Common Sense and Language

Benjamin Van Durme

Johns Hopkins University

Common Sense and Language

Common sense **for** language

…

Common sense **from** language

Common Sense and Language

Common sense **for** language

…

Common sense **from** language

Hobbs (1987)

We use words to talk about the world. Therefore, to understand what words mean, we must have a prior explication of how we view the world.

Shared Knowledge

If:

Shared Knowledge

If:

- I know **p**

Shared Knowledge

If:

- I know **p**
- You know **p**

Shared Knowledge

If:

- I know **p**
- You know **p**
- I know that you know **p**

Shared Knowledge

If:

- I know **p**
- You know **p**
- I know that you know **p**
- You know that I know **p**

Shared Knowledge

If:

– I know **p**
– You know **p**
– I know that you know **p**
– You know that I know **p**

then **p** is *shared knowledge*

Background Knowledge and Common Sense

- If **p** is shared across some group, then we say that **p** is *background knowledge*, for that group

Background Knowledge and Common Sense

- If **p** is shared across some group, then we say that **p** is *background knowledge*, for that group

- When the group is big enough, we call this common knowledge, aka *common sense*

Nunberg (1987)

[…] you have this extensive body of knowledge and assumptions – the collective sense – which underlies the use of natural-language expressions. A part of this knowledge is actually possessed by all discourse participants when they interpret utterances – this is what constitutes their "common-sense beliefs" in the accepted use of the term.

[...] the collective sense does play a role in the interpretation of all utterances, even when I am ignorant of it. Whatever my internal state vis-a-vis the world, I make certain social commitments about the world when I use an expression, and these are determined by the collective sense.

Prince (1978)

Stereotypical Tacit Assumptions (STAs)

People have: parents, siblings, a spouse, a home, a television, a clock, neighbors, …

Countries have: a leader, a president, a queen, a duke, citizens, land, a language, a history, …

Minsky (1974)

A frame is a data-structure for representing a stereotyped situation, like being in a certain kind of living room, or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed.

Schank (1975)

What is a frame anyway? It has been apparent to researchers within the domain of natural language understanding for some time that the eventual limit to our solution of that problem would be our ability to characterize world knowledge. In order to build a real understanding system it will be necessary to organize the knowledge that facilitates understanding.

*[...] a frame is a general name for a class of knowledge organizing techniques that guide and enable understanding. Two types of frames that are necessary are **SCRIPTS** and **PLANS**. Scripts and plans are used to understand and generate stories and actions, and there can be little understanding without them.*

Johns Hopkins University

Clark (1975)

These implicatures [bridging inferences], though conveyed by language and a necessary part of the intended message, draw on one's knowledge of natural objects and events that goes beyond one's knowledge of language itself.

Common Sense is Prevalent

John walked up to a house.

He knocked on the door.

Common Sense is Prevalent

John walked up to a house.

He knocked on the door.

What door?

Common Sense is Prevalent

What door?

G x : house(x)

Exists y : door(y)

has(x,y)

Generally if something is a house, then it has something which is a door.

Common Sense is Prevalent

G x : house(x)

    Exists y : door(y)

        has(x,y)

Generally if something is a house, then it has something which is a door.

The succinct generic: *Houses have doors.*

McCarthy (1959)

a program has common sense if it automatically deduces for itself a sufficiently wide class of immediate consequences of anything it is told and what it already knows

Schank (1975)

The inference process that is the core of understanding is not random but rather is guided by knowledge of the situation one is trying to understand.

Hobbs *et al* (1993)

Interpretation as abduction

The interpretation of a text is the minimal explanation of why the text would be true. More precisely, to interpret a text, one must prove the logical form of the text from what is already mutually known, allowing for coercions, merging redundancies where possible, and making assumptions where necessary.

Schubert and Hwang (2000)

John heard steps behind him.

He began to run.

Why?

*[**implicit question answering** involves] searching for corroborative or antagonistic connections between tentative explanations and predictions evoked by a new sentence and those evoked by prior sentences.*

Recognizing common sense

Most obvious method:

Is p common sense?

1. State p
2. If p sounds like it is common sense, then it is common sense

Recognizing common sense

p : A person may have a therapist

Common sense?

Recognizing common sense

Van Durme (2009),

adapted from Kai von Fintel (2004),

adapted from Benny Shannon (1976):

Hey! Wait a minute! I didn’t know that **p**!

HWM**p**

HWMp

A test for presuppositions:

I had to take my sister to the airport.

HWM I didn't know that you have a sister !

*HWM I didn't know that you have to take your sister to the airport !

HWMp for Common Sense

Is p common sense?

1. Embed p as a presupposition
2. If HWMp sounds silly, then p is common sense

Examples

I left my dog at the pet therapist on my way to work

HWM! I didn't know that a pet can have a therapist!

Examples

I left my dog at the pet therapist on my way to work

HWM! I didn't know that a pet can have a therapist!

I left my sister at her therapist on my way to work

HWM! I didn't know that your sister had a therapist!

Examples

I left my dog at the pet therapist on my way to work

HWM! I didn't know that a pet can have a therapist!

I left my sister at her therapist on my way to work

HWM! I didn't know that your sister had a therapist!

I left my sister at her therapist on my way to work

HWM! I didn't know that a person can have a therapist!

Knowledge Acquisition Bottleneck

Lack of sufficient knowledge was deemed a killer of strong AI: we (thought) we could do inference, but had nothing to reason over

The community moved to “small”, task-oriented AI, focus on incremental improvements

Dreyfus (1979)

[Feigenbaum and Feldman] define progress very carefully as "displacement toward the ultimate goal." According to this definition, the first man to climb a tree could claim tangible progress toward reaching the moon.

Bayer *et al.* (2005)

[a rich, knowledge-backed system is] a rocket ship with nothing inside: fiendishly difficult to get off the ground, and unable to fly until a wide number of things work fairly well.

Bos (2013)

The bottleneck in achieving high recall is the lack of a systematic way to produce relevant background knowledge. There is a place for logic in [textual inference], but it is (still) overshadowed by the knowledge acquisition problem.

PReLiM

Goal:

probabilistic model of natural language semantics

Assuming success:

where do we get the probabilities??

Build up a KB?

If we can recognize common sense, then perhaps we can **manually** generate in bulk? For example:

Early musings:
Hobbs and Navarretta (1993)

Professionals:
Cyc (Lenat 1995), Project Halo (Friedland and Allen 2003)

Crowdsourced:
OpenMind (Singh 2002, Havasi *et al* 2007), Learner (Chklovski 2003)

Games with a purpose:
Verbosity (von Ahn *et al* 2004)

Common Sense and Language

Common sense **for** language

or

Common sense **from** language

This won't work ?
Havasi et al (2007)

Grice's theory of pragmatics states that when communicating, people tend not to provide information which is obvious or extraneous. If someone says *"I bought groceries"*, he is unlikely to add that he used money to do so, unless the context made this fact surprising or in question. This means that it is difficult to automatically extract common-sense statements from text, and the results tend to be unreliable and need to be checked by a human.

Oren Etzioni (Tuesday)

Q: How are we going to get this knowledge from the web?

A: The knowledge we need ain't there

Relevant bits of the relevant maxims

Maxim of Quantity

Make your contribution as informative as is required

Maxim of Manner

Be brief

Relevant bits of the relevant maxims

Maxim of Quantity

Make your contribution as informative as is required

Maxim of Manner

Be brief

Information Theoretic version:

Maximize available bandwidth

Relevant bits of the relevant maxims

Maxim of Quantity

Make your contribution as informative as is required

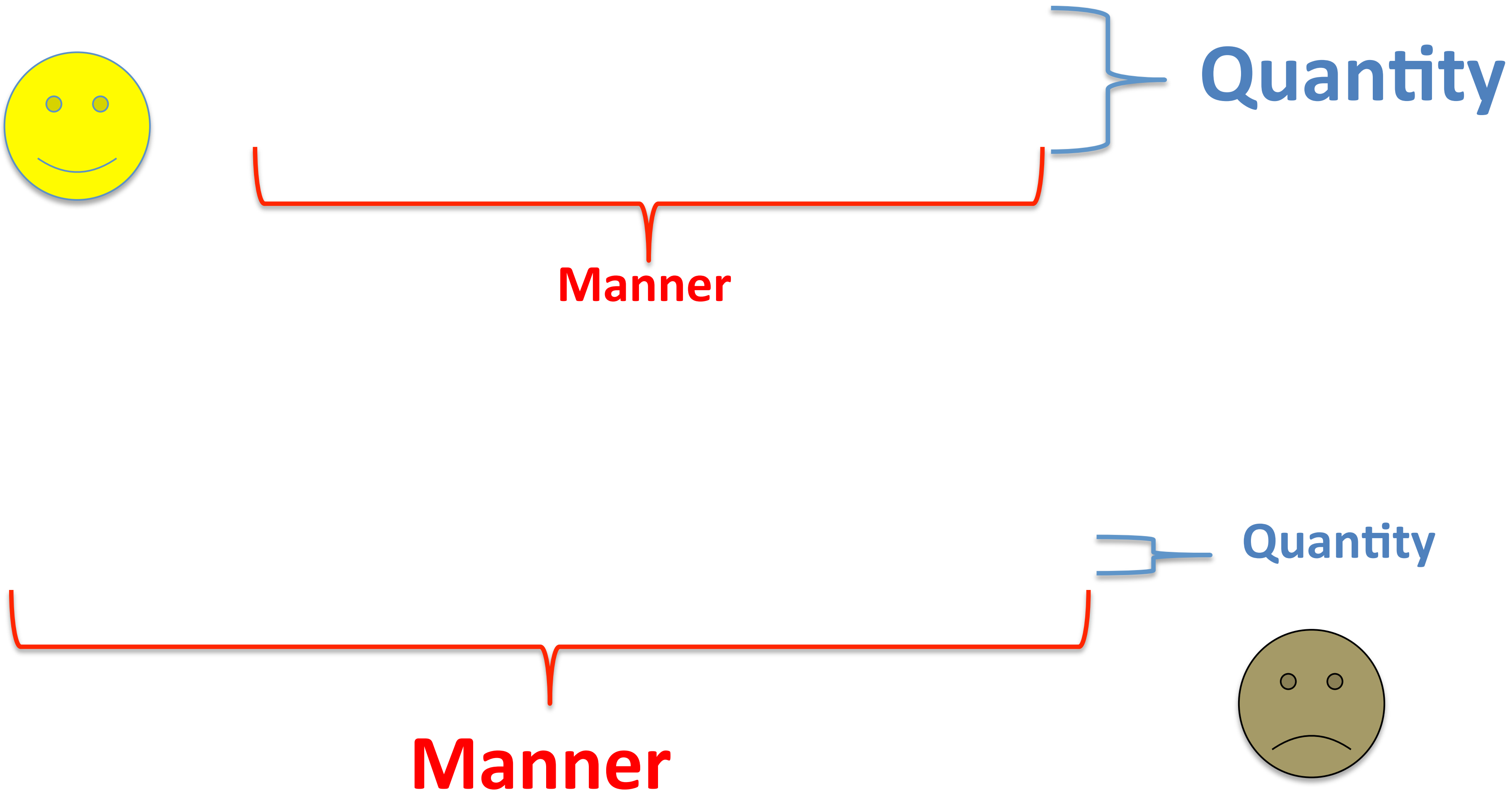
Maxim of Manner

Be brief

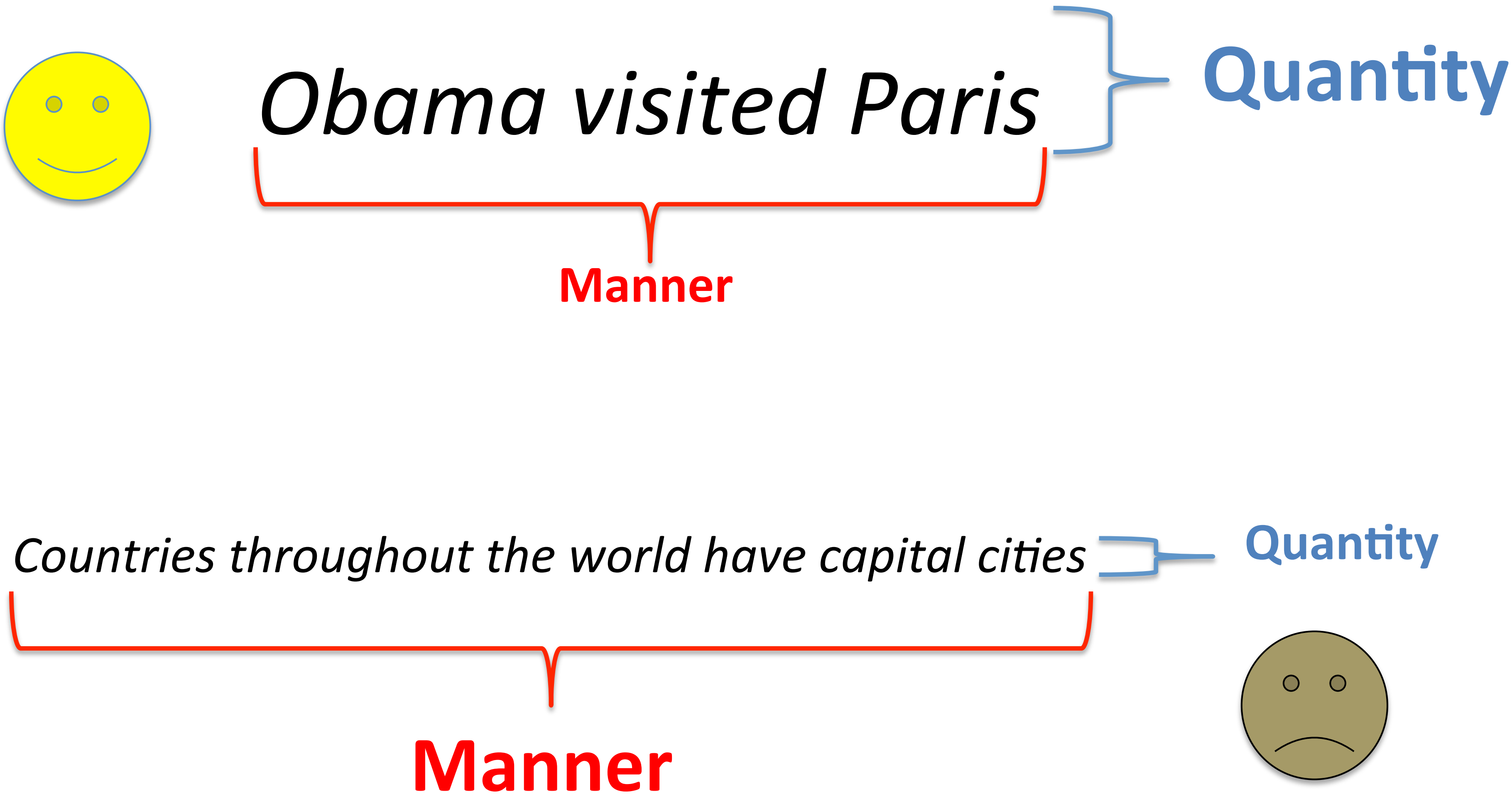
slkdjfkasjdfljaslkjdflkajsdflkjaslkdfj

Quantity

Manner



JOHNS HOPKINS
UNIVERSITY



JOHNS HOPKINS UNIVERSITY

Two approaches

Recognizing patterns of situation descriptions

Directly interpreting generic sentences

Two approaches

Recognizing patterns of situation descriptions

generalize from assertions about individuals

Directly interpreting generic sentences

such as found in dictionaries, encyclopedias

Two approaches

Recognizing patterns of situation descriptions

Directly interpreting generic sentences

Schubert (2002)

there is a largely untapped source of general knowledge in texts, lying at the level beneath the explicit assertional content

Knowledge Acquisition

"Machine Reading" often denotes open domain factoid extraction, e.g.:

Obama visited Hawaii

Knowledge Acquisition, tends to mean "common sense" knowledge acquisition:

A president may visit a US-State

Knowledge Acquisition
(Data Mining on Text)

step 1: Recognize patterns of situation descriptions in a large collection of text

step 2: Abstract from individuals (entities) to kinds (classes)

step 3: Count

steps 4…n: ?

result: rules, probabilities, …, something useful for performing inference

Recognize patterns of situation descriptions

Pick your favorite term:

Underspecified Logical Form Parsing,

Semantic Role Labeling,

Open IE,

Deep Syntactic Parsing (e.g., Stanford Deps.),

...

Run over a large collection of documents

Examples (some of many)

"Knowledge" Patterns

KNEXT (Schubert 2002)
Liakata and Pulman (2002)
Almuhareb and Poesio (2004)
ALICE (Banko and Etzioni 2007)
Pasca and Van Durme (2007)
DART (Clark and Harrison 2009)
Berant *et al* (2011)
…

"Language" Patterns

Church and Hanks (1989)
Zernik (1992)
Resnik (1993)
Rosenfeld (1994)
Abney (1996)
Lin (1998)
VerbOcean (Chklovski and Pantel 2004)
Chambers and Jurafsky (2008)
PPDB (Ganitkevitch *et al* 2013)
…

JOHNS HOPKINS UNIVERSITY

KNEXT (Schubert 2002)

Rilly or Glendora had entered her room while she slept, bringing back her washed clothes.

KNEXT (Schubert 2002)

Rilly or Glendora had entered her room while she slept, bringing back her washed clothes.

```
((S (NP (NP (NNP Rilly) )
              (CC or) (NP (NNP Glendora) ))
    (AUX (VBD had) )
    (VP (VBN entered)
        (NP (PRP$ her) (NN room) ))
    (SBAR (IN while)
          (S (NP (PRP she) ) (VP (VBD slept) )))
      ...
```

KNEXT (Schubert 2002)

Rilly or Glendora had entered her room while she slept, bringing back her washed clothes.

[(det.q person.n) enter.v (the.q room.n)]

[(det.q female-individual.n) have.v (det.q room.n)]

[(det.q female-individual.n) sleep.v]

[(det.q female-individual.n) have.v
(det.q (plur.f clothe.n))]

[(det.q (plur.f clothe.n)) washed.a]

KNEXT (Schubert 2002)

Rilly or Glendora had entered her room while she slept, bringing back her washed clothes.

[(det.q person.n) enter.v (the.q room.n)]

A person may enter a room

[(det.q female-individual.n) sleep.v]

A female may sleep

[(det.q (plur.f clothe.n)) washed.a]

Clothes may be washed

www.cs.rochester.edu/research/knext

QUERY RESULTS: APPLE EDIBLE

5 results in the active KB.

An apple can be edible.

Logical form

[⟨det apple.n⟩ edible.a]

Learned from

8 sentences.

www.cs.rochester.edu/research/knext

QUERY RESULTS: APPLE EDIBLE

QUERY RESULTS: APPLE FLESH

3 results in the active KB.

Flesh may pertain-to an apple.

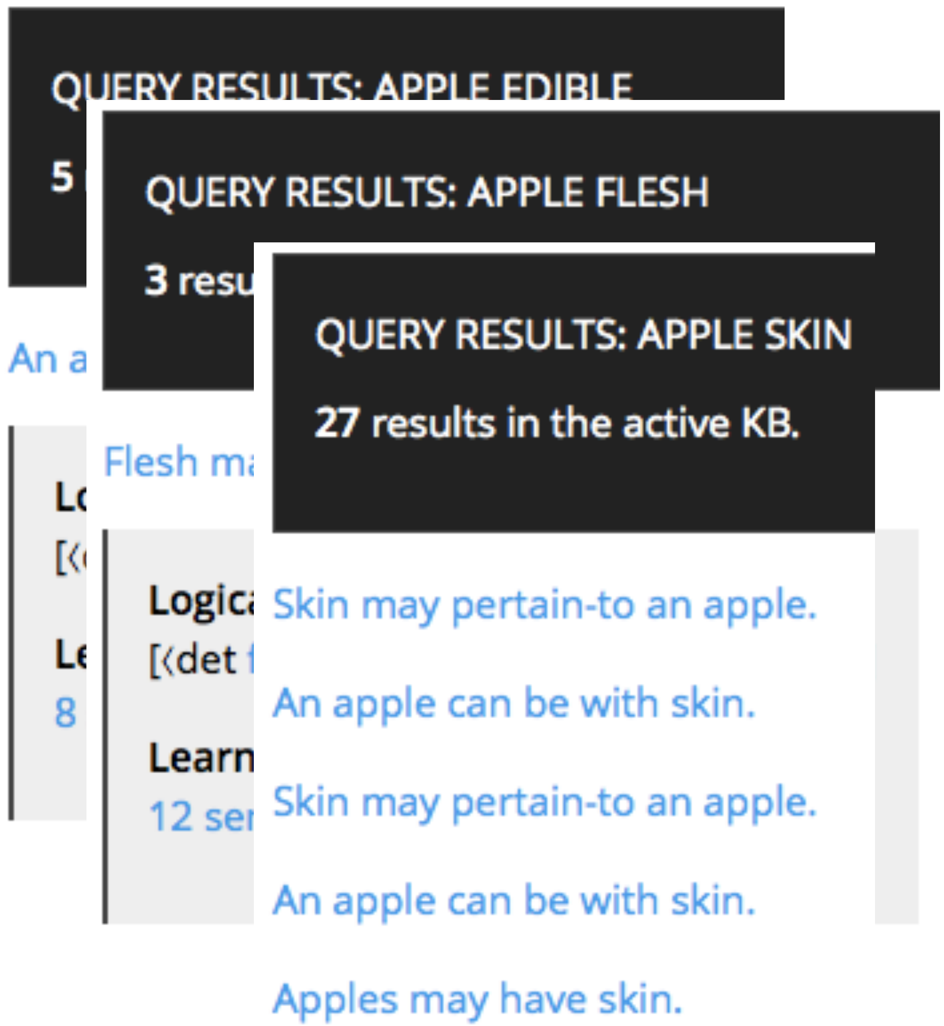
Logical form

[⟨det flesh.n⟩ pertain-to.v ⟨det apple.n⟩]

Learned from

12 sentences.

www.cs.rochester.edu/research/knext



JOHNS HOPKINS
UNIVERSITY

www.cs.rochester.edu/research/knext

Apples may pertain-to a person.

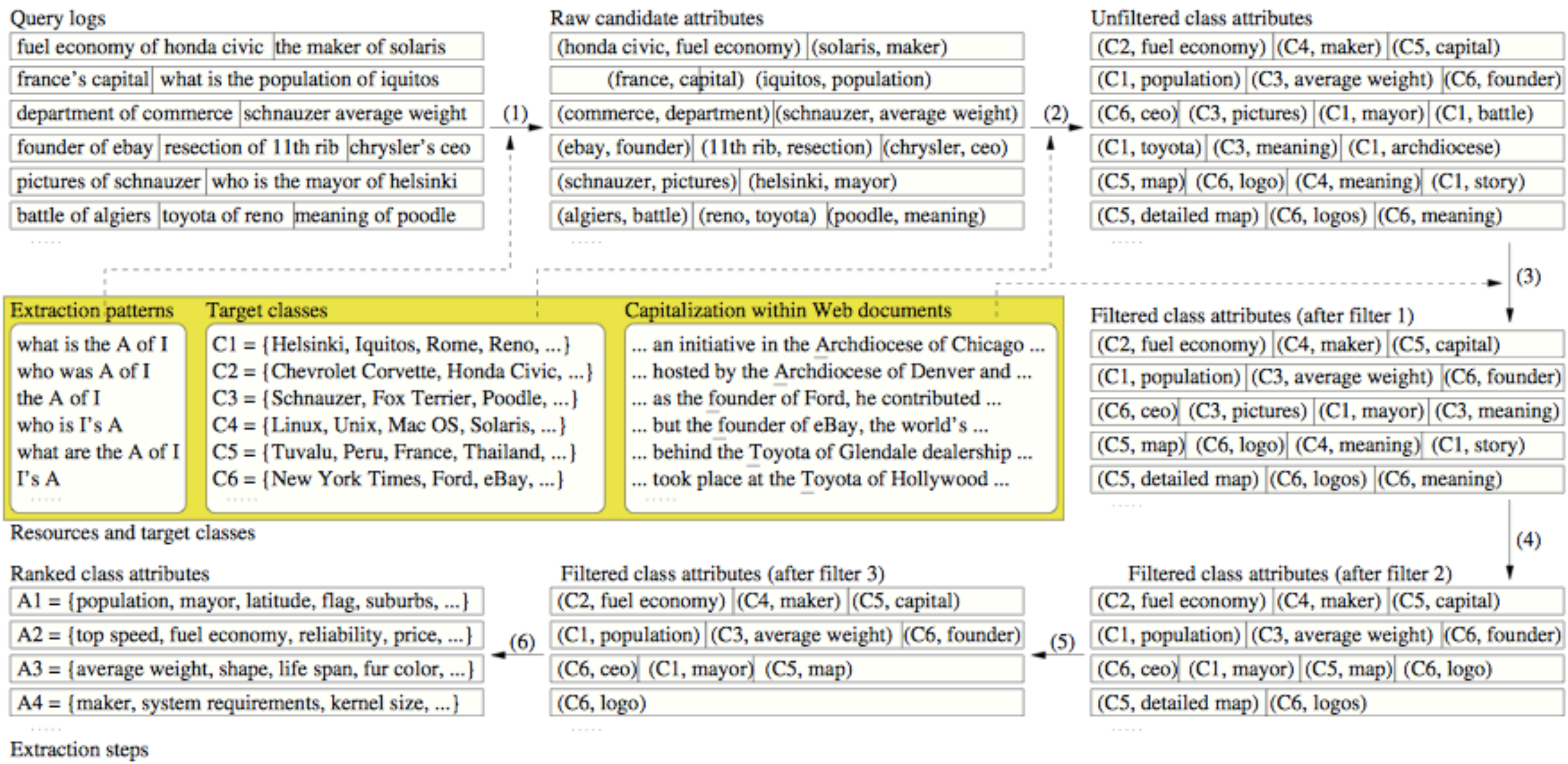
(News search results) can be for (apple computer).

Apples can be fresh.

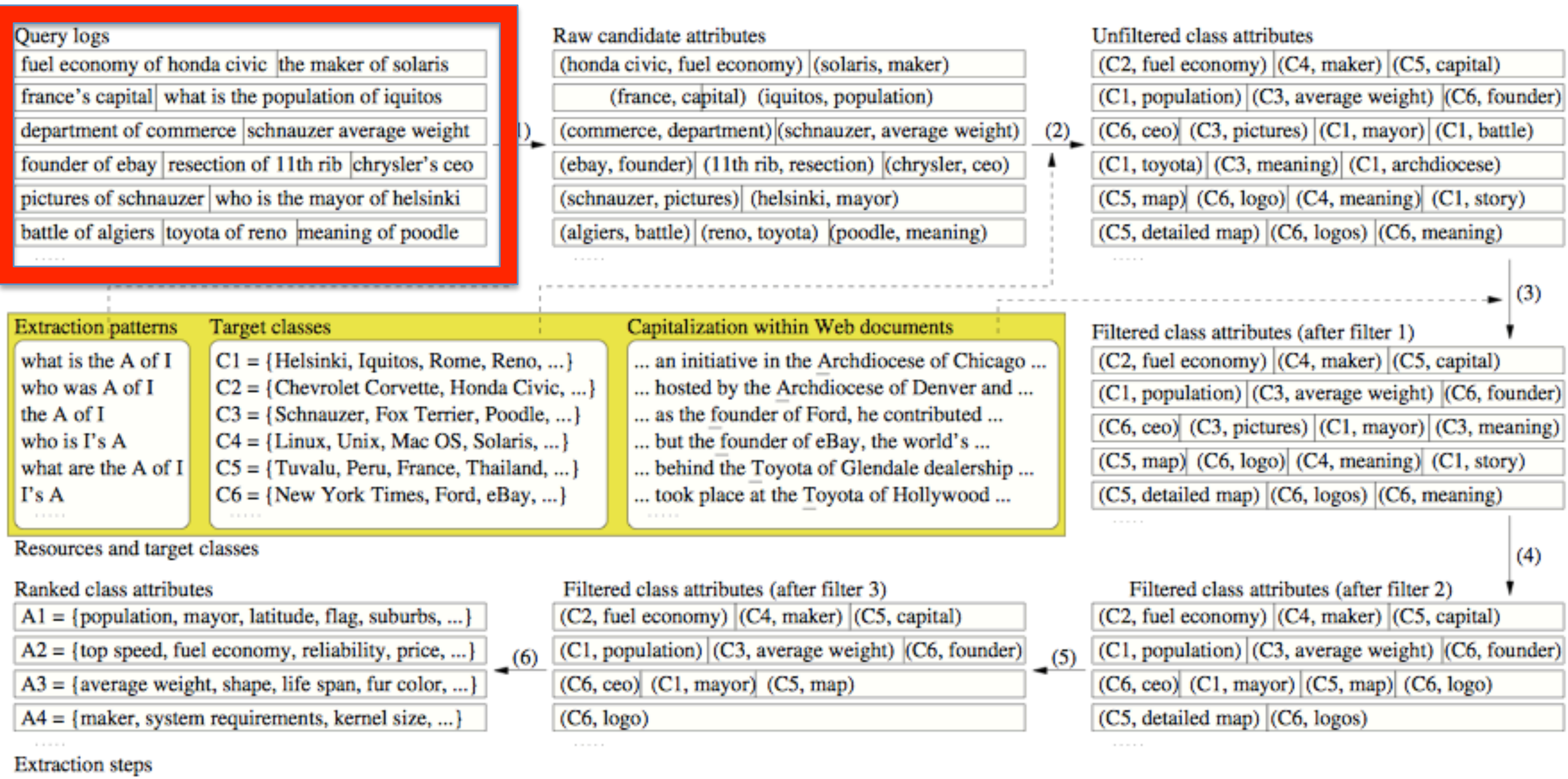
An apple may fall.

Apples can be sliced.

Pasca and Van Durme (2007)



Pasca and Van Durme (2007)



Pasca and Van Durme (2007)

Query logs

fuel economy of honda civic	the maker of solaris	
france's capital	what is the population of iquitos	
department of commerce	schnauzer average weight	
founder of ebay	resection of 11th rib	chrysler's ceo
pictures of schnauzer	who is the mayor of helsinki	
battle of algiers	toyota of reno	meaning of poodle

Pasca and Van Durme (2007)

what is the capital of Texas

what is the capital of New York

what is the capital of Florida

…

A US-State may have a capital

Church and Hanks (1989)

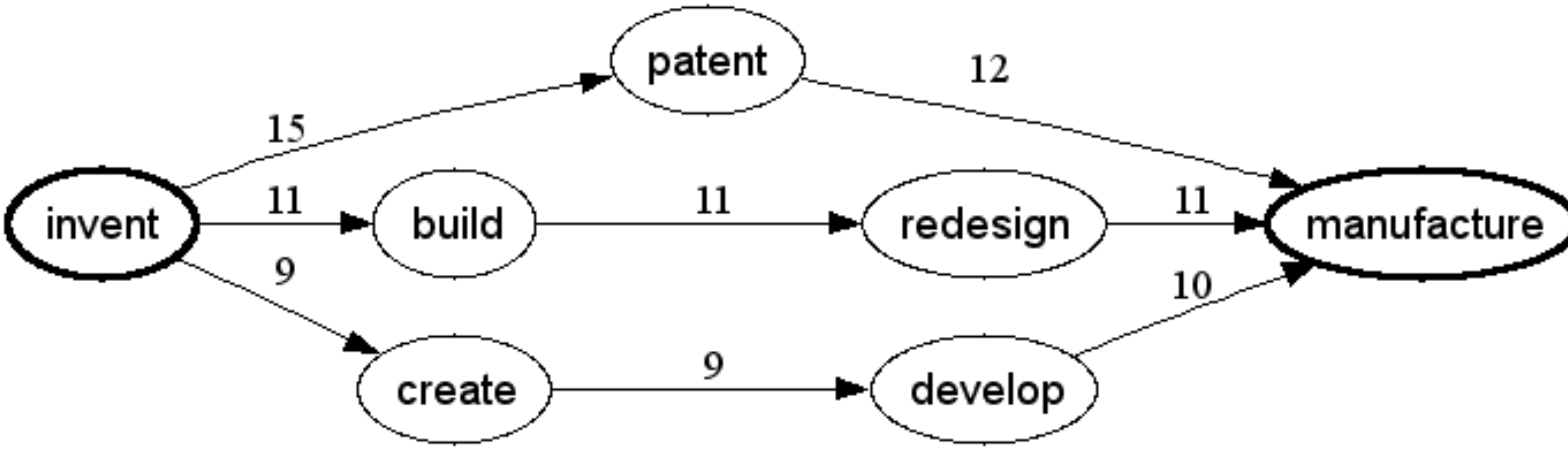
3. save ANIMAL from DESTRUCT(ION) (5 concordance lines)

give them the money to	save the dogs[ANIMAL] from being destroyed[DESTRUCT] ,
program intended to	save the giant birds[ANIMAL] from extinction[DESTRUCT] ,

VerbOcean (Chklovski and Pantel 2004)

happens-before (12)	to X and then Y to X * and then Y Xed and then Yed Xed * and then Yed to X and later Y Xed and later Yed to X and subsequently Y Xed and subsequently Yed to X and eventually Y Xed and eventually Yed

VerbOcean (Chklovski and Pantel 2004)

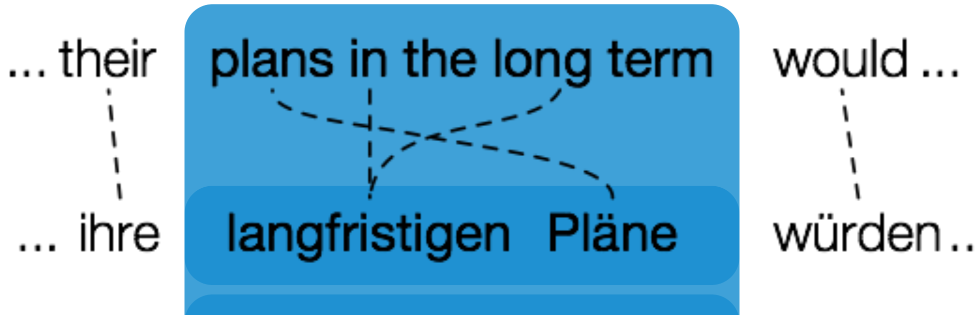


ParaPhrase DataBase (PPDB) Ganitkevitch *et al* (2013)

- Differing **textual** expressions of the same **meaning**:

cup	⟺	mug
the king's speech	⟺	His Majesty's address
X_1 talks to X_2	⟺	X_1 converses with X_2
one JJ instance of NP	⟺	a JJ case of NP

Pivoting over a Foreign Language



Johns Hopkins University

Paraphrase Quality

thrown into jail

arrested
detained
imprisoned
incarcerated
jailed
locked up
taken into custody
thrown into prison

be thrown in prison
been thrown into jail
being arrested
in jail
in prison
put in prison for
were thrown into jail
who are held in detention

arrest
cases
custody
maltreated
owners
protection
thrown

PPDB: The Paraphrase Database

http://paraphrase.org/

	All	Lexical	One-To-Many	Phrasal	Syntactic
S	Paraphrases (424MB, 6.8M rules)	Paraphrases (1.7MB, 31k rules) Identity (16MB, 437k rules)	One-To-Many (3.8MB, 47k rules) Many-To-One (3.8MB, 47k rules)	Paraphrases (42MB, 637k rules) Identity (170MB, 4.1M rules)	Constituent (38MB, 585k rules) Non-Constituent (343MB, 5.6M rules)
M	Paraphrases (757MB, 11.9M rules)	Paraphrases (1.7MB, 69k rules) Identity (16MB, 468k rules)	One-To-Many (7.6MB, 94k rules) Many-To-One (7.6MB, 94k rules)	Paraphrases (42MB, 1.2M rules) Identity (170MB, 4.3M rules)	Constituent (69MB, 1.0M rules) Non-Constituent (601MB, 9.6M rules)
L	Paraphrases (1.5GB, 23.5M rules)	Paraphrases (12MB, 198k rules) Identity (19MB, 503k rules)	One-To-Many (16MB, 188k rules) Many-To-One (16MB, 188k rules)	Paraphrases (209MB, 3.0M rules) Identity (191MB, 4.5M rules)	Constituent (148MB, 2.2M rules) Non-Constituent (1.2GB, 18.2M rules)
XL	Paraphrases (2.8GB, 43.2M rules)	Paraphrases (33MB, 548k rules) Identity (20MB, 532k rules)	One-To-Many (31MB, 376k rules) Many-To-One (31MB, 376k rules)	Paraphrases (486MB, 6.9M rules) Identity (198MB, 4.7M rules)	Constituent (300MB, 4.4M rules) Non-Constituent (2.1GB, 31.4M rules)
XXL	Paraphrases (5.7GB, 86.4M rules)	Paraphrases (125MB, 2.1M rules) Identity (21MB, 559k rules)	One-To-Many (61MB, 752k rules) Many-To-One (61MB, 752k rules)	Paraphrases (1.5GB, 20.2M rules) Identity (204MB, 4.8M rules)	Constituent (644MB, 9.3M rules) Non-Constituent (3.6GB, 54.8M rules)
XXXL	Paraphrases (12.2GB, 169M rules)	Paraphrases (451MB, 7.6M rules) Identity (22MB, 570k rules)	One-To-Many (117MB, 1.5M rules) Many-To-One (117MB, 1.5M rules)	Paraphrases (4.9GB, 68.4M rules) Identity (207MB, 4.9M rules)	Constituent (1.1GB, 16.1M rules) Non-Constituent (5.1GB, 77.4M rules)

PPDB: The Paraphrase Database

http://paraphrase.org/

	All	Lexical	One-To-Many	Phrasal	Syntactic
S	Paraphrases (424MB, 6.8M rules)	Paraphrases (1.7MB, 31k rules) Identity (16MB, 437k rules)	One-To-Many (3.8MB, 47k rules) Many-To-One (3.8MB, 47k rules)	Paraphrases (42MB, 637k rules) Identity (170MB, 4.1M rules)	Constituent (38MB, 585k rules) Non-Constituent (343MB, 5.6M rules)
M	Paraphrases (757MB, 11.9M rules)	Paraphrases (1.7MB, 69k rules) Identity (16MB, 468k rules)	One-To-Many (7.6MB, 94k rules) Many-To-One (7.6MB, 94k rules)	Paraphrases (42MB, 1.2M rules) Identity (170MB, 4.3M rules)	Constituent (69MB, 1.0M rules) Non-Constituent (601MB, 9.6M rules)
L	Paraphrases (1.5GB, 23.5M rules)	Paraphrases (12MB, 198k rules) Identity (19MB, 503k rules)	One-To-Many (16MB, 188k rules) Many-To-One (16MB, 188k rules)	Paraphrases (209MB, 3.0M rules) Identity (191MB, 4.5M rules)	Constituent (148MB, 2.2M rules) Non-Constituent (1.2GB, 18.2M rules)
XL	Paraphrases (2.8GB, 43.2M rules)	Paraphrases (33MB, 548k rules) Identity (20MB, 532k rules)	One-To-Many (31MB, 376k rules) Many-To-One (31MB, 376k rules)	Paraphrases (486MB, 6.9M rules) Identity (198MB, 4.7M rules)	Constituent (300MB, 4.4M rules) Non-Constituent (2.1GB, 31.4M rules)
XXL	Paraphrases (5.7GB, 86.4M rules)	Paraphrases (125MB, 2.1M rules) Identity (21MB, 559k rules)	One-To-Many (61MB, 752k rules) Many-To-One (61MB, 752k rules)	Paraphrases (1.5GB, 20.2M rules) Identity (204MB, 4.8M rules)	Constituent (644MB, 9.3M rules) Non-Constituent (3.6GB, 54.8M rules)
XXXL	Paraphrases (12.2GB, 169M rules)	Paraphrases (451MB, 7.6M rules) Identity (22MB, 570k rules)	One-To-Many (117MB, 1.5M rules) Many-To-One (117MB, 1.5M rules)	Paraphrases (4.9GB, 68.4M rules) Identity (207MB, 4.9M rules)	Constituent (1.1GB, 16.1M rules) Non-Constituent (5.1GB, 77.4M rules)

PPDB: The Paraphrase Database

http://paraphrase.org/

	All	Lexical	One-To-Many	Phrasal	Syntactic
S	Par[illegible] (424M[illegible]	Paraphrases (1.7MB, 31k rules) [illegible]	One-To-Many (3.8MB, 47k rules) [illegible]	Paraphrases (42MB, 637k rules) [illegible]	Constituent (38MB, 585k rules) Non-Constituent (343MB, 5.6M rules)
M	Paraphrases (757MB, 11.9M rules)	[illegible] Identity (16MB, 468k rules)	[illegible] Many-To-One (7.6MB, 94k rules)	[illegible] Identity (170MB, 4.3M rules)	Constituent (69MB, 1.0M rules) Non-Constituent (601MB, 9.6M rules)
L	Paraphrases (1.5GB, 23.5M rules)	Paraphrases (12MB, 198k rules) Identity (19MB, 503k rules)	One-To-Many (16MB, 188k rules) Many-To-One (16MB, 188k rules)	Paraphrases (209MB, 3.0M rules) Identity (191MB, 4.5M rules)	Constituent (148MB, 2.2M rules) Non-Constituent (1.2GB, 18.2M rules)
XL	Paraphrases (2.8GB, 43.2M rules)	Paraphrases (33MB, 548k rules) Identity (20MB, 532k rules)	One-To-Many (31MB, 376k rules) Many-To-One (31MB, 376k rules)	Paraphrases (486MB, 6.9M rules) Identity (198MB, 4.7M rules)	Constituent (300MB, 4.4M rules) Non-Constituent (2.1GB, 31.4M rules)
XXL	Paraphrases (5.7GB, 86.4M rules)	Paraphrases (125MB, 2.1M rules) Identity (21MB, 559k rules)	One-To-Many (61MB, 752k rules) Many-To-One (61MB, 752k rules)	Paraphrases (1.5GB, 20.2M rules) Identity (204MB, 4.8M rules)	Constituent (644MB, 9.3M rules) Non-Constituent (3.6GB, 54.8M rules)
XXXL	Paraphrases (12.2GB, 169M rules)	Paraphrases (451MB, 7.6M rules) Identity (22MB, 570k rules)	One-To-Many (117MB, 1.5M rules) Many-To-One (117MB, 1.5M rules)	Paraphrases (4.9GB, 68.4M rules) Identity (207MB, 4.9M rules)	Constituent (1.1GB, 16.1M rules) Non-Constituent (5.1GB, 77.4M rules)

http://paraphrase.org

From individuals to kinds

To generalize to a statement about an *individual* (entity) to a *kind* (semantic class), then we need a mapping.

Generalized gazetteers, or *ontologies*

A few examples

Hearst (1992)

Caraballo (1999)

Pantel and Ravichandran (2004)

Snow *et al* (2005)

Van Durme and Pasca (2008)

Talukdar *et al* (2008)

Bansal *et al* (2014)

Snow *et al* (2005)

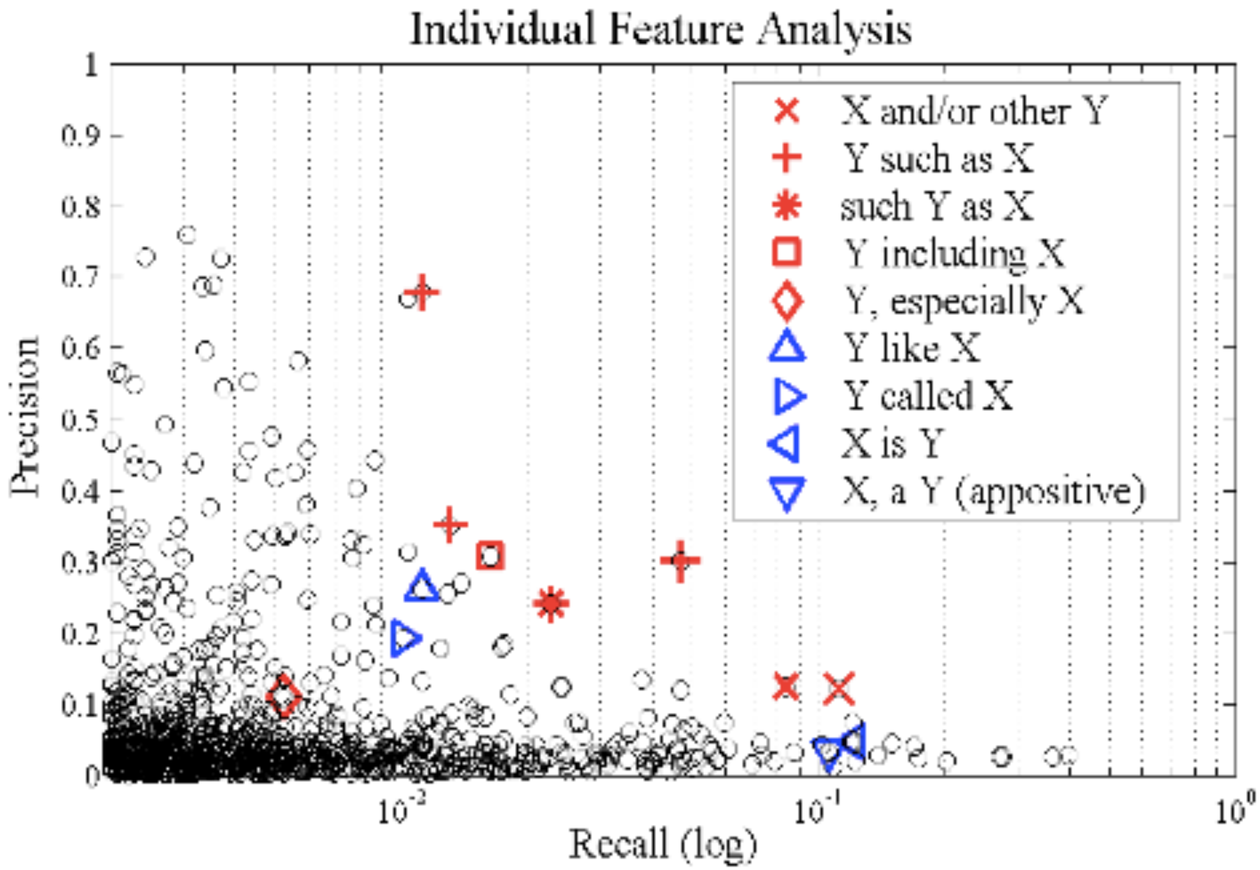


Figure 2: Hypernym pre/re for all features

3. Abstracting Existentials

$''Rover\ barked\ all\ night'',$
$''Rover\ is\ a\ dog'' \rightarrow$

$$\exists\, x.Dog(x)\,\&\,Bark(x)\ \rightarrow$$

$$It\ is\ possible\ that\ a\ dog\ will\ bark \rightarrow$$

$$''Dogs\ bark''$$

Challenges

- Extract basic situational descriptions
- Extract chained situational descriptions
- Extract statistics concerning the real world

Challenges

- Extract basic situational descriptions
  - SOLVED (kinda sorta)
  - Enables collection of basic patterns of predication
- Extract chained situational descriptions
- Extract statistics concerning the real world

Challenges

- Extract basic situational descriptions
  - SOLVED (kinda sorta)
  - Enables collection of basic patterns of predication
- Extract chained situational descriptions
  - ATTEMPTED (coref is noisy, but provides signal)
  - Enables noisy collection of script-like models
- Extract statistics concerning the real world

Challenges

- Extract basic situational descriptions
  - SOLVED (kinda sorta)
  - Enables collection of basic patterns of predication
- Extract chained situational descriptions
  - ATTEMPTED (coref is noisy, but provides signal)
  - Enables noisy collection of script-like models
- Extract statistics concerning the real world
  - **PROBLEM** (Reporting bias)

Reporting Bias

Van Durme ('09), Gordon & Van Durme ('13)

Is there a difference between a semantic language model, and a model of the world?

Blinking and Breathing (Van Durme '09)

Word	Teraword	Word	Teraword
spoke	11,577,917	breathed	725,034
laughed	3,904,519	hugged	610,040
murdered	2,843,529	blinked	390,692
inhaled	984,613	exhaled	168,985

Reporting Bias
Van Durme ('09), Gordon & Van Durme ('13)

Is there a difference between a semantic language model, and a model of the world?

And are we even looking for a model of the real world, or a model of how humans think about the world?

ask me offline about psycholinguistics, or see hints in works like:
Fine, Frank, Jaeger and Van Durme. ACL. 2014.

Two approaches

Recognizing patterns of situation descriptions

Directly interpreting generic sentences

Directly interpreting generics

Reference works such as textbooks, dictionaries, and encyclopedias are full of this

Dogs perform many roles for people, such as hunting, herding, pulling loads, protection, assisting police and military, companionship, and, more recently, aiding handicapped individuals. [*Dog*. Wikipedia. Feb. 17, 2011]

What is a generic?

Krifka *et al* (1995)

$$\text{G}{\scriptsize\text{EN}}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\mathbf{Restrictor}[x_1, ..., x_n];\ \mathbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

Van Durme (2009)

generics are used by humans to express rules that they take to underlie patterns observed in the world, and usually have strong quantificational force. In the many cases where this would seem to lead to error, such as in: A bird lays eggs, we assume the existence of implicit constraints (e.g., (female) bird) in the domain restrictor that allows for the strong reading. Cases such as: Lightning rarely strikes people, show that generics are not universally strongly quantified. The key property of a generic are their nomic, or rule-like, character.

JOHNS HOPKINS
UNIVERSITY

Dogs bark

$$\text{G{\scriptsize EN}}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\textbf{Restrictor}[x_1, ..., x_n];\ \textbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

(Gen) Dogs bark

$$\textsc{Gen}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\mathbf{Restrictor}[x_1, ..., x_n];\ \mathbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

Some, all, many, most, at least a few, ..., dogs bark

(Gen) Dogs bark

$$\textsc{Gen}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\mathbf{Restrictor}[x_1, ..., x_n];\ \mathbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

Some, all, many, most, at least a few, …

(Gen) Dogs bark

$$\textsc{Gen}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\mathbf{Restrictor}[x_1, ..., x_n];\ \mathbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

Some, all, many, most, at least a few, …

Dogs bark

Gen x : dog(x)

Exists e: bark(x,e)

For some, most or all instances of DOG, there exists an event such that the instance barks

Dogs bark

Gen e : Exists x : dog(x) & extant(x,e)
bark(x,e)

For some/all/most events such that a dog exists in that event, then that dog barks

Dogs bark

Gen k : "dog-kind"(k)
"has-ability-to-bark"(k)

Some/all/most kinds of dog (beagles, terriers, …) have the ability to bark

(Gen) Dogs bark

$$\text{G}\small{\text{EN}}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\textbf{Restrictor}[x_1, ..., x_n];\ \textbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

Some, all, many, most, at least a few, …

P("Dogs" | "bark") == Gen

(Gen) Dogs bark

$$\textsc{Gen}\ [x_1, ..., x_n; y_1, ..., y_m]\ (\mathbf{Restrictor}[x_1, ..., x_n];\ \mathbf{Matrix}[\{x_1\}, ..., \{x_n\}; y_1, ..., y_m])$$

Some, all, many, most, at least a few, …

Two approaches

Recognizing patterns of situation descriptions

=> allows for deriving generic statements

Directly interpreting generic sentences

Two approaches

Recognizing patterns of situation descriptions

=> allows for deriving generic statements

Directly interpreting generic sentences

=> … allows for deriving generic statements

Two approaches, One underlying problem

Recognizing patterns of situation descriptions

=> allows for deriving generic statements

Directly interpreting generic sentences

=> … allows for deriving generic statements

Now: simply transform generic sentences into probabilistic statements about the world

Johns Hopkins University

Ducks lay eggs

Sea turtles live long lives

Ducks lay eggs (less than 50% do)

Sea turtles live long lives (most die young)

Van Durme (2009)

The core task of knowledge acquisition is to first discover basic semantic relations along with their arguments from text, and then to abstract beyond argument instances to more general statements about the conceptual classes we expect the predications to hold over.

Van Durme (2009)

*Future work on refining these basic relations needs to address the issues of **determining proper quantifier strength**, and what, if anything, is being left implicit in a rule's **quantifier domain restrictor**. [...] the first problem involves determining whether a given rule, e.g., A dog may bark, should be strongly (Most dogs bark) or weakly (Some dogs bark) quantified, or something in between [...] Restating the second problem, which is closely related to the first: determine the proper contexts in which a given rule is applicable.*

JOHNS HOPKINS
UNIVERSITY

Scalable probabilities/inference

Streaming Pointwise Mutual Information. Van Durme and Lall. NIPS 2009.

Probabilistic Counting with Randomized Storage. Van Durme and Lall. IJCAI 2009.

Online Generation of Locality Sensitive Hash Signatures. Van Durme and Lall. ACL 2010.

Efficient Online Locality Sensitive Hashing via Reservoir Counting. Van Durme and Lall. ACL 2011.

Streaming Analysis of Discourse Participants. Benjamin Van Durme. EMNLP 2012.

Space Efficiencies in Discourse Modeling via Conditional Random Sampling. Kjersten and Van Durme. NAACL 2012.

Shared Components Topic Models. Gormley, Dredze, Van Durme, and Eisner. NAACL 2012.

Particle Filter Rejuvenation and Latent Dirichlet Allocation. May, Clemmer and Van Durme. ACL 2014.

Exponential Reservoir Sampling for Streaming Language Models. Osborne, Lall and Van Durme. ACL 2014.

Information Extraction over Structured Data: Question Answering with Freebase. Yao and Van Durme. ACL 2014.

Bayesian Script Induction. Ferraro and Van Durme (in progress).

Wide Coverage Embeddings. Rastogi, Van Durme and Arora (in progress).

Hierarchical Properties. May and Van Durme (in progress).

Questions