Deliverable #2:
Question Classification

Thomas Phan
Jason Shaw
Varden Wang
Focus: Feature Design

- Surface
  - Unigrams
- Syntactic
  - POS tags
- Semantic
  - WSD
  - NER
  - WordNet

- All classification done using Mallet’s MaxEnt
Surface Features

- Unigrams only
- No stemming or case neutralizing
- Weak tokenization

- Classification on Li and Roth data:
  - Coarse: 84.6%
  - Fine: 75.8%
Syntactic Features

- Words tagged with POS
- NLTK tagger trained on Penn Treebank

Classification on Li and Roth data:
  - Coarse: 85.6%
  - Fine: 76.8%
Semantic Features

- Word sense disambiguation
- Unsupervised Lesk disambiguator using WordNet
  - Coarse: 85.4%
  - Fine: 73.8%

- Named entity recognition
- Stanford NER package
- Identified PERSONs, ORGANIZATIONs, and LOCATIONs
  - Coarse: 37.8%
  - Fine: 12.0%
Semantic Features (cont’d)

- Classification sense lists from WordNet
- Map Question Hierarchy to high-level senses
- Test for presence of these senses as hypernyms
- Used in conjunction with other features, performs better than WSD alone
  - Coarse: 75.8%
  - Fine: 64.6%

<table>
<thead>
<tr>
<th>Class</th>
<th>Sense</th>
</tr>
</thead>
<tbody>
<tr>
<td>ENTITY</td>
<td>entity#1</td>
</tr>
<tr>
<td>food</td>
<td>food#1, food#2</td>
</tr>
<tr>
<td>plant</td>
<td>plant#1</td>
</tr>
<tr>
<td>HUMAN</td>
<td>human#2</td>
</tr>
<tr>
<td>individual</td>
<td>individual#1</td>
</tr>
<tr>
<td>LOCATION</td>
<td>location#1</td>
</tr>
<tr>
<td>city</td>
<td>city#1, city#2</td>
</tr>
</tbody>
</table>
Putting it all together

- Optimal results when using unigrams, POS tags, and classification sense lists

<table>
<thead>
<tr>
<th>Training data</th>
<th>Test data</th>
<th>Type</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Li and Roth</td>
<td>TREC 10</td>
<td>Coarse</td>
<td>88.0%</td>
</tr>
<tr>
<td>Li and Roth</td>
<td>TREC 10</td>
<td>Fine</td>
<td>80.2%</td>
</tr>
<tr>
<td>Li and Roth, TREC 2004</td>
<td>TREC 10</td>
<td>Coarse</td>
<td>88.0%</td>
</tr>
<tr>
<td>Li and Roth, TREC 2004</td>
<td>TREC 2005</td>
<td>Coarse</td>
<td>83.5%</td>
</tr>
</tbody>
</table>

- The addition of NER tagging improved fine classification to 81.6%
Potential improvements

- Finer-grained NER categories
- Head chunking
- Classification sense lists without WSD