Computational Linguistics: Semantics

Raffaella Bernardi

University of Trento

1. Recall: goals

Back to our Goals:

1. provide students with an overview of the field with focus on the syntax-semantics interface;
2. bring students to be aware of **several lexicalized formal grammars**, [**Done**]
3. bring students to be aware of **computational semantics** models and be able to combine some of them to capture the natural language syntax-semantics interface; [**next block of classes**]
4. evaluate several applications with a special focus to Interactive Question Answering and Language and Vision Models;
5. make students acquainted with writing scientific reports. (Reading, Summarize, Discussion, Proposals) [Started]

2. Recall: overall program

- 8 classes on Syntax (Sep-Oct): Formal Grammars of English, Syntactic Parsing, Statistical Parsing. [done]
- **11 classes on Semantics (Oct-Nov): Formal Semantics, Distributional Semantics Models, The Representation of Sentence Meaning**

3. Semanticists

> "It is the task of semanticists to describe the **meaning** of linguistic elements and to study the principles which allow (and exclude) the assignment of meaning to **combinations** of these elements. In addition, a complete and adequate semantic theory characterizes the systematic meaning relations between words and sentences of a language, and provides an account of the relations between linguistic expressions and the things that they can be used to talk about (i.e., the **external world**)". [de Swart 1998]

In short, Semantics is the study of **meaning** of words and their **combination** into sentences used to **comunicate** a message.

- What is meaning?
- What's the relation between meaning, mind, and the world?

https://plato.stanford.edu/entries/meaning/

3.1. The focus on Meaning

- In the '30, the behaviorism school dominated the linguistic scene (Bloomfield 1933, 1936): all behvior should be explained in terms of stimulus-responsense. Bloomfield rejected the study of meaning: it requires introspections, hence no scientifically regorous. **Meaning went away from the scene**
- **Meaning entered the scene marginally:**
  - **Chomsky** (1957, 1965) was interested in sentence structure. Hence, meaning is interesting if the structure is syntatically ambigous.
  - Interpretative Semantics (**Katz and Fodor** 1964): first we develop the syntactic structure and then turn these structures into semantic representations.
  - Generative Semantics (**Ross** 1967 and **Lakoff** 1971): interpretations were generated directly by the grammar as deep structures
  - **Lexical Semantics** frames: e.g., Fillmore 1968.

3.2. Semantics dominates the scene

1. **Formal Semantics**: Traces back to Montague. Very strong in the '70-'90, still very active, see e.g., SALT and Amesterdam Colloquium.
2. **Distributional Semantics**: very strong nowadays. Traces back to Harris 1954 and Firth 1957.

We will present and practice with both.

4. Formal semantics

The foundational work by Frege, Carnap, and Tarski had led to a rise in work on modal logic, tense logic, and the analysis of **philosophically interesting issues in natural language**. Philosophers like Kripke and Hintikka added model theory.

These developments went hand-in-hand with the **logical syntax** tradition (Peirce, Morris, Carnap), distinguishing syntax (well-formedness) from semantics (interpretation) and pragmatics (use).

Though the division was inspired by language, **few linguists attempted to apply the logician's tools in linguistics as such**.

This changed with **Montague**.

> "I reject the contention that an important theoretical difference exists between formal and natural languages." (Montague, 1974)(p.188)

A compositional approach, using a "rule-by-rule" translation (Bach) of a syntactic structure into a first-order, intensional logic. This differed substantially from transformational approaches (generative or interpretative semantics).

4.1. Frege: What's the meaning of linguistic signs?

Frege's question: What is identity? It's a relation between objects vs. between linguistic signs.

None of the two solutions can explain why the two identities below convey different information:

(i) "Mark Twain is Mark Twain" [same obj. same ling. sign]

(ii) "Mark Twain is Samuel Clemens". [same obj. diff. ling. sign]

Frege's answer: A linguistic sign consists of a:

- **reference**: the object that the expression refers to
- **sense**: mode of presentation of the referent.

Linguistic expressions with the same reference can have different senses.
Formal semanticists focus on "reference" and are inspired by Logic.

4.2. Tarski: What does a given sentence mean?

The meaning of a sentence is its truth value.

"Snow is white" is true iff snow is white.

Rephrased in: "Which is the meaning representation of a given sentence to be evaluated as true or false?"

- **Meaning Representations**: Predicate-Argument Structures are a suitable meaning representation for natural language sentences. E.g. the meaning representation of "Lori knows Alex" is $\texttt{know}(\texttt{lori}, \texttt{ale})$

  whereas the meaning representation of "A student knows Alex" is $\exists x.\texttt{student}(x) \wedge \texttt{knows}(x, \texttt{ale})$.

- **Interpretation**: a sentence is taken to be a proposition and its meaning is the truth value of its meaning representations. E.g.

  $[\![\exists x.\texttt{student}(x) \wedge \texttt{walk}(x)]\!] = 1$ iff standard FOL definitions are satisfied.

4.3. Quantifiers

FOL quantifiers Frege introduced the FOL symbols: $\exists$ and $\forall$ to represent the meaning of quantifiers ("some" and "all") precisely and to avoid ambiguities.

Natural Language Syntax-Semantics The grammatical structure:

> "A natural number is bigger than all the other natural numbers."

can be represented as:

1. $\forall x \exists y Bigger(y, x)$ true
2. $\exists y \forall x Bigger(y, x)$ false

Hence, there can be a mismatch between syntactic and semantics representations

4.4. Montague: Syntax-Semantics

Stokhof (2006) summarizes Montague's theory by highlighting two characteristics:

- Semantics is model-theoretic.
- Compositionality: Semantics is syntax-driven, syntax is semantically motivated.