

Navigation temporelle dans les bibliothèques numériques archéologiques

Tiphaine Accary

L.I.R.I.S. Laboratoire d'InfoRmatique en Image et Systèmes d'information
CNRS / INSA de Lyon / Université Lyon 1 / Université Lyon 2 / École Centrale de Lyon

bât. Blaise Pascal

20 avenue Albert Einstein

69621 Villeurbanne Cedex

tiphaine.accary@liris.cnrs.fr

Résumé

Le temps est une dimension documentaire encore peu exploitée dans les bibliothèques numériques archéologiques. Il semble pourtant que le concept du temps archéologique, par son caractère complexe et méthodique, soit particulièrement représentatif des domaines dans lesquels les techniques de modélisation de connaissances peuvent devenir le support d'une « écriture dynamique » [3].

Les descripteurs temporels, n'apparaissent alors pas comme 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.

Mots-clés

Bibliothèques numériques, modélisation du temps, outils d'assistance, documents archéologiques.

1 Introduction

Au sein d'un ensemble documentaire statique, le document est un élément fini du corpus. Chaque document peut faire référence à des publications antérieures utiles pour compléter son interprétation, mais il ne pourra pas référencer des documents plus récents. Les liens entre documents, lorsqu'ils existent, ne peuvent pas évoluer et l'utilisateur de la bibliothèque n'enrichit pas le contenu par ses recherches.

La numérisation des documents va permettre l'émergence de nouveaux comportements face à un corpus. Contrairement au document lié à son support, le document numérique peut évoluer au gré des annotations de ses lecteurs et des liens de composition, de hiérarchie et de partage d'information qui vont s'établir avec son contexte. Le document acquiert donc une certaine dynamisme, de son émergence à son exploitation.

Ainsi, dans une bibliothèque numérique, la lecture du document originel ne sera pertinente qu'à la lumière de l'ensemble auquel il est rattaché, ensemble qui n'est jamais statique dans la mesure où ces liens évoluent dans le temps.

La dématérialisation des documents pose aussi le problème de l'hétérogénéité des collections documentaires. Comme le passage au support numérique peut permettre d'assembler des documents d'une grande variété, les caractéristiques d'une collection ne dépendront plus des caractéristiques formelles des documents qui la compose.

On va alors chercher à dégager des dimensions communes qui permettent de réorganiser les collections numériques.

Dans le cadre des bibliothèques numériques pour l'archéologie, nous nous sommes intéressés aux dimensions temporelles du document, et à l'utilisation de descripteurs temporels comme support à la navigation documentaire.

Un des intérêts principaux de l'utilisation de tels descripteurs étant que, au-delà de leur aspect navigationnel, ils sont porteurs d'information et peuvent ainsi servir de base à des outils d'assistance aux experts travaillant sur le corpus. Ces outils permettant, par exemple, de détecter les désaccords entre experts [1].

Nous présentons ici un de ces outils spécifiques aux descripteurs temporels basé sur les modèles développés en « Intelligence artificielle » [5].

2 Spécificités du temps archéologique

Une chronologie n'est pas une donnée mais un résultat. Le temps archéologique est donc construit à partir d'indices stylistiques, spatiaux ou naturels en utilisant des heuristiques plus ou moins fiables.

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 ces indices sont souvent très dépendants des aspects spatiaux car, si 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.

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, les unités supérieures de stratification sont les plus jeunes et les inférieures les plus anciennes [6].

Cependant cette règle de séquence temporelle n'est pas infaillible et admet des contre-exemples dus 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 un vestige ne date pas une couche mais lui fournit seulement un *terminus post quem*¹. 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 selon laquelle les rayonnements cosmiques avaient été constants dans le temps et du ré-échelonner 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). Ceci 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.

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 [4] tenait compte du fait qu'en archéologie il n'existe pas une chronologie unique mais des chronologies valables dans une région géo-

graphique donnée, avec parfois des correspondances possibles entre elles.

3 Attentes du modèle

Pour définir un modèle temporel propre à notre domaine, d'application, il faut préalablement faire le point sur les particularités du temps archéologique ainsi que définir les besoins des chercheurs du domaine.

Il s'agira ensuite de trouver, parmi les modèles de représentation déjà existants, celui qui répond le mieux à ces attentes et chercher à l'étendre pour qu'il couvre leur totalité.

3.1 caractéristiques du temps manipulé

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.

3.2 besoins des chercheurs

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 elles se rattachent. 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. On doit donc tenir compte de ces tâches particulières dans la conception de nos outils.

En outre, en archéologie, tant qu'une théorie n'a pas été réfutée par une expérience, elle est hypothétiquement considérée comme vraie.

Notre modèle 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.

De plus, pour avoir un réel intérêt, cette détection doit s'accompagner d'un mécanisme de correction. Il nous faut donc un modèle pouvant permettre à l'utilisateur de visualiser puis de traiter les éléments facteurs d'incohérence.

4 Choix de modélisation

4.1 Modèle de référence

Au vu des contraintes énoncées, le modèle le mieux adapté nous semble être le modèle de propagation de contraintes de Allen [2] qui s'appuie sur les treize relations exhaustives et exclusives possibles entre deux intervalles (Fig.1).

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

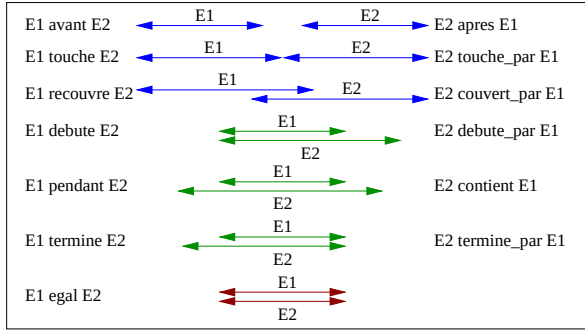


FIG. 1 – Relations de connexité, d’inclusion et de simultanéité entre événements.

Par l’expression de disjonctions, ce modèle va permettre d’exprimer les incertitudes. De plus, grâce à la propagation des contraintes, on pourra maintenir le graphe complet à tout instant et détecter ainsi les incohérences induites. Cette faculté du modèle est tout à fait adaptée au travail archéologique dont une tâche essentielle est la réfutation d’hypothèses.

Il reste ensuite à étendre ce modèle aux applications archéologiques par l’ajout de langages dédiés, externes au noyau, utilisés lors de l’interaction chercheur-machine. Ces langages, qui serviront d’interface entre l’archéologue et le système, vont permettre d’entrer les informations sous une forme plus intuitive que celle du modèle de calcul : le système, à l’aide d’un certain nombre de règles, se chargeant ensuite de les convertir.

On va ensuite pouvoir confronter le modèle local, qui est issu de la description de fouilles et qui permet l’expression de diagrammes stratigraphiques, et le modèle général, qui exprime une théorie du domaine (partagée par la communauté).

4.2 Modèle général

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

Le modèle général étant stocké dans un graphe de descripteurs munis de relations d’inclusion et de connexité, nous allons traduire celles-ci 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 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 (Fig.2).

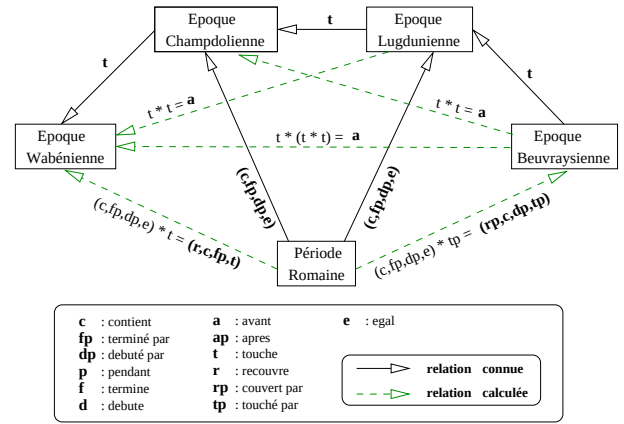


FIG. 2 – Graphe complet de Allen.

La table de composition entre relations de Allen va permettre le calcul des relations inconnues. Ici, la relation entre la période romaine et l’époque wabénienne sera la composée de la relation entre période romaine et époque champdolienne avec la relation entre époque champdolienne et époque wabénienne.

4.3 Modèle local

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.3).

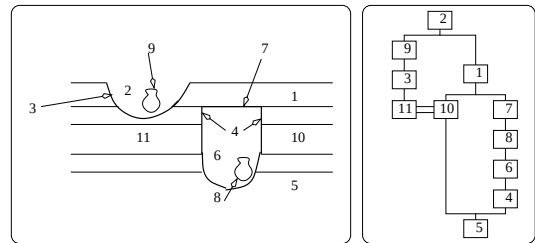


FIG. 3 – Les relations entre les éléments stratigraphiques identifiés sur la coupe sont réduites au nombre de trois sur le diagramme : **sur**, **sous** et **synchro**.

Ainsi les éléments 10 et 11, identifiés au même niveau sur la coupe, seront reliés par une relation “synchro” sur la coupe. De même, le remblais de la fosse 3 (élément 2) est plus récent que le contenu de la fosse (vase 9) : sur le diagramme on synthétisera que l’élément 2 est “sur” le 9 (respectivement, 9 “sous” 2).

On va d’abord identifier les fragments stratigraphiques et les définir par des descripteurs de type “Couche”. Ces descripteurs vont ensuite être rattachés à des styles préexistants.

Lors de cette opération, les relations stratigraphiques entre descripteurs vont être traduites en relations de Allen en utilisant la table de correspondance que nous avons définie (Tab.2).

Stratigraphie	Conversion en Allen
$A \text{ sous } B$	$avant(A,B)$ ou $touche(A,B)$
$A \text{ synchro } B$	$recouvre(A,B)$ ou $recouvre(B,A)$ ou $egal(A,B)$
$A \text{ sur } B$	$apres(A,B)$ 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.4), nos deux modèles seront donc exprimés dans le même langage et leur comparaison deviendra alors possible.

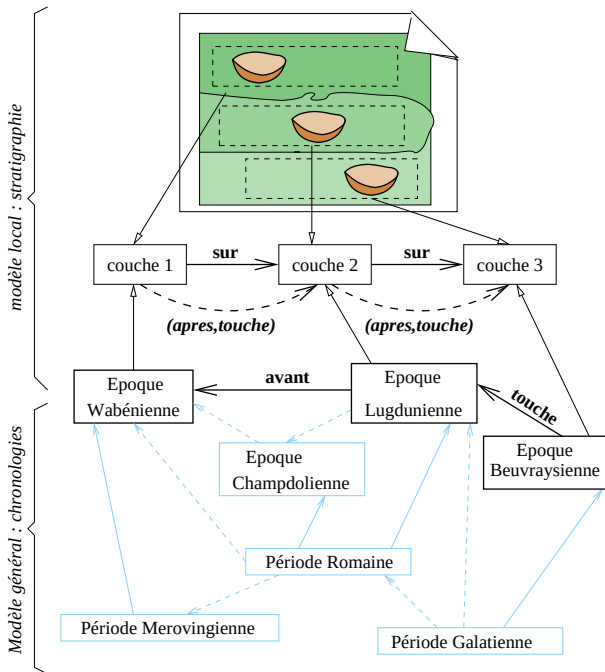


FIG. 4 – Comparaison du modèle local au modèle général.

4.4 Gestion 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, cette détection n'a d'intérêt que si elle peut s'accompagner d'une correction. Lorsque le système détecte une incohérence, il s'arrête de propager les contraintes. Une solution simpliste serait de lui faire supprimer la relation qui a induit l'incohérence puis de relancer la vérification, mais plusieurs raisons nous empêchent de le faire dans le cadre de nos outils.

Tout d'abord, il serait faux de conclure que la relation fautive est la dernière entrée, car l'erreur peut venir d'une re-

lation appartenant aux données déjà connues du système qui serait alors remise en cause. Dans ce cas, on peut avoir à mettre en doute le 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 doit pas automatiser la réaction provoquée par cette découverte.

Pour aider l'utilisateur à identifier les relations mises en cause lors de la détection d'une incohérence, on va donc garder une trace du chemin parcouru par les contraintes lors de la propagation. Lors de la détection d'une incohérence, on proposera à l'utilisateur une liste des relations qui ont conduit au résultat remis en question. 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.

5 Exemples

5.1 Mise en doute du modèle général

Sur la figure 5, 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. On 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.

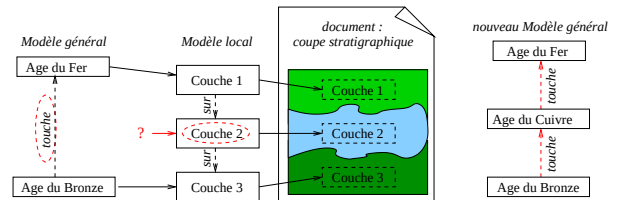


FIG. 5 – Réfutation de l'hypothèse selon laquelle l'Age du Fer suit exactement l'Age du Bronze.

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.

5.2 Mise en doute du modèle local

Sur ce second exemple (Fig.6), 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.

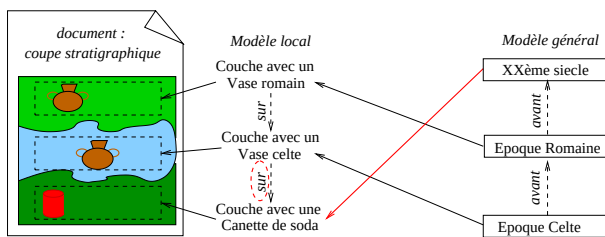


FIG. 6 – Mise en doute de la portée d'une fouille dans laquelle une canette de soda a été trouvée dans une couche inférieure.

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.

6 Bilan et perspectives

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 permettant aux archéologues d'effectuer le traitement de leurs données sans s'écarter de leurs pratiques courantes mais tout en profitant de l'aide de l'informatique pour vérifier la cohérence de l'ensemble de leurs données.

Pour compléter cette approche, il nous reste encore à définir les modalités du traitement des *terminus post quem*, ainsi qu'à trouver le mode de visualisation des résultats le plus approprié à ce type de données.

Nous sommes en effet persuadés qu'avec un 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., Bénel 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*, à paraître janvier 2004.
- [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] Demoule J.-P., Projet de bibliographie automatique en préhistoire et protohistoire européenne, *Banques de données archéologiques*, Marseille, 12-14 juin 1972.
- [5] 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.
- [6] Jockey Ph., *L'archéologie*, Paris : Belin, 1999.