

PROJ2: Building an ASR System

Julia Hirschberg

CS 4706

Goal

- Design and build your own speech understanding system for your domain
- Your system will
 - Take an input utterance
 - Transcribe it automatically
 - Convert the transcription into a semantic representation corresponding to the domain concepts (degrees of freedom) in your domain
- Your system will consist of two components: an ASR system and an Understanding system

ASR System

- You will be given a skeleton script that call an ASR system built using the HTK (an HMM toolkit)
 - Acoustic models are already trained on TIMIT, BDC, and the Columbia Games corpora
 - System input will be a wav file (audio format: mono, sample rate: 16Khz)
 - System output will be the automatic transcript in mlf file format, e.g.

```
#!MLF!  
"/test2.rec"  
5100000 5400000 I -250.811493  
5400000 6300000 NEED -767.471863  
6300000 7100000 TO -789.156311  
7100000 9100000 GO -1631.608887  
9100000 10000000 TO -913.183228
```

Grammar

- To build your system you need to create a grammar that handles your domain
 - Constrains the recognition output to conform to queries in your domain
- \$city = BOSTON | NEWYORK | WASHINGTON | BALTIMORE;
- \$time = MORNING | EVENING;
- \$day = FRIDAY | MONDAY;
(SENT-START
(((WHAT TRAINS LEAVE) | (WHAT TIME CAN I TRAVEL) | (IS THERE A TRAIN)) (FROM|TO) \$city (FROM | TO) \$city ON \$day [\$time])
SENT-END)

Multiple Acoustic Models

- ASR system has been trained with different numbers of Gaussians per HMM state
 - Experiment with these different HMMs to decide which works best in your domain.
- Detailed instructions on how to build and run the system in [PROJ2 description](#)

Generating Concept Tables

- You must write a script to transform the ASR output into a semantic representation, e.g. translate

```
#!MLF!#
```

```
"*/test2.rec"
```

```
5100000 5400000 I -250.811493
```

```
5400000 6300000 NEED -767.471863
```

```
6300000 7100000 TO -789.156311
```

```
7100000 9100000 GO -1631.608887
```

```
9100000 10000000 TO -913.183228
```

```
10000000 12400000 BALTIMORE -1923.127319
```

```
13300000 14000000 FROM -679.068176
```

```
14000000 14600000 WASHINGTON -560.649719
```

```
15900000 16500000 ON -547.398132
```

```
16500000 18500000 MONDAY -1689.119995
```

```
18500000 20200000 EVENING -1382.312256
```

- Into this

Departure city: Baltimore

Destination: Washington

Day: Monday

Time: Evening

- You'll be graded on concept accuracy and grammar coverage
- More information on the HTK toolkit, including the grammar format can be found at <http://www.csie.ntu.edu.tw/%7Eb6506053/doc/htkbook.pdf>