

# Einführung in die Computerlinguistik Semantik II

Hinrich Schütze & Robert Zangenfeind

Centrum für Informations- und Sprachverarbeitung, LMU München

2015-10-23

# Take-away

- Lexikalische Semantik:  
die Lehre von der Bedeutung der Wörter
- Lexika: ihre vielleicht wichtigste Funktion ist die Erklärung der Bedeutung von Wörtern
- Lexikographie beschäftigt sich mit Theorie und Praxis der Lexika
- Zwei verschiedene Ansätze zur Lexikographie, die aus der Linguistik/Philosophie kommen: (i) Semantische Metasprache, (ii) Merkmalsemantik
- Semantische Relationen zwischen Lexemen, z.B. Synonymie
- Logik: Aussagenlogik, Prädikatenlogik
- Anwendungen

# Overview

- 1 Lexikalische Semantik
- 2 Semantische Relationen
- 3 Logik
- 4 Anwendungen

# Outline

- 1 Lexikalische Semantik
- 2 Semantische Relationen
- 3 Logik
- 4 Anwendungen

# Lexical semantics

- Lexical semantics:  
the field concerned with the meaning of words
- As opposed to: how the meaning of complex linguistic expressions / phrases is composed from the meanings of their parts.
- Different approaches to lexical semantics
  - Traditional lexicography
  - (Perhaps the most important job of a dictionary is to explain the meaning of words.)
  - Axiomatic lexicography (this lecture)
  - Montague semantics: the meaning of “elephant” is a set of all individuals that are elephants
  - Wittgenstein: the meaning of a word is the way we use it
  - Many others

# Dictionary

- “A book that lists the words of a language in alphabetical order and gives their meaning, or that gives the equivalent words in a different language” (OED)
- This lecture: monolingual dictionaries
- In addition to “meaning”, dictionaries usually also give some grammatical information, e.g., part of speech, gender, inflectional paradigms of nouns and verbs etc
- Also: genre, register
- Also: usage notes
- Also: etymology

# Dictionary entry: Example

## Baum, der | Baum |

der Baum; Genitiv: des Baum[e]s, Bäume

HERKUNFT mittelhochdeutsch, althochdeutsch boum, Herkunft ungeklärt

- 1** Holzgewächs mit festem Stamm, aus dem Äste wachsen, die sich in Laub oder Nadeln tragende Zweige teilen

*die Bäume werden grün, verlieren ihr Laub | einen Baum fällen | er ist stark wie ein Baum (sehr stark)*

**Bäume ausreißen [können]** UMGANGSSPRACHLICH sehr viel leisten können, sodass einem nichts zu anstrengend ist

**vom Baum der Erkenntnis essen** durch Erfahrung klug, wissend werden nach 1. Mose 2, 9 einer der beiden mit Namen benannten Bäume im Garten Eden, von denen zu essen Gott Adam und Eva verboten hatte; das Essen von diesem Baum ist im A. T. das Bild für den Ungehorsam des Menschen gegen Gott, die erste Sünde des Menschen

# Lexicography

- Practical lexicography is the art or craft of compiling, writing and editing dictionaries.
- Theoretical lexicography is the scholarly discipline of analyzing and describing the semantic, syntagmatic and paradigmatic **relationships within the lexicon** (vocabulary) of a language, developing **theories of dictionary components and structures** linking the data in dictionaries, the needs for information by **users** in specific types of situation, and how users may best access the data incorporated in printed and electronic dictionaries.
- (from Wikipedia)



# Lexicography as a rigorous science

- Lexicography in practice is often more of an art and a craft: century-old traditions, best practices, staff educated in humanities departments
- Alternative: “axiomatic” approach
- Construct a metalanguage consisting of basic concepts, combination rules, axioms
- Define the words of the language in terms of this metalanguage
- Modeled on mathematics and hard sciences

# Metasprache

- exakte, genau definierte Kunstsprache
- begrenzte Anzahl von Wörtern
- idealerweise Seme, d.h. “kleinste Bedeutungseinheiten”, elementare, nicht weiter zerlegbare Bedeutungseinheiten
- spezielle Syntax

## Traditional lexicography: No metalanguage

- Duden definitions of “zusagen”, “garantieren”
- **zusagen** 1b: 'jemandem **zusichern**, sich in einer bestimmten Angelegenheit seinen Wünschen entsprechend zu verhalten, ihm etwas zuteilwerden zu lassen'
- **zusichern**: '[offiziell] etwas Gewünschtes oder Gefordertes als sicher **zusagen**; **garantieren** (a)'
- **garantieren** a: '(durch Versprechen) fest **zusichern**, **zusagen**'
- circular, something that “axiomatic” lexicography does not allow

# “Axiomatische” Lexikographie: Anforderungen

- Bedeutungserklärungen müssen zirkuläre Worterklärungen vermeiden.
- Stattdessen: Systematische Beschreibung unter Benutzung eines genau definierten Wortvorrates
- Das ist nur praktikabel, wenn dieser definierte Wortvorrat aus viel weniger Wörtern besteht als die Sprache als Ganzes.
- Ein Ansatz: **Seme** oder semantische Primitiva: Wörter mit elementarer, sehr einfacher Bedeutung
- so genannte Zwischenbegriffe (semantisch komplexere Begriffe)
- Bedeutungserklärung mit jeweils semantisch einfacheren Ausdrücken: Erklärung von Zwischenbegriffen durch Seme, Erklärung von komplexen Begriffen durch Seme und Zwischenbegriffe

## “Axiomatische” Lexikographie: Anforderungen (2)

- keine Synonyme: die führen nur zur Verwirrung, weil man für eine Bedeutung mehrere Wörter verwenden kann
- keine Homonyme: die führen zu mehrdeutigen Definitionen

## Beispiel für Bedeutungszerlegung: “versprechen”

- Schritt 1: semantische Analyse (Dekomposition): A verspricht [der Person] C, dass B = 'A teilt C mit, dass B; B hängt von A ab; C ist an B interessiert'
- Schritt 2: Ersetzung von “mitteilen” durch primitivere Ausdrücke: (A teilt [den Sachverhalt] B [der Person] C mit = 'A verursacht explizit, dass C weiß, dass B; A weiß, dass B'):
- Definition von “versprechen” nach Schritt 2: A verspricht [der Person] C, dass B = 'A verursacht explizit, dass C weiß, dass B; A weiß, dass B; B hängt von A ab; C ist an B interessiert'
- Schritt 3: Ersetzung von “wissen” durch primitivere Ausdrücke: (A weiß, dass B = 'A besitzt die Information B'):
- Definition von “versprechen” nach Schritt 3: A verspricht [der Person] C, dass B = 'A verursacht explizit, dass C die Information B besitzt; A besitzt die Information B; B hängt von A ab; C ist an B interessiert'

# Bedeutungszerlegung

- Basiert auf einer Semantischen Analyse der Bedeutung des Wortes
- Zerlegung / Explizierung, bis keine semantisch einfacheren Ausdrücke in natürlicher Sprache mehr vorgefunden werden
- Schlusspunkt der schrittweisen Analyse: Ebene der Seme
- Seme: Wörter, die sich in der betrachteten Sprache nicht weiter zerlegen lassen

# Semantische Primitiva bei A. Wierzbicka (1)

Substantives:	I, YOU, SOMEONE, PEOPLE, SOMETHING~THING, BODY
Relational substantives:	KIND, PART
Determiners:	THIS, THE SAME, OTHER~ELSE
Quantifiers:	ONE, TWO, SOME, ALL, MUCH~MANY, LITTLE~FEW
Evaluators:	GOOD, BAD
Descriptors:	BIG, SMALL



## Semantische Primitiva bei A. Wierzbicka (2)

Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech:	SAY, WORDS, TRUE
Actions, events, movement, contact:	DO, HAPPEN, MOVE, TOUCH
Location, existence, possession, specification:	BE (SOMEWHERE), THERE IS, HAVE, BE (SOMEONE/SOMETHING)
Life and death:	LIVE, DIE
Time:	WHEN~TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
Space:	WHERE~PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE
Logical concepts:	NOT, MAYBE, CAN, BECAUSE, IF
Intensifier, augmentor:	VERY, MORE
Similarity:	LIKE~AS~WAY

## Semantische Primitiva bei A. Wierzbicka (3)

- Primitiva können sprachspezifische Varianten haben (Allolexe: ~)
- z.B. MUCH~MANY, SOMETHING~THING → SOMETHING mit Determinierer wird zu THING: “this something” = this thing, “one something” = one thing

# Semantic explications (1)

someone X is contented (at this time): someone X thinks like this at this time: “something good is happening to me now I want this I don’t want anything else now” because of this, this someone feels something good at this time like someone can feel when they think like this

## Semantic explications (2)

someone X is happy (at this time): someone X thinks like this at this time: “many good things are happening to me as I want I can do many things now as I want this is good” because of this, this someone feels something good at this time like someone can feel when they think like this

Substantives:	I, YOU, SOMEONE, PEOPLE, SOMETHING~THING, BODY
Relational substantives:	KIND, PART
Determiners:	THIS, THE SAME, OTHER~ELSE
Quantifiers:	ONE, TWO, SOME, ALL, MUCH~MANY, LITTLE~FEW
Evaluators:	GOOD, BAD
Descriptors:	BIG, SMALL
Mental predicates:	THINK, KNOW, WANT, FEEL, SEE, HEAR
Speech:	SAY, WORDS, TRUE
Actions, events, movement, contact:	DO, HAPPEN, MOVE, TOUCH
Location, existence, possession, specification:	BE (SOMEWHERE), THERE IS, HAVE, BE (SOMEONE/SOMETHING)
Life and death:	LIVE, DIE
Time:	WHEN~TIME, NOW, BEFORE, AFTER, A LONG TIME, A SHORT TIME, FOR SOME TIME, MOMENT
Space:	WHERE~PLACE, HERE, ABOVE, BELOW, FAR, NEAR, SIDE, INSIDE
Logical concepts:	NOT, MAYBE, CAN, BECAUSE, IF
Intensifier, augmentor:	VERY, MORE
Similarity:	LIKE~AS~WAY

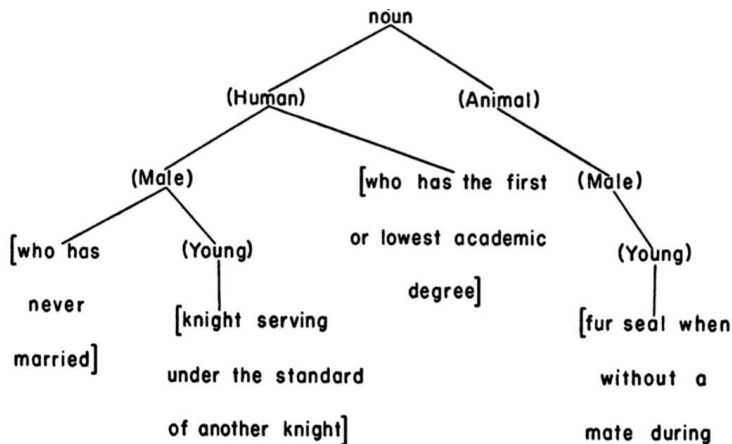
Aufgabe:  
Zerlege "A belügt B"

Beispiel: A verspricht  
[der Person] C, dass B  
= 'A verursacht explizit,  
dass C die Information  
B besitzt; A besitzt die  
Information B; B hängt  
von A ab; C ist an B in-  
teressiert'

# Andere Methode zur Bedeutungserklärung: Merkmalsemantik

- Semantisches Merkmal:  
Bedeutungsatom, Bedeutungskomponente
- Wortbedeutung als Bündel von semantischen Merkmalen; z.B.  
+ menschlich, + erwachsen, + männlich → 'Mann'
- eher Bauelemente der Erkenntnis (Ontologie) als sprachliche Einheiten
- Bisher existiert kein vollständiges System.  
Prinzipiell unmöglich?

# Katz & Fodor



## Semantic feature vectors

	moon	silver fox	fruitbat	owl	homo sapiens
animate	no	yes	yes	yes	yes
furry	no	yes	yes	no	no
silver	yes	yes	no	no	no
can-fly	no	no	yes	yes	no
carnivore	no	yes	no	yes	yes



# Konnotationen – nicht Teil der Kernsemantik

## Definition

an idea or feeling that a word invokes in addition to its literal or primary meaning Example: 'the word "discipline" has unhappy connotations of punishment and repression'

# Konnotationen – nicht Teil der Kernsemantik

- expressive Elemente einer Bedeutung
- semantische Assoziationen, assoziative Merkmale, Gefühlswert, feeling, tone
- diese pragmatischen Elemente reflektieren Verbundenheit eines Wortes mit kulturellen Vorstellungen und Traditionen
- spiegeln gesellschaftsabhängige Praxis der Verwendung bestimmter Gegenstände wider
- oft außersprachliche Faktoren

## Example of a semantic field: “proud”

Proud, arrogant, haughty, lordly, insolent, overbearing, supercilious, disdainful can mean filled with or **showing a sense of one's superiority and scorn for what one regards as in some way inferior.**

# Arrogant

Arrogant implies a disposition to claim for oneself, often domineeringly or aggressively, **more consideration or importance than is warranted** or justly due “The Junker developed into a rude, domineering, arrogant type of man, without cultivation or culture.”

# Insolent

Insolent implies both haughtiness and extreme contemptuousness; it carries a stronger implication than the preceding words of [a will to insult](#) or affront the person so treated “She could not determine whether the silent contempt of the gentlemen, or the insolent smiles of the ladies, were more intolerable.”

# Disdainful

Disdainful implies a more passionate scorn for what is beneath one than does supercilious; it as often as not suggests **justifiable pride or justifiable scorn** “A democracy smugly disdainful of new ideas would be a sick democracy.”

# Proud

Proud usually connotes a lofty or imposing manner, attitude, or appearance that **may be interpreted as dignified**, elevated, spirited, imperious, satisfied, contemptuous, or inordinately conceited according to the circumstances “She’s a stuck-up proud girl, and she hasn’t a proper decency.”

# Haughty

Haughty implies a **strong consciousness** of exalted birth, station, or character, and a more or less obvious **scorn** of those who are regarded as beneath one “His walk, his haughty, indifferent manner spoke his scorn for the two . . . men who accompanied him.”



# Lordly

Lordly usually suggests pomposity, strutting, or an arrogant display of power or magnificence “a lordly indifference to making money by his writings”

# Overbearing

Overbearing suggests a bullying or **tyrannical disposition**, or intolerable insolence “an overbearing employer”

# Supercilious

Supercilious stresses such [superficial aspects of haughtiness](#) as a lofty patronizing manner intended to repel advances. It refers to one's behavior to others rather than to one's conceit of oneself, though the latter is always implied; often it suggests not only scorn but also incivility "They have no blood, these people. Their voices, their supercilious eyes that look you up and down."

# Connotation: Summary

- “arrogant” ‘arrogant’:  
“claim too much importance”
- “insolent” ‘unverschämt’:  
“will to insult”
- “disdainful” ‘verächtlich’:  
“suggests justifiable pride or justifiable scorn”
- A connotation is an idea or feeling that a word invokes in addition to its literal or primary meaning.
- Many connotations are difficult to capture in terms of core meaning / core semantics.

# Konnotationen: Beitrag zum Sprachwandel

- z.B. *Esel*: störrisch, dumm
- vgl. Semantik: 'dem Pferd verwandtes, aber kleineres Säugetier mit grauem bis braunem Fell, kurzer Mähne, langen Ohren und Quastenschwanz' [Duden] → Bewertung der Wahrscheinlichkeit der Situation durch den Sprecher stellt ein notwendiges Element der lexikalischen Bedeutung mancher Wörter dar und muss deshalb in die Bedeutungserklärung aufgenommen werden

“someone X lied to someone Y”

someone X said something to someone else Y this someone knew that it was not true this someone said it because he/she wanted this other someone to think that it was true people think that it is bad if someone does something like this

# Outline

- 1 Lexikalische Semantik
- 2 Semantische Relationen**
- 3 Logik
- 4 Anwendungen

# Homonymie

zwei oder mehr Ausdrücke mit allen folgenden Eigenschaften:

- mit demselben Signifikanten (Ausdrucksseite), aber
- mit verschiedenen Bedeutungen
- kein gemeinsames nichttriviales semantisches Merkmal
- z.B. *Bank 1* ('Sitzgelegenheit') vs. *Bank 2* ('Geldinstitut')  
(gemeinsames triviales Merkmal: 'nicht-belebter Gegenstand')
- → 2 Lemmata



# Polysemie

ein Wort ist polysem, wenn

- es verschiedene Lexeme (d.h. mit unterschiedlichen Bedeutungen) hat,
- die ein gemeinsames nichttriviales semantisches Merkmal aufweisen
- z.B. Beispiel: *Schule 1* ('Lehranstalt als Institution') vs. *Schule 2* ('Schulgebäude') vs. *Schule 3* ('Unterricht'; "die Schule beginnt um ...") vs. ...
- → ein Lemma / Lexikoneintrag

# Synonymie

zwei oder mehr Ausdrücke mit allen folgenden Eigenschaften:

- mit derselben Bedeutung, aber
- mit verschiedenen Gestalten (Signifikant, Ausdrucksseite)
- z.B. *Aufzug*, *Fahrstuhl*, *Lift* (→ Lexikalische Funktion Syn)

# Hyponymie und Hyperonymie

## Hypernym:

- Oberbegriff für artgleiche Hyponyme, z.B. *Blume*:  
*Sonnenblume, Mohnblume, Strohblume*
- Relation ist transitiv (wenn A ein Hypernym von B ist und B ein Hypernym von C, dann ist A ein Hypernym von C)

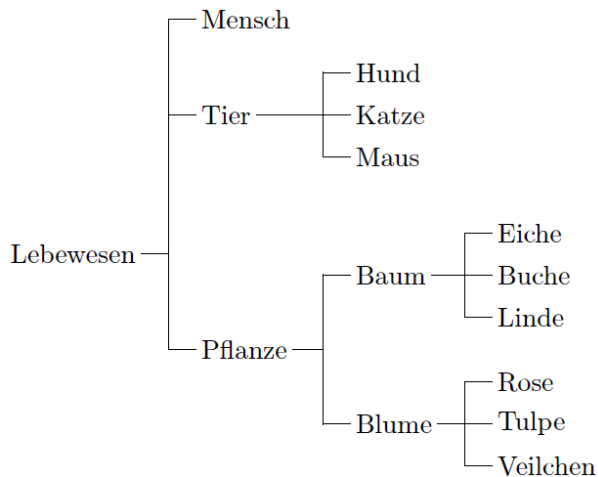
## Hyponym:

- Unterbegriff (Arten, die zu einer Gattung gehören)
- A is hyponym of B  $\Leftrightarrow$  B is hypernym of A
- A troponym can be viewed as a special kind of hyponym: it specifies manner.
- Examples: “nibble” and “gorge” are troponyms of “eat”

# Cohyponymy

- All hyponyms of a particular word are cohyponyms: apple, banana, pear are cohyponyms of fruit.

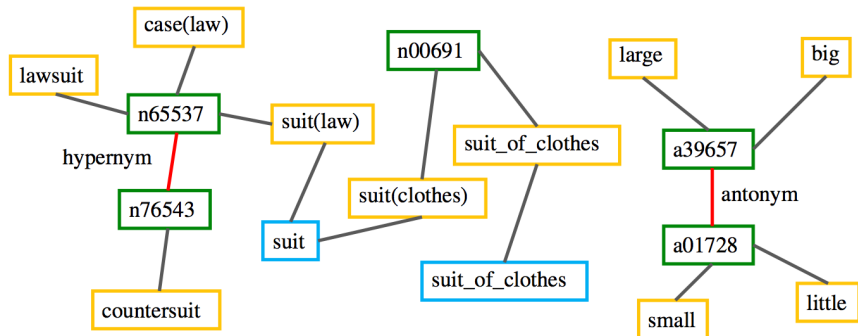
# Ein Hyponymiebaum (Ausschnitt)



# Weitere Relationen

- Antonymie → Lexikalische Funktion Anti
- Meronymie (Teil-Ganzes-Beziehung) bzw. Holonymie, z.B.  
Holonym: *Hand* – Meronym: *Finger*; *Haus* – *Dach*

# WordNet



# Exercise

- Give one example each of homonymy (“Bank”), polysemy (“Schule”), synonymy (“Fahrstuhl” – “Lift”), antonymy (“gut” – “schlecht”), hyponymy (“Hund” – “Tier”), hyperonymy (“Blume” – “Sonnenblume”), meronymy (“Finger” – “Hand”), holonymy (“Körper” – “Hand”), troponymy (“nibble” – “eat”), cohyponymy (“lily”, “rose”, “tulip”)



# Outline

- 1 Lexikalische Semantik
- 2 Semantische Relationen
- 3 Logik**
- 4 Anwendungen

Logic: Why?

# Einführendes zur Aussagenlogik

- geht zurück auf Aristoteles
- logische Verknüpfung von elementaren Sätzen (einfache Prädikation) zu komplexen Sätzen
- “ärmer” als natürliche Sprache in vielfacher Hinsicht: siehe unten
- Grundeinheit ist der Aussagesatz.
- Ein Aussagesatz hat einen Wahrheitswert: entweder “wahr” oder “falsch”

# Die Syntax der Aussagenlogik

- Satzbuchstaben: Ausdrücke, die ganzen Sätzen entsprechen, z.B. A, B, C
- 5 Satzoperatoren / Junktoren:
  - Negation (nicht)  $\neg$
  - Konjunktion (und)  $\wedge$
  - Disjunktion (nicht ausschließendes oder)  $\vee$
  - Subjunktion / Implikation (wenn, dann)  $\rightarrow$
  - Bisubjunktion / Äquivalenz (genau dann, wenn)  $\leftrightarrow$
- Hilfszeichen: ( )

# Übersetzung in Aussagenlogik: Beispiel

- Satz: “Obwohl die Aussagenlogik ein sehr grobes Modell der natürlichen Sprachen ist, kann man mit ihr einiges anfangen, falls man es geschickt anstellt”
- A: ‘die Aussagenlogik ist ein sehr grobes Modell der natürlichen Sprachen’
- B: ‘man kann mit ihr einiges anfangen’
- C: ‘man stellt es geschickt an’
- Übersetzung in Aussagenlogik:  $A \wedge (C \rightarrow B)$

# Grenzen der Aussagenlogik

- Gegeben: zwei Aussagesätze:
  - A: 'Hans isst'
  - B: 'Fritz schläft'
- Übersetzung von "Hans isst, Fritz schläft" in Aussagenlogik:
  - $A \wedge B$
- Übersetzung von "Obwohl Hans isst, schläft Fritz" in Aussagenlogik:
  - $A \wedge B$
- Dieses Beispiel zeigt: Einschätzung des Sprechers fehlt, "Pragmatik" (nächste Vorlesung) von Äußerungen geht verloren.

# Grenzen der Aussagenlogik

- “die Aussagenlogik ist für uns nützlich, aber sie ist bei weitem noch nicht die ganze Logik”
- A: ‘Die Aussagenlogik ist für uns nützlich’
- B: ‘sie ist die ganze Logik’
- Übersetzung:  $A \wedge \neg B$
- “aber”, “bei weitem”, “noch” können nicht übersetzt werden  
→ “Pragmatik” (nächste Vorlesung)

# Grenzen der Aussagenlogik

- “nicht Hans, sondern Fritz hat das getan”
- A: 'Hans hat das getan'
- B: 'Fritz hat das getan'
- Übersetzung:  $\neg A \wedge B$
- Teil der Bedeutung des Satzes (im weitesten Sinne) ist, dass angenommen wird, dass Hans es getan hat oder dass Hans der Type ist, der das machen würde. Diese Annahme bleibt unübersetzt.



# Grenzen der Aussagenlogik

- “John attends an event (if and) only if Paul attends it”
- A: “John attends the event”
- B: “Paul attends the event”
- $A \leftrightarrow B$

A	B	A only if B
t	t	t
t	f	f
f	t	f
f	f	t

# Einführendes zur Prädikatenlogik

- baut auf Aussagenlogik auf
- innere Struktur von Sätzen erkennbar
- enthält Ausdrücke, die Namen und Prädikaten der natürlichen Sprache entsprechen
- Quantoren

# Syntax der Prädikatenlogik

- Namen: 'Paris', 'Bodensee', 'Quentin Tarantino', ...
- Prädikate: '...läuft', '...ist groß', '...ist ein Bruder von ...', '...befindet sich zwischen ... und ...' (Eigenschaften von Referenten der Namen oder bestimmte Beziehungen zwischen den Referenten der Namen)
- Satzoperatoren (wie Aussagenlogik)
- quantifizierende Ausdrücke:
  - Alloperator ('alle')  $\forall$
  - Existenzoperator (es gibt ein Ding  $x$ , das ...)  $\exists$
- Namen und Prädikate: deskriptive Ausdrücke
- Satzoperatoren und quantifizierende Ausdrücke: logische Ausdrücke

# Übersetzung in Prädikatenlogik: Beispiele

- *Fritz schläft.*  $\rightarrow$  schläft (Fritz)  $\rightarrow S(f)$
- *Hans ist krank.*  $\rightarrow$  ist krank (Hans)  $\rightarrow K(h)$
- *Erna liebt Max.*  $\rightarrow$  liebt (Erna, Max)  $\rightarrow L(e, m)$
- *Alles ist rot.* (= Für jedes Ding gilt: es ist rot)  $\rightarrow \forall xR(x)$

## Exercise: Translate into predicate logic

- Every person is mortal.
- Horses do not fly.
- Every time one person is late.
- Most children must go to school.

# Outline

- 1 Lexikalische Semantik
- 2 Semantische Relationen
- 3 Logik
- 4 Anwendungen**

# Anwendungen von Semantik und Pragmatik

- Ziel der Computerlinguistik: Sprachverstehen
- Sprachverstehen erfordert komplexe Inferenzen.

# Computerlinguistische Probleme, die Inferenz erfordern

- Coreference resolution
- Beispiel: "Steve Jobs was a visionary. After he died, President Obama issued a statement from the White House."
- Worauf bezieht sich "he"?



## Computerlinguistische Probleme, die Inferenz erfordern (2)

- Question Answering
- Beispiel:
  - “die Einladung des Staatschefs wurde heftig diskutiert”
  - “die Einladung des Staatschefs wurde nach langer Diskussion nicht ausgesprochen”
  - “gegen die Einladung des Staatschefs wurde demonstriert”
  - “die Einladung des Staatschefs wurde zurückgezogen”
- Wurde der Staatschef eingeladen?

## Computerlinguistische Probleme, die Inferenz erfordern (3)

- Disambiguierung: Auflösen von Mehrdeutigkeiten auf der Ebene von Wörtern oder Sätzen
- Beispiel: “time flies like an arrow”
- “Zeitfliegen mögen einen Pfeil” vs. “Zeit fliegt wie ein Pfeil”
- Beispiel: “er kennt sich mit Kiefern aus”
- Baum vs. Teil des Schädels

## Computerlinguistische Probleme, die Inferenz erfordern (4)

- zeitliche Einordnung von Ereignissen
- “he fell asleep and had a nightmare”
- “they had dinner at an Italian restaurant and enjoyed both the food and the wine”

# Key problem: Knowledge acquisition

- Solution 1:  
Manual encoding of knowledge by experts  
(Knowledge engineering)  
→ Cyc
- Solution 2:  
Automatic acquisition of knowledge  
→ KnowItAll (UW), NELL (CMU)
- Solution 3:  
Hybrid approaches  
→ IBM Watson

# Automatic acquisition of knowledge from text

Systems: KnowItAll (UW), NELL (CMU), many others; Examples:  
<http://rtw.ml.cmu.edu/rtw>

<b>Fact</b>	<b>Confid</b>
federica_fontana is a director	91.5
illustrations_of_swollen_lymph_nodes is a lymph node	90.3
lake_triangle is a lake	100.0
louis_pasteur_and_robert_koch is a scientist	99.6
Illinois_governor_george_ryan is a politician	99.8
stephen is a person who moved to the state california	100.0
louis_armstrong is a musician who plays the trumpet	99.6
cbs_early_show is a company in the economic sector of news	93.8
knxv is a TV station in the city phoenix	100.0
broncos is a sports team that plays against new_york_jets	100.0

Cyc: Manual encoding of knowledge by experts  
= Knowledge engineering

- Cyc is a company based in Texas.
- Work going on for several decades
- OpenCyc: 239 000 concepts, 2 000 000 facts

Cyc

Watson



# Summary: Approaches to semantics

- Semiotics: Semiotic triangle
- Melchuk: Lexical functions
- Montague semantics: Compositionality
- Traditional lexicography
- Axiomatic approach: Wierzbicka's semes
- Feature semantics (Merkmalsemantik)
- Semantic relations: Synonymy, hyponymy, meronymy, . . .
- Logic: Propositional calculus, predicate logic
- Knowledge acquisition is the key bottleneck
  - Work on automatic acquisition, manual encoding and hybrids

## Exercise: Why we need more than just logic

- Reason 1: “pragmatics” (see examples above)
- Reason 2: Many inferences are not strictly logical “deductive” inferences. Rather, they are attempts to come up with the best explanation.
- **Abduction**
- “Four Venezuelan firefighters who were traveling to a training course in Texas were killed when their sport utility vehicle drifted onto the shoulder of a highway and struck a parked truck”
- Describe in detail how the firefighters died.
- Explain why these details are obvious to anybody reading this sentence, but why they cannot be logically deduced from the sentence.

# Take-away

- Lexikalische Semantik:  
die Lehre von der Bedeutung der Wörter
- Lexika: ihre vielleicht wichtigste Funktion ist die Erklärung der Bedeutung von Wörtern
- Lexikographie beschäftigt sich mit Theorie und Praxis der Lexika
- Zwei verschiedene Ansätze zur Lexikographie, die aus der Linguistik/Philosophie kommen: (i) Semantische Metasprache, (ii) Merkmalsemantik
- Semantische Relationen zwischen Lexemen, z.B. Synonymie
- Logik: Aussagenlogik, Prädikatenlogik
- Anwendungen