
La temporalité des corpus archéologiques

Tiphaine Accary-Barbier – Sylvie Calabretto

*Laboratoire d'InfoRmatique en Image et Systèmes d'information
UMR 5205 CNRS / INSA de Lyon / Université Lyon 1 /
Université Lyon 2 / École Centrale de Lyon
bât. Blaise Pascal, 20 avenue Albert Einstein, F-69621 Villeurbanne cedex
tiphaine.accary@liris.cnrs.fr*

RÉSUMÉ. La numérisation de corpus documentaires permet de comparer une grande variété de documents (textes, photos, vidéos, sons, dessins...). Les caractéristiques d'une collection documentaire numérique ne sont donc plus fondées sur les caractéristiques formelles des documents qui la composent. Il est alors nécessaire de dégager de nouvelles dimensions communes de ré-organisation propres à chaque type de collection. Au travers de ces dimensions de description, le document numérique pourra évoluer au gré des annotations pratiquées par ses lecteurs et offrir une meilleure lecture de son sens grâce aux liens de composition, de hiérarchie et de partage d'information qui vont s'établir avec son contexte. Dans le cadre des bibliothèques numériques pour l'archéologie, le concept du temps archéologique, par son caractère complexe et méthodique, constitue une de ces dimensions apte à devenir le support d'une « écriture dynamique » entre les documents. Les descripteurs temporels, ne sont alors plus de simples outils d'indexation du corpus, mais peuvent aussi, dans le cadre d'outils spécialisés, permettre la mise en évidence de nouvelles connaissances et enrichir le corpus.

ABSTRACT. Digitizing of document corpora makes it possible today to compare a large variety of documents (texts, photographs, videos, sounds, drawings...). The characteristics of a digital document collection can't thus be founded on the formal description of documents composing it. It is then necessary to find some common dimensions for each type of collection. Through these dimensions of description, the document will be able to evolve with the annotations made by readers and to offer a better reading thanks to the links which will be established with its context. Into the framework of digital libraries for archaeology, the concept of archaeological time, which is very complex and methodical, constitutes one of these dimensions ready to become the support of a "dynamic writing" between documents. Then temporal descriptors are not any more simple tools of indexing but also can, within the framework of specialized tools, allow the description of new knowledge and enrich the corpus.

MOTS-CLÉS : bibliothèques numériques, modélisation du temps, outils d'assistance, documents archéologiques.

KEYWORDS: digital libraries, time modelling, assisting tools, archaeological documents.

1. Introduction

Face au problème de l'hétérogénéité des collections documentaires numériques, il est nécessaire de mettre en évidence des dimensions communes de réorganisation des corpus, indépendantes des caractéristiques formelles du support. Dans notre domaine d'application, l'archéologie, le temps peut être considéré comme la dimension principale d'organisation documentaire. Nous nous sommes donc intéressés à l'utilisation des dimensions temporelles du document en tant que support « d'écriture dynamique » entre les documents [BAC 99].

1.1. Travaux relatifs à l'indexation temporelle de documents archéologiques

La temporalisation manuelle d'un corpus consiste à associer à des documents (ou à des fragments de ceux-ci) des annotations ayant un contenu temporel, éventuellement issues d'organisation de connaissances (thésaurus ou ontologies) [CAL 04]. L'ensemble de ces annotations forme alors une structure « temporelle » qui va enrichir les connaissances liées au document et permettre également une exploitation « temporalisée » du corpus. De nombreux systèmes de gestion documentaire intègrent la notion de date afin de temporaliser les documents mais, dans la plupart des cas, il s'agit d'informations signalétiques ou d'informations de gestion documentaire constituant une catégorie du profil du document. La temporalité que nous cherchons à exploiter dans le domaine archéologique concerne le temps du contenu documentaire. Les interprétations et les connaissances annexes d'ordre temporel que le lecteur va associer à un document lors de son indexation afin de lui rajouter du sens constituent une classe particulière d'annotations sur le corpus documentaire. Ces annotations, pour être temporellement exploitables, doivent nécessairement disposer d'un système permettant de les ordonner. Certains langages de bibliothèques numériques dédiés aux disciplines historiques tels que RAMEAU, PACTOLS ou PORPHYRY, permettent par des moyens divers d'exprimer ce type d'annotations ordonnées, tout en restant très limités au niveau de l'exploitation de cet ordre.

1.1.1. Les langages précoordonnés : RAMEAU et PACTOLS

RAMEAU (Répertoire d'Autorité-Matière Encyclopédique, Alphabétique Unifié), créé en 1980 par la Bibliothèque nationale de France [BNF99] [BNF00] à partir du fichier d'autorité des vedettes-matières de l'université Laval de Québec, est le langage d'indexation utilisé aujourd'hui par la Bibliothèque Nationale de France et par de nombreuses autres bibliothèques de lecture publique ou de recherche. Il est composé d'un vocabulaire de termes reliés entre eux et d'une syntaxe (règles de construction pour l'indexation). La subdivision chronologique permet d'exprimer la période envisagée dans le document à indexer. La vedette-matière construite ne comporte qu'une subdivision chronologique. Ainsi, pour un document portant sur deux siècles, il faudra établir deux vedettes-matières. Par ce biais, RAMEAU permet d'exprimer deux types de subdivisions chronologiques :

- spécifique : propre à un lieu ou à un sujet qui sont entrés dans la liste d'autorité sous le nom du lieu ou du sujet ;
- affranchie ou « commune » : employée lorsqu'il n'existe pas de découpage chronologique spécifique.

A la différence d'un thésaurus, la liste d'autorité encyclopédique n'est pas constituée *a priori* mais au fur et à mesure des besoins d'indexation et évolue sur la base des propositions faites par le réseau de ses utilisateurs.

Le thésaurus PACTOLS (Peuples, Anthroponymes, Chronologie, Toponymes, Œuvres, Lieux et Sujets) [PAC89] du langage FRANTIQ, permet également une prise en compte du temps du contenu documentaire. Ce langage comprend des listes de termes sélectionnés et organisés selon des relations sémantiques (synonymie, antynomie, adjacence) ou hiérarchiques (termes génériques et termes spécifiques). L'indexation d'un document se fait par un choix dans la liste des termes des thésaurus PACTOLS, mais aussi par des mots-clés libres et des mots-clés macroscopiques constitués de trois chiffres qui correspondent à un découpage de la réalité selon quatre axes : espace, temps, disciplines et objet d'études. Des relations d'équivalence permettent d'utiliser le descripteur retenu et contrôle les problèmes de synonymie. Par exemple, pour « âge de la pierre », le thésaurus précise qu'il faut utiliser « paléolithique ».

1.1.2. Les réseaux de description de PORPHYRY

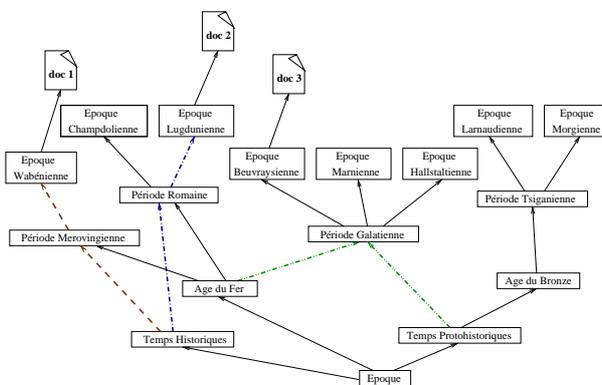


Figure 1. Réseau de descripteurs temporels

Dans l'outil PORPHYRY ¹, la description des documents s'effectue non plus à l'aide de descripteurs isolés comme dans les systèmes usuels d'indexation, mais grâce à des descripteurs organisés dans un ordre partiel conformément à des relations explicites de spécialisation [BEN 02]. L'utilisateur peut définir ses propres descripteurs et organiser ceux-ci comme il le souhaite.

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

Dans le cadre des descripteurs temporels, les relations de spécialisation permettent d'exprimer l'inclusion temporelle. Ainsi, sur la figure 1, la Période Romaine apparaît comme une partie des Temps Historiques. De plus, en extrapolant, on peut voir aisément que certains schémas de relations permettent l'expression du chevauchement entre deux périodes. Ainsi, en constatant que la Période Galatienne appartient conjointement à l'Âge du Fer et aux Temps Protohistoriques, on peut en déduire que ces deux périodes possèdent une partie commune.

1.2. Contributions

Cependant, si de tels outils permettent d'enrichir les documents par des informations temporelles, la question du traitement de ces informations reste posée. En effet, grâce à ces systèmes on peut retrouver un document en formulant une requête portant sur cette nouvelle catégorie d'information, mais on ne peut pas pour autant travailler sur la connaissance contenue au niveau de ces informations. Le principe de l'inclusion entre descripteurs dans l'outil PORPHYRY rend possible une première exploitation de la temporalité documentaire, mais celle-ci demeure trop partielle pour permettre l'émergence d'une connaissance temporelle extra-documentaire. L'information connue ne permet pas à ce niveau d'établir une « chronologie » des descripteurs. Ainsi, lorsqu'on va comparer sur la figure 1 l'Epoque Wabénienne et l'Epoque Lugdunienne, le seul élément de comparaison dont on dispose est leur appartenance commune aux Temps Historiques. Il serait intéressant de connaître également leurs positionnements temporels relatifs.

L'idée nouvelle que nous développons ici est celle d'outils d'assistance centrés sur l'exploitation du caractère temporel des annotations. Ainsi, parallèlement aux opérations d'annotations documentaires réalisées par les chercheurs, ces assistants vont par exemple permettre la détection des incohérences temporelles que le corpus annoté pourrait contenir et ainsi valider ou invalider les hypothèses temporelles de l'expert.

Pour appréhender tous les éléments nécessaires à l'exploitation de ces annotations, il convient de s'intéresser à l'importance du temps archéologique et à son mécanisme de construction. Cette approche permet entre autres de comprendre que le modèle archéologique est basé sur une multitude d'ordres temporels relatifs connectés les uns aux autres de manière très souple (section 2). Il a fallu tenir compte de ces caractéristiques pour adopter un modèle informatique de représentation du temps le plus en adéquation avec le modèle archéologique. Après avoir comparé les grandes familles de représentation utilisées dans d'autres domaines, il est apparu que les modèles utilisant la propagation des contraintes réunissaient le double avantage de permettre une représentation temporelle assez fidèle tout en offrant des possibilités de développement d'outils permettant l'émergence de nouvelles connexions (section 3.1). Afin de rendre notre modèle applicable, il convient d'imaginer comment il pourrait être intégré dans la structure d'indexation préexistante du système PORPHYRY à moindre coût. Pour cela, nous proposons dans un premier temps de munir le réseau de descripteurs de relations transversales permettant de signifier les rapports de connexité

et de simultanéité (section 3.2). Pour pouvoir étendre ce modèle aux applications archéologiques, nous utilisons ensuite des langages-interfaces entre l'archéologue et le système qui permettent de réutiliser les méthodes de description propres aux archéologues. On a ainsi pu identifier un ensemble de règles permettant de faire la jonction entre les informations stratigraphiques et les réseaux de descripteurs (section 3.3).

L'outil que nous présentons dans cet article n'a d'intérêt qu'en complément d'un système d'annotation de documents. Grâce au modèle temporel, parallèlement à la phase d'annotation par l'utilisateur, le système effectuera les vérifications de la cohérence temporelle des informations connues. Cette détection est particulièrement intéressante dans le domaine archéologique, puisque grâce à elle la connaissance va être remise en question et des nouvelles hypothèses vont pouvoir l'étayer (section 4).

2. Spécificités du temps archéologique

Pour dater un objet manufacturé, la méthode la plus utilisée est d'identifier son style (indices stylistiques). On pourra ainsi supposer que deux objets analogues dans leur forme ou leur technique sont contemporains. Cependant ces indices sont très dépendants des caractéristiques spatiales du lieu auquel se rattachent les objets. En effet, si on adhère à la théorie du diffusionnisme, un style peut apparaître avec un certain retard dans une contrée voisine. D'autre part, on peut observer des analogies de style sur des objets d'époques et de lieux différents dues à des impulsions de mode. Il faut donc prendre en compte d'autres informations externes pour comparer les objets.

L'indice spatial le plus courant est la profondeur relative des vestiges. La méthode stratigraphique s'appuie sur le principe qui veut que dans une série de couches, les unités supérieures de stratification soient les plus jeunes [JOC 99]. Cependant cette règle de séquence temporelle n'est pas infaillible et admet des contre-exemples dûs aux remaniements naturels ou humains du terrain. Pour obtenir une chronologie, il faut alors croiser les indices stylistiques et stratigraphiques recueillis tout en restant prudent. Ainsi on ne dira pas qu'un vestige date une couche mais seulement qu'elle détermine son *terminus post quem*². Lorsque les archéologues du futur découvriront le musée du Louvre, ils ne devront pas déduire 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. La position géographique d'un vestige pourra également aider à établir sa datation. Les adeptes du diffusionnisme considèrent qu'un style va apparaître en un foyer et se diffuser petit à petit dans les régions voisines, ainsi la distance au foyer pourrait être révélatrice du temps nécessaire à la diffusion.

Enfin, grâce aux physiciens et aux chimistes on dispose d'indices naturels liés aux théories quantitatives sur l'évolution dans le temps de certaines propriétés des vestiges. Cependant on a pu constater 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

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

selon laquelle les rayonnements cosmiques avaient été constants dans le temps et dû rééchelonner les datations au radiocarbone en fonction de la dendrochronologie (basée sur la succession de cernes d'arbres plus ou moins épaisses suivant les époques de sécheresse). Cette technique pose toutefois un problème de comparaison car, pour faire correspondre une suite de cernes avec une autre, il faut supposer que deux zones géographiquement distinctes ont eu au même moment un enchaînement climatique identique. 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 jamais vraiment.

En archéologie, une chronologie n'est jamais une donnée mais peut être vue comme un résultat construit à partir du croisement de ces différents indices. De plus, sa validité est perpétuellement remise en question par l'émergence de nouvelles informations puisque le chercheur ne connaît qu'une partie des informations temporelles sur les événements qu'il manipule. L'archéologie est donc un domaine où la plupart des datations manipulées sont relatives, où les événements sont des périodes et où la connaissance des informations est disparate. On mesure 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 les systèmes d'information proposés pour l'archéologie comme celui de J. P. Demoule [DEM 72] s'appuient sur le 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.

3. Proposition d'assistants

Si le chercheur en archéologie manipule le temps à la fois sur le terrain (expérimentations) et dans ses pratiques d'érudition (hypothèses), cette manipulation est toujours liée à un travail documentaire d'écriture et de lecture. En fonction du type d'activité menée par l'archéologue, les connaissances temporelles issues de ces manipulations sont appelées à prendre diverses formes. Ainsi, dans le cas d'une analyse d'un site de fouilles, elles seront probablement stockées sous forme de diagramme stratigraphique. Par contre, si le chercheur effectue plutôt des travaux de synthèse concernant une région ou une période donnée, elles prendront plus certainement la forme de frise chronologique. Il nous faut tenir compte des particularités de ces pratiques pour la conception de nos outils.

Un des principes généraux de l'archéologie établit que tant qu'une théorie n'a pas été réfutée par une expérience elle est hypothétiquement considérée comme vraie. Il est donc nécessaire que nos assistants permettent aux chercheurs de faire le point sur les incohérences que le corpus pourrait contenir entre strates (expériences) et frises (théories). Pour permettre ce type de comparaison, on aura alors besoin d'établir un modèle pivot pour passer d'un modèle de connaissances à l'autre. Cette détection présente un réel intérêt dans un domaine tel que l'archéologie, puisque c'est grâce au débat engendré par ces désaccords que de nouvelles théories peuvent émerger.

3.1. Choix de modélisation

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 [EUZ 97]. Parmi les modèles temporels utilisés dans ces domaines, les approches fondées sur la propagation des contraintes semblent être celles qui répondent le mieux aux spécificités du domaine archéologique. Comme nous venons de le montrer, le temps archéologique est relatif. Nous avons donc exclu automatiquement de notre choix les modèles de représentation fondés sur un système de datation absolue tels que les représentations élémentaires utilisant l'ordonnement de dates [ALL 91] et les modèles de type diagrammes de PERT. Par ailleurs la nécessité de manipuler des événements correspondant à des périodes de temps discrédite d'emblée les modèles dont les primitives sont des instants (événements ponctuels). Les modèles de Dean et McDermott [HAT 91] 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 comme celui défini par Vilain et Kautz [VIL 86] pourraient convenir à la relativité des représentations archéologiques, mais ils ne permettent d'exprimer qu'un faible nombre de relations (<, >, =) et ne sont donc pas assez précis pour avoir un intérêt dans la construction de connaissances. Le modèle s'appuyant sur les treize relations exhaustives et exclusives possibles entre deux intervalles (figure 2) décrit par Allen [ALL 91] constitue en cela une base solide pour un modèle temporel valide pour l'archéologie.

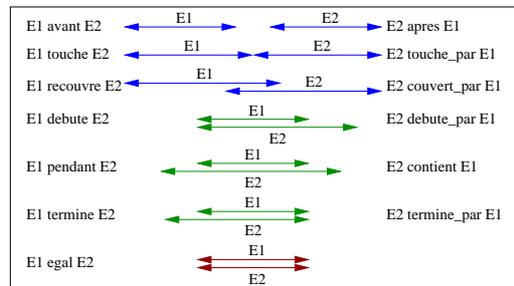


Figure 2. Relations de connexité, d'inclusion et de simultanéité entre événements

Ce modèle se constitue sous la forme d'un graphe dans lequel chaque nœud stocke un événement (défini par intervalle de durée) et chaque arc est étiqueté par la relation entre les nœuds qu'il relie. Sa particularité principale est qu'il est également possible d'étiqueter un arc avec une disjonction de relations. Ainsi, lorsqu'on n'a aucune information sur l'ordre temporel entre deux événements, on peut l'étiqueter avec la disjonction (OU) totale de toutes les relations existantes. De même, si on a une incertitude sur la relation entre deux événements (par exemple si on considère qu'ils sont de la même période mais qu'on ignore ce qu'il en est de leurs bornes), on va caractériser

tériser l'arc avec la disjonction de ces deux relations. Grâce à ce mécanisme, on va pouvoir exprimer les incertitudes et le manque de données propres aux connaissances du domaine archéologique.

La seconde raison qui nous incite à choisir ce type d'approche tient dans le principe de la propagation des contraintes. Par ce mécanisme, on va pouvoir maintenir le graphe temporel complet à tout instant et détecter les incohérences induites par l'ajout de nouvelles informations. Cette faculté du modèle est tout à fait adaptée au travail du chercheur en archéologie puisque une tâche essentielle est la réfutation d'hypothèses.

Pour pouvoir étendre ce modèle aux applications archéologiques, il va falloir définir des langages dédiés qui serviront d'interface entre l'archéologue et le système. Ces langages vont permettre d'insérer les informations sous une forme plus intuitive que celle du modèle de calcul et le système, à l'aide d'un certain nombre de règles, se chargera ensuite de leur conversion. Grâce à ce mécanisme de traduction, on va pouvoir confronter les modèles locaux (issus de la description de fouilles) exprimés sous forme de diagramme stratigraphique en fonction de relations propres et le modèle général (savoir de référence du domaine partagé par la communauté).

3.2. Traduction du modèle général

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

Tableau 1. Règles de traduction du réseau de descripteurs en graphe de Allen

Le modèle général des connaissances est constitué par l'ensemble des annotations à caractère temporel qui ont été rattachées aux documents du corpus. Dans le système PORPHYRY, il s'agit du réseau de description des documents. Actuellement notre prototype fonctionne indépendamment du système PORPHYRY qui est, quant à lui, déjà expertisé par plusieurs équipes de chercheurs en archéologie. Cependant, pour avoir accès à un plus grand nombre de données pour nos tests et pouvoir valider le modèle auprès des utilisateurs, son intégration comme module de PORPHYRY est prévue à court terme. Nous avons donc déjà étudié les conditions de cette intégration. Pour cela, il conviendrait d'enrichir le réseau de description tel qu'il existe actuellement par l'ajout de relations transversales permettant d'exprimer la connexité entre événements. Nous nous basons donc sur une version encore hypothétique de PORPHYRY dans laquelle le modèle général s'exprime sous la forme d'un graphe de descripteurs munis de relations d'inclusion, de chevauchement et de connexité. Pour pouvoir appliquer la propagation des contraintes, il nous faut donc traduire ce graphe

en un autre graphe muni des relations de Allen. Afin d'automatiser cette tâche, on a donc défini un ensemble exhaustif de règles de traduction (tableau 1).

On va ensuite compléter le graphe obtenu 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 (figure 3). La table de composition entre relations de Allen va permettre le calcul des relations inconnues. Ainsi, sur l'exemple, la relation entre la période romaine et l'époque wabénienne sera la composée de la relation entre période romaine et époque champdoliennne avec la relation entre époque champdoliennne et époque wabénienne.

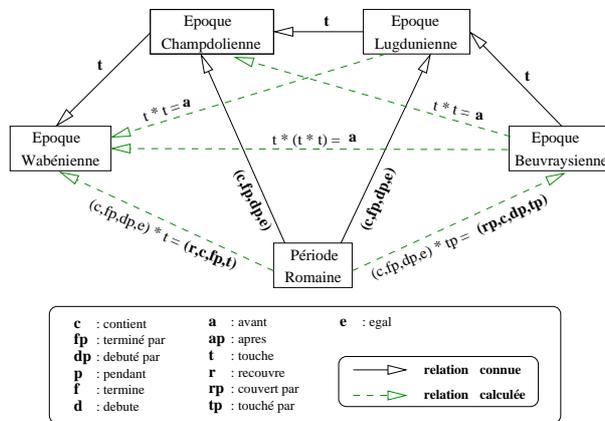


Figure 3. Graphe complet de Allen

3.3. Traduction du modèle local

Les documents tels que relevés de fouilles, coupes, rapports constituent le modèle local de l'archéologue. Il s'agit là de documents bruts, demandant une interprétation. Lors de son travail d'analyse sur une coupe stratigraphique, le chercheur va identifier les fragments stratigraphiques significatifs et établir le diagramme stratigraphique des relations entre ces fragments en fonction de leur nature. Par exemple, il établira que les murs d'un puits sont plus anciens que les objets contenus à l'intérieur. Il va ensuite consigner ces résultats sous la forme d'un diagramme permettant d'ordonner l'ensemble des fragments avec trois relations : **sur**, **sous** et **synchro**. Sur la figure 4, on peut voir qu'on a consigné dans le diagramme que le fragment 8 (vase contenu dans une fosse) était plus récent que le fragment 4 (parois de la fosse). Au niveau documentaire, cette phase d'analyse consiste à identifier les fragments sur le document et à les extraire en tant que nouveau document pouvant être annoté.

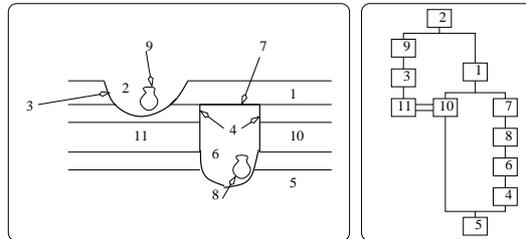


Figure 4. Identification de fragments et diagramme stratigraphique correspondant

Le système PORPHYRY permet déjà ce type d'extraction documentaire. L'ordonnement temporel entre ces fragments est hérité de la méthode stratigraphique. Il est donc particulièrement intéressant de permettre à l'utilisateur d'exprimer cet ordre stratigraphique entre les fragments. Lors de l'indexation du corpus, on va rattacher ces fragments aux descripteurs temporels qui les caractérisent (style servant à dater un objet, une couche). On va alors reporter le schéma des relations entre les fragments sur le graphe des descripteurs. Dans la mesure où ce graphe n'est pas exprimé dans le même langage, il va falloir au cours de l'opération traduire les relations stratigraphiques entre descripteurs en relations de Allen. Pour permettre cette transformation, nous avons défini une table de correspondance (tableau 2).

Stratigraphie	Conversion en Allen
<i>A sous B</i>	<i>avant(A, B) ou touche(A, B)</i>
<i>A synchro B</i>	<i>recouvre(A, B) ou recouvre(B, A) ou egal(A, B)</i>
<i>A sur B</i>	<i>apres(A, B) ou touche(A, B)</i>

Tableau 2. Relations stratigraphiques traduites en Allen

4. Confrontation et détection des incohérences

Grâce à la propagation des contraintes, chaque modèle va donc constituer un graphe complet dans lequel tout conflit entre les informations connues par le système peut être détecté. La traduction des modèles en langage de Allen va permettre de vérifier leur cohérence et de faire le point sur les informations qui la mettent en doute.

De plus, comme les deux modèles s'articulent autour des relations de Allen en tant que langage commun, on va pouvoir les confronter (figure 5). Cette confrontation va alors permettre de vérifier la cohésion entre le savoir de référence et la connaissance issue des expériences.

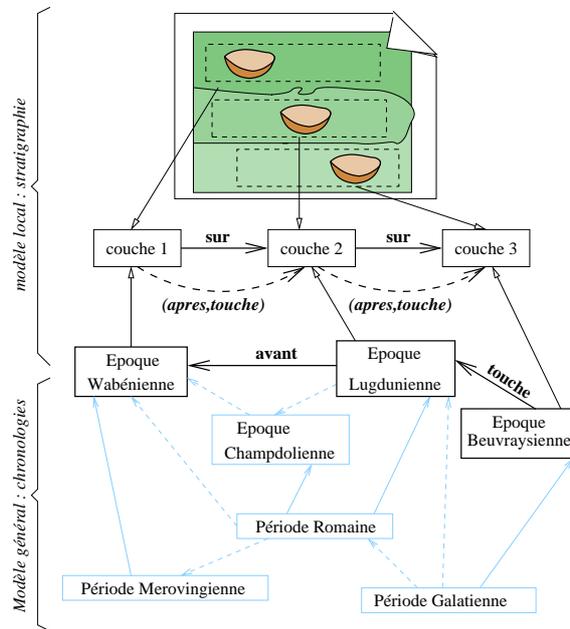


Figure 5. Comparaison du modèle local au modèle général

Il est primordial de noter qu'en archéologie, plus que la correction de ces conflits, c'est leur détection qui est importante. En faisant le point sur les incohérences qu'il peut y avoir entre les modèles, on permet au chercheur de faire évoluer la connaissance. Lorsque le système détecte une incohérence, il s'arrête de propager les contraintes et signale le conflit : c'est alors que le travail du chercheur commence. En fonction des cas de figure, il devra corriger ses modèles. S'il juge que le conflit vient d'une hypothèse du modèle général qui se trouve réfutée par ses résultats d'expérience, il peut choisir de mettre en doute le modèle général en proposant une nouvelle théorie. Il peut également vouloir remettre en doute son modèle local et même éventuellement choisir de le rejeter.

En dépit de l'aide que l'outil informatique pourrait apporter, nous avons choisi délibérément de ne pas automatiser la correction des incohérences. Dans les disciplines historiques, l'établissement de la connaissance doit obligatoirement passer par l'analyse humaine des situations. L'outil a pour rôle de préparer le travail d'analyse en détectant les informations sujettes à caution mais les conclusions résultant de cette détection sont dépendantes du mode de pensée de chaque chercheur. On attendra donc de l'outil qu'il présente à l'expert une liste des relations liées au conflit puisqu'il le laisse choisir quelle démarche adopter pour le résoudre.

Globalement, les détections de conflit les plus porteuses en matière de construction du savoir sont celles qui ouvrent le débat entre savoir relatif et résultats expérimentaux.

4.1. Exemple de mise en doute du modèle général

La figure 6 présente un exemple de type cas d'école permettant d'illustrer le mécanisme de remise en doute du modèle cognitif par la détection d'incohérence. On peut constater que le modèle général (hypothèses partagées par la communauté) proposé tient comme acquis que l'Âge de Bronze est dans le prolongement immédiat du Néolithique. On a identifié la couche 1 comme appartenant à l'Âge de Bronze, et la couche 3 comme appartenant à l'époque Néolithique. Le modèle local (relatif à une fouille) pose, grâce à l'étude de la coupe stratigraphique, une chronologie pour les couches identifiées.

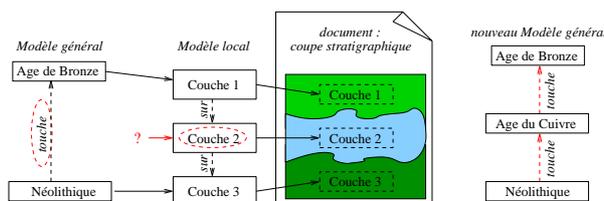


Figure 6. Réfutation de l'hypothèse selon laquelle l'Âge de Bronze suit exactement le Néolithique

Comme l'expérience prouve que la couche 2 n'appartient ni à l'époque néolithique ni à l'Âge de 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 âge entre le Néolithique et l'Âge de Bronze : l'Âge du Cuivre (Chalcolithique) inconnu jusqu'alors par la communauté. Il aurait pu aussi mettre en doute son modèle local et invalider la coupe stratigraphique.

4.2. Exemple de mise en doute du modèle local

Sur ce second exemple (figure 7), 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 général et rattachés à différentes époques.

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.

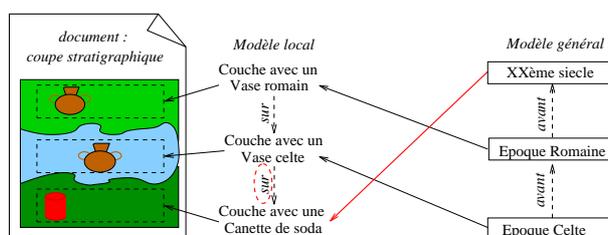


Figure 7. Mise en doute de la portée d'une fouille dans laquelle une canette de soda a été trouvée dans une couche inférieure

5. Bilan et perspectives

En partant du constat que la dimension temporelle des documents dans les bibliothèques numériques archéologiques n'était que très peu exploitée en tant que connaissance à part entière, nous avons proposé un exemple d'outil permettant un traitement constructeur de sens de ces informations temporelles. Grâce à un modèle de représentation du temps adapté aux pratiques archéologiques, nous proposons un modèle d'assistant permettant, par le mécanisme de vérification de la cohérence de l'ensemble des connaissances, de mettre en évidence des points sensibles sujets à discussion et donc d'aider au développement de la connaissance.

Cette modélisation souffre encore d'un défaut de validation en situation réelle. Pour le moment, nous n'avons effectué que des tests en interne, sur notre propre base de documents. Pour permettre une validation à plus grande échelle, l'intégration de nos outils dans le système PORPHYRY est prévue à court terme. Cette intégration permettra également de vérifier la fiabilité de nos principes de traduction entre les réseaux de descripteurs améliorés et les graphes à propagation de contraintes. Cependant les travaux actuellement menés semblent répondre parfaitement aux attentes de la communauté archéologique. Le travail collaboratif mené avec les experts utilisateurs de PORPHYRY au sein du réseau thématique de recherche ARTCADHi montre que ce type d'application est tout à fait en adéquation avec les attentes exprimées par les chercheurs. En effet les outils d'assistance proposés n'entrent pas en conflit avec les pratiques du domaine.

Comme nous l'avons montré dans cet article, ce qui importe le plus est de mettre en évidence ces points afin de favoriser le débat. Dans cette optique, nous travaillons actuellement sur la confrontation de points de vue entre des communautés d'experts [ACC 04]. Il serait en effet très intéressant de permettre aux communautés de construire

leurs propres modèles tout en scrutant les désaccords qu'ils présentent avec les modèles construits par des communautés aux thématiques proches.

Nous nous intéressons également aux différents mode de visualisation possibles pour nos résultats. Nous sommes en effet persuadés que, grâce à un mode de représentation adéquat, certaines hypothèses nouvelles pourraient être mises en évidence plus simplement.

6. Bibliographie

- [ACC 04] ACCARY T., BENEL A., CALABRETTO S., IACOVELLA A., « Confrontation de points de vue sur des corpus documentaires : Le cas de la modélisation du temps archéologique », *Actes du 14ème Congrès Francophone AFRIF-AFIA de Reconnaissance des Formes et Intelligence Artificielle [RFIA], Toulouse, 28-30 janvier 2004*, p. 197-205.
- [ALL 91] ALLEN J. F., « Time and Time Again : The Many Ways to Represent Time », *International Journal of Intelligent Systems*, vol. 6, n° 4, 1991, p. 341-355.
- [BAC 99] 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.
- [BEN 02] BENEL 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]. Lecture Notes in Artificial Intelligence 2366, Lyon, Berlin : Springer-Verlag, 26-29 juin 2002*, p. 351-361.
- [BNF99] Bibliothèque Nationale de France. Service de coordination bibliographique, « RAMEAU [Répertoire d'Autorité Matière, Encyclopédique, Alphabétique Unifié] : noms communs, noms géographiques, suivis de subdivisions chronologiques », 1999.
- [BNF00] Bibliothèque Nationale de France, « Guide d'indexation RAMEAU », 2000, ISBN 2-7177-2131-2.
- [CAL 04] CALABRETTO S., LALLICH-BOIDIN G., SÈDES F., « Les temps du document numérique : Rapport de l'action spécifique 95, CNRS/STIC », 2004.
- [DEM 72] DEMOULE J.-P., *Projet de bibliographie automatique en préhistoire et protohistoire européenne*, Banques de données archéologiques, Marseille, 12-14 juin 1972.
- [EUZ 97] EUZENAT J., BESSIÈRE C., JEANSOULIN R., REVAULT J., SCHWER S., « Dossier Raisonnement Temporel et spatial », *Bulletin de l'Association Française pour l'Intelligence Artificielle*, vol. 29, Avril 1997, p. 2-13.
- [HAT 91] HATON J.-P., BOUZID N., CHARPILLET F., HATON M.-C., LÂASRI B., LÂASRI H., MAQUIS P., MONDOT T., NAPOLI A., *Le Raisonnement en Intelligence Artificielle : Modèles, techniques et architectures pour les systèmes à bases de connaissances*, InterEditions, 1991.
- [JOC 99] JOCKEY P., *L'archéologie*, Paris : Belin, 1999.
- [PAC89] « Les thésaurus PACTOLS de FRANTIQU », 1989,
<http://frantiq.adbs.fr/prodserv/jetude/html/PACTOLS/PACTOLS.html>.
- [VIL 86] 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.