1 Instructor

Dr. Tom Nurkkala
Associate Professor, Computer Science and Engineering
Director, Center for Missions Computing
Office: Euler Science 211
E-Mail: `tnurkkala@cse.taylor.edu`
Phone: 765/998-5163
Hours: MW 2:00-4:00 or by appointment

2 Course Overview

This course in Natural Language Processing (NLP) studies the automation of textual human communication abilities. At the conclusion of this course, you will be familiar with problems encountered in NLP as well as a variety of solutions currently in vogue. You will have used standard software packages and will have written NLP software. You will also be familiar with standard language processing applications and algorithms.

3 Learning Objectives

The learning objectives for the course are as follows.

- Introduce you to the history and development of NLP.
- Give you an understanding of the topics commonly included in NLP.
- Stimulate your understanding of the algorithms used in carrying out text processing.
- Stimulate your interest in NLP so that you will continue investigation after the course ends.
- Develop in you an attitude of critical inquiry.
- Provide a setting for scientific investigation & experimentation.
- Stimulate inter-disciplinary learning.
- Practical, hands-on experience in NLP.
- Stimulate development of programming competence in languages relevant to text processing.
- Develop an awareness of the responsibilities a Christian has in the development of AI capabilities, especially those in NLP.

4 Possible Topics

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

- Regular expressions
- Finite state automata
- Finite state transducers
- N-Grams
- Part-of-speecth tagging
- Hidden Markov Models
- Formal Grammars of English
- Syntactic Parsing
- Statistical Parsing
- Features and Unification
- Representation of Meaning
- Computational Semantics
- Lexical Semantics
- Computational Lexical Semantics
- Computational Discourse
- Information Extraction
- Question Answering and Summarization
- Machine Translation

5 Required Texts

There are two texts required for the class. The standard textbook for the field is:

- Daniel Jurafsky and James Martin, *Speech and Language Processing*, Prentice Hall, 2009. ISBN 978-0-13-605234-0.

We will be using the Python-based Natural Language Toolkit, which you can find at `http://nltk.org/`. The NLTK and its latest documentation (`http://www.nltk.org/book/`) are open source and available freely on-line.

There is also a physical book about NLTK, but it is now *out of date* with the current version of the software. However, if you prefer a physical book to on-line documentation, much of the book is relevant. Bibliographic details are:

- Steven Bird, Ewan Klein, and Edward Loper, *Natural Language Processing—Analyzing Text with Python and the Natural Language Toolkit*, O'Reilly Media, 2009. ISBN 978-0-596-51649-9.

6 Moodle

The Computer Science and Engineering department uses Moodle as our Learning Management System. The URL for Moodle is `https://cms.cse.taylor.edu`. To sign on to the course site for the first time, you will need an enrollment key. The key for this course is `nerds4christ`. You are responsible for checking Moodle regularly to keep up with assignment due dates and other announcements posted to the site. For due dates, the Moodle calendar is your friend.

7 Classroom Expectations

Following are my expectations about classroom conduct.

7.1 Attendance

Attendance is required. I will be in class each day, and I expect you to be there also. I will log who attends each class session.

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

7.2 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.

7.3 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.

8 Evaluation

The grading breakdown will be as follows:

Deliverable	Weight
Homework	25%
Project	25%
Attendance and Participation	10%
Midterm	20%
Final	20%

Refer to the Periodic Table of the Grades (on Moodle) for my grading scheme. I reserve the right to award a higher grade than strictly earned; outstanding attendance and class participation figure prominently in such decisions.

9 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. Such a violation may result in your failing the course and other disciplinary action by the University. Refer to the Taylor catalog for the official statement of these ideas.