
Lexical Semantics:

WordNet and Word Senses, Ontologies, and Semantic Lexical Resources

Lexical Semantics

- Lexicons – words (or lexemes or stems) together with some information
- Dictionaries – a lexicon with definitions for each word sense
 - Most are now available online
- Thesauruses – add synonyms for each word sense
 - Roget Thesaurus
 - WordNet
- Semantic networks – add more semantic relations
 - WordNet
 - EuroWordNet
- Ontologies – add semantic relations and rules about entities, concepts and relations

Word Senses

- We say that a word has more than one word sense if there is more than one definition.

Online dictionary definitions for the noun *plant*

1. a living organism of the kind exemplified by trees, shrubs, herbs, grasses, ferns, and mosses, typically growing in a permanent site, absorbing water and inorganic substances through its roots, and synthesizing nutrients in its leaves by photosynthesis using the green pigment chlorophyll.
2. a place where an industrial or manufacturing process takes place

- Word senses may be
 - Coarse-grained, if not many distinctions are made
 - Fine-grained, if there are many distinctions of meanings

WordNet

- WordNet is a database of facts about words
 - Meanings and the relations among them
- Words are organized into clusters of synonyms
 - Synsets
- <http://wordnet.princeton.edu/>
- Organized into nouns, verbs, adjectives, and adverbs
 - Currently 170,000 synsets
 - Available for download, arranged in separate files (DBs)

Knowledge Resources - Dictionary

- For each word in the language vocabulary, a dictionary provides:
 - A list of meanings
 - Definitions (for all word meanings)
 - Typical usage examples (for most word meanings)

WordNet definitions(called glosses)/examples for synsets of the noun *plant*

1. buildings for carrying on industrial labor; "they built a large plant to manufacture automobiles"
2. a living organism lacking the power of locomotion
3. something planted secretly for discovery by another; "the police used a plant to trick the thieves"; "he claimed that the evidence against him was a plant"
4. an actor situated in the audience whose acting is rehearsed but seems spontaneous to the audience

Knowledge Resources - synonyms

- A thesaurus adds:
 - An explicit synonymy relation between word meanings

WordNet synsets for the noun “plant”

1. plant, works, industrial plant
2. plant, flora, plant life

Knowledge Resources - relations

- A semantic network adds relations for each word sense:
 - hypernymy/hyponymy (IS-A),
 - hypernyms are more general, hyponyms are more specific
 - meronymy/holonymy (PART-OF),
 - antonymy, entailment, etc.

WordNet related concepts for the meaning “plant life”

{plant, flora, plant life}

hypernym: {organism, being}

hypomym: {house plant}, {fungus}, ...

meronym: {plant tissue}, {plant part}

holonym: {Plantae, kingdom Plantae, plant kingdom}

WordNet Relations

- A more detailed list from Jurafsky and Martin

Relation	Also Called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	<i>breakfast</i> ¹ → <i>meal</i> ¹
Hyponym	Subordinate	From concepts to subtypes	<i>meal</i> ¹ → <i>lunch</i> ¹
Instance Hypernym	Instance	From instances to their concepts	<i>Austen</i> ¹ → <i>author</i> ¹
Instance Hyponym	Has-Instance	From concepts to concept instances	<i>composer</i> ¹ → <i>Bach</i> ¹
Member Meronym	Has-Member	From groups to their members	<i>faculty</i> ² → <i>professor</i> ¹
Member Holonym	Member-Of	From members to their groups	<i>copilot</i> ¹ → <i>crew</i> ¹
Part Meronym	Has-Part	From wholes to parts	<i>table</i> ² → <i>leg</i> ³
Part Holonym	Part-Of	From parts to wholes	<i>course</i> ⁷ → <i>meal</i> ¹
Substance Meronym		From substances to their subparts	<i>water</i> ¹ → <i>oxygen</i> ¹
Substance Holonym		From parts of substances to wholes	<i>gin</i> ¹ → <i>martini</i> ¹
Antonym		Semantic opposition between lemmas	<i>leader</i> ¹ ⇔ <i>follower</i> ¹
Derivationally Related Form		Lemmas w/same morphological root	<i>destruction</i> ¹ ⇔ <i>destroy</i> ¹

WordNet Hierarchies

Sense 3

bass, basso --

(an adult male singer with the lowest voice)

=> singer, vocalist, vocalizer, vocaliser

=> musician, instrumentalist, player

=> performer, performing artist

=> entertainer

=> person, individual, someone...

=> organism, being

=> living thing, animate thing,

=> whole, unit

=> object, physical object

=> physical entity

=> entity

=> causal agent, cause, causal agency

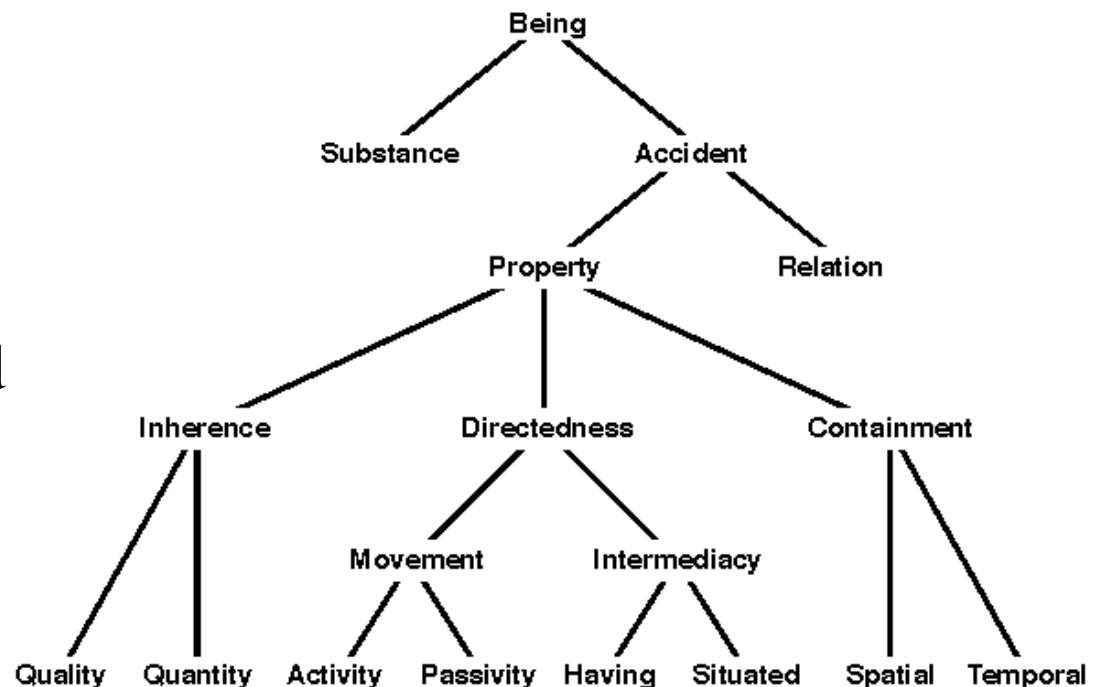
=> physical entity

=> entity

Origins of Ontology

- In philosophy, ontology studies existence/being of the world.
 - We can think of ontology as categorizing everything in the world.

- In his work “categories”, Aristotle listed ten categories that all things of the world should belong to.



Ontology in Information Science

- Ontology is an approach of knowledge organization.
- In general, ontologies are about the **representations of semantics**:
 - Concepts, e.g. *person, animal, food, table, movie, etc.*
 - Instances (or entities), e.g. Barack Obama is an *instance* of the concept “person”.
 - Properties, e.g. a person has properties of *gender, height, weight, father, mother, etc.*
 - Relations, e.g. Syracuse University is *located in* Syracuse.
 - Rules between concepts, properties, and relations, e.g. if someone is *married*, then he/she should have a *spouse*.

Ontology Example: UMLS

- The Unified Medical Language System (UMLS) aggregates various controlled vocabularies and mapped them to a comprehensive biomedical ontology. It has three knowledge sources:
 - Metathesaurus. Mapping concepts and terms in different thesaurus and organizing them in the UMLS structure
 - Semantic network. Connecting semantic types of concepts in metathesaurus by semantic relations.
 - Specialist Lexicon. Containing lexical information of biomedical terms.
- This is an example of a word and phrase level resource
- Online, but not publically available

Semantic Lexicons

- Lexicon where each word is assigned to a semantic class
- Lexical resources have been developed to assign words to semantic classes in support of applications that need to detect opinion, sentiment, or other more subjective meanings
- Three examples given here; additional examples will be given when we cover sentiment analysis

Semantic classes: Subjectivity Lexicon

- Subjectivity Lexicon from the MPQA project with Jan Wiebe
 - Gives a list of 8,000+ words that have been judged to be weakly or strongly positive, negative or neutral in **subjectivity**
 - Examples:

type=weaksubj len=1 word1=abandoned pos1=adj stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=abandonment pos1=noun stemmed1=n priorpolarity=negative
type=weaksubj len=1 word1=abandon pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abase pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abasement pos1=anypos stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=abash pos1=verb stemmed1=y priorpolarity=negative
type=weaksubj len=1 word1=abate pos1=verb stemmed1=y priorpolarity=negative
type=strongsubj len=1 word1=absolve pos1=verb stemmed1=y priorpolarity=positive
type=strongsubj len=1 word1=absolute pos1=adj stemmed1=n priorpolarity=neutral

Semantic classes: LIWC

- Linguistic Inquiry and Word Count
 - Text analysis software based on dictionaries of word dimensions
 - Dimensions can be syntactic
 - Pronouns, past-tense verbs
 - Dimensions can be semantic
 - Social words, affect, cognitive mechanisms
 - Other categories
 - See <http://liwc.wpengine.com/compare-dictionaries/>
 - James Pennebaker, Univ. of Texas at Austin
 - <http://www.liwc.net/>
- Often used for positive and negative emotion words in opinion mining

Semantic classes for words: ANEW

- Affective Norms for English Words
 - Provides a set of **emotional** ratings for a large number of words in the English language
- Participants gave graded reactions from 1-9 on three dimensions
 - Good/bad, psychological valence
 - Active/passive, arousal valence
 - Strong/weak, dominance valence
- From the NIMH Center for the Study of Emotion and Attention at the University of Florida
 - <http://csea.phhp.ufl.edu/Media.html>
 - See also the paper by Dodds and Danforth on Happiness of Large-Scale Written Expressions