

# Computational Semantics

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

February 12, 2010

# Today's lecture

- 1 Computational semantics
- 2 Logic
- 3 First-Order Predicate logic

## 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.

# Semantic representation language

## Definition

A **semantic representation** is an unambiguous expression into which natural language strings are transformed for the purposes of analyzing the meaning of natural language. A good representation language for NLP should be:

# Semantic representation language

## Definition

A **semantic representation** is an unambiguous expression into which natural language strings are transformed for the purposes of analyzing the meaning of natural language. A good representation language for NLP should be:

- **unambiguous** (many to one mapping)  
Even controlled languages are usually ambiguous: *CO attaches to hemoglobin in mammals.*

# Semantic representation language

## Definition

A **semantic representation** is an unambiguous expression into which natural language strings are transformed for the purposes of analyzing the meaning of natural language. A good representation language for NLP should be:

- **unambiguous** (many to one mapping)  
Even controlled languages are usually ambiguous: *CO attaches to hemoglobin in mammals.*
- **representationally adequate**, or expressive enough to handle all natural language phenomena:  
speaker intention, evidentiality, tense, aspect, coordination, etc.

# Semantic representation language

## Definition

A **semantic representation** is an unambiguous expression into which natural language strings are transformed for the purposes of analyzing the meaning of natural language. A good representation language for NLP should be:

- **unambiguous** (many to one mapping)  
Even controlled languages are usually ambiguous: *CO attaches to hemoglobin in mammals.*
- **representationally adequate**, or expressive enough to handle all natural language phenomena:  
speaker intention, evidentiality, tense, aspect, coordination, etc.
- compatible with natural language (**naturalness**), but still allow efficient semantic analysis  
Units of syntax should map onto units of semantic rep. in a straightforward manner.

# Semantic representation language

A good **reasoning procedure** for some representation language for NLP should be:

# Semantic representation language

A good **reasoning procedure** for some representation language for NLP should be:

- **inferentially adequate**, or able to facilitate the kind of reasoning necessary for comp. linguistics  
*The bus almost collided with the small car.*  
*The driver looked drunk.* (bus driver, car driver?)

# Semantic representation language

A good **reasoning procedure** for some representation language for NLP should be:

- **inferentially adequate**, or able to facilitate the kind of reasoning necessary for comp. linguistics  
*The bus almost collided with the small car.*  
*The driver looked drunk.* (bus driver, car driver?)
- **sound**: new knowledge actually follows from old knowledge, and is not simply created on the fly (truth preserving)

# Semantic representation language

A good **reasoning procedure** for some representation language for NLP should be:

- **inferentially adequate**, or able to facilitate the kind of reasoning necessary for comp. linguistics  
*The bus almost collided with the small car.*  
*The driver looked drunk.* (bus driver, car driver?)
- **sound**: new knowledge actually follows from old knowledge, and is not simply created on the fly (truth preserving)
- **complete**: each and every correct inference can be made (more difficult to achieve than soundness)  
incomplete reasoning is still useful

# Semantic representation language

A good **reasoning procedure** for some representation language for NLP should be:

- **inferentially adequate**, or able to facilitate the kind of reasoning necessary for comp. linguistics  
*The bus almost collided with the small car.*  
*The driver looked drunk.* (bus driver, car driver?)
- **sound**: new knowledge actually follows from old knowledge, and is not simply created on the fly (truth preserving)
- **complete**: each and every correct inference can be made (more difficult to achieve than soundness)  
incomplete reasoning is still useful
- **tractable**, or be able to make inferences in a reasonable amount of time.

# Semantic representation language

A good **reasoning procedure** for some representation language for NLP should be:

- **inferentially adequate**, or able to facilitate the kind of reasoning necessary for comp. linguistics  
*The bus almost collided with the small car.*  
*The driver looked drunk.* (bus driver, car driver?)
- **sound**: new knowledge actually follows from old knowledge, and is not simply created on the fly (truth preserving)
- **complete**: each and every correct inference can be made (more difficult to achieve than soundness)  
incomplete reasoning is still useful
- **tractable**, or be able to make inferences in a reasonable amount of time.

There is always a trade-off between inferential adequacy and tractability (FOL is intractable).

# Today's lecture

- 1 Computational semantics
- 2 Logic
- 3 First-Order Predicate logic

## Definition

A **logic** is an unambiguous formal language with:

## Definition

A **logic** is an unambiguous formal language with:

- a well defined syntax,

## Definition

A **logic** is an unambiguous formal language with:

- a well defined syntax,
- a precisely defined semantics that defines what the formulas mean,

## Definition

A **logic** is an unambiguous formal language with:

- a well defined syntax,
- a precisely defined semantics that defines what the formulas mean,
- and an accompanying proof theory whereby formulas are manipulated according to certain rules.

## Definition

A **logic** is an unambiguous formal language with:

- a well defined syntax,
- a precisely defined semantics that defines what the formulas mean,
- and an accompanying proof theory whereby formulas are manipulated according to certain rules.

Logics are often used to study valid reasoning and the properties of rational thought.

# Review for this lecture

## Chap 1 of ForAllX

You should have a grasp of this material:

- logical sentence
- validity of logical arguments
- truth values
- logical truth

# Review for this lecture

## Chap 2,3 of ForAllX

You should have a grasp of this material:

- formal language
- Sentential Logical
- logical connectives of sentential logical
- how to write sentential logical forms for simple English sentences
- wff
- truth tables for logical connectives
- logical equivalence, tautology, contradiction, consistency, validity,

# Sentential logical

Sentential logic is a logic in which the smallest (atomic) units are sentences themselves:

- John knows Jill.
- Jill knows Sue.
- John knows Sue.

- $A \rightarrow B$
- $A \wedge B$
- $A \vee C$

# Today's lecture

- 1 Computational semantics
- 2 Logic
- 3 First-Order Predicate logic

# Quantified logic

## FOPL

First-Order Predicate Logic (FOPL) refers to a family of logics.

## QL

Let's start with a simple version of one such logic, called **QL** for quantified logic (see *ForAllX* chap 4).

# Universe of Discourse

## Definition

The **universe of discourse** is the world that a formal logic refers to: individuals, properties, relations, etc.

# Universe of Discourse

## Definition

The **universe of discourse** is the world that a formal logic refers to: individuals, properties, relations, etc.

## *UD*

Sam, Houston, Sue, Moe, 'being tired', 'jumping', 'breaking', 'knowing the answer', etc.

# Universe of Discourse

## Definition

The **universe of discourse** is the world that a formal logic refers to: individuals, properties, relations, etc.

## *UD*

Sam, Houston, Sue, Moe, 'being tired', 'jumping', 'breaking', 'knowing the answer', etc.

## *UD* for NL

In NL the universe of discourse is essentially everything that we can talk about (the world and our experiences). In formal logic, we usually restrict *UD* to be some well-defined subset of the "everything we can talk about": the world of physical objects and spatial relations, people and kinship relations, etc..

# Components of QL

## QL

- constants
- variables
- predicates
- quantifiers
- wff's

# Constant terms

## Definition

A **constant** symbolizes an individual entity within the *UD*. In fact, it must pick out one and only one member of the *UD*.

In NL, a *singular term* refers to specific element of our *UD*, which can be of the form proper name or definite description.

## proper names

- *Dr. Phil*
- *Diane Sawyer*
- *Hamburg*

# Constant terms

## Definition

A **constant** symbolizes an individual entity within the *UD*. In fact, it must pick out one and only one member of the *UD*.

In NL, a *singular term* refers to specific element of our *UD*, which can be of the form proper name or definite description.

## proper names

- *Dr. Phil* DOCPHIL
- *Diane Sawyer* DIANE
- *Hamburg* HAMBURG

# Constant terms

## definite description

- *the depot*
- *that cat*
- *the largest continent*

## definite description

- *the depot LOCX1*
- *that cat BILL*
- *the largest continent ASIA*

# Constant terms

Of course not all phrases with *definite determiners* denote a single entity in the *UD*:

# Constant terms

Of course not all phrases with *definite determiners* denote a single entity in the *UD*:

- the elephants (plural)

# Constant terms

Of course not all phrases with *definite determiners* denote a single entity in the *UD*:

- the elephants (plural)
- the wolf is a stealthy animal (generic)

# Constant terms

Of course not all phrases with *definite determiners* denote a single entity in the *UD*:

- the elephants (plural)
- the wolf is a stealthy animal (generic)
- My roommate just bought a laborador.  
Those laboradors make great pets. (generic)

## Definition

We can use **variables** to stand in for constant symbols:

- $x, y, z$

Taken together, variables and constants (and functions) compose the set of **terms** in the logic.

## Definition

The simplest kind of **predicate** is a property of an individual:

- *tired(DIANE)*
- *mean(BILL)*
- *tall(EVEREST)*

## Definition

The simplest kind of **predicate** is a property of an individual:

- *tired(DIANE)*
- *mean(BILL)*
- *tall(EVEREST)*

Predicates can also represent *relations*:

- *know(SAM, DIANE)*
- *bit(BILL, SAM)*
- *climbed(BILL, EVEREST)*

A predicate may apply to some, all, or no members of the  $UD$ .

$jump(x)$  may not apply to any individual in  $UD$ . If that is the case, then given  $UD = \{DIANE, BILL, EVEREST\}$

- $jump(DIANE) = False$
- $jump(BILL) = False$
- $jump(EVEREST) = False$

## Definition

The logical vocabulary (operators or connectives) are the “closed class” elements of FOL—the function words.

## Definition

The logical vocabulary (operators or connectives) are the “closed class” elements of FOL—the function words.

Jill runs and Jack falls.

$run(JILL) \wedge fall(JACK)$

Jill is a woman or Jill is not a woman.

$woman(JILL) \vee \neg woman(JILL)$

## Definition

The logical vocabulary (operators or connectives) are the “closed class” elements of FOL—the function words.

Jill runs and Jack falls.

$run(JILL) \wedge fall(JACK)$

Jill is a woman or Jill is not a woman.

$woman(JILL) \vee \neg woman(JILL)$

Review *ForAllX*.

# Quantifiers

## Definition

Quantifiers are also operators, but they **bind** variables: they set restrictions on how a variable may be interpreted.

## Definition

Quantifiers are also operators, but they **bind** variables: they set restrictions on how a variable may be interpreted.

All students are male.

$\forall x [student(x) \rightarrow male(x)]$

## Definition

Quantifiers are also operators, but they **bind** variables: they set restrictions on how a variable may be interpreted.

All students are male.

$\forall x [student(x) \rightarrow male(x)]$

## Definition

The **universal quantifier** formalizes the notion that something may be true for everything (within a particular scope). Basically,  $\forall x.$  means something like “for all individuals in the universe of discourse the following holds”.

# Quantifiers

There is a female student.

$\exists y [student(y) \wedge female(y)]$

# Quantifiers

There is a female student.

$\exists y [student(y) \wedge female(y)]$

## Definition

The **existential quantifier** concerns the notion where at least one member of domain is relevant (given a particular scope). Basically,  $\exists x$ . means something like “for at least one individual in the universe of discourse the following holds”.

# Quantifiers

There is a female student.

$\exists y [student(y) \wedge female(y)]$

## Definition

The **existential quantifier** concerns the notion where at least one member of domain is relevant (given a particular scope). Basically,  $\exists x$ . means something like “for at least one individual in the universe of discourse the following holds”.

The symbol  $\exists$  means “at least one”, but sometimes you’ll see  $\exists!$  which means exactly one (uniqueness quantification):

$\exists!x [star(x) \wedge loc(x, SSYST)]$

# Quantifiers: General form

## QUANTIFIER VARIABLE SCOPE

- all tigers
- some tiger
- all man-eating tigers

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Quantifiers: General form

## QUANTIFIER VARIABLE SCOPE

- all tigers  
 $\forall$
- some tiger  
 $\exists$
- all man-eating tigers  
 $\forall$

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Quantifiers: General form

## QUANTIFIER **VARIABLE** SCOPE

- all tigers  
 $\forall t$
- some tiger  
 $\exists t$
- all man-eating tigers  
 $\forall t$

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Quantifiers: General form

## QUANTIFIER VARIABLE SCOPE

- all tigers  
 $\forall t \text{ tiger}(t)$
- some tiger  
 $\exists t \text{ tiger}(t)$
- all man-eating tigers  
 $\forall t [\text{tiger}(t) \wedge \text{maneater}(t)]$

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Quantifiers: General form

## QUANTIFIER VARIABLE SCOPE

- all tigers  
 $\forall t \text{ tiger}(t)$
- some tiger  
 $\exists t \text{ tiger}(t)$
- all man-eating tigers  
 $\forall t [\text{tiger}(t) \wedge \text{maneater}(t)]$

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Quantifiers: General form

## QUANTIFIER **VARIABLE** SCOPE

- all tigers  
 $\forall t \text{ tiger}(t)$
- some tiger  
 $\exists t \text{ tiger}(t)$
- all man-eating tigers  
 $\forall t [\text{tiger}(t) \wedge \text{maneater}(t)]$

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Quantifiers: General form

## QUANTIFIER VARIABLE SCOPE

- all tigers  
 $\forall t \text{ tiger}(t)$
- some tiger  
 $\exists t \text{ tiger}(t)$
- all man-eating tigers  
 $\forall t [\text{tiger}(t) \wedge \text{maneater}(t)]$

## Definition

We say that the quantifier **binds** a given variable in a quantified expression.

# Towards a formal definition of QL

See *ForAllX* for a formal definition of QL.

The equality symbol is used to indicate that two constants are the same in the *UD*. That is, they are identical.

## Example

Hillary is Mrs. Clinton.

*HILLARY = CLINTON*

The evening star is Hesperus.

*EVESTAR = HESPERUS*

# A more complex example

*The car is red.*

$\exists x [car(x) \wedge \neg \exists y [car(y) \wedge x \neq y] \wedge red(x)]$

The formula makes three claims:

# A more complex example

*The car is red.*

$$\exists x [car(x) \wedge \neg \exists y [car(y) \wedge x \neq y] \wedge red(x)]$$

The formula makes three claims:

- 1 There is a car. (an existence claim)
- 2 At most one thing is a car. (a uniqueness claim)
- 3 This car is red. (a claim of predication)