Lexical Semantics
Guest Lecture
Málvinnsla

Matthew Whelpton
whelpton@hi.is

School of Humanities
University of Iceland

Outline

Word Senses in Relationship

The founder of modern linguistics, Ferdinand de Saussure, made a fundamental distinction between two senses of "language".

- **le parole** — the particular things that we happen to say on any particular occasion.
  - A corpus of conversational speech would be a good example of *parole*.
- **la langue** — the underlying knowledge of the organisation of the language which allows us to say those things.
  - when a linguist writes a grammar, they are attempting to represent the system of relationships that underlie our ability to use language, i.e. a grammar is a theory of *la langue*
  - when a computational linguist writes a parser, they are in effect attempting to simulate one aspect of the system of relationships which allows the string of words to be properly interpreted linguistically, i.e. to simulate *la langue*.

Saussure, Signs and Structuralism

In Saussure's view, the basic building block of *la langue* 'the linguistic system' is the *sign*.

- A sign is a pairing of form and meaning

  (1) bank: bæŋk∼'a financial institution'

- Saussure's radical proposal is that signs do not take their "meaning" or *value* from the outside world directly but rather from their system of relationships to other signs in the system.

Value from Structure of Relationships

Saussure's famous example is a difference between French and English.

- French has the word *mouton* 'sheep', which was borrowed into English from Norman French.
- However it would be wrong to say that they are the "same" word because they belong to systems with different sets of relationships.
- In English, *mutton* (the food) must share space with *sheep* (the animal in the field).
- This is part of a mini-system in English: beef∼cow, pork∼pig, veal∼calf, venison∼deer.
- French does not have this system: *mouton* refers to both the food and the animal in the field.

So the *value* of *mouton* and *mutton* are different because of the different systems in which they are embedded.

Linguistic Structure∼Conceptual Structure∼The World

Much of the work in lexical semantics has gone into investigating the systematic relationships between linguistic signs.

- However, Saussure's radical reduction of "word meaning" to "structural value" is no longer generally accepted, though it remains influential and insightful.
- As we will see, word senses are often linked to more general processes of conceptualisation.
- And the degree to which "the world out there" has an influence on the organisation and "content" of the vocabulary system remains a matter of hot debate!

Synonymy

The content, meaning or sense of a word is given in dictionaries by a *definition*.

- A definition gives the meaning of a word as a statement or description

Sometimes the definition of two words is almost exactly the same, e.g. these definitions from `http://dictionary.com`.

(2) a. **sofa**: A long upholstered seat typically with a back and arms.

b. **couch**: A piece of furniture for seating from two to four people, typically in the form of a bench with a back, sometimes having an armrest at one or each end, and partly or wholly upholstered and often fitted with springs, tailored cushions, skirts, etc.; sofa.

True synonyms?

Sofa and *couch* are said to be *synonyms*.

- From Greek, *syn* 'same' + *nym* 'name'.

In fact, two words rarely mean exactly the same thing in a broad sense of "meaning".

- So, in Britain, *couch* is often used by speakers at the lower end of the social scale and *sofa* by speakers at the higher end of the social scale.
- This means that *couch* is sometimes associated with less "sophisticated" furniture.
- i.e. the *connotations* are different

A working understanding of synonymy is simply that two words can have the "same meaning" in a particular context.

Antonymy

Many words, especially adjectives, may have an *antonym*.

- **antonym** — a word with the opposite meaning.
- From Greek, *ant* 'opposite' + *nym* 'name'
- e.g. tall vs short, hot vs cold, large vs small

The interesting thing is that to count as opposite the two words have in fact to be closely related in meaning.

- Usually antonyms represent polar *values* on the same *scale*.
  - HEIGHT: high=tall/low=short
  - TEMPERATURE: high=hot/low=cold
  - SIZE: high=large/low=small

Homonymy and Ambiguity

Synonymy involved two different forms with the same meaning; *homonymy* involves two identical forms with different meanings.

(3) a. **lot** (a parking lot): a distinct portion or piece of land
b. **lot** (a lot of students): a great many or a great deal

- From Greek, *homo* 'same' + *nym* 'name'

For homonymy, the two meanings should be completely distinct.

- This is a full case of *ambiguity*, i.e. one form that represents two completely different linguistic items.

Polysemy and Ambiguity

As you may have guessed, though, it is not always clear when the two meanings are completely distinct.

(4) a. *tool*, as in *The saw is a great tool.*
b. *tool*, as in *sed is a great tool.*

- In this case the meanings are clearly closely related.
- In practical terms, they are distinct.
  - A saw is a physical object used for a certain kind of physical operation.
  - sed is a computer program used to perform certain kinds of abstract operations.
- However, in both cases, some entity (concrete or abstract) is used by an agent to achieve some purpose.

Here we say that there is really just one word but it has a number of closely related senses; it is *polysemous*.

- From Greek, *poly* 'many' + *sem* 'meaning'

Homonymy: Homography and Homophony

In our example of homonymy (*lot*), both the spelling and pronunciation are the same.

- But it is possible for the spelling to be the same and the pronunciation different.

(5) a. *wɪnd* 'wind', cf. the wind in the willows
b. *waɪnd* 'wind', cf. wind the clock up

  - **homograph**: From Greek, *homo* 'same' + *graph* 'writing'

- Or for the spelling to be different and the pronunciation the same.

(6) a. *sɔː* 'sore', cf. a running sore on his arm
b. *sɔː* 'saw', cf. cut it off with a saw

  - **homophone**: From Greek, *homo* 'same' + *phone* 'sound'

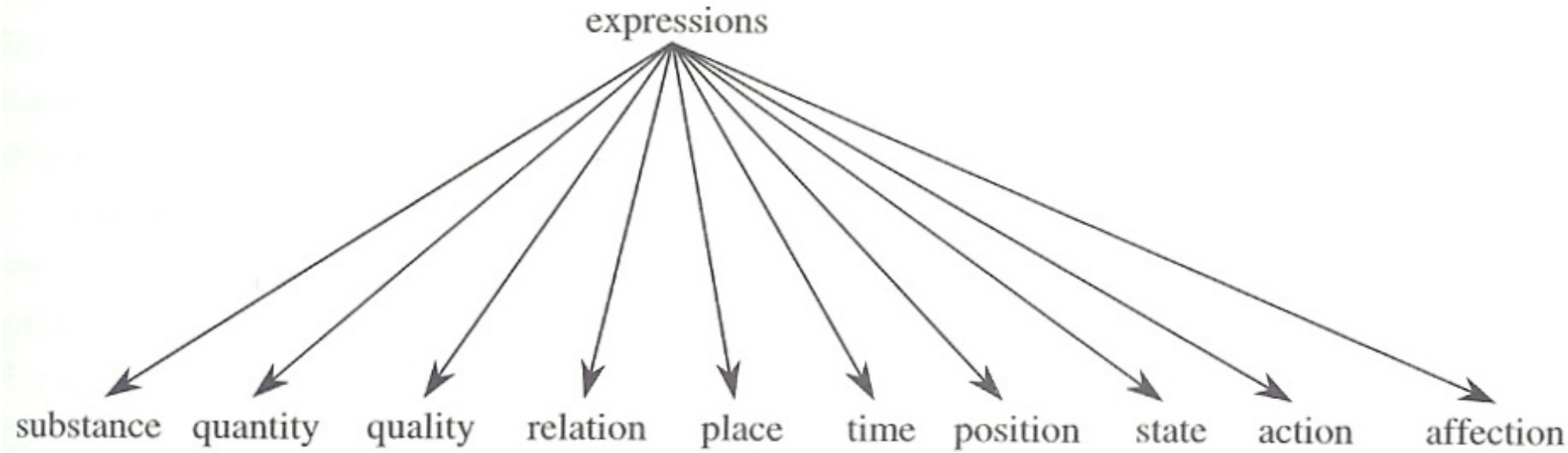
Ontologies

Computational lexical semantics deals with the organisation of the lexicon, the store of words.

- Dictionaries organise words in alphabetical lists for easy access.
- However, the focus of lexical semantics is on the meaning of words.
- How do we organise word storage in terms of sense or meaning rather than form?
- One useful way of storing words is in terms of an *ontology*.
  - An explicit formal specification of how to represent the objects, concepts and other entities that are assumed to exist in some area of interest and the relationships that hold among them. (`http://dictionary.com`)

Aristotle's Ontology

Artistotle proposed an ontology for the kinds of things that basic expressions signify.

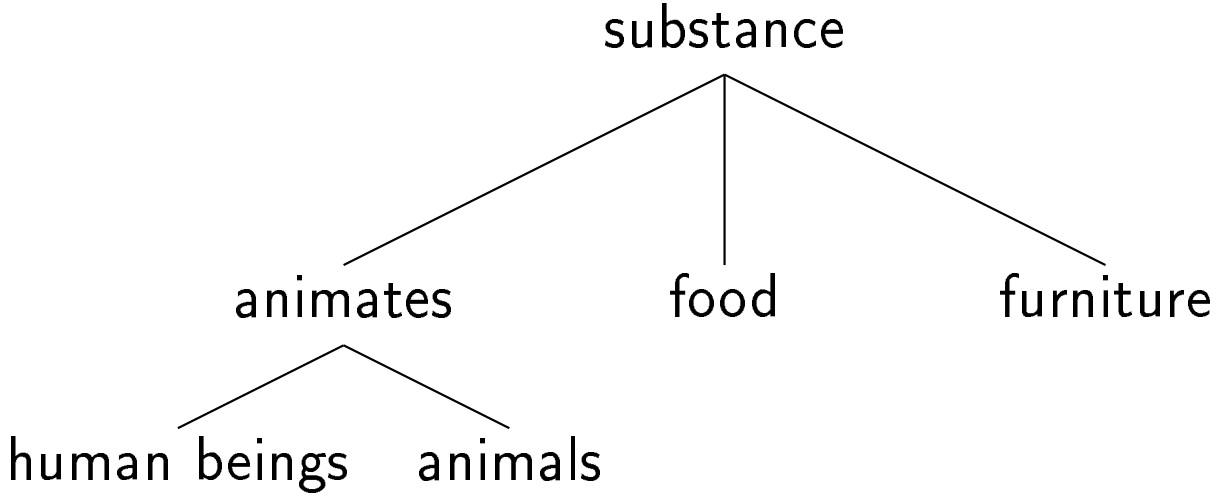


- substance (also entity, thing), e.g.
  - horse, furniture, air, mud
- quality (also attributes, properties)
  - white, tall, hot, grammatical, interesting

Extending the Ontology

This is a *flat hierarchy* but the hierarchy can be deepened to introduce subclassifications.

(7)



Hypernyms and Hyponyms

A simple hierarchy of this kind represents the relations of *hyponymy* and *hypernymy*.

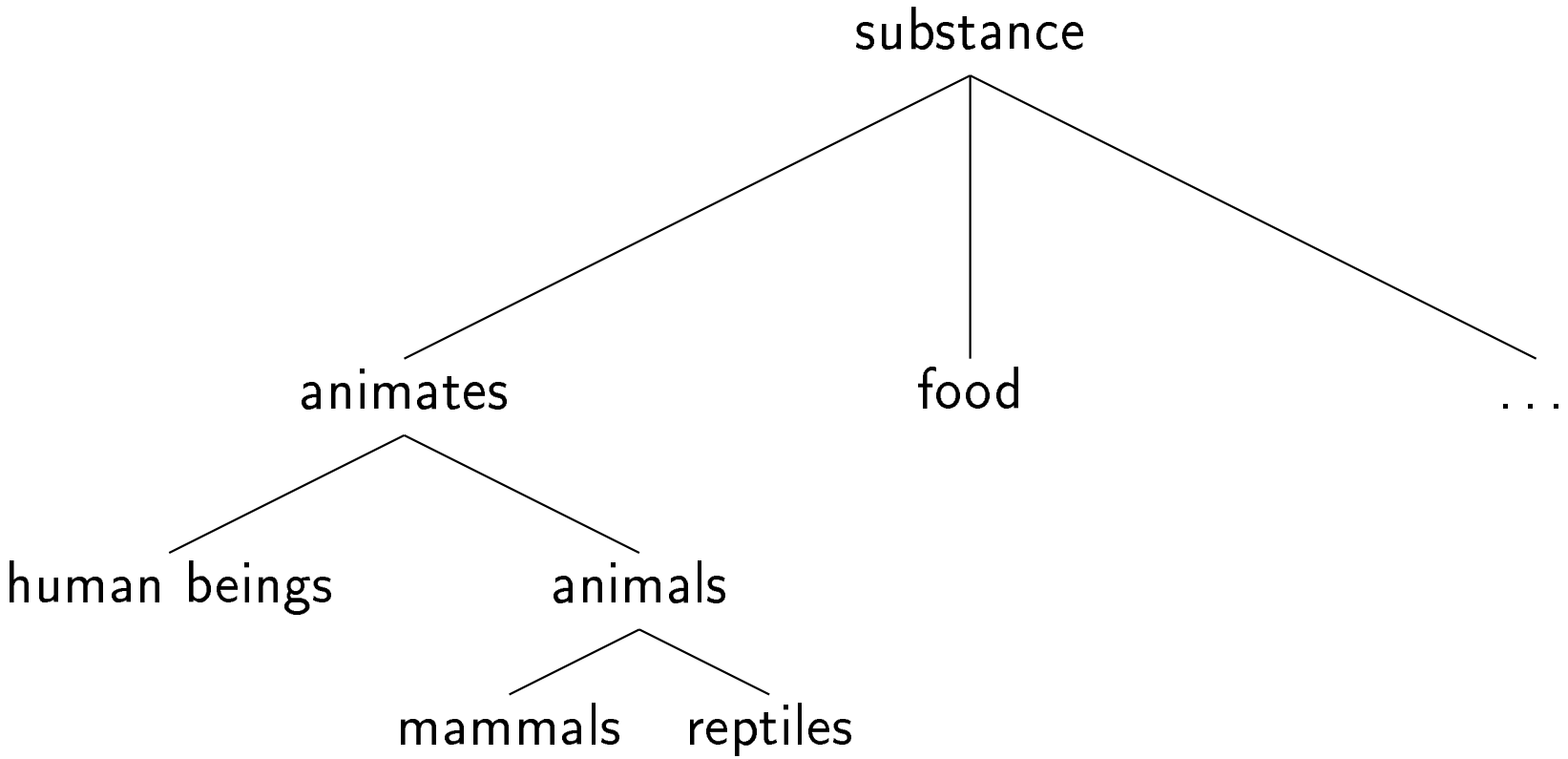
- *food* is a (kind of) *substance*, so *food* is a *hyponym* of *substance*.
  - From Greek, *hypo* 'under' + *nym* 'name'.
  - x is food $\rightarrow$ x is a substance
  - $\neg$ (x is a substance $\rightarrow$ x is food)
- Conversely, *substance* includes *food* as a sub-kind, so *substance* is a *hypernym* of *food*.
  - From Greek, *hyper* 'over' + *nym* 'name'.
- Which gives us the general patterns:
  - <hyponym> is a <hypernym>
  - x is <hyponym> $\rightarrow$ x is <hypernym>
  - $\neg$ (x is <hypernym> $\rightarrow$ x is <hyponym>)

Deep Hierarchies

This hierarchy of hyponym∼hypernym relations can be extended with new items of vocabulary as long as the relevant relation holds:

- mammal∼animal
  - a mammal is an animal
  - x is a mammal → x is an animal
  - ¬ (x is an animal → x is a mammal)
- reptile∼animal
  - a reptile is an animal
  - x is a reptile → x is an animal
  - ¬ (x is an animal → x is a reptile)
- But note, for mammal∼reptile
  - ¬(a mammal is a reptile)
  - ¬(a reptile is a mammal)
  - ¬ (x is a mammal → x is a reptile)
  - ¬ (x is reptile → x is a mammal)

Adding *mammals* and *reptiles*



Hyponymy and Hypernymy — Transitive Relations

As you can see, the relations of hyponymy and hypernymy are *transitive*.

- If we know that:
  - A mammal is an animal.
  - An animal is an animate.
- Then we can deduce that:
  - A mammal is an animate.
- Because:
  - All mammals are animals.
  - All animals are animates.

Meronymy and Holonymy

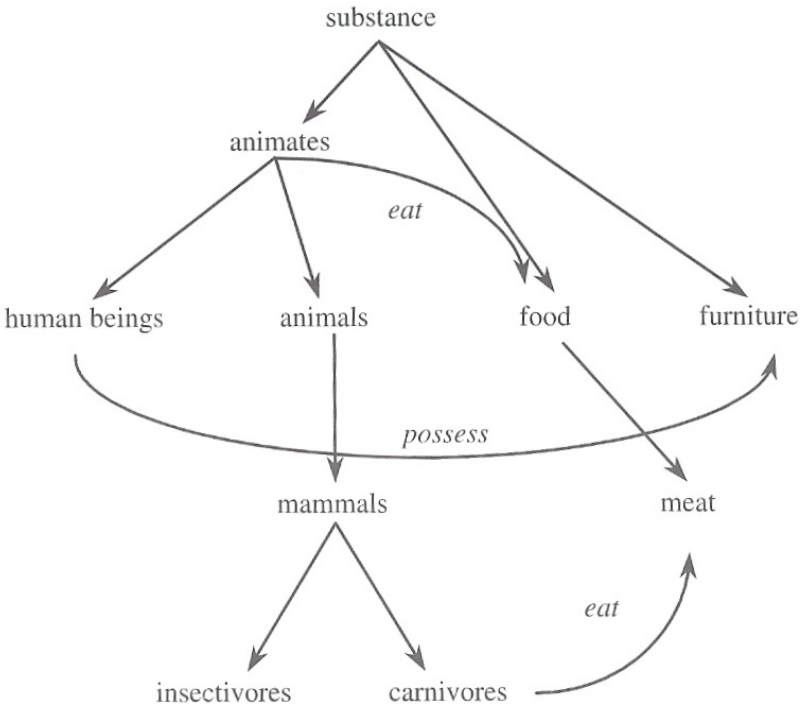
Another important relation is in terms of parts and wholes.

- Human beings have a head.
  - *head* is a meronym of *human being*
    - From Greek, *mero* 'part' + *nym* 'name'
  - *human being* is a holonym of *head*
    - From Greek, *holo* 'whole' + *nym* 'name'
- A head has a nose.
  - *nose* is a meronym of *head*
  - *head* is a holonym of *nose*
- Again, the relations are (generally) transitive.
  - If a human being has a head and a head has a nose then a human being has a nose!

But beware! This doesn't always work as you would expect.

- A house has a door.
- A door has a handle.
- !!A house has a handle.

Predicate Relations



In a more sophisticated semantic network, nodes can be linked with predicate relations of various kinds.

- is_a(food, substance)
- eat(animates, food)
- possess(human beings, furniture)

Machine-Readable Dictionaries

Lexical databases are central to work in natural language processing.

- Most are based on paper dictionaries and often retain their basic structure.
  - citation form (lemma), individual senses (polysemy), syntactic and semantic annotations, pronunciations, definitions.
- The most influential Machine-Readable Dictionary (MRD) among several options is the *Longman Dictionary of Contemporary English* (LDOCE, Procter 1978).

Ambiguity

Homonyms (same form; completely distinct sense) are usually represented in dictionaries as separate entries with identical lemmata (citation forms).

- bank[1], Noun: organization
- bank[2], Noun: raised ground
- bank[3], Verb: to turn

Polysemous lexemes (same form; closely related senses) are usually represented as one lemma (citation form) with separately defined sub-senses.

- meal, Noun:
  - 1: an occasion where food is eaten
  - 2: the food eaten on such an occasion

A major problem of text processing is therefore disambiguation.

- i.e. selecting the correct sense of a polysemous word (or the correct lexical item for homonyms), in the current context.

Problem of Disambiguation

p.349 The patron ordered a meal.

a. patron, Noun:
   (i) a person who gives money or support to a person, organisation, cause, activity
   (ii) **a customer of a shop, restaurant, theater**

b. order, Verb:
   (i) to give an order to somebody
   (ii) to request somebody to supply or make goods, etc
   (iii) **to request somebody to bring food, drink etc in a hotel, restaurant etc**
   (iv) to put something in order

c. meal, Noun:
   (i) an occasion where food is eaten
   (ii) **the food eaten on such an occasion**

Verbs — Argument Structure

Dictionaries differ in the degree to which they represent the selectional properties of the verb.

- Syntactic selection
  - English *die*: NP V
  - English *rely*: NP V on NP
  - Icelandic *sakna* 'miss': $\text{NP}^{nominative}$ V $\text{NP}^{genitive}$
- Semantic selection (e.g. restriction by ontological type)
  - English *drink*: $\text{NP}^{animate}$ V NP^{liquid}

Typical Structure of Dictionary Definition

The definition of lexical sense typically takes the form:

- lemma: **genus** *differentia specifica*
  - lemma: citation form of the lexeme
  - genus: general kind (cf. hypernym)
  - differentia specifica: the attributes which distinguish it from other members of the general kind (i.e. other hyponyms of the same hypernym)
- bank: **a land** *sloping up along each side of a canal or river*

WordNet

Not all lexicons are simply machine-readable forms of paper dictionaries, however.

- WordNet is a lexical database of English constructed to represent word sense relations.
  - Miller 1995, Fellbaum 1998
  - `http://wordnet.princeton.edu/`
- Wide coverage (120,000 words)
- Has proved to be a tractable model (see `http://www.globalwordnet.org/`)
  - WordNets for Dutch, German, Basque, Japanese etc.
  - A current RANNÍS Project of Excellence includes a work package dedicated to building up a semantic network for Icelandic.

Use it online! `http://wordnetweb.princeton.edu/perl/webwn`

Lexical Matrix

A lexical sign or word in WordNet is a pairing of a form and an individual sense or meaning. You can think of this as a matrix, with each word representing a cell.

Word Meanings	**Word Forms**				
	F_1	F_2	...	...	F_n
M_1	$E_{1,1}$	$E_{1,2}$			
M_2		$E_{2,2}$			
$\vdots$					
M_m					$E_{m,n}$

- A row in the matrix includes all word forms that can express the same sense, i.e. synonyms, i.e. sets of synonyms, i.e. synsets
- A column in the matrix includes all the senses of a word form, i.e. polysemous word forms.

Synset Relations in WordNet

WordNet then represents various semantic relations between synsets.

- WordNet keeps the major parts of speech separate
  - Nouns, Verbs, Adjectives, Adverbs
- Partly based on psycholinguistic evidence that people store words of the same syntactic part of speech together, even when there are meaning relations to words in other parts of speech.
- Partly because the words in each of the parts of speech seem to have a different organisational structure.
- Partly because they started with Nouns only!

Nouns

Nouns are organised in terms of hypernym relations of the sort we have been looking at.

- 95,000 nouns
- Nouns are organised into 25 general ontological kinds / primitive concepts / semantic primes.
- These are further clustered into a number of abstract root types.

Ontological Types for Nouns

- 25 general ontological types for Nouns

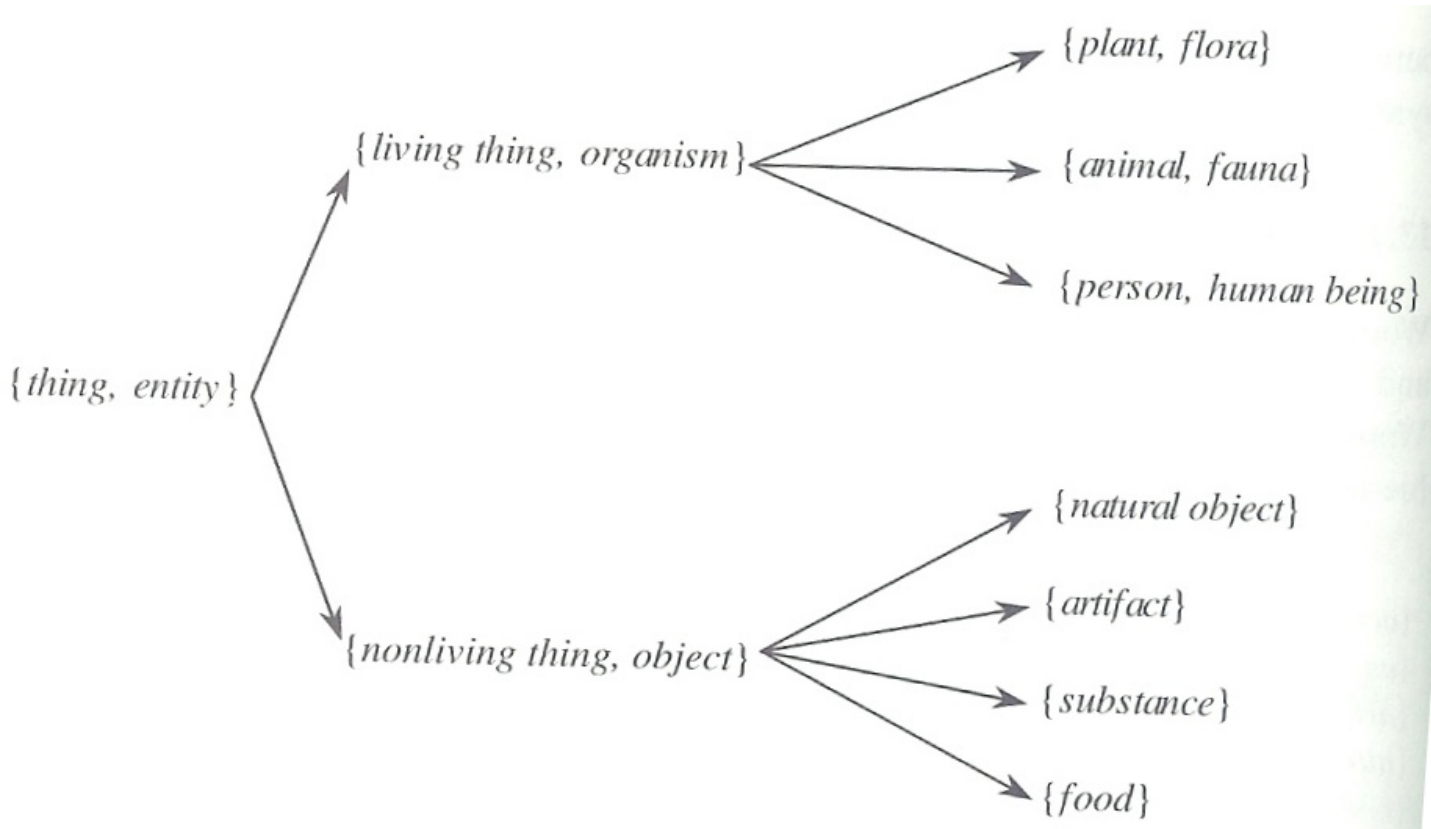
{*act, action, activity*}	{*food*}	{*possession*}
{*animal, fauna*}	{*group, collection*}	{*process*}
{*artifact*}	{*location, place*}	{*quantity, amount*}
{*attribute*}	{*motive*}	{*relation*}
{*body, corpus*}	{*natural object*}	{*shape*}
{*cognition, knowledge*}	{*natural phenomenon*}	{*state, condition*}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

- Abstract root clusters of ontological types.

{*entity, something*}	{*state*}	{*group, grouping*}
{*psychological feature*}	{*event*}	{*possession*}
{*abstraction*}	{*act, human action, human activity*}	{*phenomenon*}

Nugues, Fig. 13.7

N.B. The relation between "food" and "substance" here does not seem to be consistent with the latest online version of WordNet or his own listing in the text.



Latest Hypernym Hierarchy for "food"

- S: (n) **food**, nutrient (any substance that can be metabolized by an animal to give energy and build tissue)
  - *direct hyponym* / *full hyponym*
  - *part meronym*
  - *direct hypernym* / ***inherited hypernym*** / *sister term*
    - S: (n) substance (a particular kind or species of matter with uniform properties) "*shigella is one of the most toxic substances known to man*"
      - S: (n) matter (that which has mass and occupies space) "*physicists study both the nature of matter and the forces which govern it*"
        - S: (n) physical entity (an entity that has physical existence)
          - S: (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))

Relational Adjectives

Adjectives are divided into two major classes:

- relational
- descriptive

Relational adjectives are usually related to nouns.

- fraternal > brother
- contextual > context

This is the main information included in the entry for relational adjectives.

Descriptive Adjectives

Descriptive adjectives usually describe a property, quality or attribute of a noun.

- hot, tall, wet, heavy

An important organisational feature of adjectives is antonymy.

- hot∼cold; tall∼short; wet∼dry; heavy∼light

In each of these cases, the antonyms represent polar opposite values for a particular attribute, e.g.

- TEMPERATURE: hot∼cold
- HEIGHT: tall∼short

Descriptive adjectives are therefore listed with the attribute they restrict.

Of Adjectives and Bicycles

It turns out, however, that not every adjective has an antonym, even when it is a synonym for an adjective that does.

- synonym(hot, torrid)
- antonym(hot, cold)
- ¬[antonym(torrid, cold)]

Adjective networks in WordNet therefore often have a "bicycle" structure.

- There is an antonym pair: hot∼cold
- Radiated around each antonym are then the adjectives which have a similar meaning in some context.
  - hot: baking, blistering, scorching, sizzling, sultry, stifling, sweltering, torrid, tropical
  - cold: arctic, frigid, gelid, glacial, chilly, parky, crisp, frosty

Verbs and Entailments

An important relation for structuring the domain of verb senses is entailment.

- $p \rightarrow q$

WordNet recognises four kinds of relation between verb senses that produce entailments.

- manner entailments; extension; troponymy
  - From Greek, *tropo* 'manner' + *nym* 'name'
- inclusion
- presupposition (technically NOT an entailment!)
- causation

Troponymy

Troponymy is the equivalent in the verbal domain to hyponymy in the nominal domain.

- Ambling is a kind of walking.
- x ambles $\rightarrow$ x walks
- $\neg$(x walks $\rightarrow$ x ambles)

Here there really is a "kind" relation because *amble* represents a particular manner of performing a more general action.

- *amble* provides information which specifies or extends the description provided by *walk*

Causation

Causation is a very common and important relation between verbs in human languages.

(8) Icelandic *frysta* 'freeze (transitive)'∼*frjósa* 'freeze (intransitive)'.

a. Jón frysti kjúklinginn.
John froze chicken=the
'John froze the chicken.'

b. Kjúklingurinn fraus.
chicken=the froze
'The chicken froze.'

frysta 'freeze (transitive)' and *frjósa* 'freeze (intransitive)' are related by causation.

- John froze the chicken = John caused the chicken to freeze.
- x frysti y → y fraus
- ¬(y fraus → x frysti y)

Inclusion

Inclusion also produces entailments.

- x snores $\rightarrow$ x sleeps
- $\neg$(x sleeps $\rightarrow$ x snores)

However, this is not for either of the reasons we have seen so far.

- Snoring is not a kind of sleeping.
- Sleeping does not cause snoring (you can sleep and not snore).

The relation here is a little bit like meronymy in the nominal domain.

- Snoring is a possible part of sleeping (is included in it).
- Except that it is not a necessary part.

Presupposition

Finally, there is presupposition — a proposition that we assume to be true in order to evaluate the main proposition of the sentence.

- John managed to open the safe.
- John tried to open the safe.

You can only "manage" to do something if you "try" to do it. One of the strange properties of presupposition is that it usually remains under negation.

- John didn't manage to open the safe.
- John tried to open the safe.

Senses as Tags

A major semantic challenge for natural language processing is sense disambiguation, i.e.

- picking the right sense for a word form in a particular context

This can be recast as a kind of tagging problem.

- each word sense is represented by a semantic tag
- a word form may be associated with many tags
- disambiguation involves chosing one tag (the correct one, hopefully!)

However, the domain is challenging because

- there is no consensus on how many senses a word has
- how precisely those senses should be defined
- interannotator agreement for semantic tagging is much lower than for syntactic tagging

Numerical and Symbolic Methods

As with syntactic tagging, methods include:

- numerical
  - statistical method to optimise a sequence of semantic tags based on the probability of co-occurence derived from a training corpus
  - e.g. Naive Bayes classifier (which I will leave it to Hrafn to review!)
- symbolic
  - a constraint based approach, using constraints to exclude incorrect senses

Words in Context

As always, context is important.

- *bank* in the sense of a financial institution is likely to occur together with a quite different set of words from *bank* in the sense of the side of a river.
- information on related senses by topic can therefore help to disambiguate such cases.

Sometimes the context needs to be very specific.

- In our example *the patron ordered the meal*, the two senses of *meal* both relate to the same general topic.
  - Topic: eating something at a restaurant
  - meal^1: occasion of eating food
  - meal^2: food eaten on such an occasion

"Topic" Information

Some dictionaries such as the LDOCE tag word senses for general topic.

- agriculture, business, engineering etc.

Databases such as WordNet obviously include hypernym information that might be useful.

- hypernym(bank^1, financial institution)
- hypernym(bank^2, geological formation)

N.B. The "Tennis Problem" for WordNet

The Tennis Problem

What we are calling a "topic" is a coherent situation — like a game of tennis.

- In a game of tennis, two to four players assemble on a court and hit a ball using racquets over a net.

However, in WordNet, the various aspects of this coherent situation are distributed in unrelated parts of the network.

- tennis player is a hyponym of person
- racquet is a hyponym of artifact
- tennis court is a hyponym of location
- tennis stroke is a hyponym of act

WordNet can therefore be ill-suited to providing the sort of "topic" context that is useful for sense disambiguation.

Context and Bag of Words

One influential technique does not take the word-by-word context but rather only the nouns that occur within the relevant window.

- the bag-of-words technique
- it is a simple technique in that it ignores a lot of grammatical detail
- and exploits the fact that nouns are often "contentful" words relating to the current topic
- overlap in the senses of nouns appearing in the context window (whatever their order) can provide valuable disambiguation information.

The bag-of-words technique can be used with the Naive Bayes classifier to give a statistical estimate of most likely word sense collocations.

- Note however that this method requires hand-annotated corpora for training, which is very labour intensive.

Exploiting Verb Argument Structure

The selectional properties of verbs can also help to disambiguate senses.

(9) a. He ate a meal.
b. He scheduled a meal.

The argument structures of the verbs *eat* and *schedule* impose syntactic and semantic restrictions.

(10) a. $NP^{animate}$ eat NP^{food}
b. NP^{human} schedule NP^{event}

The argument structures pick out distinct senses of meal.

- $meal^1$: occasion of eating food (EVENT)
- $meal^2$: food eaten on such an occasion (FOOD)

Here we need a shallow parse to identify heads of phrases and relative position to the head verb.

An Unsupervised Disambiguation Algorithm

Wilks and Stevenson (1997), using a method inspired by Lesk (1986), describe an algorithm for sense disambiguation using only general dictionary definitions as a resource.

- Name recogniser picks out proper nouns.
- Lemmatiser assigns each word token its lemma. Via the lemma, each token is assigned a list of its possible senses.
- Tagger assigns each word its part of speech tag.
  - particular word senses may be associated with particular parts of speech
  - senses which clash with tagger-assigned part of speech can be discarded

An Unsupervised Disambiguation Algorithm

- for the remaining senses, each definition is analysed based on the words used in the definition
  - words (lemmas) in the definition of each sense of the target token are matched with words (lemmas) in the definitions of tokens in the context.
  - whenever there is a match, the sense with the match is assigned a point
- the sense of the target word which scores highest in the definition overlap estimation is the one that is selected.

N.B. The definition overlap estimation is obviously very memory intensive!