

CSCI 5832

Natural Language Processing

Lecture 10
Jim Martin

2/20/07

CSCI 5832 Spring 2007

1

Today: 2/20

- **Review**
 - POS Tagging
 - HMMs and Viterbi
- **Break**
- **Syntax and Context-free grammars**

2/20/07

CSCI 5832 Spring 2007

2

Review

- **Parts of Speech**
 - Basic syntactic/morphological categories that words belong to
- **Part of Speech tagging**
 - Assigning parts of speech to all the words in a sentence

2/20/07

CSCI 5832 Spring 2007

3

Probabilities

- We want the best set of tags for a sequence of words (a sentence)
- W is a sequence of words
- T is a sequence of tags

$$\arg \max P(T | W) = P(W | T)P(T)$$

2/20/07

CSCI 5832 Spring 2007

4

So...

- We start with

$$\arg \max P(T | W) = P(W | T)P(T)$$

- And get

$$\arg \max \prod_{i=1}^N P(w_i | t_i) * P(t_1) * \prod_{i=2}^N P(t_i | t_{i-1})$$

2/20/07

CSCI 5832 Spring 2007

5

HMMs

- This is an HMM

$$\arg \max \prod_{i=1}^N P(w_i | t_i) * P(t_1) * \prod_{i=2}^N P(t_i | t_{i-1})$$

- The states in the model are the tags, and the observations are the words.
 - The state to state transitions are driven by the bigram statistics
 - The observed words are based solely on the state that you're in

2/20/07

CSCI 5832 Spring 2007

6