Modéliser l'entre deux dans l'organisation spatiale des exploitations agricoles

F. Le Ber[1] *et C. Brassac*[2]

Résumé. Dans le cadre d'un projet interdisciplinaire regroupant agronomes, informaticiens, psychologues et linguistes, nous nous sommes penchés sur la modélisation des organisations spatiales agricoles, en confrontant l'approche chorématique et une approche informatique basée sur des graphes. Le travail conjoint s'est accompli pour partie lors de (trois) séances filmées qui ont permis de réaliser a posteriori *une analyse fine des problématiques de modélisation. Dans ce texte, nous focaliserons notre analyse sur une des questions surgies au cours du travail conjoint, à savoir la modélisation de l'entre-deux spatial et temporel, des surfaces "pas importantes mais néanmoins commodes" pour la gestion quotidienne de l'exploitation agricole. Nous donnerons des exemples tirés du corpus filmé et tenterons d'expliciter la problématique de modélisation sous-jacente.*

Mots-clés : territoire agricole, modélisation, graphes, chorèmes, interaction

Introduction

Nous traitons d'un projet qui a réuni agronomes et informaticiens autour de la conception d'un système à bases de connaissances, dénommé ROSA, pour *Raisonnement sur des Organisations Spatiales Agricoles*. Ce système a pour objectif de représenter et d'exploiter de façon automatique un ensemble de données, informations et connaissances sur les exploitations agricoles enquêtées et étudiées par les agronomes (Lardon et Osty, 2004). Ces enquêtes se focalisent sur les rapports entre organisations spatiale et fonctionnelle des exploitations, dans une perspective de gestion environnementale. Le système construit est un système de *raisonnement à partir de cas* (Riesbeck et Schank, 1989 ; Leake, 1996) : il est composé d'une *base de cas* constituée des données, informations et connaissances sur les exploitations enquêtées, d'une *base de connaissances* sur le domaine et d'un *module de raisonnement* à partir de cas. Les objectifs de ce module sont de comparer et d'adapter les connaissances liées à un cas (par exemple une exploitation *source*, dont on connaît à la fois la structure spatiale et le fonctionnement) à un autre cas (une exploitation *cible* dont on ne connaît que la structure). La comparaison est établie sur des mesures de similarités entre les structures spatiales des deux cas et l'adaptation permet de proposer une logique de fonctionnement à l'exploitation *cible* à partir de la logique de fonctionnement de l'exploitation *source* en s'appuyant sur la similarité entre les structures spatiales. L'hypothèse sous-jacente peut s'exprimer ainsi : *le fonctionnement de* source *est au fonctionnement de* cible *ce que la structure de* source *est à la structure de* cible (Lieber, 1997).

Le développement du système ROSA s'appuie sur et s'accompagne d'un processus d'acquisition - modélisation de connaissances impliquant agronomes et informaticiens, et confrontant leurs outils et méthodes de représentation de l'espace. Les agronomes utilisent les chorèmes (Lardon *et al.*, 2000), tandis que les informaticiens ont fait le choix d'utiliser des graphes étiquetés (Le Ber *et al.*, 2003). Dans le cadre de notre projet, nous avons utilisé un formalisme inspiré des graphes conceptuels (Sowa, 1984 ; Chein et Mugnier, 1992). Un graphe est composé de deux ensembles de sommets, l'un représente des entités, l'autre des relations. Les sommets-entités et les sommets-relations peuvent être regroupés en concepts auxquels sont associés des attributs. Les

1 *florence.leber@engees.u-strasbg.fr*, ENGEES-ULP, CEVH, Strasbourg & LORIA
2 *Christian.Brassac@univ-nancy2.fr*, Université de Nancy, Codisant, Laboratoire de Psychologie

arêtes du graphe sont étiquetées par le rôle[3] des entités dans la relation. Nous utilisons les graphes pour modéliser les chorèmes en supposant que :

- les objets du chorème peuvent être modélisés par les sommets-entités d'un graphe,
- les caractéristiques de ces objets (formes et légendes associées) peuvent être décrites par des attributs associés aux concepts,
- l'arrangement des objets peut être représenté par les sommets-relations d'un graphe.

Par ailleurs, nous nous plaçons dans le cadre d'une représentation hiérarchique des connaissances (Ducournau *et al.*, 1998, chapitres 10, 11, 12) : les concepts sont organisés selon une relation d'ordre, fondée sur une classification naturelle du domaine (par exemple *« les parcs sont des surfaces en herbe », « la bergerie est un bâtiment d'exploitation », « un puits est un aménagement »*). Cette relation d'ordre est utilisée pour établir une mesure de ressemblance entre les concepts et entre les graphes, et c'est donc sur elle que reposent les opérations de comparaison et d'adaptation du système ROSA.

Trois étapes de modélisation des connaissances

La construction du système ROSA s'est déroulé en trois étapes principales réunissant agronomes et informaticiens autour d'une problématique de modélisation.

Lors de la première étape, il s'est agi pour l'ensemble des chercheurs du projet de cerner les concepts que les agronomes utilisent pour décrire la structure spatiale et fonctionnelle d'une exploitation agricole. Ce travail s'appuie sur la confrontation des deux modes de représentation des connaissances que sont les chorèmes et les graphes (Capitaine *et al.*, 2001). En ressortent d'une part un ensemble de 'cas', i.e. des graphes représentant une organisation spatiale agricole assortie d'une explication fonctionnelle (voir Fig. 1), et, d'autre part, un ensemble de concepts décrivant les connaissances du domaine.

La deuxième étape consistait à expliciter des similarités spatio-fonctionnelles entre exploitations, sur la base de leurs représentations chorématiques et 'graphiques'. L'objectif était de mettre à jour des connaissances « d'adaptation », connaissances que les agronomes mettent en œuvre pour analyser une exploitation en référence aux exploitations précédemment étudiées. Ces connaissances doivent ensuite être formalisées et utilisées par le système pour '*adapter*' les explications d'un graphe à l'autre.

L'utilisation du système construit a fait l'objet de la troisième étape. Le couple agronome – informaticien fournit au système un graphe représentant la structure spatiale d'une exploitation. ROSA retourne un ensemble de cas qui sont soumis à évaluation. Ceci conduit tout à la fois à réviser les cas enregistrées dans la machine, à affiner le modèle des connaissances du domaine et à formaliser de nouvelles connaissances d'adaptation.

Ces différentes étapes ont donné lieu à des séances de travail dont certaines ont été filmées dans le cadre d'une collaboration avec des psychologues et des linguistes. Le corpus résultant a été partiellement analysé. Nous nous appuierons sur certains extraits pour traiter deux problématiques apparues dans le cours de ce travail, l'une concernant la modélisation des relations entre le temps et l'espace, l'autre la prise en compte des

3 Ici *rôle* est pris dans le sens des logiques des descriptions (Baader *et al.*, 2003), c'est-à-dire comme une relation entre deux concepts.

espaces 'peu importants'. Il s'agit d'exemples choisis parmi les nombreux autres problèmes soulevés dans ce projet de recherche[4].

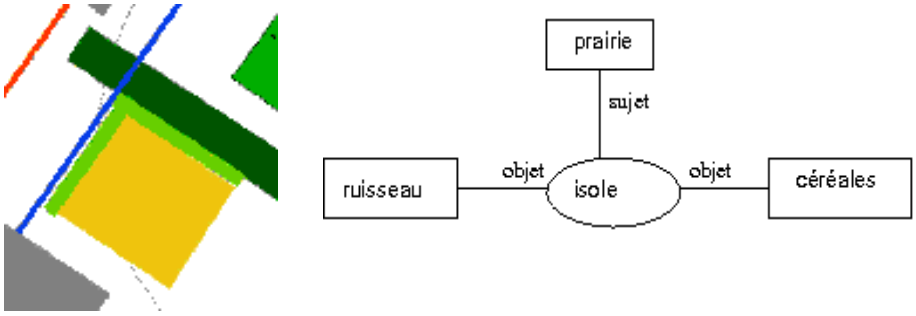


Figure 1. Partie du schéma chorématique d'une exploitation lorraine et graphe associé. Le cas obtenu est la réunion de ce graphe avec l'explication suivante : "l'agriculteur a placé une prairie pour isoler la parcelle de céréales du ruisseau afin de protéger les cultures de l'humidité".

Proche, loin et entre deux

Au cours de la séance 3 (en 2005), le groupe s'est intéressé à la modélisation du changement de statut de l'espace en fonction des modes de gestion des troupeaux (gardés et non gardés). À une moment de la séance, suivant l'utilisation du logiciel ROSA, la discussion s'est focalisée sur les notions de proche et de loin, qui peut recouvrir des réalités métriques différentes. La question précisément était de comparer les modes de gestion des parcs (troupeau non gardé) et des parcours (troupeau gardé) semblablement *proches* ou *éloignés* d'un bergerie. Cette discussion a donné lieu à la confrontation de différents modèles, *via* un tableau blanc sur lequel les partenaires peuvent intervenir tour à tour ou conjointement[5].

La première proposition est une proposition informatique (Fig. 2), sur la base de modèles de distance qualitative à granularité variable (Frank, 1992) : « *donc quand c'est gardé, je peux avoir proche repas, moyen jour et loin semaine* » « *et puis quand je suis dans le gardé je pourrais utiliser ça donc à ce moment là, j'aurais mon jour il viendrait jusque là et donc il faudrait du très loin pour avoir à la semaine* » (I1, 2005-E4, 0).

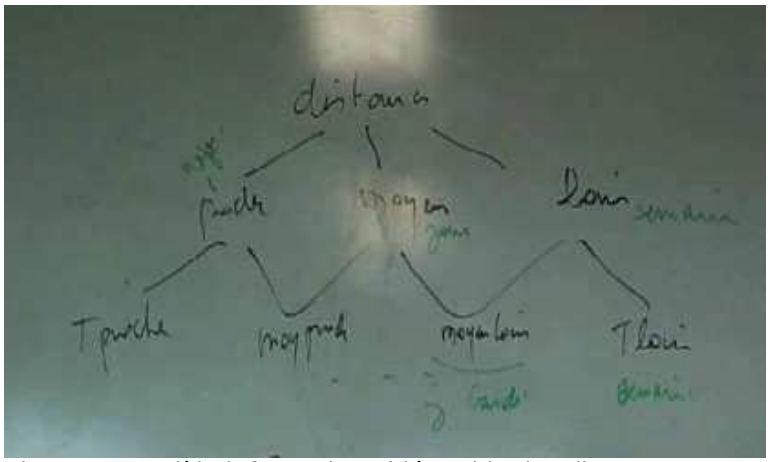


Figure 2 : modèle informatique hiérarchie des distances qualitatives et étiquetage par les temporalités, repas, jour, semaine (E4, 2'10).

4 Voir par exemple (Brassac et Le Ber, 2005) sur la modélisation des entités complexes.

5 Sont en présence deux agronomes (dénommés A1, A2), trois psycho-linguistes (P1, P2, P3), deux informaticiens (I1, I2).

Dans la foulée, le groupe valide le changement de granularité « *tu as eu besoin, toi I1, de descendre d'un grain, d'une granulatité à cause du gardiennage* » (P1, 2005-E4, 3'20) « *et du jeu de bascule* » (A1, 2005-E4, 3'26)

Mais ce modèle ne satisfait pas totalement l'agronome, qui éprouve le besoin de réécrire à sa façon les différentes zones spatio-temporelles sous le schéma précédent (Fig. 3). Il le fait en deux lignes superposées, initiées par les inscriptions G et G barré et qui représentent les zones 'repas', 'journée' et 'semaine', dans le cas de la garde (G) et de la non garde (G barré). Sur le schéma, sont indiquées des notions de distance (ordre et extension des rectangles) et sont figurées les clôtures. La zone de 'bascule' (jour, gardé vers semaine, non gardé) est entourée.



Figure 3 : Modèle agronomique : zones spatiaux-temporelles pour les modes gardé et non-gardé (E4, 5'30)

Au cours de l'interaction qui s'en suit, le modèle agronomique est complété de deux traits verticaux bien marqués, indiqués par des flèches sur la figure 4.

- Un premier trait, séparant jour et semaine est explicité ainsi « *là tu viens de redéfinir un proche-loin* » (P1, 2005-E4, 6'18) « *là je suis dans les allers-retours quotidiens, [...] c'est dans le mouvement quotidien et au-delà, c'est des séjours de plus longues périodes* » (A1, 2005-E4, 6'24).
- Le second trait, séparant les zones repas et jour intervient suite à la remarque complémentaire « *c'est pas tout à fait proche parce que dans proche il y avait suffisamment proche pour pouvoir surveiller* » (A2, 2005-E4, 6'37).

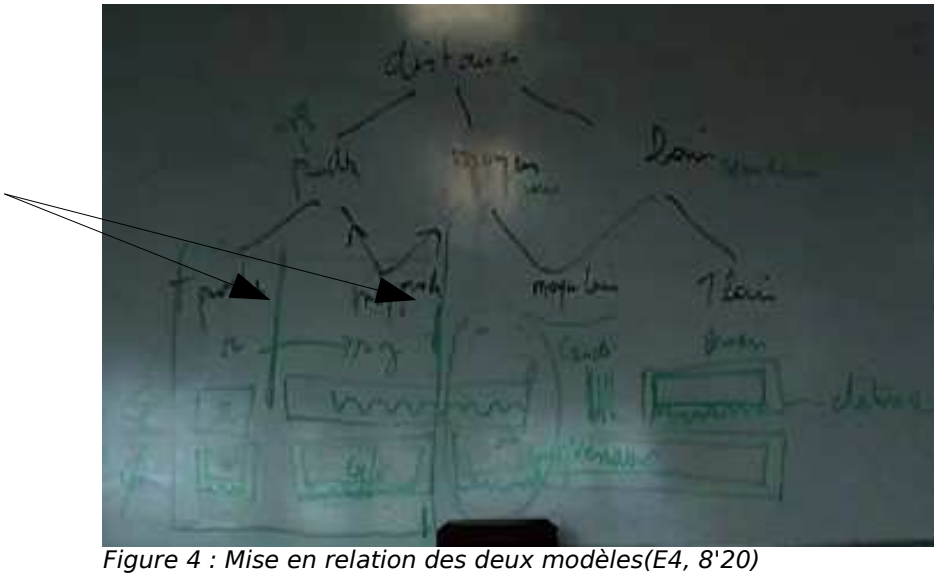


Figure 4 : Mise en relation des deux modèles(E4, 8'20)

Ces deux traits permettent d'inscrire le modèle agronomique en regard de la proposition informatique (Fig. 4). Ce qui suggère de revisiter le modèle informatique de la façon

suivante : les zones repas sont rattachées à la catégorie de distance '*très-proche*', les zones journée sont rattachées aux catégories '*moyen-proche*' dans le cas non gardé et '*moyen*' dans le cas gardé, les zones semaine sont rattachées aux catégories '*loin*' dans le cas non gardé et '*très-loin*' dans le cas gardé.

Ce qui est expliqué par A2 (2005-E4, 25'40, parlant de la zone journée gardée) : « *cet ensemble là, serait du parcours moyen-proche euh moyen* » (26'04), parlant des zones journée et semaine non gardées) « *et ça, ça deviendrait parc moyen-proche et ça ça deviendrait parc ... ben finalement ça dviendrait parc loin* ».

Ce modèle construit en commun autorise à la fois à distinguer et à relier des espaces gérés de façon différentes. Le parcours à distance *moyenne* peut en effet (doublement) se spécialiser en parc *moyen-proche*, qui peut être utilisé à la journée, ou en parc *moyen-loin*, qui est aussi un parc *loin* (par généralisation), utilisable seulement à la semaine.

Un sectionnal peu important

Nous présenterons ici rapidement un autre exemple où la confrontation des modèles s'est révélée riche pour l'ensemble du groupe. Nous nous situons ici dans le courant de la deuxième séance (en 2003) où il s'agit d'apparier des graphes d'exploitations différentes, afin de faire émerger des connaissances dites d'adaptation. La procédure est la suivante, deux agronomes et deux informaticiens disposent du graphe, G-cible, d'une exploitation à analyser. Les informaticiens piochent dans la base de cas (constitué précédemment), un graphe G-source qui leur paraît similaire à une partie de G-cible. L'appariement est discuté, puis validé ou invalidé par les agronomes.

Dans le moment présent, I1 propose d'apparier les graphes présentés en figure 5. Plus précisément, il propose de modifier le graphe cible en effectuant une inférence spatiale (Fig. 6) puis d'apparier les paires de sommets entités (parcours sectionnal – parc 4), (draille – route 1), (bergerie 1 – bergerie 1) et les deux somments-relations relie-proche. L'appariement est refusé par les agronomes, d'une part parce que l'inférence spatiale n'apparaît pas valide (« *le sectionnal il est proche quand même* » (I1, 2003-S2, 52'42) « *c'est pas dit dans la chanson* » (A1) « *c'est pas dit, je suis là tout à fait d'accord c'est pas dit* » (I2)), d'autre part parce que le sectionnal qualifié de « *pas important* » (A1, 2003-S2, 63'33) ne peut être apparié à un parc.

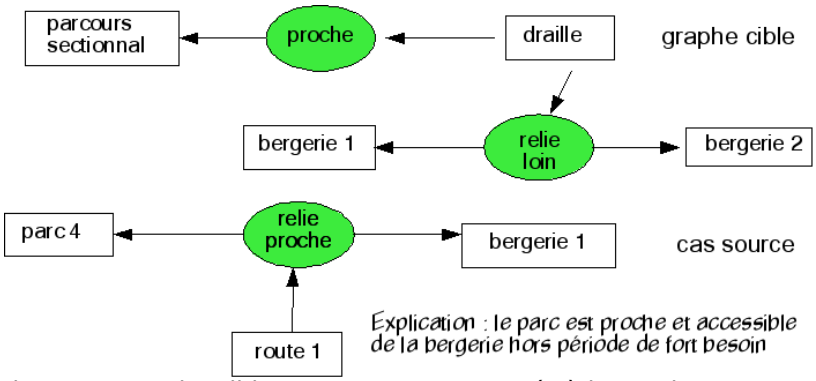


Figure 5 : Graphe cible et cas source proposés à l'appariement

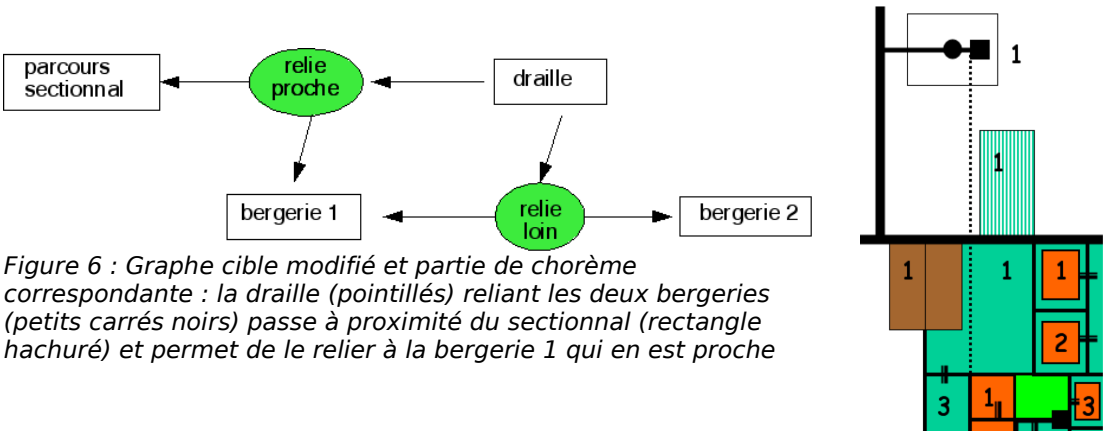


Figure 6 : Graphe cible modifié et partie de chorème correspondante : la draille (pointillés) reliant les deux bergeries (petits carrés noirs) passe à proximité du sectionnal (rectangle hachuré) et permet de le relier à la bergerie 1 qui en est proche

Toutefois, l'analyse de cette partie du graphe cible et du chorème correspondant conduit A2 à souligner l'intérêt du parcours sectionnal, situé le long de la draille reliant les deux bergeries (2003-S2, 67'00) : « *ça c'est traversable en toute saison c'est pas très important ce n'est jamais stratégique dans le bilan alimentaire mais euh c'est accessible tout le temps sans conséquences sans problèmes et c'est peut-être pas indifférent de ce point de vue donc c'est pour sauver l'intérêt de ce sectionnal [...] donc je vote une clause de suspension sur on ne peut rien conclure c'est une commodité sans enjeu perceptible mais c'est une commodité je dirais passe commodément à côté de, la draille passe commodément à côté du parcours sectionnal* »

Conformément au modèle du système ROSA, on a là l'énoncé d'un cas : le graphe cible initial est transformé – la relation 'proche' entre draille et sectionnal est remplacée par la relation 'passe à côté de' – et doté d'une explication qu'on peut résumer ainsi : « le sectionnal situé le long de la draille est sans enjeu mais peut être utisé commodément, en toutes saisons ». De plus au cours de l'interaction, les agronomes ont énoncé une règle générale, qui permet d'utiliser la connaissance de ce cas pour analyser d'autres situations. On remarquera pour finir que c'est une proposition des informaticiens invalidée par les agronomes qui conduit néanmoins à la construction d'un objet de connaissance commun.

Conclusion

Ce texte présente quelques exemples tirés d'un travail de modélisation de connaissances entrepris entre agronomes et informaticiens en collaboration avec des psycho-linguistes[6]. Quelques séances de travail ont été filmées ce qui nous permet d'analyser finement les problématiques de modélisation en jeu. Les exemples choisis mettent en exergue la question de l'entre-deux, du 'peu important' dans l'organisation des exploitations agricoles. Ils montrent que ces notions sont souvent oubliées et soulèvent des difficultés de modélisation. Toutefois l'interaction entre les différentes disciplines permet de progresser sur ces questions, parfois de faire émerger des éléments laissés de côté, et de faire évoluer les modèles de part et d'autre. Nous montrons également que l'interaction conduit effectivement à l'enrichissement du système ROSA et valide d'une certaine façon le modèle de raisonnement à partir de cas pour l'analyse des organisations spatiales d'exploitations agricoles.

6 Voir (Brassac, 2004 ; Mondada 2004).

Remerciements

S. Lardon, P.-L. Osty, M. Capitaine, J.-L. Metzger, J. Lieber, L. Mondada ont participé à ce projet. Les captations et leur transcription ont été possibles grâce au financement du programme GETM (CNRS-IGN-Cemagref).

Références bibliographiques

Baader F., Calvanese D., McGuinness D., Nardi D., Patel-Schneider P., 2003. *The Description Logic Handbook - Theory, Implementation and Applications*. Cambridge University Press.

Brassac C., 2004. Action située et distribuée et analyse du discours : quelques interrogations. *Cahiers de Linguistique Française*, vol. 26, pp. 251-268.

Brassac C., Le Ber F., 2005. Inscription spatiale d'une activité cognitive collective de représentation de l'espace. *Intellectica*, vol. 2-3, n° 41–42, p. 181–200.

Capitaine M., Lardon S., Le Ber F., Metzger J.-L., 2001. Chorèmes et graphes pour modéliser les interactions entre organisation spatiale et fonctionnement des exploitations agricoles. *In* : *Géomatique et espace rural (Journées de la recherche CASSINI Montpellier)*, Libourel T. & Maurel P. (eds), Éditions CIRAD, pp. 145-163.

Chein M., Mugnier M.-L., 1992. Conceptual graphs: Fundamental notions. *Revue d'Intelligence Artificielle*, vol. 6, n° 4, pp. 365-406.

Ducournau R., Euzenat J., Masini G., Napoli A., éditeurs, 1998. *Langages et modèles à objets — État des recherches et perspectives*. INRIA, Collection Didactique D-019.

Frank A., 1992. Qualitative Spatial Reasoning about Distances and Directions in Geographic Space, *Journal of Visual Languages and Computing*, vol. 3, pp. 343-371.

Lardon S., Capitaine M., Benoît M., 2000. Les modèles graphiques pour représenter l'organisation spatiale des activités agricoles. *In : Représentations graphiques dans les systèmes complexes naturels et artificiels*, ENST, pp. 127-150.

Lardon S., Osty P.-L., 2004. Diversité des modes de conduite des élevages et organisation spatiale des lieux-dits : des atouts pour mieux gérer le territoire ? *In : Colloque "Gérer les paysages de montagne, pour un développement concerté et durable".* CdRom CEP Florac.

Leake D. B., éditeur, 1996. *Case-Based Reasoning. Experiences, Lessons, Future Directions*. AAAI Press / The MIT Press.

Le Ber F., Napoli A., Metzger J.-L., Lardon S., 2003. Modeling and comparing farm maps using graphs and case-based reasoning. *Journal of Universal Computer Science,* vol. 9, n° 9, pp. 1073-1095.

Lieber J., 1997. Raisonnement à partir de cas et classification hiérarchique. Application à la planification de synthèse en chimie organique. Thèse de l'Université Poincaré - Nancy I.

Mondada L., 2004. Temporalité, séquentialité et multimodalité au fondement de l'organisation de l'interaction : Le pointage comme pratique de prise du tour. *Cahiers de Linguistique Française*, vol. 26, pp. 269-292.

Riesbeck C. K., Schank R. C., 1989. *Inside Case-Based Reasoning*. Lawrence Erlbaum Associates.

Sowa J.F., 1984. Conceptual Structures: Information Processing in Mind and Machine. Addison Wesley.