

Introduction to Natural Language Syntax and Parsing: L95

Lecture 1: Introduction

Ann Copestake (standing in for Simone Teufel)

Department of Computer Science and Technology
University of Cambridge

October 2019

Outline of today's lecture

NLP and linguistics

Course structure

Examples of analysis

To do list

Analysis of text

- ▶ Analysis of text is the core of NLP applications.
- ▶ Level and type of analysis attempted depends on application:
 - ▶ Depth of analysis? Meaning?
 - ▶ Explicit linguistic analysis vs surface-based techniques?
 - ▶ Precision? e.g., testing for grammaticality vs processing noisy data.
- ▶ Also work on language generation and regeneration.

NLP and linguistics

The computational modelling of human language
(NLP/computational linguistics/CL)

1. **Morphology** — the structure of words.
2. **Syntax** — the way words are used to form phrases.
3. **Semantics**
 - ▶ **Compositional semantics** — the construction of meaning based on syntax.
 - ▶ **Lexical semantics** — the meaning of individual words.
4. **Pragmatics** — meaning in context.

NLP courses in the ACS

- ▶ L90: Overview of NLP
- ▶ L95
- ▶ L101: Machine learning for language processing
- ▶ R250: Advanced topics
- ▶ Also relevant material in background module: M20 — Data Science: principles and practice

Syllabus for L95

- ▶ Linguistics for NLP — morphology, syntax, semantics and pragmatics [6 lectures, Simone Teufel (and AAC)]
- ▶ Parsing - grammars, treebanks, representations and evaluation, statistical parse ranking [8 sessions, Paula Buttery]
- ▶ Interpretation — compositional semantics and entailment, pragmatic inference [2 lectures, Simone Teufel]
- ▶ Four take home assignments: first available today, hand in by October 21.
- ▶ One practical: start in week 6, hand in January 14.

Resources for L95

On website:

- ▶ slides (on website after lecture).
- ▶ two extensive handouts: please read before the lectures (do exercises and ask questions!)
- ▶ introductory logic worksheet
- ▶ assessed exercises (and associated resources)
- ▶ Possibly useful background: L90 lecture notes

Recommended Book: Jurafsky and Martin (2009)

See also suggestions at the end of the handouts.

Assessment for L95

- ▶ Four 'ticked' assignments: first due on October 21 (hand in to student admin). Each worth 5%.
- ▶ Practical report, due next term.

Examples of analysis

Chromosomal alterations were detected in humans exposed to butadiene.

- ▶ Morphological analysis and POS assignment
- ▶ Noun phrase bracketing
- ▶ Syntax tree
- ▶ Grammatical relations
- ▶ Logical form

To do

- ▶ Before the next lecture (Monday): read sections 1–4 of ‘Introduction to Linguistics’ handout and do the associated exercises.
- ▶ Start logic worksheet (let me know if you get stuck).
- ▶ Look at first assessed exercise.