

Natural Language Processing Techniques

NOTE: This is the traditional Prolog-based course, not the Python-based version that was tried in 2010.

Students who took the Python-based version and would also like to take this version of the course for separate credit can be accommodated; please inquire.

Instructor: Dr. Michael A. Covington, mc@uga.edu, www.ai.uga.edu/mc

Class meetings: Tues./Thurs., 2:00-3:15 and Wednesdays, 2:30-3:20, 110 Boyd GSRC (AI Conference Room)
NOT as scheduled in Baldwin Hall G1.

Course web page: <http://www.ai.uga.edu/mc/8570.html>

Content: This is a technical course in computer processing and understanding of human languages. It is designed primarily for students in the M.S. in AI program but is also open to other appropriately prepared students. Topics include morphological analysis, POS tagging, parsing, unification-based grammar, and formal semantics.

Prerequisites: **Symbolic Programming and Generative Syntax, or P.O.D.**

The Generative Syntax prerequisite is often waived, but Symbolic Programming (CSCI/ARTI 6540) is necessary, and a genuine interest and background in cognitive science are presumed. Note that:

- (1) You must be able to program in Prolog prior to taking this course. You can learn Prolog on your own, but please do it *before, not during* this course. You will need to be able to do the homework exercises in chapters 4–6 of Covington, Nute, and Vellino, *Prolog Programming in Depth*. Prolog has a programming methodology all its own; it is not just a new notation for familiar concepts.
- (2) As usual in 8000-level computer science courses, **mastery of programming in a conventional language is also presumed**, along with understanding of data structures and algorithms, file handling, knowledge of how your operating system installs and launches software, etc. (Suitable conventional languages include C#, Java, Python, and others.) **Linguists with no computer background may require several semesters of preparation, including self-study and lower-level CSCI courses.**
- (3) Those who have relatively little linguistics background should **read an introductory linguistics textbook**. (Hint: Past editions of Fromkin et al., *An Introduction to Language*, are inexpensive on the secondhand market. You do not have to have the current edition.)

Textbook:

REQUIRED:

Covington, Natural Language Processing for Prolog Programmers

Prentice-Hall, 1994, now available through printing-on-demand. ISBN 0-13-629213-5.
(This book is widely available secondhand around UGA.)

This book will be supplemented by a large amount of new material on handouts and web pages, especially covering newer corpus-based methods.

Requirements:

Attendance

Homework and class participation

Midterm examination

Final examination

Term project

Academic honesty:

The rules of the University apply to this course.

As a University of Georgia student, you have agreed to abide by the University's academic honesty policy, "A Culture of Honesty," and the Student Honor Code. All academic work must meet the standards described in "A Culture of Honesty" found at www.uga.edu/honesty. Lack of knowledge of the academic honesty policy is not a reasonable explanation for a violation. Questions related to course assignments and the academic honesty policy should be directed to the instructor.