Semantic Theory
Summer 2006

M. Pinkal / A. Koller

Overview

- What is semantics?
- Why is it not trivial?
- Key challenges in semantics
- Logic in semantics
- Semantics in Language Technology

- Organisational issues

What is semantics?

- Semantics is the theory of the meaning of words, sentences, and texts.
- It is typically situated between:
  - syntax: grammatical structure of sentences (we assume that we already have syntactic analyses here)
  - pragmatics: use of utterances in extralinguistic contexts (we mostly ignore this here)

Why should we consider semantics?

- Consider the following argument:
  - John loves Mary.
  - Mary is a woman.
  - Therefore, John loves a woman.
- Is this true?
- How do you know it is true?

Why should we consider semantics?

- Consider the following argument:
  - Every student is intelligent.
  - John is not intelligent.
  - Therefore, John is not a student.
- Is this true?
- How do you know it is true?

Why should we consider semantics?

- Consider the following argument:
  - Every student is intelligent.
  - John is stupid.
  - Therefore, John is not a student.
- Is this true?
- How do you know it is true?

Why should we consider semantics?

- Consider the following argument:
  - John takes a train to Paris.
  - Therefore, he takes a train to France.
- Is this true?
- How do you know it is true?

About dolphins

About dolphins

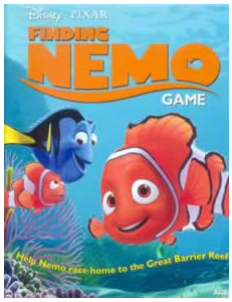
Dolphins are mammals

About dolphins

Dolphins are mammals, not fish.

About dolphins

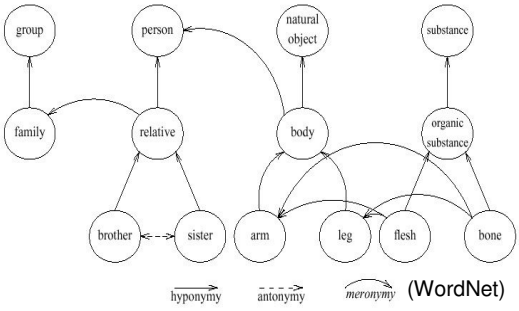
Dolphins are mammals, not fish.



About dolphins

Dolphins are mammals, not fish.

Figure 2. Network representation of three semantic relations among an illustrative variety of lexical concepts



(WordNet)

About dolphins

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time.

About dolphins

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time. At birth a bottlenose dolphin calf is about 90-130 cms long and will grow to approx. 4 metres, living up to 40 years.They are highly sociable animals, living in pods which are fairly fluid, with dolphins from other pods interacting with each other from time to time.

The two faces of semantics

- The truth of statements, and entailment between statements, can't be judged in terms of syntax alone.
- The mapping from syntax to semantics is not trivial.
- Semantics sits on the borderline between syntax and the world:
  - meaning is expressed through words that are combined grammatically
  - words and sentences are interpreted with respect to the real world

Ambiguity

- It is not trivial to determine what a sentence or text actually means.
- One challenge that comes up again and again is semantic ambiguity:
  - pronoun reference
  - genericity
  - word meaning
  - etc etc etc

WordNet Senses

The **noun** "body" has 9 senses in WordNet.

1. **body**, organic structure, physical structure -- (the entire physical structure of an organism (especially an animal or human being); "he felt as if his whole body were on fire")
2. **body**, dead body -- (body of a dead animal or person; "they found the body in the lake")
3. **body** -- (a group of persons associated by some common tie or occupation and regarded as an entity; "the whole body filed out of the auditorium")
4. torso, trunk, **body** -- (the body excluding the head and neck and limbs; "they moved their arms and legs and bodies")
5. **body** -- (an individual 3-dimensional object that has mass and that is distinguishable from other objects; "heavenly body")
6. **body** -- (a collection of particulars considered as a system; "a body of law"; "a body of doctrine"; "a body of precedents")
7. **body** -- (the external structure of a vehicle; "the body of the car was badly rusted")
8. consistency, consistence, **body** -- (the property of holding together and retaining its shape; "when the dough has enough consistency it is ready to bake")

9. **body** -- (the central message of a communication; "the body of the message was short")

WordNet Senses

1. **body**, organic structure, physical structure -- (the entire physical structure of an organism (especially an animal or human being); "he felt as if his whole body were on fire")
2. **body**, dead body -- (body of a dead animal or person; "they found the body in the lake")
3. **body** -- (a group of persons associated by some common tie or occupation and regarded as an entity; "the whole body filed out of the auditorium")
4. torso, trunk, **body** -- (the body excluding the head and neck and limbs; "they moved their arms and legs and bodies")
5. **body** -- (an individual 3-dimensional object that has mass and that is distinguishable from other objects; "heavenly body")

Key challenges in semantics

- Representation: Come up with formal descriptions that represent meaning.
- Semantics construction: Find a way to compute such semantic representations from a syntactic analysis.
- Computation: Perform computational tasks with such representations efficiently, and connect them with world knowledge.

What should a semantic theory provide?

- A framework to specify word meaning
- The composition process leading from word meanings to sentence information
- The building of a semantic discourse representation from a sequence of sentences in a text (or piece of dialogue)
- Disambiguation/ resolution mechanisms selecting the intended information of an utterance from the large number of linguistically possible interpretations
- Inference mechanisms leading from the given utterance information to other relevant information

Structure of this course

- Sentence semantics
  - representation
  - semantics construction
  - computational aspects
- Discourse semantics
- Lexical semantics

Formal representations of meaning

- The meaning of words, sentences, or texts are hard to capture in full generality.
- Standard approximation: Represent the truth conditions of the sentence.
  - What does the world have to look like in order to make the sentence true?
- We will modify this approximation when we talk about discourse.

Logic as a framework for NL semantics

- Logic with its model-theoretic, truth-conditional interpretation is a (so far) indispensable guide to the analysis of natural language meaning.
  - Check whether truth conditions of a logical expression conform with the intuitive judgment about the truth of a sentence in a given situation
- Logic allows to model inference by deduction systems, which again are controlled through a formal entailment concept.
- Logic supports the composition process.

Dolphins in logic

Dolphins are mammals, not fish.
$\forall d$ (dolphin(d)$\rightarrow$mammal(d) $\wedge\neg$fish(d))

Dolphins live-in pods.
$\forall d$ (dolphin(d)$\rightarrow \exists x$ (pod(p) $\wedge$ live-in (d,p))

Dolphins give birth to one baby at a time.
$\forall d$ (dolphin(d)$\rightarrow$
$\forall x \forall y \forall t$ (give-birth-to (d,x,t) $\wedge$ give-birth-to (d,y,t) $\rightarrow$ x=y)

How much of meaning does logic capture?

- Compositional sentence semantics ☺ Type Theory
- Inference ☺ FOL etc.
- Discourse and dialogue semantics ☺ 😐 DRT
- Disambiguation/ resolution ☺☹
- Word meaning ☺☹

Logic as a framework for NL semantics

Less convincing aspects:

- Efficiency
- Cognitive adequacy
- Coverage
- Robustness

Semantics in language technology

- Efficiency and robustness are key desiderata in language technology.
- In particular, wide-coverage grammars that robustly compute meaningful semantic representations don't really exist.
- Most LT systems therefore don't use semantics as we will discuss it here. (Yet.)

"Shallow" semantic representations

- Represent meaning as a conjunction of atoms:
  - "J. P. Bolduc, vice chairman of W. R. Grace & Co. was elected a director."

    ```
    verb(e1, elect) & funct_of('Bolduc', x1) &
    subj(e1,unspecified) &
    description(e1,x1,director,de1) & ...
    ```
  - But: "Kennedy prevented a war"

    ```
    verb(e1, prev) & patient(e1, x1) & war(x1) ...
    ```

    ????
- Use simple world knowledge such as encoded in Wordnet.

The knowledge bottleneck

- The "other" face of semantics is connected to the real world, and we need world knowledge to deal with it.
- In principle, we can try to formalise our world knowledge as logical axioms.
- But there is a lot of world knowledge! Except for limited domains, extremely challenging to formalise.
- This is a major impediment to the practical use of LT systems based on "deep" semantics.
- New approaches try to learn knowledge from corpora.

The silver lining

- Nevertheless, learning about semantic theory is important!
  - It's the truth.
  - necessary for 100% accuracy
  - principled solutions
  - the way of the future

Organisational issues

- Books
- Website
- Exercises
- Exercise vs. lecture sessions
- Exam