

## Interação em Linguagem Natural para Acesso a Banco de Dados

Para possibilitar o tratamento de perguntas feitas em linguagem natural para acesso a informações de bases de dados, este trabalho descreve a construção de um sistema de processamento de texto desenvolvido no âmbito do projeto Comunica. O foco do projeto é em disponibilizar informações sobre transferências constitucionais (vindas da FAMURS). O sistema recebe uma pergunta e identifica as informações a serem buscadas na base de dados através de uma análise sintática automática das palavras e em informações buscadas em ontologia. Os conceitos relevantes identificados são usados para selecionar padrões de perguntas previamente definidos, baseado em medida de semelhança e um limiar. Caso nenhum padrão tenha sido satisfeito por falta de informações o sistema pergunta ao usuário os conceitos faltantes. É feito então o acesso a informação em bases de dados e retornada a resposta à pergunta do usuário.

Uma das ontologias utilizadas foi gerada a partir de dados extraídos de forma semi-automática do dicionário de Língua Portuguesa da Editora Porto, que contém relações entre palavras como sinonímia e hiperonímia. Para a representação dos dados na linguagem OWL foi necessária a conversão dessas relações. Para hiperonímia foi criada uma ontologia automaticamente. Porém as relações de sinonímia não puderam ser convertidas por causa da polissemia das palavras. Assim foi criado um sistema que informa quando há conflito e permite que um linguista realize correções (usando como base dicionários da língua portuguesa).

O sistema também tem um procedimento de reconhecimento de datas, que podem estar escritas por extenso, com dígitos ou definirem períodos, como por exemplo, semana e bimestre. Desta forma o sistema é capaz de identificar quais as informações o usuário deseja saber e se estas estão, ou não, no escopo do sistema. Um sistema destes pode ser usado para facilitar acesso a informações por qualquer pessoa através de telefone.