ et } (\cdot, \cdot)^1 = \overline{(O^1, O^3)} \\
 \rightarrow L^1 &= \begin{pmatrix} [O^3, O^1] & [O^3, O^3] & [O^3, O^4] \\ & [O^1, O^3] & [O^1, O^4] \\ & & [O^3, O^4] \end{pmatrix} \text{ et } (\cdot, \cdot)^1 = \overline{(O^1, O^3)}
 \end{aligned} \tag{7.4.16}$$

$$\begin{aligned}
 L^2 &= \begin{pmatrix} [O^3, O^2] & [O^3, O^1] & [O^3, O^4] \\ & [O^2, O^1] & [O^2, O^4] \\ & & [O^1, O^4] \end{pmatrix} \text{ et } (\cdot, \cdot)^2 = \overline{(O^2, O^1)} \\
 \rightarrow L^2 &= \begin{pmatrix} [O^3, O^3] & [O^3, O^1] & [O^3, O^4] \\ & [O^3, O^1] & [O^3, O^4] \\ & & [O^1, O^4] \end{pmatrix} \text{ et } (\cdot, \cdot)^2 = \overline{(O^3, O^1)}
 \end{aligned} \tag{7.4.17}$$

$$\begin{aligned}
 L^3 &= \begin{pmatrix} [O^3, O^1] & [O^3, O^2] & [O^3, O^4] \\ & [O^1, O^2] & [O^1, O^4] \\ & & [O^2, O^4] \end{pmatrix} \text{ et } (\cdot, \cdot)^3 = \underline{(O^2, O^4)} \\
 \rightarrow L^3 &= \begin{pmatrix} [O^3, O^1] & [O^3, O^3] & [O^3, O^4] \\ & [O^1, O^3] & [O^1, O^4] \\ & & [O^3, O^4] \end{pmatrix} \text{ et } (\cdot, \cdot)^3 = \underline{(O^3, O^4)}
 \end{aligned} \tag{7.4.18}$$

$$\begin{aligned}
 L^4 &= \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^2] \\ & [O^1, O^4] & [O^1, O^2] \\ & & [O^4, O^2] \end{pmatrix} \text{ et } (\cdot, \cdot)^4 = \overline{(O^1, O^4)} \\
 \rightarrow L^4 &= \begin{pmatrix} [O^3, O^1] & [O^3, O^4] & [O^3, O^3] \\ & [O^1, O^4] & [O^1, O^3] \\ & & [O^4, O^3] \end{pmatrix} \text{ et } (\cdot, \cdot)^4 = \overline{(O^1, O^4)}
 \end{aligned} \tag{7.4.19}$$

$$\begin{aligned}
 L^5 &= \begin{pmatrix} [O^3, O^4] & [O^3, O^1] & [O^3, O^2] \\ & [O^4, O^1] & [O^4, O^2] \\ & & [O^1, O^2] \end{pmatrix} \text{ et } (\cdot, \cdot)^5 = \underline{(O^1, O^2)} \\
 \rightarrow L^5 &= \begin{pmatrix} [O^3, O^4] & [O^3, O^1] & [O^3, O^3] \\ & [O^4, O^1] & [O^4, O^3] \\ & & [O^1, O^3] \end{pmatrix} \text{ et } (\cdot, \cdot)^5 = \underline{(O^1, O^3)}
 \end{aligned} \tag{7.4.20}$$

$$\begin{aligned}
 L^6 &= \begin{pmatrix} [O^3, O^4] & [O^3, O^2] & [O^3, O^1] \\ & [O^4, O^2] & [O^4, O^1] \\ & & [O^2, O^1] \end{pmatrix} \text{ et } (\cdot, \cdot)^6 = \epsilon \\
 \rightarrow L^6 &= \begin{pmatrix} [O^3, O^4] & [O^3, O^3] & [O^3, O^1] \\ & [O^4, O^3] & [O^4, O^1] \\ & & [O^3, O^1] \end{pmatrix} \text{ et } (\cdot, \cdot)^6 = \epsilon
 \end{aligned} \tag{7.4.21}$$

3. pour chaque matrice L de paires de brins nommés de W , supprimer la première ligne. Ceci matérialise le passage d'une représentation de \mathbb{B}_4 à \mathbb{B}_3 . Nous avons finalement le résultat suivant :

$$\begin{aligned}
 W \equiv & \left(\left(\begin{pmatrix} [O^1, O^3] & [O^1, O^4] \\ & [O^3, O^4] \end{pmatrix}, \overline{(O^1, O^3)} \right), \left(\begin{pmatrix} [O^3, O^1] & [O^3, O^4] \\ & [O^1, O^4] \end{pmatrix}, \overline{(O^3, O^1)} \right) \right) \\
 & \left(\left(\begin{pmatrix} [O^1, O^3] & [O^1, O^4] \\ & [O^3, O^4] \end{pmatrix}, (O^3, O^4) \right), \left(\begin{pmatrix} [O^1, O^4] & [O^1, O^3] \\ & [O^4, O^3] \end{pmatrix}, \overline{(O^1, O^4)} \right) \right) \\
 & \left(\left(\begin{pmatrix} [O^4, O^1] & [O^4, O^3] \\ & [O^1, O^3] \end{pmatrix}, \overline{(O^1, O^3)} \right), \left(\begin{pmatrix} [O^4, O^3] & [O^4, O^1] \\ & [O^3, O^1] \end{pmatrix}, \epsilon \right) \right)
 \end{aligned} \tag{7.4.22}$$

Nous avons représenté sur la figure fig-7.28, la suppression du brin "superflu" O^2 .

Il n'existe plus ensuite ni de mouvements isotopiques de brins libres, ni de mouvements de Markov, ni de relations d'Artin, ni de passage de paires de brins nommés du début vers la fin du mot, qui permettent de réduire la taille du mot de noeud tressé obtenu ou d'améliorer sa forme "shortlex". Nous en déduisons alors que le mot de noeud tressé suivant est l'"optimal" au sens "shortlex" :

$$w^* = \overline{(O^1, O^3)} \overline{(O^3, O^1)} \underline{(O^3, O^4)} \overline{(O^1, O^4)} \underline{(O^1, O^3)} \tag{7.4.23}$$

En termes de générateurs classiques, nous obtenons le mot ci-dessous :

$$w^* = AAbAb \tag{7.4.24}$$

Nous avons vérifié que le mot de noeud tressé "shortlex" obtenu était en correspondance avec les tables d'invariants complets de noeuds donnés par T. Gittings dans [Git04]. La forme

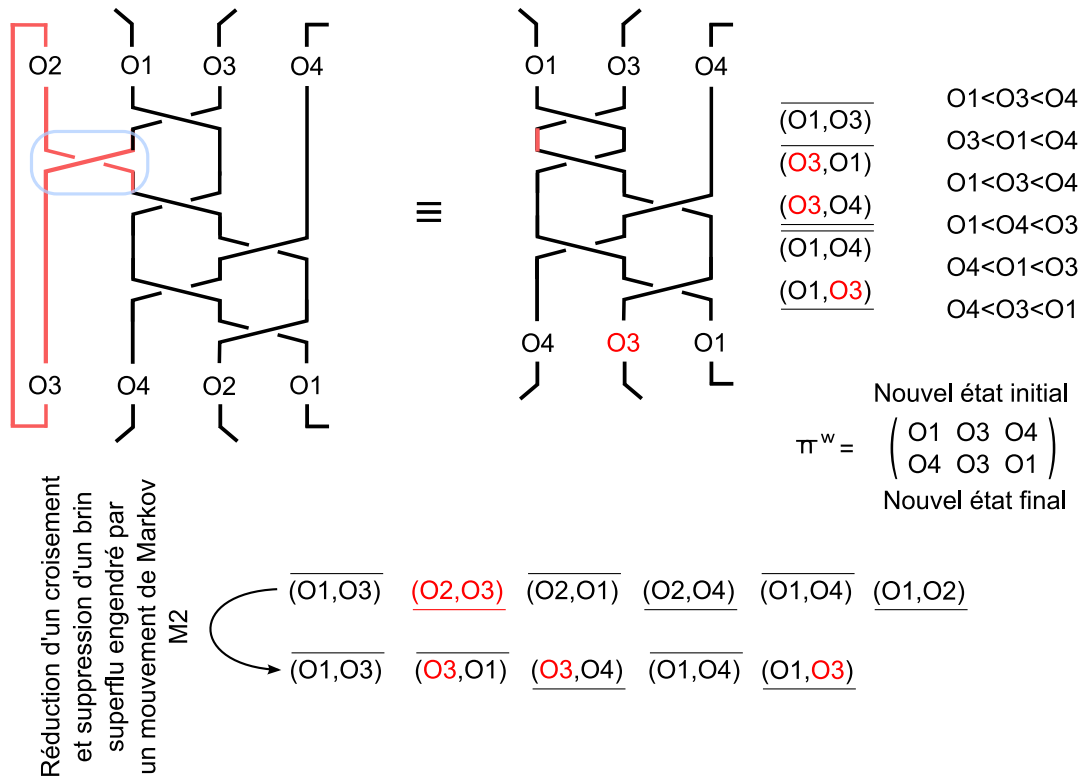


FIGURE 7.28 – Exemple d’application : résultat de la simplification un mouvement de Markov M_2 , obtention de la forme “shortlex” du mot de noeud tressé

obtenue est finalement la même que celle donnée par T. Gittings et correspond dans l’indexation établie par ce dernier au noeud référencé 5 : 2 – 01.

T. Gittings a donné la définition d’un invariant complet pour les noeuds qui est simple mais pas forcément facile à déterminer. D’ailleurs, comme il le précise lui-même, il ne donne pas de méthode permettant, à partir d’un diagramme tressé d’un noeud obtenu par l’algorithme de P. Vogel [Vog90] par exemple, de calculer cet invariant.

Nous ne savons pas si le procédé proposé ici permet rigoureusement de déterminer la forme “shortlex” d’un mot de noeud tressé ou s’il s’agit d’une heuristique donnant “au pire” une solution approchée et “au mieux” la solution optimale. Toutefois, nous pensons que les outils et les propriétés présentés pourraient, s’il s’avérait qu’il s’agissait effectivement d’une heuristique, contribuer *in fine* à la définition d’une méthode exacte.

CHAPITRE 8

Conclusion et perspectives

Nous avons abordé au cours de ce mémoire de thèse divers domaines mathématiques que nous souhaitons rappeler succinctement avec une brève mention des résultats obtenus.

Nous nous sommes tout d'abord intéressés aux statistiques appliquées et en particulier :

- à l'étude théorique des principales méthodes de classification automatique qui a donné lieu à la définition d'un nouvel algorithme aux propriétés avantageuses
- à une généralisation des indices de similarité qui a conduit à la définition des indices d'ordre t dont l'interprétation en termes géométriques est valable pour des données numériques, binaires, qualitatives, quantitatives ou relationnelles
- aux critères d'association entre variables qualitatives avec une étude et un éclaircissement apportés quant aux diverses notions d'indépendances¹ sous-jacentes
- aux critères d'association entre variable ordonnées avec une unification des formalismes relationnels entre relation d'équivalence et relations d'ordre au travers de la généralisation de l'indépendance logique ou indétermination

Nous avons ensuite rappelé l'algèbre des relations et les fonctions d'agrégation, qui nous ont conduit à :

- définir une méthode générique permettant de déterminer les contraintes linéaires de propriétés s'énonçant sous forme de proposition logique. Ceci nous a permis d'enrichir les contraintes linéaires des propriétés de relations de l'AR en déterminant par exemple les contraintes linéaires de la propriété de Ferrers
- définir une méthode d'agrégation d'inspiration combinatoire, qui est fondée sur une règle de vote et qui permet par ailleurs de modéliser les interactions entre variables au travers de l'utilisation de "t-norms". Cette méthode d'agrégation dite calcul consensuel, est proposée notamment dans le cadre de l'agrégation de relations d'équivalence débou-

1. indépendance statistique, indépendance géométrique et indétermination qui est une approche logique de l'indépendance

chant ainsi sur une méthodologie de classification de données hétérogènes

Puis, nous avons étudié la théorie du choix social qui est issue de préoccupations émanant des sciences économiques :

- nous avons défini dans ce cadre un formalisme nouveau associant AR, dioïde $\{\mathbb{Z}_M, \max, \min\}$, algèbre relationnelle multivalente et algèbre matricielle
- nous avons, à l'aide de ce formalisme réussi à définir des conditions nécessaires et suffisantes sur les domaines de préférences individuelles, caractérisant les propriétés de la relation collective obtenue. En particulier, nous donnons dans ce contexte, des conditions nécessaires et suffisantes d'acyclicité et de transitivité. Celles-ci sont valables pour plusieurs règles de vote et ne nécessitent aucune hypothèse ce qui en font des conditions plus générales que celles de Black, d'Inada et de Sen

Enfin, nous avons abordé la théorie des tresses et des noeuds, pour laquelle :

- nous avons proposé une modélisation nouvelle et originale. Elle utilise le codage sous forme de comparaisons par paires de "brins nommés". Ce codage permet de mettre en évidence des relations d'ordre topologique entre les brins d'une tresse. Ce faisant, nous avons été capable de modéliser tout type de mouvement isotopique de brin généralisant d'une certaine manière les relations d'Artin
- nous avons présenté le problème de la forme canonique "shortlex" d'un mot de tresse ou d'un mot de noeud tressé comme étant des problèmes d'optimisation combinatoire. Nous avons proposé des procédés qui, toutefois, ne sont pas identifiés, à l'heure actuelle, comme étant des méthodes de résolution exacte ou approchée. Néanmoins, la modélisation proposée permettrait, nous le pensons, d'aller plus loin dans la résolution de ces problèmes si ces algorithmes s'avéraient être des méthodes approchées et non exactes

Les domaines mathématiques auxquels nous nous sommes intéressés sont d'apparence disjointes pour beaucoup d'entre eux. Nous avons montré que ce n'était pas nécessairement le cas ! Comme J.F. Marcotorchino et P. Michaud l'avaient déjà démontré à la fin des années 1970, les problèmes de classification automatique et d'agrégation et de consensus d'ordres pouvaient se formaliser de la même manière qui en l'occurrence était le modèle en programmation linéaire en nombres bivalents.

Dans ce même esprit, nous avons donc, au cours de ce travail de thèse, cherché à étendre et unifier certains résultats : c'est le cas des unifications obtenues pour les indices de similarité et les critères d'association et de partitionnement. De plus, nous avons cherché à étendre les possibilités de l'AR en allant chercher des éléments de l'analyse combinatoire, de l'algèbre relationnelle, de la logique multivalente, de la théorie des groupes et en cherchant alors à modéliser de nouveaux problèmes de manière nouvelle et originale : c'est le cas des retombées obtenues en théorie du choix social et en théorie des tresses.

Au delà des résultats obtenus, nous pensons que les idées et approches développées pourraient engendrer des études intéressantes à entreprendre. Par ailleurs, certains travaux ne sont pas encore finalisés.

Ainsi, nous terminerons par donner ci-dessous, des perspectives de recherche relatives à chacun des domaines d'application précédemment cités.

Dans le cas des mesures de "ressemblance" qu'elles soient de similarité ou d'association ou encore de corrélation, nous avons cherché à unifier sous le formalisme des indices d'ordre t certaines d'entre elles dont les structures sont soit vectorielles, soit relationnelles, soit graphiques. Cette approche ouvre parallèlement des perspectives d'études empiriques étant donné qu'il existe souvent pour un type de structures, des aspects qui n'ont pas été investigués et qui existent pour d'autres types de structures. Par exemple, le critère de Kim est dans le cas des

matrices relationnelles d'ordre, un indice d'ordre $t = 1$. Nous pourrions dans le cas de matrices relationnelles de partition ou de vecteurs quantitatifs étudier de la même manière le cas des indices d'ordre $t = 1$ ou $t > 1$. De manière plus générale, l'indétermination pondérée valable pour les matrices relationnelles mériterait aussi des études empiriques complémentaires.

En ce qui concerne le calcul consensuel, les propositions qui ont été faites afin d'utiliser cette méthode d'agrégation pour classifier des données hétérogènes devront aussi être testées sur des exemples pratiques.

Nous pouvons par ailleurs, reconsidérer le problème de l'association maximale donnant des critères de partitionnement ou de vote dans un contexte général où, l'agrégation utilisée est une fonction d'agrégation telle que le calcul consensuel.

Par ailleurs, un des avantages de ce dernier est qu'il s'agit d'une fonction d'agrégation basé sur une règle majoritaire. Son interprétation est simple, de ce fait, l'idée d'appliquer cette méthode pour le calcul de vraisemblance dans un environnement d'incertitude nous semble intéressante. En effet, supposons que pour une quelconque raison, il existe une incertitude sur la fiabilité de l'échantillon à partir duquel nous cherchons à estimer un paramètre. Supposons que nous sachions *a priori* que 5% des données ne sont pas fiables mais qu'en revanche, nous ne sachions pas de quelles données particulières il s'agit. Une solution à ce problème d'incertitude pourrait être la maximisation de la vraisemblance consensuelle avec $m = 95\%$ de M où M est la taille de l'échantillon.

Quant aux approches développées en théorie du choix social, nous pensons que la caractérisation des propriétés d'une relation collective obtenue par l'agrégation de préférences individuelles, devrait être étendue à d'autres types de propriétés relationnelles et à d'autres relations. L'axiomatique existante par exemple autour des relations d'équivalence peut aider à mieux formaliser certaines caractéristiques en classification automatique comme les partitions emboîtées. Celles-ci définissent un ensemble de familles particulières de relations d'équivalence que nous retrouvons dans certaines approches comme la classification hiérarchique et la classification conceptuelle. De manière générale, nous pensons, qu'une meilleure appréhension des propriétés relationnelles dans un cadre consensuel pourrait servir en mathématiques des sciences humaines.

Notre approche, comme nous l'avons rappelé, a été d'associer AR, dioïde $\{\mathbb{Z}_M, \max, \min\}$, algèbre relationnelle multivalente et algèbre matricielle. Nous avons montré le rôle joué par des opérateurs matriciels tels que le produit relatif, la trace et le permanent. Certains de ces opérateurs sont propres à l'algèbre relationnelle tel que le produit relatif mais d'autres ne le sont pas comme le permanent. Peut-on parler alors d'une algèbre ou d'une géométrie de la logique du consensus ?

Dans un cadre proche de ce qui vient d'être énoncé, nous pensons, que l'appréhension de l'AR sous l'angle de la théorie des graphes pourrait certainement déboucher sur des résultats théoriques intéressants. Nous pourrions considérer bon nombre de propriétés et théorèmes définies sur des graphes en général et étudier ce qu'elles donneraient pour des graphes particuliers de relations structurées comme les partitions ou les ordres.

Enfin, la modélisation relationnelle des tresses et des noeuds proposée nécessite des expérimentations. En effet, de nombreuses questions restent en suspens et l'implémentation des procédés suggérés pourrait très certainement apporter des éclaircissements.

Sur un plan théorique, comme nous l'avons déjà mentionné, l'idée des relations d'ordre topologique engendrées par des mots de tresse de paires de brins nommés mériterait d'être exploitée davantage. Nous pensons en effet, qu'elles pourraient permettre d'identifier de nouvelles propriétés encore méconnues ou peu modélisées, de ces objets mathématiques fascinants.

D'autres résultats théoriques sont attendus par ailleurs. Il faudrait en effet, arriver à démontrer que les algorithmes proposés permettent véritablement de déterminer la forme canonique "shortlex" d'une tresse ou d'un noeud tressé. Pour cela, il nous faut montrer que cette forme unique, peut toujours être atteinte à partir de tout mot de la classe d'équivalence en utilisant des mouvements isotopiques de brins libres n'augmentant pas la taille du mot. Les méthodes de démonstration développées par P.Dehornoy pourraient certainement aider dans cette optique. Si nous arrivons à démontrer ceci, alors, les algorithmes que nous avons proposés définiraient des méthodes de détermination d'un invariant complet des noeuds.

BIBLIOGRAPHIE

- [Abd96] H. Abdallah. *Application de l'analyse relationnelle pour classifier descripteurs et modalités en mode discrimination*. PhD thesis, Thèse de l'Université de Paris VI, 1996.
- [Ada04] C.C. Adams. *The knot book - An elementary introduction to the mathematical theory of knots*. American Mathematical Society, 2004.
- [AE03] G. Amerly and C. Elkan. Learning the k in k -means. In *In proceedings of the seventeenth annual conference on neural information processing systems (NIPS)*, pages 281–288, 2003.
- [Ale28] J.W. Alexander. Topological invariants of knots and links. *Transactions of the American Mathematical Society*, 30 :275–306, 1928.
- [APLB05] J. Ah-Pine, J. Lemoine, and H. Benhadda. Un nouvel outil de classification non supervisée de documents pour la découverte de connaissances et la détection de signaux faibles : Rares text. *Colloque Ile Rousse*, 2005.
- [APM07] J. Ah-Pine and J.F. Marcotorchino. Statistical, geometrical and logical independences between categorical variables. *Applied Stochastic Models and Data Analysis 2007, Chania Crete, 2007*. A paraître.
- [AR86] K.J. Arrow and H. Raynaud. *Social choice and multicriterion decision-making*. MIT Press, 1986.
- [Arr74] K. J. Arrow. *Choix collectif et préférences individuelles*. Calmann-Lévy, 1974.
- [Art47] E. Artin. Theory of braids. *Annals of mathematics*, 48 :101–126, 1947.
- [Ban02] P.D. Bangert. *Algorithmic problems in the braid group*. PhD thesis, Thesis from University College London, 2002.
- [Bar79] J.P. Barthélémy. Caractérisations axiomatiques de la distance de la différence symétrique entre des relations binaires. *Mathématiques et sciences humaines*, 67 :85–113, 1979.
- [Ben98] H. Benhadda. *La similarité régularisée et ses applications en classification automatique*. PhD thesis, Thèse de l'Université de Paris VI, 1998.
- [Ber94] M. Berger. Minimal crossing numbers for 3-braids. *Journal of physics A*, 27 :6205–6213, 1994.
- [BF89] D. Bourigault and J.M. Fouilloux. Utilisation du critère de condorcet dans le cadre de la recherche d'une typologie des sites nucléaires français. *Rapport EDF, Direction des études et recherches, Groupe Environnement social et socio-économique*, 1989.
- [BGH89] J. P. Barthélémy, A. Guénoche, and O. Hudry. Median linear orders : Heuristics and a branch and bound algorithm. *European Journal of Operational Research*, pages 313–325, 1989.

- [BGKZ02] J. Burillo, M. Gutierrez, S. Krstic, and Nitecki Z. Crossing matrices and thurston's normal form for braids. *Topology and its applications*, 118 :293–308, 2002.
- [BKL98] J. Birman, K.H. Ko, and S.J. Lee. A new approach to the word and conjugacy problems in the braid groups. *Advances in mathematics*, 139 :322–353, 1998.
- [BL95] J.P. Barthélemy and B. Leclerc. The median procedure for partitions. *DIMACS Series in Discrete Mathematics and Theoretical Computer Science*, 19 :3–33, 1995.
- [Bla48] D. Black. On the rationale of group decision making. *Journal of Political Economy*, 56 :23–34, 1948.
- [BM79] J.P. Barthélemy and B. Monjardet. Ajustement et résumé de données relationnelles : les relations centrales. In *Data Analysis and Informatics*, pages 645–654. North Holland, 1979.
- [BM98] H. Benhadda and J.F. Marcotorchino. Introduction à la similarité régularisée en analyse relationnelle. *Revue de Statistique Appliquée*, 46 :45–69, 1998.
- [BM03] M. Beigbeder and A. Mercier. Étude des distributions de tf et de idf sur une collection de 5 millions de pages html. In *INFORSID 2003 - Atelier Recherche d'Informations*, 2003.
- [Bur98] C.J.C. Burges. Geometry and invariance in kernel based methods. In *Advances in Kernel methods - Support vector learning*, pages 89–117. MIT Press, 1998.
- [Béd89] C. Bédécarrax. *Classification Automatique en Analyse Relationnelle : La Quadridécomposition et ses Application*. PhD thesis, Thèse de l'Université de Paris VI, 1989.
- [Cha85] S. Chah. Comparaisons par triplets en classification automatique. *Etude IBM F086*, 1985.
- [Chi80] G. Chichilnisky. Social choice and the topology of spaces of preferences. *Advances in mathematics*, 37 :165–176, 1980.
- [CKVW03] D. Cheng, R. Kannan, S. Vempala, and G. Wang. Normalized cuts and image segmentation. *Technical report MIT-LCS-TR-906, Massachusetts Institute of Technology*, 2003.
- [CLR02] T. Cormen, Ch. Leiserson, and R. Rivest. *Introduction à l'algorithmique*. Dunod, 2002.
- [Com70] L. Comtet. *Analyse Combinatoire, tome 1 et tome 2*. Presses Universitaires de France, 1970.
- [Con85] M.J.A. (Marquis de) Condorcet. *Essai sur l'application de l'analyse à la probabilité des décisions rendues à la pluralité des voix*. Paris, 1785.
- [CP81] J.L. Chandon and S. Pinson. *L'Analyse Typologique : Théories et Applications*. Masson, 1981.
- [dABR92] S.G. de Amorim, J.P. Barthélemy, and C.C. Ribeiro. Clustering and clique partitioning : simulated annealing and tabu search approaches. *Journal of classification*, 9 :17–41, 1992.
- [DDRW02] P. Dehornoy, I. Dynnikov, D. Rolfsen, and B. Wiest. *Why are braids orderable?* Société Mathématiques de France, 2002.
- [Deh92] P. Dehornoy. Deux propriétés des groupes de tresses. *Comptes rendus à l'académie des sciences Paris*, 315 :633–638, 1992.
- [Deh96] P. Dehornoy. Une autre application de la théorie des ensembles. *Gazette des mathématiciens*, 69 :3–20, 1996.
- [Deh97] P. Dehornoy. A fast method for comparing braids. *Advances in mathematics*, 125 :200–235, 1997.
- [Deh00a] P. Dehornoy. *Braids and self distributivity*. Birkhauser, 2000.
- [Deh00b] P. Dehornoy. On completeness of word reversing. *Discrete Mathematics*, 225 :93–119, 2000.
- [Deh04a] P. Dehornoy. Braid-based cryptography. *Contemporary mathematics*, 360 :5–33, 2004.
- [Deh04b] P. Dehornoy. Diagram colourings. In *Proc. First East Asian School of Knots, Links and related Topics, K.Y. Ko ed., Seoul (Korea)*, pages 37–64, 2004.

- [Det00] M. Detyniecki. *Mathematical aggregation operators and their application to video querying*. PhD thesis, Thèse de l'Université de Paris VI, 2000.
- [DGK05] I.S. Dhillon, Y. Guan, and B. Kulis. A unified view of kernel k-means, spectral clustering and graph partitioning. *UTCS Technical Report TR-04-25, University of Texas at Austin*, 2005.
- [DJ90] N. Dershowitz and J.P. Jouannaud. Rewrite systems. In *Handbook of theoretical computer science Volume B : Formal methods and semantics*, pages 243–320. North-Holland, Amsterdam, 1990.
- [DL05] M. Denuit and P. Lambert. Constraints on concordance measures in bivariate discrete data. *Journal of multivariate analysis*, 93 :40–57, 2005.
- [DLS05] J.J. Droesbeke, M. Lejeune, and G. Saporta. *Modèles statistiques pour données qualitatives*. Technip, 2005.
- [dST03] V. de Silva and J. Tenenbaum. Global versus local methods in nonlinear dimensionality reduction. In *Proc. NIPS, volume 15*, pages 721–728, 2003.
- [DW06] P. Dehornoy and B. Wiest. On word reversing in braid groups. *International Journal of Algebra and Computation*, 16, 2006.
- [ECH⁺92] D.B.A. Epstein, J.W. Cannon, D.F. Holt, S.V.F. Levy, M.S. Paterson, and W.P. Thurston. *Word processing in groups*. Jones and Bartlett, 1992.
- [EL05] C. Elsholtz and C. List. A simple proof of sen's possibility theorem on majority decisions. *Elemente der Mathematik*, 60 :45–56, 2005.
- [ES00] D.B.A Epstein and P.J. Sanders. Knuth-bendix for groups with infinitely many rules. *International Journal of Algebra and Computation*, 10 :539–591, 2000.
- [FH69] J. Feldman-Högaasen. Ordres partiels et permutoèdres. *Mathématiques et sciences humaines*, 28 :27–38, 1969.
- [Fis77] P.C. Fishburn. Condorcet social choice functions. *SIAM Journal on Applied Mathematics*, 33 :469–489, 1977.
- [FLC84] B. Fichet and G. Le Calvé. Structure géométrique des principaux indices de dissimilarité sur signes de présence-absence. *Statistiques et Analyse de données*, 9 :11–44, 1984.
- [FM83] E.B. Fowlkes and C.L. Mallows. A method for comparing two hierarchical clusterings. *Journal of American Statistical Association*, 78 :553–569, 1983.
- [For65] E. Forgy. Cluster analysis of multivariate data : Efficiency vs. interpretability of classifications. *Biometrics*, 21 :768, 1965.
- [Gar69] F.A. Garside. The braid groups and other groups. *Quarterly journal of mathematics Oxford*, 20 :235–254, 1969.
- [GBG03] Y. Guan, N. Belacel, and A.A. Ghorbani. Y-means : a clustering method for intrusion detection. In *Canadian Conference on Electrical and Computer Engineering (CCECE-2003)*, 2003.
- [Gha90] S. Ghashghaie. *Agrégation relationnelle des données ordinales : généralisation de critères d'association*. PhD thesis, Thèse de l'Université de Paris VI, 1990.
- [Gir85] S. Giraud. *Comparaison de règles d'agrégation d'ordres et étude de règles multicritères d'aide à la décision*. PhD thesis, Thèse de l'Université de Paris VI, 1985.
- [Git04] T. Gittings. Minimum braids : a complete invariant of knots and links. *preprint*, 2004.
- [GK63] L. Goodman and W. Kruskal. Measures of association for cross classification iii : Approximate sample theory. *Journal of The American Statistical Association*, 58 :310–364, 1963.
- [GK99] P. Green and A. Krieger. A generalized rand-index method for consensus clustering of separate partitions of the same data base. *Journal of classification*, 16 :63–89, 1999.
- [GM87] A. Guenoche and B. Monjardet. Méthodes ordinales et combinatoires en analyse des données. *Mathématiques et sciences humaines*, 100 :5–47, 1987.
- [GM01] M. Gondran and M. Minoux. *Graphes, dioïdes et semi-anneaux*. Technique et Documentation, 2001.

- [Gow71] J.C. Gower. A general coefficient of similarity and some of its properties. *Biometrics*, 27 :857–872, 1971.
- [GR02] J. Grabmeier and A. Rudolph. Techniques of cluster algorithms in data mining. *Data Mining and Knowledge Discovery*, 6 :303–360, 2002.
- [gra80] In *Regards sur la théorie des graphes, actes du colloque de Cerisy*. Presses Polytechniques romandes, 1980.
- [HA85] L. Hubert and P. Arabie. Comparing partitions. *Journal of classification*, 2 :193–218, 1985.
- [Har75] J.A. Hartigan. *Clustering Algorithms*. John Wiley and Sons, 1975.
- [HO00] A. Hyvärinen and E. Oja. Independent component analysis : algorithms and applications. *Neural Networks*, 13 :411–430, 2000.
- [Hol97] D.F. Holt. Rewriting techniques in finitely generated groups and monoids. *Lecture given at New Mexico University*, 1997.
- [Hol98] D.F. Holt. Automatic groups, subgroups and cosets. *Geometry and Topology Monographs*, 1 :249–260, 1998.
- [How69] N. Howard. Least-squares classification and principal component analysis. In *Quantitative Ecological Analysis In The Social Sciences*. Cambridge MIT Press, 1969.
- [HSST03] D.R. Hardoon, S. Szedmak, and J. Shawe-Taylor. Canonical correlation analysis : An overview with application to learning methods. *Technical Report CSD-TR-03-02, Royal Holloway University of London*, 2003.
- [Hud96] J.F.P. Hudson. The intersection of flat subsets of a braid group. In *Geometric group theory down under*. Walter de Gruyter and Co, 1996.
- [Hud04] O. Hudry. Computation of median orders : complexity results. In *Workshop on Computer Science and Decision Theory*, pages 179–214, 2004. Annales du LAMSADE Nr 3.
- [HV05] B. Hammer and T. Villmann. Classification using non standard metrics. In *ESANN 2005, 13th European Symposium on Artificial Neural Networks, Bruges, Belgium, April 27-29, 2005, Proceedings*, 2005.
- [Ina64] K. Inada. A note on the simple majority decision rule. *Econometrica*, 32 :525–531, 1964.
- [JL69] E. Jacquet-Lagrèze. L'agrégation des opinions individuelles. *Informatiques et sciences humaines*, 4 :1–21, 1969.
- [Jon05] V.F.R. Jones. The jones polynomials. *preprint*, 2005.
- [Jor26] Ch. Jordan. Sur la probabilité des épreuves répétées, le théorème de bernoulli et son inversion. *Bulletin de la S.M.F.*, 54 :101–137, 1926.
- [Jor27] Ch. Jordan. Les coefficients d'intensité relative de körösy. *Revue de la société hongroise de statistique*, 5, 1927.
- [Jor39] Ch. Jordan. Problèmes de la probabilité des épreuves répétées dans le cas général. *Bulletin de la S.M.F.*, 67 :223–242, 1939.
- [JV92] S. Janson and J. Vegelius. The j-index as a measure of association for nominal scale response agreement. *Applied psychological measurement*, 16 :243–250, 1992.
- [Kas99] C. Kassel. L'ordre de dehornoy sur les tresses. *Séminaire Bourbaki*, 52, 1999.
- [Kau83] L.H. Kauffman. *Formal knot theory*. Princeton university press, 1983.
- [Kau88] L.H. Kauffman. New invariants in the theory of knots. *American mathematical monthly*, 95 :195–242, 1988.
- [Kau05] L.H. Kauffman. The mathematics and physics of knots. *Reports on progress in physics*, 68 :2829–2857, 2005.
- [Kee94] R.D. Keever. Minimal 3-braid representations. *Journal of knot theory and its ramification*, 3 :163–177, 1994.
- [Ken70] M. G. Kendall. *Rank correlation methods*. Griffin, Londres, 1970.
- [KG03] M. Kuss and T. Graepe. The geometry of kernel canonical correlation analysis. *Technical Report Nr 108, Max Planck Institute for Biological Cybernetics*, 2003.

- [Kim71] J. Kim. Predictive measures of ordinal association. *American journal of sociology*, 76 :891–907, 1971.
- [KMP00] E.P. Klement, R. Mesiar, and E. Pap. *Triangular norms*. Kluwer Academic Pub, 2000.
- [KS62] J.G. Kemeny and J.L. Snell. *Mathematical models in the social sciences*. Blaisdell, New York, Reprinted by MIT Press (1972), Cambridge, 1962.
- [L’A58] M. L’Abbé. Structures algébriques suggérées par la logique mathématique. *Bulletin de la S.M.F.*, 86 :299–314, 1958.
- [LBAP06] J. Lemoine, H. Benhadda, and J. Ah-Pine. Classification non supervisée de documents hétérogènes : application au corpus 20 newsgroup. *IPMU*, 2006.
- [Leb05] G. Lebanon. *Riemannian Geometry and statistical machine learning*. PhD thesis, Thesis from Carnegie Mellon University, 2005.
- [Ler79] I.C. Lerman. Les représentations factorielles de la classification. *R.A.I.R.O.*, 13 :107–128, 1979.
- [Ler81] I.C. Lerman. *Classification et analyse ordinale de données*. Dunod, 1981.
- [Lis] C. List. The epistemology of special majority voting. *Social choice and welfare*, à paraître.
- [Lis04] C. List. On the significance of the absolute margin. *British Journal for the Philosophy of Science*, 55 :521–544, 2004.
- [LPM06] L. Lebart, M. Piron, and A. Morineau. *Statistique exploratoire multidimensionnelle : visualisation et inférence en fouilles de données*. Dunod, 2006.
- [LS94] L. Lebart and A. Salem. *Statistique textuelle*. Dunod, 1994.
- [LW67] G.N. Lance and W.T. Williams. A general theory of classificatory sorting strategies : 1. hierarchical systems. *Computer Journal*, 9 :373–380, 1967.
- [Mad91] R.D. Maddux. The origin of relation algebras in the development and axiomatization of the calculus of relations. *Studia Logica*, 50(3/4) :421–455, 1991.
- [Mah] K. Mahlburg. An overview of braid group cryptography. *preprint*.
- [Mar81] J.F. Marcotorchino. Agrégation des similarités en classification automatique. *Thèse de Doctorat d’Etat de l’Université Paris VI*, 1981.
- [Mar84a] J.F. Marcotorchino. Utilisation des comparaisons par paires en statistique des contingences partie I. *Etude IBM F069*, 1984.
- [Mar84b] J.F. Marcotorchino. Utilisation des comparaisons par paires en statistique des contingences partie II. *Etude IBM F071*, 1984.
- [Mar85] J.F. Marcotorchino. Utilisation des comparaisons par paires en statistique des contingences partie III. *Etude IBM F081*, 1985.
- [Mar86] J.F. Marcotorchino. Maximal association theory as a tool of research. In *Classification as a tool of research*, pages 275–288. North-Holland, Amsterdam, 1986.
- [Mar91a] J.F. Marcotorchino. L’analyse factorielle-relationnelle : parties I et II. *Etude du CE-MAP, IBM France, MAP-03*, 1991.
- [Mar91b] J.F. Marcotorchino. Seriation problems : An overview. *Applied stochastic models and data analysis*, 7 :139–151, 1991.
- [Mar03] J.F. Marcotorchino. Présentation des indices de similarités existants en vue d’une catégorisation et d’une restructuration nouvelle. *Rapport technique, Thales Communications*, 2003.
- [Mar06] J.F. Marcotorchino. Relational analysis theory as a general approach to data analysis and data fusion. In *Cognitive Systems with interactive sensors*, 2006.
- [May52] K.O. May. A set of independent necessary and sufficient conditions for simple majority decisions. *Econometrica*, 20 :680–684, 1952.
- [MEA91] J.F. Marcotorchino and N. El Ayoubi. Paradigme logique des écritures relationnelles de quelques critères fondamentaux d’association. *Revue de statistique appliquée*, 39 :55–46, 1991.
- [Mei05] M. Meila. Comparing clustering - an axiomatic view. In *In Proceedings of the 22nd International Conference on Machine Learning (ICML’2005)*, 2005.

- [Men42] K. Menger. Statistical metrics. *Proc. Nat. Acad. of Sci., USA*, 28 :535–537, 1942.
- [Men51] K. Menger. Probabilistic theories of relations. *Proc. Nat. Acad. of Sci.*, 28 :226–229, 1951.
- [Mic82] P. Michaud. Agrégation à la majorité I : Hommage à condorcet. *Etude du Centre Scientifique IBM France*, F051, 1982.
- [Mic85a] P. Michaud. Agrégation à la majorité II : analyse du résultat d’un vote. *Etude du Centre Scientifique IBM France*, F052, 1985.
- [Mic85b] P. Michaud. Hommage à condorcet (version intégrale pour le bicentenaire de l’essai de condorcet). *Etude du Centre Scientifique IBM France*, F094, 1985.
- [Mic90] P. Michaud. L’analyse logique : le lien entre les mathématiques appliquées et l’intelligence artificielle. *Publications scientifiques et techniques d’IBM france*, 1, 1990.
- [Mil00] N. Miller. The geometry of voting cycles : theoretical developments. *Paper presented at the Workshop on mathematical models of individual and public choice*, 2000.
- [Mir74] B. Mirkin. Approximation problems in a relation space and the analysis of nonnumeric methods. *Automatika i Telemekhanika (Automaton and remote control)*, 9 :53–61, 1974.
- [Mir86] B. Mirkin. Additive clustering and qualitative factor analysis methods. In *Data Analysis and Informatics IV*, pages 71–82, 1986.
- [Mir98] B. Mirkin. Least-squares structuring, clustering, and data processing issues. *The computer journal*, 41 :518–536, 1998.
- [Mir01] B. Mirkin. Eleven ways to look at the chi-squared coefficient for contingency tables. *The American Statistician*, 55 :111–120, 2001.
- [MK99] K. Murasugi and B.I. Kurpita. *A study of braids*. Kluwer academic publishers, 1999.
- [MM78] J.F. Marcotorchino and P. Michaud. *Optimisation en analyse ordinale des données*. Masson, 1978.
- [MM79] P. Michaud and J.F. Marcotorchino. Modèles d’optimisation en analyse des données relationnelles. *Mathématiques et Sciences Humaines*, vol 17(67) :7–38, 1979.
- [MM81] J.F. Marcotorchino and P. Michaud. Heuristic approach of the similarity aggregation problem. *Methods of operations research*, 43 :395–404, 1981.
- [Mon85] B. Monjardet. Concordance et consensus d’ordres totaux : les coefficients k et w . *Revue de statistique appliquée*, 33 :55–87, 1985.
- [NCS04] J.P. Nakache, J. Confais, and G. Saporta. *Approche pragmatique de la classification : arbres hiérarchiques et partitionnements*. Technip, 2004.
- [Nel99] R.B. Nelsen. *An introduction to copulas*, volume 139 of *Lecture notes in statistics*. Springer Verlag, 1999.
- [Nel03] R.B. Nelsen. Properties and applications of copulas : A brief survey. In *Proceedings of the First Brazilian Conference on Statistical Modelling in Insurance and Finance*, 2003.
- [NI00] A. Najah Idrissi. *Contribution à l’unification de critères d’association pour variables qualitatives*. PhD thesis, Thèse de l’Université de Paris VI, 2000.
- [PA91] M.S. Paterson and Razborov A.A. The set of minimal braids is co-np-complete. *Journal of algorithms*, 12 :393–408, 1991.
- [PKK05] J. Peltonen, A. Klami, and S. Kaski. Improved learning of riemannian metrics for exploratory analysis. *Neural networks*, 18, 2005.
- [PM00] D. Pelleg and A. Moore. X-means : Extending k-means with efficient estimation of the number of clusters. In *Proceedings of the Seventeenth International Conference on Machine Learning*, pages 727–734, San Francisco, 2000. Morgan Kaufmann.
- [PP] M. Petkovsek and T. Pisanski. Combinatorial interpretation of unsigned stirling and lah numbers. *preprint*.
- [Pra92] V. Pratt. Origins of the calculus of binary relations. In *Logic in Computer Science*, pages 248–254, 1992.
- [Ran71] W. H. Rand. Objective criteria for the evaluation of clusterings methods. *Journal of the American Statistical Association*, 66, 1971.

- [Ray81a] H. Raynaud. Conditions for transitivity of majority rule with algorithmic interpretations. *Technical Report 347. Institute for mathematical studies in the social sciences, Stanford University, Stanford, CA*, 1981.
- [Ray81b] H. Raynaud. How restrictive actually are value restriction conditions. *Technical Report 348. Institute for mathematical studies in the social sciences, Stanford University, Stanford, CA*, 1981.
- [Ray81c] H. Raynaud. Paradoxical results from inada's conditions for majority rule. *Technical Report 331. Institute for mathematical studies in the social sciences, Stanford University, Stanford, CA*, 1981.
- [RS00] S. Roweis and L. Saul. Nonlinear dimensionality reduction by locally linear embedding. *Science*, 290 :2323, 2000.
- [Rég65] S. Régnier. Sur quelques aspects mathématiques des problèmes de classification automatique. *ICC Bulletin*, 4 :175–191, 1965.
- [Sap88] G. Saporta. About maximal association criteria in linear analysis and in cluster analysis. In H.H. Bock, editor, *Classification and Related Methods of Data Analysis*, pages 541–550. Elsevier Science Publishers B.V. (North-Holland), 1988.
- [Sap06] G. Saporta. *Probabilités, analyse des données et statistiques - 2nde édition*. Technip, 2006.
- [SBC03] H. Simsek, M. Bayram, and I. Can. Automatic calculation of minimum crossing numbers of 3-braids. *Appl. Math. Comput.*, 144 :507–516, 2003.
- [Sen66] A.K. Sen. A possibility theorem on majority decisions. *Econometrica*, 34 :491–499, 1966.
- [SF66] H. Solomon and J.J. Fortier. Clustering procedures. In *International Symposium on Multivariate Analysis*. P. Krishnaiah Editor, Academic Press, New York, 1966.
- [SJ00] J. Shi and Malik J. Normalized cuts and image segmentation. *IEEE Transaction on Pattern Analysis and Machine Intelligence*, 22 :888–905, 2000.
- [SKK00] M. Steinbach, G. Karypis, and V. Kumar. A comparison of document clustering techniques. In *KDD Workshop on Text mining*, 2000.
- [Som62] R.H. Somers. A new asymmetric measure of association for ordinal variables. *American sociological review*, 33 :971–977, 1962.
- [Sos99] A. Sossinsky. *Noeuds, genèse d'une théorie mathématique*. Seuil, 1999.
- [SS58] B. Schweizer and A. Sklar. Espaces métriques aléatoires. *Comptes rendus Acad. Sci. Paris*, 247 :2092–2094, 1958.
- [SS83] B. Schweizer and A. Sklar. *Probabilistic metric spaces*. North-Holland, 1983.
- [SS02] B. Schölkopf and A. Smola. *Learning with kernels*. MIT Press, 2002.
- [Tar41] A. Tarski. On the calculus of relations. *Journal of symbolic logic*, 6 :73–89, 1941.
- [Tat87] K. Tatsuoka. On geodesics in the braid group. *Preprint*, 1987.
- [TdSL00] J. Tenenbaum, V. de Silva, and J.C. Langford. A global geometric framework for nonlinear dimensionality reduction. *Science*, 290 :2319–2323, 2000.
- [Tor52] W.S. Torgerson. Multidimensional scaling : I. theory and method. *Psychometrika*, 17 :401–419, 1952.
- [Tur00] N. Turenne. *Apprentissage statistique pour l'extraction de concepts à partir de textes. Application au filtrage d'informations textuelles*. PhD thesis, Thèse de l'université Louis Pasteur, 2000.
- [Vai79] L.G. Vailiant. The complexity of computing the permanent. *Theoretical Computer Science*, 8 :189–201, 1979.
- [Vic60] W. Vickrey. Utility, strategy and social decision rules. *Quarterly journal of economics*, 74 :507–535, 1960.
- [Vog90] P. Vogel. Representation of links by braids : a new algorithm. *Commentarii Mathematici Helvetici*, 65 :104–113, 1990.
- [Wak98] Y. Wakabayashi. The complexity of computing medians of relations. *Resenhas IME-USP*, 3 :323–349, 1998.

- [War65] B. Ward. Majority voting and the alternative forms of public enterprise. In *In Margolis J (ed) The public economy urban communities*. John Hopkins University Press, Baltimore, 1965.
- [YS04] G. Youness and G. Saporta. Some measures of agreement between close partitions. *Student*, 51 :1–12, 2004.

LISTE DES TABLEAUX

1.1	Propriétés relationnelles exprimées sous forme de règles logiques	3
1.2	Propriétés relationnelles exprimées sous forme de contraintes linéaires dans le cadre de l'AR	4
1.3	Rappels des différentes similarités régularisées dans le cas de variables qualitatives	9
1.4	Formules de passage contingence/relationnel	12
1.5	Critères de partitionnement issus de l'association maximale	15
1.6	Formules de passage "concordance"/relationnel de S. Ghashghaie	19
2.1	Action des opérations de transferts sur les matrices d'affectation	47
2.2	Dualité "affectation / relationnel" lors de l'action des opérations de transferts sur des matrices relationnelles codant des partitions	50
3.1	Définition des quantités $11_{ii'}$, $00_{ii'}$, $10_{ii'}$ et $01_{ii'}$ entre deux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$	71
3.2	Définition des quantités dont sont fonctions les indices de similarité entre vecteurs binaires	71
3.3	Expressions relationnelles des quantités $11_{ii'}$, $00_{ii'}$, $10_{ii'}$, $01_{ii'}$	73
3.4	Expressions des indices de similarité en fonction de Dice et règles majoritaires correspondantes	75
3.5	Interprétations géométriques des quantités $11_{ii'}$, $10_{ii'}$, $01_{ii'}$	77
3.6	Etude analytique des indices de similarité	79

3.7	Etude des caractéristiques des indices de similarité	80
3.8	Principe d'écart à une tendance centrale vérifié par les critères de partitionnement issus de l'association maximale	89
3.9	Critères de partitionnement issus de l'association maximale en termes de produits scalaires	90
3.10	Définition des accords et des désaccords entre deux matrices relationnelles de partition	100
3.11	Tableau de correspondance des critères d'association en termes relationnels sous forme de coefficient de corrélation	105
3.12	Dualité dans le formalisme contingence / relationnel entre le critère de Belson et le critère de Janson et Vegelius correspondant à une dualité indépendance statistique / indépendance géométrique lié à la décomposition de Torgerson	106
4.1	Stabilité des propriétés relationnelles sous l'action des opérations booléennes	117
4.2	Tables de vérité de la logique trivalente de Lukasiewicz	119
4.3	Tabulation des coefficients $\alpha_k^{m,M}$ dans le cas particulier où m représente la "majorité large" de M	137
4.4	Enveloppes de Condorcet selon les coefficients combinatoires de Jordan	138
4.5	Enveloppes de Condorcet selon les nombres de Lah	139
4.6	Enveloppes de Condorcet selon les coefficients binomiaux négatifs	141
4.7	Triangle de Pascal et coefficients binomiaux négatifs	141
5.1	Action des transpositions sur les matrices d'affectation	172
5.2	Dualité affectation/relationnel lors de l'action des transpositions sur des matrices relationnelles codant des classements	174
6.1	Nouvelles formules de passage "concordance"/relationnel	187
6.2	Définition des accords et des désaccords entre deux matrices relationnelles d'ordre	189
6.3	Exemple d'ensemble de préférences individuelles qui ne vérifient pas les conditions de Sen et qui conduisent à une relation collective transitive	207
6.4	Triplet de l'exemple étudié ne vérifiant pas les conditions de Sen	207

TABLE DES FIGURES

1.1	Exemple de deux graphes quelconques	14
1.2	Exemple d'une tresse à 3 brins	20
1.3	Relations d'Artin du groupe des tresses	22
1.4	Noeud de trèfle	23
1.5	Les mouvements de Markov	25
1.6	Relations entre les différents chapitres	29
2.1	Exemple de treillis de partitions	39
2.2	Graphe orienté des partitions d'un ensemble d'objets reliées par des transferts générateurs de $\mathbb{T}(\mathbb{O})$	53
2.3	Graphe orienté des partitions d'un ensemble d'objets reliées par des transferts générateurs de $\mathbb{T}(\mathbb{D})$	54
2.4	Les 5 possibilités de partitions à partir de l'intersection de 2 classes	62
2.5	Liens entre les classes atomiques issus de l'intersection de 2 classes	63
2.6	Représentation graphique de l'algorithme AR2	67
3.1	Interprétation géométrique de la similarité d'ordre t entre deux vecteurs	84
3.2	Domaines des contributions positives des indices de similarité d'ordre t	87
4.1	Représentations des quantités intervenant dans le calcul de similarités entre ob- jets en se basant sur les triordonnances	148

5.1	Exemple de treillis de permutations	167
5.2	Graphe de Cayley des permutations d'un ensemble d'objets	168
5.3	Méthode graphique de calcul de la distance de Kendall entre deux permutations	169
5.4	Graphe orienté des permutations d'un ensemble d'objets	173
6.1	Illustration des conditions d'unimodalité de D. Black et de K. Inada	196
6.2	Représentation de l'exemple traité mettant en évidence la non vérification des conditions de Black et d'Inada	208
7.1	Mouvements isotopiques de retournement à droite	219
7.2	Mouvements isotopiques de réduction d'une σ^j -poignée	222
7.3	Chemin dans l'automate de \mathbb{B}_4 représentatif du mot de tresse bac partant de la tresse triviale	228
7.4	Automate du groupe des tresses à 3 brins	228
7.5	Différents cas de brins libres	231
7.6	Transposition de brins non juxtaposés et déplacement d'un brin libre	233
7.7	Illustration d'un déplacement à droite d'un brin libre sans tenir compte dans un premier temps de la nature topologique des croisements	234
7.8	Exemple de configuration d'un brin libre O^2vO^2	236
7.9	Exemple de mouvement à droite d'un brin libre O^2vO^2	242
7.10	Signatures du mouvement isotopique à droite d'un brin libre O^2vO^2 pour $\alpha = (1, 1, 1)$	245
7.11	Signatures du mouvement isotopique à droite d'un brin libre O^2vO^2 pour $\alpha = (-1, -1, -1)$	245
7.12	Signatures du mouvement isotopique à droite d'un brin libre O^2vO^2 pour $\alpha = (-1, -1, 1)$	246
7.13	Signatures du mouvement à droite d'un brin libre O^2vO^2 pour $\alpha = (-1, 1, -1)$ qui donne un mouvement non isotopique	247
7.14	Mouvement de brin libre faisant apparaître un brin libre de type "poignée de Dehornoy simple" permettant de "sortir d'un optimum local"	257
7.15	Mouvement de brin libre faisant apparaître un brin libre de type "poignée de Dehornoy" permettant ensuite de "sortir d'un optimum local"	258
7.16	Procédé 1 permettant de calculer dans le meilleur des cas la forme canonique "shortlex" d'un mot de tresse	260

7.17 Procédé 2 permettant de calculer dans le meilleur des cas la forme canonique "shortlex" d'un mot de tresse	262
7.18 Exemple d'application : identification du brin libre de type "poignée de Dehornoy simple" conduisant à une réduction de la taille du mot de tresse	263
7.19 Exemple d'application : détermination d'une forme "shortlex" à l'aide d'un mouvement isotopique de brin libre issue d'une poignée qui n'est pas de type de Dehornoy	265
7.20 Illustration de la terminaison du procédé 1 sur l'exemple traité	266
7.21 Exemple d'application : mot de tresse obtenue après le mouvement isotopique de brin libre et détermination de la forme canonique "shortlex"	267
7.22 Exemple d'application : mot de tresse obtenu après le mouvement isotopique de brin libre et détermination de la forme canonique "shortlex"	270
7.23 Mouvement de Markov $M1$ dans le cadre des paires de brins nommés	273
7.24 Mouvement de Markov $M2$ dans le cadre des paires de brins nommés	274
7.25 Procédé permettant de calculer dans le meilleur des cas la forme canonique "shortlex" d'un mot de noeud tressé	278
7.26 Exemple d'application : calcul de la forme canonique "shortlex" d'un mot de noeud tressé	279
7.27 Exemple d'application : résultat de la simplification par un mouvement de brin libre débouchant sur la réduction libre issue d'un mouvement de Markov $M1$	281
7.28 Exemple d'application : résultat de la simplification un mouvement de Markov $M2$, obtention de la forme "shortlex" du mot de noeud tressé	283

NOTATIONS

Notations générales présentées en introduction

AR	:	Analyse Relationnelle
\mathcal{T}	:	Tableau de données
\mathbb{O}	:	Ensemble des éléments qui sont soit des objets, soit des candidats, soit des brins
O^i	:	Element i de \mathbb{O}
N	:	Nombre d'éléments de \mathbb{O}
\mathbb{V}	:	Ensemble des éléments qui sont des variables ou des juges
V^k	:	Elément k de \mathbb{V}
M	:	Nombre d'éléments de \mathbb{V}
p_k	:	Nombre de modalités de V^k lorsque qu'il s'agit de la variable qualitative
\mathbb{D}	:	Ensemble des éléments qui sont soit des descripteurs (binaires ou non binaires) soit des rangs
D^j	:	Elément j de \mathbb{D}
P	:	Nombre d'éléments de \mathbb{D}
κ	:	Codage algébrique d'une partition ou d'une relation d'ordre
K	:	Matrice d'affectation représentative de κ
\mathbf{K}	:	Matrice disjonctive complète
\mathbf{n}^{kl}	:	Tableau de contingence des variables qualitatives V^k et V^l
\mathcal{R}^k	:	Relation binaire k
$G(\mathcal{R}^k)$:	Graphe de la relation binaire k
\mathbb{C}	:	Ensemble des matrices relationnelles
C^k	:	Matrice relationnelle d'une relation k
\mathbf{C}	:	Matrice relationnelle collective
\mathbf{S}	:	Matrice de comparaisons par paires de similarités
X	:	Matrice relationnelle inconnue (d'une relation d'équivalence ou d'ordre)

Notations liées au chapitre 2

$\text{RE}(\mathbb{O})$: Ensemble des relations d'équivalence possibles sur l'ensemble \mathbb{O}
$\mathbb{T}(\mathbb{O})$: Monoïde des transferts sur l'ensemble \mathbb{O}
τ^k	: Élément du monoïde des transferts
$\mathbb{T}^s(\mathbb{O})$: Monoïde des transferts simples sur l'ensemble \mathbb{O}
$\mathbb{T}^{sg}(\mathbb{O})$: Monoïde des transferts générateurs sur l'ensemble \mathbb{O}
$\tau_i^{i'}$ (resp. $\sigma_j^{j'}$)	: Transfert simple de l'objet O^i vers l'objet $O^{i'}$ (resp. du descripteurs D^j vers le descripteur $D^{j'}$)
$\mathbb{T}^d(\mathbb{O})$: Groupe des transferts doubles sur l'ensemble \mathbb{O}
$\mathbb{T}^{dg}(\mathbb{O})$: Groupe des transferts doubles générateurs sur l'ensemble \mathbb{O}
$\tau^{ii'}$ (resp. $\sigma^{jj'}$)	: Transfert double générateur des objets O^i et $O^{i'}$ (resp. des descripteurs D^j et $D^{j'}$)
$\tau^k \circ \tau^{k'}$: Composition à droite de deux opérations de transfert
$\tau^k(\mathcal{R})$: Action d'un transfert sur les objets d'une relation
$\sigma^k(\mathcal{R})$: Action d'un transfert sur les classes d'une relation
T^k	: Représentation matricielle d'un élément du monoïde des transferts (matrice de transfert) définies sur \mathbb{O}
S^k	: Représentation matricielle d'un élément du monoïde des transferts (matrice de transfert) définies sur \mathbb{D}

Notations liées au chapitre 3

$11_{ii'}$: Nombre de 1 en commun aux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$
$00_{ii'}$: Nombre de 0 en commun aux vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$
$01_{ii'}$: Nombre de couples (0,1) issus du croisement des vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$
$10_{ii'}$: Nombre de couples (1,0) issus du croisement des vecteurs binaires \vec{O}^i et $\vec{O}^{i'}$
$\mathcal{M}^t(\mathbb{A})$: Moyenne généralisée d'ordre t d'un ensemble de nombres \mathbb{A}
$\mathbf{S}_{ii'}^t$: Indice de similarité d'ordre t entre les vecteurs \vec{O}^i et $\vec{O}^{i'}$
$S_C^t(C^k, C^l, f, \mu^k, \mu^l)$: Indice de corrélation d'ordre t entre les matrices relationnelles C^k et C^l dépendant des paramètres (f, μ^k, μ^l)
$S_I^t(C^k, C^l, f, \mu^k, \mu^l)$: Indice d'ordre t d'écart à l'indétermination pondérée entre les matrices relationnelles C^k et C^l dépendant du système de poids $(\mu_1^k, \mu_0^k, \mu_1^l, \mu_0^l)$
M	: Métrique relationnelle

Notations liées au chapitre 4

$\mathcal{B}(\mathcal{C})$: Algèbre de Boole engendrée par \mathcal{C}
$C^k \oplus C^l$: Addition booléenne (terme à terme) ou union relationnelle de deux matrices relationnelles
$C^k \otimes C^l$: Multiplication booléenne (terme à terme) ou intersection relationnelle de deux matrices relationnelles
\otimes	: Dans le cas de données non binaires \otimes représente de manière plus générale une "t-norm"
$C^k \odot C^l$: Multiplication relative (matricielle) booléenne issue de l'algèbre des relations)
$C^k \ominus C^l$: Addition relative (matricielle) booléenne issue de l'algèbre des relations)
\check{C}^k	: Converse d'une relation égale à la transposée ${}^t C^k$
I_N	: Matrice identité de dimension N
U_N	: Matrice carrée de dimension N remplie de 1
S_k^M	: Somme des $\binom{M}{k}$ combinaisons d'intersections de k matrices relationnelles (binaires ou non) parmi M
$[S_k^M]_{ii'}$: Terme général de la matrice S_k^M
$[S_k^M]^j$: j -ème matrice ordonnée de la somme S_k^M
$[S_k^M]_{ii'}^j$: Terme général de la matrice $[S_k^M]^j$
E_m^M	: Enveloppe de Condorcet d'ordre m (paramètre de la majorité) sur M variables ou juges
H_m^M	: Enveloppe de Condorcet partielle d'ordre m sur M variables ou juges
$[E_m^M]_{ii'}$: Terme général de l'enveloppe de Condorcet E_m^M
$\alpha_k^{M,m}$: Coefficients de détermination de l'enveloppe de Condorcet par combinaison linéaire
\overline{E}_m^M	: Complémentaire de type 1 de E_m^M
$\overline{\overline{E}}_m^M$: Complémentaire de type 2 de E_m^M

Notations liées au chapitre 5

$\text{ROST}(\mathbb{O})$: Ensemble des relations d'ordre strict et total possibles sur l'ensemble \mathbb{O}
$\mathbb{S}(\mathbb{O})$: Groupe des permutations de l'ensemble \mathbb{O}
$\mathbb{S}^g(\mathbb{O})$: Ensemble des transpositions génératrices du groupe des permutations de \mathbb{O}
$\mathbb{S}_m(\mathbb{O})$: Ensemble des transpositions d'ordre m du groupe des permutations de \mathbb{O}
κ	: Codage algébrique d'un relation d'ordre strict et total
K	: Représentation matricielle d'un élément κ
π^k	: Codage algébrique d'une permutation
$\pi^{ii'}$: Transposition ou 2-cycle des objets de rangs respectifs i et i'
π^i	: Transposition ou permutation génératrice des objets de rang i et $i + 1$
$\pi^k \circ \pi^l$: Composition (à droite) de π^k par π^l
P^k	: Représentation matricielle d'un élément π^k du groupe des permutations (matrice de permutation)
L	: Représentation sous forme de matrices de paires d'objets nommés d'un ordre κ
$[O^i, O^{i'}]$: Paire d'objets nommés élément constituant des matrices L
$(O^i, O^{i'})$: Paire d'objets nommés agissant sur un ordre strict et total représenté par L qui vaut $\phi([O^i, O^{i'}])$ et qui correspond à une transposition des objets nommés
$L * (.,.)$: Action d'une paire d'objets nommés sur la matrice L
$\mathbb{I}(\pi^k)$: Ensemble des inversions de la permutation π^k
I^k	: Matrice de comparaisons par paires des inversions d'une permutation π^k
$\tau^{ii'}(\mathcal{R})$: Action d'une transposition sur les objets d'une relation
$\sigma^{j,j+1}(\mathcal{R})$: Action d'une transposition génératrice sur les rangs d'une relation
$T^{ii'}$: Représentation matricielle d'une transposition sur \mathbb{O} (matrice de permutation) agissant sur les objets
$S^{j,j+1}$: Représentation matricielle d'une transposition génératrice sur \mathbb{D} (matrice de permutation) agissant sur les rangs
\mathbf{R}	: Matrice relationnelle de passage
R_N	: Matrice relationnelle de passage standard codant un ordre strict et total $D^1 < D^2 < \dots < D^N$

Notations liées au chapitre 6

$\text{ROLT}(\mathbb{O})$: Ensemble des relations d'ordre large et total possibles sur l'ensemble \mathbb{O}
κ	: Codage algébrique d'un ordre large et total
K	: Représentation matricielle d'un élément κ
$\text{Conc}(V^k, V^l)$: Nombre de paires concordantes entre les variables V^k et V^l
$\text{Disc}(V^k, V^l)$: Nombre de paires discordantes entre les variables V^k et V^l
$\text{Tie}(V^k)$: Nombre de paires qui ont la même valeur pour V^k
$\text{Tie}(V^k, V^l)$: Nombre de paires qui ont la même valeur à la fois pour V^k et pour V^l
Q_N	: Matrice relationnelle de passage standard codant un ordre large, total et sans <i>ex-aequo</i> $D^1 \leq D^2 \leq \dots \leq D^N$
\mathbb{Z}_M	: Ensemble des entiers relatifs de $-M$ à M
\mathbf{C}	: Matrice relationnelle collective
\mathbf{C}^r	: Matrice relationnelle de contributions relative à la règle majoritaire $r = 1, 2, 3$ majoritaire $r = 1, 2, 3$
X^r	: Matrice relationnelle représentative de l'ordre collectif obtenu selon la règle majoritaire $r = 1, 2, 3$
$C^k \odot^* C^l$: Multiplication relative (matricielle) modifiée issue de l'algèbre des relations)
$C^k \oplus^* C^l$: Addition relative (matricielle) modifiée issue de l'algèbre des relations)

Notations liées au chapitre 7

\mathbb{B}_N	:	Groupe des tresses à N brins
β	:	Tresse géométrique
$\mathring{\beta}$:	Tresse fermée ou clôturée définissant un noeud
w	:	Mot de tresse représentative de la tresse β
w^*	:	Forme canonique "shortlex" d'un mot de tresse ou d'un mot de noeud tressé
$(O^i, O^{i'})$:	Paire de brins nommés
L	:	Matrice de paires de brins nommés donnant l'ordre à un instant donné engendré par une tresse sur ses brins
$L * (O^i, O^{i'})$:	Action d'une paire de brins nommés sur une matrice de paires de brins nommés
$(L, (., .))$:	Couple décrivant à un instant donné une tresse dans le cadre de la modélisation par automate du groupe \mathbb{B}_N
W	:	Successions de couples $(L, (., .))$ décrivant une tresse dans le cadre de la modélisation par automate du groupe \mathbb{B}_N
$O^i v O^i$:	Configuration de brin libre relatif à O^i , correspond à un mot de paires de brins nommés u
$\xrightarrow{O^i}$:	Mouvement isotopique à droite du brin libre O^i
i^c	:	Indice du brin O^{i^c} juxtaposé à un brin libre
j^i	:	Indice constant du rang du brin libre O^i au cours des ordres (de permutation) engendrés par v relatif à une configuration $O^i v O^i$
a	:	Mot de paires de brins nommés matérialisant la phase ascendante d'un mouvement isotopique
d	:	Mot de paires de brins nommés matérialisant la phase descendante d'un mouvement isotopique
$C^{w,e}$:	Matrice relationnelle de la partition des brins liés engendré par w
$C^{w,o}$:	Matrice relationnelle de l'ordre topologique sur les brins engendré par w
$a(\alpha)$:	Mot de paires de brins nommés a (initialement non signé) signé par le $ a $ -uplet α
$d[a(\alpha)]$:	Mot de paires de brins nommés d (initialement non signé) signé par l'intermédiaire de $a(\alpha)$ en respectant les symétries topologiques

RESUME, ABSTRACT

Résumé

L'Analyse Relationnelle est un ensemble de méthodes qui s'intéresse à l'étude des relations binaires et aux applications faisant intervenir ces dernières. La particularité de la théorie présentée ici et proposée initialement par J.F. Marcotorchino et P. Michaud, est de représenter ces relations binaires sous la forme de matrices de comparaisons par paires d'objets sur lesquels portent ces relations. Par ailleurs, elle fait appel à diverses branches mathématiques que sont la théorie des graphes, la statistique et la programmation linéaire. Deux des applications classiques de l'Analyse Relationnelle sont la classification automatique et l'agrégation multicritère qui correspondent respectivement à l'étude des relations d'équivalence et des relations d'ordre.

Nous avons cherché à enrichir l'Analyse Relationnelle de différents outils provenant de l'algèbre relationnelle, de la logique multivalente, de l'analyse combinatoire et de l'optimisation combinatoire. Cette approche nous a permis d'étendre la théorie et les champs d'application de l'Analyse Relationnelle. Outre les applications classiques citées précédemment, nous nous sommes intéressés à l'axiomatique issue de la théorie du choix social et à la théorie des tresses. Les contributions que nous avons pu obtenir finalement dans le cadre de ce mémoire de thèse sont les suivantes :

- Définition d'un nouvel algorithme de classification automatique ne nécessitant pas la fixation d'un nombre de classes
- Présentation unifiée de certains indices de similarité entre vecteurs numériques (issus de variables binaires, qualitatives ou quantitatives), de certains critères d'association entre variables qualitatives et de certains critères d'association entre variables ordonnées
- Définition d'une méthode d'agrégation appelée calcul consensuel, fondée sur une règle majoritaire et qui se base sur les formules combinatoires de H. Poincaré et de Ch. Jordan
- Généralisation des résultats de D. Black, de K. Inada et de A.K. Sen en théorie du choix social aboutissant à des conditions nécessaires et suffisantes pour que la préférence collective issue de plusieurs préférences individuelles soit transitive
- Modélisation originale des tresses permettant de définir de manière générique des mouvements isotopiques de brins et d'aborder les problèmes algorithmiques autour du groupe des tresses et celui des noeuds sous l'angle de l'optimisation combinatoire

Mots-clés : Analyse Relationnelle, classification automatique, indices de similarité, critère d'association, critère de partitionnement, agrégation multicritère, fonction d'agrégation, choix social, théorème d'impossibilité d'Arrow et restrictions de domaines, permutations, théorie des tresses et des noeuds, mot de tresse minimal et "shortlex", invariant des noeuds

Abstract

Relational Analysis is concerned with the analysis of binary relations and their applications in different mathematical fields. The approach presented here and which has been developed initially by J.F. Marcotorchino and P. Michaud is particular as it represents the binary relations as pairwise comparisons matrices, and as it is basically related to different tools from graph theory, statistics and linear programming. The most usual application domains of Relational Analysis are clustering and multicriteria decision making which are respectively based upon the analysis of equivalence and order relations.

We have been interested in extending Relational Analysis to binary relations algebra, multi-valued logics, combinatorics and combinatorial optimization. These researches have given new results in the usual application domains which have been mentioned beforehand and also to some new extensions namely, the axiomatic of social choice theory and braids and knots theories. Finally, this work presents the following contributions :

- A new clustering algorithm for which the number of clusters is not required
- A unified framework for some similarity indices between numerical vectors (which are valid for binary or qualitative or quantitative variables), some categorical variables association criteria and ordered variables association criteria
- An aggregation process called the "consensus principle", which is based on a generalized majority rule and which is derived from the combinatoric formulas of H. Poincaré and Ch. Jordan
- A generalization of the D. Black, K. Inada and A.K. Sen restriction conditions in social choice theory, which aims at defining sufficient and necessary conditions for the transitivity of the social choice obtained from several individual preferences
- A new modelling approach for braids that enables to take into account generic isotopic movements of strings and to present algorithmic problems in braids and knots theory as combinatorial optimization problems

Keywords : Relational Analysis, clustering, similarity indices, association criteria, partitioning criteria, multicriteria decision making, aggregation function, social choice theory, Arrow's impossibility theorem and domain restrictions, permutations, braids and knots theory, minimal and "shortlex" braid word, knot invariants