C SC 620
Advanced Topics in Natural Language Processing

Sandiway Fong
Natural Language Processing (NLP)

• Very broad, interdisciplinary field:
  – At the intersection of linguistics, computer science, cognitive science …

• Areas include:
  – Parsing methods and grammar formalisms
  – Machine translation
  – Large-scale resources: corpora, dictionaries, ontologies
  – Information retrieval, classification …
  – Document processing and summarization
  – etc.
Scope of Class

• Another computational linguistics seminar class this semester
  – LING 696G
  – Time: Mondays 3-5:30 pm
Administrivia

- Email: sandiway@email.arizona.edu
- Course Homepage:
  - Notes, announcements, papers, readings
  - http://linguistics.arizona.edu/~sandiway/csc620
    • not yet set up …
- Requirements:
  - Advanced class: seminar format
  - Summarize and present one paper in class for discussion
  - Term project or paper
Administrivia

• Email me:
  – Name
  – Background
  – Why are you interested in NLP?
  – Specific topics you’d like to see covered?
  – Preferences for class time
    • Available times
    • One class per week or two?
Start with…

**WordNet**

a lexical database for the English language

- WordNet:
  - a large-scale ontology-based network for nouns, verbs, adjectives and adverbs
  - Freely available
  - http://www.cogsci.princeton.edu/~wn/
  - Download WordNet to your computer
  - Get the 5 core papers
    - *we’ll begin looking at them next time*