ENG 685: Computational Linguistics
Fall 2013, W 7:20–10:00p, BA339

Instructor: Christian F. Hempelmann
Office: Hall of Languages 226
Office Hours:
online workdays 10am–10pm
office W 10a–12n and by appt.
Phone: (903) 886-5291
Fax: (903) 886-5980
Email: c.hempelmann@tamuc.edu

COURSE INFORMATION

Materials – Textbooks, Readings, Supplementary Readings:
Required textbooks:
- none

Required texts:
- Jurafsky, Daniel and James H. Martin. 2008. *Speech and Language Processing.* 2nd Ed. Prentice-Hall. Selected chapters. Online.
- Manning, Christopher D. and Hinrich Schütze. 1999. *Foundations of Statistical Natural Language Processing.* MIT. Selected chapters. Online.
- Nirenburg, Sergei and Victor Raskin. 2004. *Ontological Semantics.* MIT. Selected chapters. Online.

Optional additional reading:
- Further materials will be made available online.

Course Description:
"This course provides a general introduction to Computational Linguistics, the study of computational systems that understand and generate human language. This class will cover fundamental concepts and techniques in Computational Linguistics, such as lexical and ontological semantics, word sense disambiguation, syntactic and semantic parsing, discourse (including coreference resolution), dialogue, summarization, and generation. Throughout the class, students will be exposed to recent research that connects the concepts learned to exciting research questions that are practically motivated and application-oriented. Additional emphasis will be on the different traditions and theoretical frameworks that informed the theories and algorithms used for these solutions, namely, linguistics, statistics, and computer science, and the history of their struggle from the 1950s until today." (general catalogue description)

Course Objectives:
1. Students will become familiar with basic concepts in computational linguistics and the way of linguistically thinking about an issue through readings from texts, in-class discussions, and exercises done as homework and in class. This objective will be measured through exercises and contributions to a final project and exam.

2. Students will become active participants in the course, not only in staying current with readings and other assignments, but also in sharing their understanding of the material as assessed by weekly exercises and contributions to in-class discussions.

Student Learning Outcomes

Students who have successfully participated in this class will

1. understand the history of the discipline and its subdisciplines, in particular in relation to the fields of linguistics—phonology, syntax, semantics—and other fields of science;
2. have hands-on knowledge of theories and algorithms required to process natural language;
3. grasp the crucial differences between natural and artificial languages;
4. be able to apply this knowledge to real-world issues, theoretically and practically;
5. read scientific texts on these issues;
6. do guided research and in a topic in computational linguistics.

COURSE REQUIREMENTS

Instructional Methods, Activities, and Assessments

Readings and Topics

Many of the readings for this course will come from Jurafsky and Martin (2008). We will begin by reading introductory chapters, continue by exploring issues in the history of the discipline and the main problems it has faced, and finally discuss selected theories and topics in computational linguistics with the help of the readings and current research of the instructor. These topics include, human-computer interface design, voice recognition and production, machine translation, data mining, web search technology, computational humor, information security and assurance, and artificial intelligence. A second emphasis will be corpus work in relation to the aligned class by Dr. Reid that focuses on creating and curating a corpus of selected works by J.R.R. Tolkien. These readings from other sources will be made available on eCollege and as links to the webpages of their authors. Read all assigned readings closely before class and be prepared to discuss them in class.

Exercises and Final Project

Throughout the semester we will do a number of smaller graded (plus, check, minus) exercises, some in class, most as homework. Assignments have to be handed in on time and are dropped one letter grade for each day they are late, unless there is a documented emergency. The final project for this class will be a guided literature review on a topic in computational linguistics, a programmed solution to a real-world problem in NLP, or a solid practical and theoretical involvement in the Tolkien Corpus Project. After we have discussed the basics of the field and looked at a number of applications, you will write a proposal for your project (participation). After the proposal has been approved, you will produce a draft, which will be the basis of a conference with the instructor. The final project (participation) will be due at the end of the semester.

Computers

The class meets in a computer lab and computer literacy, as well as a little programming (or a lot, if you want to) is part of this course. Apart from in-class writings, all assignments

must be printed. Save everything you write. Make backup copies. Losing a file is no fun and no excuse for missing an assignment.

Much of our communication will be by e-mail, and I usually send e-mails after every class. It is university policy that you check your e-mail account at least once per weekday. If you send me e-mail, include the course number and the project in the subject line. There will be an eCollege shell, where materials, including this syllabus, and announcements will be posted and where assignments can be exchanged.

Attendance and Participation

You don't want to miss class! It will affect your performance and your grade. Attendance is taken. If you are not there, you cannot get feedback, participate in the discussion (one percent down for every miss after the third one), and hand-in assignments. So be there! Missing an appointment we have set is an absence. For an excused miss, and thus the possibility not to lose points, you have to let me know in advance. If you miss a class, you are responsible to find out what was going on in class. Participation facilitates your understanding of the issues we'll discuss in this class. A positive attitude and active participation are important to succeed in this class.

Grading Policy

If you have completed all assignments by their due dates, in particular the midterm and final, and satisfied the attendance requirements, and all other requirements stated in this specific or the departmental syllabus, your grade for this course will be determined as follows:

A	≥ 90%	assignments, exercises	50%
B	89-80%	final project	30%
C	79-70%	attendance, attitude, participation	20%
D	69-60%		
F	< 59%		

Attendance Policy

Students are responsible for attending class and keeping a log of their attendance. Please note that this means that no excuse will be accepted for failure to comply with the class requirements. No make-up quizzes will be given or late assignments accepted. Students needing particular attention should notify the instructor during the first week. If you miss a class you are responsible to receive the information you missed or you have to be prepared for surprises.

Tentative Calendar

week		*notes*
8-28	first day of class	
9-4		
9-11		
9-18		
9-25		
10-2		
10-9		
10-16		
10-23		
10-30		
11-6		
11-13		
11-20		

11-27 11-28 Thanksgiving
12-4 last class meeting
12-9 finals week

Note that this schedule is purely tentative and provided as a rough idea of how we will distribute class time. Changes can and will be made.

COURSE AND UNIVERSITY PROCEDURES/POLICIES

Course Specific Procedures:

Academic Honesty Policy

Texas A&M University-Commerce does not tolerate **plagiarism** and other forms of **academic dishonesty**. Conduct that violates generally accepted standards of academic honesty is defined as academic dishonesty. "Academic dishonesty" includes, but is not limited to, plagiarism (the appropriation or stealing of the ideas or words of another person and passing them off as one's own), cheating on exams or other course assignments, collusion (the unauthorized collaboration with others in preparing course assignments), and abuse (destruction, defacing, or removal) of resource material. Instructors uphold and support the highest academic standards, and students are expected to do likewise. Penalties for students guilty of academic dishonesty include disciplinary probation, suspension and expulsion (Texas A&M University-Commerce Code of Student Conduct 5.b[1, 2, 3]). Cheating, including plagiarizing papers in whole or part, will result in a grade of zero (0) on the assignment for the first offense and failure of the course for any subsequent offenses.

Attendance Policy

Because your active participation in discussions is expected, attendance is vital for success in this course. Attendance requires more than simply coming to class.

Assignment Policy

I will provide specific details during class meetings and possibly by email. Students must stay current with all readings and discussions. Students must complete and submit all assignments by their respective due dates.

Late Work

I will not accept any assignment after its due date. Assignments submitted after the due date may receive a score of zero (0).

Drop a Course

A student may drop a course by logging into his/her myLEO account and clicking on the hyperlink labeled "Drop a class" from among the choices found under the myLeo section of the Web page.

Incompletes

Incompletes (grade of "X") are granted only under rare and extraordinary circumstances.

Administrative Withdrawal

I reserve the right to drop a student from the course administratively for excessive absences or violations of student conduct codes.

University-Specific Procedures:

Students with Disabilities

The Americans with Disabilities Act (ADA) is a federal anti-discrimination statute that provides comprehensive civil rights protection for persons with disabilities. Among other things, this legislation requires that all students with disabilities be guaranteed a learning environment that provides for reasonable accommodation of their disabilities. If you have a disability requiring an accommodation, please contact:
Office of Student Disability Resources and Services
Texas A&M University-Commerce
Gee Library
Room 132
Phone (903) 886-5150 or (903) 886-5835
Fax (903) 468-8148
StudentDisabilityServices@tamuc.edu

Student Conduct

All students enrolled at the University shall follow the tenets of common decency and acceptable behavior conducive to a positive learning environment. (See *Code of Student Conduct from Student Guide Handbook).*