

302

USO DE INFORMAÇÕES SEMÂNTICAS PARA RESOLUÇÃO DE EXPRESSÕES REFERENCIAIS DA LÍNGUA PORTUGUESA. *Jorge César Barboza Coelho, Renata Vieira (orient.)* (UNISINOS).

Este trabalho apresenta uma proposta de resolução de expressões referenciais da Língua Portuguesa (LP) utilizando informações semânticas. Uma avaliação dessa proposta também é apresentada. Expressões referenciais são unidades lingüísticas com a propriedade de representar entidades do discurso (e.g., animais, lugares, fatos). Essas expressões podem estabelecer relações entre si, denominadas anáforas. O mapeamento de tais relações é importante para diversas tarefas de Processamento da Linguagem Natural (PLN), e.g., Recuperação de informação, Sumarização. Em estudos anteriores, observamos que, no processo de tratamento de anáforas, expressões referenciais anafóricas associativas (e.g., a casa – a porta) e anafóricas indiretas (e.g., a casa – a moradia) apresentam maiores dificuldades de tratamento do que as demais. Isso porque essas duas classes estabelecem relações de ordem semântica entre expressões. A identificação de relações semânticas é uma tarefa difícil do ponto de vista computacional. Recentemente, o analisador sintático da LP, o PALAVRAS, incorporou marcações semânticas no resultado da análise das sentenças. Essas marcações semânticas são do tipo traço semântico, que descrevem o significado por meio de dados superordenados; e.g., na sentença "Os cursos ensinam o japonês e modo de vida", a expressão "os cursos" apresenta a etiqueta semântica <speak-work> e "o japonês" <language>. Assim, uma solução para identificação das relações de ordem semântica seria considerar a anotação semântica no processo de tratamento das expressões referenciais. Nossa proposta visa comparar traços semânticos de expressões referenciais com os de expressões antecedentes, traços semânticos equivalentes indicariam relações anafóricas associativas e indiretas. Este trabalho faz parte do Projeto TeXto – Acesso a informações em bases textuais.