

RDBToOnto : un logiciel dédié à l'apprentissage d'ontologies à partir de bases de données relationnelles

Farid Cerbah*

*Dassault Aviation
Département des études scientifiques
farid.cerbah@dassault-aviation.fr

Résumé. RDBToOnto¹ est un logiciel extensible qui permet d'élaborer des ontologies précises à partir de bases de données relationnelles. Le processus supporté est largement automatisé, de l'extraction des données à la génération du modèle de l'ontologie et son instanciation. Pour affiner le résultat, le processus peut être orienté par des contraintes locales définies interactivement. C'est aussi un cadre facilitant la mise en oeuvre de nouvelles méthodes d'apprentissage.

Bien que les bases de données relationnelles présentent un intérêt évident pour l'apprentissage d'ontologies, les outils développés à ce jour pour exploiter ces sources de données sont d'ambition limitée (par exemple, l'outil DataMaster est restreint à l'import de tables de données dans une ontologie générale du modèle relationnel). RDBToOnto comble en partie cette lacune en offrant la possibilité de dériver à partir d'une base de données source une ontologie précise dont la structure peut s'éloigner sensiblement du schéma de la base source².

Un des principes de RDBToOnto est de permettre une automatisation complète du processus. Il suffit de fournir l'url de la base de données pour obtenir rapidement une ontologie (instanciée). Cependant, l'ontologie produite peut être affinée en attachant interactivement des contraintes sur les tables de la base source (pour, par ex., ajuster les hiérarchies identifiées automatiquement ou définir des patrons de nommage des instances). Le paramétrage du processus et la définition des contraintes se font à travers une interface finalisée. Un autre aspect facilitateur de cette plateforme est l'intégration d'extracteurs pour différents formats de base de données. De plus, un composant de normalisation permet d'améliorer la base de données avant l'apprentissage de l'ontologie. L'outil est accompagné d'une documentation et peut être étendu de différentes manières (nouveaux extracteurs, autres méthodes d'apprentissage, ...).

Summary

RDBToOnto is a tool that allows to automatically generate ontologies from relational databases. A prominent feature of this tool is the ability to produce highly structured ontologies by exploiting structuring patterns hidden in the data. Though automated to a large extent, the process can be constrained in many ways through a friendly user interface.

¹www.tao-project.eu/researchanddevelopment/demosanddownloads/RDBToOnto.html

²cf. dans ces actes, l'article du même auteur intitulé « Fouille de données dans les bases relationnelles pour l'acquisition d'ontologies riches en hiérarchies de classes ».