

Computational Linguistics: Formal Semantics

RAFFAELLA BERNARDI

UNIVERSITY OF TRENTO

Contents

1	Semantics	4
2	What's the meaning of linguistic signs?	5
	2.1 From words to sentences	6
	2.2 Answer: syntax-semantics	7
	2.3 Quantifiers	8
3	Philosophy of Language: Two lines of thoughts	9
4	Formal vs. Distributional Semantics	10
5	Computational Semantics: Main questions	11
	5.1 What does a given sentence mean?	12
	5.2 Meaning as Reference: constants	13
	5.3 Meaning as Reference: properties	14
	5.4 Meaning as Reference: relation	15
	5.5 Meaning as Reference: Linguistic example	16
	5.6 Characteristic function	17
6	Lambda Calculus	18
	6.1 Lambda Calculus: Function and lambda terms	19
	6.2 Lambda calculus: Typed lambda terms	20

6.3	Lambda Calculus: examples	21
7	Back to Meaning representation	22
7.1	Back to Meaning representation: Scope ambiguity	23
8	Formal Semantics: Summing up	24
8.1	Formal Semantics: Abstraction	25
8.2	Formal Semantics: Abstraction	26
8.3	Formal Semantics: Models	27
9	Logic Entailment	28
9.1	Lexical entailment (partially ordered domains)	29
9.2	Phrase Entailment	30
9.3	Lesson	31
10	Back to Frege's challenges	32
11	Problems with set-theoretic semantics	33
12	Exercises	34
13	Extra Reference	35

1. Semantics

Semantics: it's the study of the meaning of words and the reference of linguistic inputs. The former is studied in **Lexical Semantics** and the latter in **Computational Semantics**.

Lexical Semantics Words are seen as having a systematic structure that governs what they mean, how they **relate to actual entities** and how they can be used. Studies on this topic result into e.g. Dictionary or Ontologies like WordNET.

2. 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.

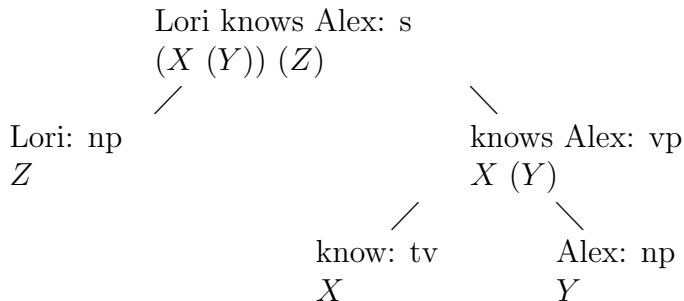
2.1. From words to sentences

Complete vs. Incomplete Expressions Frege made the following distinction:

- ▶ A sentence is a **complete** expression, it's reference is the truth value.
- ▶ A proper name stands for an object and is represented by a constant. It's a **complete** expression.
- ▶ A predicate is an **incomplete** expression, it needs an object to become complete. It is represented by a function. Eg. "left" needs to be completed by "Raj" to become the complete expression "Raj left".

Principle of Compositionality: The meaning of a sentence is given by the meaning of its parts and by the compositionality rules. This holds both at the reference and sense level.

2.2. Answer: syntax-semantics



- ▶ We want to replace the variables (X, Y, Z) with the meaning representation of the words they stand for
- ▶ Compute the meaning representation of the phrases they build: $X (Y)$ and $((X (Y)) (Z))$, viz. we want to know the operation that assembles the words, and phrases systematically.

2.3. 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. $\forall x \exists y \text{Bigger}(y, x)$ true
2. $\exists y \forall x \text{Bigger}(y, x)$ false

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

3. Philosophy of Language: Two lines of thoughts

Formal Semantics Building on Wittgenstein (Truth Tables), Tarski (model, domain, interpretation function and assignments), Montague aimed to define a model-theoretic semantics for natural language. He treats natural language as a formal language:

- ▶ Syntax-Semantics go in parallel.
- ▶ It's possible to define an algorithm to compose the meaning representation of the sentence out of the meaning representation of its single words.

Language as use (the second) Wittgenstein claims that the meaning of linguistic signs is its **use** within a context and cannot be given by a fixed set of properties since it is **vague**, but it's possible to identify the “family of expressions” to which a word/expression is similar to.

4. Formal vs. Distributional Semantics

We will show how nowadays the two trends are converging.

1. Formal Semantics Models [Today]
 - ▶ Brief intro: Reference, Model, Domain, Interpretation func.
 - ▶ Syntax-semantics and meaning of phrases/sentences
 - ▶ Lexical, phrasal and sentential entailment
2. Distributional Semantics Models [Next time: Monday]
 - ▶ Brief intro: DS assumptions, vectors, space and matrices.
 - ▶ From content words to grammatical words.
 - ▶ Lexical and phrasal entailment: results
3. DS and Compositionality [Next next time: Wednesday]
 - ▶ Composing DS representation: state-of-art methods

5. Computational Semantics: Main questions

The main questions are:

1. What does a given sentence mean?
2. How is its meaning built?
3. How do we infer some piece of information out of another?

The first and last question are closely connected.

We will consider two approaches:

1. Formal Semantics Models (Logical approach).
2. Distributional Semantics Models

5.1. What does a given sentence mean?

The meaning of a sentence is its truth value. Rephrased in: “Which is the meaning representation of a given sentence to be evaluated as true or false?”

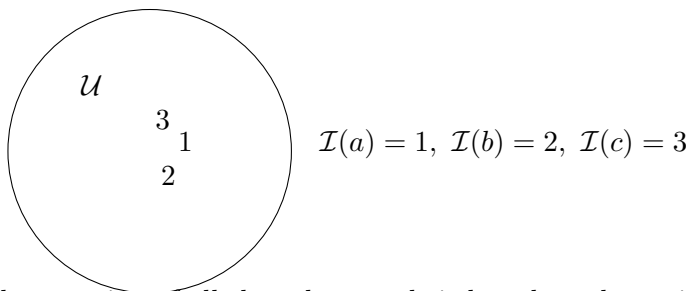
- ▶ **Meaning Representations:** Predicate-Argument Structures are a suitable meaning representation for natural language sentences. E.g. the meaning representation of “Lori knows Alex” is `konw(lori, ale)`
whereas the meaning representation of “A student knows Alex” is $\exists x.\text{student}(x) \wedge \text{knows}(x, \text{ale})$.
- ▶ **Interpretation:** a sentence is taken to be a proposition and its meaning is the truth value of its meaning representations. E.g.
 $\llbracket \exists x.\text{student}(x) \wedge \text{walk}(x) \rrbracket = 1$ iff standard FOL definitions are satisfied.

How is the MR built?

5.2. 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.

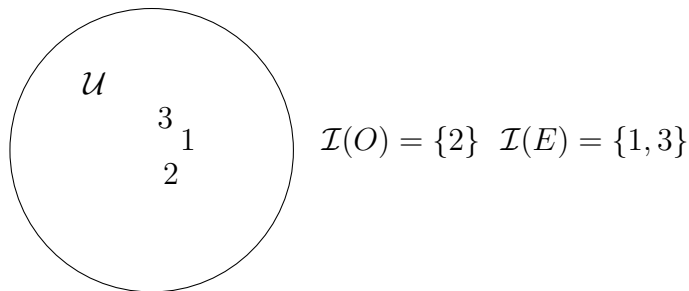
Individual constants If α is an individual constant, \mathcal{I} maps α onto one of the entities of the universe of discourse \mathcal{U} of the model $\mathcal{M} : \mathcal{I}(\alpha) \in \mathcal{U}$.



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

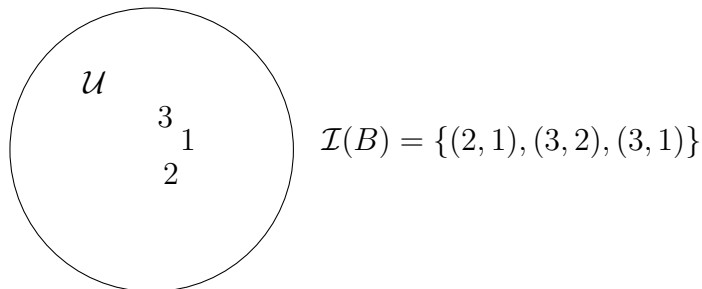
5.3. Meaning as Reference: properties

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 $\mathcal{U} : \mathcal{I}(P) \subseteq \mathcal{U}$.



5.4. 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} : $\mathcal{I}(R) \subseteq \mathcal{U} \times \mathcal{U}$



5.5. Meaning as Reference: Linguistic example

Let $\llbracket \mathbf{w} \rrbracket$ indicate the interpretation of \mathbf{w} :

$\llbracket \mathbf{sara} \rrbracket$	=	sara; ...
$\llbracket \mathbf{walk} \rrbracket$	=	{lori};
$\llbracket \mathbf{know} \rrbracket$	=	{(lori, alex), (alex,lori), (sara, lori), (lori, lori), (alex, alex), (sara, sara), (pim, pim)};
$\llbracket \mathbf{student} \rrbracket$	=	{lori, alex, sara};
$\llbracket \mathbf{professor} \rrbracket$	=	{pim};
$\llbracket \mathbf{tall} \rrbracket$	=	{lori, pim}.

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.

5.6. Characteristic function

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.

$$\llbracket student \rrbracket = \{\text{lori, alex, sara}\}$$

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

$$\llbracket student \rrbracket = \{x \mid \text{student}(x) = T\}$$

Functions can be represented by lambda terms: $\lambda x.\text{student}(x)$

6. Lambda Calculus

Lambda calculus was introduced by Alonzo Church in the 1930s as part of an investigation into the foundations of mathematics.

- ▶ It has a **variable binding operator** λ . Occurrences of variables bound by λ should be thought of as place-holders for missing information: they explicitly mark where we should substitute the various bits and pieces obtained in the course of semantic construction.
- ▶ Function can be applied to argument (**Function application**)
- ▶ An operation called **β -conversion** performs the required substitutions.
- ▶ Variables can be abstracted from a term (**Abstraction**)

6.1. Lambda Calculus: Function and lambda terms

Function $f : X \rightarrow Y$. And $f(x) = y$ e.g. $SUM(x, 2)$ if $x = 5$, $SUM(5, 2) = 7$.

▶ $\lambda x.x$

▶ $\lambda x.(x + 2)$

[SUM(x,2)]

▶ $\underbrace{(\lambda x.(x + 2))}_{\text{function}} \underbrace{5}_{\text{argument}}$

▶ $\underbrace{(\lambda x.(x + 2))}_{\text{function}} \underbrace{5}_{\text{argument}} = 5 + 2$

▶ $((\lambda y.\lambda x.(x + y)) \underbrace{5}_{\text{argument2}}) \underbrace{2}_{\text{argument1}} = (\lambda x.(x + 5)) 2 = 2 + 5$

▶ $\lambda y.\lambda x.(x + y) = \lambda(x, y).(x + y)$

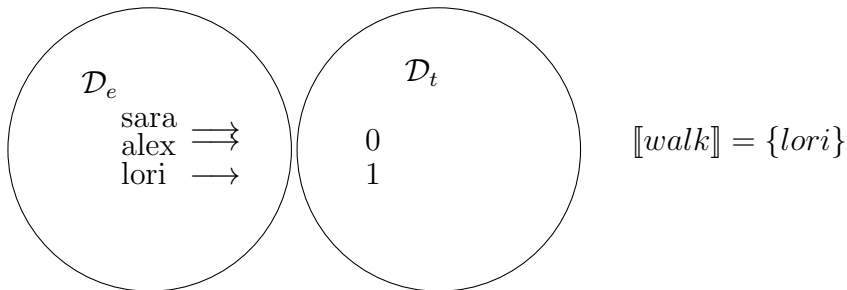
6.2. Lambda calculus: Typed lambda terms

- ▶ **Sentences** can be thought of as referring to their truth value - they denote in the the domain $D_t = \{1, 0\}$. M_t
- ▶ **Entities**: can be represented as constants denoting in the domain D_e , e.g. $D_e = \{\text{john, vincent, mary}\}$ j_e
- ▶ **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 $(a \rightarrow b)$. $\lambda x_a.P_b$

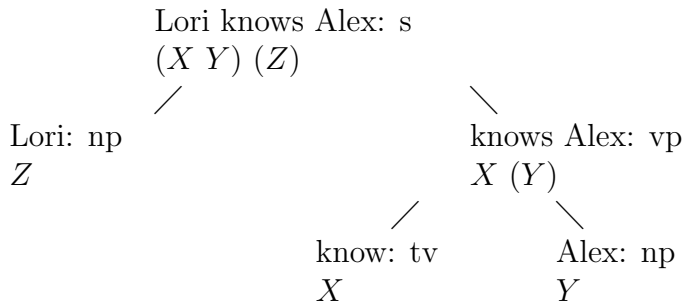
6.3. Lambda Calculus: examples

For instance “walk” is a set of entities (those entities which walk), hence it’s a function:

- ▶ denotes in $D_t^{D_e}$
- ▶ is of type $(e \rightarrow t)$,
- ▶ is represented by the term $\lambda x_e(\text{walk}(x))_t$



7. Back to Meaning representation



We call $((X Y) (Z))$ the proof-term of the tree.

- ▶ We can replace the place holders (X, Y, Z) with the λ -terms representing the words they stand for
- ▶ Compute the meaning representation of the phrases they build, $X (Y)$ and $(X Y) (Z)$, by using the λ -calculus operations.

7.1. Back to Meaning representation: Scope ambiguity

“A natural number is bigger than all the other natural numbers”:

- ▶ One syntactic tree.
- ▶ Two “proof terms”
 - ▷ $Q_{obj} \lambda x. Q_{sub} \lambda y. TV(y, x)$
 - ▷ $Q_{sub} \lambda y. Q_{obj} \lambda x. TV(y, x)$

by replacing the place holders (Q_{obj} , Q_{sub} , TV) with their MR, one obtains the MR of the sentence.

8. Formal Semantics: Summing up

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

Solution Set-theoretical interpretation, represented by λ -terms, and exploited function application and λ -abstraction to assemble meaning representation of larger expressions compositionally.

word	type	term	meaning
“lori”	e	l	lori
“walks”	$(e \rightarrow t)$	$\lambda x_e.(\mathbf{walks}(x))_t$	{lori}
“teases”	$(e \rightarrow (e \rightarrow t))$	$\lambda y_e.(\lambda x_e.(\mathbf{teases}(x, y))_t)$	{(lori, alex)}

8.1. Formal Semantics: Abstraction

For instance relative clauses need abstraction.

“the book which John read [...]”:

We know how to represent the noun phrase “John” and the verb “read”, namely, as `john` and $\lambda x.\lambda y.\text{read}(y, x)$.

What is the role of “which” in e.g. “the book which John read is interesting”?

The term representing “which” has to express the fact that it is replacing the role of a noun phrase in subject (or object position) within a subordinate sentence while being the subject (object) of the main sentence:

$$\lambda X.\lambda Y.\lambda z.Y(z) \wedge X(z)$$

The double role of “which” is expressed by the double occurrence of z .

8.2. Formal Semantics: Abstraction

$$\lambda X.\lambda Y.\lambda z.Y(z) \wedge X(z)$$

1. read u: $\lambda y(\text{read}(y, u))$
2. John read u: $\text{read}(j, u)$
3. John read: $\lambda u.\text{read}(j, u)$
4. which John read: $\lambda Y.\lambda z.Y(z) \wedge \text{read}(j, z)$
5. book which John read $\lambda z.\text{Book}(z) \wedge \text{read}(j, z)$

$$\llbracket \text{Book} \rrbracket \cap \llbracket \text{John read} \rrbracket$$

8.3. Formal Semantics: Models

The main ingredients in the FS Models are:

- ▶ The domains of the atomic types
- ▶ The domains of functional types
- ▶ The interpretation function that assigns to a reference its meaning.
- ▶ The entailment relation holding between sentence meaning.

9. Logic Entailment

$$\begin{aligned} \llbracket \phi \rrbracket \leq_t \llbracket \psi \rrbracket & \quad \text{iff} \\ \llbracket \phi \rrbracket = 0 \text{ or } \llbracket \psi \rrbracket = 1 \end{aligned}$$

$$\begin{aligned} \llbracket X \rrbracket \leq_{(a \rightarrow b)} \llbracket Y \rrbracket & \quad \text{iff } \forall \alpha \in D_a \\ \llbracket X(\alpha) \rrbracket \leq_b \llbracket Y(\alpha) \rrbracket \end{aligned}$$

9.1. Lexical entailment (partially ordered domains)

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

walk
move

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

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

$$\begin{array}{ll} 0 \leq 1 & \text{for } \llbracket \alpha \rrbracket = \text{alex} \\ 1 \leq 1 & \text{for } \llbracket \alpha \rrbracket = \text{lori} \\ 0 \leq 0 & \text{for } \llbracket \alpha \rrbracket = \text{sara} \end{array}$$

know
tease

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

$$\llbracket \text{tease} \rrbracket \leq_{(e \rightarrow (e \rightarrow t))} \llbracket \text{know} \rrbracket$$

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

9.2. Phrase Entailment

$$\begin{aligned} \llbracket \text{tall student} \rrbracket &\leq_{(e \rightarrow t)} \llbracket \text{student} \rrbracket && \text{iff } \forall \alpha \in D_e \\ \llbracket \text{tall student}(\alpha) \rrbracket &\leq_t \llbracket \text{student}(\alpha) \rrbracket && \text{iff} \\ \llbracket \text{tall student} \rrbracket(\llbracket \alpha \rrbracket) &\leq_t \llbracket \text{student} \rrbracket(\llbracket \alpha \rrbracket) && \text{iff} \\ \llbracket \text{tall student} \rrbracket(\llbracket \alpha \rrbracket) &= 0 \text{ or } \llbracket \text{student} \rrbracket(\llbracket \alpha \rrbracket) = 1. \end{aligned}$$

9.3. Lesson

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

10. Back to Frege's challenges

1. What's the meaning of linguistic signs? entity or set of ...
2. From words to sentence syntax driven composition
3. Quantifiers scope ambiguity one tree more proof terms
4. MR composition by function application and abstraction.

11. Problems with set-theoretic semantics

- ▶ Poor representation of the semantic content of words. (There is lexical semantics, but it is a semantics for very specific relations: hyponymy, synonymy, antonymy, etc)
- ▶ Do humans have sets in their heads? (Cognitive plausibility.)
- ▶ Is there truth?
- ▶ Where do models come from?

Next time, we look at the other meaning of linguistic signs, the “sense”.

12. Exercises

1. Build a model for a situation of your choice. Specify the domains of interpretation, the set-theoretical representation of words, and their corresponding typed lambda terms.
2. If an intransitive verb (e.g. “walk”) is represented by a unary-function, a transitive verb (e.g. “know”) by a binary-function, what is the function representing a ditransitive verb (e.g. “gave”)?
3. What could be the meaning representation of an adjective (e.g. “red”)?

13. Extra Reference

- ▶ Portner, P. and Partee, B. (eds.) “Formal Semantics: The essential readings”. Blackwell, 2002.