

Évolution d'une ontologie dédiée à la représentation de relations n-aires

Rim Touhami^{*,**} Patrice Buche^{**,**}
Juliette Dibie-Barthélemy^{*,****}, Liliana Ibănescu^{*,****}

^{*}INRA - Mét@risk, 16 rue Claude Bernard, F-75231 Paris Cedex 5, France
rim.touhami@agroparistech.fr

^{**}INRA - UMR IATE, 2, place Pierre Viala, F-34060 Montpellier Cedex 2, France
Patrice.Buche@supagro.inra.fr

^{***}LIRMM, Montpellier, France

^{****}AgroParisTech, 16, rue Claude Bernard, F-75231 Paris Cedex 5, France
{Juliette.Dibie,Liliana.Ibanescu}@agroparistech.fr

Résumé. Nous nous intéressons dans cet article à la problématique d'évolution d'une ontologie permettant de représenter des relations n-aires. Nous présentons la représentation formelle des changements applicables à notre ontologie permettant de modifier sa structure tout en maintenant sa cohérence structurelle. Nous illustrerons nos propos sur une ontologie dédiée à la représentation de relations n-aires entre des données expérimentales quantitatives.

1 Introduction

Dans cet article, nous nous intéressons à la problématique d'évolution d'une ontologie dédiée à la représentation de relations n-aires pour l'annotation de documents dans le cadre du Web Sémantique. Une relation n-aire est une relation qui est définie entre au moins deux arguments. Nous nous appuyerons sur la représentation de relations n-aires sans arguments différenciés telles que proposées par le W3C¹, ce qui correspond au cas le plus général d'utilisation des relations n-aires, le cas 3. Nous avons de plus choisi d'utiliser le "patron 1" qui consiste à représenter une relation n-aire à l'aide d'un concept, relié à ses arguments par des propriétés, comme dans l'exemple de la figure 1. Nous nous intéressons plus particulièrement aux relations n-aires entre des données quantitatives expérimentales, ce qui suppose d'apporter une attention particulière à la gestion des arguments numériques (i.e. les quantités) et leurs unités de mesure. Supposons que l'on veuille ajouter la relation n-aire *O2Permeability_Relation* de la figure 1 à notre ontologie, *MapOpt Ontology*, dédiée à la représentation de relations n-aires entre des données expérimentales quantitatives dans le domaine du risque alimentaire microbiologique étendu aux emballages (Touhami et al., 2011). L'existence d'interdépendances entre la relation et ses six arguments nécessite d'effectuer un nombre important de changements dans l'ontologie pour y parvenir (e.g. ajout de concepts, de propriétés, de restrictions). En outre, ces

1. <http://www.w3.org/TR/swbp-n-aryRelations/> "Use Case 3 : N-ary relation with no distinguished participant"

Évolution d'une ontologie pour relations n-aires

changements peuvent engendrer des incohérences dans l'ontologie. Nous présentons ici la représentation formelle des changements applicables à notre ontologie permettant de modifier sa structure tout en maintenant sa cohérence.

Le processus global de gestion de l'évolution de notre ontologie ainsi que de ses artefacts dépendants est celui de Stojanovic (2004), et nous présentons dans cet article l'étape de la représentation du changement de l'ontologie et la sémantique du changement. Dans notre approche une relation n-aire constitue un ensemble d'interdépendances (entre la relation et ses arguments, entre une quantité et les unités de mesures associées) à représenter et à traiter conjointement dans le processus d'évolution de l'ontologie.

Dans la section 2, nous présentons la modélisation de l'ontologie dédiée à la représentation de relations n-aires et la notion d'ontologie structurellement cohérente. Les changements élémentaires et composés permettant de faire évoluer l'ontologie sont détaillés dans la section 3. Enfin, nous concluons et présentons les perspectives de ce travail dans la section 4.

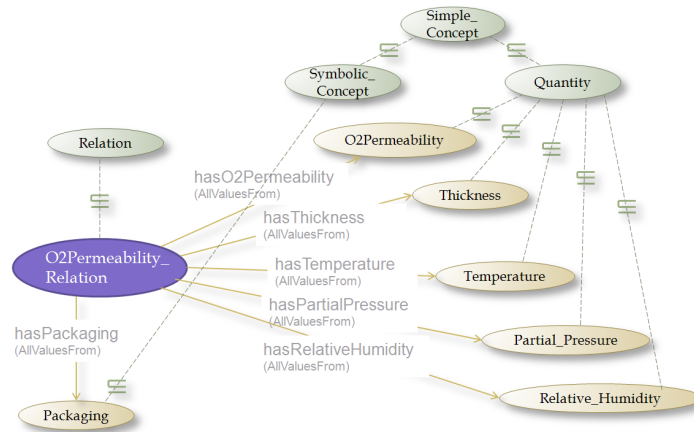


FIG. 1 – Exemple d'une relation n-aire dans le domaine de la microbiologie.

2 L'ontologie dédiée à la représentation de relations n-aires

D'après Reymonet et al. (2006), la modélisation d'une ontologie est fortement influencée par la tâche, le domaine d'intérêt et l'application pour laquelle elle a été conçue. Pour nous, la tâche à réaliser est l'annotation de relations n-aires dans des documents pour l'interrogation ; le domaine d'intérêt est l'étude de données expérimentales quantitatives des sciences du vivant ; l'application, enfin, est la construction d'un entrepôt de données ouvert sur le Web (Buche et al., 2012).

Définition 1 Pour un ensemble de concepts C , sa hiérarchie de spécialisation, notée H_C , est une relation acyclique, $H_C \subseteq C \times C$, telle que : si $(c_1, c_2) \in H_C$ alors c_1 est un sous-concept (un fils) de c_2 , c_2 est un superconcept (ou père) de c_1 , H_C^* est la fermeture réflexive, antisymétrique et transitive de H_C . $(c_1, c_2) \in H_C$ est aussi noté c_1 is-a c_2 ou $c_1 \sqsubseteq c_2$.

Définition 2 Soit C un ensemble de concepts et quatre concepts de C , $Simple_Concept$, $Quantity$, $Symbolic_Concept$ et $Relation$ tel que $Symbolic_Concept \sqsubseteq Simple_Concept$ et $Quantity \sqsubseteq Simple_Concept$. Nous définissons les ensembles des concepts suivants :

$$\begin{aligned} C_{arg} &= \{c \in C \mid (c, Simple_Concept) \in H_C^* \setminus \{Quantity, Symbolic_Concept\}, \\ C_{rel} &= \{c \in C \mid (c, Relation) \in H_C^* \setminus \{Relation\}, \\ C_{quantity} &= \{c \in C \mid (c, Quantity) \in H_C^* \setminus \{Quantity\}. \end{aligned}$$

Définition 3 Une ontologie dédiée à la représentation de relations n -aires est définie par le tuple suivant : $MO := (C, H_C, P, H_P, I, \mathcal{F})$ avec :

- C : l'ensemble des concepts de l'ontologie (cf. définition 2) et H_C sa hiérarchie (cf. définition 1) ;
- P : l'ensemble des propriétés telles que définie par le W3C², où une propriété est définie comme une relation binaire entre des instances de concepts, des littéraux et des types de données primitifs, et H_P sa hiérarchie telle que donnée dans la définition 1 où la notion de "concept" est remplacée par celle de "propriété" ;
- I : l'ensemble des instances de concepts ;
- \mathcal{F} : un ensemble de fonctions retournant un sous-ensemble d'instances ou de concepts de l'ontologie.

Nous présentons ci-dessous plusieurs fonctions de \mathcal{F} , utilisées dans la suite du papier :

- la fonction *signature* de C_{rel} dans $2^{C_{arg}}$ qui permet d'associer à une relation $r \in C_{rel}$ le domaine de valeurs de chacun de ses arguments (c_1, \dots, c_n) où $c_i \in C_{arg}$;
- la fonction *dimension* de $C_{quantity}$ dans $I_{Dimension}$ ³ qui permet d'associer à une quantité $q \in C_{quantity}$ sa dimension $d \in I_{Dimension}$;
- la fonction *linkconcepts* de $C_{rel} \times P \times \mathcal{R}$, où $\mathcal{R} = \{allValuesFrom, someValuesFrom, hasValue\}$, dans C_{arg} qui permet d'associer à un concept $c_1 \in C_{rel}$ un autre concept $c_2 \in C_{arg}$ via la propriété $p \in P$ à l'aide d'une restriction de propriétés⁴.

Pour la représentation du changement nous introduisons ci-dessous la notion d'ontologie \mathcal{CC} -cohérente, en nous inspirant de Stojanovic (2004) pour adapter ses contraintes de cohérences à la structure particulière de notre ontologie.

Définition 4 Une ontologie, définie par le tuple $(C, H_C, P, H_P, I, \mathcal{F})$ cf. à la définition 3, est dite \mathcal{CC} -cohérente si elle vérifie un ensemble de contraintes de cohérences structurelles, noté \mathcal{CC} .

Nous avons identifié 39 contraintes de cohérences structurelles dont 16 portent sur les concepts, 7 sur les propriétés et 12 sur les instances. Nous présentons ci-dessous quelques exemples de ces contraintes :

- $\mathcal{CC}1$: une relation a au moins deux arguments : $\forall r \in C_{rel}, |signature(r)| \geq 2$
- $\mathcal{CC}2$: Tous les concepts appartenant à la signature d'une relation doivent être définis : $\forall r \in C_{rel}, \forall c \in signature(r) \implies c \in C_{arg}$
- $\mathcal{CC}3$: Chaque quantité doit être associée à une seule valeur de dimension : $\forall q \in C_{quantity} \implies |dimension(q)| = 1$

2. <http://www.w3.org/TR/2004/REC-owl-guide-20040210/#SimpleProperties>

3. Les dimensions sont définies comme des instances du concept $Dimension \in C$, utilisé pour les unités de mesure.

4. <http://www.w3.org/TR/2004/REC-owl-guide-20040210/#PropertyRestrictions>

3 Évolution de l'ontologie dédiée à la représentation de relations n-aires

L'évolution de la structure de l'ontologie telle que donnée dans la définition 4, se traduit par divers changements tels que l'ajout d'une relation n-aire, l'ajout ou la suppression d'un de ses arguments, etc. Pour des raisons de simplicité et de lisibilité, nous présentons l'évolution de la structure de notre ontologie à travers un cas d'utilisation, détaillé dans la sous-section 3.1. Puis, nous présentons dans les sous-sections 3.2 et 3.3 les changements élémentaires et composés qui permettent une telle évolution.

3.1 Cas d'utilisation : ajouter une relation n-aire à *MapOpt Ontology*

Supposons que la relation n-aire *O2Permeability_Relation* présentée dans la figure 1 ne soit pas définie dans *MapOpt Ontology*. L'ontologue doit commencer par ajouter ses six arguments : le nouveau concept *Packaging* comme sous-concept du concept *Symbolic_Concept* et les nouveaux concepts *Partial_Pressure*, *Relative_Humidity*, *Thickness* et *O2Permeability* comme sous-concepts du concept *Quantity*. L'ontologue doit ensuite définir le concept *O2Permeability_Relation* comme sous-concept du concept *Relation*, ainsi que les propriétés et les restrictions de propriétés permettant de relier la relation à ses arguments.

L'ontologue doit donc, pour ajouter une nouvelle relation n-aire dans l'ontologie, effectuer un nombre important d'opérations de modification. Il doit également veiller à ce que les modifications effectuées respectent les contraintes de cohérences présentées dans la section 2. Afin d'alléger la tâche de l'ontologue et d'éviter des erreurs lors de manipulations fastidieuses de l'ontologie, nous définissons un ensemble de changements élémentaires et composés permettant de faire évoluer l'ontologie tout en préservant sa *CC*-cohérence.

3.2 Notion de changement élémentaire

Nous appelons changement élémentaire, une opération de modification de l'ontologie portant sur une seule entité : un concept, une propriété ou une instance. Pour faire évoluer l'ontologie, nous avons identifié 60 changements élémentaires dont 8 s'appliquent aux concepts (e.g. *CreateHierarchyConceptLink*), 41 aux propriétés (e.g. *CreateConceptsLinked*) et 5 aux instances de concepts (e.g. *CreateInstance*). Comme dans Stojanovic (2004), la formalisation des changements repose sur la définition d'un ensemble de pré-conditions (un ensemble d'assertions qui doivent être vraies pour pouvoir appliquer le changement) et de post-conditions (qui correspondent au(x) résultat(s) du changement) pour chaque type de changement, élémentaire ou composé. Les pré-conditions associées à un changement correspondent au sous-ensemble de contraintes de cohérences *CC* portant sur les parties de l'ontologie affectées par le changement considéré. La vérification des pré-conditions d'un changement permet d'éviter d'appliquer ce changement sur une ontologie non *CC*-cohérente. Les post-conditions garantissent que le changement a bien été appliqué, mais pas la *CC*-cohérence de l'ontologie.

Nous présentons ci-dessous l'exemple du changement élémentaire nécessaire pour faire évoluer *MapOpt Ontology* dans le cas d'utilisation décrit dans la section 3.1.

Définition 5 *CreateLinkConcepts* ($c_1, p, c_2, allValuesFrom$) :=

Sémantique : associer un concept c_2 au concept c_1 via une propriété p en utilisant la restriction $allValuesFrom$

Pré-condition : $c_1 \in C_{rel}, c_2 \in C_{arg}, p \in P, \{c \mid c \in C_{arg} \wedge linkconcepts(c_1, p, allValuesFrom) = c\} = \emptyset$

Post-condition : $c_1 \in C_{rel}, c_2 \in C_{arg}, p \in P, linkconcepts(c_1, p, allValuesFrom) = c_2$.

Ce changement est considéré comme élémentaire étant donné qu'il ne permet de manipuler qu'une seule entité : le concept c_1 . Comme pré-condition, il s'agit de vérifier : i) que les concepts c_1 et c_2 et la propriété p sont bien définies et ii) qu'aucun argument n'est déjà associé directement au concept c_1 via la propriété p à l'aide de la restriction $allValuesFrom$.

3.3 Changements composés

Un changement composé est une opération de modification de l'ontologie qui affecte plusieurs entités et peut se décomposer en une succession de changements élémentaires. Pour faire évoluer l'ontologie, nous avons identifié 12 changements composés (e.g. AddConcept, AddHierarchyOfSymbolicConcept). Nous présentons ci-dessous le changement composé nécessaire pour faire évoluer *MapOpt Ontology* dans le cas d'utilisation décrit dans la section 3.1.

Définition 6 *AddConcept*($newc, c, d, \{u_1, \dots, u_{n_u}\}, min, max, \{p_1, \dots, p_n\}, \{o_1, \dots, o_n\}, \{rs_1, \dots, rs_n\}$)

Pré-condition : $newc \notin C, (newc, c) \notin H_C,$

$[(c \in C \mid (c, Relation) \in H_C^*) \wedge (|\{o_1, \dots, o_n\}| \geq 2)] \vee (c \in C \mid (c, Simple_Concept) \in H_C^*),$

$u_1, \dots, u_{n_u} \in I_{Unit_Concept}$ ⁵, $p_1, \dots, p_n \in P, o_1, \dots, o_n \in C_{arg}, rs_1, \dots, rs_n \in \mathcal{R}$

Sémantique : ajouter un concept $newc \sqsubseteq c$, qui peut être une relation ($(c, Relation) \in H_C^*$) ou un concept simple ($(c, Simple_Concept) \in H_C^*$) avec sa hiérarchie, et ses restrictions.

Traitement du concept :

- Créer le concept $newc$
- Créer un lien hiérarchique de subsomption entre $newc$ et c : $CreateHierarchyConceptLink(newc, c)$

Traitement des restrictions :

1. $CreateLinkDimension(newc, hasDimension, d, hasValue)$
2. $\forall i \in [1, n_u], CreateLinkUnit(newc, hasUnitConcept, u_i, allValuesFrom)$
3. $CreateRestrictionMinMax(newc, hasNumericalValue, min, max)$
4. $\forall i \in [1, n], CreateLinkConcepts(newc, p_i, o_i, rs_i)$.

Post-condition : $newc \in C, (newc, c) \in H_C.$

Le changement composé **AddConcept** permet de créer soit une nouvelle relation, soit un nouveau concept, sous-concept de *Simple_Concept*, (i.e. une quantité, un concept symbolique ou un concept simple), qui sera utilisé comme argument d'une relation. Pour tous ces concepts, il y a un traitement en commun : créer le nouveau concept et le lier avec son concept père. Ensuite, il n'y a aucun traitement supplémentaire pour les concepts simples et les concepts symboliques. Pour les quantités, il s'agit de leurs associer leur dimension (traitement des restrictions 1), leurs unités de mesures (traitement des restrictions 2) et éventuellement un intervalle

5. Les unités de mesure sont définies comme des instances du concept $Unit_Concept \in C$

de valeurs (traitement des restrictions 3). Enfin pour les relations, il s'agit de créer les liens avec leurs différents arguments ($\{o_1, \dots, o_n\}$) selon les types de restriction ($\{rs_1, \dots, rs_n\}$) demandés (traitement des restrictions 4).

4 Conclusion

Nous avons présenté dans cet article les définitions des changements élémentaires et composés permettant de faire évoluer la structure d'une ontologie dédiée à la représentation de relations n-aires tout en maintenant sa cohérence structurelle. Nous nous sommes en particulier intéressés aux relations n-aires entre des données quantitatives expérimentales. Pour des questions de place, nous avons focalisés notre présentation sur les changements nécessaires pour ajouter une nouvelle relation n-aire à notre ontologie.

Dans un avenir proche, nous continuerons d'étudier l'évolution d'une ontologie dédiée à la représentation de relations n-aires en proposant des stratégies d'évolution associées aux changements. Ces stratégies permettent de prendre en considération les différents choix que peut faire un ontologue. Nous étudierons, aussi, l'évolution des unités de mesures qui sont liées aux quantités. La difficulté dans la gestion de l'évolution des unités de mesure s'explique par le fait qu'elles sont elles-même décomposables, ce qui va nous amener à devoir traiter des petits blocs d'ontologies qui sont très fortement interconnectés.

Références

- Buche, P., J. Dibie-Barthélemy, L. Ibănescu, et L. Soler (2012). Fuzzy Web Data Tables Integration Guided by an Ontological and Terminological Resource. *IEEE TKDE 99*(PrePrints).
- Reymonet, A., N. Aussenac-Gilles, et J. Thomas (2006). Tâche, domaine et application : influences sur le processus de modélisation de connaissances. In *Actes d'IC*, pp. 71–80.
- Stojanovic, L. (2004). *Methods and Tools for Ontology Evolution*. Ph. D. thesis, University of Karlsruhe, Germany.
- Touhami, R., P. Buche, J. Dibie-Barthélemy, et L. Ibănescu (2011). An Ontological and Terminological Resource for n-ary Relation Annotation in Web Data Tables. In *OTM 2011 (2)*, Volume 7045 of *LNCS*, pp. 662–679. Springer.

Summary

In this paper we are addressing the evolution's problem of an Ontological and Terminological Resource (OTR) allowing to represent n-ary relations. We present the formal definitions of changes to be applied to our OTR, modifying its structure while maintaining its structural consistence. Our examples are using an OTR representing n-ary relations between quantitative experimental data.