

T-Lex: Thesaurus com Estruturação Semântica e Operações Gerativas

Marco Gonzalez

PUCRS - Faculdade de Informática
Av.Ipiranga, 6681 – Prédio 16, térreo
90619-900 Porto Alegre, Brazil
gonzalez@inf.pucrs.br

Vera Lúcia Strube de Lima

PUCRS - Faculdade de Informática
Av.Ipiranga, 6681 – Prédio 16, térreo
90619-900 Porto Alegre, Brazil
vera@inf.pucrs.br

ABSTRACT

T-Lex is a thesaurus manually constructed from a traditional dictionary, adopting aspects from the Generative Lexicon Theory (GLT) proposed by James Pustejovsky, as well as object-oriented software modeling concepts. The Qualia structure of GLT is adapted for describing nouns, verbs and adjectives on a differentiated way, respectively, as objects, operations, and states. Generative operations which return sets of related words are included in the T-Lex prototype implementation. So, this thesaurus is a helpful tool in of automatic query expansion for information retrieval in Portuguese language.

Keywords: natural language processing, artificial intelligence, thesaurus, data mining, generative operations

RESUMO

T-Lex é um thesaurus construído manualmente a partir de dicionário tradicional, adotando aspectos da Teoria do Léxico Gerativo (TLG) de James Pustejovsky e conceitos da modelagem de software orientada a objetos. A estrutura Qualia da TLG é adaptada com o objetivo de descrever substantivos, verbos e adjetivos de forma diferenciada, respectivamente, como objetos, operações e estados. Operações gerativas que visam a geração de campos lexicais são incluídas na implementação do protótipo do T-Lex e, dessa forma, o thesaurus é utilizado com bons resultados na expansão automática de consulta em sistema de recuperação de informação em língua portuguesa.

Palavras-chave: processamento de linguagem natural, inteligência artificial, thesaurus, base de dados, operações gerativas