W. Detmar Meurers
OSU Department of Linguistics
Autumn 2002

LING 795K (11787-9)
Tue. & Thu. 10^{30}–12^{18}
029 Derby & 201 Oxley

Constraint-based grammar implementation

I. Course description

Background Constraint-based linguistic frameworks such as Head-Driven Phrase Structure Grammar make it possible to be explicit about the data structures and theories one proposes. It thereby also becomes feasible to develop these theories into grammar fragments, and several computational systems have been developed to support the implementation of such grammars, which generally focus on syntactic and semantic aspects.

From the linguistic perspective, the development of grammar fragments can be an important means of obtaining feedback on the empirical consequences of a linguistic theory and the compatibility of the various theories which are integrated in the grammar fragment. I would argue that one can go one step further by stating that comprehensive grammar fragments integrating the state-of-the-art of syntactic theorizing are essential for reestablishing the credibility of generative syntax as a science with a measurable criterion for progress.

From the computational perspective, in contrast to the heydays of deep, theory-driven syntactic analysis (where it was viewed as the only viable way to approach natural language) it is now generally recognized that the variety of language processing tasks can and probably should use a variety of approaches, which differ with respect to the kind of linguistic insight they incorporate. While this raises the question what kind of syntactic and semantic representations are needed for what kind of computational task, it seems clear that there is a need at least for syntactic dependency and high-level semantic and discourse representations.

This seminar This seminar is intended to gives students the possibility to get hands-on experience with a complex grammar implementation project. Different from previous seminars I've held, the focus of this seminar is not the theory behind grammar implementation systems, but the grammar implementation effort itself. As implementation environment, the seminar will use the TRALE system, developed as part of the Milca cooperation by Gerald Penn in Toronto, my group at OSU, and the CL group in Tübingen.

Based on input from potential participants, here are some potential grammar implementation projects that might be tackled:

- A grammar for detecting mistakes of a learner of Spanish
- Continuing work on the OSU version of the English Resource Grammar
- A German fragment using linearization domains
- Ginzburg and Sag (2001)
- Sag and Wasow (1999)

To accommodate those students who did not yet have a chance to take the Introduction to HPSG, the seminar will initially be split into two groups, with lectures on HPSG and grammar implementation used to bring the introductory group up to speed before taking on parts of Pollard and Sag (1994) as a smaller grammar implementation task.

Preparation to be done before seminar starts In preparation for this seminar, participants should (re-)read Chapters 1, 3 and 9 of Pollard and Sag (1994) as well as Meurers (1994). To run the grammar that comes with the paper, use the ALE system at /home/projects/milca/systems/ale/3.2.1/ale_256vars.pl.

II. Organization and expectations

- Course email list: 795k@ling.osu.edu
- Course web page: http://ling.osu.edu/~dm/2002/autumn/795K/
- Office hours: Tuesdays, 14^{00}–15^{30} in 201a Oxley

This is a seminar consisting of individual grammar writing projects. Joint work on a project is permitted (and actually encouraged), as long as each participant is more or less equally involved, active, and competent in the entire process. Each participant (or group of participants) is expected to take on the task of specifiying, implementing, documenting, and testing a grammar fragment. More concretely there are three parts:

a) Specify what is to be implemented and what the implementation is supposed to be able to do. (Weeks 1–3)

   This step will generally involve exploring the linguistic theory to be implemented, extracting and writing down the components of the linguistic theory and the signature they are based on explicitly and fully, and specifying a list of sentences exemplifying each of the constructions that are supposed to be covered.

   Every student is expected to complete this step during the first three weeks of the course, to turn in this specification (approx. 5 pages) before Monday, Oct. 21, and to give a 15 minute presentation on it during that week's session (Week 4). Since it makes no sense to continue without having a specification at this point, failure to hand in the specification leads to exclusion from the rest of the course.

b) Carry out, document and test the implementation (Weeks 4–9)

   This step will generally involve getting to know the Trale system and implementing the theory from the specification. The weekly Tuesday sessions are used to discuss general issues in the group, the Thursday sessions are used for instruction and discussion specific to each implementation.

c) Present the implementation result (Week 10)

   In the last two session, each grammar implementation is presented in a talk using overheads. The presentation will generally include the specification, discuss issues and problems that arose during implementation, and comment on the relation between linguistic theory, the specification and the implemented result. Contentwise the documentation (approx. 15 pages) will cover the

same material, but will generally include more specific comments on the actual grammar code.

Every student is expected to turn in the grammar code and the documentation at the end of the quarter, i.e., Dec. 13. There will be no incompletes.

III. Overview of course sessions

Week 0 (029 Derby)

- *Thu, 26. Sept*: Organizational meeting

Weeks 1 & 2 (029 Derby)

- *Tue, 1. Oct*: Discussion of projects to be tackled.
- *Thu, 3. Oct, Tue, 8. Oct, Thu, 10. Oct*
  - introductory group: Lectures on HPSG and grammar implementation
  - advanced group: Individual work on the project specification

Week 3 (029 Derby)

- *Tue, 15. Oct*: Individual work on the projectc specifications and meeting time to discuss them and their presentation
- *Thu, 17. Oct*: Lecture on the Trale system

Week 4 (029 Derby)

- *Tue, 22. Oct, Thu, 24. Oct*: Presentation of specifications

Weeks 5–9 (201 Oxley)

- *Tuesdays, 29. Oct, 5. Nov, 12. Nov, 19. Nov, 26. Nov*: General sessions discussing issues relevant for each of the different endeavors: signature, lexical and phrasal organization, implicational principles, the division of labor between theses components of a linguistic theory, the use and role of abbreviations, etc.
- *Thursdays, 31. Oct, 7. Nov, 14. Nov, 21. Nov*: Lab sessions for work on individual grammar projects and individual discussions

Week 10 (029 Derby)

- *Tue, 3. Dec, Thu, 5. Dec*: Presentation and discussion of grammar projects

References

Ginzburg, Jonathan and Ivan A. Sag (2001). *Interrogative Investigations: the form, meaning, and use of English Interrogatives*. Stanford, CA: CSLI Publications.

Meurers, Walt Detmar (1994). On Implementing an HPSG Theory – Aspects of the Logical Architecture, the Formalization, and the Implementation of Head-Driven Phrase Structure Grammars. In Erhard W. Hinrichs, Walt Detmar Meurers and Tsuneko Nakazawa (eds.), *Partial-VP and Split-NP Topicalization in German – An HPSG Analysis and its Implementation*, Tübingen: Universität Tübingen, no. 58 in Arbeitspapiere des SFB 340, pp. 47–155. http://ling.osu.edu/~dm/on-implementing.html.

Pollard, Carl and Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago, IL: University of Chicago Press.

Sag, Ivan A. and Thomas Wasow (1999). *Syntactic Theory: A Formal Introduction*. Stanford, CA: CSLI Publications.