

Topics in Language Processing: Spoken Dialog Systems (CS4390/5319)

Spring 2010

Monday & Wednesday, 4:30 – 5:50 , Computer Science room 321

Instructor: Nigel Ward
Office: Comp 206
Phone: 747-6827
E-mail: nigel@cs.utep.edu
Office Hours: Tu 3:30-4:30, Fr 10:30-11:30, also whenever my door is open, or by appointment

Motivation: In the past decade there has been an explosion in the creation and commercial deployment of spoke dialog systems user interfaces, especially for use over the telephone. These systems provide callers with voice access to information and allow them to perform transactions. Although standards, design tools, and design methods now exist, building a successful application is still a challenge. The course will briefly cover the underlying technology, speech recognition, but the focus will be on using this to produce usable products.

Course Objectives:

Students will acquire the knowledge and skills needed to create highly usable voice user interfaces: including:

- key properties of human language and communication
- constraints of the underlying technology
- the development process,
- implementation, in VoiceXML
- integration with web technologies (PHP, XML, XSLT)
- VUIs and rival technologies in the marketplace
- VUI usability fundamentals and heuristics

Students will also learn about emerging technologies and research issues, such as dialog strategy learning, adaptation to the user, and efficient turn-taking.

Students' knowledge and skills in the following areas will be reinforced:

- software engineering
- human-computer interaction
- formal language
- usability testing
- experiment design.

Format: Primarily lectures, augmented with in-class design exercises, student project presentations, and student-led discussions.

Textbook: *Voice User Interface Design*, Michael H. Cohen, James P. Giangola and Jennifer Balogh, Addison-Wesley, 2004 (henceforth VUID).

This will be supplemented by readings handed out in class.

Two other good books to own are *VoiceXML: Introduction to Developing Speech Applications*, James A. Larson, Prentice Hall, 2003, and *VoiceXML: 10 Projects to Voice-Enable Your Web Site*, Mark Miller, Wiley, 2002.

Prerequisites: This course is designed for graduate and senior-level students in Computer Science, Linguistics, and Psychology. No specific prerequisite knowledge is required, but programming experience or experience with formal notations will be very helpful. Students outside Computer Science will need permission from the instructor.

Assignments: There will be a number of structured assignments, designed to give experience with various usability engineering activities. Most assignments will be done in teams.

As a final project, individuals or teams will build a substantial dialog system. Past projects included an Auto Body Parts Finder, the SEAL Agenda and Events VUI, the Radio Paradise Song Request and Information Line, a Movie Recommendations and Rental service, a “UTEP Phonemine” Course Information Portal, an Advising Scheduling System, and a Remote Home Control system.

Grading: The weighting will be approximately 50% assignments, 30% examinations, and 20% final project. Graduate students will have one more assignment than undergraduates.

Assignments and tests will be challenging. Grading will be on a points-earned basis (points above zero), rather than a points-off basis (points below expectation). Letter grades will be assigned accordingly.

Tests will be closed-book, except that one single-sided page of hand-written notes may be brought in.

Students are expected to be punctual. Assignments due at the start of class will be collected after a one minute grace period; late assignments will receive at most two-thirds credit.

No make-up exams or assignments will be given except under the conditions set forth in the Catalog. Students are free to attend class or not, bearing in mind that absence may annoy other students, interfere with learning, and result in a lower grade.

Cooperation among students and among teams is encouraged, but not to the extent that it interferes with each individual’s understanding or with learning-by-doing. Help given and received from other students and sources should be noted in the assignment write-up. More generally, students will follow the UTEP Standards of Conduct, available at <http://studentaffairs.utep.edu/Default.aspx?tabid=4386> .

Important Dates:

Test 1	February 17
Test 2	March 29
Spring Break	March 15-19
Chavez Day	March 31
Final Exam	May 10, 4-6:45 pm

Course Web Site: <http://www.cs.utep.edu/nigel/dialog/>

Approximate Schedule

Part I Introduction (3.5 days)

1. Course Overview (1 day)

Course Objectives, Requirements
Historical Perspective
Possible Futures
Assignment A: Dissect a Voice User Interface (2hrs, 20 points)

2. Overview of VUI Design Issues (1 day)

(VUID Chapter 1)
VUIs vs GUIs
usability in voice applications

3. Technology Overview (1 day)

(VUID Chapter 2)
VUI components

4. Project Overview (.5 day)

Overview of Project
Development Case Study
Sample Projects and Ideas (think-time in class)
Assignment B: Project Proposals (2 hour; 20pts)

Part II: VoiceXML and Friends (8 days)

1. Basic VoiceXML (2 days)

Requirements for a VUI language; in-class exercise
(Larson Chapter 5)
VoiceXML Basics
BeVocal Café
Forms
Assignment U: "Hello world" (1 hour, 10pts)
Assignment C: a VUI-based drop-add form (3 hours, 40pts)

2. Advanced VoiceXML (2 days)

Modularization into multiple files
(VUID Chapter 16)
Grammar Writing; in-class exercise

3. Mixed Initiative (1 day)

(VUID Chapter 5)
Interaction Styles
Multi-slot forms and universals in VoiceXML
Multimodal Interaction, SALT
Assignment D: a restaurant bill advisor, etc. (3 hours, 20pts)

4. Dynamic Content (2 days)

Embedded Javascript
Parsing XML, using <data>
Generating VoiceXML from XML using XSLT

Assignment F: VoiceXML from an XML file or RSS feed (4 hours, 30pts)

5. Interfacing with a Back-end (1 day)

Generating VoiceXML server-side using PHP

Assignment G: VoiceXML and PHP (4 hours, 40 pts)

Part III Speech Industry Structure and Trends (2 days)

1. Customers, Markets, Companies, Teams, Opportunities (1.5 days)

Nortel Networks' Tom Chavez?

Assignment E: Examine a Speech Company or Industry Issue (2 hours, 10 pts)

2. System Quality (.5 day)

Usability

Evaluation

Metrics

Part IV: Design and Development (6 days)

1. Requirements (1 day)

(VUID Chapters 3, 4, 5, 7)

Requirements Gathering

Focus Groups; in-class exercise

Project Triage; in-class exercise

Assignment I: Project Charter (2 hr, 20 pts)

2. Detailed Design (.5 day)

(VUID Chapters 8 and 14)

Assignment J: High-Level Design or Quick Prototype (5 hours, 20 pts)

3. Design Principles, Design Elements, Usability (.5 day)

(VUID Chapters 9, 12, 13)

Cognitive load

Efficiency

Clarity

Naturalness

Accuracy

Assignment K: Quick Prototype or High-Level Design (4 hours, 20pts)

4. Prompt Design and Recording (2 days)

(VUID Chapter 17, readings)

Audio formats

Recording (KTEP studio tour?)

Prompt Design; in-class rewriting exercise

Prosody; in-class exercise on splicing and prosody

Voice Coaching

5. Persona Design and Implementation (2 days)

(VUID Chapter 6, readings)

Persona design (in-class exercise)

Social Psychology (Nass)

Assignment L Persona Tuning(3 hours, 15 pts)

6. Development (1 day)

(VUID Chapters 10, 11, 15, 16, 18, 19)
Grammar Development and Testing
Application Development and Testing
Data-driven Tuning

Part V: Technical Underpinnings (4 days)

1. Sound, Speech, and Digital Signal Processing (1 day)

Articulation
Time- and Frequency-Domain Representations
Phonemes and Features

2. Speech Recognition (1 day)

(Jurafsky&Martin 2, Chapter 9)
Probabilistic Acoustic Models
Language Models
Argmax and Search
Dictation vs. Dialog

3. Natural Language Understanding (.5 day)

Call routing; in-class exercise on decision-making

4. Dialog Managers (1.5 days)

(Jurafsky&Martin 2, Chapter 24)
Dialog Acts
Error Handling Strategies
Information State

Part VI: Other (5 days)

1. Research Topics (3 days)

Turn-taking
Tutoring Systems
Animated Agents
Other topics of interest
Assignment H: present a research paper (4 hrs, 30 pts; graduate students only)
Assignment Z: develop a hypothesis and experiment design (4 hours, 20 pts grads only)

2. Project and Experiment Presentations (1 day)

Assignment M: Final Reports and Presentations (20 hours, 60pts)

3. Review (1 day)

Assignment N: A Question for the Exam (1 hour, 5pts)

The time estimates for the assignments are for an efficient person working with a well organized team.

Graduate students will be held to a higher standard for Assignment M, and will have two additional Assignments: Y and Z.