

PROTOTIPAGEM DE CONSTRUÇÃO DE ONTOLOGIA PARA SUPORTAR CONSULTAS EM UM SUBDOMÍNIO DE NUTRIÇÃO

TULIO, Lucas Serrano¹ (lucas.s.tulio@gmail.com); **BATISTA JUNIOR, Joinville**² (joinville@ufgd.edu.br);

¹ Discente do curso de Sistemas de Informação; PIVIC/UFGD;

² Docente do curso de Sistemas de Informação; UFGD;

Ontologias interligam conceitos de um determinado domínio através de relações, a fim de representar informações de maneira estruturada. Tais relações são extraídas de fontes de conhecimento do domínio de interesse, compostas, geralmente, de textos em linguagem natural. As relações que compõem uma ontologia podem ser de especialização (representando hierarquias entre os conceitos), associação (expressando propriedades e características aos conceitos), agregação (partes que foram um todo) e instanciação (elemento pertencente a determinada classe de conceito). Atualmente, extratores de informação ainda não apresentam um bom desempenho em relação à precisão e cobertura das relações extraídas, ou seja, extraem relações mal construídas e omitem a extração de várias relações que deveriam ter sido extraídas. O objetivo deste trabalho é o desenvolvimento de ferramentas de software para suportar: (a) a extração automática de relações; e (b) a realização e visualização de consultas a partir de uma ontologia construída com base nas relações extraídas. A metodologia utilizada incorpora: (1) definição da fonte de informação para extração automática de relações; (2) definição do tipo de questões a serem suportadas e dos padrões de relações a serem extraídos; (3) implementação e teste de uma ferramenta para extração automática das relações de interesse; (4) construção de uma pequena ontologia a partir das relações extraídas; e (5) implementação e teste de outra ferramenta para suportar as consultas a partir das relações da ontologia. A fonte escolhida para extração automática de relações foram os textos explicativos sobre conceitos definidos na WordNet. Para esta pesquisa, foram escolhidos cinco conceitos principais do domínio de interesse: *symptom*, *disease*, *cause*, *nutrient*, *food*. Tais conceitos apresentam uma sequência frequentemente encontrada nas fontes de informação do domínio de nutrição: determinado sintoma é característico de uma doença, que é causada pela falta de um nutriente, o qual pode ser encontrado em algum alimento. Foi implementada uma ferramenta para extrair relações de especialização e de associação da WordNet, para o nutriente *vitamin*. A partir dos textos explicativos da WordNet, e do conceito descrito pelo texto, foram montadas sentenças, considerando as várias formas de representação que a WordNet utiliza para textos explicativos. Alguns conectivos (“that”, “or”, “and”), e verbos no particípio passado após substantivos, foram utilizados para quebrar as sentenças em sentenças mais simples, viabilizando a extração automática de relações de associação. A partir das relações extraídas automaticamente foi construída uma pequena ontologia. Também foi implementada uma ferramenta para visualizar consultas a partir dos encadeamentos das relações da ontologia, com foco nos cinco conceitos de interesse. Essa ferramenta permite visualizar consultas baseadas no encadeamento dos cinco conceitos, ou baseada em encadeamentos parciais. A conclusão deste trabalho é a viabilidade da automação da construção de ontologias, para suportar consultas baseadas em encadeamento de conceitos, a partir de textos explicativos da WordNet.

Palavra-chave: Automação de Ontologias. Extração Automática de Relações. Consultas a partir de Ontologias.