

Une nouvelle méthode d'alignement et de visualisation d'ontologies OWL-Lite

Sami Zghal^{*,**}, Karim Kamoun^{*}, Sadok Ben Yahia^{*}, Engelbert Mephu Nguifo^{**}

^{*}Département des Sciences de l'Informatique, Faculté de Sciences de Tunis, Tunisie
sadok.benyahia@fst.rnu.tn

^{**}CRIL CNRS FRE 2499, Université d'Artois, IUT de Lens, France
{sami.zghal, mephu}@cril.univ-artois.fr

Résumé. Dans ce papier, une nouvelle plate-forme d'alignement et de visualisation des ontologies, appelée POVA¹ (Prototype OWL-Lite Visual Alignment), est décrite. Le module d'alignement implémente une nouvelle approche d'alignement d'ontologies remédiant au problème de la circularité et de l'intervention de l'utilisateur.

Une seule ontologie ne suffit plus pour effectuer toutes les tâches envisageables dans un environnement distribué. Les techniques d'alignement peuvent fournir un cadre dans lequel plusieurs ontologies peuvent être exploitées. Aligner deux ontologies consiste à comparer les différences ou les ressemblances définies dans celles-ci.

La nouvelle méthode d'alignement proposée est intégrée dans un prototype d'alignement et de visualisation d'ontologies OWL-Lite, appelé **POVA** (Prototype OWL-Lite Visual Alignment). Ce prototype est constitué des trois modules : module de construction du graphe OWL-Graph, module d'alignement d'ontologies et module de visualisation. Le premier module, **BUILD OWL GRAPH**, permet la construction d'une nouvelle représentation, appelée OWL-Graph pour représenter l'ontologie décrite en OWL-Lite. Le graphe ainsi construit permet de décrire toutes les informations existantes dans une ontologie OWL-Lite. Le deuxième module, **EDOLA** (Extended Diameter OWL-Lite Alignment), implémente le nouveau algorithme automatique d'alignement. À chaque couple d'entités appartenant à une même catégorie, l'algorithme d'alignement calcule les mesures de similarité locale. Il définit un modèle global de calcul de similarité globale à travers le voisinage, tout en remédiant au problème de la circularité et de l'intervention de l'utilisateur dans le processus d'alignement. Le troisième module, **OWL-Lite VISUALIZATION**, permet de visualiser les deux ontologies à aligner, ainsi que le résultat de l'alignement produit par le deuxième module.

Dans le cadre des expérimentations menées pour évaluer la méthode d'alignement d'ontologies EDOLA, quelques tests fournis dans la base benchmark mise à la disposition de la communauté par la compétition EON (Evaluation of Ontology-based Tools), EON (2004)², sont utilisés. L'ontologie de base est constituée par un ensemble de références bibliographiques. Elle représente une version plus allégée en nombre d'entités ontologiques comparativement à des ontologies réelles. L'ontologie de base est composée en tout de 97 entités réparties comme

¹Ce travail est partiellement financé par le projet franco-tunisien CMCU 05G1412.

²<http://oaei.ontologymatching.org/2004/Contest/>

N. Test	Similarité	<i>Précision</i>		<i>Rappel</i>		<i>Fallout</i>	
		EDOLA	OLA	EDOLA	OLA	EDOLA	OLA
101	Levenshtein	1,000	0,590	1,000	0,970	0,000	0,410
103	Levenshtein	0,989	0,550	0,989	0,901	0,011	0,450
205	WORDNET	0,505	0,490	0,505	0,802	0,495	0,510
222	WORDNET	0,927	0,550	0,967	0,901	0,073	0,450
225	WORDNET	0,969	0,590	0,969	0,967	0,031	0,410
301	WORDNET	0,643	0,493	0,770	0,607	0,357	0,507
304	WORDNET	0,627	0,439	0,710	0,618	0,373	0,561

TAB. 1 – Comparaison entre EDOLA et OLA.

suit : 33 classes, 40 propriétés de nature type de données et 24 propriétés de nature objet. Afin d'évaluer la qualité des alignements obtenus, trois mesures sont retenues : *Précision*, *Rappel* et *Fallout*, Do et al. (2002). Ces mesures sont utilisées dans EON. L'objectif de ces mesures est d'évaluer la qualité des alignements et l'automatisation du processus de comparaison des alignements produits.

Afin d'évaluer les résultats de l'approche d'alignement proposée, la table 1 présente les résultats de deux algorithmes EDOLA et OLA Euzenat et al. (2004) sur des données test de la littérature. Ces résultats montrent un avantage de EDOLA sur OLA sur la majorité des ensembles de tests utilisés. Cet avantage s'explique par deux aspects : le premier relève de la propagation de la similarité qui est meilleure dans le cas de EDOLA, et le second découle du fait que EDOLA, contrairement à OLA, ne se base pas sur la stabilité de la mesure de similarité en utilisant un seuil d'arrêt ϵ défini par l'utilisateur. La valeur par défaut de ce seuil est fixé à 0.01 dans OLA.

Références

- Do, H., S. Melnik, et E. Rahm (2002). Comparison of schema matching evaluations. In *Proceedings of the 2nd Int. Workshop on Web Databases*, Erfurt. German Informatics Society.
- EON, G. (2004). EON ontology alignment contest. In *Proceedings of the 3rd Evaluation of Ontology-based Tools (EON) Workshop*.
- Euzenat, J., D. Loup, M. Touzani, et P. Valtchev (2004). Ontology alignment with OLA. In *Proceedings of the 3rd International Workshop : Semantic Web Conference EON*, Hiroshima, Japan, pp. 341–371.

Summary

In this paper, we introduce a prototype for both visualizing and aligning OWL-Lite ontologies. Its main originality is the integrated alignment module that resolves the problems of circularity and user-parameter settings.