

Proseminar Lexikalische Semantik

Manfred Pinkal
UdS



Merkmalsemantik



- In Analogie zur Phonologie: Beschreibung der Wortbedeutungen durch eine kleine Menge von „distinktiven Merkmalen“ (z.B. Jakobson 1936)

Mann Frau Kind
Hahn Huhn Küken
Hengst Stute Fohlen

Mann: [+animate, +human, +male, +adult]

Kind: [+animate, +human, -adult]

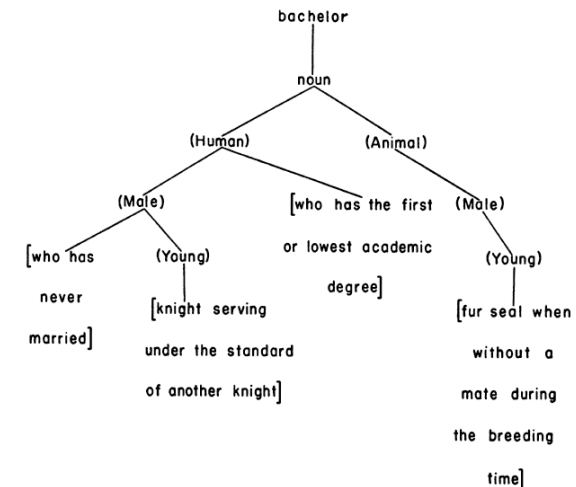
- Problem: Der Bereich der Bedeutungen ist vielfältiger und differenzierter als der Bereich der Sprachlaute.

Semantische Repräsentation



- Wie kann Wortbedeutung systematisch repräsentiert werden?
- Wortbedeutung ist vielschichtig und unterschiedliche Schichten der Bedeutung (visuelle Prototypen, Handlungsmuster, ...) erfordern völlig unterschiedliche Repräsentationen.
- Eine wichtige Schicht oder Art von semantischer Information ist „propositionale“ Information - Information, wie sie typischerweise durch sprachliche Definitionen oder Paraphrasen ausgedrückt wird.
- Wie kann propositionale Bedeutungsinformation systematisch beschrieben werden?

Katz & Fodor 1963





- Kleines Inventar von „semantischen Markern“.
- Zusätzlich eine beliebig große Menge von „Distinguishern“, die beliebig spezielle individuelle Bedeutungsunterschiede beschreiben.
- Erster Vorschlag für eine **kompositionelle Semantik** natürlicher Sprache: Berechnung der Satzbedeutung aus den Wortbedeutungen (und der syntaktischen Struktur).
- Zusätzliche Information: **Selektionsrestriktionen**



1. *The + man + hits + the + colorful + ball* → Sentence → [Some contextually definite] → (Physical Object) → (Human) → (Adult) → (Male) → (Action) → (Instancy) → (Intensity) → [Collides with an impact] → [Some contextually definite] → (Physical Object) → (Color) → [[Abounding in contrast or variety of bright colors] [Having globular shape]]



- P_4
1. *Colorful* → Adjective → (Color) → [Abounding in contrast or variety of bright colors] ⟨(Physical Object) v (Social Activity)⟩
 2. *Colorful* → Adjective → (Evaluative) → [Having distinctive character, vividness, or picturesqueness] ⟨(Aesthetic Object) v (Social Activity)⟩
- P_5
1. *Ball* → Noun concrete → (Social activity) → (Large) → (Assembly) → [For the purpose of social dancing]
 2. *Ball* → Noun concrete → (Physical Object) → [Having globular shape]
1. *hits* → Verb → Verb transitive → (Action) → (Instancy) → (Intensity) → [Collides with an impact] ⟨Subject: (Higher Animal) v (Improper Part) v (Physical Object), Object: (Physical Object)⟩²⁴



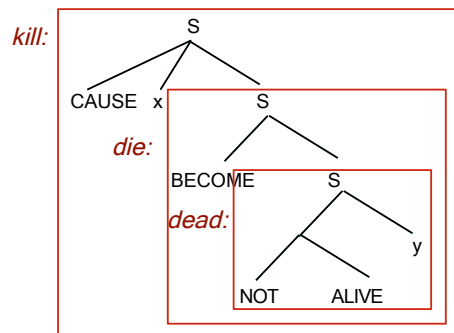
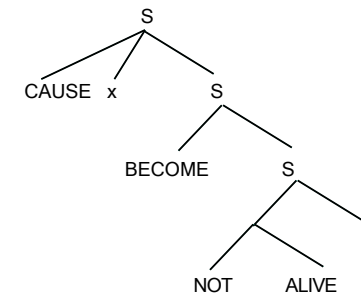
- Problem: Was bedeuten die Repräsentationen?
- Übersetzung in Prädikatenlogik:
 - $\text{physical_object}(x) \wedge \text{having_globular_shape}(x)$
 - $\forall x(\text{ball1}(x) \leftrightarrow \text{physical_object}(x) \wedge \text{having_globular_shape}(x))$
- Logische Übersetzung von *hit*?
 - $\forall x(\text{hit1}(x) \leftrightarrow \text{action}(x) \wedge \text{instancy}(x) \wedge \text{intensity}(x) \wedge \text{collides with an impact}(x))$
- Problem: Zu wenig Struktur für die Repräsentation von Wortbedeutung (und entsprechend für die Satzbedeutung).



- Ausgangsbeobachtung: Regelmäßigkeiten semantische Muster im Wortschatz zwischen Handlungsverbe, Ereignisverben und Zustandsprädikaten:

- *The door is open.*
- *The door opened.*
- *John opened the door.*

- *Bill is dead.*
- *Bill died.*
- *Somebody killed Bill.*



- Strukturierte Darstellung von Wortbedeutungsinformation
- Erklärt die systematische Beziehung zwischen lexikalischen Ausdrücken mithilfe „semantischer Primitive“ wie CAUSE, BECOME etc.
- Variablen geben Information über die Stelle, an der im Kompositionsprozess die semantische Information aus den syntaktischen Komplementen angedockt werden soll.
- Notationsvariante:
 - CAUSE(x, BECOME(NOT(ALIVE)(y)))
- Prädikatenlogische Formulierung:
 - $\forall x \forall y (\text{kill}(x,y) \leftrightarrow \text{CAUSE}(x, \text{BECOME}(\neg \text{ALIVE}(y))))$
 - $\forall y (\text{die}(y) \leftrightarrow \text{BECOME}(\neg \text{ALIVE}(y)))$
 - $\forall y (\text{dead}(y) \leftrightarrow \neg \text{ALIVE}(y))$

Generative Semantik



- Prädikatenlogische Definitionen sind „**Bedeutungspostulate**“ (bzw. eine Klasse von Bedeutungspostulaten).
- Definitionen als Axiome für die Prädikatenlogik erlauben uns systematische Inferenzen zwischen Sätzen mit unterschiedlichem Vokabular.
Somebody killed Bill \Rightarrow *Bill died* \Rightarrow *Bill is dead*.
- Der λ -Kalkül (\rightarrow Vorlesung Einführung in die Semantik) erlaubt es, strukturierte Wortbedeutungsinformation direkt im Lexikoneintrag darzustellen:
 - *kill*: $\lambda x \lambda y \text{CAUSE}(x, \text{BECOME}(\neg \text{ALIVE}(y)))$
 - *die*: $\lambda y \text{BECOME}(\neg \text{ALIVE}(y))$
- Komposition im λ -Kalkül mittels λ -Konversion:
 - *John killed Bill* $\Rightarrow \lambda x \lambda y \text{CAUSE}(x, \text{BECOME}(\neg \text{ALIVE}(y)))(\text{bill})(\text{john})$
 $\Leftrightarrow \text{CAUSE}(\text{john}, \text{BECOME}(\neg \text{ALIVE}(\text{bill})))$

Generative Semantik



- Merkmalssemantik und generative Semantik sind „**dekompositionelle**“ Ansätze der lexikalischen Semantik: Wortbedeutungen werden mittels einer kleinen Menge von Merkmalen/ semantischen Primitiven repräsentiert.
- Fragen an die dekompositionelle Semantik:
- Was ist das Inventar an semantischen Primitiven?
 - Die Frage versucht die „Kognitive Semantik“ zu beantworten (z.B. J. Lakoff, R. Jackendoff, L. Talmy)
 - Vorschläge mit praktischer Motivation (Wissensrepräsentation und Verarbeitung) wurden ebenfalls in der KI der 70er Jahre gemacht (z.B. durch R. Schank).
 - Wie verfahren wir mit der Masse an Bedeutungsinformation, die sich nicht durch Primitive beschreiben lässt?
 - Eine Antwort: Beschreibung von Bedeutungsinformation durch semantische Relationen im Wortschatz statt durch die vollständige Definition.

Semantische Relationen



- Als Beispiel die wichtigste Relation: Hyponymie (Ober-/Unterbegriffsrelation):
 - *Dackel* - *Hund*
 - *Hund* - *Haustier*
 - *Tisch* - *Möbel*
 - *Auto* - *Fahrzeug*
- Wenn wir Relationen dieser Art im semantischen Lexikon festhalten, haben wir eine für Inferenz und Wissenszugriff wichtige Ressource, ohne dass wir die beteiligten Ausdrücke komplett definieren.
- Semantische Relationen lassen sich als logische Formeln (Axiome, Bedeutungspostulate) darstellen:
 - $\forall x(\text{hund}(x) \rightarrow \text{haustier}(x))$
 - $\forall x(\text{auto}(x) \rightarrow \text{fahrzeug}(x))$
- Zu semantischen Relationen: Lyons 1977, Cruse 1986, WordNet (Referat!)

Semantische Ähnlichkeit



- „Semantische Netze“ durch Assoziationsexperimente in der Psychologie
- Korpus- und kookkurenzbasierte Ähnlichkeitsmaße für Wortbedeutung