November 18, 2003

Chapter 18.2-18.5

Text Coherence

Overview

- Leftovers: Centering Theory
- Leftout: FOPL
- Text coherence
- Coherence resolution
- Inference
- Inference-based coherence resolution
- Discourse structure

Review: Pronoun resolution

- Some “hard constraints” on possible antecedents
- Many “soft” constraints, largely salience factors and parallelism
- “Soft” constraints can conflict
- Pronoun resolution algorithms must encode constraints and mediate between them.

Centering Theory

- Discourse model keeps track of entity being "centered" on in each utterance.
- Every utterance (U_n) has:
  - $C_b(U_n)$ the '*backward looking center*', which is the entity being centered on after U_n is interpreted.
  - (Exception: For the first utterance in a discourse is undefined.)
  - $C_f(U_n)$ the '*forward looking centers*', an ordered (e.g., by grammatical role) list of all the entities referred to in U_n.

Centering Theory

- Every utterance (U_n) has:
  - $\ldots$
  - $C_b(U_{n+1}) =$ the highest ranked element of $C_f(U_n)$ mentioned in U_{n+1} 'by definition'.
  - $C_p(U_n)$ the *preferred center*, the highest ranking element of $C_f(U_n)$.

Intersentential relationships

	$C_b(U_2) = C_b(U_1)$ or undefined $C_b(U_1)$	$C_b(U_2) \neq C_b(U_1)$
$C_b(U_2) = C_p(U_2)$	Continue	Smooth-Shift
$C_b(U_2) \neq C_p(U_2)$	Retain	Rough-Shift

Rules

- Rule 1: If any element of $C_f(U_n)$ is realized by a pronoun in utterance U_{n+1}, then $C_b(U_{n+1})$ must be too.
- Rule 2: Transition states are ordered:

  Continue $\rangle$ Retain $\rangle$ Smooth-Shift $\rangle$ Rough-Shift

Algorithm

- Generate possible C_b-C_f combinations for each possible set of reference assignments.
- Filter by constraints (e.g., syntactic coreference constraints, selectional restrictions, centering rules and constraints).
- Rank by transition orderings.
- Pick the reference assignment that gives the most preferred relation in Rule 2.

First Order Predicate Logic

- Representation of sentence meanings, assuming someone else has figured out lexical meanings
- $\rightarrow$ The meaning of life is life$'$
- Represent predicate-argument relations, boolean connectives between predicates, universal and existential quantification.
- Rules of reasoning over FOPL are statements well-studied

FOPL: Building blocks

- Predicate names
- Variables
- Boolean connectives: $\wedge$, $\vee$, $\Rightarrow$, $\neg$
- Quantifiers: $\forall$, $\exists$

FOPL: Quantifier scope

- Quantifiers bind particular variables, and take scope over everything to the right of them in an expression (modulo parentheses)
- Quantifiers bind every instance of their variables in their scope.
- Variables outside the scope of any quantifier are ‘free’

FOPL: Examples

- A dog barked: $\exists x, e_i$ dog(x), barked(e_i, x), past(e_i)
- All dogs bark: $\forall x \exists e_i$ dog(x) $\Rightarrow$ bark(e_i, x)
- Kim's friend left: $\exists x, e_i$ friend-of(x, k), leave(e_i, x), past(e_i)
- Everyone loves someone: $\forall x \exists e_i, y$love(e_i, x, y)

  $\exists y \forall x \exists e_i$love($e_i, x, y$)

Text coherence

- John hid Bill’s car keys. He was drunk.
- #John hid Bill’s car keys. He likes spinach.

Coherence relations (1 of 2)

- Result: Infer that the state or event asserted by S_0 causes or could cause the state or event asserted by S_1.
- Explanation: Infer that the state or event asserted by S_1 causes or could cause the state or event asserted by S_0.
- Parallel: Infer $p(a_1, a_2, \ldots)$ from the assertion of S_0 and $p(b_1, b_2, \ldots)$ from the assertion of S_1, where a_i and b_i are similar, for all i.

Coherence relations (2 of 2)

- Elaboration: Infer the same proposition P from the assertions of S_0 and S_1.
- Occasion: A change of state can be inferred from the assertion of S_0, whose final state can be inferred from S_1, or a change of state can be inferred from the assertion of S_1, whose initial state can be inferred from S_0.

Coherence resolution

- Determine the relationships between sentences or discourse segments
- Discover inferences that should be made
- Useful for IR, text summarization, pronoun resolution

Inference

- Sound inference, e.g., modus ponens (deduction):

$$\frac{\begin{array}{c}\alpha \Rightarrow \beta \\ \alpha\end{array}}{\beta}$$

- Unsound inference, e.g., abduction:

$$\frac{\begin{array}{c}\alpha \Rightarrow \beta \\ \beta\end{array}}{\alpha}$$

- Associate ‘unsound’ conclusions with some kind of weight or cost, and make the DEFEASIBLE.

Inference-Based Coherence Resolution

- Establish axioms
  - Pertaining to coherence relations
  - Encoding world knowledge
- Represent discourse segments in the same formalism as the axioms
- Establish coherence by creating a chain of reasoning linking the sentence interpretations that is rooted the assertion of a coherence relation
- In the process, posit unprovable assumptions
- $\rightarrow$ inference

Coherence relation axioms

[1] $\forall e_i, e_j$ Explanation(e_i, e_j) $\Rightarrow$ CoherenceRel(e_i, e_j)

[2] $\forall e_i, e_j$ Result(e_i, e_j) $\Rightarrow$ CoherenceRel(e_i, e_j)

$\ldots$

[3] $\forall e_i, e_j$ cause(e_j,e_i) $\Rightarrow$ Explanation(e_i, e_j)

[4] $\forall e_i, e_j$ cause(e_i,e_j) $\Rightarrow$ Result(e_i, e_j)

$\ldots$

World knowledge axioms

[5] $\forall x, y, e_i$ 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)

[6] $\forall x, y, e_j, e_k$ 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)

[7] $\forall x, y, z, e_l, e_m$ 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)

[8] $\forall e_i, e_j, e_k$ cause$(e_i, e_j) \wedge$ cause$(e_j, e_k) \Rightarrow$
cause(e_i, e_k)

Translation of two statements

[9] $\exists e_1, ck$ hide$(e_1, j, b, ck) \wedge$ carkeys(ck, b)

[10] $\exists e_2$ drunk(e_2, h)

Reasoning from coherence to the statements

[A] Assume coherence, i.e., Coherence-Rel(e_1, e_2)

[B] Infer Explanation(e_1, e_2) [1],[A]

[C] Infer cause(e_2, e_1) [3],[B]

[D] Infer cause(e_2, e_3) $\wedge$ (e_3, e_1) [8],[C]

[E] Infer cause(e_2, e_4) $\wedge$ (e_4, e_3) [8],[D]

[F] Infer diswant(e_3, j, b) $\wedge$ have(e_5, b, ck) [9],[D],[7]

Reasoning from coherence to the statements

[G] Infer diswant(e_3, j, e_6) $\wedge$ drive(e_6, b) [9],[8],[F]

[H] Infer drunk(e_2, b) [5],[8],[G]

- But [H] equals [10], if $b = h$ (pronoun resolution)
- Chain included [1],[9], and [10], so the discourse was coherent.
- Along the way, we inferred things not explicitly stated in the dicourse: John did not want Bill to drive; This is why John hid Bill's keys.

Reasoning from coherence to the statements

- Serious search problem
- ... managing the size of the search space
- ... choosing the best possibility
- Hobbs et al (1993) deal with this by assigning assumption costs to each inference.

Discourse structure

- Always looking for coherence between adjacent pairs of sentences would give incorrect results.
- Instead, search for structure in discourse, and look for coherence between adjacent discourse segments.
- ‘Parsing’ discourse structure (a side effect of the above) is useful for summarization, IR, etc. Possibly also for pronoun resolution.

Overview

- Leftovers: Centering Theory
- Leftout: FOPL
- Text coherence
- Coherence resolution
- Inference
- Inference-based coherence resolution
- Discourse structure