

COS 380

Natural Language Processing

Department of Computer Science and Engineering
Taylor University

Three Credit Hours
Euler Science 103 • MWF 1:00–1:50

Final Exam: M/09-Dec-2019, 8:00–10:00

1 Instructor

Dr. Tom Nurkkala

Associate Professor, Computer Science and Engineering
Director, Center for Missions Computing

Office Euler Science Complex 211
Email tnurkkala@cse.taylor.edu
Phone 765/998-5163
Hours Mon, Wed 2:00–4:00
Tue, Thu 10:00–11:00
Or by appointment

2 Description

Prerequisites: COS 280

From the catalog: A study of the automation of human communication abilities, covering both textual and vocal aspects. Major topics include language parsing, understanding, representation, enhancement, generation, translation, and speaker/author recognition.

3 Learning Objectives

Upon successful completion of this course, you will be able to:

1. Articulate the history, development, and topics commonly studied in Natural Language Processing (NLP)
2. Use and apply algorithms for text processing
3. Continue investigation into NLP

4. Demonstrate critical inquiry in NLP and computational linguistics
5. Use NLP for practical, hands-on purposes
6. Create software at a high level of sophistication using languages relevant to text processing.
7. Have an awareness of how Christians should apply the capabilities of Artificial Intelligence, especially those in NLP.

4 Possible Topics

NLP is a very large field. We will cover some (but not all) of the following topics.

1. Regular expressions
2. Finite state automata
3. Finite state transducers
4. N-Grams
5. Part-of-speech tagging
6. Hidden Markov Models
7. Formal Grammars of English
8. Syntactic Parsing
9. Statistical Parsing
10. Features and Unification
11. Representation of Meaning
12. Computational Semantics
13. Lexical Semantics
14. Computational Lexical Semantics
15. Computational Discourse
16. Information Extraction
17. Question Answering and Summarization
18. Machine Translation

5 Project

During the course, you will build a non-trivial application that demonstrates your understanding of Natural Language Processing.

6 Text

Rather than a traditional textbook, this course uses selections from various texts available at [Safari Books Online](#). Safari is a subscription service that gives you access to thousands of technical titles. You can also download complete books for off-line access on your mobile devices. The cost to subscribe is \$39/month.

Why use Safari instead of a standard textbook?

1. Rather than asking you to buy an entire textbook and use only *portions* of it, we will be draw on multiple resources that are *directly related* to the topics covered throughout the semester.
2. Your total cost will be about \$120, which is less than the cost of some single textbooks. If you are taking other courses from me this semester, you only need to pay the fee once.
3. You should have experience learning from written and electronic resources, which you will do throughout your technical career. I have been a member of Safari for many years and find it to be an invaluable resource when I am learning a new technology or buffing up my understanding of a familiar one.
4. You will have immediate access to a rich collection of technical material to advance your learning in this or other computer science courses.

7 Evaluation

The grading breakdown for the course is shown in Table [1](#). Refer to my *Periodic Table of the Grades* for the grading scheme. I reserve the right to award a higher grade than strictly earned; outstanding attendance and class participation figure prominently in such decisions.

8 Course Expectations

Following are my expectations regarding the course.

Weight	
Homework	10%
Programming	20%
Participation	10%
Project	20%
Midterm	20%
Final	20%
<hr/>	
Total	100%

Table 1: Grading details

8.1 Attendance

You are required to attend all class sessions. I will be in class each day, and I expect you to be there also.

In general, I am very understanding about students who must miss class due to a sanctioned Taylor activity, medical appointment, job interview, family emergency, and the like. If possible, let me know in advance that you will not be in class; I will work with you to arrange make-up instruction, homework, exams, etc.

8.2 Late Work

All course assignments will include an unambiguous due date. Usually, assignments are due at the beginning of class on the due date. If there are multiple sections of a class, the assignment is due at the beginning of the earliest such section. Barring exceptional circumstances like those mentioned in section 8.1, I expect your work to be submitted *on the due date*. Late work will *not* be accepted.

This policy on late work is intended to prepare you for real-world experience after graduation. In the marketplace, late work is not merely an inconvenience. Missing a deadline may alienate your customer, upset your manager, ruin your project, or terminate your employment! *Now* is the time to learn the self discipline and time management skills required to complete your work when it is due.

8.3 Conduct

I expect you to be prepared, awake, aware, and participatory during class. I will not hesitate to ask you to stand or move if you are distracted or sleepy.

I expect you to join in discussions, respond to questions from me and from your colleagues, and ask questions of me. I expect you to hold my feet to the fire if I am being unclear, unkind, or contradictory.

8.4 Gizmos

You may not use a laptop, tablet, or similar device to check e-mail, engage in social networking, surf the web, or any other activity not directly relevant to current classroom activity. If you use an electronic gizmo during class for legitimate academic purposes (e.g., note taking), be prepared to demonstrate relevant use on demand at any time.

9 Course Management

We use several systems to help manage the course and for on-line communication.

9.1 Email

Electronic mail is an official channel of communication between all members of the university community. You are responsible to check your email regularly (daily) for information related to the course.

9.2 Canvas

The Computer Science and Engineering department uses Canvas as our Learning Management System. The URL for Canvas is <https://canvas.cse.taylor.edu>.

You are responsible for checking Canvas regularly to keep up with assignment due dates and other announcements. For due dates, *the Canvas calendar is your friend*.

9.3 Slack

This course will use Slack for informal communication, Q&A, last minute announcements, jokes, and the like. Find the *TU CSE Student* slack team at tucsestudents.slack.com. Look there for a *channel* dedicated to the course.

10 Academic Integrity

As a student at an institution whose goal is to honor Christ in all that it does, I expect you to uphold the strictest standards of academic integrity. You must do your own work, cite others when you present their work, and never misrepresent your academic performance in any way. Violation of these standards stains the reputations of you as a student, Taylor as an institution, and Jesus as our Lord.

Every assignment should indicate clearly that it is either:

- An **individual** assignment, to be done *entirely by you*, without any direct participation from other students.

- A **group** assignment, to be done *collectively with a group*

Unless otherwise stated, assignments are **individual** assignments.

You are *always* welcome to get help from the instructor on *any* homework assignment or project, whether an individual or group assignment.

10.1 What Constitutes Academic Dishonesty?

For purposes of this course, the following are *non-exhaustive* examples of violations of academic integrity.

1. Sharing code or other electronic files by copying, retyping, looking at, or supplying a copy of a file from this or a previous semester.
2. Sharing written assignments or exams by looking at, copying, or supplying an assignment or exam.
3. Using another student's code. Using code from this or previous offerings of the class, from courses at other institutions, or from any other source (e.g., software found on the Internet).
4. Looking at another student's code. Although mentioned above, it bears repeating: looking at other students' code or allowing others to look at yours is academic dishonesty. There is no notion of looking "too much," since no looking is allowed at all.

10.2 What Does Not Constitute Academic Dishonesty?

In contrast, the following are *non-exhaustive* examples of activities that *do not* violate academic integrity.

1. Clarifying ambiguities or vague points in class handouts or textbooks.
2. Helping others use the computer systems, networks, compilers, debuggers, profilers, or other system facilities without regard to a particular assignment or project.
3. Helping others with high-level design issues.
4. Helping others with high-level (*not* code-based) debugging.
5. Using code provided by the instructor from the course web site or elsewhere.

10.3 From the Provost

Taylor's Provost¹ defines *plagiarism* as follows:

In an instructional setting, plagiarism occurs when a person presents or turns in work that includes someone else's ideas, language, or other (not common-knowledge²) material without giving appropriate credit to the source. Plagiarism will not be tolerated and may result in failing this course, and may also result in further consequences.

The Provost goes on to say:

Academic dishonesty constitutes a serious violation of academic integrity and scholarship standards at Taylor that can result in substantial penalties, at the sole discretion of the University, including but not limited to, denial of credit in a course as well as dismissal from the University.

In short, a student violates academic integrity when he or she claims credit for any work not his or her own (words, ideas, answers, data, program codes, music, etc.) or when a student misrepresents any academic performance.

For more information, see the [Taylor University Undergraduate Catalog](#).

11 Support Services

Be aware of the following support services available to you as a Taylor student.

11.1 Academic Assistance

The Academic Enrichment Center (AEC), located in the Zondervan Library, provides individualized academic skills help (e.g. test preparation, note taking, planning, etc.). Contact **Dr. Scott Gaier**, scgaier@taylor.edu.

11.2 Tutoring

Peer Tutoring Services, located in the AEC in Zondervan Library, provides free help to students in most content areas. For further information, contact **Darci Nurkkala**, drnurkkala@taylor.edu.

¹At Taylor, the *Provost* is our Chief Academic Officer.

²*Common knowledge* means any knowledge or facts that could be found in multiple places or as defined by a discipline, department, or faculty member.

11.3 Students with Special Needs

The Academic Enrichment Center provides a variety of services for students who have physical or other disabilities. If you need accommodations due to a disability, please contact **Ken Taylor**, kntaylor@taylor.edu.

This document last updated August 27, 2019.