Computational Linguistics: Formal Semantics

Raffaella Bernardi

University of Trento

Contents

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 on the one hand of **several lexicalized formal grammars**, [**Done**]
3. on the other hand 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 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 Dependency Parsing. [done]
- **14 classes on Semantics (Oct-Nov): Formal Semantics, Distributional Semantics Models, The Representation of Sentence Meaning**
- 2 classes on Multimodal Models (end of Nov): Language and Vision

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 systemantic 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. Meaning went away from the scene

1. In US, **Breal** (1899) study on lexicon and its evolution (diachronic), mostly fieldwork (hence, focus on phonology, morphology).

2. In EU, **de Saussure** (1916): focus on synchronic study of language (vs. diachronic). Plus: “sign” as a combination of a significant (form) and an a signifie’ (meaning), whose relation is arbitrary. Still interest mostly on the lexicon.

3. In the ’30, the behaviorism school dominated the linguistic scene (Bloomfield 1933, 1936): a psychology theory that rejects the study of the mind, all behvior should be explained in terms of stimulus-responsense. **Bloomfield** rejected the study of meaning: it requires interospections, hence no scientifically regorous.

3.2. Meaning entered the scene marginally

1. **Chomsky** (1957, 1965) was interested in sentence structure. Hence, meaning is interesting if the structure is syntatically ambigous.

2. Interpretative Semantics (**Katz and Fodor** 1964): first we develop syntactic structure and then turn these structures into semantic representations. (Syntax is autonomous from semantics!)

3. Generative Semantics (**Ross** 1967 and **Lakoff** 1971): interpretations were generated directly by the grammar as deep structures, and were subsequently transformed into recognizable sentences by transformations.

4. **Lexical Semantics** frames: e.g., Fillmore 1968.

The study of meaning has been for long marginalized in linguistics.

See CS's course for work on Lexical Semantics.

3.3. Semantics dominates the scene

1. Formal Semantics: very strong in the '70-'90. Still very active (see SALT and Amsterdam 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 $\texttt{konw}(\texttt{lori}, \texttt{ale})$

  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.

Today we look at the first issue. Next time to the second.

5. Semantics is model-theoretic

The focus is on meaning as "extension":

> "The extension of an expression is the set of things it extends to, or applies to" (Wikipedia)

Ingredients:

- A model of the world
- the model consists of sets
- words in a language refer or denote parts of the model
- a proposition is true iff it corresponds to state of affairs in the model.

5.1. Propositional Logic (PL)

A **model** consists of two pieces of information:

- which collection of atomic propositions we are talking about (**domain**, D),
- and for each formula which is the appropriate **semantic value**, this is done by means of a function called **interpretation function** ($\mathcal{I}$).

Thus a model $\mathcal{M}$ is a pair: $(D, \mathcal{I})$.

Main interest: Entailment it is valid iff **For all the interpretations** for which the premise is true, the consequence is also true.

Propositional Logic (PL): represents **propositions**. Atomic ones, p, q, r and complex ones built with truth-functional connectives: $P \wedge Q, P \vee Q, P \rightarrow Q, \neg P$.

5.2. How far can we go with PL?

1. Casper is bigger than John
2. John is bigger than Peter
3. Therefore, Casper is bigger than Peter.

Questions:

How would you formalize this inference in PL?

What do you need to express that cannot be expressed in PL?

Answer:

You need to express: "relations" (is bigger than) and "entities" (Casper, John, Peter)

5.3. What else do we need?

1. Bigger(casper,john)
2. Bigger(john,peter)
3. Therefore, Bigger(casper,peter)

Question: Do you still miss something?

The knowledge that: for all x, for all y and for all z

IF Bigger(x,y) AND Bigger(y,z) THEN Bigger(x,z).

We miss the universal quantifier: $\forall$.

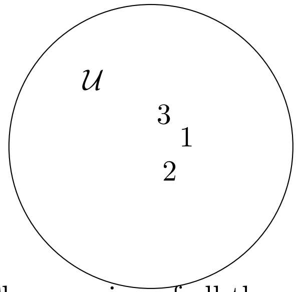
5.4. More expressive logic: First order Logic

- Just like in propositional logic, a (complex) FOL formula may be true (or false) with respect to a given interpretation.
- An interpretation specifies referents for
  constant symbols → **objects**
  predicate symbols → **relations**
- An atomic sentence $P(t_1, \ldots, t_n)$ is true in a given interpretation
  iff the *objects* referred to by $t_1, \ldots, t_n$
  are in the *relation* referred to by the predicate P.

5.5. Meaning as Reference: constants

Following Tarski, we build a Model by looking at a Domain (the set of entities) and at the **interpretation function** $\mathcal{I}$ which assigns an appropriate **denotation** in the model $\mathcal{M}$ to each individual and n-place predicate constant.

Individual constants If α is an individual constant, $\mathcal{I}$ maps α onto one of the entities of the universe of discourse $\mathcal{U}$ of the model $\mathcal{M} : \mathcal{I}(\alpha) \in \mathcal{U}$.

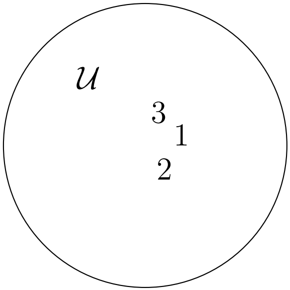


$$\mathcal{I}(a) = 1,\ \mathcal{I}(b) = 2,\ \mathcal{I}(c) = 3$$

The meaning of all the other words is based on the entities.

5.6. Meaning as Reference: properties

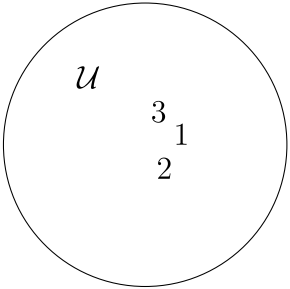
Set of entities the property of being "odd" denotes the **set of entities** that are "odd". Formally, for O (res. E) a one-place predicate, the interpretation function $\mathcal{I}$ maps O onto a subset of the universe of discourse $\mathcal{U} : \mathcal{I}(P) \subseteq \mathcal{U}$.



$\mathcal{I}(O) = \{1, 3\}$ $\mathcal{I}(E) = \{2\}$

5.7. Meaning as Reference: relation

Set of entities pairs The relation such as “bigger” denotes **sets of ordered pairs of entities**, namely all those pairs which stand in the “bigger” relation. Given the relation R, the interpretation function $\mathcal{I}$ maps R onto a set of ordered pairs of elements of $\mathcal{U}$: $\mathcal{I}(R) \subseteq \mathcal{U} \times \mathcal{U}$



$$\mathcal{I}(B) = \{(2,1),(3,2),(3,1)\}$$

5.8. Meaning as Reference: Linguistic example

Let $[\![\texttt{w}]\!]$ indicate the interpretation of **w**:

$$
\begin{array}{lcl}
[\![\texttt{sara}]\!] & = & \text{sara}; \ldots \\
[\![\texttt{walk}]\!] & = & \{\text{lori}\}; \\
[\![\texttt{know}]\!] & = & \{(\text{lori, alex}), (\text{alex,lori}), (\text{sara, lori}), \\
 & & (\text{lori, lori}), (\text{alex, alex}), (\text{sara, sara}), (\text{pim, pim})\}; \\
[\![\texttt{student}]\!] & = & \{\text{lori, alex, sara}\}; \\
[\![\texttt{professor}]\!] & = & \{\text{pim}\}; \\
[\![\texttt{tall}]\!] & = & \{\text{lori, pim}\}.
\end{array}
$$

which is nothing else to say that, for example, the relation **know** is the **set of pairs** (α, β) where α knows β; or that 'student' is the set of all those elements which are a student.

6. From sets to functions

A set and **its characteristic function** amount to the same thing:
if f_X is a function from Y to $\{F, T\}$, then $X = \{y \mid f_X(y) = T\}$.

In other words, the assertion '$y \in X$' and '$f_X(y) = T$' are equivalent.

$$[\![student]\!] = \{\text{lori, alex, sara}\}$$

`student` can be seen as a function from entities to truth values:

$$[\![student]\!] = \{x | \texttt{student}(x) = T\}$$

6.1. Types of denotations

- **Sentences** can be thought of as referring to their truth value - they denote in the the domain $D_t = \{1, 0\}$.
- **Entities**: can be represented as constants denoting in the domain D_e, e.g. $D_e = \{\texttt{john}, \texttt{vincent}, \texttt{mary}\}$
- **Functions**: The other natural language expressions can be seen as incomplete sentences and can be interpreted as **boolean functions** (i.e. functions yielding a truth value). They denote on functional domains $D_b^{D_a}$ and are represented by functional terms of type $(a \to b)$.

6.2. Exercises: Model, Types

Model

1. Harry is a wizard.
2. Hagrid scares Dudley.
3. All wizards are magical.
4. Uncle Vernon hates anyone who is magical.
5. Aunt Petunia hates anyone who is magical and anyone who scares Dudley.

Build a model for it by writing your interpretation for wizards, magical, scares, hates using the set theoretical interpretation.

Types of Denotation Translate the sets of the previous exercises into functions by assigning them the corresponding semantic type.

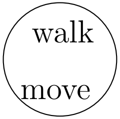
7. Back to Logic Entailment

$[\![\phi]\!] \leq_t [\![\psi]\!]$ iff
$[\![\phi]\!] = 0$ or $[\![\psi]\!] = 1$

$[\![X]\!] \leq_{(a \rightarrow b)} [\![Y]\!]$ iff $\forall \alpha \in D_a$
$[\![X(\alpha)]\!] \leq_b [\![Y(\alpha)]\!]$

7.1. Lexical entailment (partially ordered domains)

Given $D_e = \{\text{lori, alex, sara}\}$.



$$\{\text{lori}\} \subseteq \{\text{lori,alex}\}$$

$$[\![\text{walk}]\!] \leq_{(e \to t)} [\![\text{move}]\!] \text{ iff } \forall \alpha \in D_e,\ [\![\text{walk}]\!][\![\alpha]\!] \leq_t [\![\text{move}]\!][\![\alpha]\!]$$

$$\begin{array}{ll} 0 \leq 1 & \text{for } [\![\alpha]\!] = \text{alex} \\ 1 \leq 1 & \text{for } [\![\alpha]\!] = \text{lori} \\ 0 \leq 0 & \text{for } [\![\alpha]\!] = \text{sara} \end{array}$$

$$\{(\text{sara,lori})\} \subseteq \{(\text{sara,lori}),(\text{lori,alex})\}$$

$$[\![\text{tease}]\!] \leq_{(e \to (e \to t))} [\![\text{know}]\!]$$

Note, $(e \to (e \to t)) = (e \times e) \to t$

7.2. Phrase Entailment

$$
\begin{array}{ll}
[\![\text{tall student}]\!] \leq_{(e \rightarrow t)} [\![\text{student}]\!] & \text{iff} \quad \forall \alpha \in D_e \\
[\![\text{tall student}(\alpha)]\!] \leq_t [\![\text{student}(\alpha)]\!] & \text{iff} \\
[\![\text{tall student}]\!]([\![\alpha]\!]) \leq_t [\![\text{student}]\!]([\![\alpha]\!]) & \text{iff} \\
[\![\text{tall student}]\!]([\![\alpha]\!]) = 0 \text{ or } [\![\text{student}]\!]([\![\alpha]\!]) = 1. &
\end{array}
$$

7.3. Lesson

- (a) different entailment relations for different domains;
- (b) same entailment relation for words and phrases belonging to the same category (e.g. "dog $\leq_{(e \to t)}$ animal" and also "small dog $\leq_{(e \to t)}$ animal")

8. Formal Semantics: more advanced

- Intensional Semantics: http://web.mit.edu/fintel/fintel-heim-intensional.pdf
- Dynamic Semantics: https://plato.stanford.edu/entries/dynamic-semantics/
- Inquisitive Semantics https://projects.illc.uva.nl/inquisitivesemantics/

9. Formal Semantics: Summing up

Aim: Specify semantic representations for the **lexical items** based on reference and build the representation of sentence **compositionally**.

Solution We have seen that lexical meaning can be represented by sets or equivalently by functions.

Next time we are going to speak of compositionality.

Extra Reference

- Barbara Partee “Formal Semantics”. Chapter in the Handbook of Formal Semanitcs. Cambridge University (see dropbox.)