



HAL
open science

Confrontation de points de vue sur des corpus documentaires : Le cas de la modélisation du temps archéologique

Tiphaine Accary-Barbier, Aurélien Bénel, Sylvie Calabretto, Andrea Iacovella

► To cite this version:

Tiphaine Accary-Barbier, Aurélien Bénel, Sylvie Calabretto, Andrea Iacovella. Confrontation de points de vue sur des corpus documentaires : Le cas de la modélisation du temps archéologique. 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle, Jan 2004, Toulouse, France. pp.197-205. hal-01514838

HAL Id: hal-01514838

<https://hal.science/hal-01514838>

Submitted on 18 Nov 2021

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.

Confrontation de points de vue sur des corpus documentaires : Le cas de la modélisation du temps archéologique

View points conflicts on document corpora : The case of archaeological time modelling

Tiphaine Accary^{1 2}

Aurélien Bénel^{1 2}

Sylvie Calabretto¹

Andréa Iacovella²

¹ LIRIS, Laboratoire d'InfoRmatique en Image et Systèmes d'information
CNRS / INSA de Lyon / Université Lyon 1 / Université Lyon 2 / École Centrale de Lyon

² École française d'archéologie d'Athènes (EFA)

L.I.R.I.S, bât. Blaise Pascal
20 avenue Albert Einstein,
69621 Villeurbanne Cedex

École Française d'Athènes
6 rue Didotou
106 80 Athènes - GR

{taccary,abenel,scalabretto}@liris.cnrs.fr
andrea.iacovella@efa.gr

Résumé

Nos précédents travaux sur les bibliothèques numériques spécialisées montraient que chaque expert était amené à enrichir le corpus documentaire avec son point de vue (exprimé sous la forme d'un modèle de connaissance). Ces points de vue pouvant être contradictoires, il est nécessaire d'offrir aux communautés d'experts des outils, spécialisés à leur domaine, leur permettant de détecter certaines de ces contradictions.

Nous proposons ici un assistant à la modélisation du temps archéologique utilisant les méthodes de la discipline et permettant la détection et la correction interactive des incohérences.

En effet, la construction conceptuelle du temps archéologique est particulièrement représentative, par son caractère complexe et méthodique, des domaines dans lesquels les techniques de modélisation de connaissances peuvent devenir un support pour une « écriture dynamique » [3] et par là ouvrir un espace de débat entre experts.

Mots-clés

Modélisation du temps, bibliothèques numériques, détection d'incohérences, documents archéologiques.

Abstract

We have shown in our previous works that, in specialized digital libraries, the document corpus is augmented by the

point of view of every expert. These points of view may be contradictory and communities need some specialized tools to help them detecting inconsistencies.

In this paper, we propose a tool dedicated to the modelling of archaeological time enabling interactive inconsistencies detection and correction, and based on methods usually used in archaeology.

Indeed, the conceptual construction of archaeological time is a pretty good example (very complex and methodic) of domains where knowledge modelling techniques could be used for "dynamic writing" [3], and as a debate place for experts.

Keywords

Time modelling, digital library, inconsistencies detection, archeological documents.

1 Introduction

Le but de notre travail est d'offrir à la communauté des chercheurs en archéologie des outils leur permettant d'exploiter leurs corpus de documents. Mais, parce que la bibliothèque virtuelle ne peut se limiter à un « entrepôt » de documents, il est indispensable d'y prévoir des « lieux » d'interaction sociale. L'importance de cette dimension sociale s'explique par le fait qu'un document ne comporte pas un sens, mais plutôt des sens donnés par les conditions de sa lecture. Toute lecture est donc une réécriture.

Nos travaux antérieurs, matérialisés dans le système *Porphyre*¹, se sont attachés à garder trace des parcours d'interprétation des lecteurs [4].

Le point de vue du lecteur se manifeste entre autres par la restructuration du corpus. Cette restructuration se fait à l'aide de ce que nous avons appelé un réseau de description, organisant des descripteurs de contenu dans un graphe orienté acyclique .

La question qui nous préoccupe aujourd'hui est celle de la publication. En effet, par celle-ci, l'intervention de l'expert devient immédiatement visible dans une sorte d'agora virtuelle. Afin que les débats menés dans cette agora puissent en retour devenir références pour l'expert, un travail « d'édition » est nécessaire. C'est en effet la crédibilité de l'expert et celle de la communauté qui sont en jeu dans ce délicat travail de construction d'un Tout cohérent.

Par rapport à ces fonctions de régulation communautaire, deux orientations nous semblent possibles :

- des outils génériques permettant de détecter des ressemblances entre différents points de vue (et par là de permettre à l'expert de voir les limites de cette ressemblance),
- des outils spécifiques (à un champ de significations normé par une méthode), permettant de détecter directement les désaccords entre experts.

Dans cet article nous proposons un outil spécifique aux descripteurs temporels basé sur les modèles développés en « Intelligence artificielle ».

Le temps est souvent présenté par les archéologues comme l'une des trois dimensions principales de leur discipline avec l'espace et les types de vestiges. Mais comme nous le verrons dans une première partie, il ne s'agit en aucune manière de « dimensions » au sens mathématique, le temps étant construit à partir des deux autres.

Dans une deuxième partie, nous définirons un modèle de représentation du temps répondant à nos attentes en étudiant les spécificités des modèles utilisés dans les autres domaines.

Nous montrerons ensuite de quelle façon notre modèle peut coopérer avec l'existant et définirons des modèles d'interaction permettant de faire l'interface entre l'archéologue et le système.

Nous nous attacherons enfin à l'aspect principal du modèle, c'est-à-dire les processus de détection et de présentation des incohérences.

2 Un temps « déroutant »

Comme le rappellent Patrick Desfarges et Bruno Helly [7] : « contrairement à une opinion bien ancrée dans l'esprit des archéologues, le temps, la chronologie n'est pas une donnée mais un résultat ». Le temps archéologique est donc

construit à partir d'indices spatiaux, stylistiques ou naturels en utilisant des heuristiques plus ou moins fiables. Cette construction a pour but, selon Alain Gallay [9], de rechercher un ordre pour saisir l'événement et le changement.

On distingue habituellement les chronologies dites « relatives » des chronologies dites « absolues ». Dans les deux cas, ces chronologies sont construites en comparant les indices cités précédemment, la comparaison étant qualitative dans le premier cas et quantitative dans le second.

2.1 Indices stylistiques

L'une des méthodes les plus utilisées pour dater un objet manufacturé est de se baser sur son style. En effet, on suppose que deux objets analogues dans leur forme ou leur technique sont contemporains.

Cependant les indices stylistiques sont souvent très dépendants des aspects spatiaux. En effet, si l'on adhère à la théorie du diffusionnisme, on considérera qu'un style peut apparaître avec un certain retard dans une contrée voisine. De plus, deux styles peuvent parfois être analogues malgré un écart chronologique et spatial abyssal.

Par le passé, certaines heuristiques évolutionnistes ont parfois été utilisées pour donner une séquence des styles dans le temps, comme par exemple considérer qu'un style abstrait était souvent suivi d'un style figuratif puis d'un style « décadent ». Notons que ce genre d'heuristique est à prendre avec beaucoup de prudence, et que très souvent les théories qui en sont issues ont été réfutées.

2.2 Indices spatiaux

L'indice spatial le plus courant est la profondeur relative des vestiges. La méthode alors utilisée, appelée « stratigraphie », part du principe que :

« Dans une série de couches, telles qu'elles ont été créées originellement, les unités supérieures de stratification sont les plus jeunes, et les inférieures les plus anciennes, car chacune a dû être déposée sur -ou créée par enlèvement de- une masse préexistante de stratification archéologique. »

(E.-C. Harris, *Principles of Archaeological Stratigraphy*, Academic Press, Londres, 1979, cité dans [11])

Notons que cette définition s'applique à la fois aux strates d'origine naturelle (par sédimentation) et artificielle (couche de démolition ou d'incendie pour construire au-dessus). Dans les deux cas on se basera sur des différences de couleur, de texture et de granulométrie pour identifier les limites d'une couche.

La règle de séquence temporelle admet évidemment des contre-exemples dus à des remaniements naturels ou humains. La méthode peut toutefois tenir compte des cas les

1. <http://www.porphiry.org>

plus courants :

- les fosses : dont le contenu est postérieur au sol dans lequel elles sont creusées,
- les élévations (murs) : dont la construction est antérieure à toutes les couches au-dessus de la couche d'habitation primordiale.

Encore une fois des indices de différents types sont croisés. Ainsi en est-il de la stratigraphie et du style des objets contenus dans les couches. On dit parfois abusivement qu'un vestige date une couche. En fait, il lui fournit un *terminus post quem*². Par exemple, si les archéologues du futur découvrent le musée du Louvre, ils ne devront pas dire que tel pavillon date de l'époque égyptienne ou babylonienne, mais que le palais est postérieur à la fois à la civilisation égyptienne et à la civilisation babylonienne.

Si la profondeur est l'indice spatial le plus utilisé pour la datation, dans certains cas, la position géographique d'un vestige pourra également l'être. En effet, nous avons vu que les adeptes du diffusionnisme considèrent qu'un style va apparaître en un foyer et se diffuser petit à petit dans les régions voisines. La distance au foyer pourrait alors être révélatrice du temps nécessaire à la diffusion. Un tel indice est bien sûr à prendre avec beaucoup de précaution.

2.3 Indices naturels

Les physiciens et les chimistes disposent de théories quantitatives sur l'évolution dans le temps de certaines propriétés des vestiges.

La datation au carbone 14, par exemple, repose sur l'hypothèse qu'une part constante du dioxyde de carbone de l'atmosphère est rendu radioactif par les rayons cosmiques : leur atome de carbone devient l'isotope ^{14}C . Ainsi, les plantes, en fixant le carbone de l'air, et les animaux en mangeant ces plantes, contiennent la même proportion $^{14}\text{C}/^{12}\text{C}$ que l'air. Le radiocarbone ayant une durée de vie fixe (mais probabiliste), sa concentration va décroître avec le temps et permettre ainsi de mesurer le temps écoulé depuis la mort de l'être vivant.

Il est intéressant de noter que pour des époques historiques comme celle de L'Égypte ancienne, les archéologues obtenaient des datations au radiocarbone incohérentes avec les textes. Les physiciens ont dû alors accepter de réfuter leur hypothèse selon laquelle les rayonnements cosmiques avaient été constants dans le temps. Il a donc fallu ré-évaluer les datations au radiocarbone en fonction de la dendrochronologie (basée sur la succession de cernes d'arbres plus ou moins épais suivant les époques de sécheresse). Mais ceci repose un problème de comparaison : pour faire correspondre une suite de cernes avec une autre, il faut supposer que deux zones géographiquement distinctes ont eu

2. La fin d'occupation de la couche est postérieure au premier usage attesté de ce type d'objet.

au même moment un enchaînement climatique identique (ex : deux années sèches, puis une année de crues, suivie de trois années moyennes). Comment alors calibrer le radiocarbone en Europe quand on sait que l'arbre fossile le plus ancien se situe en Amérique ? Autrement dit, en archéologie, même les datations dites « absolues » ne le sont pas vraiment ?

On pèse ici toute la complexité du temps en archéologie, fruit de comparaisons dont la validité diminue avec la distance entre les deux lieux comparés. D'ailleurs le système d'information annoncé par J. P. Demoule [6] tenait compte du fait qu'en archéologie il n'existe pas une chronologie unique mais des chronologies valables dans une région géographique donnée, avec parfois des correspondances possibles entre elles.

2.4 Et le temps de Porphyre ?

Dans l'outil Porphyre [5], autour duquel s'axe le projet, on adopte pour décrire les documents non plus des descripteurs isolés comme dans les systèmes usuels d'indexation, mais des descripteurs organisés dans un ordre partiel conformément à des relations explicites de spécialisation (Fig.1).

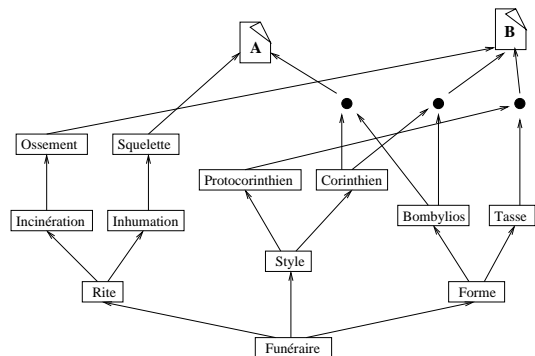


FIG. 1 – Les relations du réseau de descripteurs peuvent être des instanciations, des subsomptions ou des liens de composition.

Ces descripteurs peuvent avoir une généralisation ou une spécialisation commune. Lorsque l'on observe une spécialisation de A vers B, il faut l'interpréter comme "tout document décrit par B l'est aussi par A".

Ces relations peuvent être interprétées comme :

- des subsomptions,
- des instanciations,
- des liens d'association,
- des liens de composition.

Par exemple, sur la figure 1, le document B est caractérisé par un objet de forme Bombylios et de style Corinthien (composition), un autre de forme Tasse et de style Protocorinthien ainsi que d'ossements résultant d'un rite d'incinération.

ration.

Dans le cas de la modélisation du temps archéologique, ces descripteurs sont alors des couches ou des styles interprétés comme des intervalles de temps auxquels peuvent être rattachés les fragments recueillis.

Puisque les réseaux de description de Porphyre s'organisent conformément à des relations de spécialisation, on peut en déduire que, dans le cadre des descripteurs temporels, ils permettent d'exprimer les relations d'inclusion. Ainsi, sur la figure 2, la Période Romaine apparaît comme une partie des Temps Historiques.

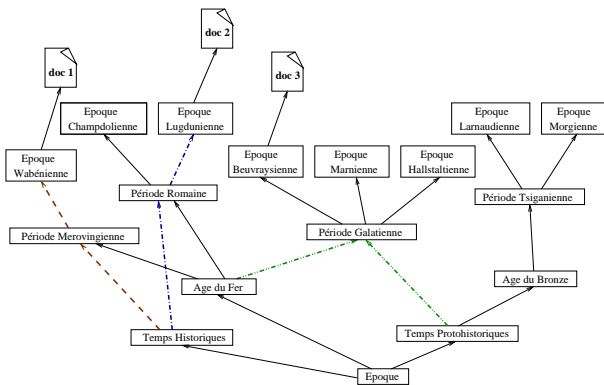


FIG. 2 – Seulement deux types de relations sont exprimées entre descripteurs temporels : l'inclusion et le chevauchement.

Le système va aussi permettre l'expression du chevauchement entre deux périodes. Ainsi, on peut constater que la Période Galatienne appartient conjointement à l'Age du Fer et aux Temps Protohistoriques. On peut en déduire que la Période Galatienne appartient à l'intersection de ces deux périodes.

Cependant, dans le système Porphyre actuel, le réseau de description ne permet pas d'exprimer d'autres types de relations. Il serait pourtant très utile de pouvoir établir une "chronologie" des descripteurs. En effet, si on compare l'Epoque Wabénienne et l'Epoque Lugdunienne, le seul élément de comparaison dont on dispose est l'appartenance commune de ces deux époques aux Temps Historiques. On aimerait alors savoir dans quel ordre ils se placent à l'intérieur de cette période.

De même, pour les couches stratigraphiques, le réseau permet d'exprimer qu'une couche appartient à une époque donnée, mais on a aucun moyen de connaître l'ordre stratigraphique des couches.

Il apparaît donc nécessaire de munir le réseau de descripteurs de relations transversales permettant de signifier les rapports de connexité et de simultanéité.

De plus, même les relations d'inclusion exprimées gagneraient à être distinguées en trois types particuliers : l'inclu-

sion, l'inclusion en borne inférieure et l'inclusion en borne supérieure.

Ainsi se dégage le besoin de définir un modèle de représentation du temps valide pour l'archéologie et capable de s'intégrer à la structure déjà existante du système Porphyre.

3 Proposition de modèle

3.1 Les attentes du modèle

Pour choisir notre modèle, nous allons définir les attentes propres à la représentation du temps archéologique.

Comme nous l'avons montré dans la partie 2, l'archéologie est un domaine où la plupart des datations manipulées sont relatives et où les événements sont des périodes et non des instants. Par conséquent, notre modèle doit permettre l'expression d'intervalles de temps ordonnés de manière relative.

De plus, le chercheur connaît rarement la totalité des informations temporelles des événements qu'il manipule. Il nous faut donc définir un modèle permettant d'exprimer des données incomplètes.

Langages d'interaction.

Le chercheur en archéologie manipule le temps à la fois sur le terrain (expérimentation) et dans ses pratiques d'érudition (hypothèses). Dans les deux cas, cette manipulation est liée à un travail documentaire d'écriture et de lecture. Les connaissances temporelles de l'archéologue prennent alors diverses formes selon l'activité à laquelle elle se rattache. Entre autres, elle pourront prendre la forme de diagramme stratigraphique dans le cas de l'analyse d'un site, ou de frise chronologique dans le cas d'une synthèse.

Nous nous sommes attachés aux pratiques conduisant à la production des diagrammes stratigraphiques et des synthèses chronologiques.

Le diagramme stratigraphique permet d'ordonner temporellement les différentes couches d'un site et de faire apparaître les phases de construction, de remblaiement, d'occupation et de démolition. De plus, chaque objet datable dans une couche, livrera pour celle-ci un *terminus post quem* duquel on pourra inférer des informations temporelles pour l'ensemble du site.

La frise chronologique, quant à elle, permet à l'archéologue de faire une synthèse graphique de ses résultats et de ceux de ses collègues afin de proposer certaines hypothèses pour une région ou une période plus étendue.

On a donc besoin de définir des langages d'interaction permettant de réaliser des assistants pour ces tâches particulières. Pour que notre modèle soit le plus évolutif possible, ces langages, externes au noyau, ne doivent être utilisés que lors de l'interaction chercheur-machine. Ces langages, servant d'interface entre l'archéologue et le système, permettent d'entrer les informations sous une forme plus in-

tuitive que celle du modèle de calcul : le système, à l'aide d'un certain nombre de règles, se chargera ensuite de les convertir.

Modèle pivot et détection d'incohérence.

Dans le cadre de l'épistémologie contemporaine, une théorie n'est considérée comme hypothétiquement vraie que tant qu'elle n'a pas été réfutée par une expérience. Le modèle qu'il nous faut doit donc permettre au chercheur de faire le point sur les incohérences que son corpus pourrait contenir entre strates (expériences) et frises (théories). Ceci nécessite donc de les traduire dans un modèle pivot. Cependant cette détection n'aura d'intérêt que si elle s'accompagne d'un mécanisme de correction. On a donc besoin d'un modèle qui puisse permettre à l'utilisateur de visualiser et de traiter les éléments facteurs d'incohérence.

Espace de travail.

Vérifier la cohérence des modèles produits par l'ensemble de la discipline aurait d'une part un coût algorithmique exorbitant et serait d'autre part d'une utilité très limitée. Aussi, nous préférons offrir à l'archéologue un "espace de travail" dans lequel il pourra importer les périodes stylistiques et les couches stratigraphiques qu'il souhaite, et c'est sur les relations entre ces seuls objets que s'effectuera la détection d'incohérence.

En résumé, l'utilisation d'un espace de travail privé présente les avantages suivants :

- elle permet la mise en évidence des pôles de recherche,
- elle limite les calculs, puisque le chercheur ne travaille que sur les données qui l'intéressent,
- elle offre la possibilité de laisser de côté les incohérences dues aux interactions avec d'autres courants de pensée.

3.2 Quel temps faut-il ?

Actuellement, la représentation du temps intervient dans de nombreux domaines d'application tels que la planification, le traitement du langage naturel et la supervision de processus [8].

Avant d'établir un modèle de représentation du temps propre au domaine archéologique, il est nécessaire d'avoir une vision d'ensemble des modèles déjà utilisés dans les autres domaines pour figurer le temps.

On a vu, dans les attentes du modèle, qu'en archéologie le temps manipulé était principalement relatif. Il paraît donc évident d'exclure de notre choix les modèles de représentation fondés sur un système de datation absolue. On peut donc éliminer les représentations élémentaires utilisant l'ordonnancement de dates [2], ainsi que les modèles comme les diagrammes de PERT.

On a aussi vu que les événements archéologiques correspondaient à des périodes de temps. On peut donc aussi éliminer les modèles dont les primitives sont des instants

(événements ponctuels). Les modèles de Dean et McDermott [2] qui présentent une approche fondée sur la notion de mesurabilité des événements ne peuvent donc pas convenir à notre domaine.

Les modèles de base utilisant la propagation des contraintes définis par Vilain et Kautz [13] présentent, grâce à leur complétude, l'avantage que toutes les relations entre les événements soient définies à tout instant (Fig.3).

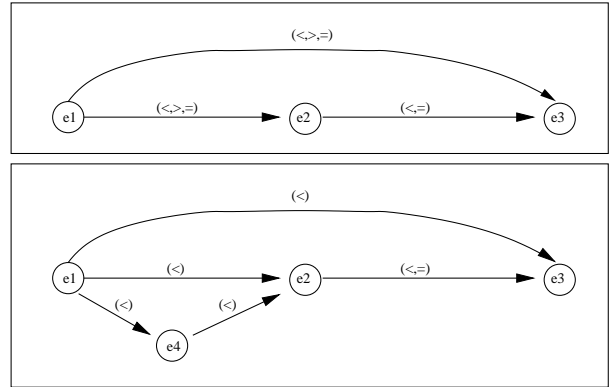


FIG. 3 – L'ajout d'une information va contraindre les relations du graphe et la propagation va permettre d'affiner les relations "floues".

Tous les arcs sont étiquetés (au moins par la disjonction totale des relations) et on ne risque donc pas de voir apparaître de fausses relations. Cependant ces modèles simples ne possèdent pas assez de relations (<,>=) pour exprimer les rapports temporels nécessaires à l'archéologie.

Le modèle qui s'adapte le mieux au temps archéologique est donc celui de Allen : une version plus complexe du modèle de Vilain, qui s'appuie sur treize relations exhaustives et exclusives possibles entre deux intervalles (Fig.4). La propagation des contraintes se fait grâce à un ensemble de règles réunies dans la table de composition des treize relations.

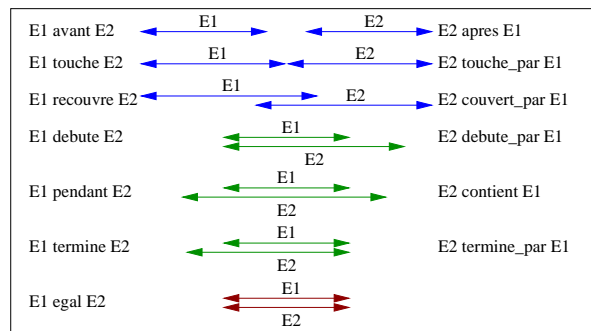


FIG. 4 – Les 13 relations du modèle de Allen permettent d'exprimer la connexité, l'inclusion et la simultanéité entre événements.

Par l'expression de disjonctions, le modèle permet d'expri-

mer les incertitudes propres à l'archéologie. De plus la propagation des contraintes le long du graphe lors de l'ajout d'une nouvelle information permet de maintenir le graphe complet à tout instant et ainsi de détecter les incohérences induites. Cette faculté du modèle est donc tout à fait adaptée au travail archéologique dont une tâche essentielle est la réfutation d'hypothèses.

Nous avons donc choisi de prendre le modèle de Allen comme modèle pivot, en l'étendant aux applications archéologiques par l'ajout de langages dédiés.

4 Mise en pratique

L'idée est de confronter d'une part le modèle local, qui est issu de la description de fouilles et qui permet l'expression de diagrammes stratigraphiques, et d'autre part le modèle général, qui exprime une théorie du domaine (partagée par la communauté).

4.1 Chronologies stylistiques

L'importation d'une partie du modèle général dans l'espace de travail nécessite sa "traduction" en un graphe muni des relations de Allen.

On a vu dans l'introduction que le réseau de descripteurs, dans sa forme première, ne permettait d'exprimer que des relations d'inclusion temporelles. La première démarche est donc de permettre l'expression d'un ordre dans les inclusions au sein du réseau.

Pour cela on a d'abord besoin de typer les descripteurs qui concernent nos éléments chronologiques. On va les marquer comme "Période stylistique" afin de signifier que ces descripteurs sont soumis à d'autres types de relations que les descripteurs standards de Porphyre.

On va ensuite ajouter la relation "juste-avant" dans le réseau, pour obtenir un réseau avec une orientation temporelle (Fig.5).

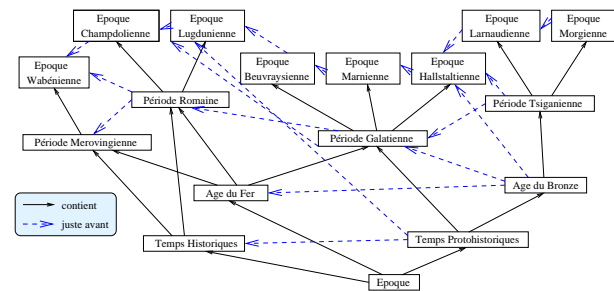


FIG. 5 – Pour exprimer la connexité, on va ajouter une nouvelle relation entre les descripteurs temporels : juste-avant.

Pour obtenir le modèle général, il faut ensuite traduire le réseau de descripteurs de Porphyre en un graphe de Allen. Nous proposons la traduction des relations d'inclusion et

de connexité maintenant exprimées sur le réseau en disjonctions de relations de Allen (Tab.1).

Relations entre descripteurs	Conversion en Allen
A juste-avant B	$touche(A,B)$
A contient B	$contient(A,B)$ ou $termine(B,A)$ ou $debute(B,A)$ ou $egal(A,B)$

TAB. 1 – Réseau de descripteurs et relations de Allen

On transformera ensuite le graphe en graphe complet en étiquetant tous les arcs non renseignés de la disjonction des 13 relations de Allen puis en appliquant la propagation des contraintes sur ces arcs (Fig.6).

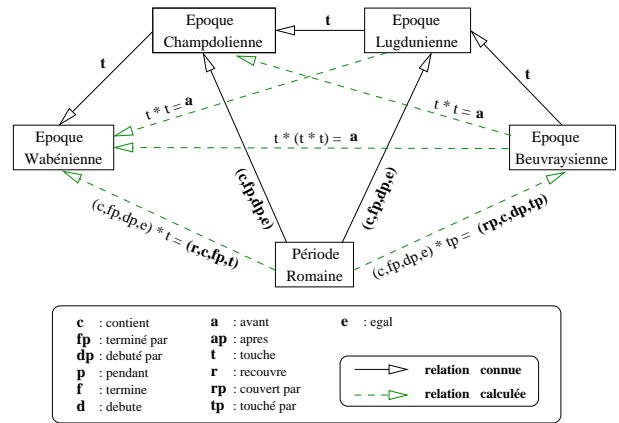


FIG. 6 – La conversion des relations entre descripteurs en relations de Allen permet d'appliquer la propagation et de calculer les relations entre tous les événements.

Notre modèle général peut alors être importé dans l'espace de travail, et comparé au modèle local.

4.2 Stratigraphies

Dans le modèle local, les fragments identifiés sur les stratigraphies sont ordonnés en utilisant les relations propres aux diagrammes stratigraphique : sur, sous et Synchro (Fig.7).

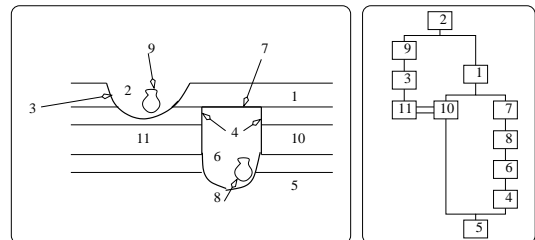


FIG. 7 – Les relations entre les éléments stratigraphiques identifiés sur la coupe sont réduites au nombre de trois sur le diagramme : sur, sous et synchro.

Les descripteurs correspondant aux fragments identifiés sur les stratigraphies vont être marqués comme descripteurs de type ‘‘Couche’’.

Ces descripteurs vont être ensuite rattachés à des styles préexistants. Lors de cette opération, les relations stratigraphiques entre descripteurs de type ‘‘couches’’ vont être traduites en relations de Allen en utilisant la table de correspondance que nous avons définie (Tab.2).

Stratigraphie	Conversion en Allen
A sous B	$avant(A,B)$ ou $touche(A,B)$
A synchro B	$recouvre(A,B)$ ou $recouvre(B,A)$ ou $egal(A,B)$
A sur B	$apres(A,B)$ ou $touche(A,B)$

TAB. 2 – Relations stratigraphiques traduites en Allen

Lorsque l’on va confronter le modèle général au modèle local dans le cadre des stratigraphies (Fig.8), nos deux modèles seront donc exprimés dans le même langage et leur comparaison deviendra alors possible.

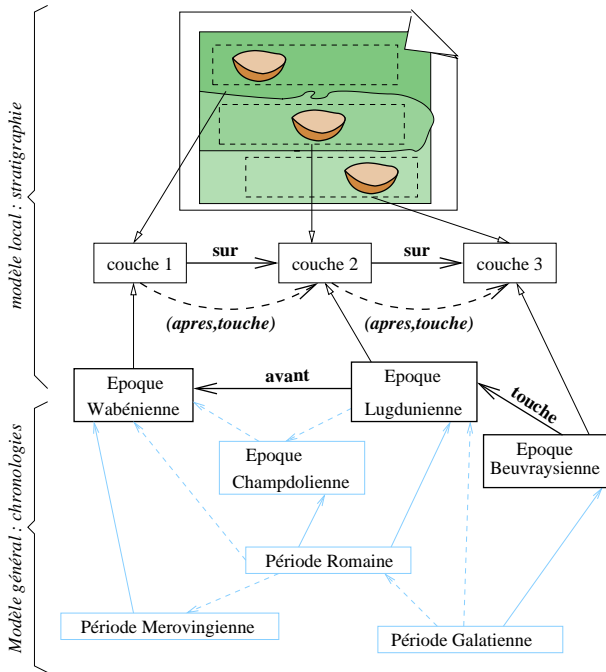


FIG. 8 – Après conversion des relations stratigraphiques en disjonction de Allen, la comparaison du modèle local au modèle général se fait dans le même langage.

5 Détection d’incohérences

La propagation va permettre, à chaque ajout d’information, d’obtenir une caractérisation plus fine de la position relative des entités. Si le système devient sur-contraint, on proposera à l’utilisateur les relations coupables potentielles en lui permettant de corriger les hypothèses.

5.1 Propagation des contraintes

Nous utilisons la table de transitivité fournie par Allen pour développer les relations possibles entre chaque paire d’intervalle. On notera I les intervalles, R les relations. Pour représenter le graphe formé par les événements i, j et k et les relations R_{ij} et R_{jk} entre ces événements, nous noterons le graphe :

$$I_i \text{ ---}(R_{ij})\text{ ---} I_j \text{ ---}(R_{jk})\text{ ---} I_k$$

Pour un système à N événements, le graphe des relations peut être représenté par un tableau de $N \times N$ disjonctions de relations élémentaires (Fig.9). Chaque cellule du tableau représente la liste des relations de l’événement i vers l’événement j . On peut constater que chaque case de la matrice triangulaire inférieure contient l’inverse des relations contenues dans la case qui lui est symétrique par rapport à la première diagonale et que celle-ci ne contient que des relations d’équivalence. Pour des raisons d’optimisation du stockage, nous n’utilisons donc que la matrice triangulaire supérieure du tableau, en faisant appel, si besoin, à la fonction $Inverse()$ ³.

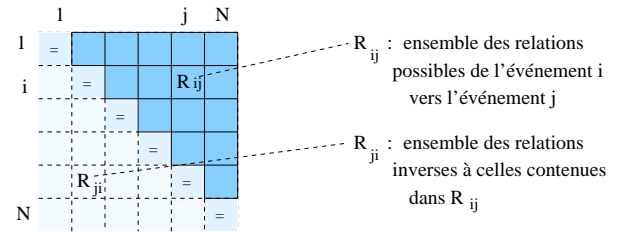


FIG. 9 – Les relations entre N événements sont stockées dans la matrice Triangulaire $(N-1) \times (N-1)$.

On cherche maintenant à déterminer les relations possibles R_{ik} entre I_i et I_k connaissant les contraintes R_{ij} et R_{jk} . On a alors $NouveauR_{ik} = R_{ik} \cap Compose(R_{ij}, R_{jk})$ où

$$\begin{aligned}
 &Compose(R_{ij}, R_{jk}) \\
 &C \leftarrow \emptyset; \\
 &\forall r_1 \in R_{ij} \\
 &\quad \forall r_2 \in R_{jk} \\
 &\quad C = C \cup TableTransitivite(r_1, r_2); \\
 &Renvoyer C;
 \end{aligned}$$

L’ajout d’une nouvelle relation permet de réduire les ensembles de relations possibles. Si on sait maintenant que la liste de relations possibles entre i et j a été modifiée, il est nécessaire de propager ces nouvelles contraintes au reste du graphe. Pour réaliser cette propagation, nous proposons une adaptation de l’algorithme de fermeture transitive des relations connues de Allen (Fig.10).

A chaque application, on fait l’intersection du résultat avec les relations possibles déjà connues. On contribue ainsi à

3. La fonction $Inverse()$ d’une disjonction renvoie la disjonction des relations inverses de chaque $r \in R_{ij}$

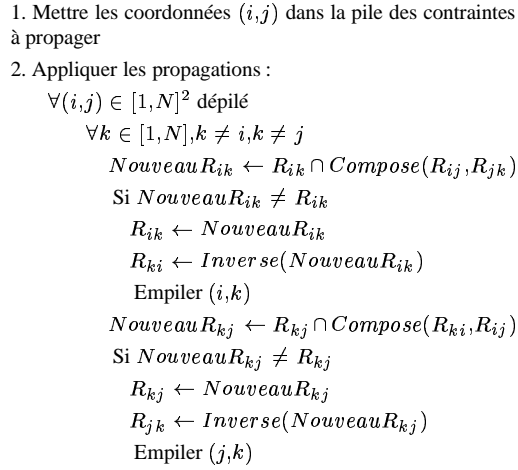


FIG. 10 – Algorithme de propagation des contraintes

réduire l'ensemble des relations possibles entre deux intervalles. Lorsqu'il ne reste qu'une relation élémentaire, on a caractérisé la position entre les deux événements correspondants. En revanche, s'il n'y a plus aucune relation sur l'arc, on peut détecter l'inconsistance de la situation.

5.2 Présentation des incohérences

La propagation des contraintes va permettre de détecter tout conflit dans les informations connues par le système. Cependant, si nous voulons apporter une aide aux archéologues, leur montrer que les résultats obtenus ne sont pas cohérents ne suffit pas. Il faut aussi leur permettre de pouvoir corriger les relations et les événements mis en cause.

Le système vérifie la cohérence du graphe à chaque ajout d'une nouvelle relation. Lorsqu'il détecte une inconsistance, il s'arrête de propager les contraintes. Une solution simple serait de lui faire supprimer la relation qui induit cette incohérence dans le graphe puis de relancer la vérification. Cependant, plusieurs raisons nous empêchent de le faire dans le cadre de nos outils.

Tout d'abord, on ne peut pas conclure que la relation fautive est la dernière entrée, car l'erreur résulte peut-être d'une relation appartenant aux données déjà connues du système. Dans ce cas, on peut faire face à une remise en doute du modèle général, qui contient les hypothèses propres à la communauté.

De plus, même lorsque l'incohérence vient du modèle local, et plus particulièrement de la dernière entrée, on ne peut pas automatiser la réaction provoquée par cette découverte.

Mise en doute du modèle général.

Sur la figure 11, on peut constater que le modèle général (hypothèses partagées par la communauté) proposé tient comme acquis que l'Age du Bronze est immédiatement antécédent à celui du Fer. Dans Porphyre, le chercheur a identifié la couche 1 comme appartenant à l'Age du Fer, et la

couche 3 comme appartenant à l'Age du Bronze.

Le modèle local (relatif à une fouille) pose, grâce à l'étude de la coupe stratigraphique, une chronologie pour les couches identifiées.

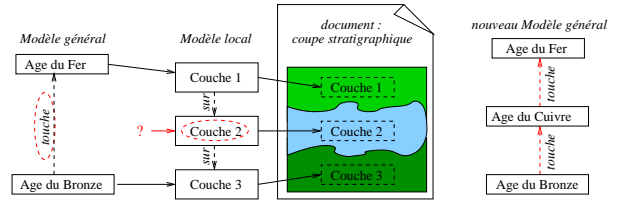


FIG. 11 – Réfutation de l'hypothèse selon laquelle l'Age du Fer suit exactement l'Age du Bronze.

Comme l'expérience prouve que la couche 2 n'appartient ni à l'Age du Fer ni à l'Age du Bronze, le système va signaler la présence d'une incohérence au chercheur. Celui-ci va ensuite identifier le problème, et proposer une correction du modèle. Sur l'exemple, le chercheur a émis une hypothèse réfutant le modèle général, à savoir l'existence d'un age entre l'Age du Fer et l'Age du Bronze : l'Age du Cuivre inconnu jusqu'alors de la communauté. Il aurait pu aussi mettre en doute son modèle local et invalider la coupe stratigraphique.

Mise en doute du modèle local.

Sur ce second exemple (Fig.12), le modèle local, s'appuyant sur la coupe stratigraphique, propose une chronologie entre les artefacts trouvés dans les différentes couches. Cependant, ces types d'objets sont aussi identifiés dans le modèle local et rattachés à différentes époques.

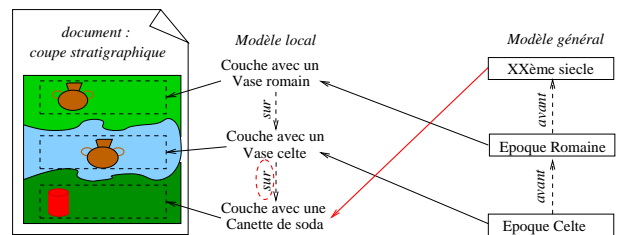


FIG. 12 – Mise en doute de la portée d'une fouille dans laquelle une canette de soda a été trouvée dans une couche inférieure.

Lorsque l'on confronte les modèles, on se retrouve face à une incohérence entre les deux chronologies obtenues. Cette incohérence met ici en défaut le modèle local et montre au chercheur qu'il y a eu remaniement du terrain et que cette coupe stratigraphique n'a, par conséquent, aucune valeur chronologique.

Exploitation des traces. Afin d'aider l'utilisateur à comprendre quelles relations sont mises en cause lors de la détection d'une incohérence, on va garder la trace du chemin parcouru par les contraintes lors de la propagation.

Lors de la détection d'une incohérence, on propose à l'utilisateur une liste des relations qui ont conduit au résultat remis en question, tout en laissant à sa charge l'action à effectuer suite à ce constat.

C'est le chercheur qui décidera alors s'il doit invalider le modèle général (en proposant sa théorie à la communauté), corriger son modèle local, ou rejeter les documents facteurs de l'incohérence.

6 Conclusion

En partant du constat que le temps était une dimension documentaire non exploitée dans les bibliothèques numériques archéologiques, nous avons proposé ici une approche innovante de la modélisation des données temporelles.

Pour cela, nous avons défini un modèle de représentation du temps adapté au domaine, et montré comment rendre son intégration possible au sein d'un système préexistant de restructuration du corpus (Porphyre). Nos outils permettent aux archéologues d'effectuer le traitement de leurs données sans s'écarter de leurs pratiques courantes et de profiter de l'aide de l'informatique pour vérifier la cohérence d'un ensemble de données.

Les travaux en cours portent essentiellement sur l'amélioration du prototype et sur la validation de nos résultats par des tests portant sur un plus grand nombre de données. Pour compléter notre approche, il nous reste aussi à définir les modalités du traitement des *terminus post quem*.

A plus long terme, nos perspectives de développement portent sur les techniques de visualisation des résultats. En effet, nous sommes persuadés qu'avec le mode de représentation adéquat, le chercheur pourrait voir apparaître certaines nouvelles hypothèses calculées par le système, comme des mises en corrélations qui seraient restées invisibles lors d'un traitement courant.

Références

- [1] Accary T., Bénéel A., Calabretto S., Modélisation de connaissances temporelles en Archéologie, *Actes des Journées francophones d'Extraction et de Gestion des Connaissances [EGC'2003]*, *Revue des Sciences et Technologies de l'Information [RSTI]*, Volume 17, Numéro spécial. Paris : Lavoisier - Hermès Sciences, 2003. p.503-508.
- [2] Allen James F., Time and Time Again: The Many Ways to Represent Time, *International Journal of Intelligent Systems*, p.341-355, volume 6(4), July 1991.
- [3] Bachimont B., L'intelligence artificielle comme écriture dynamique: de la raison graphique à la raison computationnelle, *Au nom du sens*, J. Petitot, Paris, Grasset, 1999.
- [4] Bénéel A., Egyed-Zsigmond E., Prié Y., Calabretto S., Mille A., Truth in the Digital Library: From Ontological to Hermeneutical Systems, *Proceedings of the fifth European Conference on Research and Advanced Technology for Digital Libraries [ECDL'2001]*, Darmstadt, September 4-9, 2001. *Lecture Notes in Computer Science 2163*. Berlin : Springer-Verlag. p.366-377.
- [5] Bénéel A., Calabretto S., Iacovella A., Pinon J.-M., Porphyry 2001: Semantics for scholarly publications retrieval, *Proceedings of the thirteenth International Symposium on Methodologies for Intelligent Systems [ISMIS'2002]*, Lyon, June 26-29, 2002. *Lecture Notes in Artificial Intelligence 2366*. Berlin : Springer-Verlag. p.351-361.
- [6] Demoule J.-P., Projet de bibliographie automatique en préhistoire et protohistoire européenne, *Banques de données archéologiques*, Marseille, 12-14 juin 1972.
- [7] Desfarges P., Helly B. L'archéologie, système d'information scientifique. *Aplicaciones Informaticas en Arqueologia: Teorias y sistemas*. Saint-Germain-en-Laye, 1991.
- [8] Euzenat Jérôme, Bessière Christian *et al.*, Dossier Raisonnement Temporel et spatial, *Bulletin de l'Association Française pour l'Intelligence Artificielle*, volume 29, p.2-13, Avril 1997.
- [9] Gally A., *L'archéologie demain*, Belfond, 1986.
- [10] Haton Jean-Paul, Bouzid Nadjet, Charpillat François *et al.*, *Le Raisonnement en Intelligence Artificielle: Modèles, techniques et architectures pour les systèmes à bases de connaissances*, p.183-229, Inter Edition, 1991.
- [11] Jockey Ph., *L'archéologie*, Paris : Belin, 1999.
- [12] Ubeda Thierry, *Contrôle de la qualité spatiale des bases de données géographiques: Cohérence topologique et corrections d'erreurs*, Thèse de doctorat, 1997, INSA de Lyon.
- [13] Vilain M., Kautz H., Constraint Propagation Algorithms for Temporal Reasoning, *Proceedings of the National Conference of the American Association for Artificial Intelligence (AAAI-86)*, 1986.