

Computational Linguistics

Course Information

This advanced-level course introduces symbolic concepts, tools, and methods for working with natural language in computational systems. This course complements the introductory statistical NLP course, Linguistics 539 (and is a prerequisite for 539).

Course objectives

This course introduces programming that is relevant to computational linguistics in three programming languages—Perl, Python and Prolog. This course also uses the Chomsky Hierarchy as a framework to introduce concepts and tools that are commonly used in symbolic computational linguistics: regular grammars, as represented by regular expressions, finite-state automata and finite-state transducers; and context-free grammars, as represented by Prolog definite clause grammars. We apply these tools to parsing a small range of realistic language data.

Learning outcomes

Successful students in this course will...

1. be able to read and write simple programs in Perl, Python, and Prolog.[†]
2. be able to apply basic concepts, techniques and tools in symbolic computational linguistics.[‡]
3. be able to explain more advanced concepts (of the student's choice) in computational linguistics.
4. be well prepared to take *Advanced Computational Linguistics* (Ling 581) and *Statistical Natural Language Processing* (Ling 539).

[†] relates to Linguistics HLT program outcome #1.

[‡] relates to Linguistics HLT program outcomes #2 and #3.

HLT learning outcomes addressed in this course

1. Students will demonstrate programming skills for the workplace.
2. Students will be able to use fundamental algorithms and concepts in Natural Language Processing.

3. Students will show knowledge of tools and packages used in Natural Language Processing.

Prerequisites

Although having successfully completed LING 529 and 531, or equivalent, will help students gain the most from this course, there are no formal prerequisites.

Instructor

name Eric Jackson
email ejackson1@email.arizona.edu
hours Thursdays 10:00am–12:00pm (Arizona time, UTC-7) and by appointment,
online via Zoom at <https://arizona.zoom.us/j/88591523837> (passcode 679812)

Requirements

Students are expected to actively participate in the course by watching the recorded lecture videos, reading the assigned readings, completing the assigned homework, and engaging with the instructor and other students in the course forum. You are all adults, and you are responsible for your own learning.

Lecture videos will be available on the course website (D2L). You are expected to watch all lectures and understand the content. If the content of a lecture is not clear, you are expected to send a question to the instructor by email, meet with the instructor in regular office hours or arrange another time to meet, or post a question for clarification on the course forum.

There will be twelve graded homework assignments, eleven forum discussions that are graded for completion only, and ungraded assignments. Although attempting or completing the ungraded assignments will not affect your grade, they *will* affect how much of the course content you understand and retain. I expect that students will attempt all assignments, both graded and ungraded.

Graded homework assignments will be given via the course website (D2L). Assignments will have unequal weights: the first seven assignments will be worth 5 points, while the last five will be worth 6 points. Student homework submissions will also be collected through the course website and must be in PDF format; files submitted in any other format (.doc, .docx, .rtf, .odt, .txt, or any other) will not be accepted. Freely available options to convert files to PDF include Google Docs and LibreOffice; students also have free online access to Adobe Acrobat by signing in with their UArizona NetID.

Forum discussion items will be graded for completion, not for content. There will be eleven forum prompts, but only ten will be required for full participation; one grade item from this category will be dropped. Although I would be happy for all students to participate in all forum discussions, this means that students may choose not to respond to *one* forum prompt without penalty.

All students will give a **10-15 minute presentation summarizing a chapter selected from the textbook**. Because this is an asynchronous course, each student will record themselves giving this presentation and post the video online. Students will also be required to watch and post feedback for at least two other students' presentations.

| type | number | total |
|----------------------|--------------|-------|
| homework assignments | 12 | 65% |
| forum interaction | 10 out of 11 | 10% |
| book chapter report | 1 | 25% |
| total | | 100% |

The due date for each assignment will be posted with the assignment in D2L. All times will be given in Arizona time (Mountain Standard, GMT-7). **Late work will otherwise not be accepted.**

Readings

A draft version of the textbook used in this course is available for free on-line.

- *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*, Jurafsky and Martin, <https://web.stanford.edu/%7Ejurafsky/slp3/>

Technology

This course will use Perl 5, Python 3 (I'll be using Python version 3.8), and SWI Prolog. All of these can be downloaded for free, and guidance on how to install these will be provided in the course materials in D2L. I can provide limited support to get these running in a Linux environment, but if you prefer to have them running in a different operating system, you're welcome to do so.

You are free to install any of these programming languages either directly in your system or indirectly via Docker, a virtual machine, a dual-boot linux system, etc. You just need to be able to run the commands that the homework asks of you.

Collaboration Policy

Students are encouraged to discuss problems and general approaches for solutions, but everyone must turn in their own work. You may not submit assignments that are substantially the same as your classmates. Assignments that seem suspiciously similar will be forwarded to the Dean of Students office in accordance with the Code of Academic Integrity (linked below). Please be a responsible adult and don't run the risk of losing credit for an assignment by copying *or* by allowing others to copy.

Schedule

The course is divided into four topical modules, with lectures and assignments for each module. The length of each module is *roughly* an integer number of weeks, but not exactly. On average, students will need to complete one lecture every two days.

| Module | Dates | Lectures | Homework |
|--------------------------|--------------------|-----------------|----------------|
| 1: Foundations | 1/11-1/17 (7 days) | 4 (2 are short) | HW 1, 2, 3 |
| 2: Working with Perl | 1/18-2/2 (16 days) | 8 | HW 4, 5, 6 |
| 3: Regular Languages | 2/3-2/22 (20 days) | 11 | HW 7, 8, 9, 10 |
| 4: Context-Free Grammars | 2/23-3/3 (9 days) | 4 | HW 11, 12 |

University boilerplate

All of the following items are required by the university to be included on syllabi. If you find something here that is surprising or unexpected, please bring it up with me as soon as possible.

By way of a brief summary:

Disabilities If you have a disability that affects how you will need to do the work in this class, please let me know *within the first week of class*.

Academic Code of Conduct Cheating and plagiarism are not remotely acceptable in any way. Disruptive behavior in class—which here means on any of our course websites or by email—is not acceptable. Please be respectful of others.

Sensitive Material This is a university and you are adults. It is possible that we may touch on topics that some students could find sensitive during the semester. Given the focus of this course, this seems unlikely, but I alert you nonetheless.

Covid

The university has a specific site for covid information: <http://covid19.arizona.edu>. These are extraordinary times and you may still be experiencing personal and financial challenges. Let me know if we need to make accommodations for covid-related things, and please stay safe.

Absence and Class Participation Policy

Attendance in an all-online course is not evaluated like attendance in an in-person course. For this course, attendance will be represented by active reading, completion, and participation in online course activities, including materials and activities posted on D2L, OpenClass, our course forum, and any other related websites.

The UA's policy concerning Class Attendance, Participation, and Administrative Drops is available at: <http://catalog.arizona.edu/policy/class-attendance-participation-and-administrative-drop>

The UA policy regarding absences is that any sincerely held religious belief, observance or practice will be accommodated where reasonable, <http://policy.arizona.edu/human-resources/religious-accommodation-policy>.

Absences pre-approved by the UA Dean of Students (or Dean Designee) will be honored. See: <https://deanofstudents.arizona.edu/absences>

Classroom Behavior Policy

To foster a positive learning environment, students and instructors have a shared responsibility. We want a safe, welcoming, and inclusive environment where all of us feel comfortable with each other and where we can challenge ourselves to succeed. To that end, our focus is on the tasks at hand and not on extraneous activities.

Students are asked to refrain from disruptive conversations with others in the course. Students observed engaging in disruptive activity will be asked to cease this behavior. Those who continue to disrupt the class will be asked to leave lecture or discussion and may be reported to the Dean of Students.

Threatening Behavior Policy

The UA Threatening Behavior by Students Policy prohibits threats of physical harm to any member of the University community, including to oneself. See <http://policy.arizona.edu/education-and-student-affairs/threatening-behavior-students>.

Accessibility and Accommodations

At the University of Arizona, we strive to make learning experiences as accessible as possible. If you anticipate or experience barriers based on disability or pregnancy, please contact the Disability Resource Center (520-621-3268, <https://drc.arizona.edu/>) to establish reasonable accommodations.

Code of Academic Integrity

Students are encouraged to share intellectual views and discuss freely the principles and applications of course materials. However, graded work/exercises must be the product of independent effort unless otherwise instructed. If you found a code snippet online, it's important to cite where it came from, even if that source was stackexchange.com.

Students are expected to adhere to the UA Code of Academic Integrity as described in the UA General Catalog. See: <http://deanofstudents.arizona.edu/academic-integrity/students/academic-integrity>.

UA Nondiscrimination and Anti-harassment Policy

The University is committed to creating and maintaining an environment free of discrimination; see

<http://policy.arizona.edu/human-resources/nondiscrimination-and-anti-harassment-policy>

Subject to Change Statement

Information contained in the course syllabus, other than the grade and absence policy, may be subject to change with advance notice, as deemed appropriate by the instructor.