

LING 261: Language and Computers Spring 2020

Instructor: Emma Manning, esm76@georgetown.edu

Class time/location: MW 5:00-6:15, Maguire 104

Office Hours time/location: Thursdays 2:00-3:00, Lauinger 2nd floor (outside/near Midnight MUG)

Course Description

Science fiction has promised us intelligent robots like C3P0 and HAL, but instead we're stuck with Siri. What happened? Why has getting computers to understand language proven so difficult?

In this course, we'll look at this question through the history of computational linguistics and natural language processing: what approaches have researchers taken over the last 60 years of computational linguistics? In what ways did those approaches succeed? In what ways did they fail?

Topics will include:

- The Goals and Applications of Computational Linguistics
- Pre-statistical Approaches to Natural Language Processing (NLP)
- Modern, Statistical Approaches to NLP

The class will focus on the concepts behind these topics rather than implementing them, so no programming experience is required.

No prerequisites, though many concepts will overlap with those in Introduction to Language (LING 001).

There is no textbook to purchase; any required readings will be provided via Canvas.

Objectives

By the end of this class, you should:

- Understand the basics of how linguists analyze language, and especially how this applies to computers' use of human language
- Understand the similarities and differences between historical and more recent approaches to NLP, and the pros and cons of each type of approach
- Develop intuitions about what computers can and cannot do when it comes to language, and be able to address over-hyped claims of artificial intelligence
- Develop an awareness of the place of NLP technologies in your life, and an understanding of their strengths and limitations
- Develop an awareness of ethical issues in NLP, including privacy and bias concerns

Assignments & Grading

Grade breakdown

| | |
|---------------|-----|
| Homework | 60% |
| Final Project | 30% |
| Participation | 10% |

Homework

Homework assignments will be assigned on an approximately weekly basis; typically, they will be distributed on Wednesday and due the following Monday at the start of class (5:00 pm) on Canvas. These may include a variety of ways to engage with course material, including but not limited to:

- Reading an article and writing a short reflection
- Experimenting with an NLP technology and writing a short reflection
- Practicing concepts by performing a linguistic analysis and/or some of the calculations of an NLP algorithm by hand

Homework assignments will be weighted equally in the final course grade, with the lowest assignment grade dropped.

Extensions may be granted in special circumstances; please communicate these to me well in advance whenever possible. Unexcused late work will be accepted up to a week late, but the grade will drop by 10% of the maximum grade for each day past the deadline.

Final Project

For the final project, you will identify something you learned in this course that you would like the general public to know more about, and envision how you might communicate this concept. This may be done individually or in small groups, and you will have a choice of mediums. More information will be given later in the semester.

Attendance & Participation

To receive full credit for attendance and participation, I expect you to regularly show up to class on time and engage meaningfully with the class content, including participating in class discussions.

You are allowed up to two unexcused absences without affecting your course grade. This does not include cases where you are excused due to illness, religious observances, athletic travel, etc. Please email me about such situations, well in advance when possible, to be excused from these class sessions. Make sure to get notes from a classmate whenever you miss class.

Policies

Academic Integrity & Collaboration

You are welcome to discuss homework with other students in the class, but the work you turn in should be your own. If you collaborate with classmates or consult sources other than course materials, provide attribution for these ideas (e.g. "I worked with [classmate's name] on this question" or "I looked at [website link] for more information."). For more information on the university's academic honesty policies, see <https://bulletin.georgetown.edu/regulations/honor/>.

Technology Use

Technology such as laptops in the classroom are a double-edged sword: they can be a useful tool to support course engagement, including by facilitating note-taking, but they are often frequently a distraction. Please consider the role these types of technology play in your personal classroom experience, and only use devices in class when they will support your learning.

Disability Accommodations

If you have a disability that may affect your performance in this class, please let me know as soon as possible so we can arrange for accommodations. In this case you should also contact the Academic Resource Center in Leavey 338 (<http://academicsupport.georgetown.edu/disability>).

Instructional Continuity

If campus is closed during our scheduled class time, I will send an announcement via Canvas about activities to replace the regular classroom instruction.

Notice Regarding Sexual Misconduct

Please know that as a faculty member I am committed to supporting survivors of sexual misconduct, including relationship violence, sexual harassment and sexual assault. University policy also requires me to report any disclosures about sexual misconduct to the Title IX Coordinator, whose role is to coordinate the University's response to sexual misconduct.

Georgetown has a number of fully confidential professional resources who can provide support and assistance to survivors of sexual assault and other forms of sexual misconduct. These resources include:

Jen Schweer, MA, LPC
Associate Director of Health Education Services for Sexual Assault Response and Prevention
(202) 687-0323
jls242@georgetown.edu

Erica Shirley, Trauma Specialist
Counseling and Psychiatric Services (CAPS)
(202) 687-6985
els54@georgetown.edu

More information about campus resources and reporting sexual misconduct can be found at <http://sexualassault.georgetown.edu>.

Schedule

This schedule is tentative and subject to change as the semester progresses. Check Canvas for the most recent version of the syllabus.

| Date | Topic | Assignments |
|------------------------------------|--|-------------------------|
| W 1/8 | Introduction | HW 1 Assigned |
| M 1/13 | What are computers? | HW 1 Due |
| W 1/15 | What is language? | HW 2 Assigned |
| <i>No class 1/20: MLK Day</i> | - | - |
| W 1/22 | Writing Systems & Language Representations | HW 2 Due; HW 3 Assigned |
| M 1/27 | Tokens & Regular Expressions | HW 3 Due |
| W 1/29 | Morphology & Finite State Automata | HW 4 Assigned |
| M 2/3 | Parts of Speech | HW 4 Due |
| W 2/5 | Syntax & Parsing, part 1: Constituents | HW 5 Assigned |
| M 2/10 | Syntax & Parsing, part 2: Dependencies | HW 5 Due |
| W 2/12 | N-Gram Modeling & Generation | HW 6 Assigned |
| T 2/18 Monday Schedule | Bags of Words & Vector Models | HW 6 Due |
| W 2/19 | Topic Modelling | HW 7 Assigned |
| M 2/24 | Naïve Bayes Classification | HW 7 Due |
| W 2/26 | Whose Data do we use? | HW 8 Assigned |
| M 3/2 | Creating Corpora | HW 8 Due |
| W 3/4 | Using Corpora for Linguistics | HW 9 Assigned |
| <i>Spring Break</i> | - | - |
| M 3/16 | Intro to Neural Networks | HW 9 Due |
| W 3/18 | Machine Translation | HW 10 Assigned |
| M 3/23 | Dialogue Systems | HW 10 Due |
| W 3/25 | Muppets | HW 11 Assigned |
| M 3/30 | Evaluation | HW 11 Due |
| W 4/1 | Is this AI? | Final Project Assigned |
| M 4/6 | Is this Ethical? | |
| W 4/8 | Are Computers Changing Language? | Project Proposals Due |
| <i>No class 4/13: Easter Break</i> | - | - |
| W 4/15 | Bonus Topic | |
| M 4/20 | Final Project Presentations | |
| W 4/22 | Final Project Presentations | |
| M 4/27 | Wrapping Up | |
| W 5/6 | - | Final Project Due 6pm |