Foundations of Language Science and Technology

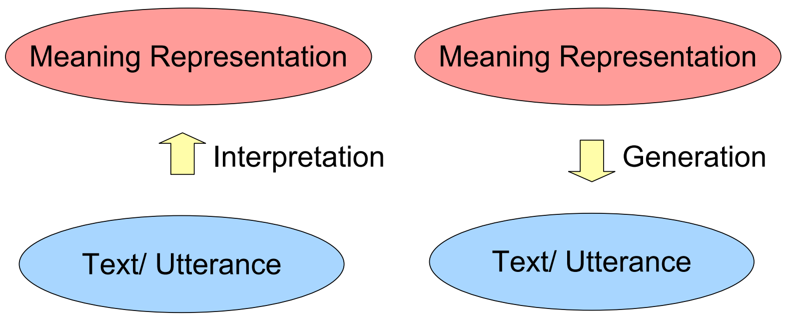
Semantics 1

Manfred Pinkal
Saarland University

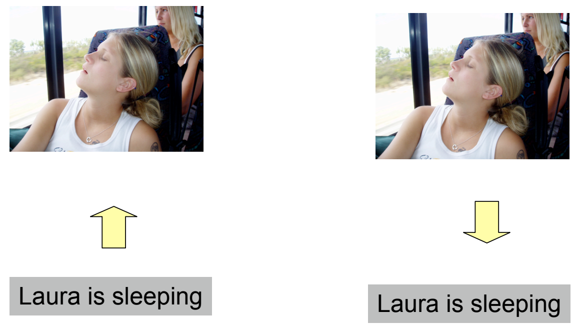
Overview

- Semantic Processing - Introduction
- Logic-based meaning representation and processing: Truth-conditional interpretation, entailment, deduction
- Word Meaning: Lexical-semantic resources, ontologies, similarity-based approaches
- Semantic Composition: Composing sentence and text meaning from word meaning
- Textual Entailment and Inference

Semantic Processing



Semantic Processing



Semantic Processing

They start at 10.

- Aspects of semantic interpretation:
  - Abstraction from linguistic surface
  - Disambiguation / Contextual resolution
  - Inference: Inferring (situationally) relevant information from meaning information encoded in an utterance.

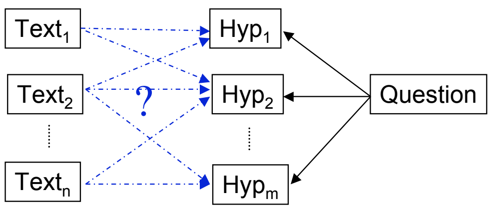
The Role of Semantics in LT

Role of Semantics in Language Technology

- *Dialogue Processing*
- *Machine Translation*
- *Information access (IR, IE, Q&A)*

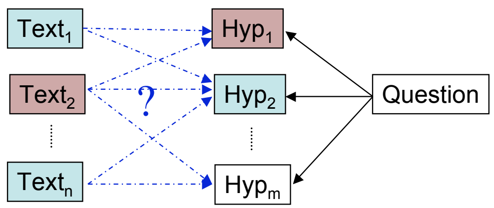
Textual Entailment

- Schematically, we can reduce information access tasks to comparisons between two pieces of natural language.
- Take Question Answering as an example:
- Find those documents/passages of text T which contain a piece of information H that can count as possible answer to question Q. Calling a possible answer a hypothesis, we reformulate the task in the following way:
- For which pairs of texts T and hypotheses H, T (most likely) entails H (H is contained in T, H can be inferred from T) ?



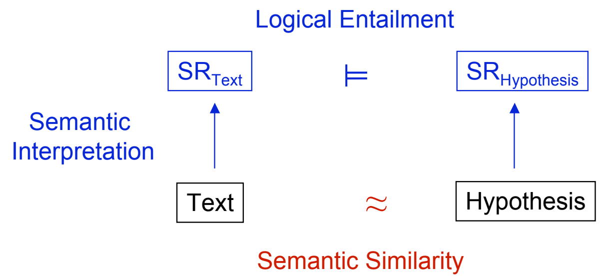
Textual Entailment

- Schematically, we can reduce information access tasks to comparisons between two pieces of natural language.
- Take Question Answering as an example:
- Find those documents/passages of text T which contain a piece of information H that can count as possible answer to question Q. Calling a possible answer a hypothesis, we reformulate the task in the following way:
- Check for all pairs of texts T and hypotheses H, whether T (most likely) entails H (H is contained in T, H can be inferred from T) ?



Two methods

- There are two basic methods in use to check entailment or containment:
  - Computing semantic similarity between T and H
- Answer type 1: Yes, if T entails H, i.e., if H is true in every circumstance (situation, state of the world) in which T is true.



Two methods

- There are two basic approaches in use to check entailment or containment:
  - Computing semantic similarity between T and H: Do T and H contain sufficiently similar lexical material?
  - Testing for logical entailment between T and H: Can H be inferred from T by rules of logic?

The similarity-based approach

Question: *Are dolphins mammals?*

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time. At birth a bottlenose dolphin calf ...

≈

Dolphins are mammals.

Bag-of-words approach
Vector-space models (TF/IDF)
Word overlap, similarity relations between words

Two methods

Two methods

Logical Entailment

SR_{Text} $\models$ $SR_{Hypothesis}$

Semantic Interpretation ↑ ↑

Text ≈ Hypothesis

Semantic Similarity

Are dolphins mammals?

Logical Entailment: ✓

$\forall d\ (\text{dolphin}(d) \rightarrow \text{mammal}(d) \wedge \neg \text{fish}(d))$ $\models$ $\forall d\ (\text{dolphin}(d) \rightarrow \text{mammal}(d))$

Semantic Interpretation ↑ ↑

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time. At birth a bottlenose dolphin calf ...

Dolphins are mammals.

Is Flipper a mammal?

$\forall d\ (\text{dolphin}(d) \rightarrow \text{mammal}(d) \wedge \neg \text{fish}(d))$

$\text{dolphin}(\text{flipper}) \rightarrow \text{mammal}(\text{flipper}) \wedge \neg \text{fish}(\text{flipper})$ $\text{dolphin}(\text{flipper})$

$\text{mammal}(\text{flipper}) \wedge \neg \text{fish}(\text{flipper})$

$\text{mammal}(\text{flipper})$

Logical Entailment: ✓

$\forall d\ (\text{dolphin}(d) \rightarrow \text{mammal}(d) \wedge \neg \text{fish}(d))$ $\models$ $\text{mammal}(\text{flipper})$

Semantic Interpretation ↑ ↑

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time. At birth a bottlenose dolphin calf ...

Flipper is a mammal.

Are dolphins fish?

Logical Entailment: ✗

$\forall d\ (\text{dolphin}(d) \rightarrow \text{mammal}(d) \wedge \neg \text{fish}(d))$ $\models$ $\text{fish}(\text{flipper})$

Semantic Interpretation ↑ ↑

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time. At birth a bottlenose dolphin calf ...

Flipper is a fish.

Are dolphins fish?

Dolphins are mammals, not fish. They are warm blooded like man, and give birth to one baby called a calf at a time. At birth a bottlenose dolphin calf ...

? $\approx$

Dolphins are fish.

Overview

- Semantic Processing - Introduction
- Logic-based meaning representation and processing: Truth-conditional interpretation, entailment, deduction
  - First-order predicate as a representation language
  - Truth-conditional interpretation
  - The logical entailment concept
  - Deduction systems and theorem provers
- Word Meaning: Lexical-semantic resources, ontologies, similarity-based approaches
- Semantic Composition: Compong sentence and text meaning from word meaning
- Textual Entailment and Inference

Predicate Logic – Vocabulary

- The vocabulary of the language of predicate logic:
  - Logical symbols
    - Connectives: ¬, ∧, ∨, →, ↔
    - Quantifiers: ∀, ∃
    - Equality: =
  - Infinite set of individual variables:
    VAR = { x, y, z, … }
  - Arbitrary set of individual constants:
    CON = { a, b, c, … }
  - For every n ≥ 0, an arbitrary, possibly empty set of n-ary predicate symbols: $PRED^n$ = { P, Q, … }

Predicate Logic – Syntax

- Terms: TERM = CON ∪ VAR
- (Well-formed) Formulae: the smallest set such that
  (1) If R is an n-ary predicate symbol, and $t_1, …, t_n$ are terms, then $R(t_1, …, t_n)$ is a wff.
  (2) If t_1, t_2 are terms, then $t_1 = t_2$ is a wff.
  (3) if φ, ψ are wff, then ¬φ, (φ ∧ ψ), (φ ∨ ψ), (φ → ψ), and (φ ↔ ψ) are wff.
  (4) if φ is a wff, and x an individual variable, then ∀xφ and ∃xφ are wff.

Predicate Logic – Atomic formulae

(1) If R is an n-ary predicate symbol, and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a wff.

- *Flipper is a dolphin*
- *Bill works*
- *Saarbrücken is fascinating*
- *Mary likes John*
- *John is taller than Bill*
- *John introduces Bill to Mary*
- *Saarbrücken is closer to Paris than Munich is to London*

(2) If t_1, t_2 are terms, then $t_1 = t_2$ is a wff.

- *The Morning Star is the Evening Star*
- *The German chancellor is Angela Merkel*

Predicate Logic – Complex Formulae

(3) if φ, ψ are wff, then $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$ are wff.

Name	Connective	NL Paraphrase
negation	$\neg p$	it is not the case that p
conjunction	$(p \wedge q)$	p and q
disjunction	$(p \vee q)$	p or q
implication	$(p \rightarrow q)$	if p then q
equivalence	$(p \leftrightarrow q)$	p if and only if q

Flipper is not a fish.
If Flipper is a dolphin, he is a mammal.
Saarbrücken is a fascinating city.

Predicate Logic – Complex Formulae

(4) if φ is a wff, and x an individual variable, then $\forall x\varphi$ and $\exists x\varphi$ are wff.

Bill reads an interesting book.
$\exists b\, (book(b) \wedge interesting(b) \wedge read(bill,b))$

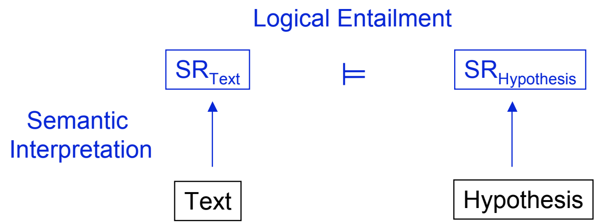
Dolphins are mammals, not fish.
$\forall d\, (dolphin(d) \rightarrow mammal(d) \wedge \neg fish(d))$

Dolphins live in pods.
$\forall d\, (dolphin(d) \rightarrow \exists x\, (pod(p) \wedge live\text{-}in(d,p))$

Dolphins give birth to one baby at a time.
$\forall d\, (dolphin(d) \rightarrow \forall x\, \forall y\, \forall t\, (give\text{-}birth\text{-}to(d,x,t) \wedge give\text{-}birth\text{-}to(d,y,t) \rightarrow x=y)$

Logical Entailment

- (The semantic representation of) T „logically entails“ (the semantic interpretation of H means that H is true in all circumstances or possible states of the worlds in which T is true.



Semantic Interpretation of FOL

- FOL expressions are interpreted with respect to certain situations or states of the world.
- FOL expressions of certain types (terms, relation symbols, formulae) are assigned specific kinds of objects (denotations) by an interpretation function.
- In particular, formulae denote truth values.
- Situations or states of the world (more precisely: the relevant properties of situations and states of the world) are formally represented by model structures.

Model Structures

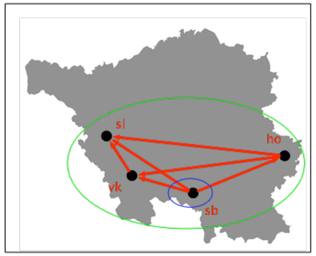
- A model structure is a pair $M = (U_M, V_M)$, where
  - U_M is a non-empty set (the "model universe"), and
  - V_M is an value assignment function for basic expressions, which assigns

    n-ary relations (over U_M) to n-ary predicate symbols, and elements of U_M to predicate constants:

    $V_M(P) \subseteq U_M^n$, if P is an n-ary predicate symbol

    $V_M(c) \in U_M$, if c is a constant

A Model of Saarland Towns



$M = (U_M, V_M)$

$U_M = \{ sl, vk, ho, sb \}$

V_M defined by:

- V_M (saarbrücken) = sb
- V_M (völklingen) = vk
- V_M (saarlouis) = sl
- V_M (homburg) = ho
- V_M(larger_than) = { (sb, sl), (sb, vk), (sb, ho), (vk, sl), … }
- V_M(town) = { sl, vk, ho, sb }
- V_M(capital) = { sb }

Interpretation of Atomic Formulae

- Terms: TERM = CON $\cup$ VAR
- (Well-formed) Formulae: the smallest set such that
  (1) If R is an n-ary predicate symbol, and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a wff.
  (2) If t_1, t_2 are terms, then $t_1 = t_2$ is a wff.
  (3) if φ, ψ are wff, then $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$ are wff.
  (4) if φ is a wff, and x an individual variable, then $\forall x\varphi$ and $\exists x\varphi$ are wff.

Interpretation of Atomic Formulae

- An interpretation function $[\![\ ]\!]^M$ recursively assigns semantic values $[\![\alpha]\!]^M$ to all expressions α.
- Interpretation of terms (preliminary):

  $[\![c]\!]^M = V_M(c)$ for all individual constants c

  $[\![x]\!]^M = ?$
- Variables are assigned members of U_M randomly as temporary, preliminary values by a variable assignment function g.
- Quantifiers overwrite the values of g for the variable they bind (modified variable assignment, see below).
- All FOL expression are interpreted with respect to a model structure M plus a variable assignment g.

Interpretation of Atomic Formulae

- An interpretation function $[\![\ ]\!]^{M,g}$ recursively assigns semantic values $[\![\alpha]\!]^{M,g}$ to all expressions α with respect to a model structure and a variable assignment g.
- Interpretation of terms:

  $[\![c]\!]^{M,g} = V_M(c)$ for all individual constants c

  $[\![x]\!]^{M,g} = g(x)$
- Interpretation of atomic expressions:

  $[\![R(t_1, ..., t_n)]\!]^{M,g} = 1$ iff $([\![t_1]\!]^{M,g}, ..., [\![t_n]\!]^{M,g}) \in V_M(R)$

  $[\![t_1 = t_2]\!]^{M,g} = 1$ iff $[\![t_1]\!]^{M,g} = [\![t_2]\!]^{M,g}$

  Example:

  larger_than(saarbrücken, homburg) =1

  iff ($[\![$ saarbrücken $]\!]^{M,g}$, ..., $[\![$ homburg $]\!]^{M,g} \in V_M$(larger_than)

  iff (V_M(saarbrücken), ..., V_M(homburg)) $\in V_M$(larger_than)

  iff (sb, ho) $\in V_M$(larger_than)

Predicate Logic – Syntax

- Terms: TERM = CON ∪ VAR
- (Well-formed) Formulae: the smallest set such that
  - (1) If R is an n-ary predicate symbol, and $t_1, ..., t_n$ are terms, then $R(t_1, ..., t_n)$ is a wff.
  - (2) If t_1, t_2 are terms, then $t_1 = t_2$ is a wff.
  - (3) if φ, ψ are wff, then $\neg\varphi$, $(\varphi \wedge \psi)$, $(\varphi \vee \psi)$, $(\varphi \rightarrow \psi)$, and $(\varphi \leftrightarrow \psi)$ are wff.
  - (4) if φ is a wff, and x an individual variable, then $\forall x\varphi$ and $\exists x\varphi$ are wff.

Interpretation of connectives

$[\![\neg\varphi]\!]^{M,g} = 1$ iff $[\![\varphi]\!]^{M,g} = 0$

$[\![\varphi \wedge \psi]\!]^{M,g} = 1$ iff $[\![\varphi]\!]^{M,g} = 1$ and $[\![\psi]\!]^{M,g} = 1$

$[\![\varphi \vee \psi]\!]^{M,g} = 1$ iff $[\![\varphi]\!]^{M,g} = 1$ or $[\![\psi]\!]^{M,g} = 1$

$[\![\varphi \rightarrow \psi]\!]^{M,g} = 1$ iff $[\![\varphi]\!]^{M,g} = 0$ or $[\![\psi]\!]^{M,g} = 1$

$[\![\varphi \leftrightarrow \psi]\!]^{M,g} = 1$ iff $[\![\varphi]\!]^{M,g} = [\![\psi]\!]^{M,g}$

- Connectives in predicate logic are truth-functional: Their truth-value is completely determined by the truth-values of their constituent clauses.
- The interpretation of connectives can be represented by truth-tables.

Truth Tables for Connectives

A	¬A
0	1
1	0

A	B	(A∧B)
0	0	0
0	1	0
1	0	0
1	1	1

A	B	(A∨B)
0	0	0
0	1	1
1	0	1
1	1	1

A	B	(A→B)
0	0	1
0	1	1
1	0	0
1	1	1

A	B	(A↔B)
0	0	1
0	1	0
1	0	0
1	1	1

Composite Truth Tables

	A	B	¬A	¬B	¬A ∧ ¬B	¬(¬A ∧ ¬B)
M_1	1	1	0	0	0	1
M_2	1	0	0	1	0	1
M_3	0	1	1	0	0	1
M_4	0	0	1	1	1	0

Predicate Logic – Syntax

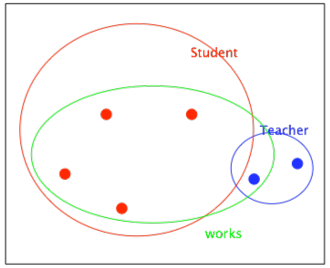
- Terms: TERM = CON ∪ VAR
- (Well-formed) Formulae: the smallest set such that
  (1) If R is an n-ary predicate symbol, and $t_1, \ldots, t_n$ are terms, then $R(t_1, \ldots, t_n)$ is a wff.
  (2) If t_1, t_2 are terms, then $t_1 = t_2$ is a wff.
  (3) if φ, ψ are wff, then ¬φ, (φ ∧ ψ), (φ ∨ ψ), (φ → ψ), and (φ ↔ ψ) are wff.
  (4) if φ is a wff, and x an individual variable, then ∀xφ and ∃xφ are wff.

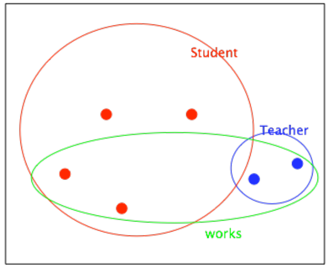
Quantifier Interpretation– Preliminary!

- Interpretation of quantifiers should return value 1:
  for ∃xF(x) iff $a \in V_M(F)$ for at least one $a \in U_M$.
  for ∀xF(x) iff $a \in V_M(F)$ for all $a \in U_M$.
- A preliminary formulation of a general interpretation function for quantified formulae:
  $[[\exists xA]]^{M,g} = 1$ iff
  there is at least one variable assignment g' such that $[[A]]^{M,g'} = 1$
  $[[\forall xA]]^{M,g} = 1$ iff
  $[[A]]^{M,g'} = 1$ for all variable assignments g'.

An Example

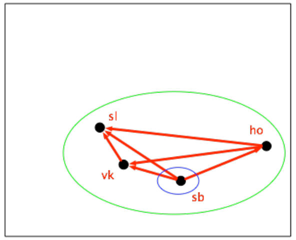
- "Every student works" $\Rightarrow \forall x(student(x) \rightarrow work(x))$
- True in model M1, false in model M2.
- M1 M2





Another example

- $[[\ \exists x(town(x) \wedge larger\_than(x, völklingen))\ ]]^{M,g} = 1$
  iff there is g' such that
  $[[\ town(x) \wedge larger\_than(x, völklingen))\ ]]^{M,g'} = 1$
  $\Leftrightarrow [[\ town(x)]]^{M,g'} = 1 \wedge [[\ larger\_than(x, völklingen))\ ]]^{M,g'} = 1$
  $\Leftrightarrow [[x]]^{M,g'} \in V_M(town)$ and
  $([[x]]^{M,g'}, [[völklingen]]^{M,g'}) \in V_M(larger\_than)$
  $\Leftrightarrow g'(x) \in V_M(town)$ and
  $(g'(x), V_M(völklingen)) \in V_M(larger\_than)$
- $\exists x(town(x) \wedge larger\_than(x, völklingen))$
  is true in the Saarland model: Saarbrücken and Homburg are verifying instantiations for x



Variable Assignments

- Attention: The interpretion function for quantifiers is incorrect for the general case of formulae containing several nested quantifiers. We need a notion of a modified variable assignment function. We do not have the time to treat it in the course. Definitions and examples ar added for completeness. They will not be part of the exam.
- Let $M = (U_M, V_M)$ be a model structure.
- A variable assignment is a function g:
  $VAR \rightarrow U_M$ that maps variables to elements of U_M.
- g[x/u] stands for the assignment g' which differs from g at most in that g'(x) = u
  - g[x/u](y) = u if x=y
  - g[x/u](y) = g(y) otherwise

Variable Assignment, Examples

	x	y	z	u	…
g	a	b	c	d	…
g[x/a]	a	b	c	d	…
g[y/a]	a	a	c	d	…
g[y/g(z)]	a	c	c	d	…
g[y/a][u/a]	a	a	c	a	…
g[y/a][y/b]	a	b	c	d	…

Interpretation of Terms

- Let $M = (U_M, V_M)$ be a model structure for some language L of predicate logic.
- The function $[[\ ]]^{M,g}$ interprets the terms of L as follows:
  - $[[x]]^{M,g} = g(x)$, if x is a variable
  - $[[c]]^{M,g} = V_M(c)$, if c is a constant

Interpretation of Formulae

- $[[R(t_1, ..., t_n)]]^{M,g} = 1$ iff $([[t_1]]^{M,g}, ..., [[t_n]]^{M,g}) \in V_M(R)$
- $[[s = t]]^{M,g} = 1$ iff $[[s]]^{M,g} = [[t]]^{M,g}$
- $[[\neg\varphi]]^{M,g} = 1$ iff $[[\varphi]]^{M,g} = 0$
- $[[\varphi \wedge \psi]]^{M,g} = 1$ iff $[[\varphi]]^{M,g} = 1$ and $[[\psi]]^{M,g} = 1$
- $[[\varphi \vee \psi]]^{M,g} = 1$ iff $[[\varphi]]^{M,g} = 1$ or $[[\psi]]^{M,g} = 1$
- $[[\varphi \rightarrow \psi]]^{M,g} = 1$ iff $[[\varphi]]^{M,g} = 0$ or $[[\psi]]^{M,g} = 1$
- $[[\varphi \leftrightarrow \psi]]^{M,g} = 1$ iff $[[\varphi]]^{M,g} = [[\psi]]^{M,g}$
- $[[\exists x\Phi]]^{M,g} = 1$ iff there is an $a \in U_M$ s.t. $[[\Phi]]^{M,g[x/a]} = 1$
- $[[\forall x\Phi]]^{M,g} = 1$ iff for all $a \in U_M$, $[[\Phi]]^{M,g[x/a]} = 1$