

CSC 2501F and CSC 485F: Computational Linguistics
Course Information, Fall 2006
Department of Computer Science, University of Toronto

Instructor and Teaching Assistant

Instructor: Suzanne Stevenson

Contact Info: PT 283C (D.L. Pratt building), (416)978-6277, suzanne@cs.toronto.edu

Office hours: Wed, 2-3:30pm, or by appointment. (Check course website for exceptions to this regular time.)

Teaching assistant: Afsaneh Fazly

Contact Info: afsaneh@cs.toronto.edu

Course home page

<http://www.cs.toronto.edu/~suzanne/2501>

Most handouts for the class, including lecture notes, will appear on this page. You are responsible for checking there for class announcements as well.

Class meetings

The class will meet every Wednesday, 10am–Noon, in BA B025 (Bahen). Tutorials will be held in the same room, on Mondays, 11–12, **only on announced dates**. Tutorials cover complementary material to the class, and are not optional.

Recommended textbooks

This book will be available for 2-hour loan in the SF library, and drafts of new chapters are available at www.cs.colorado.edu/~martin/slp2.html:

Jurafsky, Daniel, and Martin, James H. *Speech and Language Processing*. Prentice-Hall, 2000.

Note: Graduate students in computational linguistics are expected to own the above book, but may prefer to wait for the 2nd edition, projected to be available next summer.

This book is on-line at www.uni-giessen.de/~g91062/Seminare/gk-cl/Allen95/a11995co.htm:

Allen, James. *Natural Language Understanding*, 2nd Edition. Benjamin/Cummings, 1995.

Computer accounts

Accounts on CDF will be allocated shortly and posted on the course website. NSERC-funded research machines (such as *dup*, *gew*, *allen*, and the Computer Science workstations) should not be used for coursework.

Evaluation

Marks in the course are based on: two programming projects (45% of your final mark); two take-home problem sets (40%); and five “write-ups” on assigned research papers and contribution to discussion on them (10%). In addition, 5% of your final mark will be determined by additional class participation.

Students whose main area of study is not computer science, cognitive science, or Linguistics and Computing may write a term paper instead of doing the programming projects. See the instructor for details.

Plagiarism

All assignments are to be done individually. If you have questions about the definition and harm of plagiarism, please see <http://www.cs.toronto.edu/~fpitt/plagiarism.html> and <http://www.cs.toronto.edu/~clarke/acoffences/>. If you have specific questions about what constitutes plagiarism in this course, **please see me for clarification**.

Syllabus

Week	Date	Topic	Reading	
1	13	Sept	Intro to computational linguistics.	OPT: JM 1
2	20	Sept	Grammars and parsing.	OPT: JM-N 12.1–12.3 REQ: JM-N 5.1, 11.1–11.3.3, 11.3.7
3	27	Sept	Chart parsing.	OPT: JM-N 12.4 REQ: Allen 3.4, 3.6
4	4	Oct	Parsing with features.	OPT: JM 11.1–11.2, 11.5; Allen 4.1–4.4 REQ: JM-N 11.3.4, 11.3.5; Allen 4.5
5	11	Oct	Extending grammars with features.	OPT: JM 11.3 REQ: JM-N 11.3.6
6	18	Oct	Ambiguity resolution.	OPT: JM 12.5
7	25	Oct	Statistical attachment disambiguation.	REQ: Research reading.
8	1	Nov	Statistical parsing.	OPT: JM 12.1 REQ: Research reading.
9	8	Nov	Advances in statistical parsing.	OPT: JM 12.2–12.4 REQ: Research reading.
10	15	Nov	Semantic representation and parsing.	OPT: JM 15.1–15.3; JM-N 16.5.3–16.5.6 REQ: JM-N 16.2, 16.4, 16.5.1, 16.5.2
11	22	Nov	Lexical semantics, emphasizing verbs.	OPT: JM 16.3, 17.1 REQ: Research reading.
12	29	Nov	Acquisition of verb semantic properties.	REQ: Research reading.
13	6	Dec	Statistical semantic role assignment.	REQ: Research reading.

“JM” refers to the 2000 edition of the Jurafsky and Martin text, “JM-N” to chapters from the new edition that are available on-line, and “Allen” to the Allen textbook (also on-line). Readings listed as OPT are recommended, and as REQ are required. Any supplemental material handed out in class or posted on the website is also required.

Assignments and Due Dates: [Finalized]

Weight	Assignment	Due time	Due date	Where to hand in
25%	Project 1	Noon	MON , 23 Oct	PT 283
2%	“Write-Up”	10:10am	Wed, 25 Oct	class
2%	“Write-Up”	10:10am	Wed, 1 Nov	class
20%	Problem Set 1	Noon	FRI , 3 Nov	PT 283
2%	“Write-Up”	10:10am	Wed, 8 Nov	class
20%	Project 2	Noon	MON , 27 Nov	PT 283
2%	“Write-Up”	10:10am	Wed, 29 Nov	class
2%	“Write-Up”	10:10am	Wed, 6 Dec	class
20%	Problem Set 2	Noon	FRI , 8 Dec	PT 283

Note: Projects and problem sets should be given to the secretary in the AI office (PT 283) by noon on the due date. One copy of the “write-ups” on research papers should be given to the instructor at the beginning of the class in which they are due; students should keep a copy for themselves for use during the discussion.

Late submissions will be penalized 10% per actual day (including weekend days), and will be accepted no later than **four** actual days late. You are responsible for arranging a time and location for submitting late work to the TA or instructor. The lateness penalty may be waived at the discretion of the instructor.

Emailing the instructor

Please email me only from a cs.toronto.edu, cdf.toronto.edu, or utoronto.ca address; other email may be caught in my spam filter. You must put “CSC 2501” (grad students) or “CSC 485” (undergrads) in the subject header. Normal turnaround time is within one weekday, but I may save general questions for the next class period. If you don’t hear back from me and your question cannot wait for the next class, try to email again, or call my office or stop by.