

Méthodologie assistée de conception d'une ontologie à partir d'une conceptualisation consensuelle semi-formelle

Michel Héon¹, Gilbert Paquette¹, Josianne Basque¹

¹ Centre de recherche LICEF, Télé-Université, Montréal, Canada
(michel.heon, gilbert.paquette, josianne.basque)@licef.ca

Résumé : Cet article présente une méthodologie assistée de conception d'une ontologie à travers trois méthodes, soit une méthode d'élicitation des connaissances d'un domaine résultant en un modèle semi-formel de ces connaissances, une méthode de formalisation conduisant à la production d'une ontologie et une méthode de validation syntaxique et sémantique de l'ontologie. Les processus de formalisation et de validation sont assistés par un système expert à la formalisation dont la base de connaissances est une ontologie de transformation.

Mots-clés : Formalisation des connaissances, Ingénierie des connaissances, Ingénierie ontologique, Élicitation, Méta-représentation, Ontologie de transformation, Système expert, Conception assistée d'ontologie.

1 Introduction

Cet article apporte une contribution au domaine de l'ingénierie ontologique (Dietz, 2006 ; Gašević *et al.*, 2006 ; Gómez-Pérez *et al.*, 2003 ; Gruber, 1995 ; Uschold et Gruninger, 1996). Dans ses grandes lignes, la construction d'une ontologie nécessite une étape d'élicitation des connaissances du domaine visé, suivie d'une étape de formalisation et d'une étape de validation. La nature des connaissances à représenter, l'hétérogénéité du support des connaissances à formaliser (documents plus ou moins formels, communication orale ou écrite, représentations graphiques, connaissances tacites, etc.) et le manque d'assistance automatisée, rendent laborieuse la démarche de conception d'une ontologie. À l'intérieur des courants actuels, notre méthodologie préconise la conception d'ontologies génériques et réutilisables tel que le propose (Gangemi *et al.*, 2007) du projet NeOn (www.neon-project.org) et peut facilement s'intégrer à des infrastructures technologiques à base d'ontologies comme KAON (kaon.semanticweb.org), On-To-Knowledge (Davies *et al.*, 2003) et TelOs du projet LORNET (www.lornet.org).

Nous adoptons la définition d'une ontologie proposée par Gruber (1995) qui stipule qu'une ontologie est un document formel (au sens de Uschold et Gruninger (1996)) dont le contenu et la sémantique sont traitables par des systèmes informatiques et dont la connaissance qui y est exprimée est obtenue de manière consensuelle. La stratégie

de co-modélisation des connaissances à l'aide d'un langage graphique semi-formel expérimentée par Basque *et al.* (2008b) s'accorde bien à l'aspect consensuel de cette définition d'ontologie. Cette activité nous semble ainsi pouvoir être mise à profit pour la conception d'une ontologie. Ainsi, nous pensons que le processus de construction d'une ontologie doit être décomposé en deux phases bien distinctes: une phase *d'élicitation* de la connaissance dans un formalisme de degré semi-formel relativement évolué, puis une phase de *formalisation* des connaissances où le modèle semi-formel est transformé dans un formalisme ontologique.

Malgré le fait qu'un modèle semi-formel peut comporter des éléments d'ambiguïté, la souplesse et le caractère moins artificiel d'un langage semi-formel permettent d'accéder plus facilement à l'expression de connaissances tacites, car la spontanéité n'est pas bloquée par la charge cognitive associée à une formalisation plus poussée (Basque *et al.*, 2008b). L'usage d'un système de représentation plus convivial permet aussi d'élargir le bassin des personnes aptes à représenter leurs connaissances en même temps qu'il évite la démobilité des experts de leur tâche principale, laquelle coûte cher à une organisation. L'élicitation dans un formalisme de degré semi-formel peut ainsi représenter une économie de temps et surtout un gain de qualité dans la représentation des connaissances. De plus, la convivialité du langage semi-formel fait en sorte que des modèles de degré semi-formel peuvent être conçus par les experts de contenu sans l'assistance d'un ingénieur de la connaissance, lequel peut ensuite les formaliser avec la participation minimale des experts les ayant conçus ou encore d'autres experts. Les opérations d'élicitation peuvent ainsi être intégrées de manière plus souple dans les activités des experts de contenu.

Partant de l'hypothèse qu'un modèle semi-formel peut être formalisé sous la forme d'une ontologie sans perdre les distinctions entre les différents types de connaissances (notamment entre les connaissances de nature déclarative, procédurale et stratégique au sens de (Paquette, 2002 ; Paris *et al.*, 1983)), cet article présente une méthodologie de conception d'une ontologie, assistée par un système expert, qui inclut une méthode d'élicitation, une méthode de formalisation et une méthode de validation. Nous comptons, par cette méthodologie, rendre plus accessible et plus rapide la construction d'une ontologie.

2 Le système de représentation des connaissances de la méthodologie

Issu du domaine de l'ingénierie pédagogique, le langage de modélisation par objets typés MOT, de degré semi-formel, est celui qui a été utilisé pour développer la méthodologie proposée. Notre but est toutefois de faire en sorte que celle-ci puisse s'appliquer à d'autres langages semi-formels tels que ceux utilisés pour les cartes conceptuelles, les réseaux sémantiques ou les diagrammes UML. Le langage et le logiciel qui l'implémente (MOT*Plus*) ont été conçus par une équipe sous la direction du second auteur (Paquette, 2002) au Centre de recherche LICEF de la Télé-Université (www.liceef.ca).

Le langage semi-formel MOT différencie les types de connaissances au moyen de symboles graphiques (voir le **tableau 1** et le **tableau 2**). Les connaissances peuvent être combinées au sein d'un même schéma de manière à produire des modèles mixtes de connaissances.

Tableau 1. Catégories des connaissances dans le langage MOT

Catégorie de connaissance	Connaissance abstraite		Connaissance factuelle	
Déclarative <i>le quoi des choses</i>	Concept		Exemple	
Action <i>Le comment de choses</i>	Procédure		Trace	
Stratégique <i>Le pourquoi, le quand</i>	Principe		Énoncé	

Le *concept* représente « le quoi » des choses. Il sert à décrire l'essence d'un objet concret. Il peut être associé à l'idée de classe ou de catégorie. En ce sens, il est l'abstraction d'un objet concret. L'*exemple* représente l'un de ces objets en énonçant un certain nombre de faits qui le décrivent. La *procédure* permet de décrire « le comment » des choses. Elle désigne des opérations, des actions pouvant être accomplies. La *trace* représente l'ensemble des faits concrets obtenus lors de l'exécution d'une procédure. Le *principe* désigne « le pourquoi », « le quand » d'une chose. Il est une connaissance stratégique qui permet de nommer une relation qui existe entre des objets, que ce soit des concepts, des procédures ou d'autres principes. Il sert notamment à représenter une condition pouvant s'appliquer à l'exécution d'une action. L'*énoncé* représente l'instanciation d'un principe à propos d'objets concrets.

Tableau 2. Sémantique des relations typées dans MOT

Type de lien	Signification
S	Le lien de <i>spécialisation</i> associe deux connaissances abstraites de même type dont la première est une spécialisation de la seconde. Ce lien est notamment utile dans la description des taxonomies.
I	Le lien d' <i>instanciation</i> associe à une connaissance abstraite l'ensemble des faits qui caractérisent une instance de cette connaissance.
I/P	Le lien <i>intran/produit</i> sert à associer une connaissance procédurale à une connaissance conceptuelle afin de représenter l'intrant ou le produit d'une procédure. Ce lien est notamment utile dans la description des algorithmes, des processus et des méthodes.
P	Le lien de <i>précédence</i> associe une connaissance à une autre qui la suit dans une séquence temporelle de procédures ou de règle de décision (principes).
R	Le lien de <i>régulation</i> associe une connaissance stratégique (un <i>Principe</i> ou un <i>Énoncé</i>) à une autre connaissance afin de préciser une contrainte, une restriction ou une règle qui régit la connaissance.
C, C*	Les liens de <i>composition</i> et de <i>composition multiple</i> permettent de représenter l'association entre une connaissance et des connaissances qui la composent.

Les *relations* sont des liens directionnels qui unissent des connaissances. Le langage MOT offre un ensemble de liens qui sont typés (voir le **tableau 2**). Chaque

type de lien possède une sémantique propre qui respecte des règles d'intégrité. Par exemple, un lien de spécialisation (lien S) unit deux connaissances abstraites qui doivent être de même nature. L'ensemble de ces règles d'intégrité est décrit dans Paquette (2002).

3 Méthodologie de conception d'une ontologie

La **fig. 1** présente les trois principales méthodes qui composent la méthodologie que nous proposons. Dans l'ordre, ces méthodes concernent les procédures : *concevoir un modèle semi-formel*, *formaliser en ontologie* et *valider l'ontologie*.

La méthode *concevoir un modèle semi-formel*, basée sur le principe de co-modélisation, a pour objectif la conception d'un *modèle semi-formel de domaine* impliquant la collaboration d'au moins deux *experts de contenu* et optionnellement d'un *ingénieur de la connaissance* qui pilote l'activité de modélisation. Un ou plusieurs novices du domaine peuvent remplacer l'un des experts d'un contenu ou se joindre à ceux-ci afin de favoriser à la fois le processus d'explicitation des connaissances par les experts (suscité par les questions pointues et contextualisées des novices) et le transfert d'expertise chez les novices (Basque *et al.*, 2008). À ce jour, cette étape a été réalisée en présence dans nos travaux, mais nous envisageons de développer une version collaborative de l'outil *MOTPlus* qui permettrait de la réaliser à distance.

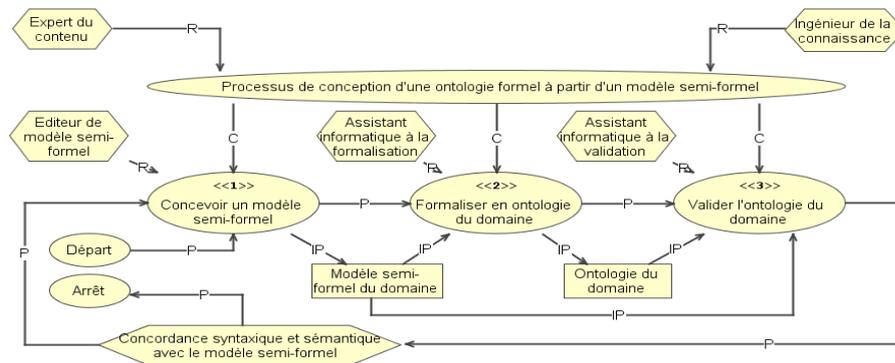


Fig. 1: Schéma MOT de la méthodologie de conception d'une ontologie

Lorsque le modèle semi-formel est jugé satisfaisant pour les acteurs participant aux séances d'élicitation, l'ingénieur a la responsabilité de piloter la *méthode de formalisation* qui a pour extrant une *ontologie du domaine*. Assisté par un *assistant informatique à la formalisation*, qui est en fait un système expert utilisant une *ontologie de transformation*, l'ingénieur formalise le modèle semi-formel de domaine en une ontologie de domaine prête à être validée.

Finalement, la *méthode de validation*, pilotée cette fois conjointement par l'ingénieur de la connaissance et l'expert de contenu, permet la *validation syntaxique et sémantique* de l'ontologie de domaine. La validation syntaxique a pour objectif

d'assurer une concordance entre les éléments déclarés dans le modèle semi-formel et la représentation de ces éléments dans l'ontologie. La validation sémantique a pour objectif de valider la concordance sémantique entre le modèle semi-formel et l'ontologie. *L'assistant informatique à la validation* permet de tirer des conclusions automatiques à partir de l'ontologie du domaine, qui sont comparées avec celles obtenues humainement à partir de l'interprétation du modèle semi-formel.

3.1 Méthode d'élicitation

L'élicitation des connaissances est un domaine de recherche vaste et complexe qui a fait l'objet de nombreux travaux depuis plus d'une vingtaine d'années (Hart, 1986). De récentes recherches de Basque *et al.* (2008a ; 2008b) ont montré que des activités de transfert d'expertise impliquant la co-modélisation entre experts et novices permettent de produire des modèles de connaissances dont le contenu ont les propriétés d'être consensuelles et riches en connaissances du domaine.

Afin de supporter et de stimuler la démarche intellectuelle des acteurs dans l'activité de co-modélisation, il importe de mettre à leur disposition un langage de représentation qui, par sa sémantique, impose une certaine structure à la pensée, de manière à en favoriser son déploiement, et ce, sans en entraver la créativité. Les langages de modélisation de degré semi-formel, tel que MOT, répondent à ces exigences Basque *et al.* (2008a ; 2008b). Des recherches menées dans des contextes d'apprentissage ont mis en évidence les médiations cognitives suscitées par l'usage d'un tel langage lors de l'élaboration de représentations graphiques de connaissances (Basque et Pudelko, sous presse).

Le schéma de la **fig. 2** présente la méthode utilisée pour concevoir un modèle semi-formel selon cette approche.

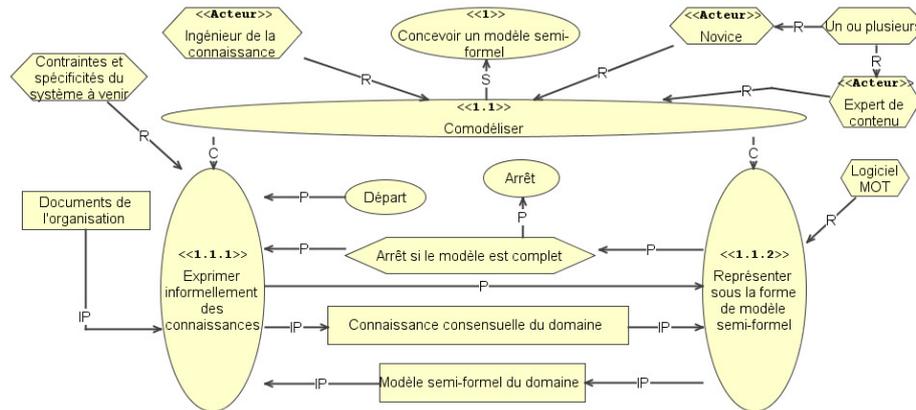


Fig. 2: La méthode de conception d'un modèle semi-formel fondée sur une stratégie de comodélisation

Le processus de *comodélisation*, qui est ici représenté en tant que sorte de processus d'*élicitation*, se compose d'une activité d'*expression informelle des*

connaissances et d'une activité de *représentation structurée semi-formellement des connaissances consensuelles de domaine*, ce qui conduit à la production, avec le logiciel MOT, d'un *modèle semi-formel du domaine*. Ces deux activités sont itératives et le produit de l'un sert d'intrant à l'autre. L'expression informelle est régie par les *contraintes et les spécifications du système à venir*. Des *documents contenant des savoirs et des savoir-faire de l'organisation* peuvent également servir en intrant à l'activité. Le processus de *comodélisation* se termine lorsque les éléments d'intérêts du domaine ont été représentés à la satisfaction des acteurs en présence.

3.2 Méthode de formalisation

La formalisation d'un modèle semi-formel en une ontologie est un processus délicat dont l'objectif est de produire une ontologie du domaine. Pour qu'elle soit opérable, l'ontologie du domaine doit à la fois respecter la syntaxe du langage formel et représenter la sémantique du domaine contenue dans le modèle semi-formel

L'application de la méthode de formalisation est sous la responsabilité de l'*ingénieur de la connaissance* qui est assisté par un *assistant informatique à la formalisation* (voir la **fig. 3**). La méthode de formalisation se décompose en trois processus: *importer dans l'espace de modélisation ontologique, désambiguïser et transformer en ontologie*.

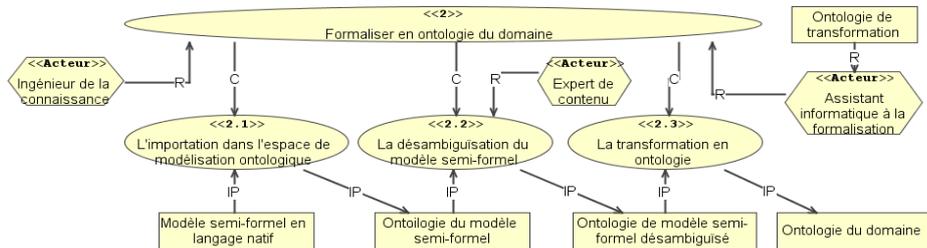


Fig. 3: Méthode de transformation du modèle semi-formel de domaine en ontologie de domaine

3.2.1 Architecture des composants de la méthode

Le diagramme de composants de la **fig. 4** présente les différents éléments et les relations qui composent l'assistant informatique à la formalisation. D'un point technologique, l'assistant est une application Eclipse (www.eclipse.org) qui hérite des propriétés de gestion des données, de conception d'interface graphique, d'intégration de "plug-in", d'interopérabilité par l'utilisation de Java et de capacité de développement des aspects procéduraux nécessaires à la tâche de formalisation et de communication avec l'utilisateur.

Une base de connaissances incarnée par l'*Ontologie de transformation* assure la cohérence entre les divers modules de l'assistant afin d'offrir un support intelligent à l'ingénieur. Quatre domaines distincts composent l'ontologie de transformation. Le domaine de *l'ontologie des langages semi-formels* contient un ensemble de

connaissances sur les langages semi-formels qui permettent aux modules d'importation d'intégrer les entités du modèle semi-formel dans l'espace de modélisation ontologique (transposition d'un modèle de format EMF en format OWL). L'ontologie contient les métamodèles des langages semi-formels supportés par la tâche de formalisation. L'ontologie de désambiguïsation contient les connaissances nécessaires à la désambiguïsation des éléments du modèle semi-formel qui seront transformés en une classification générique dans l'ontologie de référence. À chaque schéma de langage correspond une ontologie de désambiguïsation spécifique. Le domaine représenté par l'ontologie de référence permet de classifier de façon générique et non ambiguë les différents éléments du modèle semi-formel. L'ontologie de référence sert d'adaptateur entre les diverses représentations possibles du modèle semi-formel et sa représentation finale sous forme d'ontologie du domaine.

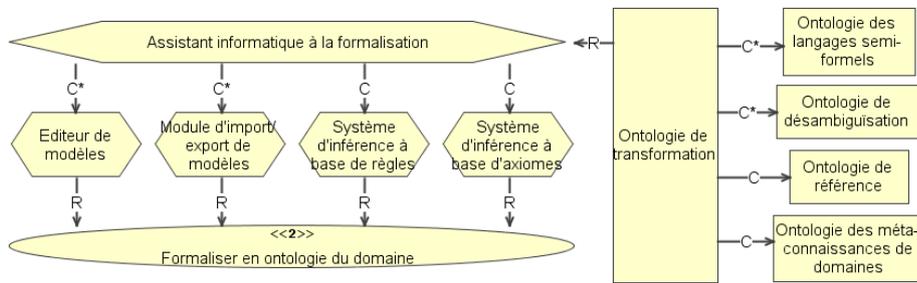


Fig. 4: Composants de l'assistant à la formalisation

Finalement, le domaine représenté par l'ontologie des méta-connaissances de domaine, qui est une ontologie de niveau générique au sens de Oberle (2006 p. 46), englobe un ensemble de méta-connaissances concernant les domaines d'application. Elle contient notamment des méta-connaissances au sujet de connaissances procédurales, déclaratives et stratégiques et des méta-propriétés telles que A-POUR-COMPOSANT, EST-LE-PRECEDENT-DE (voir la fig. 5). Cette ontologie générique est considérée comme un langage de représentation formel de domaines d'application et elle est importée par l'ontologie du domaine faisant ainsi partie résultat de la

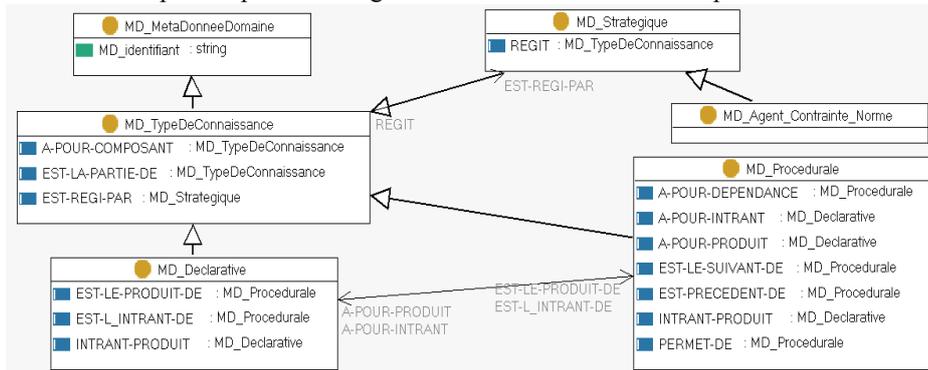


Fig. 5: Ontologie générique des méta-connaissances de domaines

méthode de formalisation.

3.2.2 Processus d'importation, de désambiguïsation et de transformation

Le processus automatique d'*importation* de modèles semi-formels assure la traduction d'un modèle de format natif à l'éditeur semi-formel vers un modèle de format OWL. La manipulation des éléments ontologiques (classes, propriétés, axiomes et individus) est alors interfacée par les classes Java générées. Nos travaux antérieurs ont permis de confirmer la validité de cette approche pour l'importation de modèles semi-formels MOT dans l'espace ontologique OWL (Héon *et al.*, 2008).

Le processus de *désambiguïsation* consiste à supprimer les ambiguïtés contenues dans le modèle semi-formel. Trois catégories de test sont réalisées, soit les tests *typologique*, *topologique* et *sémantique*. La *désambiguïsation typologique* est la plus automatique des désambiguïsations. Elle associe le type d'un composant du modèle semi-formel à un élément ontologique. Par exemple (voir le cas 1 du **tableau 3** et du **tableau 4**), le composant « lien S » (*sorte-de*) s'interprète comme une relation d'hyponymie et le composant « lien I » (*instance*) s'interprète comme une relation d'instanciation. La *désambiguïsation topologique* est plus complexe et peut, dans quelques cas, n'être que semi-automatisée. Cette désambiguïsation s'établit en caractérisant les composants du modèle par l'identification d'un patron de disposition. À l'exemple du cas 2 du **tableau 3** et du **tableau 4**, les concepts C4 et C5 sont interprétés comme des classes et le principe P1 comme une propriété (binaire) mettant en relation ces classes, grâce à l'identification du patron de disposition suivant: un *principe* est uni par un *lienR* à un *concept* en intrant et est uni par un *lienR* à un *concept* extrant. Également, les connaissances stratégiques P1 et A1 qui sont représentées par des *Principes* dans le modèle semi-formel sont désambiguïsées en *Propriété* pour P1 et en *connaissance stratégique* pour A1.

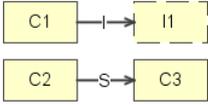
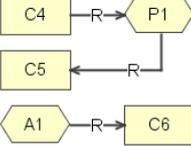
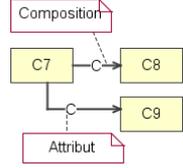
Du point de vue de l'ingénieur de la connaissance, l'étape de *désambiguïsation selon la sémantique du domaine* est délicate, car elle nécessite une compréhension du domaine qui est modélisé, ce qui implique la collaboration de l'expert de contenu afin de répondre aux questions de l'ingénieur. L'exemple d'interprétation du *lien de composition* dans le langage MOT est une bonne illustration de ce type d'ambiguïté (voir le cas 3 du **tableau 3**). Le lien de composition s'interprète de deux façons. La première interprétation possible est la composition entre concepts (exemple: *C7 a pour partie C8*). La deuxième interprétation possible est l'utilisation du lien de composition pour assigner des attributs à un concept (exemple : *C7 a pour attribut C9*), c'est-à-dire que la relation unit une classe à un objet *DataType*. Seule une connaissance adéquate du domaine de connaissances permet de lever l'ambiguïté liée à cette double interprétation du *LienC*.

Le dernier processus de la méthode de formalisation est l'*application des règles de transformation* afin de transformer le modèle semi-formel désambiguïsé en ontologie formelle. Ce processus est réalisé par un système expert utilisant des axiomes et des règles du *Semantic Web Rule Language* (SWRL) (Horrocks *et al.*, 2004). Le **tableau 4** présente le résultat de la formalisation en ontologie du domaine du modèle semi-formel présenté au **tableau 3**.

3.2.3 Expérimentation

Le **tableau 3** présente trois cas de figure de modélisation semi-formelle à formaliser. Les résultats de la réalisation des étapes de modélisation (**tableau 3a**); d'importation (**tableau 3b**), de désambiguïsation (**tableau 3c**) et de transformation (**tableau 4**) y sont respectivement représentés en langage MOT et OWL dans la Notation 3 (N3) (Berners-Lee, 1998).

Tableau 3. Études de cas représentés en langage MOT et en langage OWL-N3.

Cas 1) Relations d'instanciation et de subsumption	Cas 2) Un principe en tant que propriété ou agent	Cas 3) Relations de composition ou d'attribution
A) Représentation semi-formelle du domaine		
		
B) Représentation après importation dans l'espace de modélisation ontologique		
<pre> :C1 a metaMot:MOT_Concept . :C2 a metaMot:MOT_Concept . :C3 a metaMot:MOT_Concept . :I1 a metaMot:MOT_Exemple . :LienI_C1_I1 a metaMot:MOT_LienI ; metaMot:MOT_connDestination :I1 ; metaMot:MOT_connSource :C1 . :LienS_C2_C3 a metaMot:MOT_LienS ; metaMot:MOT_connDestination :C3 ; metaMot:MOT_connSource :C2 . </pre>	<pre> :C4 a metaMot:MOT_Concept . :C5 a metaMot:MOT_Concept . :C6 a metaMot:MOT_Concept . :P1 a metaMot:MOT_Principe . :A1 a metaMot:MOT_Principe . :LienR_C4_P1 a metaMot:MOT_LienR ; metaMot:MOT_connDestination :P1 ; :LienR_C5_P1 a metaMot:MOT_LienR ; metaMot:MOT_connDestination :P1 ; :LienR_A1_C6 a metaMot:MOT_LienR ; metaMot:MOT_connDestination :C6 ; metaMot:MOT_connSource :A1 . </pre>	<pre> :C7 a metaMot:MOT_Concept . :C8 a metaMot:MOT_Concept . :C9 a metaMot:MOT_Concept . :LienC_C7_C8 a metaMot:MOT_LienC ; metaMot:MOT_connDestination :C8 ; metaMot:MOT_connSource :C7 . :LienC_C7_C9 a metaMot:MOT_LienC ; metaMot:MOT_connDestination :C9 ; metaMot:MOT_connSource :C7 . </pre>
C) Représentation après la désambiguïsation		
Cas 1)	<pre> :C1 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :C2 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :C3 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :I1 a oRef:OR_Entite_Observable_Declaratif , metaMot:MOT_Exemple . :LienI_C1_I1 a oRef:OR_Relation_Instance , metaMot:MOT_LienI ; ... :LienS_C2_C3 a oRef:OR_Relation_Hyponyme , metaMot:MOT_LienS ; ... </pre>	
Cas 2)	<pre> :C4 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :C5 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :C6 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :P1 a oRef:OR_Relation_Propriete_Objet , metaMot:MOT_Principe . :A1 a oRef:OR_Entite_Abstraite_Strategique_Agent , metaMot:MOT_Principe . :LienR_P1_C5 a oRef:OR_Relation_Propriete_Regulation , metaMot:MOT_LienR ; ... :LienR_C4_P1 a oRef:OR_Relation_Propriete_Regulation , metaMot:MOT_LienR ; ... :LienR_A1_C6 a oRef:OR_Relation_Propriete_Regulation , metaMot:MOT_LienR ; ... </pre>	
Cas 3)	<pre> :C7 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :C8 a oRef:OR_Entite_Abstraite_Declaratif , metaMot:MOT_Concept . :C9 a oRef:OR_Entite_Schema_String , metaMot:MOT_Concept ; :LienC_C7_C8 a oRef:OR_Relation_Holonyme , metaMot:MOT_LienC ; ... :LienC_C7_C9 a oRef:OR_Relation_Propriete_Attribut , metaMot:MOT_LienC ; ... </pre>	

Chacun des cas est associé à un type de désambiguïsation particulier. Ainsi, le premier cas, qui présente un modèle composé des liens de subsumption et d'instanciation, fait appel à une désambiguïsation de type *typologique* qui classe: les *Concepts* C1, C2, C3 en tant qu'*owl:Class*, avec C2 qui subsume C3; et l'*Exemple* I1 en tant qu'individu OWL appartenant à C1. Le deuxième cas, qui fait référence à une

désambiguïisation *topologique*, permet de désambiguïser les *Principes* P1 et A1 respectivement en *owl:ObjectProperty* et *owl:class* de catégorie *Agent_Contrainte_Norme*. Finalement, le troisième cas nécessite une compréhension du domaine afin de désambiguïser de façon *sémantique* l'interprétation du *LienDeComposition*, qui: entre C7 et C8 se formalise par une *owl:ObjectProperty* de catégorie *A-POUR-COMPOSANT* dont le domaine est l'*owl:class* C7 et l'*owl:class* C8; et entre C7 et C9 qui se formalise par une *owl:DatatypeProperty* de catégorie *A-POUR-ATTRIBUT* dont le domaine est l'*owl:class* C7 et l'image une *xsd:string*.

Tableau 4. Modèle semi-formel formalisés en ontologie et représentés en OWL-N3

Cas 1)	Cas 2)
<pre> :C1 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :C3 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :C2 a owl:Class ; rdfs:subClassOf :C3 . :I1 a :C1 . </pre>	<pre> :A1 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Agent_Contrainte_Norme . :C4 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :C5 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :C6 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :P1 a owl:ObjectProperty ; rdfs:domain :C4 ; rdfs:range :C5 . :A1_regit_C6 a owl:ObjectProperty ; rdfs:domain :A1 ; rdfs:range :C6 ; rdfs:subPropertyOf metaDom:REGIT . </pre>
Cas 3)	
<pre> :C7 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :C8 a owl:Class ; rdfs:subClassOf owl:Thing , metaDom:MD_Declarative . :C7_aPourComposant_C8 a owl:ObjectProperty ; rdfs:subPropertyOf metaDom:A-POUR-COMPOSANT . :C7_aPourAttribut_C9 a owl:DatatypeProperty ; rdfs:domain :C7 ; rdfs:range xsd:string ; rdfs:subPropertyOf metaDom:A-POUR-ATTRIBUT . </pre>	

3.3 Méthode de validation

La méthode de validation se divise en deux thèmes, soit la validation syntaxique et la validation *sémantique*. La validation syntaxique a pour but d'assurer que tous les éléments du modèle semi-formel sont représentés dans l'ontologie selon les règles du langage ontologique utilisé, dans notre cas, OWL-DL. La réalisation de cette validation se décompose en deux étapes, soit *générer un modèle semi-formel* à partir de l'ontologie du domaine et *comparer les éléments* du modèle semi-formel généré avec les éléments de modèle semi-formel d'origine. Si la concordance est exacte, alors il est véridique de conclure que l'ontologie du domaine représente syntaxiquement l'ensemble des éléments du modèle semi-formel du domaine.

La validation sémantique s'attarde sur le sens des représentations qui sont consignées dans l'ontologie. À titre d'exemple, une erreur occasionnellement commise par les concepteurs de modèles peu familiers au langage MOT est issue d'une confusion entre le lien S ("sorte de" [*is-a*]) et le lien C ("composé de" [*part-of*]). Par exemple, ils pourraient être tentés d'exprimer que *planète* se compose de *vénus*, *terre*, *mars*, etc., au lieu d'exprimer que *vénus*, *terre*... sont des sortes de *planètes*. Une inférence sur l'ontologie de domaine qui déterminerait que la propriété inverse de *composeDe* est *faitPartieDe* permettrait de conclure, que *vénus*, *terre*, *mars*... font partie de *planète*. Ce type de conclusion erronée devrait servir de signal à l'ingénieur et aux experts concernant une possible erreur de sémantique dans la représentation du domaine.

La première étape à réaliser pour une validation sémantique est la production de *conclusions formelles* en appliquant des inférences sur l'ontologie du domaine. Des scénarios de test de déductions peuvent être utilisés afin de formaliser d'avantage l'étape. L'étape de *comparaison des conclusions* entre les réponses qu'auraient données l'ingénieur et l'expert aux déductions et les conclusions automatiques permettent d'entreprendre une réflexion qui peut, d'une part, amener l'expert de contenu à revoir la façon de représenter ce qu'il exprime, ou d'autre part, mettre fin au processus de construction de l'ontologie du domaine. Gómez-Pérez (2004) présente les critères selon lesquels l'ontologie devrait être validée: la *consistance* identifie s'il y a des contradictions entre éléments ontologiques; la *complétude* assure que tous les éléments ontologiques sont soit explicitement déclarés ou soit inférables; la *concision* est un principe qui stipule que seuls les éléments à être définis doivent être définis; l'*expansibilité* est la capacité d'ajouter des nouvelles connaissances sans modifier les anciennes; la *sensibilité* est la capacité à réagir à des modifications.

4 Conclusion

Dans cet article, nous avons présenté une méthodologie de conception d'une ontologie à partir d'une représentation semi-formelle de domaine. Son originalité réside notamment dans le fait qu'elle est assistée par un système expert à la formalisation. Ce système expert assiste l'ingénieur au cours du processus de formalisation et assiste l'ingénieur et l'expert dans la validation syntaxique et sémantique de l'ontologie du domaine. La méthodologie intègre les principes de consensualité et de formalité propres aux ontologies par l'utilisation de la méthode de modélisation pour la production de modèles semi-formels à l'étape d'élicitation. Finalement, il est démontré que la méthode permet la formalisation de connaissances procédurale et stratégique.

Dans cet article, la méthodologie utilise les langages MOT et OWL dans son application. Cependant, bien que ce ne soit pas ici démontré, on notera que la méthodologie est conçue afin qu'elle soit généralisée dans son application à d'autres langages semi-formel et formel. Cette particularité sera démontrée dans de futurs travaux.

Présentement, la méthodologie est validée de manière fonctionnelle à partir de modèles fictifs. Dans un avenir proche, nous comptons appliquer la méthodologie à une évaluation en laboratoire avec des cas, des experts non fictifs, et des modèles issus d'activités de co-modélisations.

Références

- BASQUE J., DESJARDINS, C., PUDELKO, B. & LÉONARD, M. (2008a). Gérer les connaissances stratégiques dans des entreprises manufacturières de la Montérégie: expérimentation de la co-modélisation des connaissances dans 3 PME. Montréal/Canada, Rapport de recherche. Montréal: CEFRIO.: 118 p En ligne. <https://www.cefr.io.qc.ca/upload/1599_rapportcefrionalotechfinal12nov08.pdf>.

- BASQUE J., PAQUETTE, G., PUDELKO, B. & LÉONARD, M. (2008b). Collaborative Knowledge Modeling with a Graphical Knowledge Representation Tool: A Strategy to Support the Transfer of Expertise in Organizations. In, Alexandra Okada, Simon Buckingham Shum et Tony Sherborne Eds. *Knowledge Cartography. Mapping Techniques and Software Tools*. London: Springer-Verlag.
- BASQUE J. & PUDELKO, B. (sous presse). Intersubjective Meaning-Making in Dyads Using Object-Typed Concept Mapping. In, P.L. Torres et R.C.V. Marriott Eds. *In Handbook of Research on Collaborative Learning Using Concept Mapping*: IGI Global.
- BERNERS-LEE T. (1998). «Notation 3». En ligne. <<http://www.w3.org/DesignIssues/Notation3.html>>.
- DAVIES J., FENSEL, D. & HARMELEN, F. V. (2003). *Towards The Semantic Web : Ontology-Driven Knowledge Management*: John Wiley & Sons.
- DIETZ J. L. G. (2006). *Entreprise Ontology: Theory and Methodology*. Coll. «Computer Science». Berlin Heidelberg: Springer-Verlag.
- GANGEMI A., GOMEZ-PEREZ, A., PRESUTTI, V. & SUAREZ-FIGUEROA. (2007). «Towards a Catalog of OWL-based Ontology Design Patterns». In *12 Conference of the Spanish Association for Artificial Intelligence* (12-16 November 2007): Springer.
- GAŠEVIĆ D., DJURIĆ, D. & DEVEDŽIĆ, V. (2006). *Model Driven Architecture and Ontology Development*. New York, Inc.: Springer-Verlag.
- GÓMEZ-PÉREZ A. (2004). Ontology Evaluation. In, R. Studer S. Staab Eds. *Handbook on Ontologies*, p. 251-274. New York: Springer.
- GÓMEZ-PÉREZ A., FERNÁNDEZ-LÓPEZ, M. & CORCHO, O. (2003). *Ontological Engineering : with examples from the areas of Knowledge Management, e-Commerce and the Semantic Web*, First edition. New York: Springer.
- GRUBER T. R. (1995). «Toward principles for the design of ontologies used for knowledge sharing». *Int. J. Hum.-Comput. Stud.* vol. 43, no 5-6, p. 907-928.
- HART A. (1986). *Knowledge acquisition for expert systems*. New York: McGraw-Hill.
- HÉON M., PAQUETTE, G. & BASQUE, J. (2008). «Transformation de modèles semi-formels en ontologies selon les architectures conduites par les modèles». In *2èmes Journées Francophones sur les Ontologies* (1-2 Décembre).
- HORROCKS I., BOLEY, H., TABELT, S., GROSOFF, B. & DEAN, M. (2004). «SWRL: A Semantic Web Rule Language Combining OWL and RuleML». W3C. En ligne. <<http://www.w3.org/Submission/SWRL/>>. Consulté le 29 mai.
- OBERLE D. (2006). *Semantic Management of Middleware* Coll. «Semantic Web And Beyond Computing for Human Experience»: Springer US.
- PAQUETTE G. (2002). *Modélisation des connaissances et des compétences : un langage graphique pour concevoir et apprendre*. Sainte-Foy: Presses de l'UQ.
- PARIS S., LIPSON, M. Y. & WIXSON, K. K. (1983). «Becoming a Strategic Reader». *Contemporary Educational Psychology*. vol. 8, p. 293-316.
- USCHOLD M. & GRUNINGER, M. (1996). «Ontologies: Principles, Methods and Applications». *Knowledge Engineering Review*. vol. 11, no 2, p. 93-155