December 9, 2004

Wrap-up/review

Notes on the exam/final projects

- Friday 12/17, 4:30-6:20 pm, Denny 109

- Open book, open notes, closed web

- Cumulative – with an emphasis on topics covered since the midterm.

- Projects/papers are due by 10:00 on Friday 12/17, via ESubmit

- Documentation guidelines are on the web

- I will be holding office hours next Tuesday

Notes on the review

- Remind ourselves where we've been.

- Look for connections and generalizations across the different topics we've covered.

- Congratulate ourselves on how much we've learned.

Topics covered before the midterm

- Regular expressions

- Finite state automata

- Finite state transducers

- Morphology & morphological parsing

- CFG

- Syntactic parsing

- Feature structures

- Unification

- Parsing with unification

Topics covered since the midterm (1/2)

- Probabilistic parsing:

  - PCFGs, probabilistic chart parsing, acquiring probabilities, probabilistic lexicalized PCFGs

- Evaluating parsers: precision and recall

- Phonetic alphabets, FSTs for phonological rules, two-level phonology

- TTS, pronunciation dictionaries/lexica

- Pronoun resolution:

  - Hard and soft constraints

- Algorithms: Salience factors, tree search, centering theory

Topics covered since the midterm (2/2)

- Text coherence, coherence resolution

- Lexical semantics:

  - homonymy, polysemy, synonymy, hyponymy

  - WordNet, FrameNet

  - theta roles, selectional restrictions

- Dialogue managers, turn taking, grounding

Synthesis (1/3)

- Computational linguistics research results usually fall into one of three categories:

  - Algorithms (often associated with software)

  - Resources

  - Methods of evaluation

- Examples?

Synthesis (2/3)

- The algorithms we've seen this quarter seem to fall into two classes:

    - Those that stress conceptual elegance/conciseness (and possibly efficiency)

    - Those that seem to consist of a baroque set of constraints/procedures carefully put together

- Examples?

- Why might this be?

Synthesis (3/3)

- How is computational linguistics research influenced by practical applications?

- How is computational linguistics research influenced by practical considerations?

- Are these influences positive or negative?

- Should they be extended to other areas of linguistics?

Course goals (1/2)

- Become familiar with computational linguistic resources, and how they are applied in research in compling and other subfields

- Have a rough sense of the state of the art in this subfield

- Be able to conceptualize problems from the perspective of computational linguistics

Course goals (2/2)

- Examples of how resources are used? Might be used?

- Any cases where the state of the art is more or less advanced than what you expected coming in?

- How does compling conceptualize problems differently from other subfields of linguistics/CS?

To learn more...

- Compling lab ("TreeHouse") meetings – subscribe to compling mailing list on mailman for information about next quarter.

- Winter 05: Intro syntax for compling, Spring 05: Grammar engineering

- Colloquia, conferences, and summer schools (ESSLLI, NASSLLI)