

Parsing Beyond Context-Free Grammar

Laura Kallmeyer, Wolfgang Maier

Summer term 2009

1 General Information

Seminar: Wednesday, 8.30am, room 1.13
Webpage: <http://www.sfs.uni-tuebingen.de/emmy/parsing>
Credits: 3 CP (BA), 5 (MA)
Instructor: Laura Kallmeyer, Wolfgang Maier
Office hours: On appointment
Office: Hausserstr. 11, 2nd floor
Phone: 29-77513, 29-77474
Email: {lk,wmaier}@sfs.uni-tuebingen.de

Examination regulations of the Neuphilologische Fakultät require that students attend courses regularly. If students do not attend a course meeting on more than two occasions in one semester without proper excuse (e.g., doctor's note), the course instructor has to give them a failing grade.

You are expected to come on time. Being late without good reasons will count as not having attended a course meeting.

Your grade will be based on three components: The midterm exam (50%), the final exam (50%). In addition, to obtain the grade, you have to hand in 50% of the homework exercises.

2 Course Schedule

| | |
|-------|--|
| 29.4. | General Introduction: Grammar formalisms and parsing as deduction. Hopcroft and Ullman (1979); Shieber (1985); Shieber et al. (1995) |
| 6.5. | TAG 1: Introduction to the formalism, CYK parsing for TAG. Joshi et al. (1975); Joshi and Schabes (1997); Vijay-Shanker and Joshi (1985); Kallmeyer and Satta (2009) |
| 13.5. | TAG 2: Earley parsing Joshi and Schabes (1997); Schabes and Joshi (1988); Nederhof (1997) |
| 20.5. | TAG 3: LR parsing. Nederhof (1998); Prolo (2003) |
| 27.5. | MCFG/LCFRS 1: Introduction to the formalism, basic CYK parsing Seki et al. (1991); Vijay-Shanker et al. (1987); Weir (1988); Burden and Ljunglöf (2005) |
| 10.6. | Midterm exam |
| 17.6. | Discussion of midterm exam MCFG/LCFRS 2: CYK parsing with active items Burden and Ljunglöf (2005) |
| 24.6. | Simple RCG: Incremental Earley parsing Boullier (1998) |
| 1.7. | Thread Automata Villemonte de La Clergerie (2002) |
| 8.7. | RCG 1: Introduction to the formalism, top-down parsing. Boullier (1998, 2000) |
| 15.7. | RCG 2: CYK parsing, Earley-style parsing. Boullier (2000); Kallmeyer et al. (2009) |
| 22.7. | Exam |

References

- Boullier, P. (1998). A Proposal for a Natural Language Processing Syntactic Backbone. Technical Report 3342, INRIA.
- Boullier, P. (2000). Range Concatenation Grammars. In *Proceedings of the Sixth International Workshop on Parsing Technologies (IWPT2000)*, pages 53–64, Trento, Italy.
- Burden, H. and Ljunglöf, P. (2005). Parsing linear context-free rewriting systems. In *IWPT'05, 9th International Workshop on Parsing Technologies*, Vancouver, Canada.
- Hopcroft, J. E. and Ullman, J. D. (1979). *Introduction to automata theory, languages and computation*. Addison Wesley.
- Joshi, A. K., Levy, L. S., and Takahashi, M. (1975). Tree Adjunct Grammars. *Journal of Computer and System Science*, 10:136–163.
- Joshi, A. K. and Schabes, Y. (1997). Tree-Adjoining Grammars. In Rozenberg, G. and Salomaa, A., editors, *Handbook of Formal Languages*, pages 69–123. Springer, Berlin.
- Kallmeyer, L., Maier, W., and Parmentier, Y. (2009). Un algorithme d'analyse de type earley pour grammaires à concaténation d'intervalles. In *Actes de la 16ème conférence sur le Traitement Automatique des Langues Naturelles (TALN 2009)*, Senlis, France.
- Kallmeyer, L. and Satta, G. (2009). A Polynomial-Time Parsing Algorithm for TT-MCTAG. *Proceedings of ACL*.
- Nederhof, M.-J. (1997). Solving the correct-prefix property for TAGs. In Becker, T. and Krieger, H.-U., editors, *Proceedings of the Fifth Meeting on Mathematics of Language*, pages 124–130, Schloss Dagstuhl, Saarbrücken.
- Nederhof, M.-J. (1998). An alternative LR algorithm for TAGs. In *Proceedings of ACL*, Montreal, Canada.
- Prolo, C. (2003). *LR Parsing for Tree Adjoining Grammars and Its Application to Corpus-Based Natural Language Parsing*. PhD thesis, University of Pennsylvania.
- Savitch, W. J., Bach, E., Marxh, W., and Safran-Naveh, G., editors (1987). *The Formal Complexity of Natural Language*. Studies in Linguistics and Philosophy. Reidel, Dordrecht, Holland.
- Schabes, Y. and Joshi, A. K. (1988). An Earley-type parsing algorithm for Tree Adjoining Grammars. In *Proceedings of the 26th Annual Meeting of the Association for Computational Linguistics*, pages 258–269.
- Seki, H., Matsumura, T., Fujii, M., and Kasami, T. (1991). On multiple context-free grammars. *Theoretical Computer Science*, 88(2):191–229.
- Shieber, S. M. (1985). Evidence against the context-freeness of natural language. *Linguistics and Philosophy*, 8:333–343. Reprinted in Savitch et al. (1987).
- Shieber, S. M., Schabes, Y., and Pereira, F. C. N. (1995). Principles and implementation of deductive parsing. *Journal of Logic Programming*, 24(1&2):3–36.
- Vijay-Shanker, K. and Joshi, A. K. (1985). Some computational properties of Tree Adjoining Grammars. In *Proceedings of the 23rd Annual Meeting of the Association for Computational Linguistics*, pages 82–93.
- Vijay-Shanker, K., Weir, D. J., and Joshi, A. K. (1987). Characterizing structural descriptions produced by various grammatical formalisms. In *Proceedings of ACL*, Stanford.
- Villemonte de La Clergerie, E. (2002). Parsing mildly context-sensitive languages with thread automata. In *Proc. of COLING'02*.
- Weir, D. J. (1988). *Characterizing mildly context-sensitive grammar formalisms*. PhD thesis, University of Pennsylvania.