Semantics

Dr. G. Bharadwaja Kumar

Three Perspectives on Meaning

1. **Lexical Semantics**
   - The meanings of individual words
2. **Formal Semantics** (or Compositional Semantics or Sentential Semantics)
   - How those meanings combine to make meanings for individual sentences or utterances
3. **Discourse or Pragmatics**

   How those meanings combine with each other and with other facts about various kinds of context to make meanings for a text or discourse
   - **Dialog or Conversation** is often lumped together with Discourse

Why semantics

- Question

  What plants are native to Scotland?

- Text available to the machine

  A new chemical plant was opened in Scotland.

- What is hard?
  - words may have different meanings (**senses**)
  - we need to be able to disambiguate between them

Slides barrowed from Sharon Goldwater

- Question

  Where did Theresa May go on vacation?

- Text available to the machine

  Theresa May spent her holiday in Cornwall

- What is hard?
  - words may have the same meaning (**synonyms**)
  - we need to be able to match them

- Question

  Which animals love to swim?

- Text available to the machine

  Polar bears love to swim in the freezing waters of the Arctic.

- What is hard?
  - words can refer to a subset (**hyponym**) or superset (**hypernym**) of the concept referred to by another word
  - we need to have database of such **A is-a B** relationships, called an **ontology**

- Question

  Did Poland reduce its carbon emissions since 1989?

- Text available to the machine

  Due to the collapse of the industrial sector after the end of communism in 1989, all countries in Central Europe saw a fall in carbon emissions.

  Poland is a country in Central Europe.

- What is hard?
  - we need to do inference
  - a problem for sentential, not lexical, semantics

Terminology

- **Lexeme:** a pairing of meaning and form
- **Lemma:** the word form that represents a **lexeme**
  - ***Carpet*** is the lemma for ***carpets***
  - ***Dormir*** is the lemma for ***duermes***

- The lemma *bank* has two **senses:**
  - *Financial insitution*
  - *Soil wall next to water*

- A **sense** is a discrete representation of one aspect of the meaning of a word

Relationships between word meanings

- Homonymy
- Polysemy
- Synonymy
- Antonymy
- Hypernomy
- Hyponomy
- Meronomy
- Troponymy

Homonymy

Homonymy:

- Lexemes that share a form
  - Phonological, orthographic or both
- But have unrelated, distinct meanings
- Clear example:
  - Bat (wooden stick-like thing) vs
  - Bat (flying scary mammal thing)
  - Or bank (financial institution) versus bank (riverside)
- Can be homophones, homographs, or both:
  - Homophones:
    - Write and right
    - Piece and peace

Homonymy

The **bank** is constructed from red brick
I withdrew the money from the **bank**
Are those the same sense?

- Sarah climbed **down** the ladder.
- Sarah bought a **down** blanket.

- My dog would always **bark** at mailmen.
- The tree's **bark** was a rusty brown.

Homonymy causes problems for NLP applications

- Text-to-Speech
  - Same orthographic form but different phonological form
    - bass vs bass
- Information retrieval
  - Different meanings same orthographic form
    - QUERY: bat care
- Machine Translation
- Speech recognition

Polysemy

A word is polysemous if it can be used to express different meanings. The difference between the meanings can be obvious or subtle.

- The **newspaper** got wet in the rain.
- The **newspaper** fired some of its editing staff.

- John was a **good** man. He donated a lot of money to charity.
- Bill was a **good** painter. His drawings always were exciting to look at.

- Laura was a very **bright** student and always got good grades.
- The lights in this room are very **bright.**

Synonyms

Word that have the same meaning in some or all contexts.

big / huge
automobile / car
vomit / throw up
sick / ill
Hot/ warm

Synonyms

But there are few (or no) examples of perfect synonymy.

Even if many aspects of meaning are identical

Still may not preserve the acceptability based on notions of politeness, slang, register, genre, etc.

Language register is the level of formality with which you speak.

Synonymy is a relation between senses rather than words

Consider the words *big* and *large*

Are they synonyms?

How **big** is that plane?

Would I be flying on a **large** or small plane?

How about here:

Miss Nelson, for instance, became a kind of **big** sister to Benjamin.

?Miss Nelson, for instance, became a kind of **large** sister to Benjamin.

Why?

big has a sense that means being older, or grown up

large lacks this sense

Antonyms

Senses that are opposites with respect to one feature of their meaning

Otherwise, they are very similar!

- dark / light
- short / long
- hot / cold
- up / down
- in / out

More formally: antonyms can

- define a binary opposition or at opposite ends of a scale (*long/short, fast/slow*)
- Be **reversives**: *rise/fall, up/down*

Hyponymy

One sense is a **hyponym** of another if the first sense is more specific, denoting a subclass of the other

car is a hyponym of *vehicle*

dog is a hyponym of *animal*

mango is a hyponym of *fruit*

Conversely

vehicle is a hypernym/superordinate of *car*

animal is a hypernym of *dog*

fruit is a hypernym of *mango*

superordinate	vehicle	fruit	furniture	mammal
hyponym	car	mango	chair	dog

Hypernymy more formally

Extensional:

The class denoted by the superordinate extensionally includes the class denoted by the hyponym

Entailment:

A sense A is a hyponym of sense B if being an A entails being a B

Hyponymy is usually transitive

(A hypo B and B hypo C entails A hypo C)

Metonymy

Metonymy is a figure of speech used in rhetoric in which a thing or concept is not called by its own name, but by the name of something intimately associated with that thing or concept

“Hollywood" is used as a metonym (an instance of metonymy) for American cinema

“scepter” for “sovereignty,” or “count heads (or noses)” for “count people.”.

Metonymy

The **White House** announced yesterday.

The term *White House* actually refers to the authorities who work in the building called the White House.

Synecdoche

A synecdoche is a figure of speech in which a term for a part of something refers to the whole of something or vice versa

Examples from common English expressions include "bread and butter" (for "livelihood"), "suits" (for "businessmen"), and "boots" (for "soldiers") (pars pro toto), or conversely "America" (for "the United States of America")

Meronymy

A meronym denotes a constituent part of, or a member of something. That is,

"X" is a meronym of "Y" if Xs are parts of Y(s), or
"X" is a meronym of "Y" if Xs are members of Y(s).

For example, finger is a meronym of hand because a finger is part of a hand.

Similarly, wheels is a meronym of automobile.

Troponym

A verb that indicates more precisely the manner of doing something by replacing a verb of a more generalized meaning.

"Stroll" is a troponym of "walk", since it involves a slow and relaxed sort of walking.

to gorge is to eat in a different manner (something that is swallowed; contents of the stomach.)

Some examples

Pattern	Participating Senses	Example Sentences
Animal for fur	Mink, chinchilla, rabbit, beaver, raccoon*, alpaca*, crocodile*	The *mink* drank some water / She likes to wear *mink*
Animal/Object for personality	Chicken, sheep, pig, snake, star*, rat*, doll*	The *chicken* drank some water / He is a *chicken*
Animal for meat	Chicken, lamb, fish, shrimp, salmon*, rabbit*, lobster*	The chicken drank some water / The *chicken* is tasty
Artifact for activity	Shower, bath, sauna, baseball,	The *shower* was leaking / The *shower* was relaxing
Body part for object part	Arm, leg, hand, face, back*, head*, foot*, shoulder*, lip*,	John's *arm* was tired / The *arm* was reupholstered
Building for people	Church, factory, school, airplane,	The *church* was built 20 years ago / The *church* sang a song
Complement Coercion	Begin, start, finish, try	John *began* reading the book / John *began* the book
Container for contents	Bottle, can, pot, pan, bowl*, plate*, box*, bucket*	The *bottle* is made of steel / He drank half of the *bottle*
Word for question	Price, weight, speed	The *price* of the coffee was low / John asked the *price* of the coffee

Pattern	Participating Senses	Example Sentences
Figure for Ground	Window, door, gate, goal	The window is broken / The cat walked through the window
Grinding	Apple, chair, fly	The apple was tasty / There is apple all over the table
Instrument for action	Hammer, brush, shovel, tape, lock*, bicycle*, comb*, saw*	The hammer is heavy / She hammered the nail into the wall
Instance of an entity for kind	Tennis, soccer, cat, dog, class*, dinner*, chair*, table*	Tennis was invented in England / Tennis was fun today
Location / Place at location	Bench, land, floor, ground, box*, bottle*, jail*	The bench was made of pine / The coach benched the player
Object for placing at goal	Water, paint, salt, butter, frame*, dress*, oil*	The water is cold / He watered the plant.
Object for taking from source	Milk, dust, weed, peel, pit*, skin*, juice*	The milk tastes good / He milked the cow
Material for artifact	Tin, iron, china, glass, linen*, rubber*, nickel*, fur*	Watch out for the broken glass / He filled the glass with water
Occupation for role in action	Boss, nurse, guard, tutor	My boss is nice / He bossed me around

Pattern	Participating Senses	Example Sentences
Place for an event	Vietnam, Korea, Waterloo, Iraq	It is raining in *Vietnam* / John was shot during *Vietnam*
Place for an institution	White House, Washington, Hollywood, Pentagon, Wall Street*, Supreme Court	The *White House* is being repainted / The *White House* made an announcement
Plant for food or material	Corn, broccoli, coffee, cotton, lettuce*, eggs*, oak*, pine*	The large field of *corn* / The *corn* is delicious
Portioning	Water, beer, jam	She drank some *water* / She bought three *waters*
Publisher for product	Newspaper, magazine, encyclopedia, Wall Street Journal*, New York Times*,	The *newspaper* is badly printed / The *newspaper* fired three employees
Artist for product	Writer, artist, composer, Shakespeare, Dickens*, Mozart*, Picasso*	The *writer* drank a lot of wine / The *writer* is hard to understand
Object for contents	Book, CD, DVD, TV*, magazine*, newspaper*	The heavy, leather- bound *book* / The *book* is funny.
Visual Metaphor	Beam, belt, column, stick, bug*, leaf*	Most of the weight rests on the *beam* / There was a *beam* of light

- How many senses does the word interest have?
  - She pays 3% **interest** on the loan.
  - He showed a lot of **interest** in the painting.
  - Microsoft purchased a controlling **interest** in Google.
  - It is in the national **interest** to invade the Bahamas.
  - I only have your best **interest** in mind.
  - Playing chess is one of my **interests**.
  - Business **interests** lobbied for the legislation.
- Are these seven different senses? Four? Three?
- Also note: distinction between polysemy and homonymy not always clear!

senses for interest

S1: a sense of concern with and curiosity about someone or something, Synonym: involvement

S2: the power of attracting or holding one's interest (because it is unusual or exciting etc.), Synonym: interestingness

S3: a reason for wanting something done, Synonym: sake

S4: a fixed charge for borrowing money; usually a percentage of the amount borrowed

S5: a diversion that occupies one's time and thoughts (usually pleasantly), Synonyms: pastime, pursuit

S6: a right or legal share of something; a financial involvement with something, Synonym: stake

S7: (usu. plural) a social group whose members control some field of activity and who have common aims, Synonym: interest group

II. WordNet

- A hierarchically organized lexical database
- is a hand-built ontology containing 117,000 synsets: sets of synonymous words

POS	Unique Strings	Synsets	Total Word-Sense Pairs
Noun	117,798	82,115	146,312
Verb	11,529	13,767	25,047
Adjective	21,479	18,156	30,002
Adverb	4,481	3,621	5,580
Totals	155,287	117,659	206,941

Wordnet 3.0 statistics

POS	Monosemous	Polysemous	Polysemous
	Words and Senses	Words	Senses
Noun	101,863	15,935	44,449
Verb	6,277	5,252	18,770
Adjective	16,503	4,976	14,399
Adverb	3,748	733	1,832
Totals	128,391	26,896	79,450

What's special about WordNet?

- Traditional dictionaries are organized alphabetically: words that are found together are not related by *meaning*
- WordNet is organized by meaning: words in close proximity are semantically similar
- Human users and computers can browse WordNet and find words that are meaningfully related to their queries
- Meaning similiarity can be measured and quantified to support Natural Language Understanding

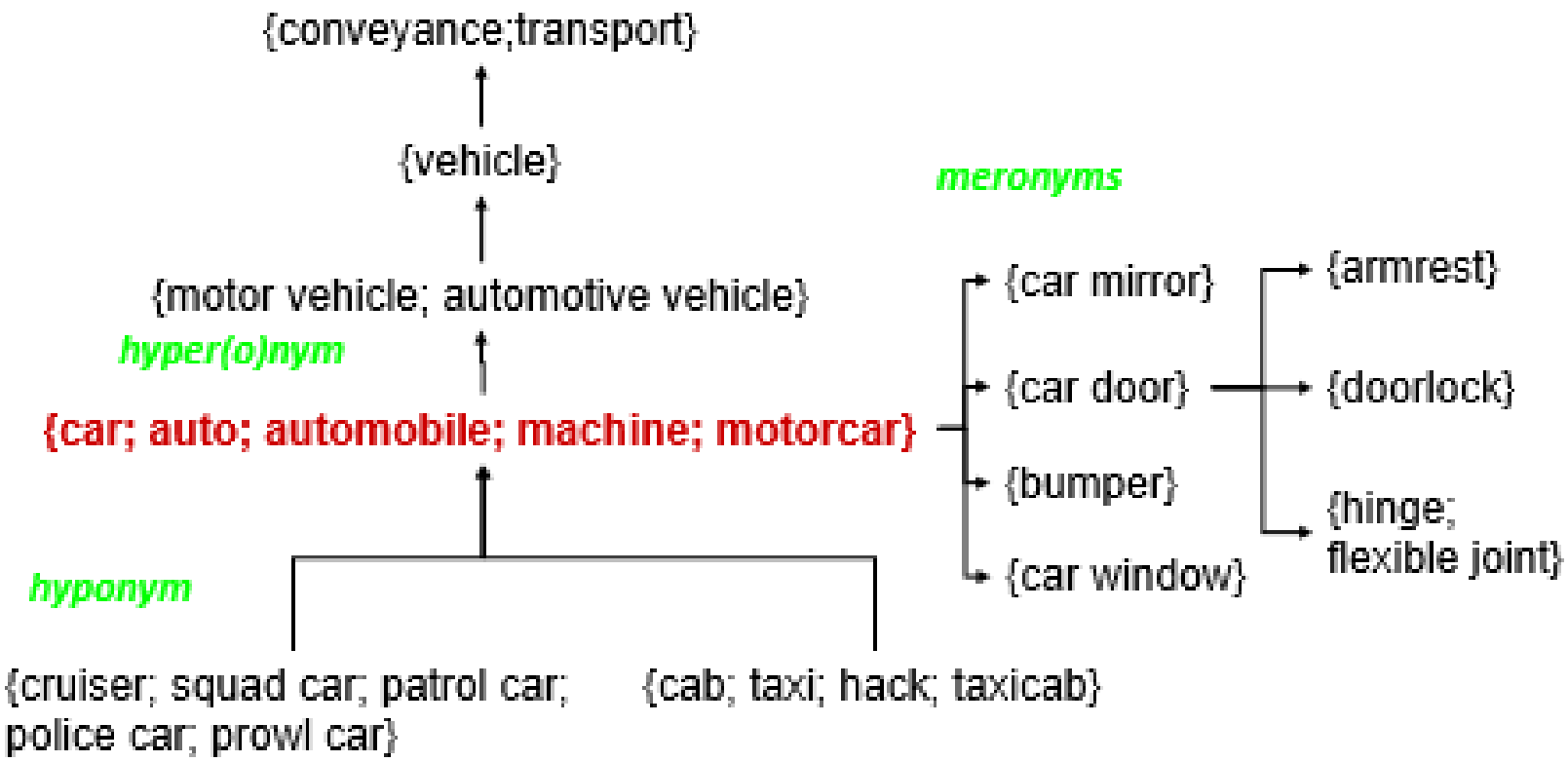
Relation to ontology

WordNet is sometimes called an ontology, because the hypernym/hyponym relationships among the noun synsets can be interpreted as specialization relations among conceptual categories.

What is ontology

In computer science and information science, an ontology is a formal naming and definition of the types, properties, and interrelationships of the entities that really or fundamentally exist for a particular domain of discourse. It is thus a practical application of philosophical ontology, with a taxonomy.

Example Representation



Structure of WordNet

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
Synonym	A concept that means exactly or nearly the same as another. *WordNet* considers immediate hypernyms to be synonyms.	×	×	×	×	{ ***sofa, couch, lounge*** } are all synonyms of one another. { ***seat*** } is the immediate hypernym of the synset.
Antonym	A concept opposite in meaning to another.	×	×	×	×	{ ***love*** } is the antonym of { ***hate, detest*** }.
Hypernym	A concept whose meaning denotes a superordinate.	×	×			A { ***feline, felid*** } is a hypernym of { ***cat, true cat*** }.
Hyponym	A concept whose meaning denotes a subordinate.	×	×			A { ***wildcat*** } is a hyponym of { ***cat, true cat*** }.
Substance meronym	A concept that is a substance of another concept.	×				A { ***snowflake, flake*** } is substance of { ***snow*** }.

Structure of WordNet

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
Part meronym	A concept that is a part of another concept.	×				A { ***crystal, watch crystal, watch glass*** } is a part of a { ***watch, ticker*** }.
Member meronym	A concept that is a member of another concept.	×				An { ***associate*** } is a member of an { ***association*** }.
Substance of holonym	A concept that has another concept as a substance.	×				A { ***tear, teardrop*** } has { ***water, H20*** } as a substance.
Part of holonym	A concept that has another concept as a part.	×				A { ***school system*** } has a { ***school, schoolhouse*** } as a part.
Member of holonym	A concept that has another concept as a member.	×				{ ***organized crime, gangland, gangdom*** } has { ***gang, pack, ring, mob*** } as a member.
Attribute	An adjective that is the value of a noun.	×				{ ***fast (vs. slow)*** } is a value of { ***speed, swiftness, fastness***}

Structure of WordNet

Semantic relation	Description	Part of speech				Example
		N	V	Adj	Adv	
Cause to	A verb that is the cause of a result.		X			{ ***give*** } is the cause of the result { ***have, have got, hold*** }
Entailment	A verb that involves unavoidably a result.		X			To { ***die, decease, perish, go, exit, pass away, expire*** } involves unavoidably to { ***leave, leave behind*** }.
Troponym	A verb that is a particular way to do another.		X			To { ***samba*** } is a particular way to { ***dance, trip the light fantastic*** }.
Pertainym	An adjective or adverb that relates to a noun.			X	X	{ ***criminal*** } relates to { ***crime*** }.
Attribute	An adjective that is the value of a noun.	X				{ ***fast (vs. slow)*** } is a value of { ***speed, swiftness, fastness***}
Value	A noun that has an adjective for a value.			X		{ ***weight*** } has { ***light (vs. heavy)*** } as a value.

Unique Beginners

Entity, something
(anything having existence (living or nonliving))

Psychological_feature
(a feature of the mental life of a living organism)

Abstraction
(a general concept formed by extracting common features from specific examples)

State
(the way something is with respect to its main attributes; "the current state of knowledge"; "his state of health"; "in a weak financial state")

Event
(something that happens at a given place and time)

Unique Beginners

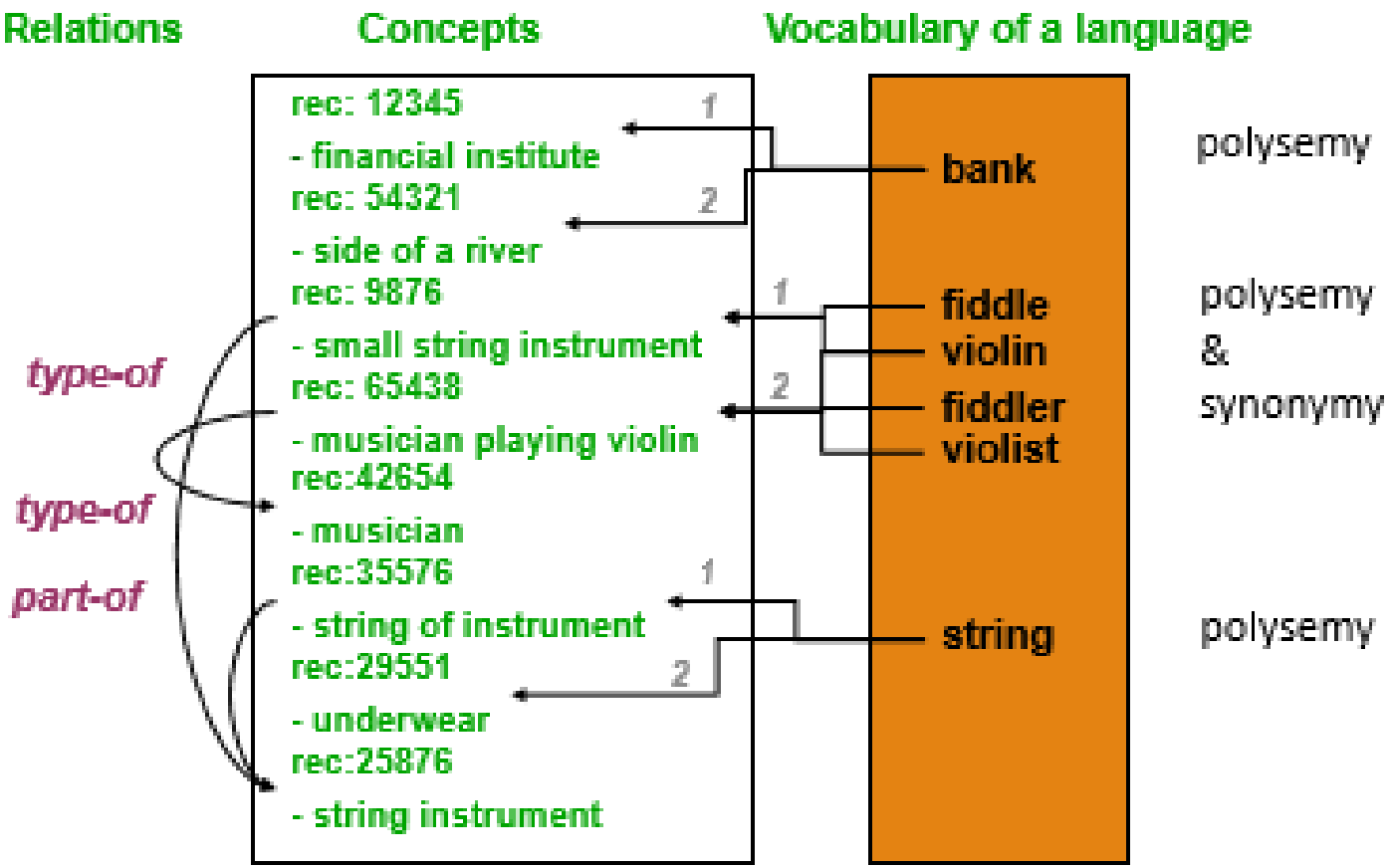
Act, human_action, human_activity
(something that people do or cause to happen)

Group, grouping
(any number of entities (members) considered as a unit)

Possession
(anything owned or possessed)

Phenomenon
(any state or process known through the senses rather than by intuition or reasoning)

WordNet


Relations
Concepts
Vocabulary of a language
rec: 12345
- financial institute
rec: 54321
- side of a river
rec: 9876
- small string instrument
rec: 65438
- musician playing violin
rec:42654
- musician
rec:35576
- string of instrument
rec:29551
- underwear
rec:25876
- string instrument
type-of
type-of
part-of
1
2
1
2
1
2
bank
fiddle
violin
fiddler
violist
string
polysemy
polysemy
&
synonymy
polysemy

Format of Wordnet Entries

The noun "bass" has 8 senses in WordNet.

1. bass[1] - (the lowest part of the musical range)
2. bass[2], bass part[1] - (the lowest part in polyphonic music)
3. bass[3], basso[1] - (an adult male singer with the lowest voice)
4. sea bass[1], bass[4] - (the lean flesh of a saltwater fish of the family Serranidae)
5. freshwater bass[1], bass[5] - (any of various North American freshwater fish with lean flesh (especially of the genus Micropterus))
6. bass[6], bass voice[1], basso[2] - (the lowest adult male singing voice)
7. bass[7] - (the member with the lowest range of a family of musical instruments)
8. bass[8] - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

The adjective "bass" has 1 sense in WordNet.

1. bass[1], deep[6] - (having or denoting a low vocal or instrumental range)
   "a deep voice"; "a bass voice is lower than a baritone voice"; "a bass clarinet"

WordNet Noun Relations

Relation	Also called	Definition	Example
Hypernym	Superordinate	From concepts to superordinates	$breakfast^1 \rightarrow meal^1$
Hyponym	Subordinate	From concepts to subtypes	$meal^1 \rightarrow lunch^1$
Member Meronym	Has-Member	From groups to their members	$faculty^2 \rightarrow professor^1$
Has-Instance		From concepts to instances of the concept	$composer^1 \rightarrow Bach^1$
Instance		From instances to their concepts	$Austen^1 \rightarrow author^1$
Member Holonym	Member-Of	From members to their groups	$copilot^1 \rightarrow crew^1$
Part Meronym	Has-Part	From wholes to parts	$table^2 \rightarrow leg^3$
Part Holonym	Part-Of	From parts to wholes	$course^7 \rightarrow meal^1$
Antonym		Opposites	$leader^1 \rightarrow follower^1$

WordNet Verb Relations

Relation	Definition	Example
Hypernym	From events to superordinate events	$fly^9 \rightarrow travel^5$
Troponym	From a verb (event) to a specific manner elaboration of that verb	$walk^1 \rightarrow stroll^1$
Entails	From verbs (events) to the verbs (events) they entail	$snore^1 \rightarrow sleep^1$
Antonym	Opposites	$increase^1 \Longleftrightarrow decrease^1$

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

Sense 7
bass --
(the member with the lowest range of a family of
musical instruments)
=> musical instrument, instrument
    => device
        => instrumentality, instrumentation
            => artifact, artefact
                => whole, unit
                    => object, physical object
                        => physical entity
                            => entity
```

How is "sense" defined in WordNet?

The set of near-synonyms for a WordNet sense is called a **synset (synonym set)**; it's their version of a sense or a concept

Example: chump as a noun to mean

'{chump[1], fool[2], gull[1], mark[9], patsy[1], fall guy[1], sucker[1], soft touch[1], mug[2]}

Each of these senses share this same gloss

Thus for WordNet, the meaning of this sense of chump *is* this list.

a synset contains a brief definition ("gloss") and, in most cases, one or more short sentences illustrating the use of the synset members.

Word Similarity

Synonymy is a binary relation
- Two words are either synonymous or not

We want a looser metric
- Word similarity or
- Word distance

Two words are more similar
- If they share more features of meaning

Actually these are really relations between **senses**:
- Instead of saying "bank is like fund"
- We say
  - Bank1 is similar to fund3
  - Bank2 is similar to slope5

We'll compute them over both words and senses

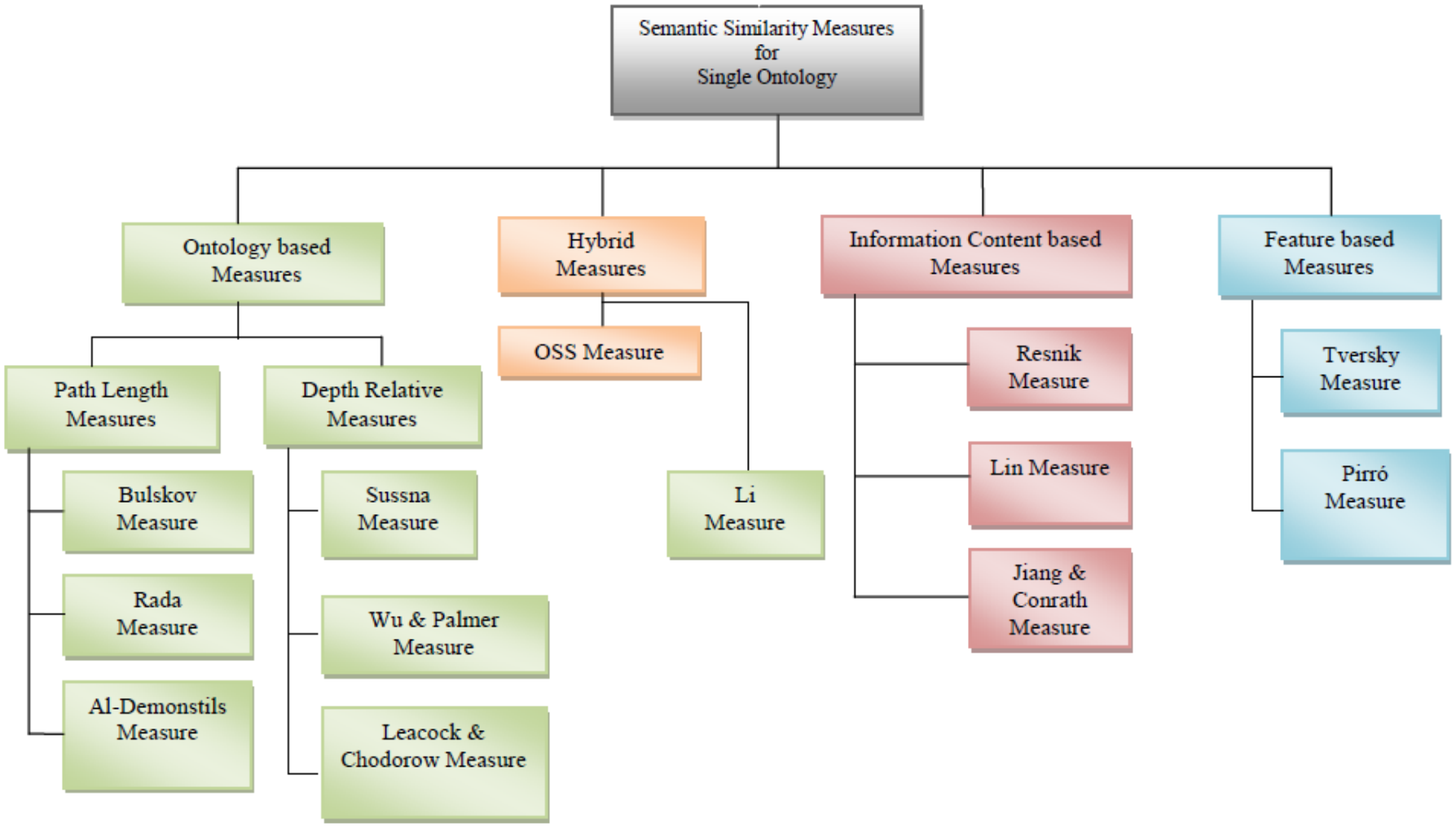
Why word similarity

Spell Checking
Information retrieval
Question answering
Machine translation
Natural language generation
Language modeling
Automatic essay grading

Two classes of algorithms

- Thesaurus-based algorithms
  - Based on whether words are "nearby" in Wordnet
- Distributional algorithms
  - By comparing words based on their distributional context in corpora

Classification Semantic Similarity Measures



Thesaurus-based word similarity

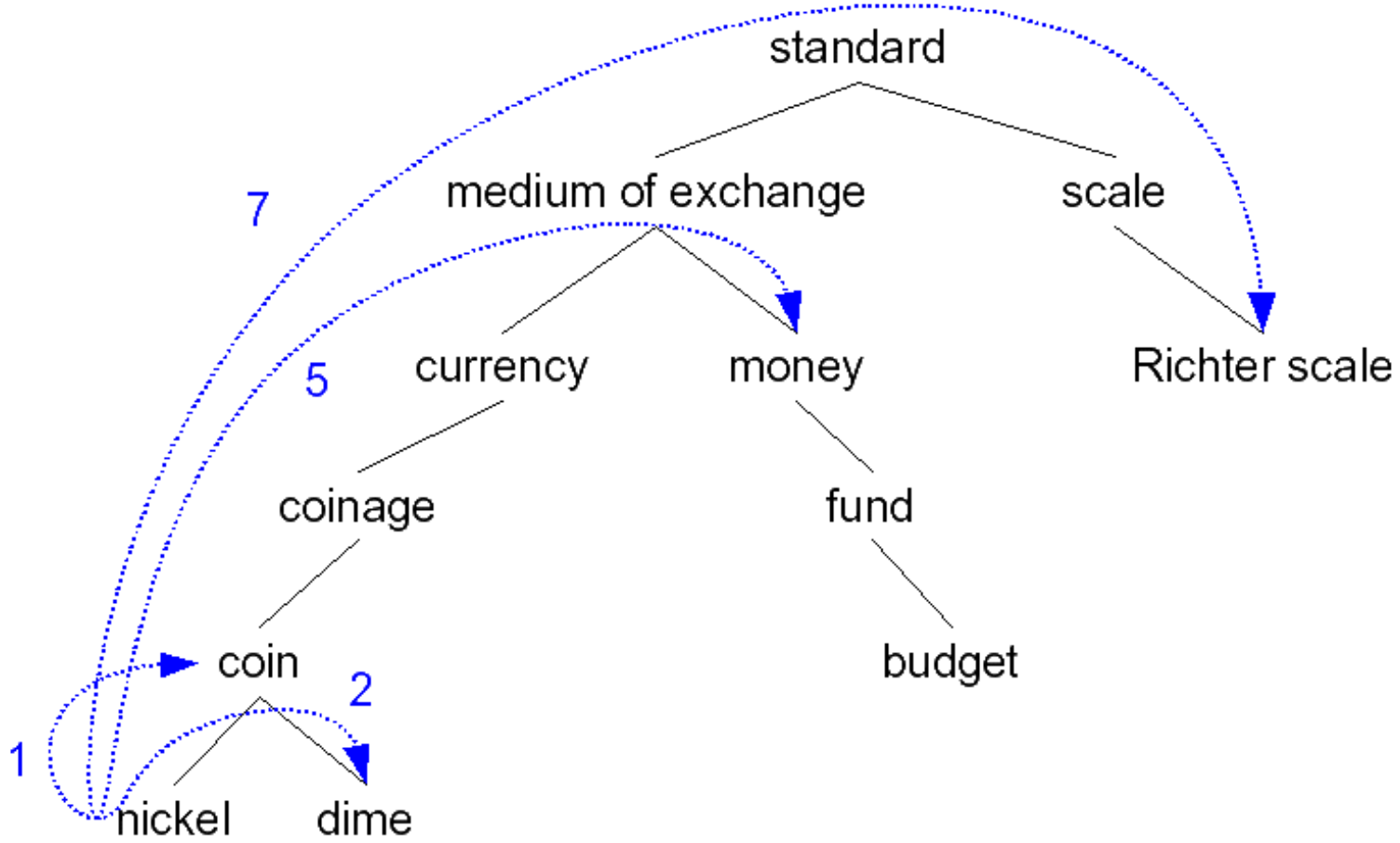
- Find words that are connected in the thesaurus
  - Synonymy, hyponymy, etc.
  - Glosses and example sentences
  - Derivational relations and sentence frames

- Word similarity versus word relatedness

  Similar words are near-synonyms
  Related could be related any way
  Car, gasoline: related, not similar
  Car, bicycle: similar

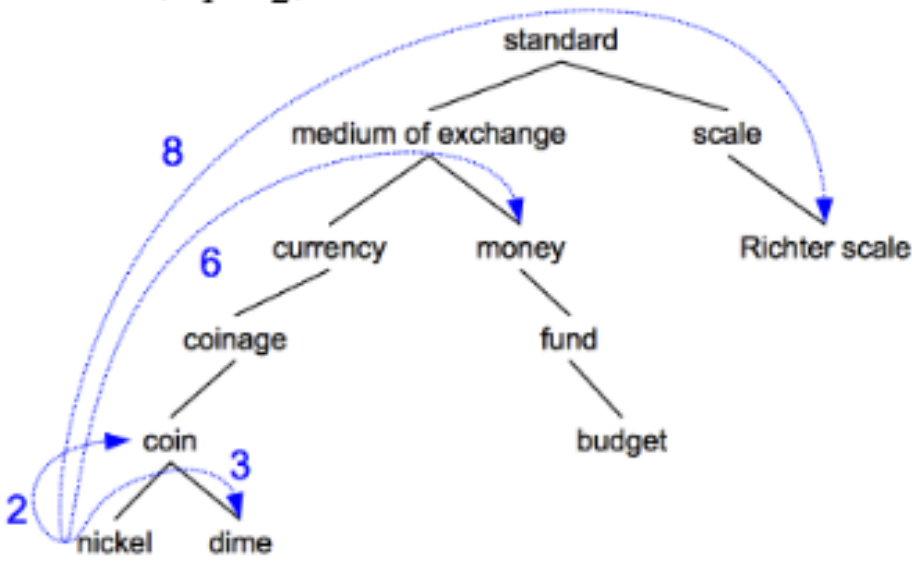
Path based similarity

Two words are similar if nearby in thesaurus hierarchy (i.e. short path between them)



Example: path-based similarity

$\text{simpath}(c_1,c_2) = 1/\text{pathlen}(c_1,c_2)$



simpath(*nickel,coin*) = $1/2 = .5$

simpath(*fund,budget*) = $1/2 = .5$

simpath(*nickel,currency*) = $1/4 = .25$

simpath(*nickel,money*) = $1/6 = .17$

simpath(*coinage,Richter scale*) = $1/6 = .17$

Refinements to path-based similarity

pathlen(c1,c2) = number of edges in the shortest path in the thesaurus graph between the sense nodes c1 and c2

simpath(c1,c2) = -log pathlen(c1,c2)

wordsim(w1,w2) =

$$\max_{c1 \in senses(w1), c2 \in senses(w2)} sim(c1,c2)$$

Problem with basic path-based similarity

Assumes each link represents a uniform distance

Nickel to *money* seem closer than *nickel* to *standard*

Instead:

Want a metric which lets us assign different "lengths" to different edges

Represent the cost of each edge independently

From paths to probabilities

Don't measure paths. Measure probability?

Define P(c) as the probability that a randomly selected word is an instance of concept (synset) c

P(ROOT) = 1

The lower a node in the hierarchy, the lower its probability

Information content similarity

Train by counting in a corpus

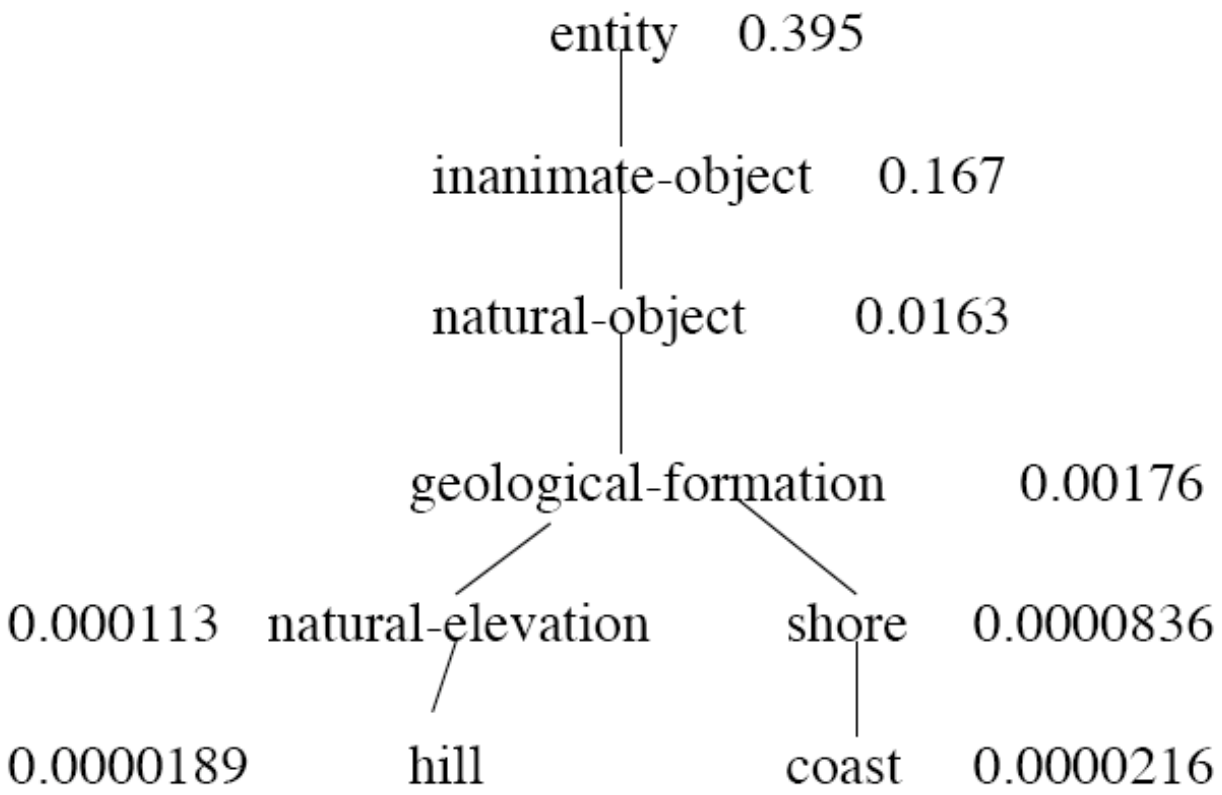
Each instance of "dime" could count toward frequency of *coin*, *currency*, *standard*, etc

More formally:

$$P(c) = \frac{\sum_{w \in words(c)} count(w)}{N}$$

Information content similarity

WordNet hierarchy augmented with probabilities P(C)



Information content: definitions

- Information content:
  IC(c)=-logP(c)

- Lowest common subsumer
  LCS(c1,c2) = the lowest common subsume

  i.e. the lowest node in the hierarchy
  That subsumes (is a hypernym of) both c1 and c2

Resnik method

The similarity between two words is related to their common information

The more two words have in common, the more similar they are

Resnik: measure the common information as:

The info content of the lowest common subsumer of the two nodes

$sim_{resnik}(c1,c2) = -\log P(LCS(c1,c2))$

Dekang Lin method

Similarity between A and B needs to do more than measure common information

The more **differences** between A and B, the less similar they are:

- Commonality: the more info A and B have in common, the more similar they are
- Difference: the more differences between the info in A and B, the less similar

Commonality: IC(Common(A,B))

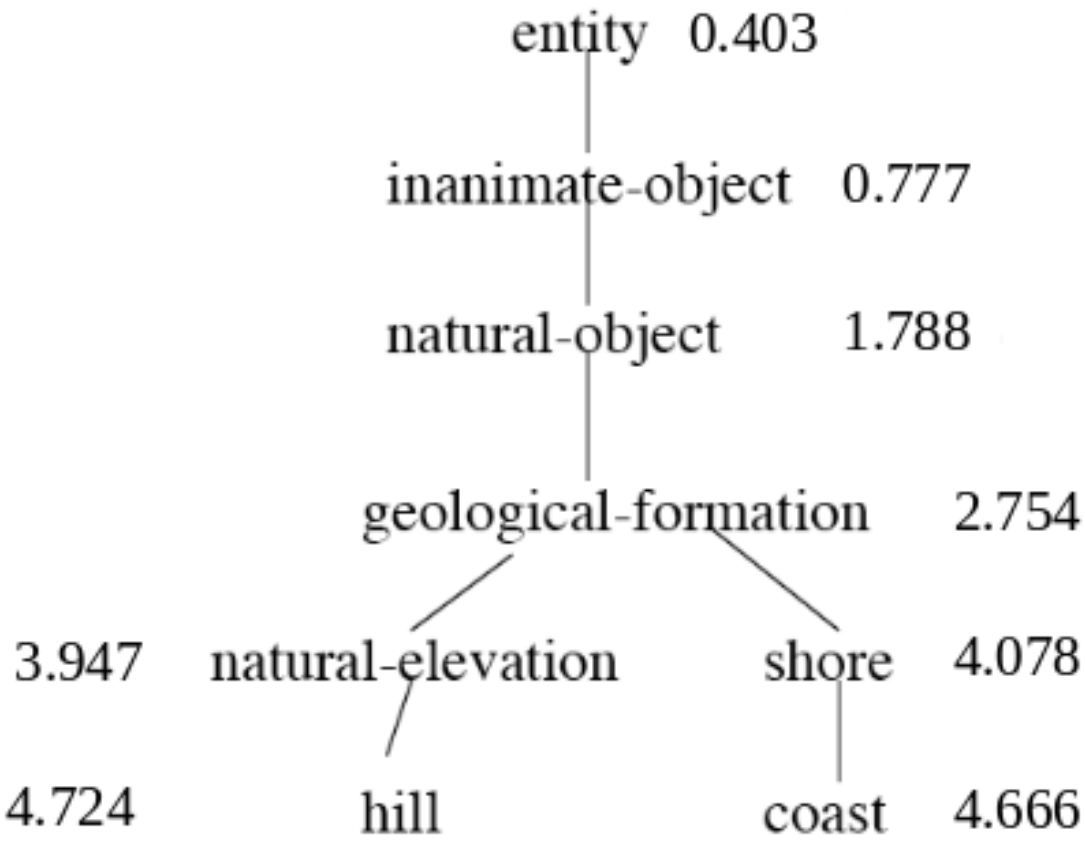
Difference: IC(description(A,B)-IC(common(A,B))

Dekang Lin method

Similarity theorem: The similarity between A and B is measured by the ratio between the amount of information needed to state the commonality of A and B and the information needed to fully describe what A and B are

$$= \frac{2 \times \log P(\mathrm{LCS}(A, B))}{\log P(A) + \log P(B)}$$

$sim_{resnik}(hill, coast) = ?$



Lin similarity function

$$\text{SimLin}(c1,c2) = \frac{2 \times \log P(\text{LCS}(c1,c2))}{\log P(c1) + \log P(c2)}$$

$$\text{SimLin}(\text{hill},\text{coast}) = \frac{2 \times \log P(\text{geological-formation}))}{\log P(\text{hill}) + \log P(\text{coast})}$$

$$= 0.612$$

Extended

Two concepts are similar if their glosses contain similar words

Drawing paper: **paper** that is **specially prepared** for use in drafting

Decal: the art of transferring designs from **specially prepared paper** to a wood or glass or metal surface

For each *n*-word phrase that occurs in both glosses

Add a score of n^2

Paper and *specially prepared* for 1 + 4 = 5…

Few measures for thesaurus-based similarity

$$\text{sim}_{\text{path}}(c_1, c_2) = -\log \text{pathlen}(c_1, c_2)$$

$$\text{sim}_{\text{Resnik}}(c_1, c_2) = -\log P(\text{LCS}(c_1, c_2))$$

$$\text{sim}_{\text{Lin}}(c_1, c_2) = \frac{2 \times \log P(LCS(c_1, c_2))}{\log P(c_1) + \log P(c_2)}$$

$$\text{sim}_{\text{jc}}(c_1, c_2) = \frac{1}{2 \times \log P(\text{LCS}(c_1, c_2)) - (\log P(c_1) + \log P(c_2))}$$

$$\text{sim}_{\text{eLesk}}(c_1, c_2) = \sum_{r,q \in \text{RELS}} \text{overlap}(\text{gloss}(r(c_1)), \text{gloss}(q(c_2)))$$

Problems with thesaurus-based methods

We don't have a thesaurus for every language

Even if we do, many words are missing

They rely on hyponym info:

- Strong for nouns, but lacking for adjectives and even verbs

Alternative

- Distributional methods for word similarity

Distributional methods for word similarity

Firth (1957): "You shall know a word by the company it keeps!"

Nida example noted by Lin:

A bottle of ***tezgüino*** is on the table

Everybody likes ***tezgüino***

Tezgüino makes you drunk

We make ***tezgüino*** out of corn.

Intuition:

just from these contexts a human could guess meaning of tezguino

So we should look at the surrounding contexts, see what other words have similar context.

Context vector

Consider a target word w

Suppose we had one binary feature f_i for each of the N words in the lexicon v_i

Which means "word v_i occurs in the neighborhood of w"

$w=(f_1,f_2,f_3,\ldots,f_N)$

If w=tezguino, v1 = bottle, v2 = drunk, v3 = matrix:

w = (1,1,0,...)

Intuition

Define two words by these sparse features vectors
Apply a vector distance metric
Say that two words are similar if two vectors are similar

	arts	boil	data	function	large	sugar	summarized	water
apricot	0	1	0	0	1	1	0	1
pineapple	0	1	0	0	1	1	0	1
digital	0	0	1	1	1	0	1	0
information	0	0	1	1	1	0	1	0

Distributional similarity

So we just need to specify 3 things

1. How the co-occurrence terms are defined
2. How terms are weighted
   (frequency? Logs? Mutual information?)
3. What vector distance metric should we use?
   Cosine? Euclidean distance?

Defining co-occurrence vectors

We could have windows of neighboring words

- Bag-of-words
- We generally remove **stopwords**

But the vectors are still very sparse

So instead of using ALL the words in the neighborhood

Let's just the words occurring in particular relations

Defining co-occurrence vectors

Zellig Harris (1968)

> The meaning of entities, and the meaning of grammatical relations among them, is related to the restriction of combinations of these entitites relative to other entities

Idea: parse the sentence, extract syntactic dependencies:

I discovered dried tangerines:

discover (subject I)	I (subj-of discover)
tangerine (obj-of discover)	tangerine (adj-mod dried)
dried (adj-mod-of tangerine)	

2. Weighting the counts ("Measures of association with context")

We have been using the frequency of some feature as its weight or value

But we could use any function of this frequency

Let's consider one feature

f=(r,w') = (obj-of,*attack*)

P(f|w)=count(f,w)/count(w)

$Assoc_{prob}(w,f)=p(f|w)$

Weighting: Mutual Information

Mutual information: between 2 random variables X and Y

$$I(X,Y) = \sum_x \sum_y P(x,y) \log_2 \frac{P(x,y)}{P(x)P(y)}$$

Pointwise mutual information: measure of how often two events x and y occur, compared with what we would expect if they were independent:

$$I(x,y) = \log_2 \frac{P(x,y)}{P(x)P(y)}$$

Weighting: Mutual Information

Pointwise mutual information: measure of how often two events x and y occur, compared with what we would expect if they were independent:

$$I(x,y) = \log_2 \frac{P(x,y)}{P(x)P(y)}$$

PMI between a target word w and a feature f :

$$\text{assoc}_{\text{PMI}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(f)}$$

Mutual information intuition

Objects of the verb *drink*

Object	Count	PMI assoc	Object	Count	PMI assoc
bunch beer	2	12.34	wine	2	9.34
tea	2	11.75	water	7	7.65
Pepsi	2	11.75	anything	3	5.15
champagne	4	11.75	much	3	5.15
liquid	2	10.53	it	3	1.25
beer	5	10.20	<SOME AMOUNT>	2	1.22

Lin is a variant on PMI

Pointwise mutual information: measure of how often two events x and y occur, compared with what we would expect if they were independent:

$$I(x,y) = \log_2 \frac{P(x,y)}{P(x)P(y)}$$

PMI between a target word w and a feature f:

$$\text{assoc}_{\text{PMI}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(f)}$$

Lin measure: breaks down expected value for P(f) differently:

$$\text{assoc}_{\text{Lin}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(r|w)P(w'|w)}$$

Summary: weightings

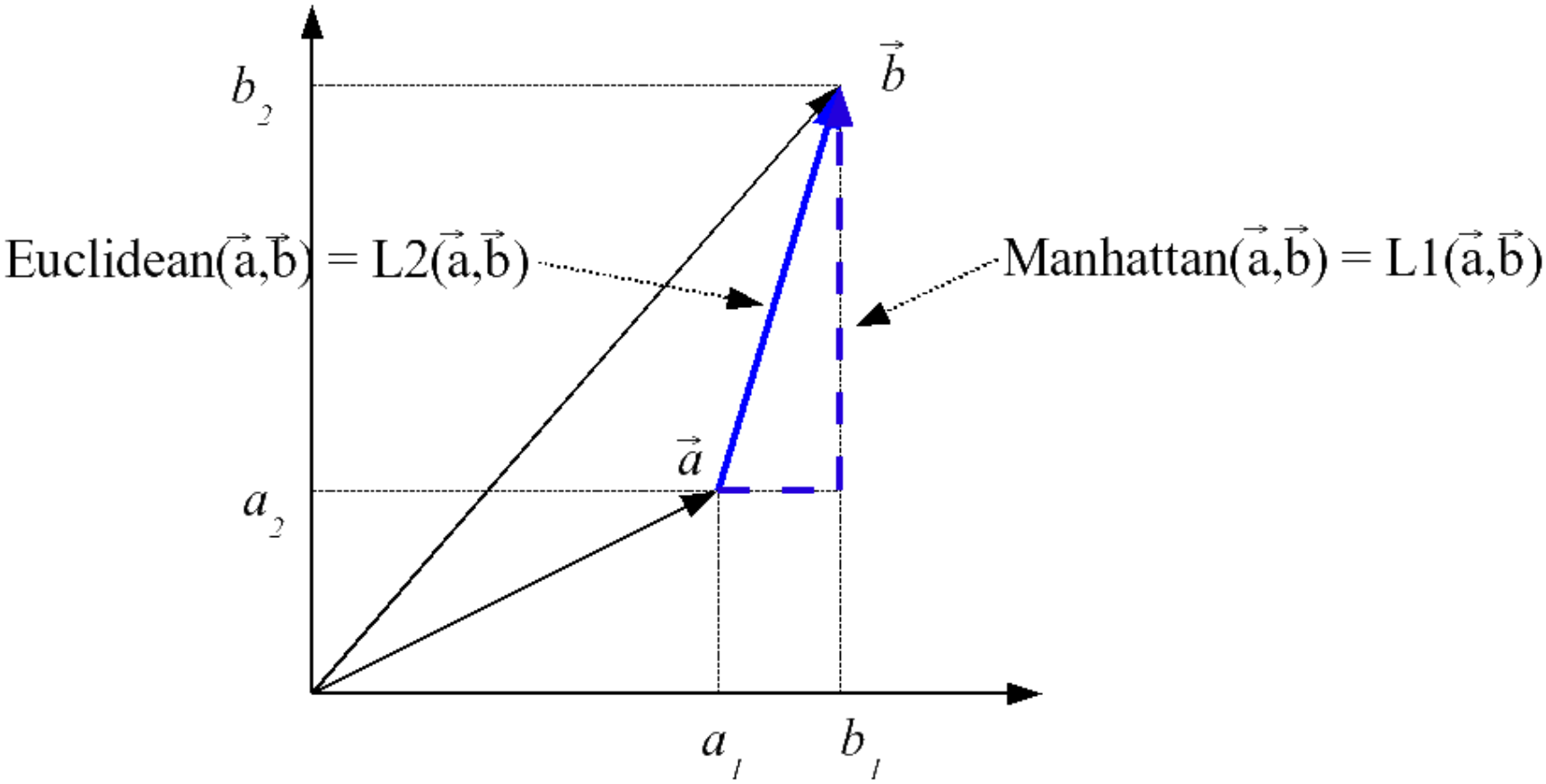
$$\text{assoc}_{\text{prob}}(w,f) = P(f|w)$$

$$\text{assoc}_{\text{PMI}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(f)}$$

$$\text{assoc}_{\text{Lin}}(w,f) = \log_2 \frac{P(w,f)}{P(w)P(r|w)P(w'|w)}$$

$$\text{assoc}_{\text{t-test}}(w,f) = \frac{P(w,f)-P(w)P(f)}{\sqrt{P(f)P(w)}}$$

3. Defining similarity between vectors



Summary of similarity measures

$$\text{sim}_{\text{cosine}}(\vec{v},\vec{w}) = \frac{\vec{v}\cdot\vec{w}}{|\vec{v}||\vec{w}|} = \frac{\sum_{i=1}^{N} v_i \times w_i}{\sqrt{\sum_{i=1}^{N} v_i^2}\sqrt{\sum_{i=1}^{N} w_i^2}}$$

$$\text{sim}_{\text{Jaccard}}(\vec{v},\vec{w}) = \frac{\sum_{i=1}^{N} \min(v_i,w_i)}{\sum_{i=1}^{N} \max(v_i,w_i)}$$

$$\text{sim}_{\text{Dice}}(\vec{v},\vec{w}) = \frac{2\times\sum_{i=1}^{N} \min(v_i,w_i)}{\sum_{i=1}^{N} (v_i+w_i)}$$

$$\text{sim}_{\text{JS}}(\vec{v}||\vec{w}) = D(\vec{v}|\frac{\vec{v}+\vec{w}}{2}) + D(\vec{w}|\frac{\vec{v}+\vec{w}}{2})$$

What about other relations?

Similarity can be used for adding new links to a thesaurus, and Lin used thesaurus induction as his motivation

But thesauruses have more structure than just similarity

In particular, hyponym/hypernym structure

Detecting hyponymy and other relations

Could we discover new hyponyms, and add them to a taxonomy under the appropriate hypernym?

Why is this important? Some examples from Rion Snow:

"**insulin**" and "**progesterone** are in WN 2.1, but "**leptin**" and "**pregnenolone**" are not.

"**combustibility**" and "**navigability**", but not "**affordability**", "**reusability**", or "**extensibility**".

"**HTML**" and "**SGML**", but not "**XML**" or "**XHTML**".

"**Google**" and "**Yahoo**", but not "**Microsoft**" or "**IBM**".

This **unknown word problem** occurs throughout NLP

Hearst Approach

Agar is a substance prepared from a mixture of red algae, such as Gelidium, for laboratory or industrial use.

What does Gelidium mean? How do you know?

$$NP_0 \; such \; as \; NP_1 \{, NP_2 \ldots, (and|or) NP_i\}, i \geq 1$$

implies the following semantics

$$\forall NP_i, i \geq 1, \text{hyponym}(NP_i, NP_0)$$

allowing us to infer

$$\text{hyponym}(\text{Gelidium}, \text{red algae})$$

Hearst's hand-built patterns

$NP\{,NP\}*\{,\}$ (and|or) other NP_H	. . . temples, treasuries, and other important civic buildings.
NP_H such as $\{NP,\}$* (or|and) NP	red algae such as Gelidium
such NP_H as $\{NP,\}$* (or|and) NP	works by such authors as Herrick, Goldsmith, and Shakespeare
NP_H $\{,\}$ including $\{NP,\}$* (or|and) NP	All common-law countries, including Canada and England
NP_H $\{,\}$ especially $\{NP,\}$* (or|and) NP	. . . most European countries, especially France, England, and Spain

Word Sense Disambiguation

- Given a *fixed* set of senses associated with a lexical item, determine which sense applies to a particular instance of the lexical item in running text.
- Two fundamental approaches
  - WSD occurs during semantic analysis as a side-effect of the elimination of ill-formed semantic representations
  - Stand-alone approach
    - » WSD is performed independent of, and prior to, compositional semantic analysis
    - » Makes minimal assumptions about what information will be available from other NLP processes
    - » Applicable in large-scale practical applications

OVERLAP BASED APPROACHES

Require a ***Machine Readable Dictionary*** *(MRD)*.

Find the overlap between the features of different senses of an ambiguous word (sense bag) and the features of the words in its context (context bag).

These features could be sense definitions, example sentences, hypernyms etc.

The features could also be given weights.

The sense which has the maximum overlap is selected as the contextually appropriate sense.

LESK'S ALGORITHM

Sense Bag: *contains the words in the definition of a candidate sense of the ambiguous word.*

Context Bag: *contains the words in the definition of each sense of each context word.*

E.g. "On burning ***coal*** we get ***ash***."

Ash

Sense 1

Trees of the olive family with pinnate leaves, thin furrowed bark and gray branches.

Sense 2

The ***solid*** residue left when ***combustible*** material is thoroughly ***burn***ed or oxidized.

Sense 3

To convert into ash

Coal

Sense 1

A piece of glowing carbon or ***burn***t wood.

Sense 2

charcoal.

Sense 3

A black ***solid combustible*** substance formed by the partial decomposition of vegetable matter without free access to air and under the influence of moisture and often increased pressure and temperature that is widely used as a fuel for ***burn***ing

In this case Sense 2 of ash would be the winner sense.

WALKER'S ALGORITHM

A Thesaurus Based approach.

Step 1: *For each sense of the target word find the thesaurus category to which that sense belongs.*

Step 2: *Calculate the score for each sense by using the context words. A context words will add 1 to the score of the sense if the thesaurus category of the word matches that of the sense.*

E.g. The money in this ***bank*** *fetches an interest of 8% per annum*

Target word: ***bank***

Clue words from the context: ***money, interest, annum, fetch***

	Sense1: Finance	Sense2: Location
Money	+1	0
Interest	+1	0
Fetch	0	0
Annum	+1	0
Total	3	0

Context words add 1 to the sense when the topic of the word matches that of the sense

WSD as classification

- Given word token in context, which sense (class) is it?
- Just train a classifier, if we have sense-labeled training data:
  - She pays 3% **interest/INTEREST-MONEY** on the loan.
  - He showed a lot of **interest/INTEREST-CURIOSITY** in the painting.
  - Playing chess is one of my **interests/INTEREST-HOBBY**.

Naïve Bayes Classification

- P(c) is the prior probability of that sense
  - Counting in a labeled training set.
- P(w|c) conditional probability of a word given a particular sense
  - P(w|c) = count(w,c)/count(c)
- We get both of these from a tagged corpus like SemCor

- Can also generalize to look at other features besides words.
  - Then it would be P(f|c)
    - Conditional probability of a feature given a sense

$$\hat{P}(c) = \frac{N_c}{N}$$

$$\hat{P}(w \mid c) = \frac{count(w,c)+1}{count(c)+|V|}$$

	Doc	Words	Class
Training	1	fish smoked fish	f
	2	fish line	f
	3	fish haul smoked	f
	4	guitar jazz line	g
Test	5	line guitar jazz jazz	?

Priors:

$P(f) = \frac{3}{4}$

$P(g) = \frac{1}{4}$

V = {fish, smoked, line, haul, guitar, jazz}

Conditional Probabilities:

P(line|*f*) = (1+1) / (8+6) = 2/14

P(guitar|*f*) = (0+1) / (8+6) = 1/14

P(jazz|*f*) = (0+1) / (8+6) = 1/14

P(line|*g*) = (1+1) / (3+6) = 2/9

P(guitar|*g*) = (1+1) / (3+6) = 2/9

P(jazz|*g*) = (1+1) / (3+6) = 2/9

Choosing a class:

$P(f|d5) \propto 3/4 * 2/14 * (1/14)^2 * 1/14$

≈ 0.00003

$P(g|d5) \propto 1/4 * 2/9 * (2/9)^2 * 2/9$

≈ 0.0006

Exemplar Based WSD (k-nn)

An exemplar based classifier is constructed for each word to be disambiguated.

Step1: From each ***sense marked sentence*** containing the ambiguous word , a training example is constructed using:

- POS of ***w*** as well as POS of neighboring words.
- Local collocations
- Co-occurrence vector
- Morphological features
- Subject-verb syntactic dependencies

Step2: Given a test sentence containing the ambiguous word, a test example is similarly constructed.

Step3: The test example is then compared to all training examples and the k-closest training examples are selected.

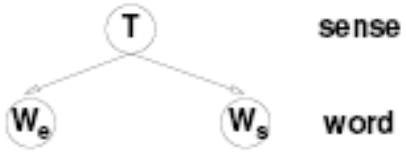
Step4: The sense which is most prevalent amongst these "k" examples is then selected as the correct sense.

WSD Using Parallel Corpora

A word having multiple senses in one language will have distinct translations in another language, based on the context in which it is used.

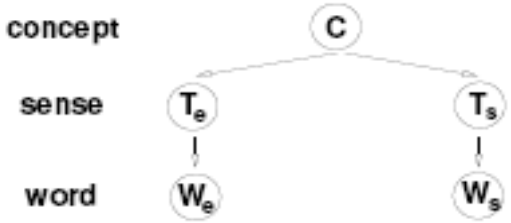
The translations can thus be considered as contextual indicators of the sense of the word

Sense Model



$$P(W_e, W_s, T) = P(T).P(W_e|T).P(W_s|T) \quad (5.1)$$

Concept Model



$$P(W_e, W_s, T_s, T_e, C) = P(C).P(T_e|C).P(T_s|C).P(W_e|T_e).P(W_s|T_s) \quad (5.2)$$

Thematic Roles

Thematic relations were introduced in generative grammar during the mid-1960s and early 1970s (Gruber, 1976; Fillmore, 1968; Jackendoff, 1972) as a way of classifying the arguments of natural language predicates into a closed set of participant types which were thought to have a special status in grammar.

Thematic roles or relations

- Thematic relations is a term used to express the role that a noun phrase plays with respect to the action or state described by a sentence's verb.
- Thematic relations concern the nature of the relationship between the **meaning** of the verb and the **meaning** of the noun.
- They serve to capture the relationship between syntax and semantics with respect to predicates and their arguments

Thematic roles or relations

- Thematic roles, also called thematic relations are characterizations of certain semantic relationships which hold between a verb and its complements (and adjuncts).
- Allow to represent the semantic correspondence between (uses of) relational concepts in a systematic way – thereby supporting basic lexical-semantic inference.
- Support a systematic representation of the mapping between syntactic complements and semantic argument positions (role-linking).

List of Basic Thematic Roles

- Agent/Actor
- Theme/ Patient/ Object
- Recipient
- Instrument
- Source / From-Location
- Goal/ To-Location
- Beneficiary
- Experiencer
- Force / Natural Cause
- Cause
- Co-agent
- Co-theme

Agent

- The initiator of some action, capable of acting with wish
- Deliberately performs the action / as subject of active sentences
  - **The boys** caught some fish.
  - **My mother** wrote me a letter
  - **Jack** ate the beans.

Experiencer

- The entity which is aware of the action or state described by the predicate but which is not in control of the action or state.
- The entity that receives sensory or emotional input
  - **Seeta** Saw the deer.
  - **Susan** heard the song**.**
  - **She** likes flowers**.**

Theme

- The entity which is moved by an action, or whose location is described.
- Undergoes the action but does not change its state
- The thing that is said to move or remain at rest
  - I put **the book** on the table.
  - He gave **the gun** to the police officer
  - Fred threw **the rock**

Note: Sometimes used interchangeably with patient

Patient

- The entity undergoing the effect of some action, often undergoing some change of state.
  - The falling rocks crushed **the car**.
  - **The ice** melted.
  - Sam chewed **the apple**.
  - Sue mowed **the lawn**

Note: Sometimes used interchangeably with theme

Instrument

- Means by which an action is performed or something comes about.
- Used to carry out the action(preposition *with*).
  - Jamie cut the ribbon **with a pair of scissors**.
  - Jane ran **with a crutch** for Susan.
  - They opened the door **with the key**.
  - The squirrel bit the acorn **with its teeth**.

Force or Natural Cause

- Mindlessly performs the action
  - An avalanche destroyed the ancient temple.
  - Hurricane Katrina destroyed New Orleans

Location

- The place in which something is situated or takes place
- Where the action occurs / prepositions *in*, *on*, *beyond*
  - Johnny and Linda played carelessly **in the park**.
  - I'll be **at Julie's house** studying for my test.
  - I am **in the closest community hall**
  - The picture hangs **above the fireplace**.

Source / Origin

- The entity from which something moves, either literally / metaphorically
- Where the action originated (preposition from)

  - The rocket was launched **from Central Command**.
  - She walked **away from him**.
  - The butterfly began its migration **from Mexico.**
  - Sue ran **from the policeman**

Direction or Goal

- The entity towards which something moves, either literally or metaphorically.
- where the action is directed towards (preposition to)

  - The caravan continued on **toward the distant oasis**.
  - He walked **to school**.

Recipient

- A subtype of goal involved in actions describing changes of possession
- A special kind of goal associated with verbs expressing a change in ownership, possession.
  - I sent **John** the letter.
  - He gave the book **to her**.
  - Bill sold the car **to Mary**

Beneficiary

- The entity for whose benefit the action was performed (preposition *for)*
  - He bought a car **for me**.
  - I fought **for the king**.
  - I gave the book to Jack **for Susan**.
  - Mary studied hard for **her mother**

Time , Manner

Time

- The time at which the action occurs
  - The rocket was launched **yesterday**.

Manner

- The way in which an action is carried out
  - **With great urgency**, Tabitha phoned 108.
  - **With great speed**, the cat ran across the yard to escape the dog.

Purpose, Cause

Purpose

The reason for which an action is performed

- Tabitha phoned 911 right away **in order to get some help**.

Cause

What caused the action to occur in the first place; not *for what*, rather *because of what*

- **Because Clyde was hungry**, he ate the cake.

Co-agent

In some sentences he acts with agent to perform actions

- Henry lifted the piano **with Jack**.

Note:
There are no clear boundaries between these relations.
For example, in "the hammer broke the window", some linguists treat ***hammer*** as an agent, some others as instrument

Example

They filled the pool with water. (Agent + Theme + Location)

Jack ran. (Agent only)

Jack ran with crutch. (Agent + INSTR)

Jack ran with crutch for Susan. (Agent + INSTR +Beneficiary)

Jack destroyed the car. (Agent + Theme)

Jack put the car through the wall. (Agent + Theme + Path)

Go through these links for more

http://www.ling.upenn.edu/~beatrice/syntax-textbook/box-thematic.html

http://www.linguisticsnetwork.com/semantics-thematic-roles/