Intelligent Systems (AI-2)

Computer Science cpsc422, Lecture 24

Nov 7, 2016

Slide credit: Satanjeev Banerjee Ted Pedersen 2003, Jurfsky & Martin 2008

Lecture Overview

- Semantic Similarity/Distance
- Concepts: Thesaurus/Ontology Methods
- Words: Distributional Methods

Why words/concepts similarity is important ?

"**fast**" is similar to "**rapid**"

"**tall**" is similar to "**height**"

Question answering:

*Q: "How **tall** is Mt. Everest?"*
*Candidate A: "The official **height** of Mount Everest is 29029 feet"*

- Extends to sentence/paragraph similarity
- **Summarization**: identify and eliminate redundancy, aggregate similar phrase/sentences
- ....

WordNet: entry for "table"

gloss

The **noun** "table" has 6 senses in WordNet.

1. table, tabular array -- (a set of data …)
2. table -- (a piece of furniture …)
3. table -- (a piece of furniture with tableware…)
4. mesa, table -- (flat tableland …)
5. table -- (a company of people …)
6. board, table -- (food or meals …)

The **verb** "table" has 1 sense in WordNet.

1. postpone, prorogue, hold over, put over, **table**, shelve, set back, defer, remit, put off - (hold back to a later time; "let's postpone the exam")

WordNet Relations (between synsets!)

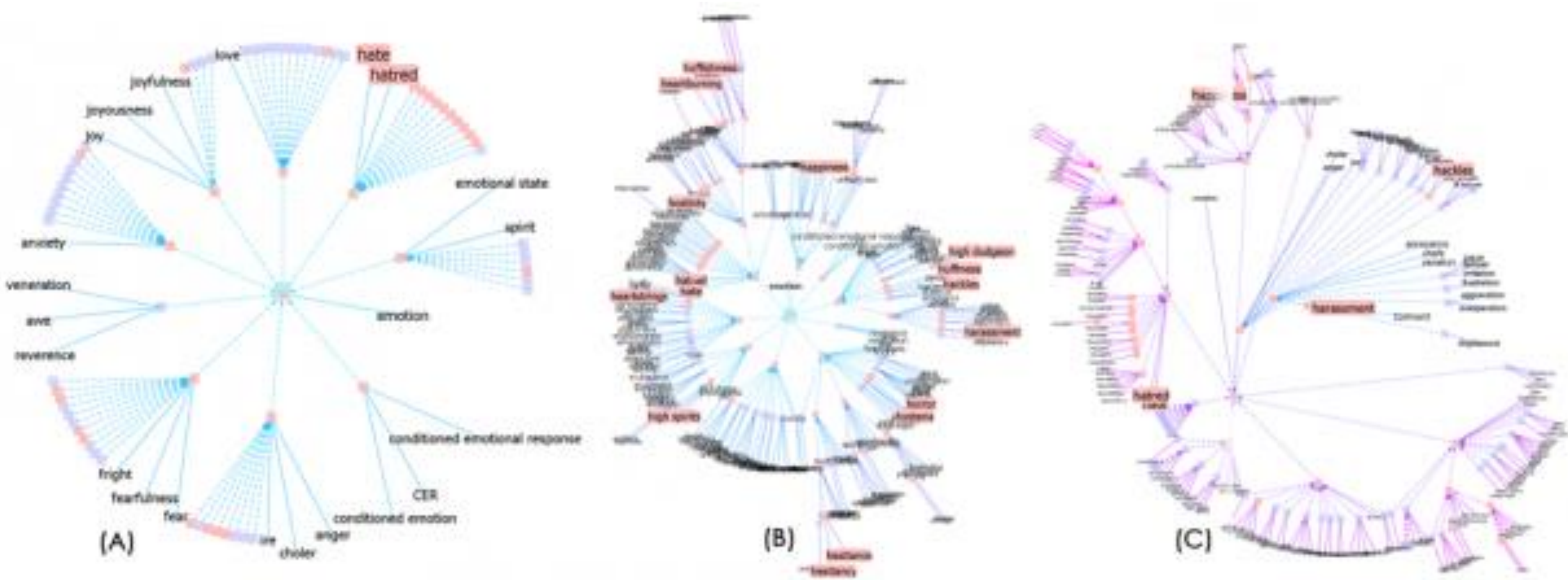
Nouns

Relation	Definition	Example
Hypernym	From concepts to superordinates	*breakfast* $\rightarrow$ *meal*
Hyponym	From concepts to subtypes	~~*meal* $\rightarrow$~~ *lunch*
Has-Member	From groups to their members	*faculty* $\rightarrow$ *professor*
Member-Of	From members to their groups	*copilot* $\rightarrow$ *crew*
Has-Part	From wholes to parts	*table* $\rightarrow$ *leg*
Part-Of	From parts to wholes	*course* $\rightarrow$ *meal*
Antonym	Opposites	*leader* $\rightarrow$ *follower*

Verbs

Relation	Definition	Example
Hypernym	From events to superordinate events	*fly* $\rightarrow$ *travel*
Troponym	From events to their subtypes	*walk* $\rightarrow$ *stroll*
Entails	From events to the events they entail	*snore* $\rightarrow$ *sleep*
Antonym	Opposites	*increase* $\Longleftrightarrow$ *decrease*

Visualizing Wordnet Relations



C. Collins, “WordNet Explorer: Applying visualization principles to lexical semantics,” University of Toronto, Technical Report kmdi 2007-2, 2007.

Semantic Similarity/Distance: example

(n) table -- (a piece of furniture having a smooth flat top that is usually supported by one or more vertical legs)

(n) mesa, table -- (flat tableland with steep edges)

sim

(n) hill (a local and well-defined elevation of the land)

(n) lamp (a piece of furniture holding one or more electric light bulbs)

dissimilar

Semantic Similarity/Distance

Between two concepts in an ontology, e.g., between two senses in Wordnet

What would you use to compute it ?

iclicker.

A. The distance between the two concepts in the underlying hierarchies / graphs

B. The glosses of the concepts

C. None of the above

D. Both of the above

Gloss Overlaps ≈ Relatedness

concepts

- Lesk's (1986) idea: Related word senses are (often) defined *using the same words.* E.g:
  - bank(1): "a financial institution"
  - bank(2): "sloping land beside a body of water"
  - lake: "a body of water surrounded by land"

Gloss Overlaps ≈ Relatedness

- Lesk's (1986) idea: Related word senses are (often) defined *using the same words.* E.g:
  - bank(1): "a financial institution"
  - bank(2): "sloping land beside a body of water"
  - lake: "a body of water surrounded by land"

Gloss Overlaps ≈ Relatedness

- Lesk's (1986) idea: Related word senses are (often) defined *using the same words.* E.g:
  - bank(1): "a financial institution"
  - bank(2): "sloping land beside a body of water"
  - lake: "a body of water surrounded by land"
- Gloss overlaps = # content words common to two glosses ≈ relatedness
  - Thus, relatedness (bank(2), lake) = 3
  - And, relatedness (bank(1), lake) = 0

Limitations of (Lesk's) Gloss Overlaps

- Most glosses are very short.
  - So not enough words to find overlaps with.

- Solution?

Extended gloss overlaps
  - Add glosses of synsets connected to the input synsets.

Extending a Gloss

sentence: “the penalty meted out to one adjudged guilty”

bench: “persons who hear cases in a court of law”

overlapped words = 0

Extending a Gloss

final judgment: “a judgment disposing of the case before the court of law”

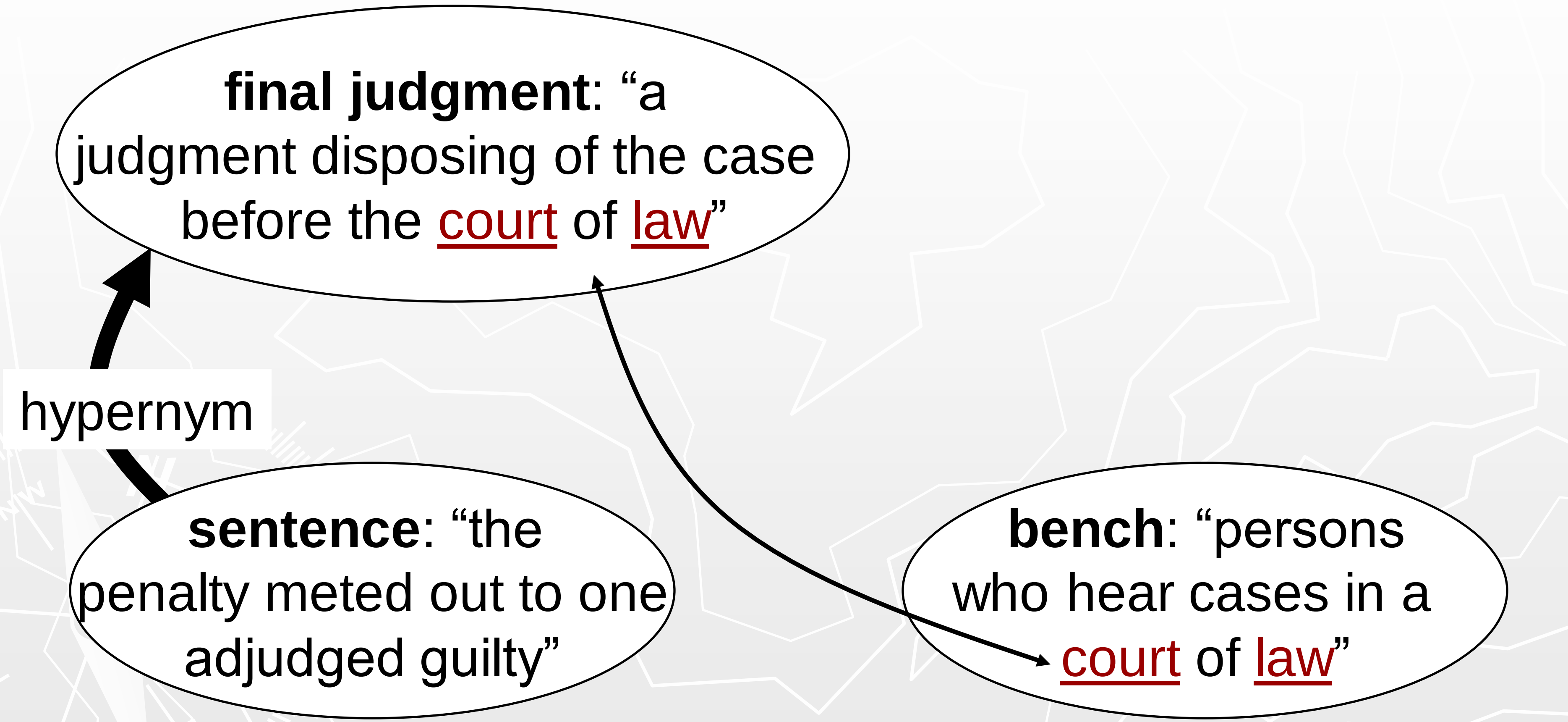
hypernym

sentence: “the penalty meted out to one adjudged guilty”

bench: “persons who hear cases in a court of law”

overlapped words = 0

Extending a Gloss



\# overlapped words = 2

Creating the Extended Gloss Overlap Measure

- How to measure overlaps?
- Which relations to use for gloss extension?

How to Score Overlaps?

- Lesk simply summed up overlapped words.
- But matches involving phrases – phrasal matches – are rarer, and more informative
  - E.g. "court of law" "body of water"
- Aim: Score of *n* words in a phrase > sum of scores of *n* words in shorter phrases
- Solution: Give a phrase of *n* words a score of n^2
  - "court of law" gets score of 9.

overlap of
9+1=10

- bank(2): "sloping land beside a body of water"
- lake: "a body of water surrounded by land"

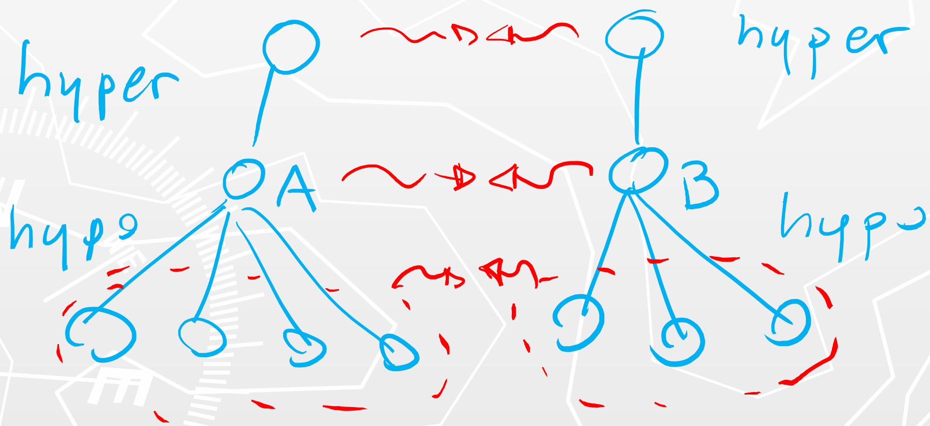
Which Relations to Use?

Typically include…

- Hypernyms [“car” → “vehicle”]
- Hyponyms [“car” → “convertible”]
- Meronyms [“car” → “accelerator”]
- Holonym [“car” → “train”]
- …

Extended Gloss Overlap Measure

- Input two synsets A and B
- Find phrasal gloss overlaps between A and B
- For *each relation*, compute phrasal gloss overlaps between every synset connected to A, and every synset connected to B



compute phrasal
score overlap

- Add phrasal scores to get relatedness of A and B

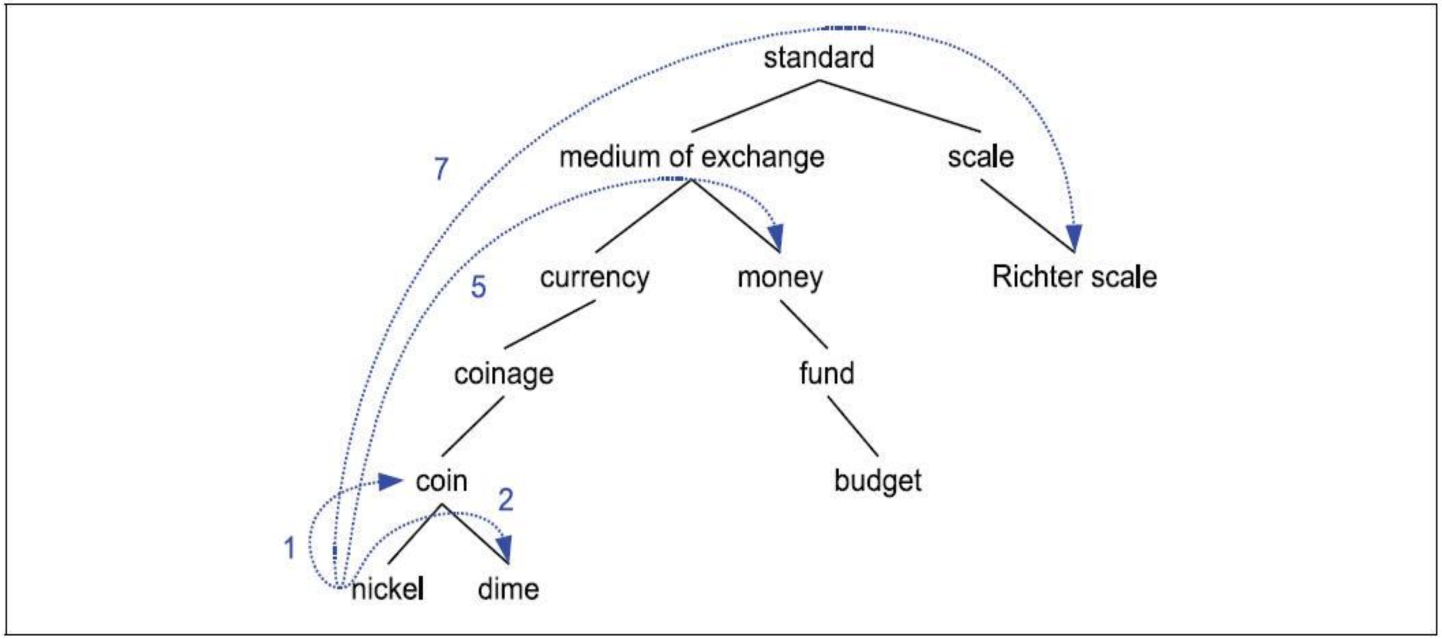
A and B can be from different parts of speech!

Distance: Path-length

Path-length sim based on is-a/hypernyms hierarchies

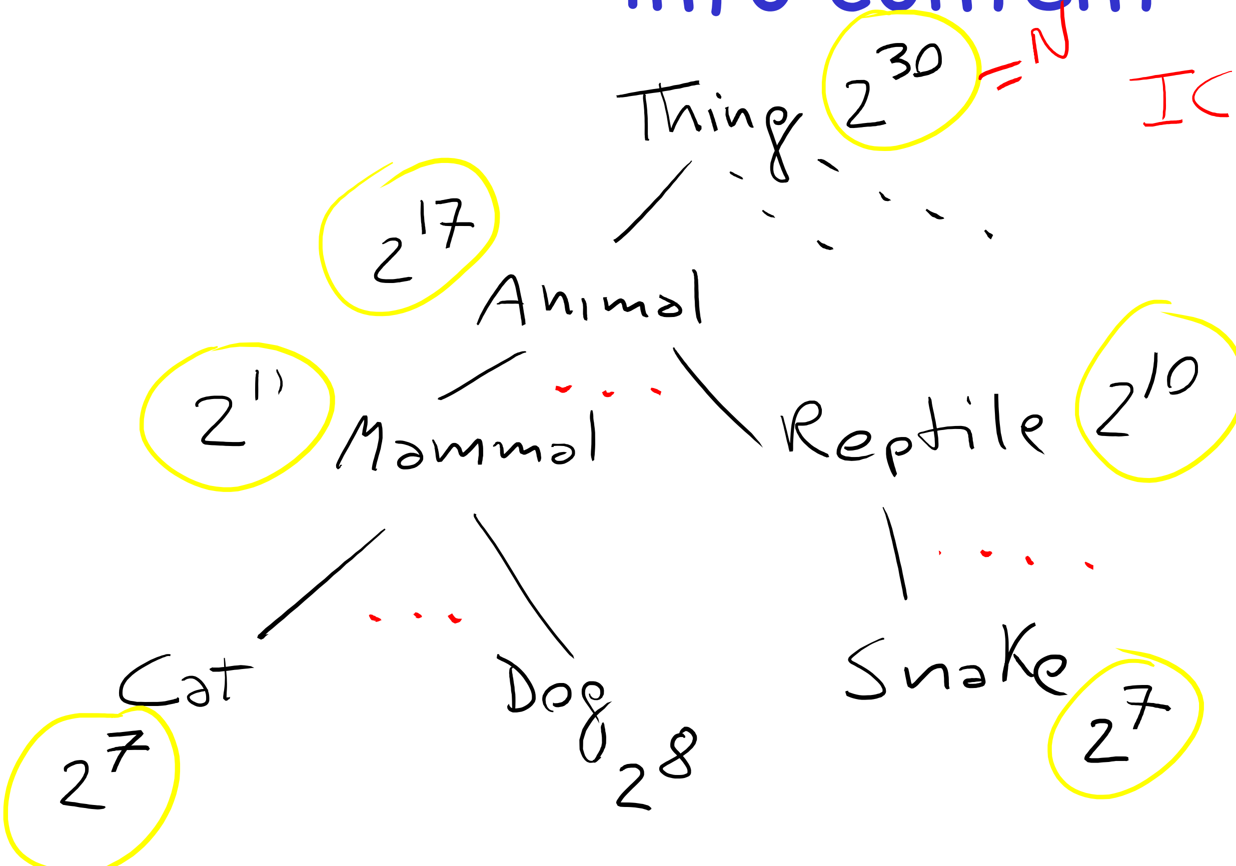
$$\boxed{\mathrm{sim}_{\mathrm{path}}(c_1, c_2) = 1/\ pathlen(c_1, c_2)}$$

c_1, c_2 are senses



But this is assuming that all the links are the same…. Encode the same semantic distance….

Probability of a concept/sense and its info content



IC(Thing) = 0

IC(Snake) = 23

probability

$$P(c) = \frac{count(c)}{N}$$

count of all Things

$$IC(c) = -\log P(c)$$

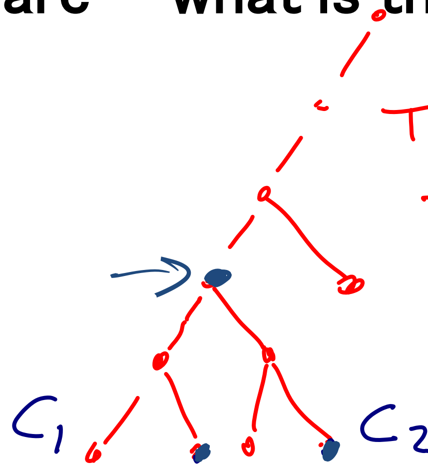
Information Content

Concept Distance: info content

- Similarity should be proportional to the information that the two concepts share… what is that?

$P(\text{root}) = 1$

The lower the concept/sense the lower its prob.



probability

$$P(c) = \frac{\sum_{c_i \in subsenses(c)} count(c_i)}{N}$$

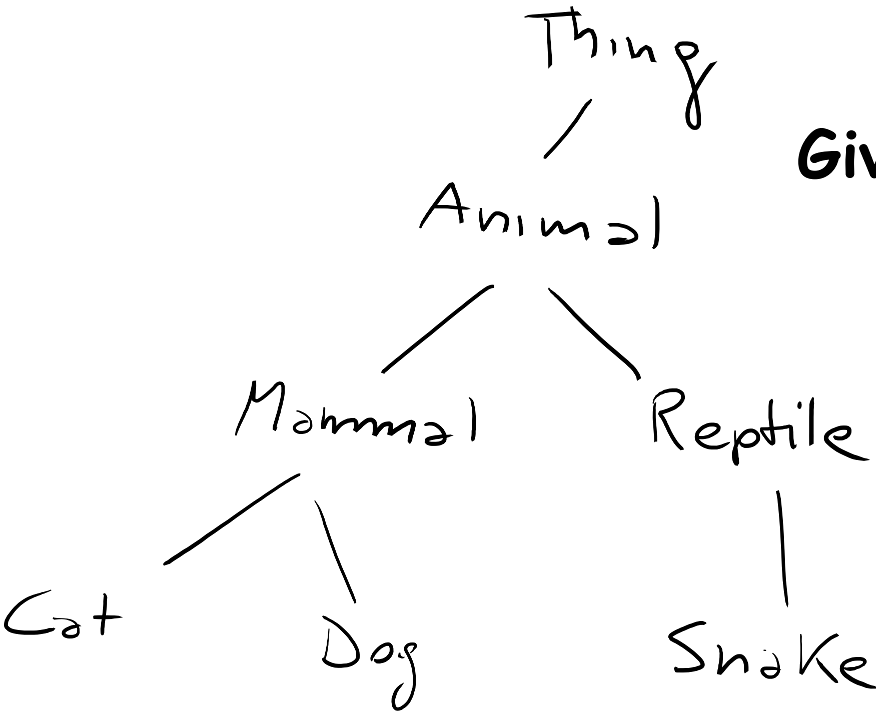
$$IC(c) = -\log P(c)$$

Information Content

$$LCS(c_1, c_2)$$

Lowest Common Subsumer

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



Given this measure of similarity

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

Are these two the same?

$\mathrm{sim}_{\mathrm{resnik}}(Dog, Snake)$ $\mathrm{sim}_{\mathrm{resnik}}(Mammal, \mathrm{Reptile})$

A. Yes B. No C. Cannot tell

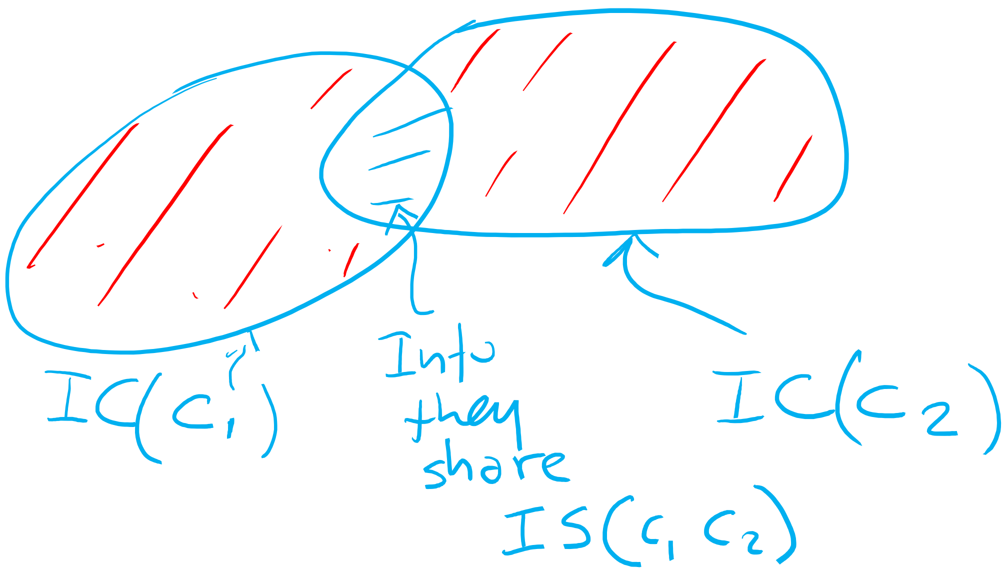
Is this reasonable?

A. Yes B. No C. Cannot tell

as well ... we can consider better alternatives ...

Concept Distance: info content

- One of best performers – Jiang-Conrath distance
- How much information the two DO NOT share ?



$$\left(IC(c_1) - IS(c_1, c_2)\right) + \left(IC(c_2) - IS(c_1, c_2)\right)$$

$$IC(c_1) + IC(c_2) - 2 * IS(c_1, c_2)$$

Concept Distance: info content

- One of best performers – Jiang-Conrath distance
- How much information the two DO NOT share

$$\boxed{dist_{JC}(c_1,c_2) = ((-\log P(c_1)) + (-\log P(c_2))) - (2\times -\log P(LCS(c_1,c_2)))}$$

$$\boxed{dist_{JC}(c_1,c_2) = 2\times\log P(LCS(c_1,c_2)) - (\log P(c_1) + \log P(c_2))}$$

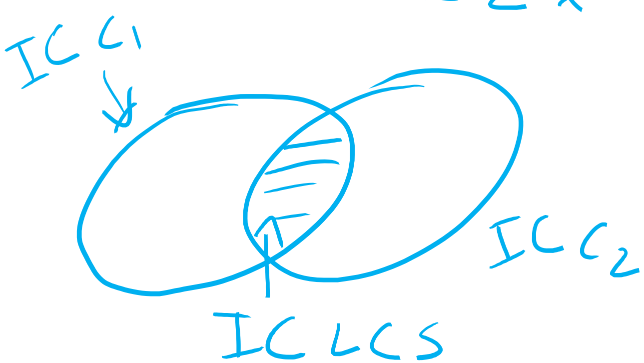
- This is a measure of distance. Reciprocal for similarity! $1/dist_{JC}$
- Problem for measures working on hierarchies/graphs: only compare concepts associated with words of part part-of speech (typically nouns) one

Concept Distance: info content

- One of best performers – Jiang-Conrath distance
- How much information the two DO NOT share

$$\boxed{\mathrm{dist}_{JC}(c_1, c_2) = 2 \times \log P(LCS(c_1, c_2)) - (\log P(c_1) + \log P(c_2))}$$

$$= 2 \times -(-\log P(LCS(c_1, c_2))) + (-\log P(c_1) - \log P(c_2))$$



Info Content they share

Info content of c_1

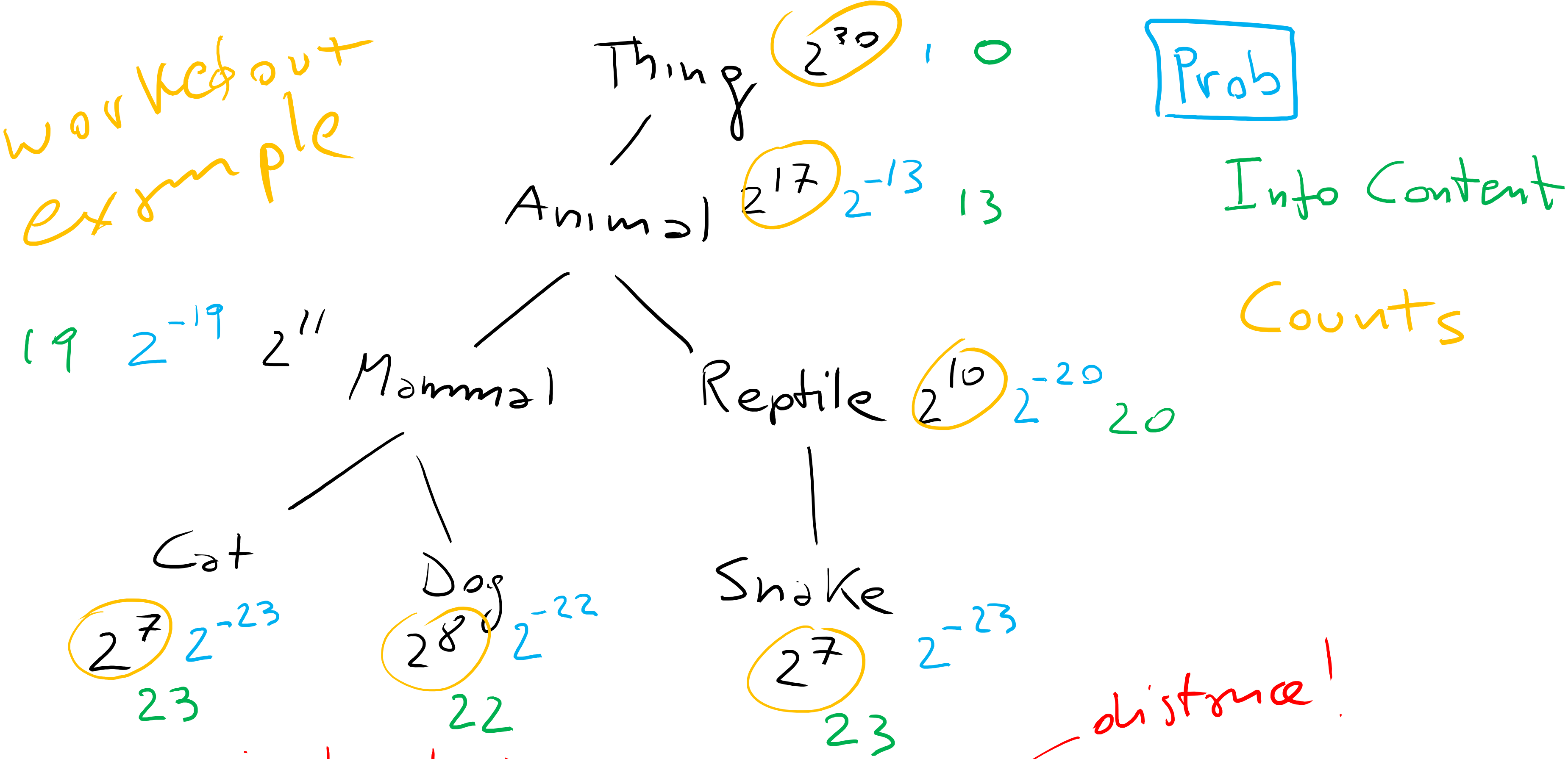
Info content of c_2

- This is a measure of distance. Reciprocal for similarity!

$1/\mathrm{dist}_{JC}$

- Problem for measures working on hierarchies/graphs: only compare concepts associated with words of part part-of speech (typically nouns)

one



similarity!

$$sim_{res}(Dog, Snake) = 13$$

$$sim_{res}(Mammal, Reptile) = 13$$

$$dist_{JC}(Dog, Snake) = (2 \times -13) + (24 + 23) = 21$$

$$dist_{JC}(Mammal, Reptile) = (2 \times -13) + (19 + 20) = 13$$

Lecture Overview

- Semantic Similarity/Distance
- Concepts: Thesaurus/Ontology Methods
- **Words: Distributional Methods – Word Similarity (WS)**

Word Similarity: Distributional Methods

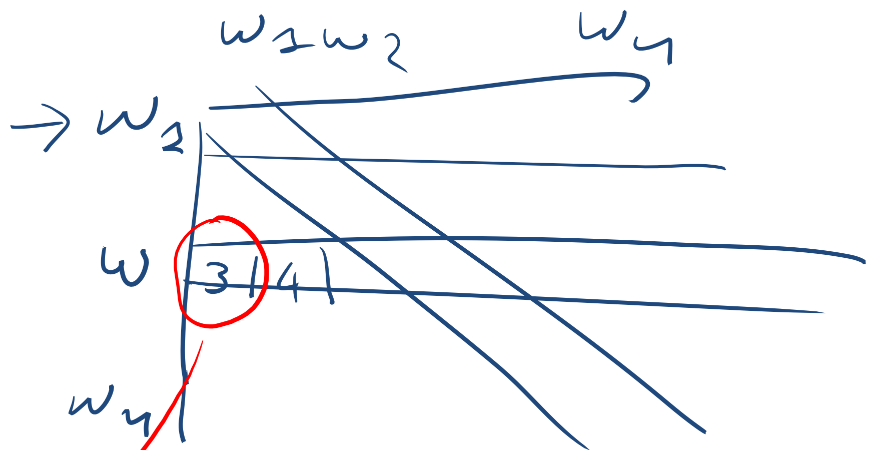
- Do not have any thesauri/ontologies for target language (e.g., Russian)
- If you have thesaurus/ontology, still
  - Missing domain-specific (e.g., technical words)
  - Poor hyponym knowledge (for V) and nothing for Adj and Adv
  - Difficult to compare senses from different hierarchies (although extended Lesk can do this)
  - **Solution:** extract similarity from corpora
  - **Basic idea:** two words are similar if they appear in similar contexts

WS Distributional Methods (1)

- **Context: feature vector**

Stop list

$$\vec{w} = (f_1, f_2, \ldots, f_N)$$



Example: f_i how many times w_i appeared in the neighborhood of w

eg w and w1 appeared 3 times in the same sentence

WS Distributional Methods (2)

- More informative values (referred to as weights or measure of association in the literature)
  - Point-wise Mutual Information

P(w) P(wi)

$$assoc_{PMI}(w, w_i) = \log_2 \frac{P(w, w_i)}{P(w)P(w_i)}$$

- t-test

$$assoct_{t-test}(w, w_i) = \frac{P(w, w_i) - P(w)P(w_i)}{\sqrt{P(w)P(w_i)}}$$

$$\sqrt{\frac{s^2}{N}}$$

PMI example

$$assoc_{PMI}(w, w_i) = \log_2 \frac{P(w, w_i)}{P(w)P(w_i)}$$

Assume w, w_i appear with equal frequency $\frac{1}{2^{10}}$

in a large set of documents

$P(w) = 2^{-10}$

$P(w_i) = 2^{-10}$

$P(w, w_i) =$

A $2^{-10} * 2^{-10} = 2^{-20}$ if the words are completely independent

B 2^{-10} if the words appear always together

A $assoc_{PMI} = \log_2 \frac{2^{-20}}{2^{-10} * 2^{-10}} = \log_2 1 = 0$

B $assoc_{PMI} = \log_2 \frac{2^{-10}}{2^{-10} * 2^{-10}} = \log_2 2^{10} = 10$

Other popular vector representations

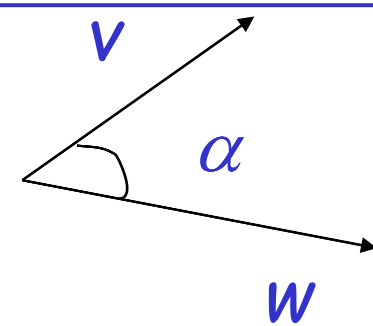
Dense vector representations (less dimensions):

1. **Singular value decomposition applied to word-word PointWise-MI matrix**
2. **Neural-Network-inspired models (skip-grams, CBOW)**

WS Distributional Methods (3)

- **Similarity between vectors**

$$sim_{cosine}(\vec{v},\vec{w}) = \frac{\vec{v}}{\|\vec{v}\|} \bullet \frac{\vec{w}}{\|\vec{w}\|} = \frac{\vec{v} \bullet \vec{w}}{\|\vec{v}\| \times \|\vec{w}\|} = \cos(\alpha)$$



Not sensitive to extreme values

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

Normalized (weighted) number of overlapping features

e.g.

$\vec{v}$ 2 1 0 3

$\vec{w}$ 3 1 1 2

$$\frac{2+1+0+2}{3+1+1+3} = \frac{5}{8}$$

Learning Goals for today's class

You can:

- Describe and Justify metrics to compute the similarity/distance of two concepts in an ontology
- Describe and Justify distributional metrics to compute the similarity/distance of two words (or phrases) in a Natural Language

Assignment-3 out – due Nov 21
(8-18 hours – working in pairs on programming parts is strongly advised)

Next class Wed

- Natural language Processing: Context free grammars and parsing

Sim/Distance: from concepts to words

- If we do not have

Word Sense Disambiguation

$$\text{wordsim}(w_1, w_2) = \max_{\substack{c_1 \in \text{senses}(w_1) \\ c_2 \in \text{senses}(w_2)}} \text{sim}(c_1, c_2)$$

WordSim: Thesaurus Methods(Extended Lesk)

- For each n-word phrase that occurs in both glosses, Extended Lesk adds in a score n^2

$$sim_{eLesk}(c_1, c_2) = \sum_{r,q \in RELS} \underline{overlap}\left(gloss(r(c_1)), gloss(q(c_2)) \right)$$

set of possible Wordnet relations whose glosses we compare

Semantic Similarity/Distance

Between two concepts in an ontology, e.g., between two senses in Wordnet

What would you use

- **Thesaurus methods**: measure distance in online thesauri (e.g., Wordnet)
- **Distributional methods**: finding if the two words appear in similar contexts

WDS: Dictionary and Thesaurus Methods

Most common: Lesk method

- **Choose the sense whose *dictionary gloss* shares most words with the target word's neighborhood**
- **Exclude stop-words**

~~We~~ move ~~the~~ table ~~to~~ . . .

Def: Words in gloss for a sense is called the *signature*

Lesk: Example

Two SENSES for channel

S1: (n) **channel** (a passage for water (or other fluids) to flow through) *"the fields were crossed with irrigation channels"; "gutters carried off the rain water into a series of channels under the street"*

S2: (n) **channel, television channel, TV channel** (a television station and its programs) *"a satellite TV channel"; "surfing through the channels"; "they offer more than one hundred channels"*

…..

" most streets closed to the TV station were flooded because the main **channel** was clogged by heavy rain ."

Corpus Lesk

Best performer

- **If a corpus with annotated senses is available**
- **For each sense: add to the signature for that sense, words "that frequently appear" in the sentences containing that sense**

CORPUS

……

"most streets closed to the TV station were flooded because the main <S1> channel </S1> was clogged by heavy rain.

…..

Word Similarity/Semantic Distance

Actually relation between two *senses*

sun vs. moon – mouth vs. food – hot vs. cold

Applications?

- **Thesaurus methods: measure distance in online thesauri (e.g., Wordnet)**
- **Distributional methods: finding if the two words appear in similar contexts**

WS: Thesaurus Methods(1)

- **Path-length based sim on hyper/hypo hierarchies**

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

- **Information content word similarity (not all edges are equal)**

probability

$$\boxed{\mathrm{P}(c) = \frac{\sum_{c_i \in subsenses(c)} count(c_i)}{N}}$$

$$\boxed{\mathrm{IC}(c) = -\log P(c)}$$

Information

$$\boxed{LCS(c_1, c_2)}$$

Lowest Common Subsumer

$$\boxed{\mathrm{sim}_{\mathrm{resnik}}(c_1, c_2) = -\log P(LCS(c_1, c_2))}$$

Ontologies

Given a logical representation (e.g., FOL)

What individuals and relations are there and we need to model?

In AI an **Ontology** is a specification of what individuals and relationships are assumed to exist and what terminology is used for them

- What **types** of individuals
- What **properties** of the individuals

Ontologies: inspiration from Natural Language
:

How do we refer to individuals and relationship in the world in NL e.g., English?

words

Where do we find definitions for words?

Dictionary

Most of the definitions are circular? They are descriptions.

Red / Blood

Fortunately, there is still some useful semantic info (Lexical Relations):

plant
bat

w_1 w_2 same Form and Sound, different Meaning Homonymy

w_1 w_2 same Meaning, different Form Synonymy big/large

w_1 w_2 "opposite" Meaning Antonymy good/bad

w_1 w_2 Meaning_1 subclass of Meaning_2 Hyponymy dog/animal

Polysemy

Def. The case where we have a set of words with the same form and multiple related meanings.

Consider the homonym:

bank → commercial **bank**$_1$ vs. river **bank**$_2$

- **Now consider:** "*A PCFG can be trained using derivation trees from a tree bank annotated by human experts*"
  - **Is this a new independent sense of bank?**

Synonyms

Def. Different words with the same meaning.

Substitutability- if they can be substituted for one another in *some* environment without changing meaning or acceptability.

Would I be flying on a large/big plane?

?... became kind of a large/big sister to...

? You made a large/big mistake

Hyponymy

Def. Pairings where one word denotes a subclass of the other

- Since dogs are canids
  - ✓Dog is a *hyponym* of canid and
  - ✓Canid is a *hypernym* of dog

car/vehicle

doctor/human

……

Lexical Resources

Databases containing all lexical relations among all words

- **Development:**
  - **Mining info from dictionaries and thesauri**
  - **Handcrafting it from scratch**

- **WordNet: fist developed with reasonable coverage and widely used, started with [Fellbaum… 1998]**
  - **for English (versions for other languages have been developed – see MultiWordNet)**

WordNet 3.0

POS	Unique Strings	Synsets	Word-Sense Pairs
Noun	117798	82115	146312
Verb	11529	13767	25047
Adjective	21479	18156	30002
Adverb	4481	3621	5580
Totals	155287	117659	206941

- **For each word: all possible senses** (no distinction between homonymy and polysemy)
- **For each sense:** a set of synonyms (synset) and a gloss

WordNet: entry for "table"

gloss

The noun "table" has 6 senses in WordNet.

1. table, tabular array -- (a set of data …)
2. table -- (a piece of furniture …)
3. table -- (a piece of furniture with tableware…)
4. mesa, table -- (flat tableland …)
5. table -- (a company of people …)
6. board, table -- (food or meals …)

The **verb** "table" has 1 sense in WordNet.

1. postpone, prorogue, hold over, put over, **table**, shelve, set back, defer, remit, put off – (hold back to a later time; "let's postpone the exam")

WordNet Relations (between synsets!)

N

Relation	Definition	Example
Hypernym	From concepts to superordinates	*breakfast* → *meal*
Hyponym	From concepts to subtypes	*meal* → *lunch*
Has-Member	From groups to their members	*faculty* → *professor*
Member-Of	From members to their groups	*copilot* → *crew*
Has-Part	From wholes to parts	*table* → *leg*
Part-Of	From parts to wholes	*course* → *meal*
Antonym	Opposites	*leader* → *follower*

V

Relation	Definition	Example
Hypernym	From events to superordinate events	*fly* → *travel*
Troponym	From events to their subtypes	*walk* → *stroll*
Entails	From events to the events they entail	*snore* → *sleep*
Antonym	Opposites	*increase* ⟺ *decrease*

WordNet Hierarchies: "Vancouver"

WordNet: example from ver1.7.1

For the three senses of "Vancouver"

⇒(city, metropolis, urban center)
    ⇒ (municipality)
        ⇒ (urban area)
            ⇒ (geographical area)
                ⇒ (region)
                    ⇒ (location)
                        ⇒ (entity, physical thing)

    ⇒ (administrative district, territorial division)
        ⇒ (district, territory)
            ⇒ (region)
                ⇒ (location
                    ⇒ (entity, physical thing)

⇒ (port)
    ⇒ (geographic point)
        ⇒ (point)
            ⇒ (location)
                ⇒ (entity, physical thing)

Wordnet: NLP Tasks

- First success in Probabilistic Parsing (PP-attachments): words + word-classes extracted from the hypernym hierarchy increase accuracy from 84% to 88% [Stetina and Nagao, 1997]
- Word sense disambiguation
- Lexical Chains (summarization)

- …… and *many others* !

More importantly starting point for larger Ontologies!

More ideas from NLP….

Relations among words and their meanings (paradigmatic)

Internal structure of individual words (syntagmatic)

Predicate-Argument Structure

- **Represent relationships among concepts, events and their participants**

"I ate a turkey sandwich for lunch"

$\exists$ *w: Isa(w,Eating)* $\wedge$ *Eater(w,Speaker)* $\wedge$ *Eaten(w,TurkeySandwich)* $\wedge$ *MealEaten(w,Lunch)*

"Nam does not serve meat"

$\exists$ *w: Isa(w,Serving)* $\wedge$ *Server(w, Nam)* $\wedge$ $\neg$*Served(w,Meat)*

Semantic Roles: Resources

- **Move beyond inferences about single verbs**

  " IBM hired John as a CEO "

  " John is the new IBM hire "

  " IBM signed John for 2M$"

- **FrameNet:** Databases containing **frames** and their syntactic and semantic argument structures
- (book online Version 1.5-update Sept, 2010)
  - for English (versions for other languages are under development)

FrameNet Entry

Hiring

- **Definition:** An Employer hires an Employee, promising the Employee a certain Compensation in exchange for the performance of a job. The job may be described either in terms of a Task or a Position in a Field.
- **Inherits From:** **Intentionally affect**
- **Lexical Units:** ***commission.n, commission.v, give job.v, hire.n, hire.v, retain.v, sign.v, take on.v***

FrameNet Annotations

Some roles..
Employer Employee Task Position

- **np-vpto**
  - **In 1979 , singer Nancy Wilson HIRED him to open her nightclub act .**
  - ….
- **np-ppas**
  - **Castro has swallowed his doubts and HIRED Valenzuela as a cook in his small restaurant .**

Includes counting: How many times a role was expressed with a particular syntactic structure…

Lecture Overview

- **Ontologies** – what objects/individuals should we represent? what relations (unary, binary,..)?
- Inspiration from **Natural Language**: WordNet and FrameNet
- Extensions based on Wikipedia and mining the Web (YAGO, ProBase, Freebase)
- Domain Specific Ontologies (e.g., Medicine: MeSH, UMLS)

YAGO2: huge semantic knowledge base

Derived from **Wikipedia**, **WordNet** and **GeoNames**. (started in 2007, paper in www conference)

10^6 entities (persons, organizations, cities, etc.)

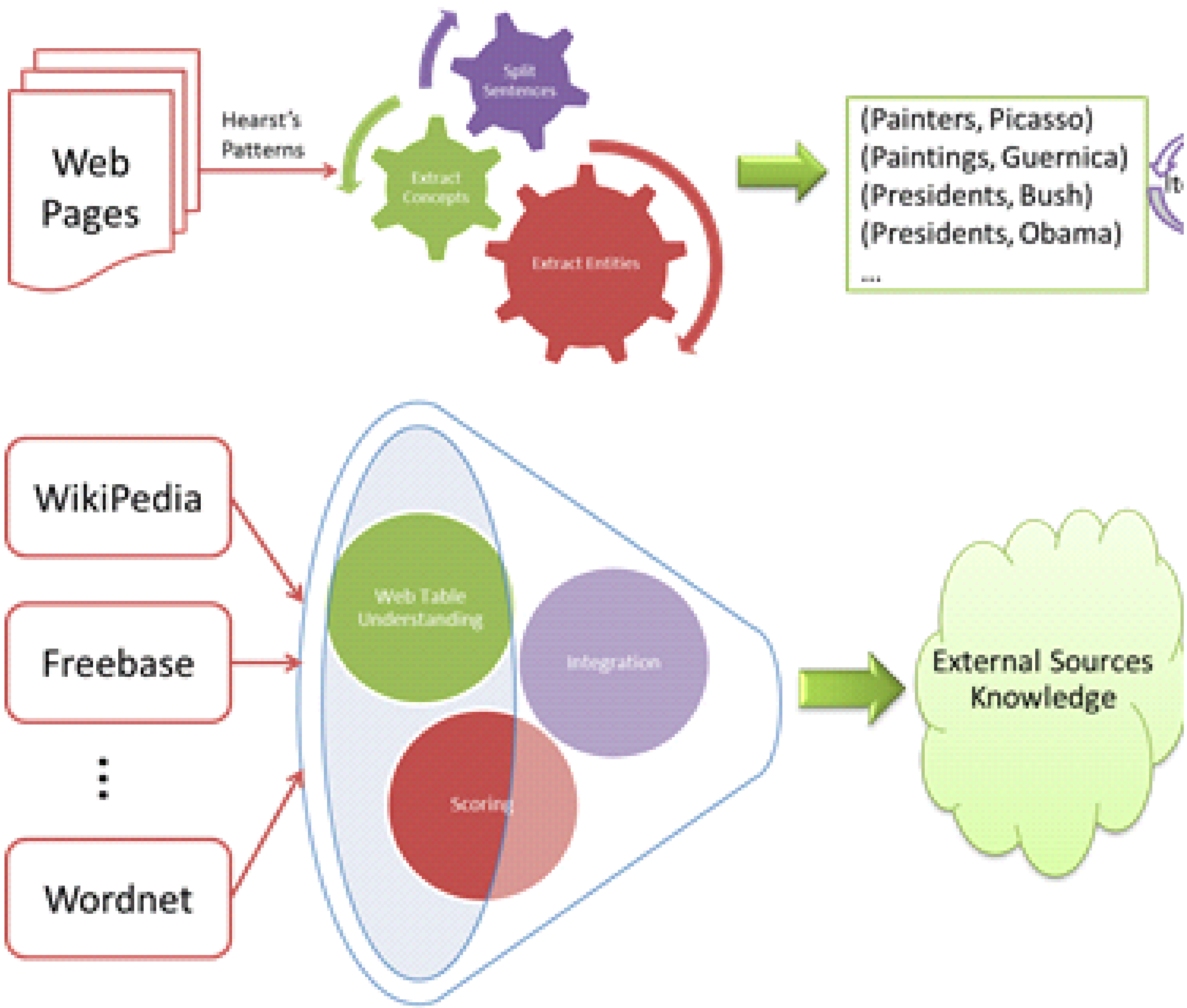
>120* 10^6 facts about these entities.

- **YAGO accuracy of 95%. has been manually evaluated.**
- **Anchored in time and space. YAGO attaches a temporal dimension and a spatial dimension to many of its facts and entities.**

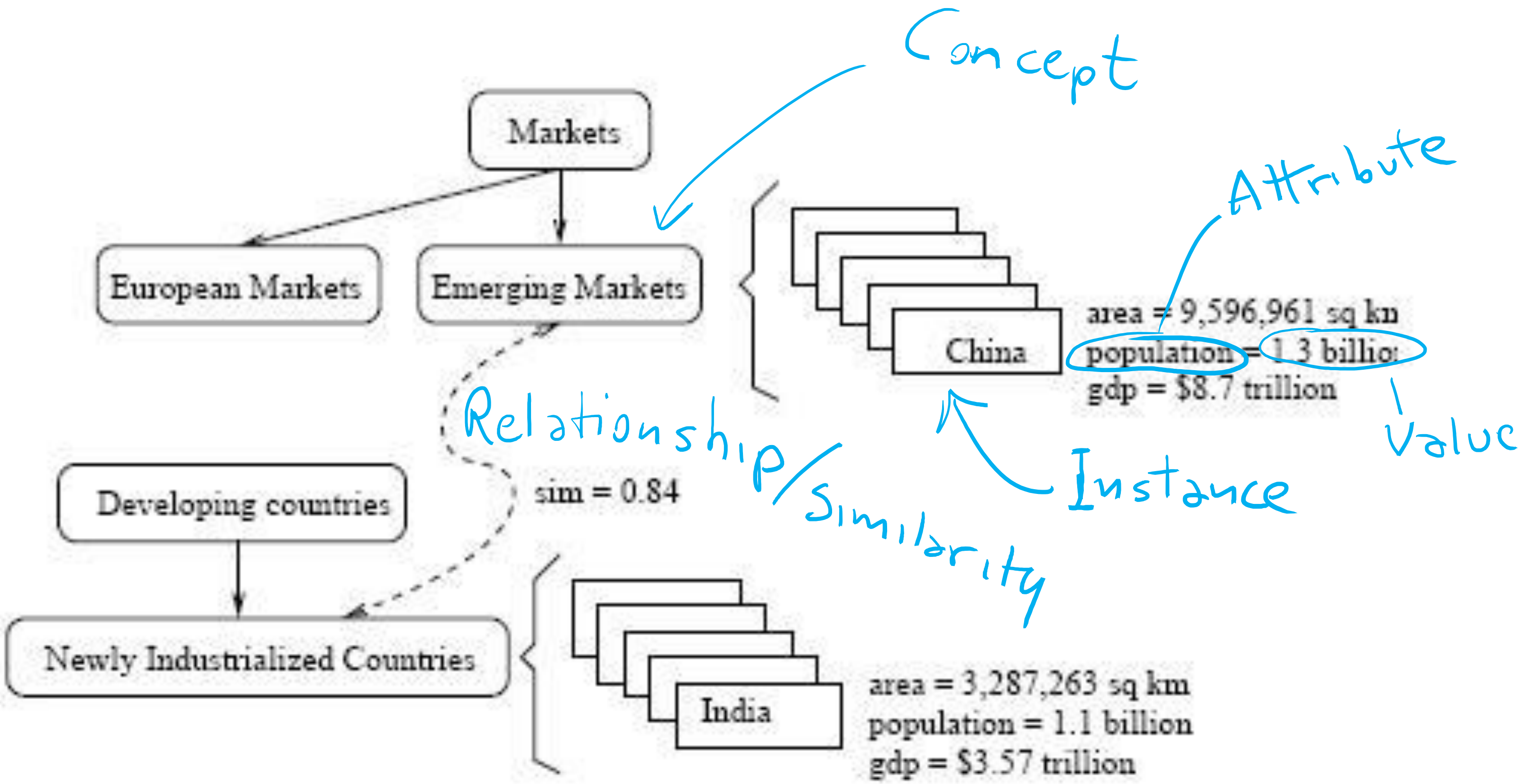
Probase (MS Research)

- Harnessed from billions of **web pages** and years worth of **search logs**
- Extremely large concept/category space (2.7 million categories).
- Probabilistic model for correctness, typicality (e.g., between concept and instance)

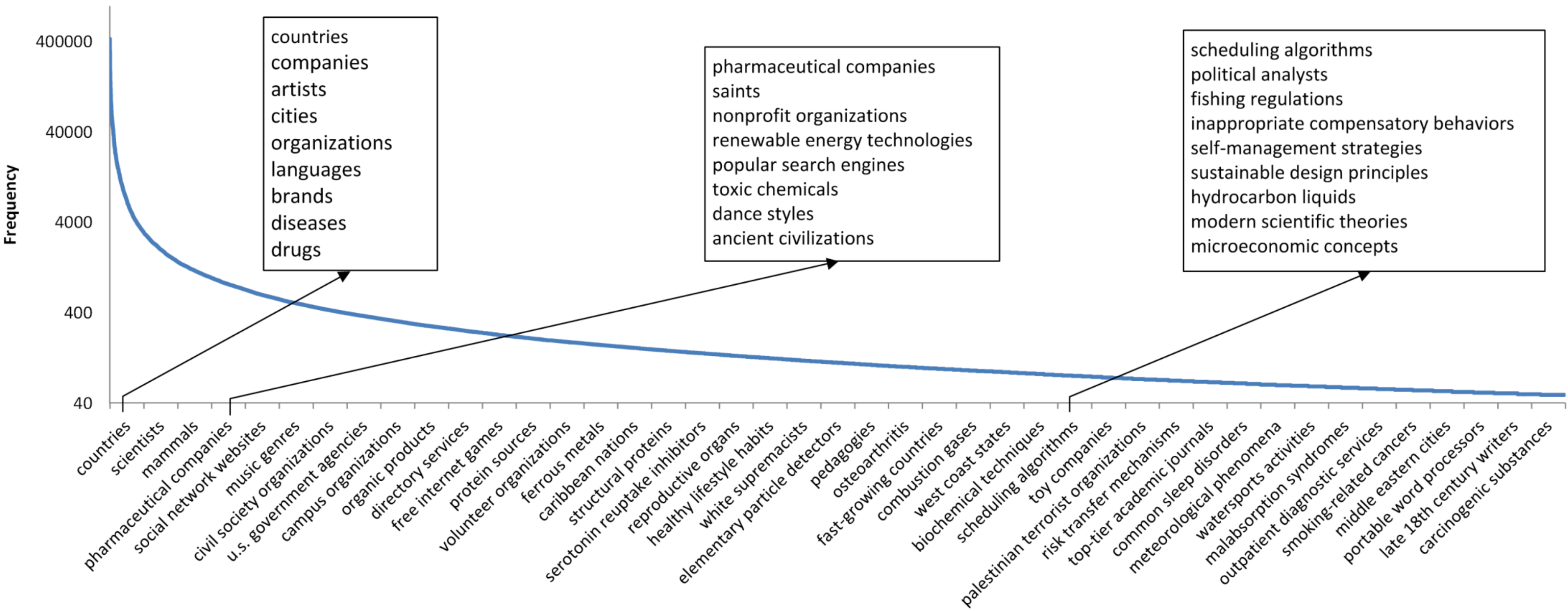
Infrastructure



A snippet of Probase's core taxonomy

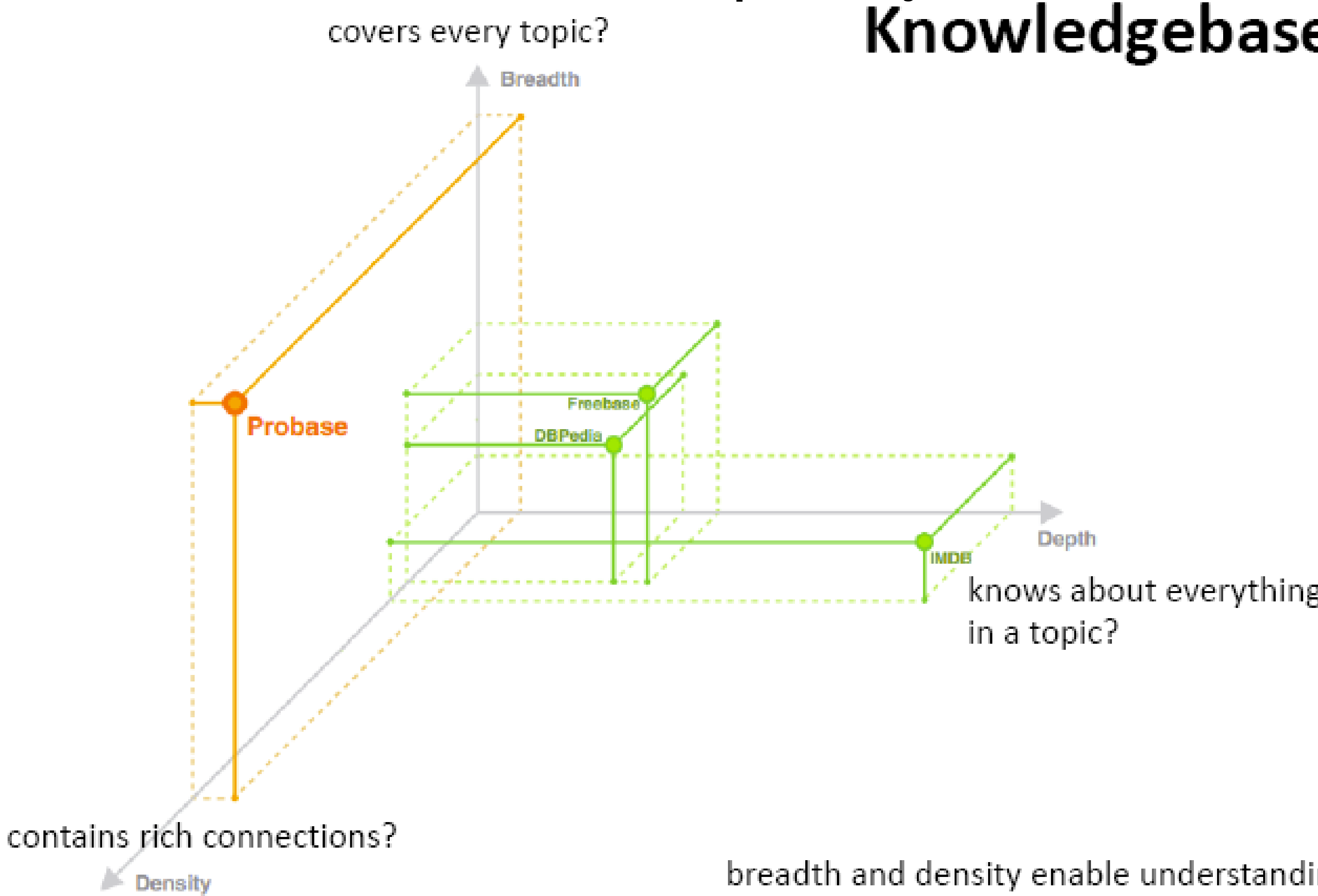


Frequency distribution of the 2.7 million concepts



Interesting dimensions to compare Ontologies (but form Probase so possibly biased)

Knowledgebases



breadth and density enable understanding

Freebase

- "Collaboratively constructed database."
- Freebase contains tens of millions of topics, thousands of types, and tens of thousands of properties and over a billion of facts
- Automatically extracted from a number of resources including **Wikipedia**, **MusicBrainz**, and **NNDB**
- as well as the **knowledge contributed by the human volunteers.**
- Each Freebase entity is assigned a set of human-readable unique keys, which are assembled of a value and a namespace.
- All **available for free** through the APIs or to download from our weekly data dumps

Lecture Overview

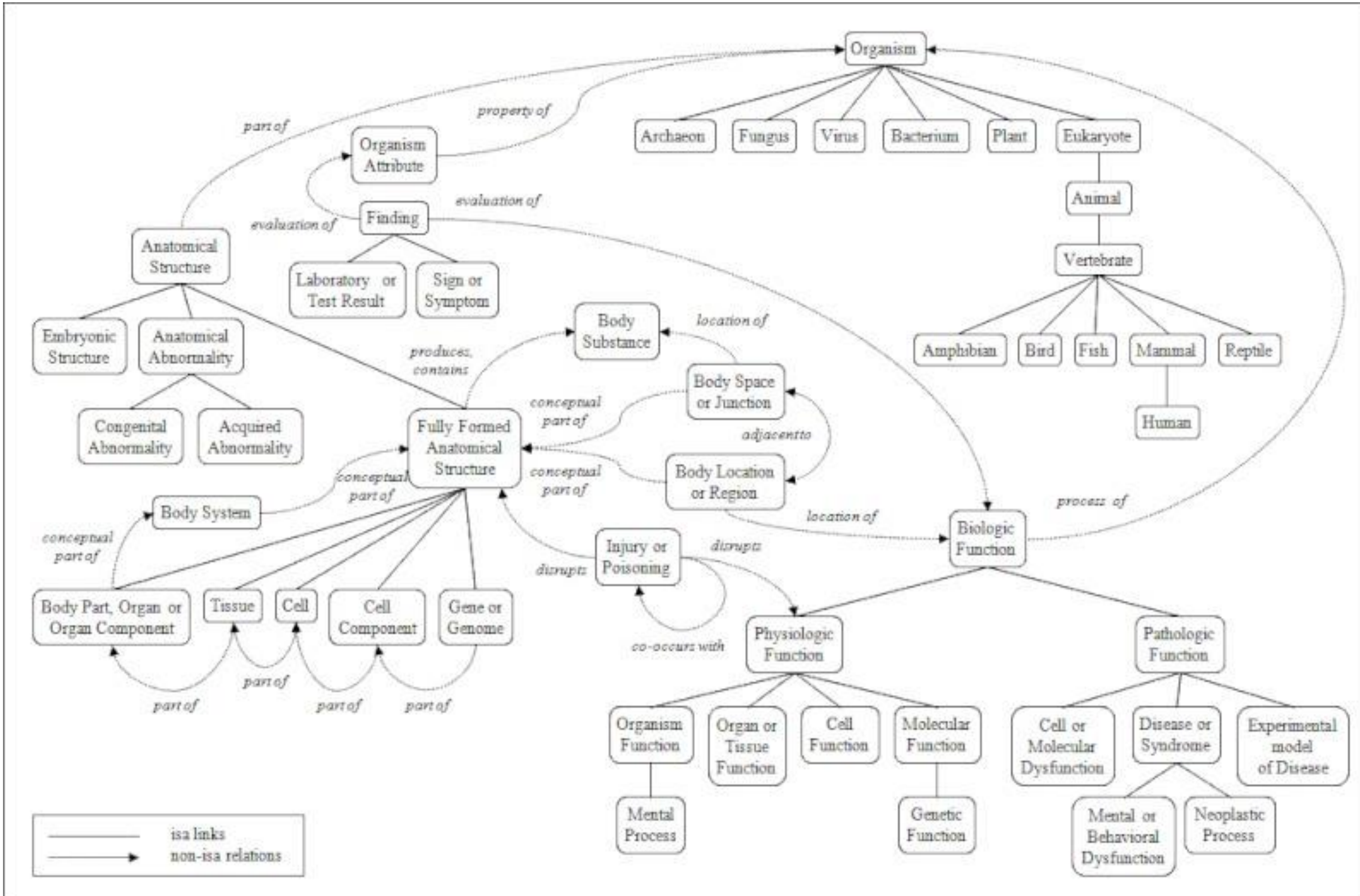
- **Ontologies** – what objects/individuals should we represent? what relations (unary, binary,..)?
- Inspiration from **Natural Language**: WordNet and FrameNet
- Extensions based on Wikipedia and mining the Web (YAGO, ProBase, Freebase)
- Domain Specific Ontologies (e.g., Medicine: MeSH, UMLS)

Domain Specific Ontologies: UMLS, MeSH

Unified Medical Language System: brings together many health and biomedical vocabularies

- Enable interoperability (linking medical terms, drug names)
- Develop electronic health records, classification tools
- Search engines, data mining

Portion of the UMLS Semantic Net



WSD: More Recent Trends
SemEval workshops - Cross Language Evaluation Forum (CLEF)

- **Better ML techniques** **(e.g., Combining Classifiers)**
- **Combining ML and Lesk** (Yuret,2004)
- **Other Languages**
- **Building better/larger corpora**

State-of-the-art systems and current literature

- **For online (certified) public systems see course web page**

- **R. Navigli. Word Sense Disambiguation: a Survey. ACM Computing Surveys, 41(2), ACM Press, 2009, pp. 1-69.**