

Einführung in die Computerlinguistik: Semantik

Manfred Pinkal / Sebastian Pado

Semantik

Wann fängt heute die CL-Vorlesung an?

Sie beginnt um 11.

Semantik ... für Dialogverarbeitung [2]

- *Ist der Flug um 2 nach London pünktlich?*
- *Ihr Flug nach London Heathrow geht pünktlich um 14:05 ab und steht zum Einsteigen bereit.*
- *Und der nächste?*

Semantik ... für Dialogverarbeitung [2]

- *Ist der Flug um 2 nach London pünktlich?*
- *Ihr Flug nach London Heathrow geht pünktlich um 14:05 ab und steht zum Einsteigen bereit.*
- *Und der nächste?*
- *Der nächste Flug nach London Gatwick geht um 15.20.*
- *Gibt es auch einen nach Heathrow?*

Semantik ... für Maschinelle Übersetzung

- *Geht es um 2 bei Ihnen?*
- *Is 2 p.m. ok for you?*
- *Can we meet at 2 in your office?*

Semantik ... für Informationszugriff und Informationsmanagement

Der US-Flugzeughersteller Lockheed hat von Grossbritannien den Auftrag fuer 25 Transportflugzeuge des Typs Hercules C130J erhalten. Vertreter des Verteidigungsministeriums bezifferten den Wert des Auftrags mit umgerechnet 2,5 Mrd. DM.

- *Wieviel kosten die Flugzeuge, die Lockheed an Großbritannien verkauft hat?*

Semantik: Grundlegende Fragestellungen

- Wie kann Bedeutungsinformation grundsätzlich dargestellt werden? (Semantische Repräsentation)
- Wie wird die Bedeutung von komplexen Ausdrücken (Sätzen) auf der Grundlage von Wortbedeutungen und der syntaktischen Struktur berechnet? (Satzsemantik, Semantikkonstruktion)

Semantische Phänomene [1]

- Quantoren und Skopus
 - (1) *Zwei Fremdsprachen spricht jeder Linguist.*
 - (2) *Alle sitzen in einem Boot*
 - (3) *Jeder sitzt in einem Boot*
 - (4) *Ein Teilnehmer von jedem Seminar hält zwei Referate.*
 - (5) *800 000 Besucher tranken 2 Millionen Tassen Kaffee (kollektive und kumulative Lesarten)*

Semantische Phänomene [2]

- Sorten- /Selektionsbeschränkungen

(6) Fahren wir nach Stuttgart

(7) Treffen wir uns nach Ostern!

- Aber:

(8) Treffen wir uns nach Stuttgart!

Nicht-wörtliche Bedeutung, Re-Interpretation

Semantische Phänomene [3]

- Nicht-wörtliche Bedeutung: Metapher und Metonymie
 - (9) *Börsenkurse auf Talfahrt*
 - (10) *Der Picasso hängt im ersten Stock*
 - (11) *Ich stehe hinter dem Haus*

Semantik: Grundlegende Fragestellungen

- Wie kann Bedeutungsinformation grundsätzlich dargestellt werden? ([Semantische Repräsentation](#))
- Wie wird die Bedeutung von komplexen Ausdrücken (Sätzen) auf der Grundlage von Wortbedeutungen und der syntaktischen Struktur berechnet? ([Satzsemantik](#), [Semantikkonstruktion](#))
- [Wie ergibt sich die spezifische Bedeutung einer Äußerung im Kontext? Wie verändert eine Äußerung den Kontext?](#)
([Diskurssemantik](#))

Semantische Phänomene [4]

- Diskurssemantik: Anaphern
 - (12) *Die nächste Computerlinguistik-Vorlesung macht Sebastian Pado. Sie findet am Dienstag statt.*
 - (13) *Hans hat Peter nicht begrüßt. Er ist beleidigt.*

Semantische Phänomene [5]

- Diskurssemantik: Ellipsen
 - (14) *Hans verwöhnt seinen Hund, und Peter seine Katze.*
 - (15) *Hans verwöhnt seinen Hund, und Peter auch.*
 - (16) *Und der nächste?*

Semantik: Grundlegende Fragestellungen

- Wie kann Bedeutungsinformation grundsätzlich dargestellt werden? ([Semantische Repräsentation](#))
- Wie wird die Bedeutung von komplexen Ausdrücken (Sätzen) auf der Grundlage von Wortbedeutungen und der syntaktischen Struktur berechnet? ([Satzsemantik](#), [Semantikkonstruktion](#))
- Wie ergibt sich die spezifische Bedeutung einer Äußerung im Kontext? Wie verändert eine Äußerung den Kontext? ([Diskurssemantik](#))
- [Wie wird Mehrdeutigkeit aufgelöst? \(Ambiguitätsresolution\)](#)

Semantische Phänomene [6]

- Anaphern
(17) Die nächste Computerlinguistik-Vorlesung macht Sebastian Pado. Sie findet am Dienstag statt.
- Mehrdeutigkeit des anaphorischen Bezuges ...
(18) Hans hat Peter nicht begrüßt. Er ist beleidigt.
- ... und der komplexe Mechanismus der Anaphernresolution:
(19) Hans mag seinen Hund, obwohl er ihn manchmal beißt.

Semantik: Grundlegende Fragestellungen

- Wie kann Bedeutungsinformation grundsätzlich dargestellt werden? ([Semantische Repräsentation](#))
- Wie wird die Bedeutung von komplexen Ausdrücken (Sätzen) auf der Grundlage von Wortbedeutungen und der syntaktischen Struktur berechnet? ([Satzsemantik](#), [Semantikkonstruktion](#))
- Wie ergibt sich die spezifische Bedeutung einer Äußerung im Kontext? Wie verändert eine Äußerung den Kontext? ([Diskurssemantik](#))
- Insbesondere: Wie wird Mehrdeutigkeit aufgelöst? ([Ambiguitätsresolution](#))
- Wie kommt man von der spezifischen Äußerungsbedeutung zur relevanten Information? ([Inferenz](#), [semantische Auswertung](#))
- [Was ist Wortbedeutung und wie kann sie erfasst werden?](#) ([lexikalische Semantik](#))

Was ist Wortbedeutung?



Lexikalische Semantik

*Der US-Flugzeughersteller Lockheed hat von Grossbritannien den Auftrag fuer 25 **Flugzeuge** des Typs Hercules C130J erhalten. Vertreter des Verteidigungsministeriums bezifferten den Wert des Auftrags mit umgerechnet 2,5 Mrd. DM.*

*Wieviel kosten die **Maschinen**, die Lockheed an Großbritannien verkauft hat?*

Lexikalische Semantik

*Der US-Flugzeughersteller Lockheed hat von Grossbritannien den Auftrag fuer 25 **Abfangjäger** erhalten. Vertreter des Verteidigungsministeriums bezifferten den Wert des Auftrags mit umgerechnet 2,5 Mrd. DM.*

- *Wieviel kosten die **Flugzeuge**, die Lockheed an Großbritannien verkauft hat?*

Funktionswörter

FOL (Prädikatenlogik 1.Stufe → Satzsemantik)

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 **is** about 90-130 cms long **and** will grow to approx. 4 metres, living up to 40 years. They **are** highly sociable animals, living in pods **which are** fairly fluid, with dolphins from other pods interacting with each other from time to time.

Funktionswörter

FOL+Anaphern (→ Diskurssemantik)

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 **is** about 90-130 cms long **and** will grow to approx. 4 metres, living up to 40 years. **They are** highly sociable animals, living in pods **which are** fairly fluid, with dolphins from **other** pods interacting with **each other** from time to time.

Funktionswörter

FOL+Anaphern+Präpositionen

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 **is** about 90-130 cms long **and** will grow **to** approx. 4 metres, living **up to** 40 years. **They are** highly sociable animals, living **in** pods **which are** fairly fluid, **with** dolphins **from other** pods interacting **with each other from** time **to** time.

Inhaltswörter: Einstellige Prädikate

FOL+Anaphern+Präpositionen+Anderes

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 **is about** 90-130 cms long **and will** grow **to approx.** 4 metres, living **up to** 40 years. **They are highly** sociable animals, living **in** pods **which are fairly** fluid, **with** dolphins **from other** pods interacting **with each other from** time **to** time.

Inhaltswörter: Einstellige Prädikate

Gattungssubstantive (→ Lexikalische Semantik)

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 is about 90-130 cms long and will grow to approx. 4 metres, living up to 40 years. They are highly sociable animals, living in pods which are fairly fluid, with dolphins from other pods interacting with each other from time to time.

Lexikalische Semantik: WordNet

- eine große lexikalische Datenbank für das Englische (z.Zt. >150.000 Wörter; dt. Version: GermaNet, >100.000 Wörter)

WordNet

- eine große lexikalische Datenbank für das Englische (z.Zt. >150.000 Wörter; dt. Version: GermaNet, >100.000 Wörter)
- Hierarchische Struktur des Lexikons, Ober- und Unterbegriffe (Hyponyme/Hyperonyme)
- Einheiten von WordNet sind „**synsets**“: Mengen von synonymen Ausdrücken, die Lesarten von Wörtern charakterisieren...

Beispiel: engl. *case*

- {*case, carton*}
- {*case, bag, suitcase*}
- {*case, pillowcase, slip*}
- {*case, cabinet, console*}
- {*case, casing* (the enclosing frame around a door or window opening)}
- {*case* (a small portable metal container)}

Beispiel: engl. *tree*

{ tree, plant @, conifer ~, oak ~ ... }

{ plant, flora, organism @, tree ~ ... }

WordNet: Oberbegriffe im Nominalbereich

Table 1
List of 25 unique beginners for WordNet nouns

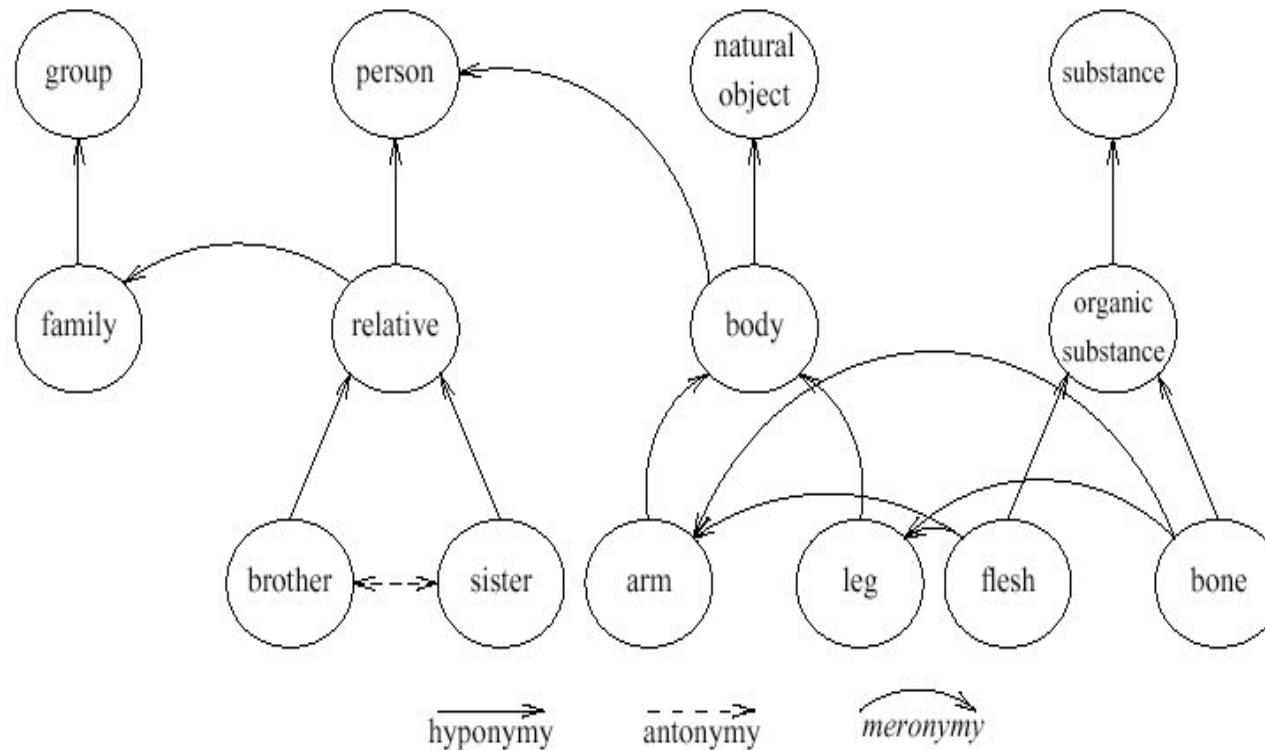
{ <i>act, action, activity</i> }	{ <i>natural object</i> }
{ <i>animal, fauna</i> }	{ <i>natural phenomenon</i> }
{ <i>artifact</i> }	{ <i>person, human being</i> }
{ <i>attribute, property</i> }	{ <i>plant, flora</i> }
{ <i>body, corpus</i> }	{ <i>possession</i> }
{ <i>cognition, knowledge</i> }	{ <i>process</i> }
{ <i>communication</i> }	{ <i>quantity, amount</i> }
{ <i>event, happening</i> }	{ <i>relation</i> }
{ <i>feeling, emotion</i> }	{ <i>shape</i> }
{ <i>food</i> }	{ <i>state, condition</i> }
{ <i>group, collection</i> }	{ <i>substance</i> }
{ <i>location, place</i> }	{ <i>time</i> }
{ <i>motive</i> }	

Weitere Relationen

- Meronymie
 - Teil-Ganzes (*Ast - Baum*)
 - Gruppe - Element (*Baum - Wald*)
 - Stoff - Objekt (*Holz - Baum*)
- Antonymie
 - *Mann - Frau*
 - *Vater - Mutter*
- Funktion
 - Messer - schneiden
 - Wolle - stricken

WordNet-Teilhierarchie, Beispiel

Figure 2. Network representation of three semantic relations among an illustrative variety of lexical concepts



Weitere Ressourcen für Wortbedeutungsinformation

- Einsprachige Wörterbücher (Duden, Wahrig)
- Thesaurus (Dornseiff, Wehrle-Eggers)
- Ontologien (SUMO)

Ontologie

- „Hierarchisch geordnetes Begriffsinventar“
 - Elemente sind nicht Wörter, sondern Konzepte
 - Anordnungsprinzip ist nicht alphabetisch, sondern semantisch (Hierarchien, weitere Relationen)
 - Domänenspezifische Fachontologien vs. allgemeinsprachliche/ domänenunabhängige/ upper-level Ontologien

Weitere Ressourcen für Wortbedeutungsinformation

- Einsprachige Wörterbücher (Duden, Wahrig)
- Thesaurus (Dornseiff, Wehrle-Eggers)
- Ontologien (SUMO, MILO)
- Ressourcen, die Prädikat-Argument-Struktur einbeziehen, Beispiele: FrameNet, SALSA