Distributional Compositionality
Intro to Formal Semantics

Raffaella Bernardi

University of Trento

February 13, 2012

Challenges and Outline

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.

Challenges and Outline

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.

Challenges and Outline

Answer: syntax-semantics

Lori knows Alex: s
$(X\ (Y))\ (Z)$

Lori: np
Z

knows Alex: vp
$X\ (Y)$

know: tv
X

Alex: np
Y

- 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 systematicaly.

Challenges and Outline

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

Challenges and Outline

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.

Challenges and Outline

Course Outline

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 [Tomorrow am]
   - Brief intro: DS assumptions, vectors, space and matrices.
   - From content words to grammatical words.
   - Lexical and phrasal entailment: results
3. DS and Compositionality [Tomorrow pm]
   - Composing DS representation: state-of-art methods

Formal 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 questions are closely connected.
In fact, since we are ultimately interested in understanding, explaining and accounting for the entailment relation holding among sentences, following Frege we can think of *the meaning of a sentence as its truth value* and use logical entailment ($\phi \vdash \psi$ iff $\mathcal{I}(\phi) = 0$ or $\mathcal{I}(\psi) = 1$.)

Formal Semantics

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 $\texttt{konw}(\texttt{lori}, \texttt{ale})$
  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.

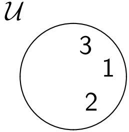
How is the MR built?

Set theoretical meaning

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 $\mathcal{M}$: $\mathcal{I}(\alpha) \in \mathcal{U}$.



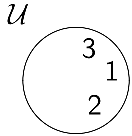
$\mathcal{I}(a) = 1,\ \mathcal{I}(b) = 2,\ \mathcal{I}(c) = 3$

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

Set theoretical meaning

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

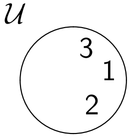


$$\mathcal{I}(O) = \{2\} \quad \mathcal{I}(E) = \{1, 3\}$$

Set theoretical meaning

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



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

Set theoretical meaning

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.

Meaning representation

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.

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

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

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

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

Lambda Calculus

Intro

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*)

Function $f : X \to Y$. 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)]
- $\underbrace{(\lambda x.(x + 2))}_{function} \underbrace{5}_{argument}$
- $\underbrace{(\lambda x.(x + 2))}_{function} \underbrace{5}_{argument} = 5 + 2$
- $((\lambda y.\lambda x.(x + y)) \underbrace{5}_{argument2}) \underbrace{2}_{argument1} = (\lambda x.(x + 2))\ 5 = 5 + 2$
- $\lambda y.\lambda x.(x + y) = \lambda(x, y).(x + y)$

Lambda calculus

Typed lambda terms

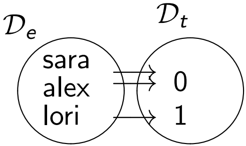
- *Sentences* can be thought of as referring to their truth value - they denote in the the domain $D_t = \{1, 0\}$. M_t
- *Entities*: can be represented as constants denoting in the domain D_e, e.g. $D_e = \{\mathtt{john}, \mathtt{vincent}, \mathtt{mary}\}$ j_e
- *Functions*: The other natural language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

Lambda Calculus

Typed lambda terms: example

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 $\lambda x_e(\texttt{walk}(x))_t$



$[\![walk]\!] = \{lori\}$

Back to Meaning representation

Building meaning representation

Lori knows Alex: s
$(X\ Y)\ (Z)$

Lori: np
Z

knows Alex: vp
$X\ (Y)$

know: tv
X

Alex: np
Y

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.

Back to Meaning representation

Scope ambiguity

"A natural number is bigger that 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.

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	$\texttt{l}$	lori
"walks"	$(e \rightarrow t)$	$\lambda x_e.(\texttt{walks}(x))_t$	{lori}
"teases"	$(e \rightarrow (e \rightarrow t))$	$\lambda y_e.(\lambda x_e.(\texttt{teases}(x, y))_t)$	{(lori, alex)}

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 $\lambda x.\lambda y.\mathtt{read}(y, x)$.

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.

Formal Semantics

Abstraction

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

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

$$[\![Book]\!] \cap [\![\text{John read}]\!]$$

Formal Semantics

Summing up: 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.

Entailment

Logic entailment

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

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

Entailment

Lexical entailment (partially ordered domains)

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

walk

move

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

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

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

know

tease

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

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

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

Entailment

Phrase Entailment

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

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")

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.

Tomorrow we look at the other meaning of linguistic signs, the "sense".

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")?

References

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