

012

USO DE RELAÇÕES SEMÂNTICAS NA CLASSIFICAÇÃO DE ANÁFORAS. *Vinicius Magnus Müller, Renata Vieira (orient.)* (UNISINOS).

Em Sistemas de Processamento da Linguagem Natural (PLN), a análise de expressões anafóricas é um componente fundamental na interpretação do sentido do texto. A resolução de anáforas é importante para várias tarefas no Processamento de Linguagem Natural, tais como Extração de Informação, Sumarização de Textos, Tradução Automática, Sistemas de Respostas a Perguntas, entre outras. A resolução de anáforas busca identificar as relações entre expressões evocadas no discurso quando estas se referem à mesma entidade. Desta forma, o completo entendimento de uma expressão anafórica, depende de expressões anteriores encontradas no discurso, o que chamamos de antecedentes. Existem três tipos de anáfora: -Direta: a expressão anafórica e seu antecedente possuem o mesmo nome núcleo. Ex. (Uma casa, a casa). -Indireta: a expressão anafórica possui uma relação de identidade com seu antecedente, mas contém um nome núcleo diferente. Ex. (Uma casa, a moradia). -Associativa: a expressão possui conexões semânticas com seu antecedente, sendo que este é necessário para a sua interpretação. Ex (Uma casa, a porta). Em um trabalho anterior apresentamos uma abordagem para resolução de anáforas para a Língua Portuguesa baseada nas relações semânticas existentes entre as entidades envolvidas, tendo como foco as anáforas nominais definidas, sintagmas nominais com um artigo definido (o, a, os, as). Este trabalho propõe-se a estender o trabalho previamente desenvolvido, adicionando-lhe capacidade de classificação das expressões segundo os seus tipos anafóricos. O processo de classificação utilizará o resultado da resolução e classificará as anáforas encontradas através de técnicas de mineração de texto e aprendizado de máquina. (PIBIC).