

Capítulo 6

Conclusões

“O maior objetivo de toda a ciência é cobrir o maior número de fatos empíricos por dedução lógica a partir do menor número possível de hipóteses ou axiomas”

Albert Einstein (1879-1955): Lincoln Barnett *The Universe and Dr Einstein* (1950 ed.).

6.1 Introdução

Há vários sistemas de Processamento de Linguagem Natural (PLN) que usam a noção do modelamento de papéis temáticos (por exemplo, McClelland e Kawamoto, 1986; McClelland *et al.*, 1989; St. John e McClelland, 1990; Jain, 1991 e Miikkulainen, 1996). O presente sistema se diferencia de todos estes na forma que faz uso do conhecimento teórico a partir da Lingüística. Várias decisões relacionadas ao modelo computacional e sua implementação são baseadas no papel dos acarretamentos semânticos nas relações temáticas.

O sistema HTRP – *Hybrid Thematic Role Processor* – Processador de Papel Temático Híbrido (Rosa e Françoze, 1999) – implementa uma abordagem híbrida simbólico-conexionista ao PLN. Nesta abordagem as vantagens do sistema simbólico (facilidade de representação do conhecimento, entendimento através da inferência lógica, etc.) são combinadas com as vantagens do connexionismo (aprendizado, generalização, tolerância a falhas, etc.). E a representação distribuída de características semânticas adotada neste sistema, permite também que novas palavras sejam acrescentadas, dado que seus vetores de microcaracterísticas semânticas sejam fornecidos.

Duas versões do sistema HTRP são implementadas: uma versão *com* “conhecimento” inicial, isto é, conhecimento simbólico a cerca das relações temáticas é introduzido na rede por meio de pesos de conexão, e uma versão *sem* “conhecimento” inicial, sendo que neste caso, valores aleatórios são atribuídos aos pesos de conexão da rede. A primeira versão é chamada de BIW – *biased initial weight version* – e a segunda de RIW – *random initial weight version*. Nas duas versões um único conjunto de redes dá conta de todas as sentenças; portanto, não apenas substantivos mas também verbos podem ser generalizados. De fato, isto é crucial no tratamento de papéis temáticos, pois eles são a generalização das relações semânticas entre verbos e substantivos no presente sistema. Um outro resultado interessante que deve ser enfatizado diz respeito à RIW. Mesmo em um sistema sem conhecimento inicial, as regras finais extraídas da rede têm alto grau de coerência com o conhecimento lingüístico. Isto é, parece que a arquitetura do HTRP juntamente com o treinamento é suficiente para o sistema chegar à grade semântica correta de uma sentença.

6.2 *Conclusões e Trabalhos Futuros*

O HTRP, com sua abordagem adaptativa, pode explicar como a sintaxe e a semântica, como comportamentos humanos categoriais, estabelecem uma base para um sistema conexionista distribuído como a versão RIW. Na versão BIW, o mesmo sistema pode ser empregado para várias línguas, com diferentes efeitos de polarização (isto é, diferentes regras simbólicas iniciais). Pode-se dizer que este sistema pode dar conta da teoria lingüística de princípios e parâmetros (os princípios são fixos, eles correspondem à arquitetura; os parâmetros podem variar, eles correspondem às regras iniciais). Em resumo, acredita-se que este sistema contribui para a pesquisa em PLN, pois oferece uma alternativa aos sistemas puramente simbólicos.

O sistema HTRP utiliza a noção de acarretamento semântico para o processamento de papéis temáticos (de Franchi e Cançado, 1998). As relações temáticas são representadas através de acarretamentos. E, além dos acarretamentos, os efeitos composicionais resultantes dos itens lexicais são igualmente importantes. Aqui torna-se extremamente clara a importância de uma abordagem conexionista, pois neste tipo de abordagem, há uma relação composicional intensa entre todos os elementos representados na rede.

Cumprir observar que, na verdade, as microcaracterísticas (ou traços) utilizadas para o verbo no sistema HTRP são representações das funções de acarretamento, ou seja, o HTRP tenta dar conta das funções por meio dos traços semânticos. Uma versão futura, e muito mais complexa, poderá acrescentar mais níveis de processamento em forma de camadas na arquitetura conexionista, para assim permitir que suas entradas possam ser caracterizadas como resultantes das funções de acarretamento – dessa forma, é até possível prever um sistema que aprenda as propriedades semânticas das palavras.

As considerações que levam a uma opção não-lexicalista dos papéis temáticos têm como finalidade a construção de processadores computacionais para a língua portuguesa, a partir de representações distribuídas para as palavras. O sistema híbrido simbólico-conexionista HTRP é apresentado como uma aplicação bem sucedida para este tipo de repre-

sentação, pois une as vantagens da abordagem lógica com as da conexionista. As duas versões do HTRP permitem, com mais precisão, a verificação das teorias lingüísticas no sistema. Por exemplo, a arquitetura do HTRP pode ser aproximada a teorias lingüísticas que postulam conhecimentos inatos como base para a competência sintática nos seres humanos. Desta forma, o sistema HTRP permite modelar diferentes hipóteses lingüísticas.