

CSCI 5832

Natural Language Processing

Jim Martin
Lecture 27

5/2/08

1

Today 5/1

- Quiz 3 review
- Review
 - ◆ Reprise of the first three quizzes
 - ◆ MT

5/2/08

2

Quiz 3

2 John requested....

- ◆ Introduce an event variable with an \exists quantifier
 - $\exists e \text{ Request}(e)$
- ◆ Introduce a role for reach thematic role specified
 - $\text{Requester}(e, \text{John})$

5/2/08

3

Quiz 3

3. VP \rightarrow V NP NP $V.\text{sem}(\text{NP}.\text{sem}, \text{NP}.\text{sem})$

V \rightarrow booked

$\lambda x, y \lambda z$

$\exists e \text{ Booking}(e) \wedge \text{Booker}(e, z) \wedge$

$\text{Bookee}(e, x) \wedge \text{Booked}(e, y)$

5/2/08

4

Quiz 3

4. True

5. Named entities, locations, temporals,
amounts, events,...

sequence classification for NER,
locations, temporals, amounts

capitalization, lists, lemmas of
surrounding words, etc.

5/2/08

5

Quiz 3

6a Hobbs.

6b Clustering

5/2/08

6

Final

- Right here. Monday May 5, 1:30 to 4:00
- You can bring 3 pages of cheat sheets
- Major parts
 - ◆ Words and word sequence models
 - ◆ Syntax and parsing
 - ◆ Semantics
 - ◆ Discourse
 - ◆ MT

5/2/08

7

Quiz 1

- Readings
 - ◆ Chapter 2: All
 - ◆ Chapter 3:
 - Skip 3.4.1 and 3.12
 - ◆ Chapter 4
 - Skip 4.7, 4.9, 4.10 and 4.11
 - ◆ Chapter 5
 - Read 5.1 through 5.5

5/2/08

8

Quiz 1

- Finite state methods
 - ◆ Recognition
 - ◆ Parsing
 - ◆ Cascades of multiple tapes
- Some morphology
 - ◆ Derivational vs. inflectional
 - ◆ Regulars vs. Irregulars
- Parts of speech and tagging
 - ◆ HMM tagging
 - ◆ Sequence labeling

5/2/08

9

Quiz 1

- Basic probability stuff
 - ◆ Chain rule
 - ◆ Markov assumption
 - ◆ Hidden states
 - ◆ Parts lists
 - Transition probs
 - Observation probs
 - Initial state probs

5/2/08

10

Quiz 2

- Quiz
 - ♦ Chapter 12: 12.1 through 12.6
 - CFGs, Major English phrase types, problems with CFGs, relation to finite-state methods
 - ♦ Chapter 13: All except 13.4.3
 - CKY, Earley, partial parsing, sequence labeling
 - ♦ Chapter 14: 14.1 through 14.6.1
 - Basic prob CFG model, getting the counts, prob CKY, problems with the model, lexicalization, and grammar rewriting

5/2/08

11

Quiz 3 Material

- Ch 17
 - ♦ Basic FOL Event representations
- Ch 18 (18.1 to 18.3 and 18.6)
 - ♦ Rule to rule semantic attachments
- Ch 20 (20.1 to 20.5)
 - ♦ WSD
- Ch 22 (all)
 - ♦ IE
- Ch 21 (21.3 to 21.8)
 - ♦ Co-reference

5/2/08

12

Big Picture Stuff

- Paradigms
 - ◆ State-space search
 - ◆ Dynamic programming
 - ◆ Probability models
 - Bayesian/Noisy channel model
- Frameworks
 - ◆ Cascades of transducers
 - ◆ IOB encoding
 - ◆ Rule to rule semantics

5/2/08

13

Algorithms

- Deterministic and non-deterministic recognition
- HMMs
 - ◆ Viterbi
 - ◆ Forward
 - ◆ EM
- Sequence classification

5/2/08

14

Algorithms

- CKY
- Earley
- IOB labeling for chunking

5/2/08

15

Algorithms

- WSD
 - ◆ Training classifiers
 - ◆ Using dictionaries
 - ◆ Clustering
- IE
 - ◆ Sequence classification
 - ◆ Relational classifiers

5/2/08

16

MT

- Noisy channel model
 - ◆ Bayesian inversion
 - ◆ Word based models
 - ◆ Phrase based models
 - ◆ EM for alignment