Semantic Parsing with Combinatory Categorial Grammars

Yoav Artzi, Nicholas FitzGerald and **Luke Zettlemoyer**
Computer Science & Engineering
University of Washington
Seattle, WA 98195
{yoav,nfitz,lsz}@cs.washington.edu

1 Abstract

Semantic parsers map natural language sentences to formal representations of their underlying meaning. Building accurate semantic parsers without prohibitive engineering costs is a long-standing, open research problem.

The tutorial will describe general principles for building semantic parsers. The presentation will be divided into two main parts: modeling and learning. The modeling section will include best practices for grammar design and choice of semantic representation. The discussion will be guided by examples from several domains. To illustrate the choices to be made and show how they can be approached within a real-life representation language, we will use λ-calculus meaning representations. In the learning part, we will describe a unified approach for learning Combinatory Categorial Grammar (CCG) semantic parsers, that induces both a CCG lexicon and the parameters of a parsing model. The approach learns from data with labeled meaning representations, as well as from more easily gathered weak supervision. It also enables *grounded* learning where the semantic parser is used in an interactive environment, for example to read and execute instructions.

The ideas we will discuss are widely applicable. The semantic modeling approach, while implemented in λ-calculus, could be applied to many other formal languages. Similarly, the algorithms for inducing CCGs focus on tasks that are formalism independent, learning the meaning of words and estimating parsing parameters. No prior knowledge of CCGs is required. The tutorial will be backed by implementation and experiments in the University of Washington Semantic Parsing Framework (UW SPF).[1]

[1] http://yoavartzi.com/spf

2 Outline

1. Introduction to CCGs
2. Modeling
   (a) Questions for database queries
   (b) Plurality and determiner resolution in grounded applications
   (c) Event semantics and imperatives in instructional language
3. Learning
   (a) A unified learning algorithm
   (b) Learning with supervised data
      i. Lexical induction with templates
      ii. Unification-based learning
   (c) Weakly supervised learning without labeled meaning representations

3 Instructors

Yoav Artzi is a Ph.D. candidate in the Computer Science & Engineering department at the University of Washington. His research studies the acquisition of grounded natural language understanding within interactive systems. His work focuses on modeling semantic representations and designing weakly supervised learning algorithms. He is a recipient of the 2012 Yahoo KSC award.

Nicholas FitzGerald is a Ph.D. student at the University of Washington. His research interests are grounded natural language understanding and generation. He is a recipient of an Intel Science and Technology Center Fellowship and an NSERC Postgraduate Scholarship.

Luke Zettlemoyer is an Assistant Professor in the Computer Science & Engineering department at the University of Washington. His research interests are in the intersections of natural language processing, machine learning and decision making under uncertainty. Honors include best paper awards at UAI 2005 and ACL 2009, selection to the DARPA CSSG, and an NSF CAREER Award.

Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics, page 2,
Sofia, Bulgaria, August 4-9 2013.