

RESOLUÇÃO DE DESCRIÇÕES DEFINIDAS EM TEXTOS DA LÍNGUA PORTUGUESA. *Daniela Rossi, Renata Vieira* (Informática – Centro de Ciências Exatas e Tecnológicas – Unisinos).

O objetivo deste trabalho é classificar automaticamente descrições definidas, encontradas em textos da língua Portuguesa. Descrições definidas são sintagmas nominais iniciados por artigos definidos (exemplo: a árvore, o governador). Elas podem ser classificadas como: novas no discurso (quando as descrições definidas aparecem pela primeira vez no discurso, introduzindo um novo referente), anáforas diretas (possuem idêntico nome-núcleo e referem-se a um mesmo objeto), anáforas indiretas (não possuem o mesmo nome-núcleo, mas mantêm relação de sinonímia com este, referindo-se ao mesmo objeto, evento ou entidade no mundo) ou associativas (introduzem um elemento novo no discurso que é dependente de um elemento anteriormente mencionado). Em um primeiro momento, as descrições definidas e os possíveis antecedentes foram extraídas dos textos. Posteriormente, efetuou-se a classificação manual que serve de base para a classificação automática e o tratamento computacional de correferência nominal. O sistema desenvolvido compara os núcleos das descrições definidas com possíveis antecedentes. Quando uma expressão antecedente com o mesmo núcleo é encontrada, a descrição definida é classificada como anáfora direta. Pré-modificação e pós-modificação também são analisadas. Quando não existe um antecedente e existe uma pós-modificação com preposição, ou quando o núcleo for um nome próprio (presença de letra maiúscula), a descrição definida é classificada como nova no discurso. O sistema em questão foi desenvolvido linguagem PROLOG. Em uma próxima etapa, serão abordadas as descrições associativas. Após efetuada a classificação, o resultado final será comparado com os textos anotados manualmente e também com os resultados obtidos pela resolução automática realizada com a língua Inglesa. (CNPq).