Syllabus COSI 115: Spoken Dialog Design

Overview

Interactive applications have become ubiquitous around the world on phones and other devices. Since voice is the most natural medium for human communication, spoken dialog is becoming an essential part of the interface. However, creating an effective spoken dialog application requires more than just programming skills. It requires knowledge from many disciplines including linguistics, artificial intelligence, computer-human interaction, and computational linguistics. This course will bring together the essential elements of these fields and the software skills and tools required to build an effective dialog system and guide students through hands-on projects applying that knowledge to real applications.

Learning Objectives

At the end of the course students will

- Understand the basic principles of the fields that are underlie spoken dialog, including:
  - Fundamental linguistic principles of discourse,
  - Artificial Intelligence techniques for plan recognition and task execution,
  - Computational models for recognizing intentions and coreference resolution
  - Human-computer Interaction (HCI) and Voice User Interface (VUI) design
- Understand the architecture of spoken dialog systems and the capabilities and limitations of the software components required to execute the application, such as
  - speech recognition
  - speech synthesis
  - dialog modules
- Be able to apply this knowledge to building spoken dialog applications using industry and research tools.

Required Reading

There is no required textbook for the course. The course will rely mostly on published papers and online resources ranging from early papers on the fundamentals of dialog to current research. The instructor will also make available lecture notes/slides on the topics covered on class. Example of published articles to be covered include the following:


Schedule

<table>
<thead>
<tr>
<th>Week</th>
<th>Morning: Theory</th>
<th>Afternoon: Applications</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Overview: Speech Recognition Applications, Speech Industry</td>
<td>Components of a spoken dialog system: Speech recognition, Speech Synthesis</td>
</tr>
<tr>
<td></td>
<td>Human Conversation: Discourse and Dialog</td>
<td>Speech recognition architecture, Dialog manager</td>
</tr>
<tr>
<td>2</td>
<td>Discourse structure</td>
<td>Dictionaries and grammars</td>
</tr>
<tr>
<td></td>
<td>Anaphora and reference resolution</td>
<td>Statistical language modeling</td>
</tr>
<tr>
<td>3</td>
<td>Plan recognition and task structure</td>
<td>Speech performance evaluation</td>
</tr>
<tr>
<td></td>
<td>Dialog and belief representation</td>
<td>Data vs. Knowledge</td>
</tr>
<tr>
<td>4</td>
<td>Dialog Design: Use cases</td>
<td>Natural Language Processing</td>
</tr>
<tr>
<td></td>
<td>Dialog Design: Clarification Dialogs and error recovery</td>
<td>Advanced dialog architectures</td>
</tr>
<tr>
<td>5</td>
<td>Dialog system Evaluation</td>
<td>Multimodal applications</td>
</tr>
<tr>
<td></td>
<td>Case studies</td>
<td>Case studies</td>
</tr>
</tbody>
</table>

Grading

50% Programming assignments: There will be 4-5 programming assignment exercising the principles covered in the lectures that will expose students to a variety of programming languages and tools that are typically used in spoken dialog development in research and industry.

30% Homework and take home quizzes: Periodic homework assignments and take home quizzes will allow students to synthesize the knowledge from readings and lectures and consider the application of the principles in multiple contexts.

20% Class participation: Students will be required to participate in class discussions, work in groups, and submit to class blog discussions.