Learning Open Domain Knowledge from Text

Gabor Angeli

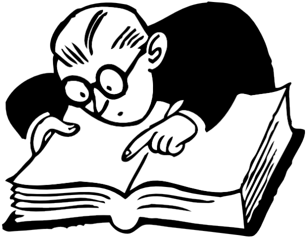
Stanford University

November 5, 2015

Learning "*Knowledge*?"

Learning "*Knowledge*?"

Knowledge = *True Statements*



Harder For Computers Than Humans

...for a human

Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School.

...for a computer

Harder For Computers Than Humans

...for a human

Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School.

...for a computer

Rattled for Austin, Alaska, Jesus is the mouse in Microsoft Google but Facebook Twitter Snapchat.

Harder For Computers Than Humans

...for a human

Born in Honolulu, Hawaii, Obama is a graduate of Columbia University and Harvard Law School.

...for a computer

Rattled *for Austin, Alaska, Jesus is the mouse in Microsoft Google but Facebook Twitter Snapchat.*

- Need to learn lexical items.

Harder For Computers Than Humans

...for a human

*Born in **Honolulu**, Hawaii, **Obama** is a graduate of Columbia University and Harvard Law School.*

...for a computer

Rattled for Austin, Alaska, Jesus is the mouse in Microsoft Google but Facebook Twitter Snapchat.

- Need to learn lexical items.
- Syntax is often non-trivial.

Harder For Computers Than Humans

...for a human

Born *in Honolulu, Hawaii, Obama is a* ***graduate*** *of Columbia University and Harvard Law School.*

...for a computer

Rattled for Austin, Alaska, Jesus is the mouse in Microsoft Google but Facebook Twitter Snapchat.

- Need to learn lexical items.
- Syntax is often non-trivial.
- Many "facts" in same sentence.

How do we Represent Knowledge?

Unstructured Text


The New York Times
The Brisket Speaks
With a Texas Accent

$\Longrightarrow$

Fixed-Schema Knowledge Bases

Barack Obama

44th President of the United States

Personal details	
Born	Barack Hussein Obama II August 4, 1961 (age 52) Honolulu, Hawaii, U.S.
Political party	Democratic
Spouse(s)	Michelle LaVaughn Robinson (m. 1992–present)
Children	Malia Ann Obama (b. 1998) Natasha Obama (b. 2001)

How do we Represent Knowledge?

Unstructured Text

$\Longrightarrow$

Fixed-Schema Knowledge Bases

(OBAMA; born_in; HONOLULU)
(OBAMA; born_in; HAWAII)
(OBAMA; born_on; 1961-8-4)
(OBAMA; spouse; MICHELLE)
(OBAMA; children; MALIA)
(OBAMA; children; NATASHA)

How do we Represent Knowledge?

Active area of research:

- Supervised relation extractors [Doddington et al., 2004, Surdeanu and Ciaramita, 2007].
- Distantly supervised extractors [Wu and Weld, 2007, Mintz et al., 2009].
- Weakly+distantly supervised extractors [Hoffmann et al., 2011, Surdeanu et al., 2012].
- Partially+weakly+distantly supervised extractors [Angeli et al., 2014a, Angeli et al., 2014b].

More to Life Than Fixed Relation Schema

How do we Represent Knowledge?

Unstructured Text

1. **Fixed-Schema Knowledge Bases**
2. **Open-Domain KBs (Open IE)**

$\Longrightarrow$ $(\textsf{SUBJECT};\ \text{relation};\ \textsf{OBJECT})$

How do we Represent Knowledge?

Unstructured Text

$\Longrightarrow$

1. **Fixed-Schema Knowledge Bases**
2. **Open-Domain KBs (Open IE)**

(CATS; have; TAILS)
(RABBITS; eat; CARROTS)
(OBAMA; enjoys playing; BASKETBALL)

Unstructured Text

1. **Fixed-Schema Knowledge Bases**
2. **Open-Domain KBs (Open IE)**
3. **Unstructured Text**

$\Longrightarrow$

cats have tails
rabbits eat carrots
Obama enjoys playing basketball

How do we Represent Knowledge?

Unstructured Text

1. **Fixed-Schema Knowledge Bases**
2. **Open-Domain KBs (Open IE)**
3. **Unstructured Text**

$\Longrightarrow$

cats have tails
rabbits eat carrots
Obama enjoys playing basketball

A graduated cylinder is best to measure the volume of a liquid

This Thesis

Store Information as Text (easier)
Query Information as Text (hard!)

This Thesis

Store Information as Text (easier)
Query Information as Text (hard!)

Build a system that:

- **Takes as input** a candidate textual statement.
- **Produces as output** the truth of that statement.

This Thesis

Store Information as Text (easier)
Query Information as Text (hard!)

Build a system that:

Takes as input a candidate textual statement.
Produces as output the truth of that statement.

- Generalizes Fixed-Schema KBs
  ✓ *Obama was born in Hawaii*
  ✗ *Obama was born in Kenya*

This Thesis

Store Information as Text (easier)
Query Information as Text (hard!)

Build a system that:

Takes as input a candidate textual statement.
Produces as output the truth of that statement.

- Generalizes Fixed-Schema KBs
  ✓ *Obama was born in Hawaii*
  ✗ *Obama was born in Kenya*
- Generalizes Open IE
  ✓ *Rabbits eat carrots*

This Thesis

Store Information as Text (easier)
Query Information as Text (hard!)

Build a system that:

Takes as input a candidate textual statement.
Produces as output the truth of that statement.

- Generalizes Fixed-Schema KBs
  ✓ *Obama was born in Hawaii*
  ✗ *Obama was born in Kenya*
- Generalizes Open IE
  ✓ *Rabbits eat carrots*
- More precise than web search
  ✗ *A stopwatch is best to measure the volume of a liquid.*

Prior Work

Flexibility

Prior Work

Relation Extraction: (BARACK OBAMA; born_in; ???)

Flexibility

[Hoffmann et al., 2011, Surdeanu et al., 2012, Angeli et al., 2014b]

Prior Work

Open Relation Extraction: (RABBITS; eat; ???)

Flexibility

[Banko et al., 2007, Fader et al., 2011, Mausam et al., 2012]

Prior Work

Entailment: If *a watch measures time*, does *it measure volume*?

Flexibility

[Glickman et al., 2006, MacCartney, 2009]

Textual Entailment

Single Premise:

Mitsubishi Motors Corp.'s new vehicle sales in the US fell 46 percent in June.

Single Hypothesis:

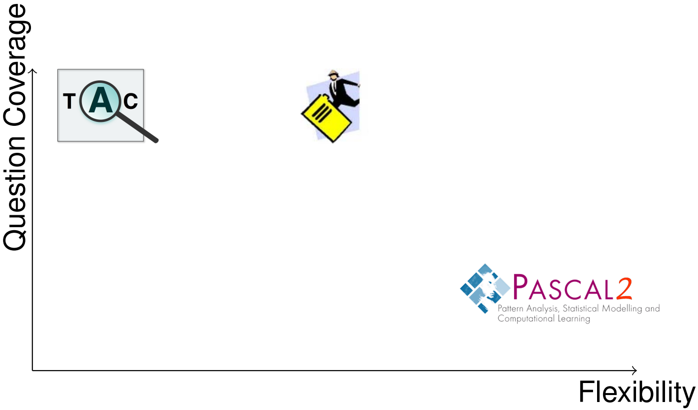
Mitsubishi sales rose 46 percent.

Classification Task: If you accept the premise, would you accept the hypothesis?

Prior Work

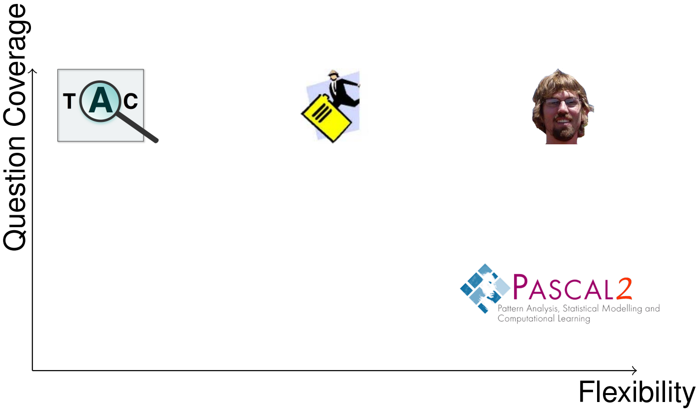
Flexibility

Prior Work


Question Coverage
TAC
PASCAL2
Pattern Analysis, Statistical Modelling and Computational Learning
Flexibility

Prior Work

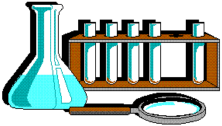
This Thesis: Formal reasoning with text over large corpora



Roadmap

Common Sense Reasoning:
Cats have tails
[Angeli and Manning, 2013, Angeli and Manning, 2014]

Complex premises:
Born in Hawaii, Obama is a graduate of Columbia
[Angeli et al., 2015]

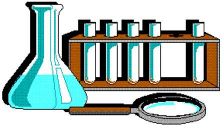


Lexical + Logical Reasoning:
A graduated cylinder would be best to measure the volume of a liquid

Roadmap

Common Sense Reasoning:
Cats have tails
[Angeli and Manning, 2013, Angeli and Manning, 2014]

Complex premises:
Born in Hawaii, Obama is a graduate of Columbia
[Angeli et al., 2015]



Lexical + Logical Reasoning:
A graduated cylinder would be best to measure the volume of a liquid

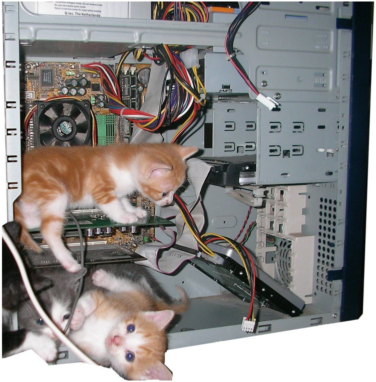
Reasoning About Common Sense Facts

✓ *Kittens play with yarn* ✗ *Kittens play with computers*

Reasoning About Common Sense Facts

✓ *Kittens play with yarn*

✗ *Kittens play with computers*

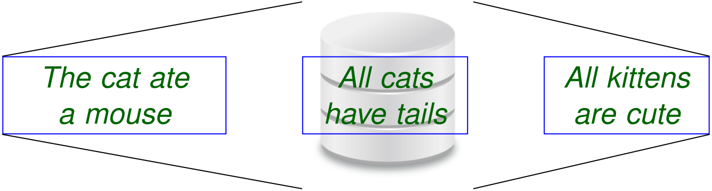


Common Sense Reasoning for NLP

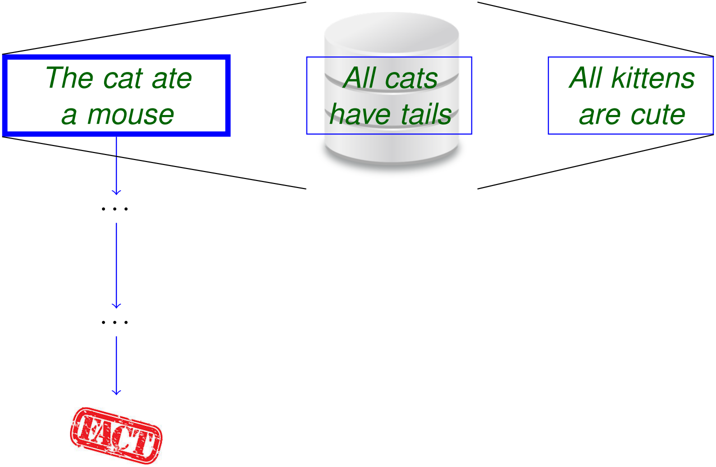
They ate the pizza with anchovies

Start with a large knowledge base

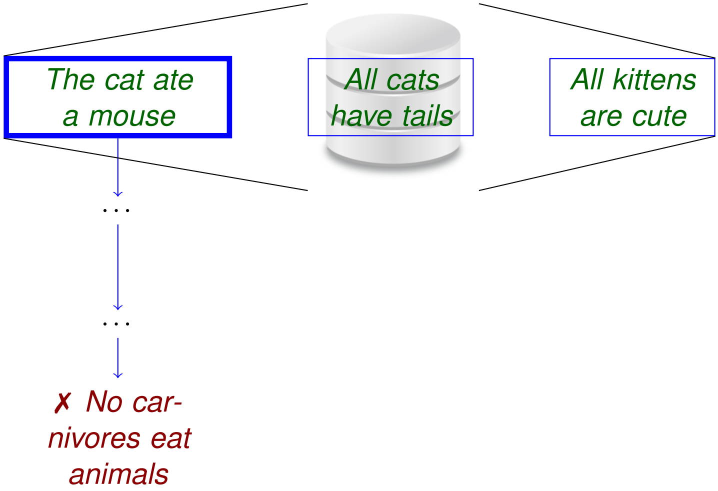
Start with a large knowledge base



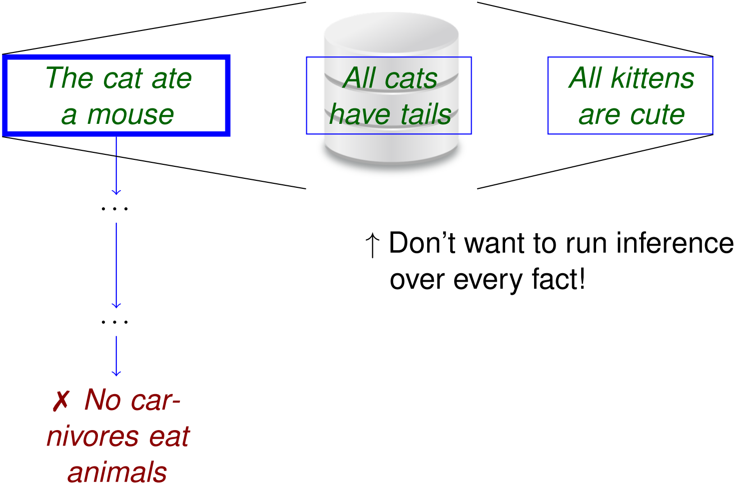
Infer new facts...



Infer new facts...



Infer new facts...



Infer new facts...

The cat ate a mouse

All cats have tails

All kittens are cute

. . .

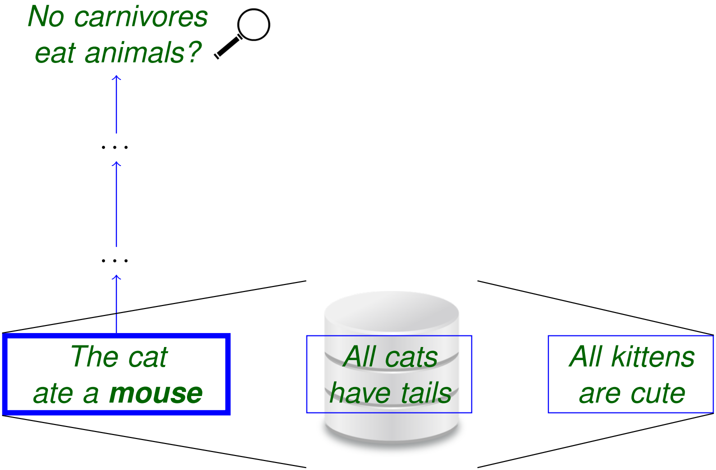
. . .

✗ *No carnivores eat animals*

↑ Don't want to run inference over every fact!

← Don't want to store all of these!

Infer new facts...on demand from a query...



...Using text as the meaning representation...

No carnivores eat animals?

***The** carnivores eat animals*

*The **cat** eats animals*

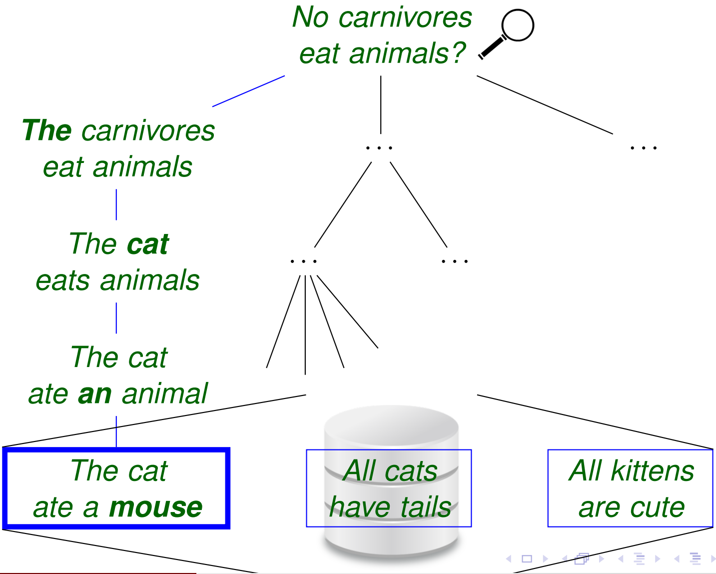
*The cat ate **an** animal*

*The cat ate a **mouse***

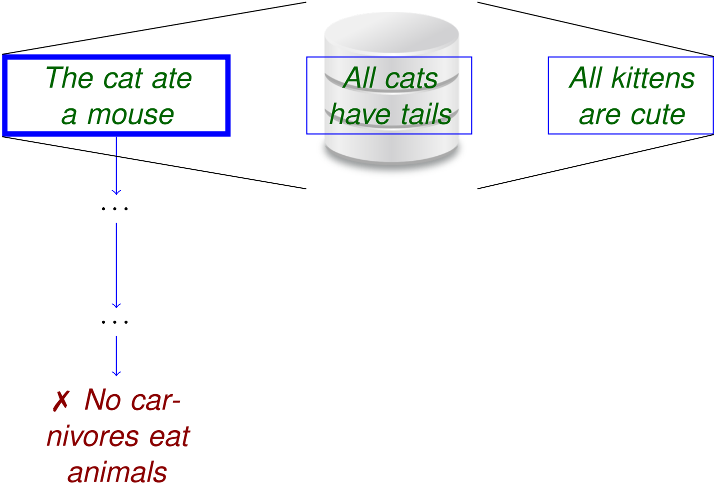
All cats have tails

All kittens are cute

...Without aligning to any particular premise.



Infer new facts...



We're Doing Logical Inference

The cat ate a mouse $\vDash \neg$ *No carnivores eat animals*

We're Doing Logical Inference

The cat ate a mouse $\models \neg$ *No carnivores eat animals*

Recall: Inference on every query: **speed is important!**

We're Doing Logical Inference

The cat ate a mouse $\vDash \neg$ *No carnivores eat animals*

Recall: Inference on every query: **speed is important!**

Recall: Both premise and query are sentences.

We’re Doing Logical Inference

The cat ate a mouse $\models \neg$ *No carnivores eat animals*

Recall: Inference on every query: **speed is important!**

Recall: Both premise and query are sentences.

Detour: Let’s talk about logic!

First Order Logic is **Intractable**

Theorem Provers

- Propositional logic is already NP-complete!

First Order Logic is **Intractable**

Theorem Provers

- Propositional logic is already NP-complete!

Markov Logic Networks

- Grounding 3,800 rules takes 7 hours (Alchemy).

First Order Logic is **Intractable**

Theorem Provers

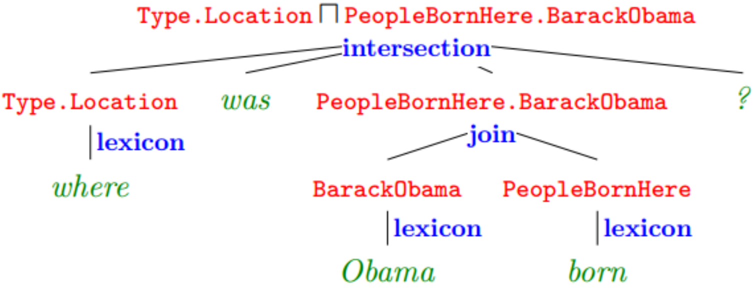
- Propositional logic is already NP-complete!

Markov Logic Networks

- Grounding 3,800 rules takes 7 hours (Alchemy).
- Add Chris Ré + decades of DB research: 106 seconds.
- ... but still slow for open-domain inference.

First order logic is **an unnatural language**

$$\exists x\,(\text{location}(x) \wedge \text{born\_in}(\text{Obama}, x))$$



[Berant et al., 2013]

First Order Logic is **Inexpressive**

Some people think that Obama was born in Kenya.

- Second order logic:
  $\exists x \exists P [P = \mathrm{born} \land \textbf{think(x, P)} \land P(\mathrm{Obama}, \mathrm{Kenya})]$

First Order Logic is **Inexpressive**

Some people think that Obama was born in Kenya.

- Second order logic:
  $\exists x \exists P [P = \text{born} \wedge \textbf{think(x, P)} \wedge P(\text{Obama}, \text{Kenya})]$
- But, can still infer: *Some people think that Obama is from Kenya.*

First Order Logic is **Inexpressive**

Some people think that Obama was born in Kenya.

- Second order logic:
  $\exists x \exists P [P = \text{born} \wedge \textbf{think(x, P)} \wedge P(\text{Obama}, \text{Kenya})]$
- But, can still infer: *Some people think that Obama is from Kenya.*

Most students who learned a foreign language learned it at a university.

- *Most* is not a first-order quantifier.
- Scoping ambiguities everywhere!

First Order Logic is **Inexpressive**

Some people think that Obama was born in Kenya.

- Second order logic:
  $\exists x \exists P [P = \mathrm{born} \wedge \mathbf{think(x, P)} \wedge P(\mathrm{Obama}, \mathrm{Kenya})]$
- But, can still infer: *Some people think that Obama is from Kenya.*

Most students who learned a foreign language learned it at a university.

- *Most* is not a first-order quantifier.
- Scoping ambiguities everywhere!
- But, can still infer: *Most students learned it at a school.*

Natural Logic

[Sánchez Valencia, 1991, MacCartney and Manning, 2008, Icard and Moss, 2014]

Natural Logic

Does a given mutation to a sentence preserve its truth?

Does a given mutation to a sentence preserve its truth?

Logic over natural language

- *Instantaneous* and *perfect* semantic parsing!
- Plays nice with lexical methods

Natural Logic

Does a given mutation to a sentence preserve its truth?

Logic over natural language

- *Instantaneous* and *perfect* semantic parsing!
- Plays nice with lexical methods

Tractable

- Polynomial time entailment checking [MacCartney and Manning, 2008].

Natural Logic

Does a given mutation to a sentence preserve its truth?

Logic over natural language

- *Instantaneous* and *perfect* semantic parsing!
- Plays nice with lexical methods

Tractable

- Polynomial time entailment checking [MacCartney and Manning, 2008].

Expressive (for common inferences)

- Second-order phenomena; *most*; quantifier scoping

Does a given mutation to a sentence preserve its truth?

Logic over natural language

- *Instantaneous* and *perfect* semantic parsing!
- Plays nice with lexical methods

Tractable

- Polynomial time entailment checking [MacCartney and Manning, 2008].

Expressive (for common inferences)

- Second-order phenomena; *most*; quantifier scoping
- No free lunch: shallow quantification; single-premise only

Natural Logic as Syllogisms

s/Natural Logic/Syllogistic Reasoning/g

Some cat ate a mouse
(all mice are rodents)
$\therefore$ *Some cat ate a **rodent***

Natural Logic as Syllogisms

s/Natural Logic/Syllogistic Reasoning/g

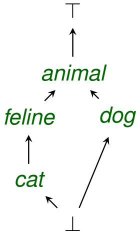
Some cat ate a mouse
(all mice are rodents)
$\therefore$ *Some cat ate a **rodent***

Beyond syllogisms

- General-purpose logic
  - Compositional grammar
  - Arbitrary quantifiers
- Model-theoretic soundness + completeness proof [Icard and Moss, 2014]

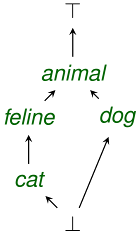
Natural Logic and Polarity

Treat hypernymy as a *partial order*.



Natural Logic and Polarity

Treat hypernymy as a *partial order*.



Polarity is the direction a lexical item can move in the ordering.

animal

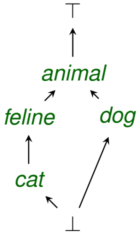
feline

$\boxed{\text{cat}}$

house cat

Natural Logic and Polarity

Treat hypernymy as a *partial order*.



***Polarity* is the direction a lexical item can move in the ordering.**

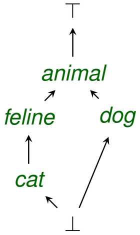
animal

feline

$\boxed{\uparrow \text{cat}}$

house cat

Treat hypernymy as a *partial order*.



***Polarity* is the direction a lexical item can move in the ordering.**

living thing

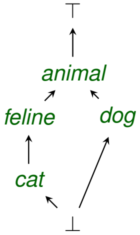
animal

$\boxed{\uparrow \text{feline}}$

cat

Natural Logic and Polarity

Treat hypernymy as a *partial order*.



***Polarity* is the direction a lexical item can move in the ordering.**

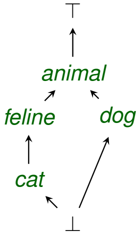
thing

living thing

↑ animal

feline

Treat hypernymy as a *partial order*.



***Polarity* is the direction a lexical item can move in the ordering.**

thing

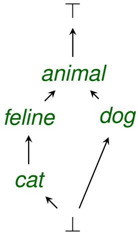
living thing

$\boxed{\downarrow \text{animal}}$

feline

Natural Logic and Polarity

Treat hypernymy as a *partial order*.



Polarity **is the direction a lexical item can move in the ordering.**

living thing

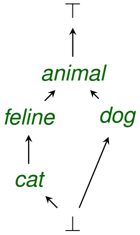
animal

$\boxed{\downarrow \text{feline}}$

cat

Natural Logic and Polarity

Treat hypernymy as a *partial order*.



***Polarity* is the direction a lexical item can move in the ordering.**

animal

feline

$\boxed{\downarrow \text{cat}}$

house cat

An Example Inference

Quantifiers determines the *polarity* (↑ or ↓) of words.

	carnivores		placentals
	felines	consume	rodents
$\uparrow$ **All**$_{\downarrow\uparrow}$	$\downarrow$ cats	$\uparrow$ eat	$\uparrow$ mice
	house cats	slurp	fieldmice

An Example Inference

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

↑ All$_{\downarrow\uparrow}$	↓ house cats	↑ eat	↑ mice
	felines		placentals
	cats	consume	rodents
	kitties	slurp	fieldmice

An Example Inference

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

↑ **All**$_{\downarrow\uparrow}$	↓ house cats	↑ consume	↑ mice
	felines		placentals
	cats		rodents
	kitties	eat	fieldmice

An Example Inference

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

	felines		rodents
	cats		mice
$\uparrow$ **All**$_{\downarrow\uparrow}$	$\downarrow$ house cats	$\uparrow$ consume	$\uparrow$ fieldmice
	kitties	eat	pine vole

An Example Inference

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

	felines		placentals
	cats		rodents
↑ **All**$_{\downarrow\uparrow}$	↓ house cats	↑ consume	↑ mice
	kitties	eat	fieldmice

An Example Inference

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

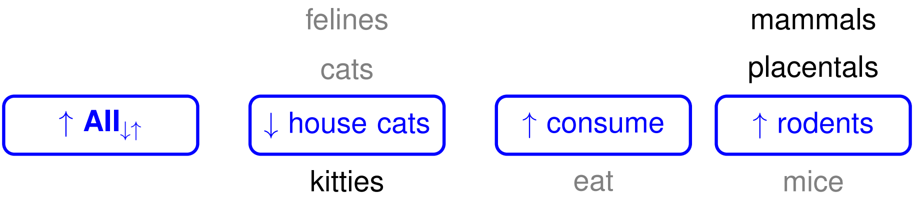
	felines		mammals
	cats		placentals
↑ **All**$_{\downarrow\uparrow}$	↓ house cats	↑ consume	↑ rodents
	kitties	eat	mice

An Example Inference

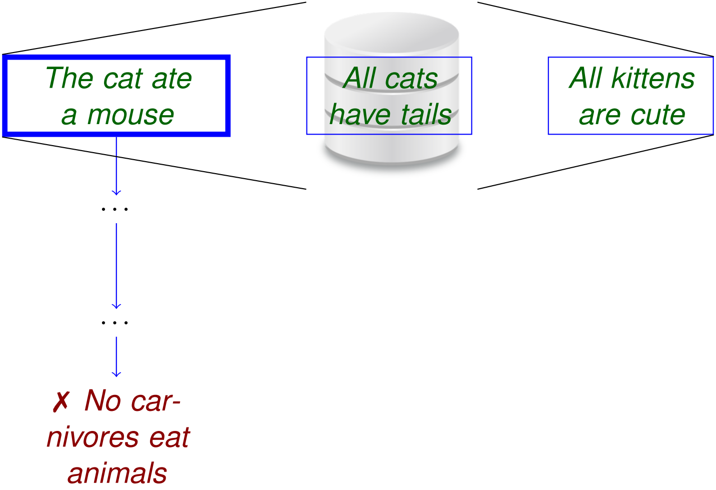
Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

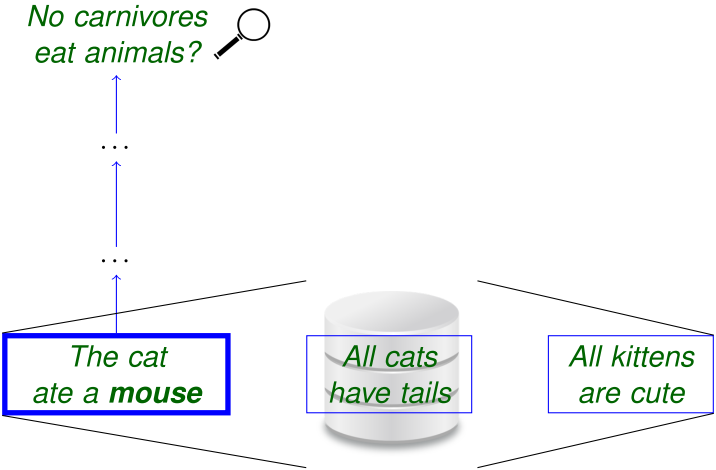
Inference is reversible.



Infer new facts...



Infer new facts...on demand from a query...



...Using text as the meaning representation...

No carnivores eat animals?

__The__ carnivores eat animals

The __cat__ eats animals

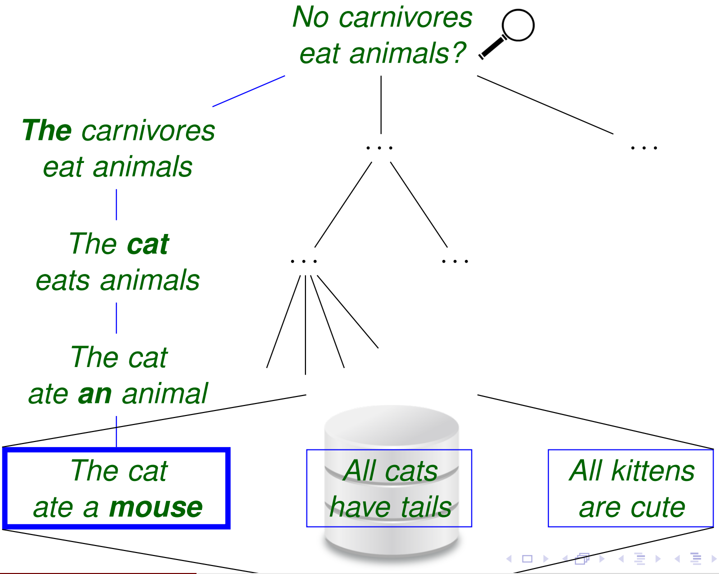
The cat ate __an__ animal

The cat ate a __mouse__

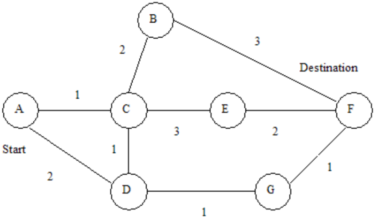
All cats have tails

All kittens are cute

...Without aligning to any particular premise.

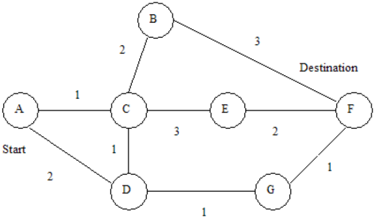


Natural Logic Inference is Search



Nodes (*fact*, truth maintained $\in \{\text{true}, \text{false}\}$)

Natural Logic Inference is Search

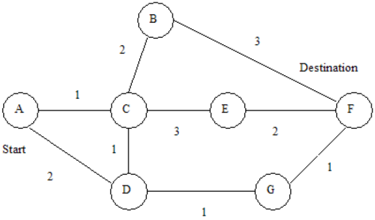


Nodes (*fact*, truth maintained $\in \{\text{true}, \text{false}\}$)

Start Node (*query fact*, ✓ *true*)

End Nodes *any known fact*

Natural Logic Inference is Search



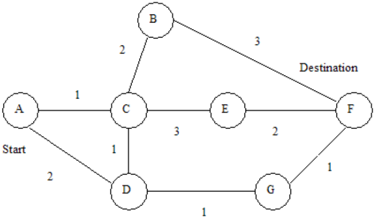
Nodes (*fact*, truth maintained $\in \{\text{true}, \text{false}\}$)

Start Node (*query fact*, ✓ *true*)

End Nodes *any known fact*

Edges Mutations of the current fact

Natural Logic Inference is Search



Nodes	(*fact*, truth maintained $\in \{\text{true}, \text{false}\}$)
Start Node	(*query fact*, ✓ *true*)
End Nodes	*any known fact*
Edges	Mutations of the current fact
Edge Costs	How "wrong" an inference step is (learned)

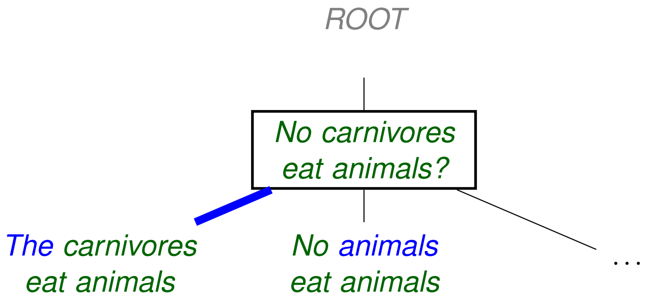
An Example Search

Shorthand for a node:

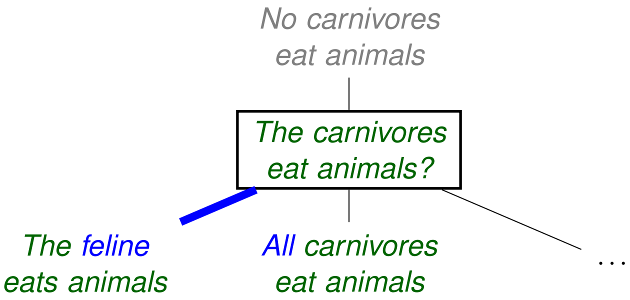
	organism		living thing
	animal	consume	organism
↑ **No**$_{\downarrow\downarrow}$	↓ carnivores	↓ eat	↓ animals
	felines	slurp	chordate

$$\boxed{\textit{No carnivores eat animals?}}$$

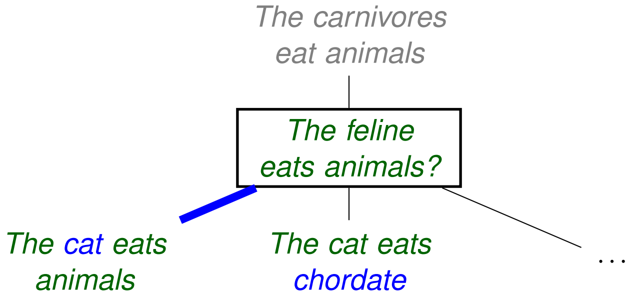
An Example Search



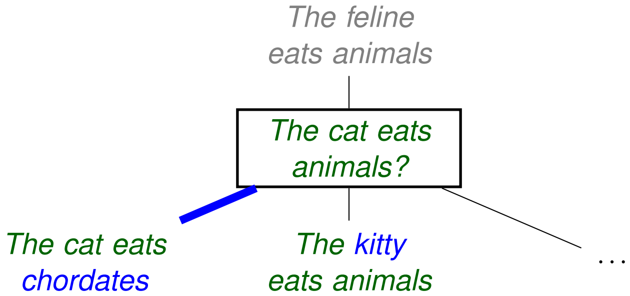
An Example Search



An Example Search



An Example Search



An Example Search

The cat eats animals

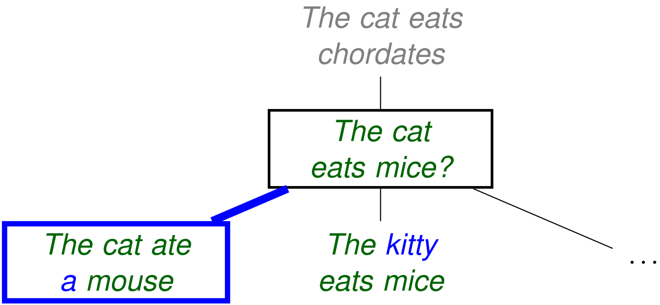
The cat eats chordates?

The cat eats mice

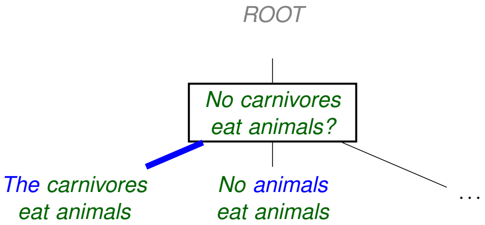
The cat eats dogs

...

An Example Search

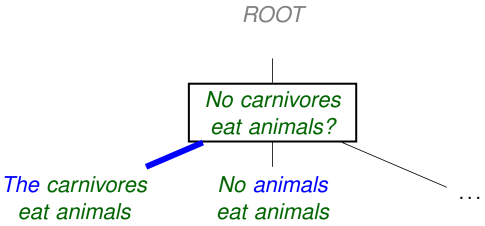


An Example Search (with edges)



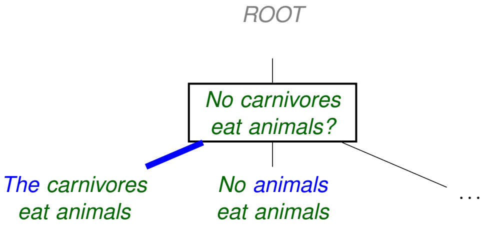
Template	Instance	Edge
Operator Negate		

An Example Search (with edges)



Template	Instance	Edge
Operator Negate	*No* $\rightarrow$ *The*	

An Example Search (with edges)



Template	**Instance**	**Edge**
Operator Negate	*No* $\rightarrow$ *The*	*No carnivores eat animals* $\rightarrow$ *The carnivores eat animals*

Edge Templates

Template	Instance
Hypernym	*animal* $\rightarrow$ *cat*
Hyponym	*cat* $\rightarrow$ *animal*
Antonym	*good* $\rightarrow$ *bad*
Synonym	*cat* $\rightarrow$ *true cat*
Add Word	*cat* $\rightarrow \cdot$
Delete Word	$\cdot \rightarrow$ *cat*
Operator Weaken	*some* $\rightarrow$ *all*
Operator Strengthen	*all* $\rightarrow$ *some*
Operator Negate	*all* $\rightarrow$ *no*
Operator Synonym	*all* $\rightarrow$ *every*
Nearest Neighbor	*cat* $\rightarrow$ *dog*

"Soft" Natural Logic

Want to make likely (but not certain) inferences.

- Same motivation as Markov Logic, Probabilistic Soft Logic, etc.

"Soft" Natural Logic

Want to make likely (but not certain) inferences.

- Same motivation as Markov Logic, Probabilistic Soft Logic, etc.
- Each *edge template* has a cost $\theta \geq 0$.

“Soft” Natural Logic

Want to make likely (but not certain) inferences.

- Same motivation as Markov Logic, Probabilistic Soft Logic, etc.
- Each *edge template* has a cost $\theta \geq 0$.

Detail: Variation among *edge instances* of a template.

- WordNet: *cat* $\rightarrow$ *feline* **vs.** *cup* $\rightarrow$ *container*.
- Nearest neighbors distance.
- Each *edge instance* has a distance f.

“Soft” Natural Logic

Want to make likely (but not certain) inferences.

- Same motivation as Markov Logic, Probabilistic Soft Logic, etc.
- Each *edge template* has a cost $\theta \geq 0$.

Detail: Variation among *edge instances* of a template.

- WordNet: *cat* $\rightarrow$ *feline* **vs.** *cup* $\rightarrow$ *container*.
- Nearest neighbors distance.
- Each *edge instance* has a distance f.

Cost of an edge is $\theta_i \cdot f_i$.

“Soft” Natural Logic

Want to make likely (but not certain) inferences.

- Same motivation as Markov Logic, Probabilistic Soft Logic, etc.
- Each *edge template* has a cost $\theta \geq 0$.

Detail: Variation among *edge instances* of a template.

- WordNet: *cat* $\rightarrow$ *feline* **vs.** *cup* $\rightarrow$ *container*.
- Nearest neighbors distance.
- Each *edge instance* has a distance f.

Cost of an edge is $\theta_i \cdot f_i$.
Cost of a path is $\theta \cdot \mathbf{f}$.

"Soft" Natural Logic

Want to make likely (but not certain) inferences.

- Same motivation as Markov Logic, Probabilistic Soft Logic, etc.
- Each *edge template* has a cost $\theta \geq 0$.

Detail: Variation among *edge instances* of a template.

- WordNet: *cat* $\rightarrow$ *feline* **vs.** *cup* $\rightarrow$ *container*.
- Nearest neighbors distance.
- Each *edge instance* has a distance f.

Cost of an edge is $\theta_i \cdot f_i$.
Cost of a path is $\theta \cdot \mathbf{f}$.
Can learn parameters θ.

Experiments

ConceptNet:

- A semi-curated collection of common-sense facts.
  - ✓ *not all birds can fly*
  - ✓ *noses are used to smell*
  - ✓ *nobody wants to die*
  - ✓ *music is used for pleasure*
- Negatives: ReVerb extractions marked false by Turkers.
- Small (1378 train / 1080 test), but fairly broad coverage.

Experiments

ConceptNet:

- A semi-curated collection of common-sense facts.

  ✓ *not all birds can fly*

  ✓ *noses are used to smell*

  ✓ *nobody wants to die*

  ✓ *music is used for pleasure*

- Negatives: ReVerb extractions marked false by Turkers.
- Small (1378 train / 1080 test), but fairly broad coverage.

Our Knowledge Base:

- 270 million lemmatized Ollie extractions.

ConceptNet Results

Systems

Direct Lookup: Lookup by lemmas.

Thesis: This thesis.

ConceptNet Results

Systems

Direct Lookup: Lookup by lemmas.

Thesis: This thesis.

Thesis - Lookup: Remove query facts from KB.

ConceptNet Results

Systems

Direct Lookup: Lookup by lemmas.

Thesis: This thesis.

Thesis - Lookup: Remove query facts from KB.

System	P	R
Direct Lookup	100.0	12.1

ConceptNet Results

Systems

Direct Lookup: Lookup by lemmas.

Thesis: This thesis.

Thesis - Lookup: Remove query facts from KB.

System	P	R
Direct Lookup	100.0	12.1
Thesis	90.6	49.1
Thesis - Lookup	88.8	40.1

ConceptNet Results

Systems

Direct Lookup: Lookup by lemmas.

Thesis: This thesis.

Thesis - Lookup: Remove query facts from KB.

System	P	R
Direct Lookup	100.0	**12.1**
Thesis	90.6	**49.1**
Thesis - Lookup	88.8	40.1

- 4x improvement in recall.

Success?

Not Yet!

The Internet doesn't speak in atomic utterances

Not Yet!

The Internet doesn't speak in atomic utterances

- Where was Obama born?
  - *Born in Hawaii, Obama is a graduate of Columbia University and Harvard Law School.*

Not Yet!

The Internet doesn't speak in atomic utterances

- Where was Obama born?
  - *Born in Hawaii, Obama is a graduate of Columbia University and Harvard Law School.*
    $\Rightarrow$ *Obama was born in Hawaii*.

Not Yet!

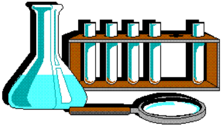
The Internet doesn't speak in atomic utterances

- Where was Obama born?
  - *Born in Hawaii, Obama is a graduate of Columbia University and Harvard Law School.*
    $\Rightarrow$ *Obama was born in Hawaii*.
- Let's store the inferred fact instead

Roadmap

Common Sense Reasoning:
Cats have tails
[Angeli and Manning, 2013, Angeli and Manning, 2014]

Complex premises:
Born in Hawaii, Obama is a graduate of Columbia
[Angeli et al., 2015]

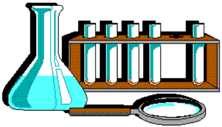


Lexical + Logical Reasoning:
A graduated cylinder would be best to measure the volume of a liquid

Roadmap

Common Sense Reasoning:
Cats have tails
[Angeli and Manning, 2013, Angeli and Manning, 2014]

Complex premises:
Born in Hawaii, Obama is a graduate of Columbia
[Angeli et al., 2015]

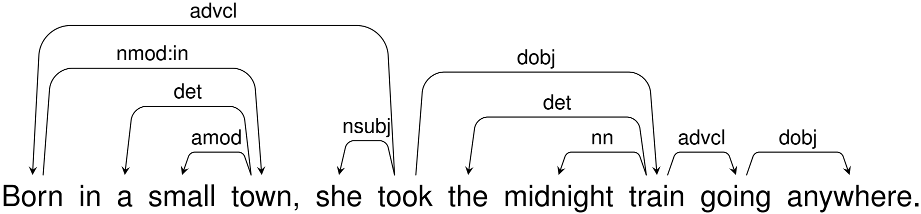


Lexical + Logical Reasoning:
A graduated cylinder would be best to measure the volume of a liquid

Atomic Clauses from Sentences

Input: Long sentence.
Born in a small town, she took the midnight train going anywhere.

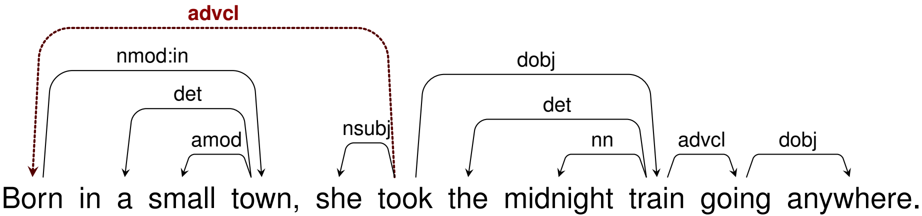
Output: Short clauses.
she was born in a small town.



Atomic Clauses from Sentences

Input: Long sentence.
Born in a small town, she took the midnight train going anywhere.

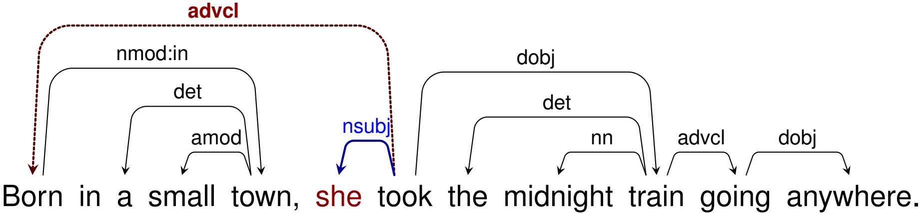
Output: Short clauses.
she was born in a small town.



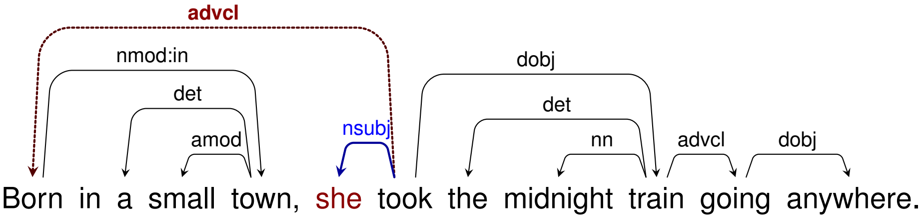
Atomic Clauses from Sentences

Input: Long sentence.
Born in a small town, she took the midnight train going anywhere.

Output: Short clauses.
she was born in a small town.

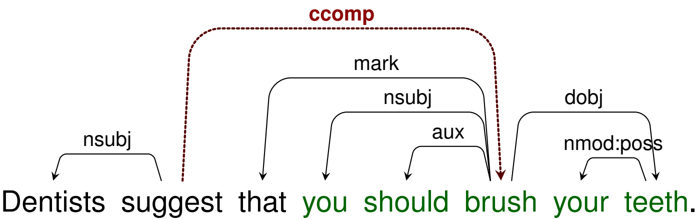


Clause Classifier



Input: Dependency arc.
Output: *Action* to take.

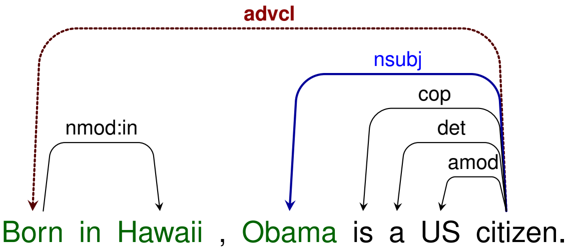
Clause Classifier



Input: Dependency arc.
Output: *Action* to take.

- **Yield** (*you should brush your teeth*)

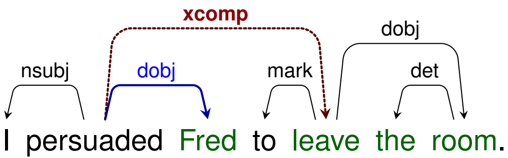
Clause Classifier



Input: Dependency arc.
Output: *Action* to take.

- **Yield** (*you should brush your teeth*)
- **Yield (Subject Controller)** (*Obama Born in Hawaii*)

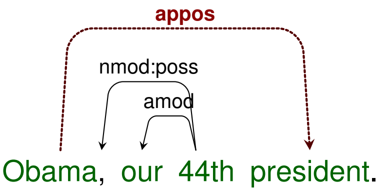
Clause Classifier



Input: Dependency arc.
Output: *Action* to take.

- **Yield** (*you should brush your teeth*)
- **Yield (Subject Controller)** (*Obama Born in Hawaii*)
- **Yield (Object Controller)** (*Fred leave the room*)

Clause Classifier



Input: Dependency arc.
Output: *Action* to take.

- **Yield** (*you should brush your teeth*)
- **Yield (Subject Controller)** (*Obama Born in Hawaii*)
- **Yield (Object Controller)** (*Fred leave the room*)
- **Yield (Parent Subject)** (*Obama is our 44th president*)

Classifier Training

Training Data Generation

1. Label the Penn Treebank with Open IE triples using traces.
2. Run exhaustive search over possible clause splits.

Classifier Training

Training Data Generation

1. Label the Penn Treebank with Open IE triples using traces.
2. Run exhaustive search over possible clause splits.
3. Positive Labels: A sequence of actions which yields a relation (33.5k examples).
   Negative Labels: All other sequences of actions (1.1M examples).

Classifier Training

Training Data Generation

1. Label the Penn Treebank with Open IE triples using traces.
2. Run exhaustive search over possible clause splits.
3. Positive Labels: A sequence of actions which yields a relation (33.5k examples).
   Negative Labels: All other sequences of actions (1.1M examples).

Features:

- Edge label; incoming edge label.
- Neighbors of governor + dependent; number of neighbors.
- Existence of subject/object edges at governor; dependent.
- POS tag of governor; dependent.

Maximally Shorten Clauses

Some strange, nuanced function:

*Heinz Fischer **of Austria***	$\Longrightarrow$	✓ *Heinz Fischer*
***United States president** Obama*	$\Longrightarrow$	✓ *Obama*
*All **young** rabbits drink milk*	$\not\Longrightarrow$	✗ *All rabbits drink milk*
*Some **young** rabbits drink milk*	$\Longrightarrow$	✓ *Some rabbits drink milk*
*Enemies give **fake** praise*	$\not\Longrightarrow$	✗ *Enemies give praise*
*Friends give **true** praise*	$\Longrightarrow$	✓ *Friends give praise*

Maximally Shorten Clauses

An entailment function:

*Heinz Fischer **of Austria***	$\Longrightarrow$	✓ *Heinz Fischer*
***United States president** Obama*	$\Longrightarrow$	✓ *Obama*
*All **young** rabbits drink milk*	$\not\Longrightarrow$	✗ *All rabbits drink milk*
*Some **young** rabbits drink milk*	$\Longrightarrow$	✓ *Some rabbits drink milk*
*Enemies give **fake** praise*	$\not\Longrightarrow$	✗ *Enemies give praise*
*Friends give **true** praise*	$\Longrightarrow$	✓ *Friends give praise*

Maximally Shorten Clauses

A natural logic entailment function:

*Heinz Fischer **of Austria***	$\Longrightarrow$	✓ *Heinz Fischer*
***United States president** Obama*	$\Longrightarrow$	✓ *Obama*
*All **young** rabbits drink milk*	$\not\Longrightarrow$	✗ *All rabbits drink milk*
*Some **young** rabbits drink milk*	$\Longrightarrow$	✓ *Some rabbits drink milk*
*Enemies give **fake** praise*	$\not\Longrightarrow$	✗ *Enemies give praise*
*Friends give **true** praise*	$\Longrightarrow$	✓ *Friends give praise*

Natural Logic For Clause Shortening

Quantifiers determines the *polarity* ($\uparrow$ or $\downarrow$) of words.

Mutations must respect *polarity*.

Polarity determines valid deletions.

	mammals		something
	rabbits	consume	liquid
$\uparrow$ **Some**$_{\uparrow\uparrow}$	$\uparrow$ young rabbits	$\uparrow$ drink	$\uparrow$ milk
	baby rabbits	slurp	Lucerne

Natural Logic For Clause Shortening

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

Polarity determines valid deletions.

	animals		something
	mammals	consume	liquid
↑ **Some**$_{\uparrow\uparrow}$	↑ rabbits	↑ drink	↑ milk
	young rabbits	slurp	Lucerne

Natural Logic For Clause Shortening

Quantifiers determines the *polarity* ($\uparrow$ or $\downarrow$) of words.

Mutations must respect *polarity*.

Polarity determines valid deletions.

	mammals		something
	rabbits	consume	liquid
$\uparrow$ **All**$_{\downarrow\uparrow}$	$\downarrow$ young rabbits	$\uparrow$ drink	$\uparrow$ milk
	baby rabbits	slurp	Lucerne

Natural Logic For Clause Shortening

Quantifiers determines the *polarity* (↑ or ↓) of words.

Mutations must respect *polarity*.

Polarity determines valid deletions.

	animals		something
	mammals	consume	liquid
$\uparrow$ **All**$_{\downarrow\uparrow}$	$\downarrow$ rabbits	$\uparrow$ drink	$\uparrow$ milk
	young rabbits	slurp	Lucerne

Bonus: Knowledge Base Triples

Heinz Fischer visited US $\Longrightarrow$ (HEINZ FISCHER; visited; US)

Bonus: Knowledge Base Triples

Heinz Fischer visited US	$\Longrightarrow$	(HEINZ FISCHER; visited; US)
Obama born in Hawaii	$\Longrightarrow$	(OBAMA; born in; HAWAII)

Bonus: Knowledge Base Triples

Heinz Fischer visited US	$\Longrightarrow$	(HEINZ FISCHER; visited; US)
Obama born in Hawaii	$\Longrightarrow$	(OBAMA; born in; HAWAII)
Cats are cute	$\Longrightarrow$	(CATS; are; CUTE)

Bonus: Knowledge Base Triples

Heinz Fischer visited US	$\Longrightarrow$	(HEINZ FISCHER; visited; US)
Obama born in Hawaii	$\Longrightarrow$	(OBAMA; born in; HAWAII)
Cats are cute	$\Longrightarrow$	(CATS; are; CUTE)
Cats are sitting next to dogs	$\Longrightarrow$	(CATS; are sitting next to; DOGS)

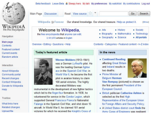
Bonus: Knowledge Base Triples

Heinz Fischer visited US	$\Longrightarrow$	(HEINZ FISCHER; visited; US)
Obama born in Hawaii	$\Longrightarrow$	(OBAMA; born in; HAWAII)
Cats are cute	$\Longrightarrow$	(CATS; are; CUTE)
Cats are sitting next to dogs	$\Longrightarrow$	(CATS; are sitting next to; DOGS)
	...	

5 dependency tree patterns (+ 8 nominal patterns)

Extrinsic Evaluation: Knowledge Base Population

Unstructured Text



$\Longrightarrow$

Structured Knowledge Base

Barack Obama

44th President of the United States

Personal details	
Born	Barack Hussein Obama II August 4, 1961 (age 52) Honolulu, Hawaii, U.S.
Political party	Democratic
Spouse(s)	Michelle LaVaughn Robinson (m. 1992–present)
Children	Malia Ann Obama (b. 1998) Natasha Obama (b. 2001)

Extrinsic Evaluation: Knowledge Base Population

Relation Extraction Task:

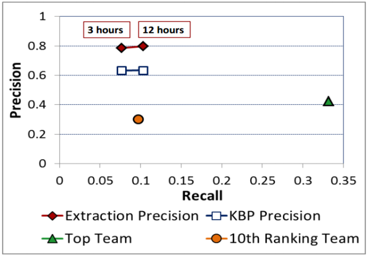
- Fixed schema of 41 relations.
- Precision: answers marked correct by humans.
- Recall: answers returned by any team (including LDC annotators).

Extrinsic Evaluation: Knowledge Base Population

Relation Extraction Task:

- Fixed schema of 41 relations.
- Precision: answers marked correct by humans.
- Recall: answers returned by any team (including LDC annotators).

Comparison: *Open Information Extraction to KBP Relations in 3 Hours*. [Soderland et al., 2013]



Map Triples to Structured Knowledge Base

KBP Relation	Text	PMI^2
Per:Date_Of_Birth	*be bear on*	1.83
	bear on	1.28
Per:Date_Of_Death	*die on*	0.70
	be assassinate on	0.65
Per:LOC_Of_Birth	*be bear in*	1.21
Per:LOC_Of_Death	**elect president of*	2.89
Per:Religion	*speak about*	0.67
	popular for	0.60
Per:Parents	*daughter of*	0.54
	son of	1.52
Per:LOC_Residence	*of*	1.48
	**independent from*	1.18

Results

TAC-KBP 2013 Slot Filling Challenge:

- End-to-end task: includes IR + consistency.
- **Precision:** facts LDC evaluators judged as correct.
  Recall: facts other teams (including LDC annotators) also found.

System	P	R	F_1
UW Submission	69.8	11.4	19.6
Ollie	57.7	11.8	19.6

Results

TAC-KBP 2013 Slot Filling Challenge:

- End-to-end task: includes IR + consistency.
- **Precision:** facts LDC evaluators judged as correct.
  Recall: facts other teams (including LDC annotators) also found.

System	P	R	F_1
UW Submission	69.8	11.4	19.6
Ollie	57.7	11.8	19.6
Our System	61.9	13.9	22.7

TAC-KBP 2013 Slot Filling Challenge:

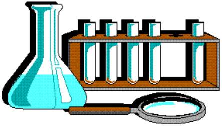
- End-to-end task: includes IR + consistency.
- **Precision:** facts LDC evaluators judged as correct.
  Recall: facts other teams (including LDC annotators) also found.

System	P	R	F_1
UW Submission	69.8	11.4	19.6
Ollie	57.7	11.8	19.6
Our System	61.9	13.9	22.7
Median Team			18.6
Our System +  + 	58.6	18.6	28.3
Top Team	45.7	35.8	40.2

Roadmap

Common Sense Reasoning:
Cats have tails
[Angeli and Manning, 2013, Angeli and Manning, 2014]

Complex premises:
Born in Hawaii, Obama is a graduate of Columbia
[Angeli et al., 2015]

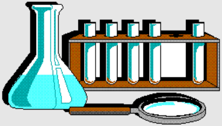


Lexical + Logical Reasoning:
A graduated cylinder would be best to measure the volume of a liquid

Roadmap

Common Sense Reasoning:
Cats have tails
[Angeli and Manning, 2013, Angeli and Manning, 2014]

Complex premises:
Born in Hawaii, Obama is a graduate of Columbia
[Angeli et al., 2015]



Lexical + Logical Reasoning:
A graduated cylinder would be best to measure the volume of a liquid

An Example Search

ROOT

Rain and snow are types of precipitation?

Rain and snow are forms of precipitation

Heavy rain and snow are types of precipitation

. . .

An Example Search

ROOT

Rain and snow are types of precipitation?

Rain and snow are forms of precipitation

Heavy rain and snow are types of precipitation

. . .

An Example Search

Rain and snow are types of precipitation

Rain and snow are forms of precipitation?

Rain and snow are forms of weather

Heavy rain and snow are forms of precipitation

. . .

An Example Search

Rain and snow are types of precipitation

Rain and snow are forms of precipitation?

Rain and snow are forms of weather

Heavy rain and snow are forms of precipitation

. . .

Forms of precipitation include rain and sleet

An Example Search

Rain and snow are types of precipitation

***Rain** and snow are **forms** of **precipitation**?*

Rain and snow are forms of weather

Heavy rain and snow are forms of precipitation

. . .

***Forms** of **precipitation** include **rain** and sleet*

Lexical Alignment Classifier

Forms of precipitation include rain and sleet

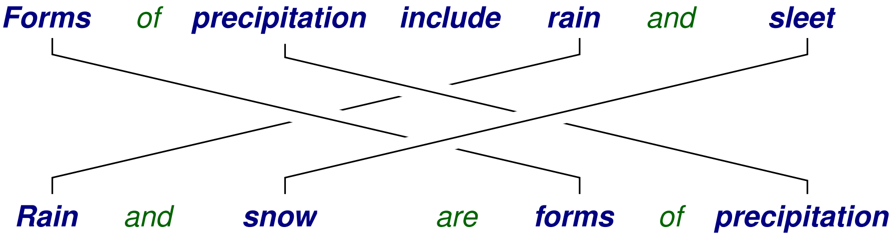
Rain and snow are forms of precipitation

Lexical Alignment Classifier

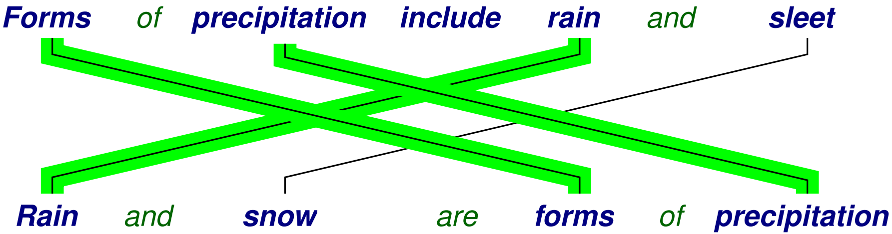
Forms *of* ***precipitation*** ***include*** ***rain*** *and* ***sleet***

Rain *and* ***snow*** *are* ***forms*** *of* ***precipitation***

Lexical Alignment Classifier



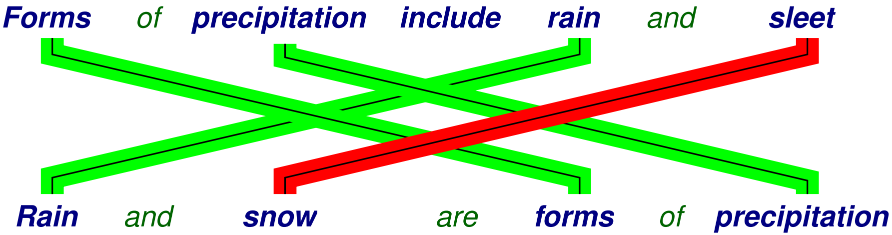
Lexical Alignment Classifier



Features

1. Matching words

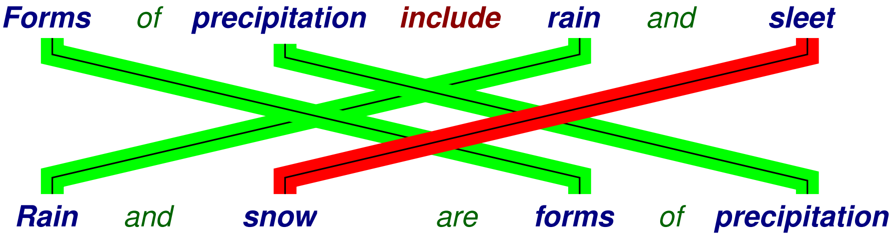
Lexical Alignment Classifier



Features

1. Matching words
2. Mismatched words

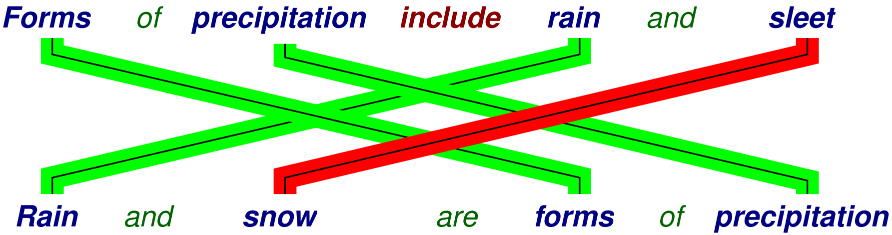
Lexical Alignment Classifier



Features

1. Matching words
2. Mismatched words
3. Unmatched words in premise/consequent

Lexical Alignment Classifier



Features

1. Matching words
2. Mismatched words
3. Unmatched words in premise/consequent

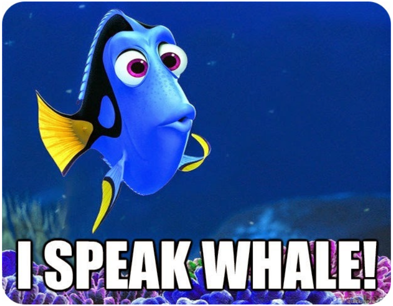
Competitive with Stanford RTE system (63% on RTE3)

Old Problem: Logic + Lexical Classifiers

FOL and lexical classifiers don't speak the same language

Old Problem: Logic + Lexical Classifiers

FOL and lexical classifiers don't speak the same language
...but natural logic does!



Big Picture

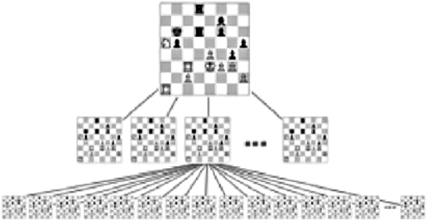
Run our usual search

1. If we find a premise, great!

Big Picture

Run our usual search

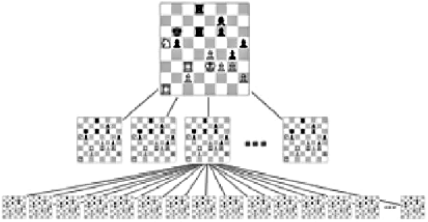
1. If we find a premise, great!
2. If not, use lexical classifier as an *evaluation function*



Big Picture

Run our usual search

1. If we find a premise, great!
2. If not, use lexical classifier as an *evaluation function*



Visit 1M nodes / second: We have to be fast!

Dissecting Our Classifier

Anatomy of a Classifier

- Features f (matching / mismatched / unmatched words)
- Weights w
- Entailment pair x

$$p(\text{entail} \mid x) = \frac{1}{1+\exp(-w^{\mathrm{T}} f(x))}$$

Dissecting Our Classifier

Anatomy of a Classifier

- Features f (matching / mismatched / unmatched words)
- Weights w
- Entailment pair x

$$p(\text{entail} \mid x) = \frac{1}{1+\exp(-w^{\mathrm{T}} f(x))}$$

$$p(\text{entail} \mid x) \text{ monotone w.r.t. } \left(w^{\mathrm{T}} f(x)\right)$$

Dissecting Our Classifier

Anatomy of a Classifier

- Features f (matching / mismatched / unmatched words)
- Weights w
- Entailment pair x

$$p(\text{entail} \mid x) = \frac{1}{1+\exp(-w^{\mathrm{T}} f(x))}$$

$$p(\text{entail} \mid x) \text{ monotone w.r.t. } \left(w^{\mathrm{T}} f(x)\right)$$

- Only need $w^{\mathrm{T}} f(x)$ during search to compute $\max p(\text{entail} \mid x)$
- $w^{\mathrm{T}} f(x)$ is our evaluation function

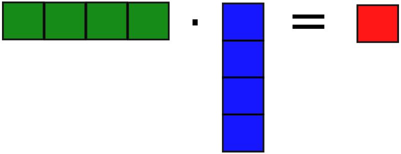
Incorporating our Evaluation Function

Anatomy of a Search Step

1. Mutate a word, or
2. Delete a word, or
3. Insert a word.

Each step updates a small number of features

$$w^{\mathrm{T}} f(x) = v$$



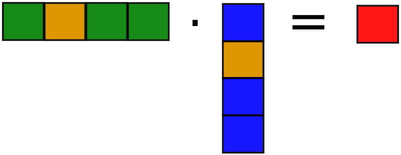
Incorporating our Evaluation Function

Anatomy of a Search Step

1. Mutate a word, or
2. Delete a word, or
3. Insert a word.

Each step updates a small number of features

$$w^{\mathrm{T}} f(x) = v$$



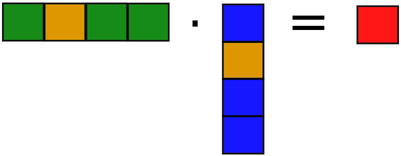
Incorporating our Evaluation Function

Anatomy of a Search Step

1. Mutate a word, or
2. Delete a word, or
3. Insert a word.

Each step updates a small number of features

$$v' = v - w_i \cdot f_i + w_i \cdot f_i$$



Why is this Important?

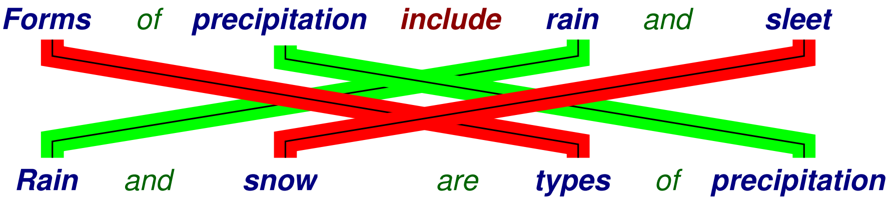
Faster Search ⇒ Deeper Reasoning

- **Speed:** Around 1M search states visited per second
- **Memory:** 32 byte search states

Speed: Don't re-featurize at every timestep.

Memory: Never store intermediate fact as String.

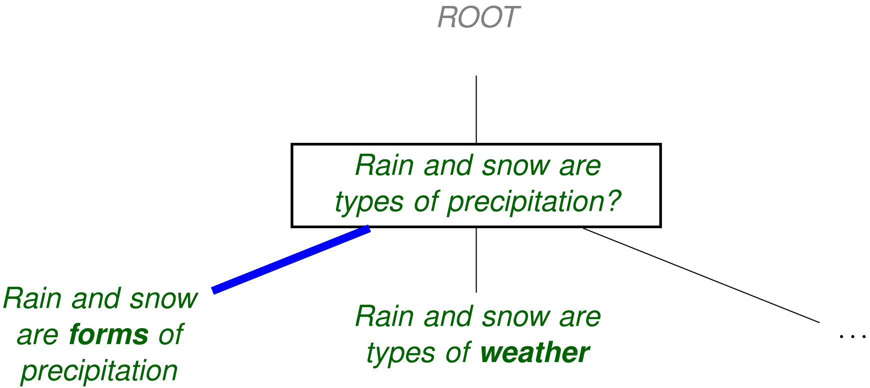
An Example Search



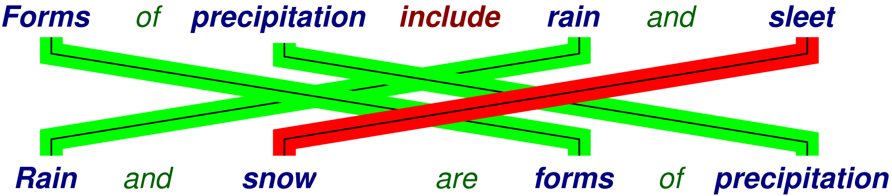
Score $w^{\mathrm{T}} f(x)$**:** -0.5

Feature	w	$f(x)$
Matching words	2.0	2
Mismatched words	-1.0	2
Unmatched premise	-0.5	1
Unmatched consequent	-0.75	0
Bias	-2.0	1

An Example Search



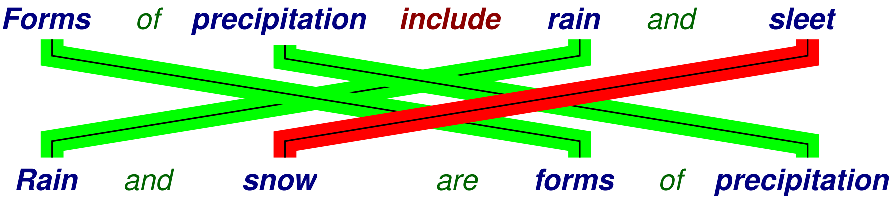
An Example Search



Score $w^{\mathrm{T}} f(x)$: -0.5 + 2 − -1

Feature	w	$f(x)$
Matching words	2.0	**3**
Mismatched words	-1.0	**1**
Unmatched premise	-0.5	1
Unmatched consequent	-0.75	0
Bias	-2.0	1

An Example Search



Score $w^{\mathrm{T}} f(x)$**:** 2.5

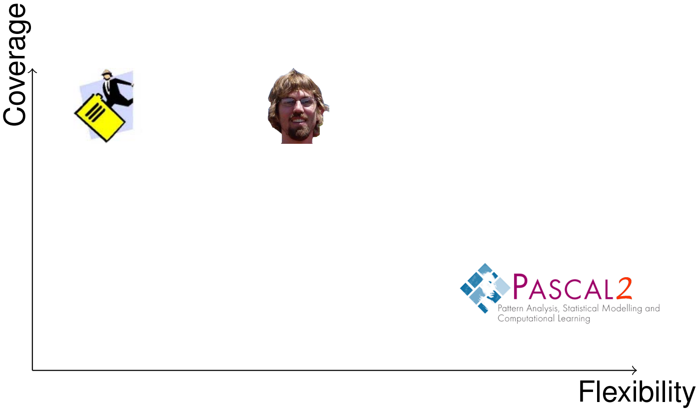
Feature	w	$f(x)$
Matching words	2.0	**3**
Mismatched words	-1.0	**1**
Unmatched premise	-0.5	1
Unmatched consequent	-0.75	0
Bias	-2.0	1

The Full System

Common Sense Reasoning

The Full System

+ Complex Premises



The Full System

+ Evaluation Function

The Full System

Common Sense Facts

- Natural logic inference as search
- Soft relaxation of “strict” inference
- 4x improvement in recall

The Full System

Common Sense Facts

- Natural logic inference as search
- Soft relaxation of “strict” inference
- 4x improvement in recall

Complex Premises

- Split the premise into atomic clauses
- Shorten each clause w/ natural logic
- 3 F_1 improvement on knowledge base population

The Full System

Common Sense Facts

- Natural logic inference as search
- Soft relaxation of “strict” inference
- 4x improvement in recall

Complex Premises

- Split the premise into atomic clauses
- Shorten each clause w/ natural logic
- 3 F_1 improvement on knowledge base population

Evaluation Function

- Use lexical classifier as evaluation function
- Detect likely entailment / contradictions

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

Which activity is an example of a good health habit?

(A) Watching television
(B) Smoking cigarettes
(C) Eating candy
(D) Exercising every day

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

Which activity is an example of a good health habit?

(A) Watching television
(B) Smoking cigarettes
(C) Eating candy
(D) Exercising every day

In our corpus:

- *Plasma TV's can display up to 16 million colors ... great for watching TV ... also make a good screen.*
- *Not smoking or drinking alcohol is good for health, regardless of whether clothing is worn or not.*
- *Eating candy for diner is an example of a poor health habit.*
- *Healthy is exercising*

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

System	Train	Test
KNOWBOT	45	
KNOWBOT (ORACLE)	57	

[Hixon et al., 2015]

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

System	Train	Test
KnowBot	45	
KnowBot (oracle)	57	
IR Baseline	49	
This Thesis	52	

[Hixon et al., 2015]

Multiple choice questions from real 4^{th} grade science exams

System	Train	Test
KNOWBOT	45	
KNOWBOT (ORACLE)	57	
IR Baseline	49	
This Thesis	52	
More Data + IR Baseline	62	
More Data + This Thesis	**65**	

[Hixon et al., 2015]

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

System	**Train**	**Test**
KNOWBOT	45	–
KNOWBOT (ORACLE)	57	–
IR Baseline	49	42
This Thesis	52	51
More Data + IR Baseline	62	58
More Data + This Thesis	**65**	**61**

[Hixon et al., 2015]

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

System	Train	Test
KNOWBOT	45	–
KNOWBOT (ORACLE)	57	–
IR Baseline	49	42
This Thesis	52	51
More Data + IR Baseline	62	58
More Data + This Thesis	65	61
This Thesis +  + 	**74**	**67**

[Hixon et al., 2015]

Solving 4^{th} Grade Science

Multiple choice questions from real 4^{th} grade science exams

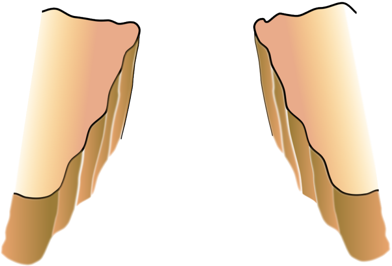
System	Train	Test
KNOWBOT	45	–
KNOWBOT (ORACLE)	57	–
IR Baseline	49	42
This Thesis	52	51
More Data + IR Baseline	62	58
More Data + This Thesis	65	61
This Thesis +  + 	**74**	**67**

We're able to pass 4^{th} grade science!

[Hixon et al., 2015]

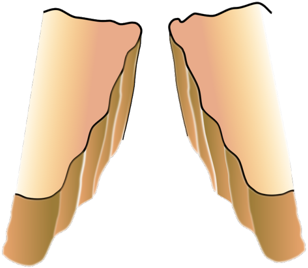
Remaining Work

First Order Logic **Lexical Methods**



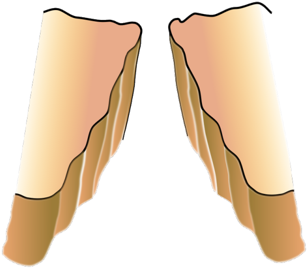
Remaining Work

Natural Logic **Lexical Methods**



Remaining Work

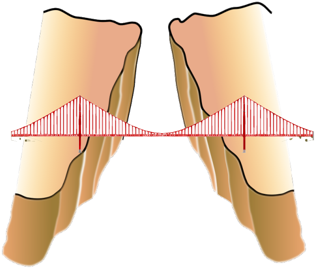
Natural Logic **Lexical Methods**



- Already useful for textual entailment [MacCartney and Manning, 2008, MacCartney, 2009]

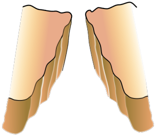
Remaining Work

Natural Logic **Lexical Methods**



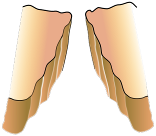
- Already useful for textual entailment [MacCartney and Manning, 2008, MacCartney, 2009]
- **This thesis:** Useful for question answering
  This thesis: We can bridge the two methods

Remaining Work



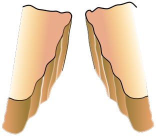
1. Encode logic in traditionally lexical representations [Bowman, 2014, Bowman et al., 2015]

Remaining Work



1. Encode logic in traditionally lexical representations [Bowman, 2014, Bowman et al., 2015]
2. Make natural logic more expressive

Remaining Work



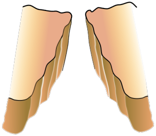
1. Encode logic in traditionally lexical representations [Bowman, 2014, Bowman et al., 2015]
2. Make natural logic more expressive
    - Propositional + Natural logics:

      Apples are red $\vee$ *Bananas are red*

      Bananas are not red

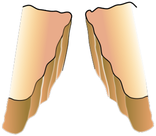
      $\therefore$ *Apples are red*

Remaining Work



1. Encode logic in traditionally lexical representations [Bowman, 2014, Bowman et al., 2015]
2. Make natural logic more expressive
    - Propositional + Natural logics:
      Apples are red $\vee$ *Bananas are red*
      Bananas are not red
      $\therefore$ *Apples are red*

Remaining Work



1. Encode logic in traditionally lexical representations [Bowman, 2014, Bowman et al., 2015]
2. Make natural logic more expressive
   - Propositional + Natural logics:
     Apples are red $\lor$ *Bananas are red*
     Bananas are not red
     $\therefore$ *Apples are red*
   - Syntactic and idiomatic entailment models: SNLI corpus?

Thanks!

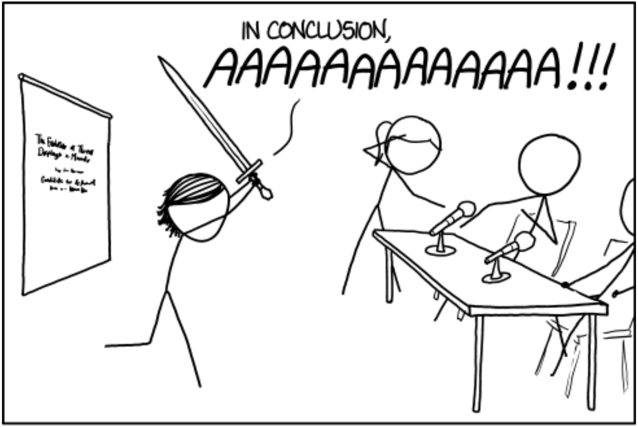




Thanks!

Thanks!

Questions?



THE BEST THESIS DEFENSE IS A GOOD THESIS OFFENSE.

References I

Angeli, G., Chaganty, A., Chang, A., Reschke, K., Tibshirani, J., Wu, J. Y., Bastani, O., Siilats, K., and Manning, C. D. (2014a).
Stanford's 2013 KBP system.
In *TAC-KBP*.

Angeli, G. and Manning, C. D. (2013).
Philosophers are mortal: Inferring the truth of unseen facts.
In *CoNLL*.

Angeli, G. and Manning, C. D. (2014).
Naturalli: Natural logic inference for common sense reasoning.
In *EMNLP*.

Angeli, G., Premkumar, M. J., and Manning, C. D. (2015).
Leveraging linguistic structure for open domain information extraction.
In *ACL*.

References II

Angeli, G., Tibshirani, J., Wu, J. Y., and Manning, C. D. (2014b).
Combining distant and partial supervision for relation extraction.
In *EMNLP*.

Banko, M., Cafarella, M. J., Soderland, S., Broadhead, M., and Etzioni, O. (2007).
Open information extraction for the web.
In *IJCAI*.

Berant, J., Chou, A., Frostig, R., and Liang, P. (2013).
Semantic parsing on freebase from question-answer pairs.
In *Proceedings of EMNLP*.

Bowman, S. (2014).
Can recursive neural tensor networks learn logical reasoning?
ICLR (arXiv:1312.6192).

References III

Bowman, S. R., Potts, C., and Manning, C. D. (2015).
Recursive neural networks can learn logical semantics.
ACL-IJCNLP 2015.

Doddington, G. R., Mitchell, A., Przybocki, M. A., Ramshaw, L. A., Strassel, S., and Weischedel, R. M. (2004).
The automatic content extraction (ACE) program–tasks, data, and evaluation.
In *LREC*.

Fader, A., Soderland, S., and Etzioni, O. (2011).
Identifying relations for open information extraction.
In *EMNLP*.

References IV

Glickman, O., Dagan, I., and Koppel, M. (2006).
A lexical alignment model for probabilistic textual entailment.
Machine Learning Challenges. Evaluating Predictive Uncertainty, Visual Object Classification, and Recognising Tectual Entailment.

Hixon, B., Clark, P., and Hajishirzi, H. (2015).
Learning knowledge graphs for question answering through conversational dialog.
NAACL.

Hoffmann, R., Zhang, C., Ling, X., Zettlemoyer, L., and Weld, D. S. (2011).
Knowledge-based weak supervision for information extraction of overlapping relations.
In *ACL-HLT*.

References V

Icard, III, T. and Moss, L. (2014).
Recent progress on monotonicity.
Linguistic Issues in Language Technology.

MacCartney, B. (2009).
Natural Language Inference.
PhD thesis, Stanford.

MacCartney, B. and Manning, C. D. (2008).
Modeling semantic containment and exclusion in natural language inference.
In *Coling*.

Mausam, Schmitz, M., Bart, R., Soderland, S., and Etzioni, O. (2012).
Open language learning for information extraction.
In *EMNLP*.

References VI

Mintz, M., Bills, S., Snow, R., and Jurafsky, D. (2009).
Distant supervision for relation extraction without labeled data.
In *ACL*.

Sánchez Valencia, V. M. S. (1991).
Studies on natural logic and categorial grammar.
PhD thesis, University of Amsterdam.

Soderland, S., Gilmer, J., Bart, R., Etzioni, O., and Weld, D. S. (2013).
Open information extraction to KBP relations in 3 hours.
In *Text Analysis Conference*.

Surdeanu, M. and Ciaramita, M. (2007).
Robust information extraction with perceptrons.
In *ACE07 Proceedings*.

References VII

Surdeanu, M., Tibshirani, J., Nallapati, R., and Manning, C. D. (2012).
Multi-instance multi-label learning for relation extraction.
In *EMNLP*.

Wu, F. and Weld, D. S. (2007).
Autonomously semantifying wikipedia.
In *Proceedings of the sixteenth ACM conference on information and knowledge management*. ACM.