CSE 5525 (Approved): Foundations of Speech and Language Processing

Course Description

Fundamentals of natural language processing, automatic speech recognition and speech synthesis; lab projects concentrating on building systems to process written and/or spoken language.

Prior Course Number: CSE 732 and CSE 733
Transcript Abbreviation: Spch & Lang Proc
Grading Plan: Letter Grade
Course Deliveries: Classroom
Course Levels: Undergrad, Graduate
Student Ranks: Senior, Masters, Doctoral
Course Offerings: Spring
Flex Scheduled Course: Never
Course Frequency: Every Year
Course Length: 14 Week
Credits: 3.0
Repeatable: No
Time Distribution: 3.0 hr Lec
Expected out-of-class hours per week: 6.0
Graded Component: Lecture
Credit by Examination: No
Admission Condition: No
Off Campus: Never
Campus Locations: Columbus
Prerequisites and Co-requisites: (CSE 3521 or CSE 5521) and (CSE 5522 or Stat 3460 or Stat 3470)
Exclusions: Not open to students with credit for CSE 733
Cross-Listings:

The course is required for this unit's degrees, majors, and/or minors: No
The course is a GEC: No
The course is an elective (for this or other units) or is a service course for other units: Yes

Subject/CIP Code: 14.0901
Subsidy Level: Doctoral Course

Programs

Abbreviation	Description
BS CSE	BS Computer Science and Engineering
MS CSE	MS Computer Science and Engineering
PhD CSE	PhD Computer Science and Engineering

Course Goals

Master the fundamentals of symbolic methods in language processing tasks, such as natural language parsing.
Be competent with fundamental concepts for natural language processing and automatic speech recognition, such as "hidden Markov models".
Be competent with fundamental concepts in text-to-speech synthesis, such as concatenative synthesis and text analysis.
Be familiar with a finite state framework integrating all of speech processing.

Be familiar with a toolkit for text classification, part-of-speech tagging and sentiment mining.
Be familiar with methods of constructing speech recognition and synthesis systems.
Be exposed to current speech and language processing research.
Be exposed to toolkits for speech recognition and speech synthesis.

Course Topics

Topic	Lec	Rec	Lab	Cli	IS	Sem	FE	Wor
Course introduction, part-of-speech tagging	3.0							
HMMs, expectation maximization and search	3.0							
Parsing	3.0							
Word senses	3.0							
Language modeling	3.0							
Text classification and opinion mining	3.0							
Human hearing, acoustics, and phonetics	3.0							
Finite state transducers and automatic speech recognition toolkits	3.0							
Dynamic time warping and acoustic modeling	3.0							
Text analysis and speech synthesis	3.0							
Language processing in context (systems)	3.0							
Quizzes and in-class assignments	2.0							
Project presentations	4.0							

Representative Assignments

Exercises in part-of-speech labeling using a statistical tagger
Modifying a bottom-up chart parser to use a different rule invocation strategy.
Evaluating the effectiveness of sentiment classification.
Implementing Lesk's word sense disambiguation system.
Exercises using the AT&T FSM toolkit, building pronunciation models for a digit recognizer
Rescore word hypothesis lattice using different language models
Text normalization using FSMs

Grades

Aspect	Percent
Labs and homeworks	40%
Final project	30%
Exams (2 x 10%)	20%
Class Participation	10%

Representative Textbooks and Other Course Materials

Title	Author
Speech and Language Processing, 2nd Edition	D. Jurafsky and J. Martin
Speech Synthesis and Recognition, 2nd edition	J. Holmes and W. Holmes
Spoken Language Processing: A guide to theory, algorithms, and system development	X. Huang, A. Acero, and H.-W. Hon

ABET-EAC Criterion 3 Outcomes

Course Contribution		College Outcome
***	a	An ability to apply knowledge of mathematics, science, and engineering.
**	b	An ability to design and conduct experiments, as well as to analyze and interpret data.
***	c	An ability to design a system, component, or process to meet desired needs.
*	d	An ability to function on multi-disciplinary teams.
*	e	An ability to identify, formulate, and solve engineering problems.
	f	An understanding of professional and ethical responsibility.
*	g	An ability to communicate effectively.
**	h	The broad education necessary to understand the impact of engineering solutions in a global and societal context.
*	i	A recognition of the need for, and an ability to engage in life-long learning.
	j	A knowledge of contemporary issues.
***	k	An ability to use the techniques, skills, and modern engineering tools necessary for engineering practice.

BS CSE Program Outcomes

Course Contribution		Program Outcome
***	a	an ability to apply knowledge of computing, mathematics including discrete mathematics as well as probability and statistics, science, and engineering;
**	b	an ability to design and conduct experiments, as well as to analyze and interpret data;
***	c	an ability to design, implement, and evaluate a software or a software/hardware system, component, or process to meet desired needs within realistic constraints such as memory, runtime efficiency, as well as appropriate constraints related to economic, environmental, social, political, ethical, health and safety, manufacturability, and sustainability considerations;
*	d	an ability to function on multi-disciplinary teams;
*	e	an ability to identify, formulate, and solve engineering problems;
	f	an understanding of professional, ethical, legal, security and social issues and responsibilities;
*	g	an ability to communicate effectively with a range of audiences;
**	h	an ability to analyze the local and global impact of computing on individuals, organizations, and society;
*	i	a recognition of the need for, and an ability to engage in life-long learning and continuing professional development;
	j	a knowledge of contemporary issues;
***	k	an ability to use the techniques, skills, and modern engineering tools necessary for practice as a CSE professional;
*	l	an ability to analyze a problem, and identify and define the computing requirements appropriate to its solution;
**	m	an ability to apply mathematical foundations, algorithmic principles, and computer science theory in the modeling and design of computer-based systems in a way that demonstrates comprehension of the tradeoffs involved in design choices;
**	n	an ability to apply design and development principles in the construction of software systems of varying complexity.

Prepared by: Christopher Brew