



Bibliography list

Morphology

The Morphology section of the course is based on Sproat (1992). Finite-state technology is introduced in great detail in Beesley and Karttunen (2003), and some of the algorithms, as well as applications, are described in Roche and Schabes (1997).

Implementations of FST include XFST (Beesley and Karttunen, 2003), FSA utilities (van Noord and Gerdemann, 2001), the FSM package (Mohri, Pereira, and Riley, 1998; Mohri, Pereira, and Riley, 2000) and Intex (Silberztein, 1993), among others.

Syntax

Context-free grammars are due to Chomsky (1956) (in more detail in Chomsky (1957)), and their mathematical properties are discussed in Aho and Ullman (1972), for example. The adequacy of CFGs for describing natural languages is discussed by Pullum and Gazdar (1982). Later, however, it was shown that some natural languages are not context-free (Higginbotham, 1984; Shieber, 1985; Manaster-Ramer, 1987).

Unification grammars are described informally by Shieber (1986), but better texts are Carpenter (1992) for the underlying logic or Sag and Wasow (1999) for a linguistic theory based on UG.

Empirical NLP

Two good sources are Manning and Schütze (1999) and Jurafsky and Martin (2000).

References

- Aho, Alfred V. and Jeffrey D. Ullman. 1972. *The Theory of Parsing, Translation and Compiling*, volume 1: Parsing. Prentice-Hall, Inc., Englewood Cliffs, N.J.
- Beesley, Kenneth R. and Lauri Karttunen. 2003. *Finite-State Morphology: Xerox Tools and Techniques*. CSLI, Stanford.



- Carpenter, Bob. 1992. *The Logic of Typed Feature Structures*. Cambridge Tracts in Theoretical Computer Science. Cambridge University Press.
- Chomsky, Noam. 1956. Three models for the description of language. In *I. R. E. transactions on information theory, Proceedings of the symposium on information theory*, volume IT-2, pages 113–123, September.
- Chomsky, Noam. 1957. *Syntactic Structures*. Mouton & Co., The Hague, The Netherlands.
- Higginbotham, James. 1984. English not a context-free language. *Linguistics Inquiry*, 15:225–234.
- Jurafsky, Daniel and James H. Martin. 2000. *Speech and Language Processing*. Prentice Hall Series in Artificial Intelligence. Prentice Hall, New Jersey.
- Manaster-Ramer, Alexis. 1987. Dutch as a formal language. *Linguistics and Philosophy*, 10:221–246.
- Manning, Christopher D. and Hinrich Schütze. 1999. *Foundations of statistical natural language processing*. The MIT Press, Cambridge, Mass.
- Mohri, Mehryar, Fernando Pereira, and Michael Riley. 1998. *A Rational Design for a Weighted Finite-State Transducer Library*. Number 1436 in Lecture Notes in Computer Science. Springer.
- Mohri, Mehryar, Fernando Pereira, and Michael Riley. 2000. The design principles of a weighted finite-state transducer library. *Theoretical Computer Science*, 231(1):17–32, January.
- Pullum, Geoffrey K. and Gerald Gazdar. 1982. Natural languages and context-free languages. *Linguistics and Philosophy*, 4:471–504.
- Roche, Emmanuel and Yves Schabes, editors. 1997. *Finite-State Language Processing*. Language, Speech and Communication. MIT Press, Cambridge, MA.
- Sag, Ivan A. and Thomas Wasow. 1999. *Syntactic Theory: A Formal Introduction*. CSLI, Stanford.
- Shieber, Stuart M. 1985. Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8:333–343.



Computational Linguistics Group
Department of Computer Science
University of Haifa

בלשנות חישובית
החוג למדעי המחשב
אוניברסיטת חיפה

Shieber, Stuart M. 1986. *An Introduction to Unification Based Approaches to Grammar*. Number 4 in CSLI Lecture Notes. CSLI.

Silberztein, Max. 1993. *Dictionnaires électroniques et analyse automatique de textes : le système INTEX*. Masson, Paris.

Sproat, R. W. 1992. *Morphology and Computation*. MIT Press, Cambridge, MA.

van Noord, Gertjan and Dale Gerdemann. 2001. Finite state transducers with predicates and identity. *Grammars*, 4(3).