

Introduction to Computational Linguistics

CFG and UBG Parsing Exercises

1 Left-Corner Relation

Describe an algorithm to compute the (transitive closure of the) left-corner relation for a given CFG, which is needed in Left-Corner Parsing.

2 Extraction of parse trees

Describe an algorithm that extracts the *complete* parse trees from the chart of the CYK algorithm using the \mathcal{C} and \mathcal{B} arrays and the context-free productions.

Treat the construction of trees from grammar symbols and smaller trees as black boxes.

3 Parse-tree extraction – run time

What is the worst-case complexity of your parse-tree extraction algorithm?

4 Bottom-up vs. Earley/Left-Corner parsing

In what situation is it advantageous to use pure bottom up parsing, resp. predictive bottom up (aka Earley or Left-Corner). Why?

5 Subsumption

Define (in pseudocode or implemented) a subsumption algorithm for (typed) feature structures. Assume the function `subsumes(type0, type1)` to be given.

6 Parsing

Modify a chart parsing algorithm to deal with unification-based grammars (with subsumption packing) and discuss points of inefficiency. Try to present a solution in pseudo-code. Assume the lexicon lookup, returning strings for input words, to be given.

Literature

- Jay Earley (1970), *An efficient context-free parsing algorithm*, Communications of the ACM, Volume 15, Number 2, pp 94–102
URL: search at <http://portal.acm.org/>
 - Klaas Sikkel: *Parsing Schemata: A Framework for Specification and Analysis of Parsing Algorithms*, Springer, 1997
Available in CoLi library, quite theoretical
 - Klaas Sikkel: *Parsing of context-free languages*
URL: <http://wwwhome.cs.utwente.nl/sikkel/papers/ps/amilp95.ps.gz> Covers parts of the abovementioned book.
 - Bob Carpenter. *The Logic of Typed Feature Structures*. Tracts in Theoretical Computer Science. Cambridge University Press, Cambridge, 1992.
 - Ann Copestake. *Implementing typed feature structure grammars*. CSLI Lecture Notes, December 2001.
 - Bernd Kiefer, Hans-Ulrich Krieger, John Carroll, and Rob Malouf. A bag of useful techniques for efficient and robust parsing. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics on Computational Linguistics*, pages 473–480, Morristown, NJ, USA, 1999. Association for Computational Linguistics.
 - Stuart M. Shieber. *An Introduction to Unification-Based Approaches to Grammar*. CSLI Lecture Notes, Number 4. Center for the Study of Language and Information, Stanford, 1986.
-