

Computational Semantics

Scott Farrar
CLMA, University of Washington
farrar@u.washington.edu

February 8, 2010

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Today's lecture

1 Background

2 Computational semantics

- Major subtasks
- Resources for computational semantics

Natural language semantics

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Natural language semantics

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

What can have meaning?

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Background

Computational semantics

Major subtasks
Resources for
computational
semantics

What can have meaning?

- a sentence/utterance of natural language

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Background

Computational semantics

Major subtasks
Resources for
computational
semantics

What can have meaning?

- a sentence/utterance of natural language
- a piece of code in a computer language

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Background

Computational semantics

Major subtasks
Resources for
computational
semantics

What can have meaning?

- a sentence/utterance of natural language
- a piece of code in a computer language
- a street sign

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Background

Computational semantics

Major subtasks
Resources for
computational
semantics

What can have meaning?

- a sentence/utterance of natural language
- a piece of code in a computer language
- a street sign
- a picture

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Background

Computational semantics

Major subtasks
Resources for
computational
semantics

What can have meaning?

- a sentence/utterance of natural language
- a piece of code in a computer language
- a street sign
- a picture
- a finger print (to a detective)

Definition

natural language semantics: the study of the meaning of natural language utterances, cf. syntax, the study of the *structure* of natural language utterances.

Background

Computational semantics

Major subtasks
Resources for
computational
semantics

What can have meaning?

- a sentence/utterance of natural language
- a piece of code in a computer language
- a street sign
- a picture
- a finger print (to a detective)
- a (para-linguistic) gesture

Meaning of a sentence can be unbounded

Example

The idea is to stop the EU from encroaching on the rights of member states other than in areas where the members have given them away.

Valid **entailments** from the example:

Meaning of a sentence can be unbounded

Example

The idea is to stop the EU from encroaching on the rights of member states other than in areas where the members have given them away.

Valid **entailments** from the example:

- There are EU member states.

Meaning of a sentence can be unbounded

Example

The idea is to stop the EU from encroaching on the rights of member states other than in areas where the members have given them away.

Valid **entailments** from the example:

- There are EU member states.
- Member states have rights.

Meaning of a sentence can be unbounded

Example

The idea is to stop the EU from encroaching on the rights of member states other than in areas where the members have given them away.

Valid **entailments** from the example:

- There are EU member states.
- Member states have rights.
- Member states can give their rights away.

Meaning of a sentence can be unbounded

Example

The idea is to stop the EU from encroaching on the rights of member states other than in areas where the members have given them away.

Valid **entailments** from the example:

- There are EU member states.
- Member states have rights.
- Member states can give their rights away.
- Member states haven't given all their rights away.

Meaning of a sentence can be unbounded

Example

The idea is to stop the EU from encroaching on the rights of member states other than in areas where the members have given them away.

Valid **entailments** from the example:

- There are EU member states.
- Member states have rights.
- Member states can give their rights away.
- Member states haven't given all their rights away.
- etc.

Methodological approaches to meaning

Computational
Semantics

Scott Farrar
CLMA, University
of Washington far-
rar@u.washington.edu

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Methodological approaches to meaning

- Meaning can only exist by virtue of a **cognitive agent** that has the ability of perception and cognitive process. Meanings are mental entities, elements of cognitive structure in the minds of the speakers. The mind, then, is **intermediate between the world and language**.

Methodological approaches to meaning

- Meaning can only exist by virtue of a **cognitive agent** that has the ability of perception and cognitive process. Meanings are mental entities, elements of cognitive structure in the minds of the speakers. The mind, then, is **intermediate between the world and language**.
- Meaning is **action**. Determining what bit of NL maps onto what action is called the **symbol grounding** problem. This is popular in robotics and other branches of AI.

Methodological approaches to meaning

- Meaning can only exist by virtue of a **cognitive agent** that has the ability of perception and cognitive process. Meanings are mental entities, elements of cognitive structure in the minds of the speakers. The mind, then, is **intermediate between the world and language**.
- Meaning is **action**. Determining what bit of NL maps onto what action is called the **symbol grounding** problem. This is popular in robotics and other branches of AI.
- Meanings are mapped onto “worlds”, or **model structures** M (after Tarski). This is known as **model theoretic** semantics. Montague, Lewis and many other **formal** semanticists use model theory.

Aspects of Meaning

There are at least three important aspects of the meaning of a NL utterance.

Aspects of Meaning

There are at least three important aspects of the meaning of a NL utterance.

- **Information structure:** what is important or emphasized by the speaker, e.g., old vs. new information.

There are at least three important aspects of the meaning of a NL utterance.

- **Information structure:** what is important or emphasized by the speaker, e.g., old vs. new information.
- **Speaker intention:** what does the speaker want to do by using language, e.g., whether the speaker believes the statement.

There are at least three important aspects of the meaning of a NL utterance.

- **Information structure:** what is important or emphasized by the speaker, e.g., old vs. new information.
- **Speaker intention:** what does the speaker want to do by using language, e.g., whether the speaker believes the statement.
- **Propositional content:** what ideas about the world are being communicated (e.g., objects, actions, spatial relations) and how those ideas are organized into discrete packets of information, i.e., **conceptualization**.

Aspects of meaning

Neutral

I want that report on my desk by Monday.

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Aspects of meaning

Neutral

I want that report on my desk by Monday.

Emphasize report

It's the report that I want on my desk by Monday.

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Aspects of meaning

Neutral

I want that report on my desk by Monday.

Emphasize report

It's the report that I want on my desk by Monday.

Empasize speaker desires, intentions, etc.

I think that I want that report on my desk by Monday.

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Key problem areas in NL semantics

Computational
Semantics

Scott Farrar
CLMA, University
of Washington far-
rar@u.washington.edu

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Key problem areas in NL semantics

- **semantic representation:** What language can we use to express propositions about semantic entities:

$$\forall x[\textit{sheep}(x) \rightarrow \textit{count}(\textit{Bill}, x)]$$

Key problem areas in NL semantics

- **semantic representation:** What language can we use to express propositions about semantic entities:

$$\forall x[\textit{sheep}(x) \rightarrow \textit{count}(\textit{Bill}, x)]$$

- **entailment:** What are the valid conclusions from a natural language utterance, e.g., an utterance *A* entails *B* when given that *A* is true, *B* can be concluded.

My oldest dog is 7. \models I have more than one dog., ...

Key problem areas in NL semantics

Computational
Semantics

Scott Farrar
CLMA, University
of Washington far-
rar@u.washington.edu

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Key problem areas in NL semantics

- **reference:** How do natural language expressions refer to real-world entities (or to entities in a model of the real world).

the devil, Satan, Lucifer

Key problem areas in NL semantics

- **reference**: How do natural language expressions refer to real-world entities (or to entities in a model of the real world).

the devil, Satan, Lucifer

- **compositionality**: The meaning of some utterance is composed of the meaning of its parts, aka *Frege's Principle*.

devil stick means what?

Key problem areas in NL semantics

- **reference**: How do natural language expressions refer to real-world entities (or to entities in a model of the real world).
the devil, Satan, Lucifer
- **compositionality**: The meaning of some utterance is composed of the meaning of its parts, aka *Frege's Principle*.
devil stick means what?
- **semantic analysis**: Deriving a semantic representation from an utterance.
from *eating a pie*, to:

Key problem areas in NL semantics

- **reference**: How do natural language expressions refer to real-world entities (or to entities in a model of the real world).
the devil, Satan, Lucifer
- **compositionality**: The meaning of some utterance is composed of the meaning of its parts, aka *Frege's Principle*.
devil stick means what?
- **semantic analysis**: Deriving a semantic representation from an utterance.
from *eating a pie*, to:

Key problem areas in NL semantics

- **reference**: How do natural language expressions refer to real-world entities (or to entities in a model of the real world).

the devil, Satan, Lucifer

- **compositionality**: The meaning of some utterance is composed of the meaning of its parts, aka *Frege's Principle*.

devil stick means what?

- **semantic analysis**: Deriving a semantic representation from an utterance.

from *eating a pie*, to:

$$\exists e \exists p [EatingEvent(e) \wedge Pie(p) \wedge patient(e, p)]$$

Today's lecture

1 Background

2 Computational semantics

- Major subtasks
- Resources for computational semantics

Definition

Computational semantics refers to the task whereby the meanings of natural language utterances are automatically computed and manipulated according to some logical system.

Definition

Computational semantics refers to the task whereby the meanings of natural language utterances are automatically computed and manipulated according to some logical system.

Key problem areas in computational semantics:

Definition

Computational semantics refers to the task whereby the meanings of natural language utterances are automatically computed and manipulated according to some logical system.

Key problem areas in computational semantics:

- defining a semantic representation (**formalism**)

Definition

Computational semantics refers to the task whereby the meanings of natural language utterances are automatically computed and manipulated according to some logical system.

Key problem areas in computational semantics:

- defining a semantic representation (**formalism**)
- defining algorithms for deriving semantics representations from NL input (**semantic analysis**)

Definition

Computational semantics refers to the task whereby the meanings of natural language utterances are automatically computed and manipulated according to some logical system.

Key problem areas in computational semantics:

- defining a semantic representation (**formalism**)
- defining algorithms for deriving semantics representations from NL input (**semantic analysis**)
- defining procedures for performing inferences using those representations (**automated inferencing, reasoning**)

Definition

Computational semantics refers to the task whereby the meanings of natural language utterances are automatically computed and manipulated according to some logical system.

Key problem areas in computational semantics:

- defining a semantic representation (**formalism**)
- defining algorithms for deriving semantics representations from NL input (**semantic analysis**)
- defining procedures for performing inferences using those representations (**automated inferencing, reasoning**)

Much overlap with artificial intelligence (AI) research (e.g., SHRDLU, KL-ONE, STRIPS). The task of robust semantic analysis is seen as an *AI-complete* problem.

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

- speech processing (recognition)

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

- speech processing (recognition)
- morphological processing (stemming)

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

- speech processing (recognition)
- morphological processing (stemming)
- syntactic processing (parsing, generation)

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

- speech processing (recognition)
- morphological processing (stemming)
- syntactic processing (parsing, generation)
- semantic processing (meaning derivation)

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

- speech processing (recognition)
- morphological processing (stemming)
- syntactic processing (parsing, generation)
- semantic processing (meaning derivation)

What about all those question-answering (Q/A) or information extraction (IE) systems?

Semantics processing

Semantics processing is perhaps the **least developed** of all of the major computational linguistics sub-fields, cf:

- speech processing (recognition)
- morphological processing (stemming)
- syntactic processing (parsing, generation)
- semantic processing (meaning derivation)

What about all those question-answering (Q/A) or information extraction (IE) systems?

An enterprising graduate student could make their name in CL if ...

Major subtasks

Major sub-tasks that have received the most research lately
(probably not AI-complete):

Major subtasks

Major sub-tasks that have received the most research lately (probably not AI-complete):

- computational **lexical semantics**: the processing of word meaning, word-word relations, and how words relate/influence syntax.

Major subtasks

Major sub-tasks that have received the most research lately (probably not AI-complete):

- computational **lexical semantics**: the processing of word meaning, word-word relations, and how words relate/influence syntax.
- word sense disambiguation (**WSD**): the task of selecting the correct **sense** given a particular word form in context

Major subtasks

Major sub-tasks that have received the most research lately (probably not AI-complete):

- computational **lexical semantics**: the processing of word meaning, word-word relations, and how words relate/influence syntax.
- word sense disambiguation (**WSD**): the task of selecting the correct **sense** given a particular word form in context
- semantic **role labeling**: the task of finding all semantic roles for each predicate in a sentence

Lexical semantics

Example

redolent, fragrant (synonyms)

Lexical semantics

Example

redolent, fragrant (synonyms)

Example

bitter, sour (antonyms)

Lexical semantics

Example

redolent, fragrant (synonyms)

Example

bitter, sour (antonyms)

Example

leg, table (meronyms)

Example

redolent, fragrant (synonyms)

Example

bitter, sour (antonyms)

Example

leg, table (meronyms)

Example

Spray the house with paint. Load the truck with bricks.
*Put the shelf with books.

Word Sense Disambiguation

Example

tank:

Word Sense Disambiguation

Example

tank:

(3) a. The tank has a 50mm gun.

Word Sense Disambiguation

Example

tank:

- (5) a. The tank has a 50mm gun.
b. Please fill my tank.

Word Sense Disambiguation

Example

tank:

- (7) a. The tank has a 50mm gun.
b. Please fill my tank.

Example

Word Sense Disambiguation

Example

tank:

- (9) a. The tank has a 50mm gun.
b. Please fill my tank.

Example

- (10) a. My caddy has new chrome wheels.

Word Sense Disambiguation

Example

tank:

- (11) a. The tank has a 50mm gun.
b. Please fill my tank.

Example

- (12) a. My caddy has new chrome wheels.
b. Don't you like my new wheels.

Semantic role labeling

Example

(13) The captain caught a fish.

Example

(14) The captain gave the fish to his mate .

Example

(15) The fish tasted good.

Semantic role labeling

Example

(16) The captain_{agent} caught a fish.

Example

(17) The captain gave the fish to his mate .

Example

(18) The fish tasted good.

Semantic role labeling

Example

(19) The captain_{agent} caught a fish.

Example

(20) The captain gave the fish to his mate_{benefactor}.

Example

(21) The fish tasted good.

Semantic role labeling

Example

(22) The captain_{agent} caught a fish.

Example

(23) The captain gave the fish to his mate_{benefactor}.

Example

(24) The fish_{theme} tasted good.

Computational semantics research

As with other subdomains of CL, computational semantics may be approached from a variety of perspectives:

Computational semantics research

As with other subdomains of CL, computational semantics may be approached from a variety of perspectives:

- **shallow** semantics processing
Given some text, find the named entities (company names, model numbers, etc.).

As with other subdomains of CL, computational semantics may be approached from a variety of perspectives:

- **shallow** semantics processing
Given some text, find the named entities (company names, model numbers, etc.).
- **deep** semantics processing
E.g., discovering how the meaning of some utterance relates to an entire discourse

Computational semantics research

As with other subdomains of CL, computational semantics may be approached from a variety of perspectives:

- **shallow** semantics processing
Given some text, find the named entities (company names, model numbers, etc.).
- **deep** semantics processing
E.g., discovering how the meaning of some utterance relates to an entire discourse
- **symbolic**, logic-based approaches (roots in AI)
new tools for manipulating meaning, shot in the arm from ontology research

Computational semantics research

As with other subdomains of CL, computational semantics may be approached from a variety of perspectives:

- **shallow** semantics processing
Given some text, find the named entities (company names, model numbers, etc.).
- **deep** semantics processing
E.g., discovering how the meaning of some utterance relates to an entire discourse
- **symbolic**, logic-based approaches (roots in AI)
new tools for manipulating meaning, shot in the arm from ontology research
- **stochastic** (and hybrid) approaches
recent, cutting edge focus, but not as developed as symbolic approaches

Machine-readable dictionaries/lexicons: resources which include lexical semantic information, sense information:

- LDOCE: Longman Dictionary of Contemporary English
- WordNet: rich in sense information and lexical relations
- word nets: many languages now have such resources

How many senses are there for the English word **pod**?

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

How many senses are there for the English word **pod**?

- 1 (n) {**pod**, cod, seedcase} (the vessel that contains the seeds of a plant (not the seeds themselves))

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

How many senses are there for the English word **pod**?

- 1 (n) {**pod**, cod, seedcase} (the vessel that contains the seeds of a plant (not the seeds themselves))
- 2 (n) {**pod**, seedpod} (a several-seeded dehiscent fruit as e.g. of a leguminous plant)

How many senses are there for the English word **pod**?

- 1 (n) {**pod**, cod, seedcase} (the vessel that contains the seeds of a plant (not the seeds themselves))
- 2 (n) {**pod**, seedpod} (a several-seeded dehiscent fruit as e.g. of a leguminous plant)
- 3 (n) {**pod**} (a group of aquatic mammals)

How many senses are there for the English word **pod**?

- 1 (n) {**pod**, cod, seedcase} (the vessel that contains the seeds of a plant (not the seeds themselves))
- 2 (n) {**pod**, seedpod} (a several-seeded dehiscent fruit as e.g. of a leguminous plant)
- 3 (n) {**pod**} (a group of aquatic mammals)
- 4 (n) {pod, fuel pod} (a detachable container of fuel on an airplane)

How many senses are there for the English word **pod**?

- 1 (n) {**pod**, cod, seedcase} (the vessel that contains the seeds of a plant (not the seeds themselves))
- 2 (n) {**pod**, seedpod} (a several-seeded dehiscent fruit as e.g. of a leguminous plant)
- 3 (n) {**pod**} (a group of aquatic mammals)
- 4 (n) {pod, fuel pod} (a detachable container of fuel on an airplane)
- 5 (v) {**pod**} (take something out of its shell or pod) *pod peas or beans*

How many senses are there for the English word **pod**?

- 1 (n) {**pod**, cod, seedcase} (the vessel that contains the seeds of a plant (not the seeds themselves))
- 2 (n) {**pod**, seedpod} (a several-seeded dehiscent fruit as e.g. of a leguminous plant)
- 3 (n) {**pod**} (a group of aquatic mammals)
- 4 (n) {pod, fuel pod} (a detachable container of fuel on an airplane)
- 5 (v) {**pod**} (take something out of its shell or pod) *pod peas or beans*
- 6 (v) {**pod**} (produce pods, of plants)

WordNet also contains a lexical hierarchy (ontology of sorts)

Sense 3

bass, basso --

(an adult male singer with the lowest voice)

=> singer, vocalist, vocalizer, vocaliser

=> musician, instrumentalist, player

=> performer, performing artist

=> entertainer

=> person, individual, someone...

=> organism, being

=> living thing, animate thing,

=> whole, unit

=> object, physical object

=> physical entity

=> entity

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Sense 7

bass --

(the member with the lowest range of a family of
musical instruments)

=> musical instrument, instrument

=> device

=> instrumentality, instrumentation

=> artifact, artefact

=> whole, unit

=> object, physical object

=> physical entity

=> entity

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

Available resources for computational semantics

Semantically tagged corpora: resources annotated for some component of semantic structure

Available resources for computational semantics

Semantically tagged corpora: resources annotated for some component of semantic structure

- **SemCor** corpus: a subset of the Brown Corpus consisting of over 234,000 words which were manually tagged with WordNet senses

Available resources for computational semantics

Semantically tagged corpora: resources annotated for some component of semantic structure

- **SemCor** corpus: a subset of the Brown Corpus consisting of over 234,000 words which were manually tagged with WordNet senses
- **SENSEVAL** corpora: group of small but rich sense-labeled resources, e.g., 2081 tagged content word tokens, from 5,000 total running words of English from the WSJ and Brown corpora

Available resources for computational semantics

Semantically tagged corpora: resources annotated for some component of semantic structure

- **SemCor** corpus: a subset of the Brown Corpus consisting of over 234,000 words which were manually tagged with WordNet senses
- **SENSEVAL** corpora: group of small but rich sense-labeled resources, e.g., 2081 tagged content word tokens, from 5,000 total running words of English from the WSJ and Brown corpora
- **FrameNet**: semantic-role tagged resource (multilingual versions)

Semantically tagged corpora: resources annotated for some component of semantic structure

Background

Computational
semantics

Major subtasks
Resources for
computational
semantics

- **SemCor** corpus: a subset of the Brown Corpus consisting of over 234,000 words which were manually tagged with WordNet senses
- **SENSEVAL** corpora: group of small but rich sense-labeled resources, e.g., 2081 tagged content word tokens, from 5,000 total running words of English from the WSJ and Brown corpora
- **FrameNet**: semantic-role tagged resource (multilingual versions)
- **Proposition Bank** ('Propbank'): semantic-role tagged resource based on PTB (English and Chinese versions)

FrameNet

Domain	Sample Frames	Sample Predicates
Transaction	Basic	buy, spend
Body	Action	flutter, wink
Cognition	Awareness	attention, obvious
	Judgment	blame, judge
	Invention	coin, contrive
Communication	Conversation	bicker, confer
	Manner	lisp, rant
Emotion	Directed	angry, please
	Experiencer-Obj	bewitch, rile
General	Imitation	bogus,forge
Motion	Arriving	enter, visit
	Filling	annoint, pack
Perception	Active	glance, savour
	Noise	snort, whine
Society	Leadership	emperor, sultan
Space	Adornment	cloak, line
Time	Duration	chronic, short
	Iteration	daily, sporadic

Propbank

Frameset agree.01

Arg0: Agreeer

Arg1: Proposition

Arg2: Other entity agreeing

Ex1: [Arg0 The group] agreed [Arg1 it wouldn't make an offer unless it had Georgia Gulfs consent].

Ex2: [ArgM-Tmp Usually] [Arg0 John] agrees [Arg2 with Mary] [Arg1 on everything].

WSJ text has been annotated with such role information.