



Alinhamento entre Ontologias de Topo e Ontologias de Domínio via WordNet

Eduardo Cardoso Ferreira, Daniela Schmidt, Dra. Renata Vieira (orientadora)

Escola Politécnica, Curso de Engenharia de Software, PUCRS;

Laboratório de Processamento de Linguagem Natural (PLN), PUCRS

Resumo

Ontologias são amplamente utilizadas na computação como uma forma para representar o conhecimento através de conceitos, relações, instâncias e axiomas. Ontologias podem ser classificadas de acordo com seu nível de abstração. Ontologias de Domínio descrevem conceitos de um domínio específico e Ontologias de Topo descrevem conceitos universais que são comuns a qualquer domínio. Existem diversas ontologias de topo, dentre elas, a: DOLCE (*Descriptive Ontology for Linguistic and Cognitive Engineering*) e a SUMO (*Suggested Upper Merged Ontology*) que utilizamos em nosso trabalho. O processo de encontrar correspondências entre ontologias é chamado *Ontology Matching* e tem como saída um conjunto de correspondência (um alinhamento). Alinhar ontologias de domínio é uma área de pesquisa bastante explorada nos últimos anos, no entanto, alinhar ontologias de diferentes níveis de abstração, como domínio e topo, ainda está em fases iniciais de pesquisa, sendo que, na maioria das vezes, essa tarefa é desenvolvida manualmente por especialistas. Nesse trabalho, propomos uma abordagem para alinhar ontologias de domínio e topo de forma automática. Para realizar esse alinhamento, utilizamos o WordNet (WN) que é uma base de dados lexical composto por substantivos, verbos adjetivos e advérbios, agrupados em conjuntos de sinônimos (*synsets*). O WN possui alinhamentos entre *synsets* e conceitos de ontologias de topo. Nós exploramos esse recurso para identificar as correspondências entre conceitos de domínio e topo. Para isso, dividimos a abordagem

em duas etapas: (i) identificação do *synset* no WordNet que representa o sentido mais adequado para o conceito de domínio; (ii) após, percorre-se os conceitos da ontologia de topo para identificar o conceito que corresponde ao *synset* do conceito de domínio. Com relação a (i), que é o foco deste trabalho, avaliamos técnicas de similaridade baseadas no WN que visam auxiliar na identificação do *synset* mais adequado para cada conceito. Durante esse período, foram implementados algoritmos para técnicas como *Lin*, *Resnik*, *Lesk*, entre outras. Para os experimentos, utilizamos 7 ontologias do domínio de conferência que são amplamente utilizadas na literatura e as ontologias de topo citadas. Como resultado, um protótipo implementando as técnicas citadas será disponibilizado para a comunidade acadêmica. O protótipo foi avaliado comparando-o com ferramentas estado da arte e obteve resultados superiores em termos de medida F.

Palavras-chave

Ontologia; Topo; Domínio; WordNet; Alinhamento;