University of Illinois at Chicago
Spring 2019

CS 421 — Natural Language Processing
Course Syllabus

Room: 138 SES
Time: TuTh   9:30 – 10:45
URL: via Blackboard for all materials; via Piazza for discussions and questions
(No class materials will be posted on Piazza)

Staff

Instructor: Barbara Di Eugenio
Office: 1138 SEO, 6-7566
E-mail: bdieugen@uic.edu
Office Hours: Tue & Th 3:00-4:30

Teaching Assistant: Mehrdad Alizadeh
Office: TBD
E-mail: maliza2@uic.edu
Office Hours: Mon & Wed 2-4pm

Course Objectives

The aim of this course is to introduce students to the field of Natural Language Processing (NLP), also called Human Language Technology, or Natural Language Engineering, or Computational Linguistics. NLP studies algorithms to enable the computer to interpret and produce *natural languages*, i.e. English, Russian, Mandarin Chinese, Italian, Turkish, etc. By now, applications of NLP include successful systems such as Google Translate, Siri, Alexa (Amazon Echo) and Cortana.

The course will provide students with the linguistic foundations that underlie NLP, introduce them to the algorithms used in the field, and provide practice in building components of NLP systems.

Reading Materials

Required Textbook: Daniel Jurafsky and James H. Martin. Speech and Language Processing (**Second edition**). Prentice Hall, 2008.

1. **Important.** The international edition is different enough that if you choose to use it, **you are responsible** for matching sections and pages across the two books.

2. A third edition of the textbook is in the making and several chapters are available on line. You are required to use the older published version, unless informed otherwise. As appropriate I will use an occasional chapter from the new edition, which I will post on Blackboard.

Prerequisites: CS 301.

Notes

- I use email (via Blackboard) a lot to communicate with the whole class. Please check your email frequently, especially around deadlines (homeworks and exams).

- The web page (Blackboard) will contain all materials relevant to the class, syllabus, assignments etc. You can also see you own grades.

- When you have a question of general interest, rather than sending me or the TA mail, post your questions to the discussion board on Piazza – **after you've checked the text of the homework, the notes posted on Blackboard, etc**. We will provide an answer within 24 hours, and most often, much faster than that.

Tentative Schedule

Dates	Topic	Readings
Week 1	Introduction	Ch. 1
Weeks 2-3	Words: Morphology, n-grams	Ch. 3-4
Weeks 4-5	Intro to NLP tools; Words: Part-of-speech tagging	Ch. 5-6 (excerpts)
Week 6	Word meaning	Ch. 6 **(3rd ed.)**
Weeks 7-8	Syntax: Parsing	Ch. 12-16 (excerpts)
Weeks 9-10	Semantics	Ch. 17-18 (excerpts)
Week 11	Discourse: Referring expression resolution	Ch. 21
Weeks 12-13	Dialogue systems and chatbots	Ch. 24 **(3rd ed.)**
Weeks 14-15	Applications and Catch-up	

Exam Dates

Date	Event
Th 3/7 (class time)	**Midterm**
Wed 5/8 (10:30-12:30)	**Final** (to be confirmed)

Grading Criteria

The class will be graded out of 1000 points, half for the two exams, half for homeworks and the project.
IMPORTANT NOTE: To pass the class you must get at least 60% of the total available points for the exams, i.e, your cumulative score across the two exams must be at least 300 points.

- **2 Exams** (500 points): 1 midterm (worth 200 points), 1 final (300 points).

- **Homework 0 (Zero)** (20 points): during weeks 1-2 of class, come to introduce yourself to the instructor or TA **during office hours**. You can come with your friend(s). The purpose of this "assignment" is to make students comfortable with coming to office hours later in the semester.

- **2-3 Homework Assignments** (between 40 and 70 points each): These will be fast assignments.

- **Project** (between 300-350 points): The project will comprise two parts, for a total of 30-35% of the grade. The goal of the project is to design and implement a simple NLP system, by using existing tools to do some task. You will do the project in **pairs**. Graduate students will have additional requirements to satisfy; hence, it is preferable that pairs are homogeneous (both ugrads or grads).

Letter grades will be decided **only at the end**. However, the following guidelines will be adhered to:

Overall Score (undergraduate)	Overall Score (graduate)	Letter grade
88%	92%	A
78%	82%	B
68%	72%	C

What these guidelines mean is that the lower bounds guarantee you that specific letter grade; thresholds may be adjusted lower, depending on the general distribution of scores.

Policies on homeworks and exams

General Policies

1. Late homeworks will not be accepted in any case, unless there is a **documented** personal emergency. Arrangements must be made with the instructor as soon as possible after the emergency arises, preferably before the homework due date.
   Advice: If for whatever reason you don't manage to finish an assignment, hand in what you have. Partial credit will be given.

2. Statute of Limitations: **Two weeks!** No grading questions or complaints — **no matter how justified** — will be listened to two weeks after the item in question has been returned.

Exams

1. The midterm will be given during class time; consequently, **no make-up** will be given.

2. Exams will be closed-book.

3. The final will be cumulative, although it will stress the materials covered after the midterm.

Policy on Academic Integrity

Unless stated otherwise, all work submitted for grading **must** be done individually. While we encourage you to talk to your peers and learn from them, this interaction must be superficial with regards to all work submitted for grading. This means you **cannot** work in teams unless told so, you cannot work side-by-side, you cannot submit someone else's work (partial or complete) as your own. The University's policy is available at https://dos.uic.edu/conductforstudents.shtml.

In particular, note that you are guilty of academic dishonesty if you extend or receive any kind of unauthorized assistance. Absolutely no transfer of program code between students is permitted (paper or electronic), and you may not solicit code from family, friends, or online forums. Other examples of academic dishonesty include emailing your program to another student, copying-pasting code from the internet, working in a group on a homework assignment, and allowing a tutor, TA, or another individual to write an answer for you. Academic dishonesty is unacceptable, and penalties range from a letter grade drop to expulsion from the university; cases are handled via the official student conduct process described at the URL above.