MPhil in Advanced Computer Science
Syntax and Semantics of Natural Language

Leaders: Prof. Ted Briscoe and Dr. Stephen Clark
Timing: Lent
Prerequisites: Introduction to NLP core module or Equivalent Background
Structure: 16 Lectures (2 per week for 8 weeks)

AIMS

This module provides an introduction to the formal syntax and semantics of natural language, in particular Montague semantics using a Categorial Grammar model of syntax. The first part of the module will focus on the theory of syntax, followed by an example of how recent advances in parsing technology allow such a theory to be implemented in practice, operating on naturally occurring text. The second part of the module focuses on the compositional semantics of sentences, and the final part will briefly consider lexical semantics at the word level, using a radically different distributional approach. The final lecture will speculate as to how a new combined theory of natural language semantics could be formulated from the compositional and distributional approaches.

SYLLABUS

1. Introduction to (Combinatory) Categorial Grammar (1L)

2. English syntax in the CCG framework (3Ls)

3. Introduction to Statistical Parsing (1L)

4. Constructing a wide-coverage CCG English grammar (1L)

5. Wide-coverage robust statistical parsing with CCG (2Ls)

6. Introduction to natural language semantics (1L)

7. Type-driven models of semantics: (Combinatory) Categorial Grammar (1L)

8. Montague semantics; intensional semantics (3L)

9. Robust, wide-coverage implementation (1L)

10. Distributional models of lexical semantics; proposals for a compositional distributional model (2L)

All lectures will be given by Prof. Briscoe or Dr. Clark.

OBJECTIVES

On completion of this module students should:

- understand how the syntax of natural language sentences can be modelled using a type-driven (Combinatory) Categorial Grammar;

- understand how a wide-coverage grammar of English can be constructed

- have studied one approach to statistical parsing in detail

- understand how the meaning of natural language sentences can be modelled using a logical, model-theoretic approach;

- understand how the meaning of natural language sentences can be constructed using Frege's principle of compositionality;

- gain an appreciation of how syntactic and semantic theory can be implemented in practice;

- understand how the meanings of words can be represented using a distributional model;

- appreciate the current divergence between the compostional and distributional models.

PRACTICAL WORK

There will be no practical work associated with this module.

ASSESSMENT

- Four ticked take-home tests at end of weeks 1,2,5,7. Questions set and marked by Prof. Briscoe and Dr. Clark. Each test is worth 5% of the final mark. Tests will be available on each Friday and due in the following Friday.

- One final take-home exam covering all the material. Final take-home exam will contribute 80% to the final mark. Questions set and marked by Prof. Briscoe and Dr. Clark.

RECOMMENDED READING

Formal Semantics. Ronnie Cann. Cambridge University Press. 1993

Combinatory Categorial Grammar, Mark Steedman (with Jason Baldridge), to appear in *Non-transformational Syntax* (eds. Borsley, R and Borjars, K.) Available at: http://www.iccs.informatics.ed.ac.uk/ steedman/papers.html

Wide-Coverage Efficient Statistical Parsing with CCG and Log-Linear Models, Stephen Clark and James R. Curran, *Computational Linguistics*, 33(4), pp.493-552, 2007. Available at: http://www.cl.cam.ac.uk/ sc609/pubs.html

Wide-Coverage Semantic Representations from a CCG Parser. Johan Bos, Stephen Clark, Mark Steedman, James R. Curran and Julia Hockenmaier. Proceedings of COLING-04, pp.1240-1246, Geneva, Switzerland, 2004. Available at: http://www.cl.cam.ac.uk/ sc609/pubs.html

Combining Symbolic and Distributional Models of Meaning. Stephen Clark and Stephen Pulman. Proceedings of QI-07, pp.52-55, Stanford, CA, 2007. Available at: http://www.cl.cam.ac.uk/ sc609/pubs.html

Last updated: October 2009