

---

# Raisonner en logique modale sur l'incertitude des données archéologiques

Élisabeth Gavignet, Éric Leclercq, Nadine Cullot,  
Marinette Savonnet

Laboratoire LE2I - UMR6306 - CNRS - ENSAM  
Univ. Bourgogne Franche-Comté  
9, Avenue Alain Savary  
F-21078 Dijon - France  
Prenom.Nom@u-bourgogne.fr

---

*RÉSUMÉ.* La prise en compte de données imparfaites est un enjeu essentiel pour les systèmes d'information archéologiques, notamment parce que la constitution d'un corpus est un travail intrinsèquement collaboratif et interdisciplinaire. L'imperfection des données collectées est en partie liée à la fiabilité des sources qui contribuent à les enrichir. Notre contribution porte sur l'utilisation de plusieurs logiques modales pour modéliser et tester les effets de la prise en compte de données incertaines, mais aussi pour faciliter la vérification de la qualité du corpus dans une plateforme ouverte et extensible permettant de gérer les données archéologiques. Sur le plan application le raisonneur utilisé, basé sur la logique du premier ordre, fournit aux archéologues les moyens de valider ou non leurs hypothèses. Plusieurs raisonnements avec différentes modalités sont proposés autour d'un exemple simple qui rassemble les types d'incertitudes habituellement rencontrés.

*ABSTRACT.* The consideration of imperfect data is a key issue for archaeological information systems, in particular because the elaboration of a corpus is intrinsically a collaborative and interdisciplinary task. The imperfect property of collected data is partly related to the reliability of the sources contributing to enrich them. Our contribution brings the use of several modal logics to model and verify the effects of the consideration of uncertain data, but also to check the quality of the corpus in an open and extensible platform allowing the management of archaeological data. From the application point of view the reasoner used, based on the first order logic, provides the archaeologists with tools to validate or not their hypotheses. Some case studies of reasoning using different modalities are proposed around a simple example which gathers the kinds of uncertainty usually encountered.

*MOTS-CLÉS :* raisonnement, logiques modales, incertitudes, annotations sémantiques, ontologies  
*KEYWORDS:* reasoning, modal logics, uncertainty, semantic annotations, ontology

---

DOI:10.3166/RIG.28.1-20 © 2016 Lavoisier

## 1. Introduction

Les recherches en archéologie sont de plus en plus interdisciplinaires et collaboratives (Costa, 2012). Les données proviennent de sources multiples (fouilles, données cartographiques anciennes et actuelles, sources historiques, etc.) et leur collecte a recours à différentes techniques (de datation, de caractérisation des matériaux, etc.) opérées par de nombreux spécialistes. Les données sont exploitées dans des cadres variés comme par exemple par des spécialistes de la restauration des bâtiments ou par des organismes du ministère de la culture, pour des missions d'inventaire, d'étude, de protection ou de diffusion du patrimoine archéologique. Pour cela, l'archéologie a donc besoin de systèmes évolutifs, d'outils de détection d'incohérence (Leclercq, Savonnet, 2012) et de modèles de représentation et d'interprétation de données incertaines.

D'un point de vue application, les différents acteurs manipulant les données archéologiques ont, par conséquent, besoin d'une plateforme qui va bien au-delà d'une simple base de données couplée à une application. Cette plateforme doit proposer des solutions pour répondre à deux problématiques essentielles : la première est l'extensibilité des modèles de données sans remettre en cause l'opérabilité des applicatifs et la seconde est la préservation de la qualité des données et des connaissances.

Cependant, la multiplicité des acteurs impliqués accroît la difficulté d'obtenir un consensus sur les données et les connaissances représentées. Les ontologies proposent une spécification formelle et explicite d'une conceptualisation partagée d'un domaine (Gruber, 1993). À ce titre, elles sont un moyen facilitant la représentation des connaissances et la contextualisation des données. Les raisonneurs qui leur sont associés peuvent prendre en charge la détection d'incohérences tout en inférant de nouvelles connaissances. La modélisation des ontologies est généralement basée sur des langage, sous-ensembles de la logique du premier ordre, issus des logiques de description (Baader, 2003) comme OWL (Ontology Web Language) qui n'offrent pas d'opérateurs adaptés à la représentation des connaissances imparfaites. Seuls des systèmes plus complexes mixant plusieurs paradigmes pour la modélisation des connaissances permettent de représenter et d'exploiter cette variété d'information.

La suite de l'article est organisée comme suit. La section 2 présente un état de l'art sur les typologies de l'incertitude à partir de classifications reconnues, montre l'intérêt des logiques modales pour modéliser certaines formes de connaissances incertaines et s'intéresse aux méthodes de raisonnement sur ces connaissances. La section 3 décrit l'architecture de la plateforme WikiBridge, présente une synthèse du modèle d'annotation (Leclercq, Savonnet, 2012) et précise le rôle des ontologies dans la plateforme. La section 4 développe l'utilisation des logiques modales pour des cas distincts de raisonnement : la gestion de la localisation, la datation et l'interprétation multi-acteurs (ou agents). La section 5 dresse le bilan de nos travaux et présente des perspectives plus larges d'usage des logiques modales comme base pour la modélisation, l'interrogation et le raisonnement sur des données incertaines dans les systèmes d'information scientifiques.

## 2. État de l'art

Dans cette partie nous reprenons les classifications de l'incertitude les plus connues ainsi que les outils de modélisation qui peuvent être mis en œuvre pour représenter les données incertaines.

### 2.1. Typologies des incertitudes

L'usage dans la littérature de nombreux synonymes est caractéristique de la difficulté à définir ce que l'incertitude recouvre. Les terminologies pour exprimer l'incertitude sont vagues et aussi nombreuses que les disciplines qui l'étudient. Les principales définitions de l'incertitude que l'on retrouve dans les dictionnaires tournent autour de deux axes : (1) état de quelqu'un qui ne dispose pas de la connaissance nécessaire pour prendre une décision, (2) propriété physique de l'information consécutive aux limites de perception de la réalité. Ces deux définitions ne sont bien évidemment pas complètement indépendantes puisqu'une information incertaine peut induire de l'incertitude dans l'esprit d'une personne. Plusieurs classifications de l'incertitude ont été proposées dans la littérature, il n'est pas dans notre intention d'en donner un inventaire complet mais d'en citer quelques unes qui peuvent permettre de mieux cibler la place de notre contribution.

De nombreux travaux se réfèrent à (Klir, Yuan, 1995) qui proposent une classification s'appuyant sur les différentes théories mathématiques de l'incertitude existantes. Pour Klir et Yuan, l'incertitude peut être imprécision ou ambiguïté (deux types d'incertitude). Cette dernière est elle-même décomposée en deux sous-catégories (conflit et non-spécificité). L'imprécision relève du caractère vague, approximatif de l'information (le mur mesure plus de deux mètres) alors que l'ambiguïté implique la notion de doute. Le conflit est dû à la présence de deux informations contradictoires (le mur est en pierre alors que le bâtiment est en bois). Quant à la non-spécificité elle englobe une information qui pourrait avoir plusieurs sens (le mur est sur la droite - droite de l'objet observé ou à la droite du bâtiment).

(Skeels *et al.*, 2010) décrivent une hiérarchisation en plusieurs niveaux (trois niveaux hiérarchiques et deux niveaux transversaux). Le niveau le plus faible d'incertitude est relatif à la précision de la mesure, le niveau moyen fait référence à la complétude des informations (celles dont on dispose, celles dont on ne dispose pas mais dont on connaît l'existence et celles dont on ignore l'existence) et le niveau le plus élevé correspond aux données produites par des inférences (supposition, extrapolation). Quant aux niveaux transversaux, ils concernent les notions de conflit et de crédibilité.

D'autres travaux comme (Gershon, 1998) et (Smets, 1997) élaborent une classification non pas de l'incertitude mais de l'imperfection. Gershon considère l'incertitude comme l'une des six causes de l'imperfection au même titre que l'incohérence, l'incomplétude, les données erronées, les données complexes et les présentations imparfaites. Smets a établi une classification de l'imperfection (figure 1) éclatée en trois

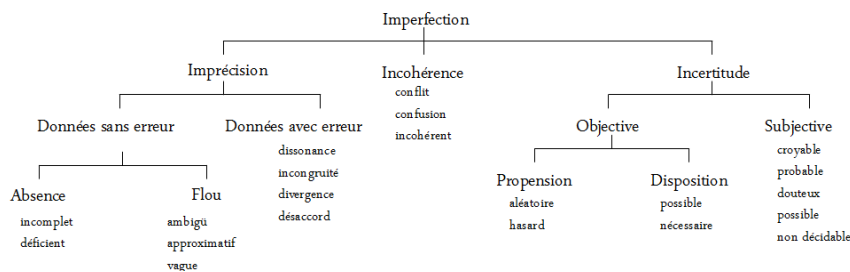


Figure 1. Classification proposée par Smets

principaux groupes : imprécision, incohérence et incertitude. Pour l’auteur, l’imprécision est relative au contenu des déclarations qui peuvent contenir ou non des erreurs. L’incohérence regroupe les déclarations conflictuelles. Quant à l’incertitude, elle est induite soit par des informations objectives (les propriétés de la donnée) soit par des informations subjectives (l’avis de l’observateur).

(Thomson *et al.*, 2005) proposent une répartition de neuf critères d’incertitude dans trois catégories : qualité de l’information, cohérence de l’information et objectivité de l’information (tableau 1). Ces catégories mettent en avant trois aspects de la connaissance : l’information prise indépendamment de son contexte qui doit répondre à certains critères pour être considérée de qualité, l’ensemble des informations sur un sujet donné qui doivent être cohérentes entre elles et enfin, l’importance de la source qui se doit d’être la plus objective possible.

Tableau 1. Synthèse de la typologie de l’incertitude de Thomson *et al.*

Qualité	Cohérence	Objectivité
exactitude de la collecte	concordance entre données	fiabilité des sources
précision de la mesure	transformation	subjectivité
exhaustivité des informations	lacune temporelle	indépendance de la source

Les classifications évoquées précédemment peuvent être séparées en deux groupes conformément aux deux définitions énoncées. Un premier groupe fait nettement apparaître l’importance des sources à l’image de ce qui se retrouve dans (Thomson *et al.*, 2005) avec l’objectivité, dans (Smets, 1997) avec l’aspect subjectif de l’information ou dans (Skeels *et al.*, 2010) avec la notion de crédibilité. À l’inverse dans le second groupe les classifications sont davantage orientées vers la représentation de l’incertitude portant sur l’information elle-même. Ce qui est aussi le cas de (De Runz, 2008), dans le domaine archéologique. Pour De Runz, l’ambiguïté est au même niveau que l’incertitude alors que Smets la considère comme une forme d’imprécision. Cette distinction est tout à fait légitime dans la mesure où De Runz propose cette classification comme support au choix de la théorie de représentation la mieux adaptée pour modéliser l’information. Il conforte l’idée selon laquelle l’ambiguïté reste délicate à traiter. Elle relève à la fois de l’imprécision et de l’incertitude. (Fisher, 1999) introduit la notion d’objet bien défini et d’objet mal défini. La classification des objets mal

définis est similaire à celle de Klir. Dans son approche Fisher associe une théorie de représentation (théorie des probabilités, théorie des ensembles flous) à chaque type d'incertitude. Les objets bien définis sont des objets considérés comme valides par un agent. Dans certains cas, en archéologie par exemple mais pas seulement, plusieurs agents peuvent émettre des avis différents sur le même objet. La prise en compte de la fiabilité de la source est alors indispensable. Une classification proposée à l'occasion de la première table ronde « L'incertitude : regards pluridisciplinaires » organisée par F. Bertonecello en 2013 va dans ce sens. Elle distingue les données lacunaires, les données ambiguës et les données imprécises. À ces catégories d'incertitude liées aux données viennent s'ajouter des incertitudes liées à l'analyse ou à l'interprétation qui est faite de ces dernières.

Chacune de ces classifications reste discutable et selon le domaine, certaines correspondent mieux à une réalité que d'autres. Néanmoins tous les travaux soulignent l'impact de l'incertitude sur la qualité des données et des décisions qui sont prises. La cohérence (respectivement l'incohérence) se retrouve dans une majorité de classifications avec parfois des variantes au niveau du vocabulaire. L'intervention humaine (l'interprétation, la prédiction, l'extrapolation, etc.) est souvent mentionnée comme ayant des conséquences sur la qualité de l'information et difficile à quantifier et à représenter. L'incertitude revêt plusieurs formes. Elle doit être modélisée en choisissant des cas d'utilisation clairement identifiés. Les outils proposés sont majoritairement liés aux théories mathématiques et donnent des résultats intéressants sur les données quantitatives. Mais l'incertitude se manifeste à différents niveaux dans la production de connaissances et peu de travaux semblent prendre en compte la fiabilité des sources (agents ou ressources) et la conduite de raisonnement sur ce type d'incertitude. Aussi pour répondre à cet état de fait, cet article se propose d'illustrer comment l'utilisation d'un cadre théorique formel comme les logiques modales avec ses mécanismes de raisonnement peut permettre de répondre, dans des contextes bien spécifiques, à cette problématique.

## 2.2. *De la logique classique aux logiques modales*

Dans ce paragraphe, nous définissons les fondamentaux de la logique des prédicats du premier ordre puis nous détaillons les concepts des logiques modales utiles à la compréhension de cet article.

La **logique des prédicats** du premier ordre permet de représenter des connaissances à l'aide de formules basées sur la définition de prédicats  $n$ -aires, l'utilisation d'opérateurs logiques (conjonction  $\wedge$ , disjonction  $\vee$ , négation  $\neg$ , implication  $\rightarrow$ ) et de quantificateurs. Une formule composée uniquement d'un prédicat est une formule atomique ou atome. Les formules non atomiques sont construites à partir d'autres formules, des opérateurs logiques et des quantificateurs universel ( $\forall$ ) et existentiel ( $\exists$ ). Le calcul des prédicats s'intéresse d'une façon générale à déterminer la véracité d'une formule. Il n'est généralement pas décidable, aussi est-il souhaitable de restreindre la description des formules à des formes permettant la mise en œuvre de moteurs de rai-

sonnement complet, capables de prouver la véracité ou non d'une formule. Cette restriction consiste à travailler sur une forme clausale particulière : les clauses de Horn. Une clause de Horn a la forme suivante :  $B_1, \dots, B_n \rightarrow H$  où  $B_1, \dots, B_n$  représente le corps de la clause et  $H$  la tête. Chacun des symboles  $H, B_i$  désigne un prédicat. Une clause dont le corps est vide énonce un fait. Une clause avec une tête et un corps non vides traduit une règle. Cette dernière indique que si les conditions  $B_1$  et ... et  $B_n$  sont vraies alors la conclusion  $H$  est aussi vraie. La forme logique d'une telle clause est  $\forall x_1, \dots, x_k (H \vee \neg B_1 \vee \dots \vee \neg B_n)$  où  $x_i$  désigne les variables référencées dans les prédicats.

Les **logiques modales** permettent, par l'introduction de nouveaux opérateurs, appelés modaux (le nécessaire  $\Box$  et son dual le possible  $\Diamond$ ), d'enrichir leur pouvoir d'expression et de modéliser des formules plus nuancées. On peut par exemple, vouloir décrire la proposition « Il est possible que l'église Notre Dame de Dijon ait été une simple chapelle » à l'aide d'une modalité de *possibilité*, ou encore, utiliser la modalité *on sait que* pour affirmer « On sait que l'église Notre Dame de Dijon a été une simple chapelle ».

Il existe plusieurs catégories de logiques modales classées selon les interprétations des opérateurs modaux définis dont les principales sont : logique aléthique (possible ou nécessaire), temporelle (parfois ou toujours), épistémique (croit que, sait que), déontique (est permis, est obligatoire). La compréhension de ces modalités nécessite une extension de la sémantique de la logique des prédicats du premier ordre qui est basée sur des modèles de Kripke (Gasquet *et al.*, 2013). Ces derniers considèrent l'existence de différents mondes reliés entre eux par une relation d'accessibilité. Ainsi, en logique aléthique, une formule  $\varphi$  pourra être interprétée comme *possible*, notée  $\Diamond\varphi$ , si elle est vraie dans un monde et fausse dans un autre et comme *nécessaire*, notée  $\Box\varphi$ , si elle est vraie dans tous les mondes.

Les **systèmes de déduction** s'appuient sur les axiomes de la logique considérée pour réaliser des déductions à partir des formules décrites et connaître la véracité d'une formule. Ils utilisent les mécanismes de substitution pour remplacer une variable par une autre expression et d'unification qui est une succession de substitutions rendant deux formules égales.

La méthode de résolution SLD (*Selective Linear Definite*) (Lloyd, 1987) est une méthode de preuve complète de la logique des prédicats du premier ordre sur un ensemble  $P$  de clauses de Horn. Pour prouver que le but  $F$  est vrai, il faut prouver que  $P \cup \{\neg F\}$  est insatisfiable. Soit  $G = \neg F$ , la démonstration est faite en dérivant à partir de  $G$  une séquence de formules jusqu'à obtenir la clause vide. Les formules  $G_i$  sont telles que  $G_0 = G$  et pour tout  $i \geq 0$ ,  $G_{i+1}$  est obtenu à partir de  $G_i$ , de la manière suivante :

- $A_m$  est un atome sélectionné dans le corps de  $G_i$
- $B_1, \dots, B_n \rightarrow H$  est une clause du programme  $P$
- $\theta$  est l'unificateur le plus général de  $A_m$  et  $H$
- $G_{i+1} = \rightarrow (A_1, \dots, A_{m-1}, B_1, \dots, B_n, A_{m+1}, \dots, A_p)\theta$

La preuve par réfutation peut être schématisée par un arbre de réfutation (figure 2). La clause résultante est reliée aux deux clauses parentes par deux arrêtes étiquetées par l'unificateur employé. Considérons l'ensemble des propositions suivantes : « Saint Léontius est un évêque », « Saint Léontius a pour ville épiscopale Fréjus », « Saint Léontius est mort dans sa ville épiscopale », « tout évêque exerce dans sa ville épiscopale ». Dans l'exemple on choisit pour constantes  $stL$ ,  $eveque$ ,  $frejus$ , pour variables  $x$ ,  $y$ ,  $z$  et pour prédicats  $E(x,y)$  :  $x$  est un  $y$ ,  $V(x, z)$  :  $x$  a pour ville épiscopale  $z$ ,  $M(x, z)$  :  $x$  est mort dans la ville  $z$ . La traduction de ces propositions sous forme clause donne l'ensemble  $P$  des clauses suivantes :  $P = \{ E(stL, eveque), V(stL, frejus), \neg V(stL, z) \vee M(stL, z) \}$  Si l'on cherche à déduire que «  $stL$  est mort à Fréjus », il faut montrer que l'ensemble  $P \cup \{ \neg M(stL, frejus) \}$  est inconsistant.

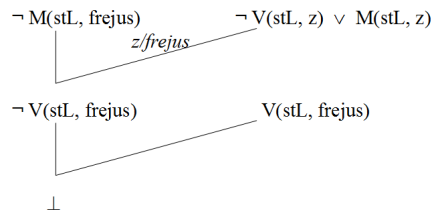


Figure 2. Exemple de résolution SLD : arbre de réfutation

Ce principe peut être étendu pour raisonner sur des formules de la logique modale. Plusieurs approches ont été proposées pour ce faire, comme la méthode des tableaux sémantiques (Fitting, 2012), la méthode des séquents de Gentzen (Szabo, 1969) ou la programmation logique. La méthode des tableaux sémantiques est comme la SLD résolution, une méthode de preuve par réfutation. La négation de la formule à prouver est décomposée en sous-formules de sorte que si la formule est satisfaite alors les sous-formules qui la composent le sont aussi. La décomposition permet la construction d'un arbre et s'effectue par application de règles d'expansion (de prolongation, de ramification, de double négation, etc.) qui tiennent compte de la sémantique des mondes possibles. Une formule a une preuve si toutes les branches de l'arbre sont fermées, c'est-à-dire qu'une formule et sa négation apparaissent dans chaque branche. En programmation logique, les approches consistent à étendre les langages de programmation logique comme Prolog, pour prendre en compte les opérateurs modaux. Elles sont détaillées dans la section 4 et notamment le langage MProlog (Nguyen, 2004) utilisé dans nos travaux.

### 3. Plateforme pour la gestion de données archéologiques

Les données archéologiques sont complexes à gérer car elles sont spatio-temporelles, hétérogènes avec un grand nombre d'attributs, évoluent, proviennent de plusieurs sources et sont sujettes à des interprétations différentes. La constitution et la publication d'un corpus dans un environnement collaboratif représentent une demande forte des archéologues. Le nombre important des acteurs impliqués pose néanmoins le problème de la qualité des informations collectées et de leur concordance les unes avec

les autres. Ce sont ces raisons qui ont conduit au développement de la plateforme WikiBridge<sup>1</sup> (Chevalier *et al.*, 2010) et à la mise en place de moyens pour contrôler la qualité des informations collectées.

### 3.1. Présentation de l'architecture

Le cœur de WikiBridge est constitué par le couplage entre des bases de données relationnelles et des annotations sémantiques pour produire des connaissances dans un environnement collaboratif évolutif. Ce couplage permet de réaliser une extensibilité dynamique des schémas de données tout en conservant les mécanismes de contrôle de la qualité des informations et le fonctionnement des applicatifs. L'interface utilisateur prend la forme d'un wiki pour décrire des objets, des édifices, des lieux, des personnages au moyen d'une description riche incluant textes, sources bibliographiques, plans, localisations, photographies. Les différents éléments d'une page du wiki peuvent être complétés par des annotations construites à partir d'une ontologie sous la forme de triplets *(sujet, prédicat, objet)* (Leclercq *et al.*, 2013).

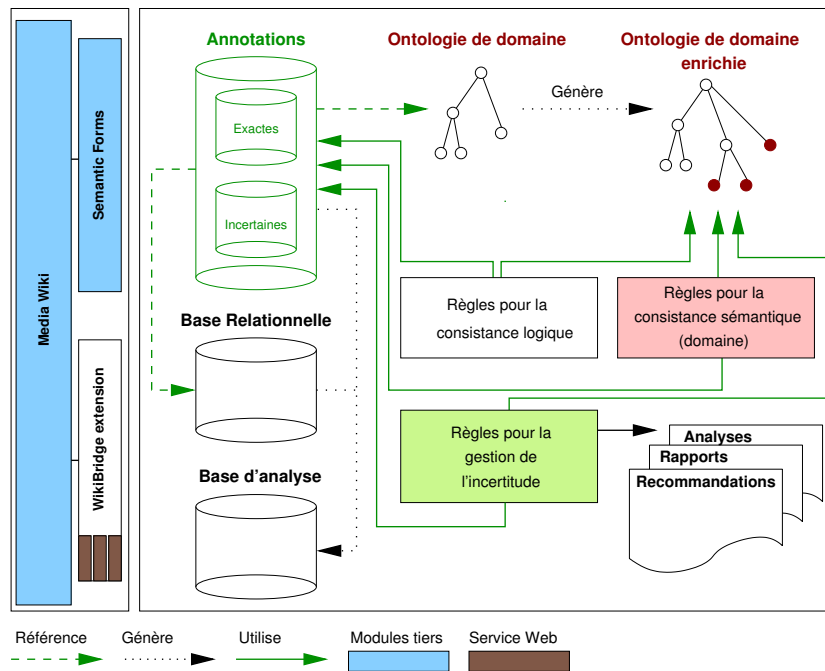


Figure 3. Architecture de WikiBridge

1. développée puis validée dans le cadre de deux projets : l'ANR CARE *Corpus Architecturae Religiosae Europaeae - IV-X saec.* - ANR-07-CORP-011 (<http://care.tge-adonis.fr>) et le projet de Plateforme numérique et mutualiste sur l'usage de la pierre bourguignonne (<http://lapierre.u-bourgogne.fr>).



La figure 3 présente les principaux modules de WikiBridge et leurs interactions. Chaque document du corpus se traduit sous la forme d'une page du wiki. La saisie ainsi que la mise en forme des documents sont implantées en utilisant Semantic Forms<sup>2</sup>, une extension de MediaWiki. Semantic Forms a permis de développer une interface d'acquisition de données basée sur des formulaires autorisant l'ajout de données structurées. Les limites de cette approche apparaissent lorsque les utilisateurs veulent rechercher des informations très précises ou réparties dans plusieurs pages du wiki, la notion de catégories dans MediaWiki n'est pas suffisante car elle a une forte granularité. En conséquence un modèle et un mécanisme d'annotations sémantiques à base ontologique ont été définis pour compléter la base de données relationnelle héritée de MediaWiki. Un assistant d'annotation permet de formuler des annotations : les utilisateurs sélectionnent dans des listes des termes provenant de l'ontologie extraits *via* des requêtes SPARQL et leur associent des propriétés et des valeurs. Un mécanisme de contrôle des annotations incluant la prise en compte du contexte et des contraintes sémantiques a été mis en place. La consistance de ces annotations est soumise à deux types de vérification : l'une logique s'assure que les propriétés associées aux éléments annotés sont cohérentes avec la description ontologique, l'autre sémantique vérifie la cohérence des annotations entre elles (non contradictoires) et la cohérence des annotations par rapport aux règles du domaine (connaissances spécifiques non nécessairement consensuelles). Les annotations ne comportant pas d'incertitude (annotations exactes) sont ajoutées dans l'ontologie afin de constituer l'ontologie enrichie et permettre l'utilisation des mécanismes d'inférence pour effectuer les contrôles de consistance. Les connaissances incertaines sont combinées avec les informations extraites de l'ontologie enrichie pour être analysées par les règles de gestion de l'incertitude et rendre compte des analyses opérées, des déductions réalisées et des recommandations à formuler pour les utilisateurs. Pour l'analyse spatiale, un ensemble de service web offre une interaction avec des outils externes et un module de cartographie s'appuyant sur OpenLayers. L'interfaçage avec le système d'information géographique est réalisé grâce à une base de données spécifique (dite base d'analyse) dont le schéma est établi à partir des champs des formulaires et des annotations.

### 3.2. Ontologies et modèle d'annotations

La figure 4 présente en exemple deux ressources provenant de Wikipédia ayant comme sujet "Léonce de Fréjus". Leur contenu est exploité afin d'enrichir les connaissances au moyen d'annotations sémantiques à base ontologique. L'ontologie de domaine utilisée par WikiBridge est une extension de l'ontologie CIDOM-CRM<sup>3</sup>. Elle est implémentée en OWL<sup>4</sup>. La figure 5 présente un extrait de cette ontologie, qui comporte des classes (Person, Place, PlaceAppellation, etc.) et des propriétés (has-

2. [https://www.mediawiki.org/wiki/Extension:Semantic\\_Forms/fr](https://www.mediawiki.org/wiki/Extension:Semantic_Forms/fr)

3. Comité International pour la Documentation - Conceptual Reference Model qui a été élaborée depuis 1994 et publiée en 2006 par l'ISO (<http://www.cidoc-crm.org>)

4. Ontology Web Language (<http://www.w3.org/TR/owl-features/>)

### Leontius of Fréjus

From Wikipedia, the free encyclopedia

**Saint Leontius** (French: *Léonce de Fréjus*) (ca. 419 – ca. 488) was a **bishop of Fréjus**, in **Provence**. He was born probably at **Nîmes**, towards the end of the fourth century; he died in his episcopal town in 488, according to some authorities, though others say 443 or even 448. His **feast day** is December 1.

(a) Extrait fiche wikipédia anglais

### Léonce de Fréjus

**Saint Léonce de Fréjus** († vers 432), **évêque de Fréjus** fêté le 1<sup>er</sup> décembre. Issu d'une noble famille gallo-romaine, Léonce serait originaire de **Nîmes** et aurait pour frère **Castor**, évêque d'Apt.

En 431, Léonce est soupçonné par le pape **Célestin I<sup>er</sup>**, de **semi-pélagianisme**, une doctrine défendue par Jean Cassien et les moines de Lérins.

(b) Extrait fiche wikipédia français

Figure 4. Ressources portant sur Léonce de Fréjus

CurrentLocation, isIdentifiedBy, etc.) issues de CIDOC-CRM et des classes et des propriétés qui étendent ces spécifications. L'extension présentée concerne les personnages religieux et des informations sur ces personnages et leurs actes. Par exemple, une personne peut être soupçonnée d'un méfait ayant une nature et une date, et ayant pu être dénoncé par une autre personne. Le formalisme G-OWL (Héon, Nkambou, 2013) est utilisé pour représenter graphiquement cet extrait d'ontologie. Les individus qui sont des instances des classes et leurs propriétés ne sont pas décrits sur la figure 5. Ils sont créés ou utilisés pour spécifier les annotations. L'ontologie enrichie est constituée de l'ontologie de domaine et des annotations sur les ressources.

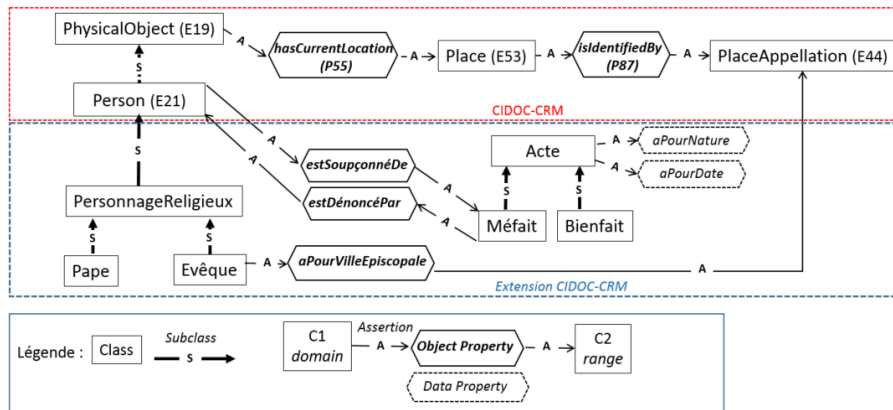


Figure 5. Extrait de l'ontologie de domaine en G-OWL

Dans WikiBridge, le mécanisme d'annotation permet d'annoter n'importe quel élément (portion de texte, image, lien, etc.) d'une page du wiki. Chaque page est identifiée par une URL et les annotations utilisent cette URL comme préfixe pour identifier les éléments annotés. Une annotation est un triplet  $\langle s, p, o \rangle$  liant un sujet  $s$ , un prédicat  $p$  et un objet  $o$ . Le sujet correspond à l'élément que l'on souhaite annoter. Le couple  $(p, o)$  représente le contenu de l'information que l'on souhaite ajouter au sujet. Le prédicat exprime la relation entre le sujet et l'objet. Il donne du sens à l'annotation et peut être un concept ou une propriété de l'ontologie.

Le corpus permet généralement d'extraire de nombreuses informations relatives à un même sujet. Certaines sont directement en relation avec le sujet comme par exemple "Saint Léonce est un évêque" ou "Saint Léonce est fêté le 1<sup>er</sup> décembre". Elles qualifient le sujet identifié "Saint Léonce" (#stL) alors que d'autres peuvent qualifier l'objet en liaison avec le sujet. C'est le cas du soupçon émis par Célestin 1<sup>er</sup> envers Saint Léonce (extrait fiche 4b). Dans ce cas Saint Léonce est soupçonné d'un méfait. Ce dernier est sa croyance en liaison avec le semi-pélagianisme. Il a été dénoncé par Célestin 1<sup>er</sup>. Les deux affirmations précédentes qualifient le méfait dont Saint Léonce est soupçonné. Dans ce cas il est primordial de ne pas perdre le lien avec le sujet (Saint Léonce).

Pour permettre de représenter tous ces cas de figures, trois types d'annotation sont proposés :

1. une annotation *simple* permet d'annoter un sujet *s* représenté par son identifiant en lui associant un prédicat *p* et un objet *o*. Par exemple, l'information "Saint Léonce est un évêque" est traduite par l'annotation simple  $\langle \#stL, estUn, Evêque \rangle$ ;

2. une annotation *réflexive* est représentée par l'imbrication de plusieurs triplets. Elle permet de mettre en relation plusieurs annotations liées au même sujet mais qualifiant l'objet du sujet. Dans l'exemple, ce cas s'applique au soupçon émis par Célestin 1<sup>er</sup> à propos de Saint Léonce. L'annotation réflexive correspondante est la suivante :

$$\langle \#stL, estSoupçonnéDe, acteStL \\ \langle \langle acteStL, aPourNature, croyanceSemiPelagianisme \rangle, \\ \langle acteStL, aPourDate, 431 \rangle, \\ \langle acteStL, estDénoncéPar, Célestin1er \rangle \\ \rangle \rangle$$

Les trois derniers triplets ayant comme sujet *acteStL* ne doivent pas être dissociés du sujet initial. Ils constituent un complément d'information sur l'objet *acteStL* relié au sujet #stL par le prédicat *estSoupçonnéDe*;

3. une annotation *complexe* est un ensemble d'annotations portant sur le même sujet et qui peuvent être simples ou réflexives. Pour l'exemple de Saint Léonce, une annotation complexe pourrait être l'ensemble suivant d'annotations :

$$\langle \langle \#stL, estUn, Evêque \rangle, \\ \langle \#stL, estSoupçonnéDe, acteStL \\ \langle \langle acteStL, aPourNature, croyanceSemiPelagianisme \rangle, \\ \langle acteStL, aPourDate, 431 \rangle, \\ \langle acteStL, estDénoncéPar, Célestin1er \rangle \\ \rangle \rangle, \\ \langle \#stL, aFête, '01-12' \rangle \rangle$$

Dans ce modèle d'annotation, le prédicat qui est un terme de l'ontologie ne peut pas être le sujet d'une annotation car la connaissance consensuelle sur cet élément est exprimée dans l'ontologie.

L'objectif des travaux que nous décrivons dans la suite consiste à compléter les outils de raisonnement présents dans la plateforme WikiBridge par des outils de raisonnement sur des données et des connaissances incertaines représentées par des annotations. En effet, l'incertitude peut se manifester à différents niveaux dans la production de connaissance et elle peut, en particulier, apparaître à la suite d'une détection d'incohérence.

#### 4. Incertitude : de la théorie à l'implémentation

La complexité intrinsèque de l'incertitude rend très difficile la généralisation d'une modélisation universelle. Plusieurs travaux ont utilisé les logiques modales pour traiter de données aussi variées que la sécurité des informations personnelles (Piolle, Demazeau, 2008) en logique déontique, les incidences géométriques (Balbiani *et al.*, 1997) avec des relations temporelles, la réglementation dans des systèmes multi-agents (Garion *et al.*, 2010) reformulée dans un cadre modal du premier ordre, la définition de régions dans l'espace et de raisonnement spatial (Bentham, Bezhanishvili, 2007; Bennett, 1996) mais aucun, à notre connaissance, n'a fait l'objet d'une implémentation en programmation logique modale.

##### 4.1. Raisonnement dans les systèmes de logique modale

De nombreux travaux ont proposé des extensions de la programmation logique avec des approches différentes mais seuls le groupe de Fariñas del Cerro avec le système Molog (Del Cerro, 1986) et Linh Anh NGuyen avec MProlog (Nguyen, 2003) ont abouti à des implémentations. Si les deux systèmes optent pour la même démarche, celui de Nguyen n'impose aucune restriction quant à l'utilisation des opérateurs modaux. Il est basé sur le modèle de Kripke et intègre le calcul de résolution SLD. Ce modèle étendu prend en compte l'axiome de base K ( $(\Box(\varphi \rightarrow \psi) \rightarrow (\Box\varphi \rightarrow \Box\psi))$  où  $\varphi$  et  $\psi$  sont des formules) ainsi que d'autres axiomes (tableau 2) enrichissant le système de base (K-système). Les nouveaux systèmes de logique, obtenus par combinaison des axiomes avec l'axiome K, portent généralement comme nom les initiales des axiomes (par exemple KD pour K + D).

Tableau 2. Axiomes caractérisant des relations entre les mondes de Kripke

Axiome	Formule	Propriété	Relation $R$ entre mondes de Kripke $\omega_i$
D	$\Box\varphi \rightarrow \Diamond\varphi$	sérielle	$\forall\omega_0 \exists\omega_1 R(\omega_0, \omega_1)$
T	$\Box\varphi \rightarrow \varphi$	réflexive	$\forall\omega_0 R(\omega_0, \omega_0)$
B	$\varphi \rightarrow \Box\Diamond\varphi$	symétrique	$\forall\omega_0, \omega_1 R(\omega_0, \omega_1) \wedge R(\omega_1, \omega_0)$
4	$\Box\varphi \rightarrow \Box\Box\varphi$	transitive	$\forall\omega_0, \omega_1, \omega_2 (R(\omega_0, \omega_1) \wedge R(\omega_1, \omega_2)) \rightarrow R(\omega_0, \omega_2)$
5	$\Diamond\varphi \rightarrow \Box\Diamond\varphi$	euclidienne	$\forall\omega_0, \omega_1, \omega_2 (R(\omega_0, \omega_1) \wedge R(\omega_0, \omega_2)) \rightarrow R(\omega_1, \omega_2)$

Un programme MProlog est un ensemble de formules restreintes aux clauses de Horn. Une clause avec modalité est une formule de la forme  $[\Delta] : \varphi$  où  $\Delta$  est une liste d'opérateurs modaux séparés par des virgules et  $\varphi$  une clause de Horn pouvant comporter des opérateurs modaux. Les formules proposées dans le tableau 3 illustrent des possibilités d'expression des modalités et montrent qu'un opérateur modal peut indifféremment porter sur une clause ou sur un prédicat, dans le corps ou la tête de la clause.

Tableau 3. Exemples de modalités avec MProlog

ref	formule avec modalités	traduction en MProlog
$\varphi_1$	$\diamond m(\text{jade}, \text{alain})$	$[\text{d}]:\text{m}(\text{jade}, \text{alain}).$
$\varphi_2$	$\square(\square f(x) \leftarrow m(x, y))$	$[\text{b}]:([\text{b}]:\text{f}(\text{X}):- \text{m}(\text{X}, \text{Y})).$
$\varphi_3$	$\square(p(x, y) \leftarrow m(x, y), \square f(x))$	$[\text{b}]:(\text{p}(\text{X}, \text{Y}) :- \text{m}(\text{X}, \text{Y}), [\text{b}]: \text{f}(\text{X})).$
$\varphi_4$	$\square e(x) \leftarrow \diamond p(y, x)$	$[\text{b}]:\text{e}(\text{X}) :- [\text{d}]:\text{p}(\text{Y}, \text{X}).$

Dans (Nguyen, 2003) l'auteur détaille l'extension du calcul de la résolution SLD aux logiques modales dans une approche fidèle à la programmation logique classique. Lors de la construction de l'univers des mondes, pour rendre compte d'une formule  $\diamond\varphi$  dans le monde actuel  $w_0$  il convient de connecter  $w_0$  à un monde  $w_1$  et d'ajouter  $\varphi$  dans  $w_1$ . Le fait que  $\varphi$  soit vrai dans le monde  $w_1$  est traduit sous la forme suivante  $\langle\varphi\rangle\varphi$ . De la même manière, tout ce qui pourra être déduit dans le monde  $w_1$  sera étiqueté de  $\langle\varphi\rangle$ . Cette étiquette indique que les propositions déduites le sont dans un monde où la proposition  $\varphi$  est supposée vraie. Cela reflète la notion de possible : possible dans  $w_1$  mais pas forcément ailleurs dans un autre monde accessible depuis  $w_0$ . À l'inverse, la prise en compte d'une formule  $\square\varphi$  dans le monde actuel  $w_0$  va se traduire par l'ajout de  $\varphi$  dans tous les mondes accessibles depuis  $w_0$ .

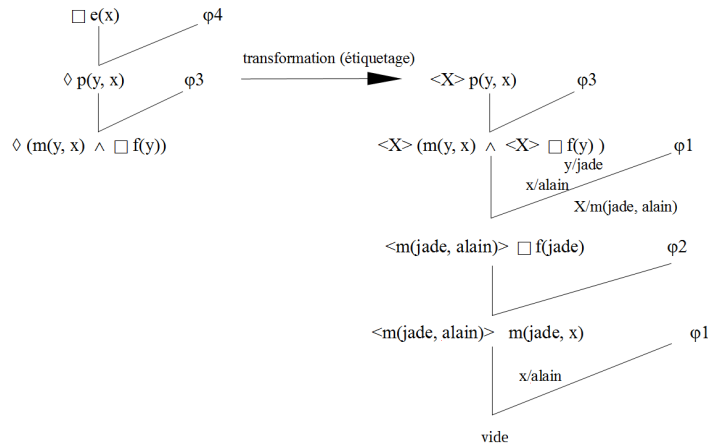


Figure 6. Exemple de résolution étendue à la logique monomodale KD

Pour illustrer les étapes du raisonnement, nous proposons de prendre l'exemple de la résolution du but  $G = \leftarrow \square e(x)$  dans la logique monomodale KD. Le processus

de résolution est schématisé dans la figure 6. Il s'appuie sur le programme P constitué des formules incluant des modalités présentées dans le tableau 3. En considérant  $G$  et  $\varphi_4$  nous dérivons un nouveau but  $G_1 = \leftarrow \Diamond p(y, x)$ . À partir de  $G_1$  et  $\varphi_3$ , nous déduisons le but  $\Diamond(m(y, x) \wedge \Box f(y))$  qui est constitué d'une séquence de deux opérateurs modaux. Or dans les logiques monomodales, la séquence d'opérateurs modaux a une longueur au plus égale à un. Pour permettre néanmoins la résolution de la formule  $\Diamond p(y, x)$ , cette dernière va être étiquetée par  $\langle X \rangle$  où  $X$  est une variable libre (non quantifiée) pour devenir  $\langle X \rangle p(y, x)$ . L'application de  $\varphi_3$  et  $K$  à ce but étiqueté va donner le nouveau but  $G_2 = \leftarrow \langle X \rangle m(y, x) \wedge \langle X \rangle \Box f(y)$ . Le rapprochement de  $G_2$  et  $\varphi_1$  avec l'hypothèse  $\langle m(\text{jade}, \text{alain}) \rangle m(\text{jade}, \text{alain})$  fournit  $G_3 = \leftarrow \langle m(\text{jade}, \text{alain}) \rangle \Box f(\text{jade})$  grâce à l'unificateur  $y/\text{jade}, x/\text{alain}, X/m(\text{jade}, \text{alain})$ . De  $G_3$  et  $\varphi_2$  nous obtenons  $G_4 = \leftarrow \langle m(\text{jade}, \text{alain}) \rangle m(\text{jade}, y)$ . En appliquant  $\varphi_1$  à  $G_4$  nous dérivons  $G_5 = \leftarrow \langle m(\text{jade}, \text{alain}) \rangle m(\text{jade}, \text{alain})$ . Et finalement, en unifiant  $G_5$  et  $\varphi_1$  nous arrivons à la clause vide. La solution au but recherché est donc alain. Cet exemple détaille les principaux mécanismes mis en œuvre par MProlog pour résoudre le but, lesquels seront très utiles pour comprendre les cas d'utilisation traités.

#### 4.2. Modélisation et implémentation de différentes formes d'incertitude

Dans les sous sections suivantes, nous montrons comment disposer d'un formalisme basé sur la logique modale et la logique des prédicats pour modéliser certaines formes d'incertitude afin de rendre compte des effets de ces données incertaines sur l'ensemble des données existantes. Nous considérerons l'incertitude liée d'une part aux données représentées et d'autre part à l'observateur qui en fait l'interprétation.

Parmi les ressources proposées (Figure 4), nous nous appuyerons sur celle qui renferme le plus d'incertitudes (extrait fiche 4a). Dans cette fiche, les informations considérées comme certaines sont : (a) St Léontius est évêque de Fréjus en Provence, (b) il est né vers la fin du 4<sup>e</sup> siècle, (c) il est mort dans sa ville épiscopale, (d) sa fête est le 1<sup>er</sup> décembre. Il est important d'attirer l'attention sur le fait que l'extrait ne précise pas qu'il est mort à Fréjus. Cependant cette information fera partie des déductions possibles. Concernant les informations incertaines, seule celle relative au lieu de naissance est concernée : (e) il est probablement né à Nîmes. Enfin, la présence de trois dates différentes (f) pour la mort de Léonce de Fréjus soulève des questions. Plusieurs solutions sont envisageables : ne modéliser aucune donnée au risque de les perdre, modéliser une seule d'entre elles ou mémoriser toutes ces données. C'est cette dernière alternative que nous retiendrons. Le stockage de plusieurs valeurs pour une caractéristique usuellement monovaluée pose problème dans bon nombre de représentations. Dans une base de données relationnelle cette difficulté peut être contournée mais systématiser la solution à tous les attributs ne serait pas pertinent. Au niveau de l'ontologie, si la propriété a été définie comme fonctionnelle une incohérence sera automatiquement détectée. En logique des prédicats du premier ordre il est possible d'énoncer trois dates mais disposer de trois valeurs pour un même événement relève d'un cas d'incertitude. Il y a forcément des valeurs non correctes parmi celles fournies et tout au plus une valeur correcte. C'est à partir de ces informations que nous

proposons d'étudier deux modalités (aléthique et épistémique) pour illustrer une modélisation différenciée d'une même source.

**Cas 1 : implémentation en logique modale aléthique.** Par la suite, pour qu'une annotation puisse être traitée par le système MProlog elle devra être traduite par la forme clausale correspondante. Celle-ci est constituée du prédicat d'arité trois *an* (abrégé de annotation) et des valeurs présentes dans le triplet. Ce qui donne, par exemple, *an(#stL, aPourVilleEpiscopale, 'Fréjus')* pour le triplet  $\langle \#stL, aPourVilleEpiscopale, Fréjus \rangle$ .

Dans ce cas l'ensemble des annotations sont considérées comme fournies par la même personne à partir de plusieurs ressources ou par plusieurs personnes à partir de plusieurs ressources sans qu'il soit utile de différencier la provenance. L'objectivité de la personne qui rédige les annotations n'est pas remise en cause. Les modalités permettent ainsi de qualifier les données par rapport à leur véracité sans tenir compte de la fiabilité de la personne à l'origine de l'information. Le tableau 4 propose la traduction des annotations de la fiche 4a concernant Léonce de Fréjus. Pour montrer

Tableau 4. Formules en logique modale aléthique

Information	Formule avec ou sans modalité	Ref
(a)	<i>an(#stL, estUn, Evêque)</i>	(f1)
	<i>an(#stL, aPourVilleEpiscopale, 'Fréjus')</i>	(f2)
	<i>an('Fréjus', estLocaliseEn, 'Provence')</i>	(f3)
(e)	$\diamond an(\#stL, estNeA, 'Nimes')$	(f4)
(b)	$\square an(\#stL, estNePdt, fin4s)$	(f5)
(c)	$an(\#stL, aPourVilleEpiscopale, V) \rightarrow an(\#stL, estMortA, V^a)$	(r1)
(f)	$\diamond an(\#stL, estMortEn, 488)$	(f6)
	$\diamond an(\#stL, estMortEn, 443)$	(f7)
	$\diamond an(\#stL, estMortEn, 448)$	(f8)
(d)	<i>an(#stL, aFete, '01-12')</i>	(f9)

a. Une valeur identifiée par un nom commençant par une majuscule représente une variable. La variable V représente la ville où est mort #stL et qui est aussi sa ville épiscopale.

la prise en compte des incertitudes (possible et nécessaire) dans le raisonnement, nous rajoutons aux formules précédentes (tableau 4) deux règles :

- (r2) :  $an(\text{Suj}, \text{Pred}, \text{Obj}) \rightarrow \square an(\text{Suj}, \text{Pred}, \text{Obj})$
- (r3) :  $\square (an(\text{Suj}, \text{estMortEn}, D_m), an(\text{Suj}, \text{estNeEn}, D_n),$   
 $Age = D_m - D_n \rightarrow an(\text{Suj}, \text{estMortA}, Age))$

La règle r2 permet de traiter les formules sans modalité comme des formules toujours nécessairement vraies. La règle se lit « Si la formule est vraie alors il est nécessaire qu'elle soit vraie ». Ce qui est tout à fait pertinent par rapport à l'interprétation qui a été proposée car un seul point de vue est considéré. Ce qui est vrai dans un monde doit l'être aussi dans tous les mondes accessibles depuis ce monde. Les faits f1, f2 et la règle r1 permettent de déduire que Léonce est mort à Fréjus. Avec r2, cette nouvelle formule est vraie non seulement dans le monde actuel mais dans tous les mondes ac-

cessibles depuis le monde actuel. La règle r2 permet de pouvoir propager les formules à tous les mondes.

La règle r3 permet de déduire l'âge d'une personne au moment de son décès en fonction de son année de mort ( $D_m$ ) et de son année de naissance ( $D_n$ ). La modalité *nécessaire* ( $\square$ ) s'applique, dans r3, à l'ensemble de la formule. La formule r3 sera donc vraie dans tous les mondes accessibles depuis le monde courant.

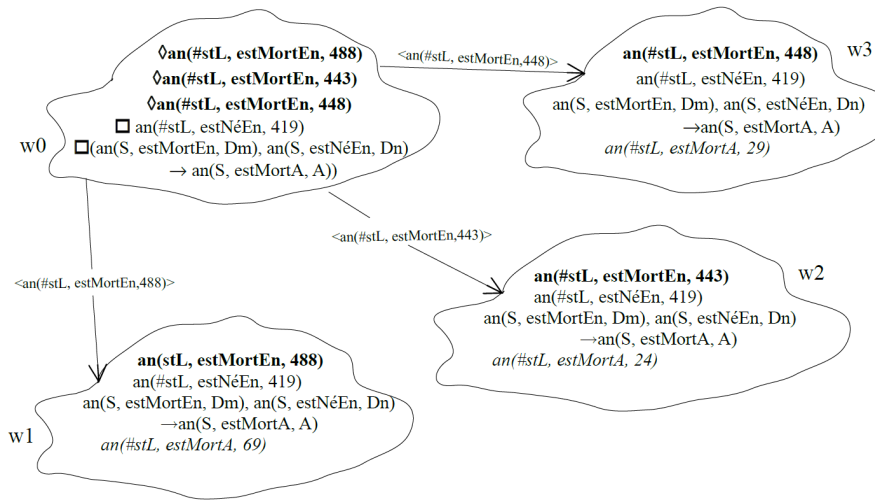


Figure 7. Représentation de la sémantique des mondes

À des fins de vérification de cohérence, il est intéressant d'obtenir toutes les données incertaines et notamment celles qui ont été déduites. La résolution du but  $\diamond an(S, P, O)$  va répondre à cette attente. À partir du monde actuel w0, comportant toutes les règles, un ensemble de mondes (figure 7) accessibles vont être générés à partir des faits possibles. Par exemple, l'unification de la formule (r3) dans le monde (w1) où la formule (f6), stipulant que Léonce de Fréjus est mort en 488, est considérée comme vraie va inférer le fait que Léonce de Fréjus est mort à l'âge de 69 ans. Cette déduction est reproduite, à l'identique, dans chacun des mondes et la requête initiale va restituer les formules qui étaient déjà incertaines (f6, f7, f8) auxquelles vont être rajoutées les trois nouvelles propositions déduites et notées en italique dans la figure 7. Dans la figure seuls les éléments utiles ont été reproduits.

L'exemple proposé à partir des dates possibles montre à quel point la multiplication des doutes peut générer des déductions tout aussi incertaines. La visualisation des déductions opérées va permettre à l'archéologue de s'interroger sur leur véracité, la nécessité ou non de les conserver en l'état ou de les transformer. Il pourra ainsi faire des simulations pour modéliser au plus juste le monde réel. Cependant certaines fois ce choix reste délicat et d'autres besoins peuvent se faire sentir comme la possibilité de différencier les sources (ressources, agents).



**Cas 2 : implémentation en logique multimodale épistémique.** La multiplicité des acteurs impliqués dans la collecte des données peut à un moment ou à un autre justifier de pouvoir différencier les sources. Dans ces situations particulières, quand les circonstances justifient un intérêt à conserver toutes les informations pour opérer des déductions différenciées, il est nécessaire de passer à une autre forme de logique : la logique multimodale  $K_{(m)}$  avec  $m$  paires d'opérateurs modaux  $\Box_i$  et  $\Diamond_i$ . Un extrait des axiomes relatifs à ces logiques est représenté dans le tableau 5.

Tableau 5. Extrait des axiomes de la logique multimodale

Nom	Règle	Commentaire
(D)	$\Box_i \varphi \rightarrow \neg \Box_i \neg \varphi$	consistance
(I)	$\Box_i \varphi \rightarrow \Box_j \varphi$ if $i > j$	$i, j$ : degré de croyance
(4s)	$\Box_i \varphi \rightarrow \Box_j \Box_i \varphi$	introspection positive
(5s)	$\neg \Box_i \varphi \rightarrow \Box_j \neg \Box_i \varphi$	introspection négative

La communauté des archéologues à laquelle nous nous intéressons collabore à la constitution de bases de connaissances communes au sein d'un groupe de confiance. Chacun veut pouvoir exprimer ses croyances et raisonner en exploitant celles des autres. D'une certaine manière, c'est bien le point de vue des agents (sources) que nous cherchons à représenter et non plus les connaissances de manière globale indépendamment de leur source. Ayant besoin de représenter une croyance individuelle, nous avons retenu la logique  $KD4_s5_s$  pour laquelle  $\Box_i \varphi$  signifie « l'agent  $i$  croit que  $\varphi$  est vraie » et  $\Diamond_i \varphi$  signifie «  $\varphi$  est considérée comme possible par l'agent  $i$  » où  $\varphi$  est une formule.

Pour illustrer ce qui peut être fait avec cette logique, nous proposons d'exprimer le recoupement de trois sources différentes (S1, S2, S3) alimentant la connaissance sur le même sujet Léonce de Fréjus. S1 et S2 sont considérées comme existantes avant la prise en compte de S3. Les croyances attribuées à S1 et S2 ont été données de manière à couvrir un maximum de combinaisons possibles. En revanche la source S3 correspond au contenu de l'extrait décrit figure 4a. Par ailleurs, comme nous nous focalisons sur la représentation des données incertaines, seul un sous-ensemble (f2, r1) des formules sans modalité du tableau 4 ainsi que deux des trois formules portant sur les dates (f6, f7) sont reprises ici. Ces formules sont reproduites dans le tableau 6. La première ligne du tableau schématise en réalité deux croyances portant sur la même formule : d'une part que S1 sait que Léonce est mort en 488 et d'autre part que S3 croit que Léonce soit mort en 488. De manière similaire la deuxième ligne fait état que S2 sait que Léonce est mort en 443 alors que S3 croit seulement que c'est possible. En effet, en présence des deux dates proposées par S1 et S2, S3 ne sait pas laquelle des deux dates est la bonne. Pour lui les deux dates sont envisageables puisque S1 et S2 l'affirment et qu'il les retrouve dans S3. À l'inverse, toutes les sources savent que Léonce est né en 419 (m3) et que le calcul de l'âge d'un individu au moment de son décès est la différence entre l'année de mort et celle de naissance (r3). Quant à la règle relative au lieu de sa mort (m6), les trois sources émettent un doute. Ainsi représentées, les données peuvent être consultées par rapport à leur fiabilité.

Tableau 6. Formulation des données en  $KD4_s5_s$ 

Ref	Formule	S1	S2	S3
(m1)	$\text{an}(\#stL, \text{estMortEn}, 488)$	$\square_1$		$\diamond_3$
(m2)	$\text{an}(\#stL, \text{estMortEn}, 443)$		$\square_2$	$\diamond_3$
(m3)	$\text{an}(\#stL, \text{estNeEn}, 419)$	$\square_1$	$\square_2$	$\square_3$
(m4)	$\text{an}(\#stL, \text{estNeA}, 'Nimes')$	$\square_1$	$\square_2$	$\diamond_3$
(m5)	$\text{an}(\#stL, \text{aPourVilleEpiscopale}, 'Fréjus')$	$\square_1$	$\diamond_2$	$\square_3$
(r3)	$\text{an}(S, \text{estMortEn}, D_m), \text{an}(S, \text{estNeEn}, D_n),$ $\text{Age} = D_m - D_n \rightarrow \text{an}(S, \text{estMortA}, \text{Age})$	$\square_1$	$\square_2$	$\square_3$
(m6)	$\text{an}(\#stL, \text{aPourVilleEpiscopale}, V) \rightarrow \text{an}(\#stL, \text{estMortA}, V)$	$\diamond_1$	$\diamond_2$	$\diamond_3$

Dans l'optique d'un travail collaboratif, il peut être intéressant d'utiliser cette logique multimodale pour relier la donnée à sa source, disposer de moyens pour vérifier des convergences respectivement des divergences. À tout moment S3 peut consulter les données qu'il confirme, celles qui deviennent probables et les autres qui sont uniquement possibles. À titre d'illustration, la règle suivante :

$$\square_3(\text{an}(S, P, O), \text{for\_all}(I, [1, 2], \square_I \text{an}(S, P, O))) \rightarrow \text{confirm}(S, P, O)$$

retrouve toutes les annotations confirmées par S3. À savoir celles considérées comme vraies pour S3 et aussi considérées comme vraies pour les sources S1 et S2. Seule l'annotation (m3) répond à ce critère. Une formule logique similaire pourrait être écrite pour retrouver les annotations probables c'est-à-dire possibles pour S3 mais nécessairement vraies pour toutes les autres sources. Ce qui donnerait ici les annotations (m3) et (m4). Les annotations (m1) et (m2) ne font pas partie du résultat car elles sont effectivement possibles pour S3 mais ne sont nécessaires que dans une seule des deux autres sources. En logique modale KD, les annotations confirmées sont forcément des annotations probables. En effet, ce qui est vrai partout est forcément vrai dans un des mondes. De façon analogue, une règle peut détecter toutes les annotations considérées comme possibles pour au moins une source. Avec ce type de modalités, il devient possible de déduire des informations, opérer des recoupements afin d'aider l'archéologue à juger de la qualité des données collectées.

## 5. Conclusion

La modélisation et le raisonnement sur des données incertaines sont des tâches complexes qui nécessitent l'usage d'environnement multi-paradigmes. Le travail présenté dans l'article s'intègre dans une plateforme collaborative de gestion de données archéologiques qui supporte un modèle d'annotation sémantique à base ontologique acceptant l'extensibilité du modèle de données. Nos travaux sur le raisonnement en présence d'incertitude rejoignent et complètent les travaux sur la qualité de données déjà développés au sein de la plateforme.

L'approche décrite dans cet article présente et illustre la logique monomodale aléthique avec les opérateurs modaux de *possibilité* et de *nécessité*, et la logique multimodale épistémique qui élargit la précédente avec la notion d'agents qui peuvent avoir

des croyances différentes. Les logiques modales sont utilisées pour représenter les connaissances incertaines mais aussi pour définir des règles de gestion de ces données permettant de faire émerger des connaissances utiles aux archéologues.

Les travaux et l'expérimentation réalisés ont permis d'envisager plusieurs perspectives de développement en relation avec différents points de la proposition. La première est liée aux mécanismes de recommandation qui peuvent être ajoutés à la plateforme afin de contrôler la qualité des données, y compris multi-sources, mais aussi pour confronter les points de vue ou les interprétations des archéologues, historiens, etc. La seconde, plus théorique, concerne la spécification d'un langage d'interrogation de données avec des opérateurs de plus haut niveau permettant de prendre en compte l'incertitude et d'interroger les annotations exactes et incertaines.

#### Remerciements

*Ces travaux ont été partiellement financés par le CNRS au travers du projet PEPS LogModArchéo en 2013 et 2014*

#### Bibliographie

- Baader F. (2003). *The description logic handbook: theory, implementation, and applications*. Cambridge university press.
- Balbani P., Cerro L. F. del, Tinchev T., Vakarelov D. (1997). Modal logics for incidence geometries. *Journal of Logic and Computation*, vol. 7, n° 1, p. 59–78.
- Bennett B. (1996). Modal logics for qualitative spatial reasoning. *Logic Journal of IGPL*, vol. 4, n° 1, p. 23–45.
- Benthem J. van, Bezhanishvili G. (2007). Modal logics of space. In *Handbook of spatial logics*, p. 217–298. Springer.
- Chevalier P., Leclercq E., Millereux A., Sapin C., Savonnet M. (2010). WikiBridge: a Semantic Wiki for Archaeological Applications. In *Proceedings of the 38th conference on computer applications and quantitative methods in archaeology (caa)*, p. 193–196.
- Costa L. (2012). La construction de référentiels géohistoriques: un enjeu pour l'interdisciplinarité dans les sciences historiques. *L'Espace géographique*, vol. 41, n° 4, p. 340–351.
- Del Cerro L. F. (1986). Molog: A system that extends prolog with modal logic. *New Generation Computing*, vol. 4, n° 1, p. 35–50.
- De Runz C. (2008). *Imperfection, temps et espace: modélisation, analyse et visualisation dans un sig archéologique*. Thèse de doctorat. Université de Reims Champagne-Ardenne.
- Fisher P. F. (1999). Models of uncertainty in spatial data. *Geographical information systems*, vol. 1, p. 191–205.
- Fitting M. (2012). *First-order logic and automated theorem proving*. Springer Science & Business Media.
- Garion C., Roussel S., Cholvy L. (2010). Une logique modale pour raisonner sur la cohérence et la complétude de réglementations. *Revue d'intelligence artificielle*, vol. 24, n° 3, p. 267–290.

- Gasquet O., Herzig A., Said B., Schwarzenruber F. (2013). *Kripke's worlds: An introduction to modal logics via tableaux*. Springer Science & Business Media.
- Gershon N. (1998). Visualization of an imperfect world. *Computer Graphics and Applications, IEEE*, vol. 18, n° 4, p. 43–45.
- Gruber T. R. (1993). A translation approach to portable ontology specifications. *Knowledge acquisition*, vol. 5, n° 2, p. 199–220.
- Héon M., Nkambou R. (2013). G-owl: Vers un langage de modélisation graphique, polymorphe et typé pour la construction d'une ontologie dans la notation owl. In *1c-24èmes journées francophones d'ingénierie des connaissances*.
- Klir G., Yuan B. (1995). *Fuzzy sets and fuzzy logic* (vol. 4). Prentice Hall.
- Leclercq E., Savonnet M. (2012). Système d'information pour la production de connaissances. l'approche wiki sémantique. *Ingénierie des Systèmes d'Information*, vol. 17, n° 3, p. 143–166.
- Leclercq E., Savonnet M., Troya-Galvis A., Buttner S. (2013). Investigating a multi-paradigm system for the management of archaeological data: Corpus lapidum burgundiae. In *Digital heritage international congress (digitalheritage), 2013*, vol. 1, p. 679–682.
- Lloyd J. W. (1987). *Foundations of logic programming. symbolic computation: Artificial intelligence*. Springer-Verlag, 2nd, extended edition edition.
- Nguyen L. A. (2003). A fixpoint semantics and an sld-resolution calculus for modal logic programs. *Fundamenta Informaticae*, vol. 55, n° 1, p. 63–100.
- Nguyen L. A. (2004). The modal logic programming system mprolog. In *Logics in artificial intelligence*, p. 266–278. Springer.
- Piolle G., Demazeau Y. (2008). Une logique pour raisonner sur la protection des données personnelles. In *16e congrès francophone africain sur la reconnaissance de formes et l'intelligence artificielle (rfia'08)*, p. 8p.
- Skeels M., Lee B., Smith G., Robertson G. G. (2010). Revealing uncertainty for information visualization. *Information Visualization*, vol. 9, n° 1, p. 70–81.
- Smets P. (1997). Imperfect information: Imprecision and uncertainty. In *Uncertainty management in information systems*, p. 225–254. Springer.
- Szabo M. E. (1969). *The collected papers of gerhard gentzen* (vol. 160). North-Holland Amsterdam.
- Thomson J., Hetzler E., MacEachren A., Gahegan M., Pavel M. (2005). A typology for visualizing uncertainty. In *Electronic imaging 2005*, p. 146–157.