Natural Language Processing

Lecture 18

Today's Outline

Announcements (exams, Friday ISP seminar)

Homework 3

Finish Reference Resolution Algorithms

Discourse Analysis

- text coherence
- discourse structure

ISP Seminar

Nov 30 (THIS Friday!): Antal van den Bosch (WhizBang! Labs-Research and Tilburg University, The Netherlands)

Current issues in machine learning of natural language: memory and representation

In machine learning, natural language processing tasks have attracted a markedly large amount of interest as applications, with mutual benefits. For machine learning, language data has been used for symbolic feature representation and classification research. For computational linguistics, machine learning has offered a way out of its knowledge acquisition bottleneck. This talk centers around two current issues relevant to both communities: memory and representation. As for memory, empirical evidence exists showing that retaining full memory of experienced language data in machine-learned language processing systems outperforms abstraction-based (minimal-description-length driven) learning, As for representation, there is a growing feeling that discovering good feature representations yields much shorter pathways to developing good machine-learned NLP systems than comparing machine learning algorithms. The talk will summarize empirical findings and open up discussion about their meaning to both scientific communities.

HW3

Due: December 3, 2001

- (5 pts) Simplify the following formulas using lambda reduction:
  - ((lambda x (P x)) A)
  - ((lambda x ((lambda y (P y)) x)) A)
- (15 pts) Exercise 15.4, Jurafsky and Martin (p. 586). Be sure to show a derivation to validate your answer.
- (20 pts) Exercise 16.6, Jurafsky and Martin (p. 629), but just for the verb "drink".
- (20 pts) Exercise 16.7, Jurafsky and Martin (p. 629). For your examples, use the American portion of the CoBuildDirect Corpus at http://titania.cobuild.collins.co.uk/form.html. Try queries such as drink@/VERB.
- (20 pts) Exercise 17.2, Jurafsky and Martin (p. 664). Use as your corpus sentences 16.1, 16.2, 16.3, 16.4, 16.7, and 17.11 from Chapters 16 and 17.
- (20 pts) Exercise 17.3, Jurafsky and Martin (p. 664).

Review: Reference Resolution

John went to Bill's car dealership to check out an Acura Integra.
He looked at it for about half an hour.

Reference Phenomena

Constraints on Coreference

Pronoun Interpretation Preferences

Pronoun Resolution Algorithm

- discourse model update
- Lappin and Leass
- BFP centering

Text Coherence

Cohesion

- referring expressions
- ellipsis, conjunction, lexical, ...
- see *Cohesion in English* by Halliday and Hasan

Coherence

- John hid Bill's car keys. He was drunk
- ? John hid Bill's car keys. He likes spinach.
- ? Bill went to see his mother. The trunk is what makes the bonsai, it gives it both its grace and power.

Coherence Relations: Hobbs

Result

- Infer that the state or event asserted by the meaning of the first sentence (S_0) causes or could cause the state or event asserted by the meaning of the second sentence (S_1).
- John bought an Acura. His father went ballistic.

Explanation

- Infer that the state or event asserted by S_1 causes or could cause the state or event asserted by S_0.
- John hid Bill's car keys. He was drunk.

Elaboration

- Infer the same proposition P from the assertions of S_0 and S_1.
- John bought an Acura this weekend. He purchased a beautiful new Integra for 20 thousand dollars at Bill's dealship on Saturday afternoon.

Hobbs Relations (continued)

Occasion

- A change of state can be inferred from the assertion of S_0, whose final state can be inferred from S_1, or vice-versa (with initial state).
- John bought an Acura. He drove to the ballgame.

An Inference-Based Algorithm

Abduction

- A implies B
- B
- infer A

Example

- "All Acuras are fast. John's car is fast."
- abductively infer "John's car is an Acura."

Abduction is unsound

- defeasible
- "John's car is a Porsche though."

Furthermore, want the *best* explanation

- probabilities
- heuristics
- hybrid (Hobbs)

Example Trace

John hid Bill's car keys. He was drunk.

Coherence Relation Axioms

- $\forall e_i, e_j Explanation(e_i, e_j) \Rightarrow CoherenceRel(e_i, e_j)$
- $\forall e_i, e_j cause(e_j, e_i) \Rightarrow Explanation(e_i, e_j)$

World Knowledge Axioms

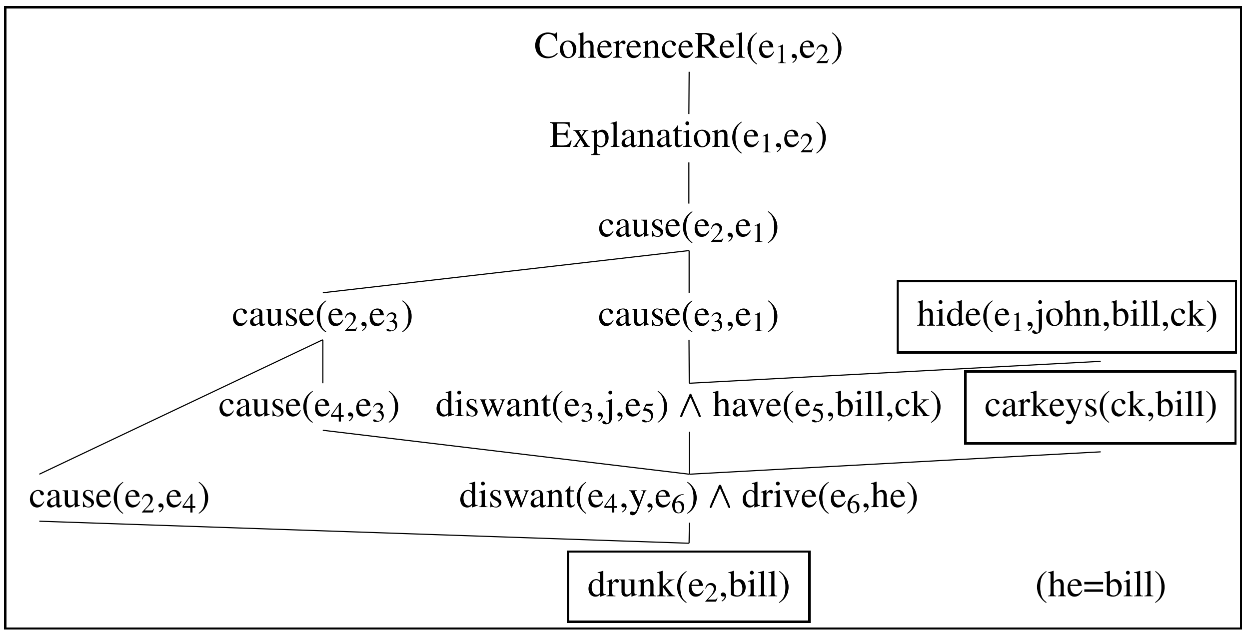
- $\forall x, y, e_i drunk(e_i, x) \Rightarrow \exists e_j, e_k diswant(e_j, y, e_k) \wedge drive(e_k, x) \wedge cause(e_i, e_j)$
- $\forall x, y, e_i, e_k diswant(e_j, y, e_k) \wedge drive(e_k, x)$
- $\Rightarrow \exists z, e_l e_m diswant(e_l, y, e_m) \wedge have(e_m, x, z) \wedge carkeys(z, x) \wedge cause(e_j, e_l)$
- $\forall x, y, z, e_i, e_j diswant(e_l, y, e_m) \wedge have(e_m, x, z)$
- $\Rightarrow \exists e_n hide(e_n, y, x, z) \wedge cause(e_l, e_n)$
- $\forall e_i, e_j, e_k, cause(e_i, e_j) \wedge cause(e_j, e_k) \Rightarrow cause(e_i, e_k)$

Utterance Axioms

- $hide(e_i, John, Bill, ck) \wedge carkeys(ck, Bill)$
- $drunk(e2, he)$

Deriving Coherence

John hid Bill's car keys. He was drunk.



Observations

The coherence derivation causes the inference of implicit information.

The given axioms do not support the coherence of "John hid Bill's car keys. He likes spinach."

Other Formalizations

Rhetorical Structure Theory (RST) (Ch. 20)

- primarily used for text generation

Plan-Based Relations (Ch. 19)

- Could you mount a magtape for me?
  - a discourse plan to introduce a domain plan to mount a tape
- It's tape 1.
  - specification of a parameter in the domain plan
- see "Linguistic Coherence: A Plan-Based Alternative" (Litman, 86)

More Observations

Coherence and Coreference

- John hid *Bill's* car keys. *He* was drunk
- *John* lost Bill's car keys. *He* was drunk

Coherence and Connectives (aka discourse markers, cue phrases)

- linguistic signals can be used to *constrain* the search
  - *because*: explanation
  - *and*: parallel, occasion, result
- however, connectives do not *create* coherence
  - ?? John hid Bill's car keys because he likes spinach.

Coherence and Applications

- e.g., summarization could merge or eliminate elaborations

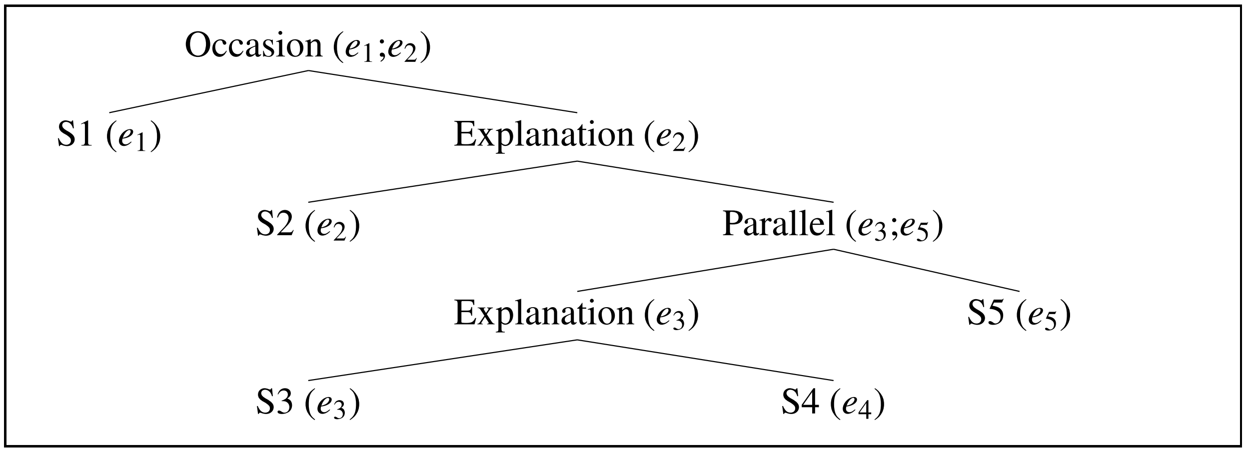
Unfortunately, the knowledge-based approach is not yet robust, efficient, practical, ...

Discourse Structure

A (tree) structure of discourse segments (locally coherent utterances)

Coherence Relation-Based Example

- John went to the bank to deposit his paycheck. (S1)
- He then took a train to Bill's car dealership. (S2)
- He needed to buy a car. (S3)
- The company he works for now isn't near any public transportation. (S4)
- He also wanted to talk to Bill about their softball league. (S5)



Discourse Structure (cont).

Structure of discourse segments (task and an interruption):

- E: So you have the engine assembly finished.
- E: Now attach *the rope*.
  - E: By the way, did you buy <u>the gas can</u> today?
  - A: Yes.
  - E: Did <u>it</u> cost much?
  - A: No.
  - E: OK, good.
- E: Have you got *it* attached yet?

Recall how this was mentioned when discussing recency for reference resolution.

Note presence of cue phrases (*by the way, ok*)

New Axioms

$\forall w, esentence(w, e) \Rightarrow Segment(w, e)$

$\forall w2, w2, e1, e2, eSegment(w1, e1) \wedge Segment(w2, e2) \wedge$
$CoherenceRel(e1, e2, e) \Rightarrow Segment(w1w2, e)$

$\exists eSegment(W, e)$

Observations

Original motivation: linguistic necessity

Recent applications

- summarization
- information retrieval
- audio browsing (broadcast news, email)
- text generation
- text to speech

Non-hierarchical structural analyses are also useful

Current topics include tagging reliability, statistical approaches

Local expertise

- research of Litman, Wiebe
- papers in my class next semester

Will be discussed further in Chapter 19

For Next Time

Chapter 19

Acknowledgements

Some of this material was taken from the course web pages of Drago Radev.