

Putting Pieces Together: Combining FrameNet, VerbNet and WordNet for Robust Semantic Parsing

Lei Shi and Rada Mihalcea

University of North Texas
Computer Science Department
Denton, TX, 76203-1366
leishi@unt.edu, rada@cs.unt.edu

Abstract. This paper describes our work in integrating three different lexical resources: FrameNet, VerbNet, and WordNet, into a unified, richer knowledge-base, to the end of enabling more robust semantic parsing. The construction of each of these lexical resources has required many years of laborious human effort, and they all have their strengths and shortcomings. By linking them together, we build an improved resource in which (1) the coverage of FrameNet is extended, (2) the VerbNet lexicon is augmented with frame semantics, and (3) selectional restrictions are implemented using WordNet semantic classes. The synergistic exploitation of various lexical resources is crucial for many complex language processing applications, and we prove it once again effective in building a robust semantic parser.

1 Introduction

The goal of a semantic parser is to identify semantic relations between words in a text, resulting in structures that reflect various levels of semantic interpretation. Such structures can be used to improve the quality of natural language processing applications by taking into account the meaning of the text. Automatic techniques for semantic parsing have been successfully used in Information Extraction and Question Answering, and are currently evaluated in other applications such as Machine Translation and Text Summarization.

The process of semantic parsing typically implies a learning stage, where the semantic structures to be identified are acquired from an existing lexical resource, which explicitly identifies the range of possible semantic relations between words in a text. While there are several lexical resources suitable for semantic parsing, built with extensive human effort over years of work – including FrameNet [3], VerbNet [5], WordNet [7], or PropBank [4] – all previous approaches to semantic parsing have relied exclusively on only one of them, as there are no connections between these resources that would enable their exploitation in a unified way. However, each resource encodes a different kind of knowledge and has its own advantages, and thus combining them together can eventually result in a