Advanced Speech Lab

Course 11-753
Tanja Schultz & Thomas Schaaf
Goals

- Learn how a large vocabulary speech recognition system works
- Learn the steps involved in building an acoustic model
- Learn how to build a language model
- Learn how to decode an audio signal
- Learn how to measure performance
General overview

- Learn basic Tcl/Tk
- Build a training database
  - What data?
- Components of a speech recognition system
  - Preprocessing
  - Phonemes, Dictionary
  - HMM topology
  - Probability density for acoustic modeling
- Choose preprocessing of audio signal
- Initial description for context independent system
General overview (2)

• Train a context independent system
  – Linear Discriminant Analysis
  – Sample extraction
  – K-means initialization
  – Label/Viterbi training

• Build a 3-gram language model
  – What is back-off
  – How to use classes

• Do a first decoding with context independent acoustic models
General overview (3)

- Move to context dependent phoneme models
- Build a context dependent phonetic decision tree
  - Optimization criteria
  - Statistic required
  - How splitting is performed
- Train context dependent acoustic model
- Test context dependent model
General overview (4 optional)

- Adaptation methods
  - VTLN
  - MLLR
- Cross Adaptation
- Speaker Adaptive Training (SAT)
- Confidence Measure
- Rover
## Schedule (tentative)

<table>
<thead>
<tr>
<th>Date</th>
<th>Topic</th>
</tr>
</thead>
<tbody>
<tr>
<td>11/17/2005</td>
<td>Introduction, get started</td>
</tr>
<tr>
<td>11/24/2005</td>
<td>Thanksgiving</td>
</tr>
<tr>
<td>12/01/2005</td>
<td>Janus and Tcl/Tk Introduction, Objects</td>
</tr>
<tr>
<td>12/08/2005</td>
<td>Pronunciation dictionary, database, word error rate, more Tcl/Tk</td>
</tr>
<tr>
<td>12/15/2005</td>
<td>Preprocessing</td>
</tr>
<tr>
<td>12/22/2005</td>
<td>Preprocessing, Acoustic Modeling</td>
</tr>
<tr>
<td>12/29/2005</td>
<td>Winter Break</td>
</tr>
<tr>
<td>01/05/2006</td>
<td>Acoustic Modeling</td>
</tr>
<tr>
<td>01/12/2006</td>
<td>Training CI</td>
</tr>
<tr>
<td>01/19/2006</td>
<td>Training CD</td>
</tr>
<tr>
<td>01/26/2006</td>
<td>Language Modeling</td>
</tr>
<tr>
<td>02/02/2006</td>
<td>Search, Scoring</td>
</tr>
<tr>
<td>02/09/2006</td>
<td>Advanced Topics (optional)</td>
</tr>
</tbody>
</table>