Instructions for ACL-2017 Proceedings

Anonymous ACL submission

Abstract

This document contains the instructions for preparing a camera-ready manuscript for the proceedings of ACL-2017. The document itself conforms to its own specifications, and is therefore an example of what your manuscript should look like. These instructions should be used for both papers submitted for review and for final versions of accepted papers. Authors are asked to conform to all the directions reported in this document.

In this section, we describe our efforts to create a new, large-scale linguistics resource: a semantically annotated grammar with wide-coverage of English, accompanied by a semantically annotated WSJ corpus. This research seeks to build upon the grammar created for the S-STRUCT experiments in Section **??**, which consisted of 30 unlexicalized XTAG trees with hand-annotated semantics.

While this grammar was sufficient for the scaling experiments, it still suffered from a few problems. First, it still does not have enough coverage of English; only the 30 most frequently used XTAG trees have semantics, whereas there are over 1000 trees available in the XTAG corpus. Secondly, the semantics were still ad-hoc in the sense that different predicates have no similarity measure; either they are the same or they are different. For example, $run(x)$ and $chase(x, y)$ are different predicates, but both express some form of motion. In the previously discussed semantics, these are treated as entirely dissimilar, but ideally we would have some level of abstraction allowing us to express their similarity. Finally, the current semantic system cannot represent higher-order semantic concepts such as adverbs, which modify verbs (predicates).

In this section, we attempt to improve upon all of these shortcomings. We explore the alternative semantic representations of Davidsonian and Neo-Davidsonian semantics, which reify verbal predicates with an explicit "event" entity. We leverage the existing linguistics resources of VerbNet and PropBank to expand our semantic coverage of declarative verb trees in the XTAG grammar. Finally, we use ideas from grammar metarules and metagrammars to increase our semantic coverage of XTAG verb trees by an order of magnitude. The end result of this process is a higher-quality semantically annotated grammar with vastly wider coverage. We then show the utility of this grammar by using it to create a semantically annotated version of the WSJ corpus.

1 Background

Before discussing our linguistic resource, we must review some linguistics background. In this section, we consider alternative semantic representations to those used in Section **??**. We then explore existing linguistics resources, VerbNet and PropBank, to see if we can leverage existing work. Finally we investigate automated ways of expanding grammar coverage through metagrammars and metarules.

1.1 Davidsonian and Neo-Davidsonian Semantics

In Section **??**, we introduced a first-order logical semantic representation, augmented with λ-calculus notation to allow for compositionality. In this representation, variables represent entities in the world (with either existential or universal quantification) and predicates represent the relations between entities in the world. We will refer to these semantics as "Classical Semantics", to maintain convention with (**?**). If we wanted