

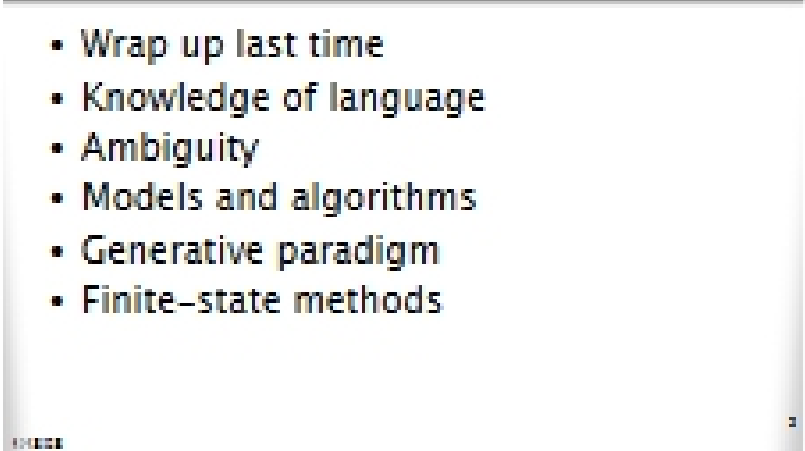
CSCI 5832 Natural Language Processing

Lecture 2
Jim Martin



Today 1/17

- Wrap up last time
- Knowledge of language
- Ambiguity
- Models and algorithms
- Generative paradigm
- Finite-state methods



Course Material

- We'll be intermingling discussions of:
 - + Linguistic topics
 - E.g. Morphology, syntax, discourse structure
 - + Formal systems
 - E.g. Regular languages, context-free grammars
 - + Applications
 - E.g. Machine translation, information extraction



Linguistics Topics

- Word-level processing
- Syntactic processing
- Lexical and compositional semantics
- Discourse processing

1/1/2008

4

Topics: Techniques

- Finite-state methods
- Context-free methods
- Augmented grammars
 - Unification
 - Lambda calculus
- First order logic
- Probability models
- Supervised machine learning methods

1/1/2008

5

Topics: Applications

- Small
 - Spelling correction
 - Hyphenation
- Medium
 - Word-sense disambiguation
 - Named entity recognition
 - Information retrieval
- Large
 - Question answering
 - Conversational agents
 - Machine translation
- Stand-alone
- Enabling applications
- Funding/Business plans

1/1/2008

6

Just English?

- The examples in this class will for the most part be English
 - + Only because it happens to be what I know.
- This leads to an (over?)-emphasis on certain topics (syntax) to the detriment of others (morphology) due to the properties of English
- We'll cover other languages primarily in the context of machine translation

11/10/04

Commercial World

- Lot's of exciting stuff going on...



11/10/04

Google Translate



11/10/04
