Named Entity Resolution

Rodrigo Agerri

Outline

Outline

Semantics/Pragmatics

What we usually refer to informally as meaning is the combination of the semantics and the pragmatics. We will treat the semantics as the part of the meaning that is relatively fixed and minimal. Pragmatics is the component of meaning that is context-sensitive and depends on both the application and the social structure within which is applied.

Approaches

- Rule-based: Logic-based tradition.
- Corpus-based: Probabilistic.

Outline

PTQ

- Montague (1974) changed the view of logic-based analisis to natural language: it is possible to build a formal system/methodology to explain/understand natural language.
- Proper Treatment of Quantification in English:
  - Compositionality: Compositional building the meaning of sentences that constitute a discourse and its combination (Frege 1896).
  - On Denoting and Descriptions by Russell (1905).
  - Algorithmic Process to build semantic representations/structures starting from syntax.

Frege

- **Sense and Reference**: “The current king of France is bald”. (Reference: false???)
- **Compositionality Principle**: the referent of a sentence depends on the references of its parts. (Marx brothers joke).
- **Sustituibility Principle**: if the reference of a sentence is its truth-value, a given truth-value should not change when replaced by a linguistic expression of the same reference, but different sense:
  - “Aristotle wrote the Nicomachean Ethics”
  - “Plato’s disciple wrote the Nicomachean Ethics”.
  - Exception when propositional attitudes or indirect style appear: “George VI wanted to know if Scott was the author of Waverley”.

Russell: On Denoting (1905)

- Proper names with sense but without reference (Frege): “The current king of France is bald”.
- Proper names are hidden Definite Descriptions: expressions within the scope of a quantifier on which the existence of the object over which we quantify is required.
- “The current king of France is not bald”. (which is which?)

Formal Representation of Natural Language

- To make explicit meaning in natural language it is required to represent it in an non-ambiguous construct:
  - First-order logic (compositional)
  - Constants, functions, quantification, inference, etc.

PTQ problems

If Sam owns a donkey he beats it.

- Target: $\forall x(donkey(x) \wedge owns(Sam, x) \rightarrow beats(Sam, x))$
- **Quantification type**: Descriptions create quantified expressions.

  $\exists x(donkey(x) \wedge owns(Sam, x)) \rightarrow beats(Sam, x)$
- **Quantification scope**:

  $\exists x(donkey(x) \wedge owns(Sam, x) \rightarrow beats(Sam, x))$

PTQ and Anaphora

It is not the case that every delegate failed to arrived. She signed in.

- Pronominal anaphora is not accounted for by logical conjunction.
- PTQ (and by extension) formal logic is not an adequate construct to represent semantics (or pragmatics) in natural language.
- The process is not mechanic.

Outline

Discourse Representation Theory

- Generative Syntax: Generalized Phrased Structure Grammar.
- Algorithm to derive semantic structures from syntax:
  - Anaphora.
  - Scope and Type of quantification.
  - Proper names.
  - The first algorithm for Named Entity Resolution.

Dynamic Semantics

- Anaphora and quantification
- NL discourse interpreted in context of a representation R
- Interpretation updates R to generate R'
- All noun phrases are discourse referents
- Quantification is controlled by context (when the referent is introduced)
- Processing of discourse is incremental new sentences update or add structure to DRS
- Inference is supported by a set of rules (broadly equivalent to FOPC but with a model-theoretic semantics rather than truth conditional interpretation)

Anaphora (pronominal)

Sam owns a downkey. He beats it.

x y u v
sam(x)
donkey(y)
owns(x,y)
u=y
v=x
beats(u,v)

Quantification

If a man eats a big kahuna burger, he enjoys it.

x y		*v w*
man(x) big_kahuna_burger(y) eat(x,y)	⇒	enjoy(v,w) v=x w=y