

COMP/LING 445, LING 645: Computational Linguistics, Fall 2020

Tue/Thu 8:35 AM – 9:55 AM

Instructor	Tim O'Donnell timothy.odonnell@mcgill.ca Office hours: By Appointment
Teaching Assistants	Jacob Hoover jacob.hoover@mail.mcgill.ca Office hours: By Appointment Michaela Socolof michaela.socolof@mail.mcgill.ca Office hours: By Appointment

1 Overview

The fields of computer science and modern generative linguistics have grown together since their very beginnings. In this course, we will learn about the tools that have been developed over the years for the analysis of natural language in a computational setting. The emphasis will be on learning the methods and logic underlying the computational generative approach, rather than on learning recent applications from natural language processing.

2 Goals

The goal of this course is to learn how to think about linguistic structure formally, define computational models of linguistic structure, and evaluate their ability to account for empirical linguistic data rigorously.

Students will:

- Learn how to formalize notions like languages and grammars
- Learn how to define probability distributions over formal languages and grammars
- Learn how to work with such probability distributions algorithmically
- Learn how to evaluate models empirically
- Develop practical programming skills and concepts that support the above goals

3 Prerequisites

Programming background at the level of COMP 250 or equivalent. Mathematics background at the level of MATH 240 or equivalent. Basic calculus will be helpful but not critical. Introductory level linguistics at the level of LING 201, especially introductory syntax will be helpful but not critical. We will emphasize building up all basic tools from scratch, but the class will move very fast.

4 Course structure

The course will be taught using an interactive textbook with runnable code samples using the Clojure programming language. The class is structured around this textbook. Problem sets will involve implementing concepts from the book. We will cover approximately the first half of the book, depending on the speed with which we make it through the material.

There will be five term-time problem sets, one problem set due after term, and a short online multiple-choice quiz each week, administered through the myCourses website—starting the week of September 15th and going through the week of November 24th (11 quizzes).

Problem set questions will generally consist of: (i) programming problems in Clojure (ii) mathematical problems and (iii) short answer problems. Problems will be submitted as ClojureScript files or L^AT_EX-based PDFs via myCourses. Problem sets are due before the beginning of class at 8:35 in the morning on the due dates specified below. Note that we cannot debug code that does not run and problem sets whose code does not run will receive a 0.

Participation and interaction is encouraged in this course. We will sometimes use 15–20 minutes of class time to cover questions on preceding problem set or to review other difficult material.

Class discussion and announcements will take place through the myCourses site.

5 Evaluation

Note that details below are subject to change.

- Weekly Quizzes (28%): 11 short, multiple choice quizzes administered through myCourses.
- Problem Sets (72%): 6 problem sets equally weighted (12% each).

6 Logistics

Course Website: <https://mycourses2.mcgill.ca/d2l/home/466603>

Course Webbook: <https://foundations-computational-linguistics.github.io/>

7 Conferences

There will be two weekly conferences beginning the week of September 15th, 2020. Each conference will be run by one of the two TAs above at a set time which will be determined during the first weeks of class. Students are welcome to attend either conference, or both. Conferences will generally be a venue to review material from lecture, go over new examples, and answer student questions. Conferences will not be recorded.

8 Lecture Recordings

This fall this course will be delivered online for the first time. Lectures will be delivered live, at the scheduled time on Tuesdays and Thursdays. However, in order to make sure that the course content is available for students located in other time zones who are unable to be in Montréal this term, lectures will be recorded using Zoom and uploaded to the course website on myCourses. Note that these videos may only be available on myCourses for a limited time after the delivery of each lecture.

Attendance: Although these videos will be available, students are strongly encouraged to attend the live lectures, if it is at all possible. The live lectures provide an opportunity for interaction that will otherwise be lacking and will allow the lecturer to more finely tune the material depending on questions and feedback.

Cameras: Although it is not required, we ask that students please leave on their cameras whenever possible. Having more members of the class visible greatly increases the feeling of interaction, feedback, and community that can sometimes be missing from online course delivery. It also allows the lecturer to better gauge how students are following the material.

Participation: Students are strongly encouraged to participate in class by asking questions and participating in discussions during lectures. This can be done by using the hand raising function in Zoom, and then once the lecturer has called on you, turning on your mic and asking the question. Students may also ask questions in the Zoom chat during lecture, although to encourage discussion the first option is preferred.

Video Privacy: To protect the privacy of students, lecturers, and teaching assistants, all class participants must ensure that videos and associated materials are not reproduced or placed in the public domain. This means that the materials can be used for educational and research purposes, but cannot be shared with others by putting them up on the internet, by giving them away or selling them, or by allowing others to copy them or otherwise make them available. Please refer to McGill's Guidelines for Instructors and Students on Remote Teaching and Learning for further information.

Changes: Since this is the first time this course has been given entirely online, many of the details are experimental. We may make changes to delivery, participation, or recording if we discover better ways of doing things.

9 Readings

The class will make use of the interactive textbook above.

10 McGill Policy Statements

McGill University values academic integrity. Therefore, all students must understand the meaning and consequences of cheating, plagiarism and other academic offences under the Code of Student Conduct and Disciplinary Procedures (see www.mcgill.ca/students/srr/honest/ for more information).

In accord with McGill University's Charter of Students' Rights, students in this course have the right to submit in English or in French any written work that is to be graded.

Instructor generated course materials including recorded lectures, course book materials, notes, summaries, quiz questions, etc. are protected by law and may not be copied or distributed in any form or in any medium without explicit permission of the instructor. Note that infringements of copyright can be subject to follow up by the University under the Code of Student Conduct and Disciplinary Procedures.

11 Course schedule

Note that the exact dates below are subject to change, depending on how quickly we make progress through topics in the course.

- Thursday, September 3, 2020
 - 📎: Problem Set 1 Released
- Tuesday, September 8, 2020
- Thursday, September 10, 2020
- Tuesday, September 15, 2020
 - 📎: Problem Set 1 Due
 - 📎: Problem Set 2 Released
- Thursday, September 17, 2020
- Tuesday, September 22, 2020
- Thursday, September 24, 2020
- Tuesday, September 29, 2020
- Thursday, October 1, 2020
 - 📎: Problem Set 2 Due
 - 📎: Problem Set 3 Released
- Tuesday, October 6, 2020
- Thursday, October 8, 2020
- Tuesday, October 13, 2020
- Thursday, October 15, 2020
- Tuesday, October 20, 2020
 - 📎: Problem Set 3 Due
 - 📎: Problem Set 4 Released
- Thursday, October 22, 2020

- Tuesday, October 27, 2020
- Thursday, October 29, 2020
- Tuesday, November 3, 2020
- Thursday, November 5, 2020
 - 📎: Problem Set 4 Due
 - 📎: Problem Set 5 Released
- Tuesday, November 10, 2020
- Thursday, November 12, 2020
- Tuesday, November 17, 2020
- Thursday, November 19, 2020
- Tuesday, November 24, 2020
 - 📎: Problem Set 5 Due
 - 📎: Problem Set 6 Released
- Thursday, November 26, 2020
- Tuesday, December 1, 2020 [Last Class]
- Thursday, December 3, 2020 [After Term]
- Tuesday, December 8, 2020 [After Term]
- Thursday, December 10, 2020 [After Term]
 - 📎: Problem Set 6 Due