

Une nouvelle approche de découverte des correspondances complexes entre ontologies

Fatma Kaâbi*,* Faiez Gargouri*,**

*Laboratoire MIRACL, Institut Supérieur d'Informatique et du Multimédia de Sfax
BP 242 - 3021, Sakiet Ezzit, Sfax, Tunisie

* kaabifatma@yahoo.fr

** faeiz.gargouri@fsegs.rnu.t

Résumé. Les correspondances complexes ont été étudiées à plusieurs reprises dans le domaine d'alignement de schémas de bases de données. Par contre, dans le domaine d'alignement des ontologies, elles ont été peu étudiées. Nous proposons, dans ce papier, une nouvelle approche de découverte de correspondances complexes entre deux ontologies. L'approche proposée est extensionnelle, terminologique et implicative. Dans cette approche, nous utilisons le modèle des règles d'association afin de découvrir des correspondances de type $x \Rightarrow y_1 \wedge \dots \wedge y_n$ entre deux ontologies.

1 Introduction

Les correspondances complexes ont été peu étudiées dans le domaine des ontologies. Un des seuls travaux allant dans ce sens a été celui de A. Doan [Doan et al. (2003)], qui s'est restreint d'une part, aux disjonctions de concepts, et d'autre part, à la relation d'équivalence.

Dans ce papier, nous proposons une approche de découverte de correspondances complexes entre ontologies. Cette approche a l'originalité d'être à la fois *extensionnelle* (la découverte de correspondances complexes s'appuie principalement sur le contenu indexé aux hiérarchies (instances, textes, etc.)), *terminologique* (la comparaison de deux entités se fait à partir d'une sélection de termes extraits du contenu textuel) et *implicative* (découverte des relations d'implication entre entité). Dans cette approche, nous inspirons du paradigme des règles d'association, afin de découvrir des correspondances de type $x \Rightarrow y_1 \wedge \dots \wedge y_n$.

2 Alignement des correspondances complexes

Notre méthode prend en entrée deux ontologies supportant l'héritage multiple et possédant une structure hiérarchique. Ces ontologies ont subi une phase de prétraitement qui permet de les préparer en redéfinissant leurs relations d'association sur une extension commune [David et al. (2007)]. Nous considérons aussi un alignement pré-calculé entre ces deux ontologies. L'alignement considéré est fonctionnel et contient des relations d'alignement de type implication (\Rightarrow ou \Leftarrow) et/ou équivalence (\Leftrightarrow). Une ontologie est une hiérarchie de concepts (ou entités) représentée par le quintuplet $\mathcal{H} = (C, \leq, \mathcal{A}_n, I, \sigma)$ où C désigne l'ensemble des concepts,

\leq est la relation de subsomption entre les concepts, I dénote l'ensemble des instances des concepts et σ est la relation associant à chaque concept, son ensemble d'instances. La fonction \mathcal{A}_n associe à chaque entité ses annotations telles qu'un identifiant, un ensemble de labels et des commentaires. Le prétraitement des ontologies consiste à, pour chaque entité à lui réassocier un ensemble d'objets contenant les termes extraits à partir de ses annotations propres et celles de ses instances [David et al. (2007)]. Cette hiérarchie pourra être alors réécrite par : $\mathcal{H}' = (C, \leq, \sigma', T)$. T représente les ensembles des termes et des données associés aux entités. Chaque entité sera associée à un sous-ensemble de T par la relation d'association σ' .

À partir de deux ontologies $\mathcal{H}'_1 = (C_1, \leq_1, \sigma'_1, T_1)$ et $\mathcal{H}'_2 = (C_2, \leq_2, \sigma'_2, T_2)$, la deuxième phase de notre approche permet d'extraire l'ensemble des correspondances complexes de la forme $x \Rightarrow y_1 \wedge \dots \wedge y_n$ ou $x \in C_1$ et $y_i \in C_2$ ($i \in [1..n]$) avec n est le nombre de prédécesseurs de x). Le processus de découverte des correspondances complexe repose sur l'extraction et l'évaluation des règles d'association entre entités. Pour cela, nous adaptons le principe suivant : pour chaque entité $x \in C_1$ qui possède plusieurs prédécesseurs x'_i dans \mathcal{H}_1 , on cherche pour chaque x'_i de C_1 l'entité y_i de C_2 qui la correspond mieux en se référant à l'alignement pré-calculé (l'entité y_i qui est en relation d'alignement de type équivalence ou implication avec x'_i et qui maximise la valeur d'intensité d'implication [Blanchard et al. (2003)] de la règle d'association $x'_i \rightarrow y_i$ ($\varphi(x'_i \rightarrow y_i)$)). Ensuite, la règle $x \rightarrow y_1 \wedge \dots \wedge y_n$ est évaluée. Cette règle est retenue si elle a une valeur d'intensité d'implication $\varphi(x \rightarrow y_1 \wedge \dots \wedge y_n)$ supérieure à un seuil de sélection φ_r .

3 Conclusion

Dans ce papier, nous avons présenté une nouvelle approche de découverte de correspondances complexes entre deux ontologies. L'approche proposée est extensionnelle, terminologique et implicative.

Références

- Blanchard, J., F. Kuntz, P. and Guillet, et R. Gras (2003). Implication intensity : from the basic statistical definition to the entropic version. *CRC Press chapitre 28*, 473–485.
- David, J., F. Guillet, et H. Briand (2007). Association rule ontology matching approach. *International Journal on Semantic Web and Information Systems* 2, 27–49.
- Doan, A. and Madhavan, J., R. Dhamankar, P. Domingos, et A. Halevy (2003). Learning to match ontologies on the semantic web. *The VLDB Journal* 12 4, 303–319.

Summary

Complex matching were very few studied for ontology's matching. This paper presents an hybrid, extensional and asymmetric matching approach designed to find out complex matching between ontologies. In this approach, we use the association rule to discover complexes relation between two ontologies.