

Integrating Semantic Frames from Multiple Sources

Namhee Kwon and Eduard Hovy

Information Sciences Institute
University of Southern California
Marina del Rey, CA 90292, USA
{[nkwon](mailto:nkwon@isi.edu),[hovy](mailto:hovy@isi.edu)}@isi.edu

Abstract. Semantic resources of predicate-argument structure have high potential to enable increased quality in language understanding. Several alternative frame collections exist, but they cover different sets of predicates and use different role sets. We integrate semantic frame information given a predicate verb using three available collections: FrameNet, PropBank, and the LCS database. For each word sense in WordNet, we automatically assign the corresponding FrameNet frame and align frame roles between FrameNet and PropBank frames and between FrameNet and LCS frames, and verify the results manually. The results are available as part of ISI's Omega ontology.

1 Introduction

With more accurate semantic analysis, systems should obtain higher performance in many applications such as machine translation, question answering, and summarization. Thanks to the release of annotated corpora with semantic argument structures and manually constructed lexical-semantic information such as FrameNet [1], PropBank [10], LCS database [3], and VerbNet [11], many models inducing semantic frames have been developed ([7], [6], [13], [17], [18]).

Such data collections cover different sets of predicates. Unfortunately, no collection covers all (or most) of the (English) predicates, and the roles and other definitional aspects of the collections differ. Due to these differences, most approaches to semantic analysis using these available resources (semantic role tagging) are specific to only one of these resources and their results are not comparable and usable over other resources.

We believe that we can build a broader and consistent semantic resource by integrating all semantic frame information from these disparate collections. The value of the integrated resource is apparent at many levels: first, as a theoretical device to highlight differences and generate further refinements in lexical semantic theory; second, as a practical resource that can be used by semantic analysis and other applications; third, as a testbed for an automatic aligning method between different resources that can also be applied to more general integration of lexical information. As more such annotated collections are created