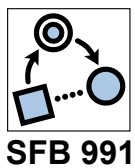
Computational Lexical Semantics

Lexical Semantics – Pustejovsky (2016)

Timm Lichte & Younes Samih

HHU Düsseldorf

2. Sitzung, 19. April 2018



Outline

Worum geht's?

Aus der Kursbeschreibung:

> *Das Lexikon enthält diejenigen semantischen Informationen, die für die Bedeutungskomposition notwendig sind.*

1. Welche Einheiten enthält das Lexikon?
   - ⇒ Morpheme, **Worte** (Wortformen, Lexeme), Phrasen, …
2. Was ist "die Bedeutung"?
   - ⇒ Problem: Mehrdeutigkeit, Abstraktheit
3. Wie repräsentieren wir (lexikalische) Bedeutung?
   - ⇒ Paraphrasen, logische Formeln, Merkmalsstrukturen, Typenhierarchien, Vektoren, …

In diesem Seminar

eine **Auswahl** (2 SWS!) computerlinguistischer "Antworten"

Stationen des Seminars

1. Theorie: Lexikon & lexikalische Semantik
2. Lexikalische Ressourcen: WordNet, FrameNet, VerbNet
3. Semantisch annotierte Korpora: SemCor, PropBank, OntoNotes
4. Anwendungen:
   - Word Sense Disambiguation
   - Semantic Role Labeling
   - …

Letzte Woche: Cruse (2001)

1. Welche Einheiten enthält das Lexikon?

- lexicon = "basic units" = "words" ("no succinct definition")
- words = word forms | lexemes | lexical units
- lexeme = "a set of **related** meanings associated with a set of **related** word forms", "units listed in a lexicon"

⇒ **Wann sind die Bedeutungen so unterschiedlich, dass man separate Lexeme annehmen muss?**

- Unterspezifikation versus Ambiguität (***teacher*** versus ***position***)
  - ⇒ unterschiedliche Bedeutungen?
- Ambiguität: Polysemie versus Homonymie (***position*** versus ***bank***)
  - ⇒ unterschiedliche Lexeme?

⇒ **Wann wird eine Bedeutung ins Lexikon aufgenommen?**

- wörtliche versus übertragene Bedeutung (Metonymie, Metapher)
- etablierte versus Spontanbedeutungen

(Beispiel ***bank*** in Merriam-Webster)

Letzte Woche: Cruse (2001)

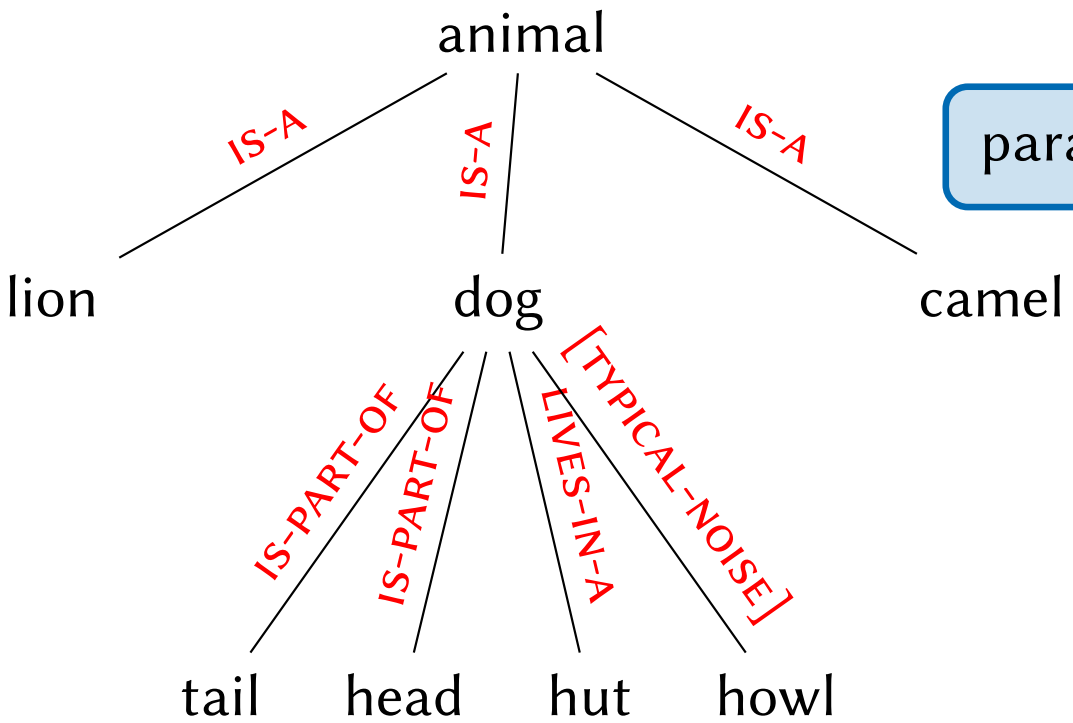
2. Was ist “die Bedeutung”?

1. Contextual/holistic approach (sense relations)
2. Componential/localist approach (semantic decomposition)
3. Conceptual approach

Letzte Woche: Cruse (2001)

2. Was ist "die Bedeutung"?

1. Contextual/holistic approach (sense relations)



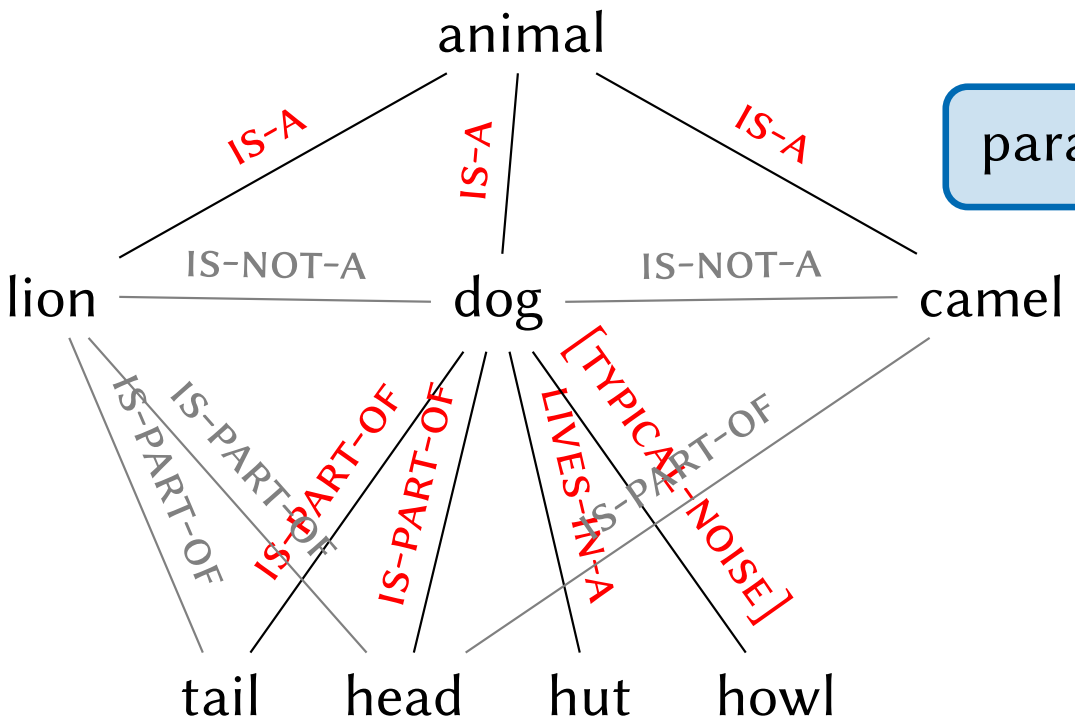
paradigmatic relations

2. Componential/localist approach (semantic decomposition)
3. Conceptual approach

Letzte Woche: Cruse (2001)

2. Was ist "die Bedeutung"?

1. Contextual/holistic approach (sense relations)



paradigmatic relations

2. Componential/localist approach (semantic decomposition)
3. Conceptual approach

Letzte Woche: Cruse (2001)

2. Was ist “die Bedeutung”?

1. Contextual/holistic approach (sense relations)

   (1) The Prime Minster attended the White House reception accompanied by his Dad/father.
   (2) John drank the wine / filing cabinet.
   (3) a male/female aunt

   syntagmatic relations

2. Componential/localist approach (semantic decomposition)
3. Conceptual approach

Letzte Woche: Cruse (2001)

2. Was ist "die Bedeutung"?

1. Contextual/holistic approach (sense relations)
2. Componential/localist approach (semantic decomposition)

filly	=	[HORSE] [FEMALE] [YOUNG]
boy	=	[HUMAN] [MALE] [YOUNG]
kill	=	[CAUSE] [BECOME] [NOT] [ALIVE]
chair	=	[OBJECT] [FURNITURE] [FOR SITTING] [FOR ONE PERSON] [WITH BACK]

3. Conceptual approach

Letzte Woche: Cruse (2001)

2. Was ist “die Bedeutung”?

1. Contextual/holistic approach (sense relations)
2. Componential/localist approach (semantic decomposition)
3. Conceptual approach

???

Pustejovsky (2016)

Pustejovsky, James. 2016. Lexical Semantics. In Maria Aloni & Paul Dekker (eds.), The Cambridge Handbook of Formal Semantics (Cambridge Handbooks in Language and Linguistics), 33–64. Cambridge, UK: Cambridge University Press.

wieder ein Handbuch-Artikel ⇒ Terminologie, Grundbegriffe, Beispiele

Leseaufgabe:
Geben Sie für jede der “strategies to lexical specification” (S.34) je ein Beispiel aus dem Text.

Pustejovsky (2016): Outline

- Introduction
- The history of lexical semantics
- Issues in lexical semantics
- Event semantics
- Lexical decomposition
- Semantic roles
- Qualia structure
- Type theory and the lexicon
- Open issues in lexical semantics

Pustejovsky (2016): Introduction

> *Lexical semantics is the study of what words mean and how their meanings contribute to the compositional interpretation of natural language utterances. The lexicon can be seen as that component of the grammar that encodes both* ***the information required for composition*** *in the syntax and the knowledge for multiple levels and types of semantic interpretation.*

Key questions:

1. What are the semantic components that constitute word meaning?

Pustejovsky (2016): Introduction

> *Lexical semantics is the study of what words mean and how their meanings contribute to the compositional interpretation of natural language utterances. The lexicon can be seen as that component of the grammar that encodes both* ***the information required for composition*** *in the syntax and the knowledge for multiple levels and types of semantic interpretation.*

Key questions:

1. What are the semantic components that constitute word meaning?
2. How are word meanings differentiated and how are they related to each other?

Pustejovsky (2016): Introduction

> *Lexical semantics is the study of what words mean and how their meanings contribute to the compositional interpretation of natural language utterances. The lexicon can be seen as that component of the grammar that encodes both* ***the information required for composition*** *in the syntax and the knowledge for multiple levels and types of semantic interpretation.*

Key questions:

1. What are the semantic components that constitute word meaning?
2. How are word meanings differentiated and how are they related to each other?
3. How does the meaning of individual words drive the compositional process to make semantically coherent sentences?

Pustejovsky (2016): Introduction

> *Lexical semantics is the study of what words mean and how their meanings contribute to the compositional interpretation of natural language utterances. The lexicon can be seen as that component of the grammar that encodes both* ***the information required for composition*** *in the syntax and the knowledge for multiple levels and types of semantic interpretation.*

Key questions:

1. What are the semantic components that constitute word meaning?
2. How are word meanings differentiated and how are they related to each other?
3. How does the meaning of individual words drive the compositional process to make semantically coherent sentences?
4. When is a component of word meaning considered "lexical" rather than "world" knowledge?

Pustejovsky (2016): Introduction

Theoretical assumptions:

- Lexical meaning involves a kind of **componential analysis**, either through predicative primitives or a system of types.
- The **selectional properties** of words must be explained in terms of the lexical semantics of predication.
- An understanding of the semantics of nouns and the contribution that verbal arguments play in composition is crucial for an adequate model of how lexical semantics contributes to compositionality.

Pustejovsky (2016): Introduction

Four strategies to lexical specification of the “interface”:

1. **Atomic predication:** No interface language is provided, and there is direct interpretation of an object expression into the model.

Pustejovsky (2016): Introduction

Four strategies to lexical specification of the "interface":

1. **Atomic predication:** No interface language is provided, and there is direct interpretation of an object expression into the model.
2. **Relational models:** The interface is a relational structure, which is interpreted in the model.

Pustejovsky (2016): Introduction

Four strategies to lexical specification of the "interface":

1. **Atomic predication:** No interface language is provided, and there is direct interpretation of an object expression into the model.
2. **Relational models:** The interface is a relational structure, which is interpreted in the model.
3. **Feature-based decomposition:** Component-based features are used to classify an expression in the object language into distinct concepts in the model.

Pustejovsky (2016): Introduction

Four strategies to lexical specification of the "interface":

1. **Atomic predication:** No interface language is provided, and there is direct interpretation of an object expression into the model.
2. **Relational models:** The interface is a relational structure, which is interpreted in the model.
3. **Feature-based decomposition:** Component-based features are used to classify an expression in the object language into distinct concepts in the model.
4. **Structural decomposition:** Component-based features are organized as a graph structure, with associated compositional interpretations in the model.

Pustejovsky (2016): Introduction

Atomic predication

- non-decompositional approaches to lexical semantic interpretation
- as many lexical entries as there are senses for it in the object language

$\Rightarrow$ type-theoretic semantics (Montague 1974)

$\Rightarrow$ language of thought (Fodor 1975)

Pustejovsky (2016): Introduction

Relational models

- words are treated as primitives
- arbitrarily complex relational structures that facilitate semantic inferences

$\Rightarrow$ Davidson's addition of the event variable to action predicates (Davidson 1967)

Pustejovsky (2016): Introduction

Feature-based decomposition

All expressions in the object language are decomposed into sets of binary-valued features, distinguishing concepts such as gender, number, age, marital status, and so on.

$\Rightarrow$ distributional semantic models

Pustejovsky (2016): Introduction

Structural decomposition

Words are defined as restricted algebraic structures, with **primitive predicates as atomic elements**.

- ⇒ Dowty (1979)
- ⇒ Generative Semantics
- ⇒ Jackendoff's Conceptual Structure
- ⇒ Generative Lexicon (Pustejovsky 1995)
- ⇒ HPSG

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic
  - feature-based decomposition: man as [+adult, +male], woman as [+adult, -male], girl as [-adult, -male]

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic
  - feature-based decomposition: man as [+adult, +male], woman as [+adult, -male], girl as [-adult, -male]
- Generative Linguistics: feature-based decomposition

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic
  - feature-based decomposition: man as [+adult, +male], woman as [+adult, -male], girl as [-adult, -male]
- Generative Linguistics: feature-based decomposition
- Valence Grammar: relational structure of sentence meaning

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic
  - feature-based decomposition: man as [+adult, +male], woman as [+adult, -male], girl as [-adult, -male]
- Generative Linguistics: feature-based decomposition
- Valence Grammar: relational structure of sentence meaning
- Case Grammar/Frame Semantics: relational model

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic
  - feature-based decomposition: man as [+adult, +male], woman as [+adult, -male], girl as [-adult, -male]
- Generative Linguistics: feature-based decomposition
- Valence Grammar: relational structure of sentence meaning
- Case Grammar/Frame Semantics: relational model
- Generative Semantics: mainly feature-based, but with relational add-ons

The history of lexical semantics

- Artistotle: a systematic set of dimensions called *aitia*
- late 19th: psychological perspective on polysemy (German, France)
- early 20th: European Structuralism
  - paradigmatic versus syntagmatic
  - feature-based decomposition: man as [+adult, +male], woman as [+adult, -male], girl as [-adult, -male]
- Generative Linguistics: feature-based decomposition
- Valence Grammar: relational structure of sentence meaning
- Case Grammar/Frame Semantics: relational model
- Generative Semantics: mainly feature-based, but with relational add-ons
- Montague Semantics: structural decomposition; synthesis of lexicon and syntax

Issues in lexical semantics: Ambiguity vs. polysemy [sic!]

Ambiguity is unavoidable: size of lexicon < number of objects and relations in the world

- **homonymy** ("unrelated senses")

  (1) a. Mary strolled along the ***bank*** of the river.
  b. This company is the largest ***bank*** in the city.

  ⇒ "represented as separate lexical entries"

- **polysemy** ("related senses")

  (2) a. Mary carried the ***book*** home.
  b. Mary doesn't agree with the ***book***.

  ⇒ [represented in one lexical entry?]

"It is the role of the lexicon to distinguish such ambiguities, and to establish what this logical relation is."

"simply listing the senses of these words [nouns and adjectives] will not always account for their **creative use** in novel contexts in language"

Issues in lexical semantics: Selectional constraints

Lexical information impacting the grammar:

- **argument structure**: number of associated co-phrases
- **syntactic category**: “ identifies the actual syntactic phrase associated with the argument”
- **selectional constraints**: grammatical and semantic features

(3) laugh(arg 1 [cat=NP,animacy=+])
  a. *The man laughed the ball.
  b. The man / *the rock laughed.

Issues in lexical semantics: Verb meaning and mapping to syntax

Verbal polysemy ⇒ alternation

(4) a. The glass ***broke***.
    b. Mary ***broke*** the glass.
(5) a. The ball ***rolled*** down the slide.
    b. The boy ***rolled*** the ball down the slide.
(6) a. The ship ***sank***.
    b. The torpedo ***sank*** the ship.

Not all verbs have alternations such as this (***die***, ***arrive***).
Where is it stored? In the lexicon? "Clear" cases: ***eat***, ***devour***[, ***dine***]
[How does this reflect the meaning of the verb?]

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

   (7) Mary ate the soup with a spoon in the kitchen at 3:00pm.
   $\exists e[\text{eat}(e, m, \text{the\_soup}) \wedge \text{with}(e, \text{a\_spoon}) \wedge \text{in}(e, \text{the\_kitchen}) \wedge \ldots]$

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

   (7) Mary ate the soup with a spoon in the kitchen at 3:00pm.

   $\exists e[\text{eat}(e, m, \text{the\_soup}) \wedge \text{with}(e, \text{a\_spoon}) \wedge \text{in}(e, \text{the\_kitchen}) \wedge \ldots]$

2. **Event types**: Aktionsarten

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

   (7) Mary ate the soup with a spoon in the kitchen at 3:00pm.

   $\exists e[\text{eat}(e, m, \text{the\_soup}) \wedge \text{with}(e, \text{a\_spoon}) \wedge \text{in}(e, \text{the\_kitchen}) \wedge \ldots]$

2. **Event types**: Aktionsarten
   - state

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

   (7) Mary ate the soup with a spoon in the kitchen at 3:00pm.

   $\exists e[\text{eat}(e, m, \text{the\_soup}) \wedge \text{with}(e, \text{a\_spoon}) \wedge \text{in}(e, \text{the\_kitchen}) \wedge \ldots]$

2. **Event types**: Aktionsarten
   - state
   - activity (atelic)

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

   (7) Mary ate the soup with a spoon in the kitchen at 3:00pm.

   $\exists e[\text{eat}(e, m, \text{the\_soup}) \wedge \text{with}(e, \text{a\_spoon}) \wedge \text{in}(e, \text{the\_kitchen}) \wedge \ldots]$

2. **Event types**: Aktionsarten
   - state
   - activity (atelic)
   - accomplishment (telic)

Event semantics

Two characterizations of events in lexical semantics:

1. **Event variable** in FOL: $\lambda y \lambda x \lambda e[\text{eat}(e, x, y)]$

   (7) Mary ate the soup with a spoon in the kitchen at 3:00pm.

   $\exists e[\text{eat}(e, m, \text{the\_soup}) \wedge \text{with}(e, \text{a\_spoon}) \wedge \text{in}(e, \text{the\_kitchen}) \wedge \ldots]$

2. **Event types**: Aktionsarten
   - state
   - activity (atelic)
   - accomplishment (telic)
   - achievement

Lexical decomposition

*Lexical decomposition is concerned with the **internal semantic structure** of lexical items within a lexicon. The focus of lexical decomposition is on how the lexical items are semantically similar and distinct by virtue of shared knowledge structures or **semantic primitives**.*

kill: (CAUSE(x,(BECOME(NOT(ALIVE y)))))[1]

Inventory of primitve categories?
Complete and exhaustive description of the meaning?
Status of decomposition?

[1] Following Dowty (1979).

Semantic roles

"partial decomposition" with respect to argument structure
(aka theta role assignment)

Semantic roles

"partial decomposition" with respect to argument structure
(aka theta role assignment)

(8) a. put⟨agent,theme,location⟩
    b. borrow⟨recipient,theme,source⟩

⇒ Case Grammar, Frame Theory

Qualia structure

Qualia structure in Generative Lexicon Theory

1. **formal role**: the basic category that distinguishes the object within a larger domain

Example: ***rock*** versus ***chair***

Qualia structure

Qualia structure in Generative Lexicon Theory

1. **formal role**: the basic category that distinguishes the object within a larger domain
2. **constitutive role**: the relation between an object and its constituent parts

Example: ***rock*** versus ***chair***

Qualia structure

Qualia structure in Generative Lexicon Theory

1. **formal role**: the basic category that distinguishes the object within a larger domain
2. **constitutive role**: the relation between an object and its constituent parts
3. **telic role**: purpose and function

Example: ***rock*** versus ***chair***

Qualia structure

Qualia structure in Generative Lexicon Theory

1. **formal role**: the basic category that distinguishes the object within a larger domain
2. **constitutive role**: the relation between an object and its constituent parts
3. **telic role**: purpose and function
4. **agentive role**: origin or "coming into being"

Example: ***rock*** versus ***chair***

Type theory and the lexicon

Richer type system than Montague Grammar, with type composition:

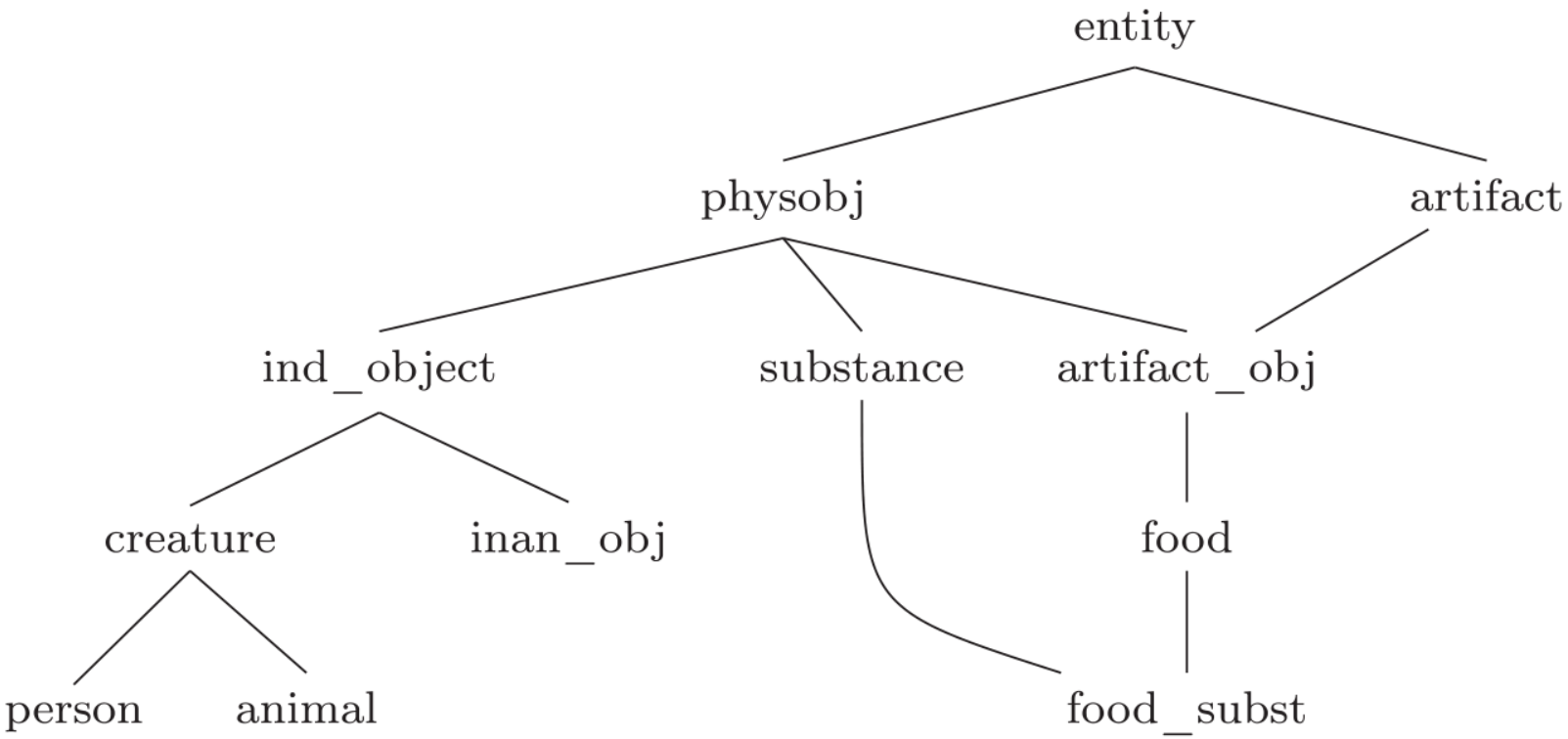


Figure 2.1 *Fragment of a type hierarchy*

Benefits: “explaining” selectional constraints; type coercion

References

[1] Davidson, Donald. 1967. The logical form of action sentences. In Nicholas Rescher (ed.), *The logic of decision and action*. University of Pittsburgh Press.

[2] Dowty, David R. 1979. *Word meaning and Montague Grammar*. Reprinted 1991 by Kluwer Academic Publishers. Dordrecht, Boston, London: D. Reidel Publishing Company.

[3] Fodor, Jerry A. 1975. *The language of thought*. Cambridge, MA: Harvard University Press.

[4] Montague, Richard. 1974. *Formal philosophy: Selected papers of Richard Montague. Ed. and with an introd. by Richmond H. Thomason*. New Haven, CT: Yale University Press.

[5] Pustejovsky, James. 1995. *The generative lexicon*. Cambridge, MA: MIT Press.