Modeling Linguistic Structure (And What We Can Do With One)

Vivek Srikumar
School of Computing

Joint work with many people at Universities of Illinois, Colorado, Georgetown University, Stanford University and IHMC Florida

This talk

- The Ins and Outs of Preposition Semantics
- Predicting Preposition Relations
- Learning to Read

What are structures and why are they useful?

- For the purpose of this talk
  - A labeled graph

- Computer science knows how to deal with discrete objects
  - Databases, queries, graph algorithms, etc

- But, natural language is *unstructured*

The agenda: Convert unstructured text into linguistically motivated structured representations

Discrete Linguistic Structures

- Frame Semantics, case grammars
  - Rich history in linguistics starting with Fillmore (1966)
- Large datasets to train classifiers
  - PropBank [Palmer, et al 2005-]
  - FrameNet [Fillmore, Ruppenhofer,…]
- The general premise: Words evoke frames
  - Which words? Most commonly, verbs and eventive nouns

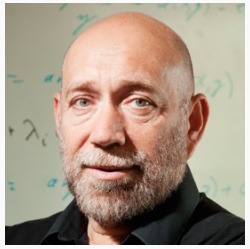
This talk

- The Ins and Outs of Preposition Semantics
- Predicting Preposition Relations
- Learning to Read

The Ins and Outs of Preposition Semantics

Nathan Schneider

Jena Hwang



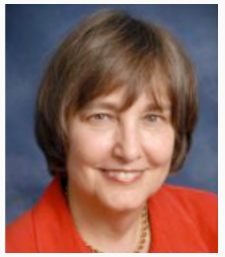
Dan Roth

Meredith Green

Abhijit Suresh

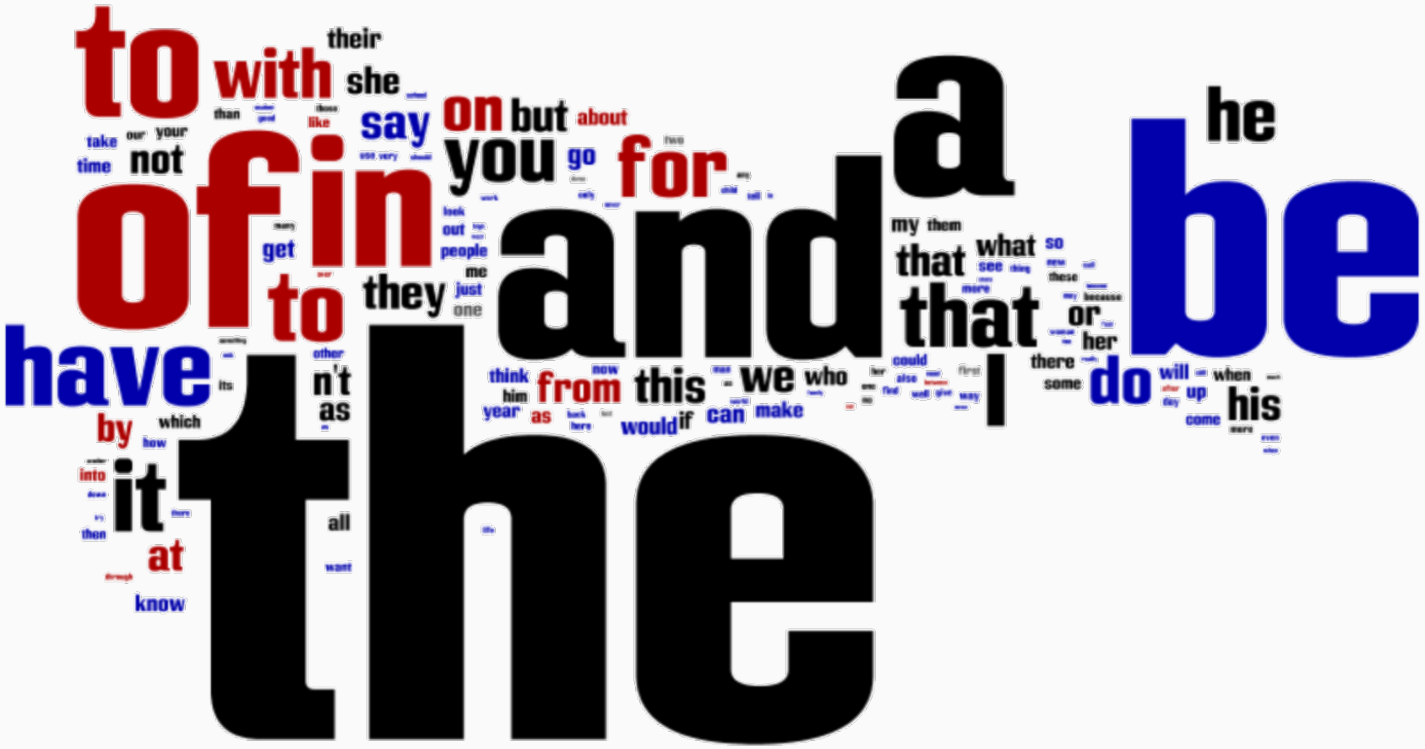
Tim O'Gorman

Katie Conger



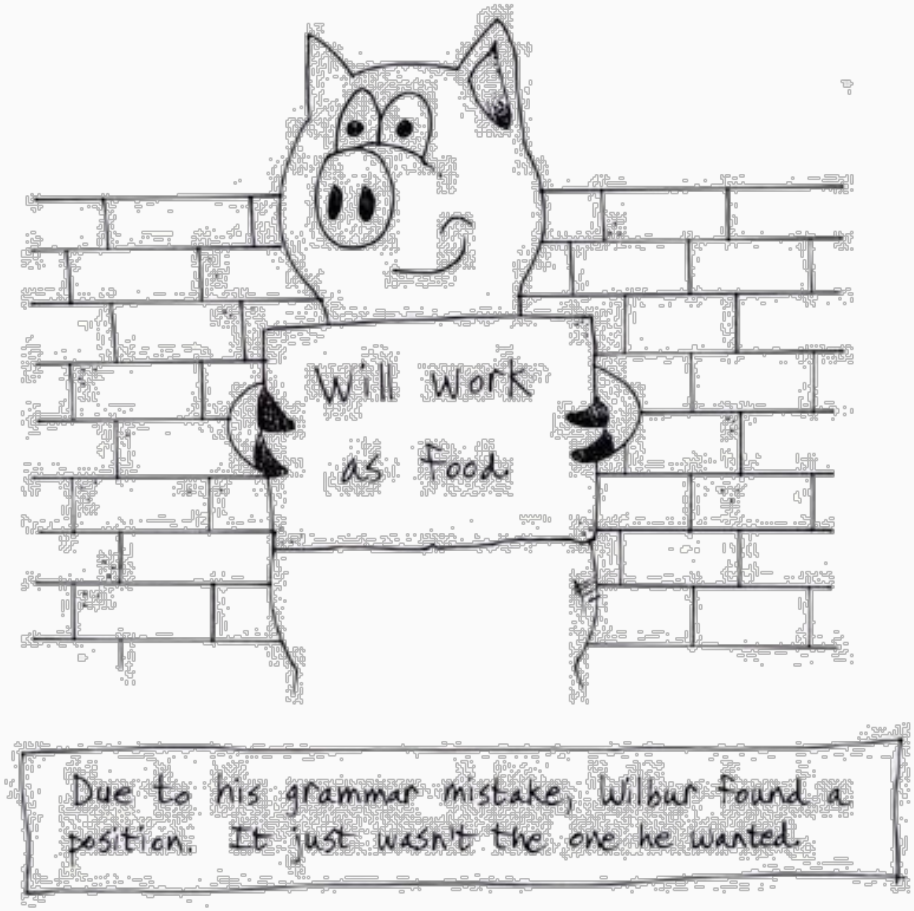
Martha Palmer

Prepositions abound in English



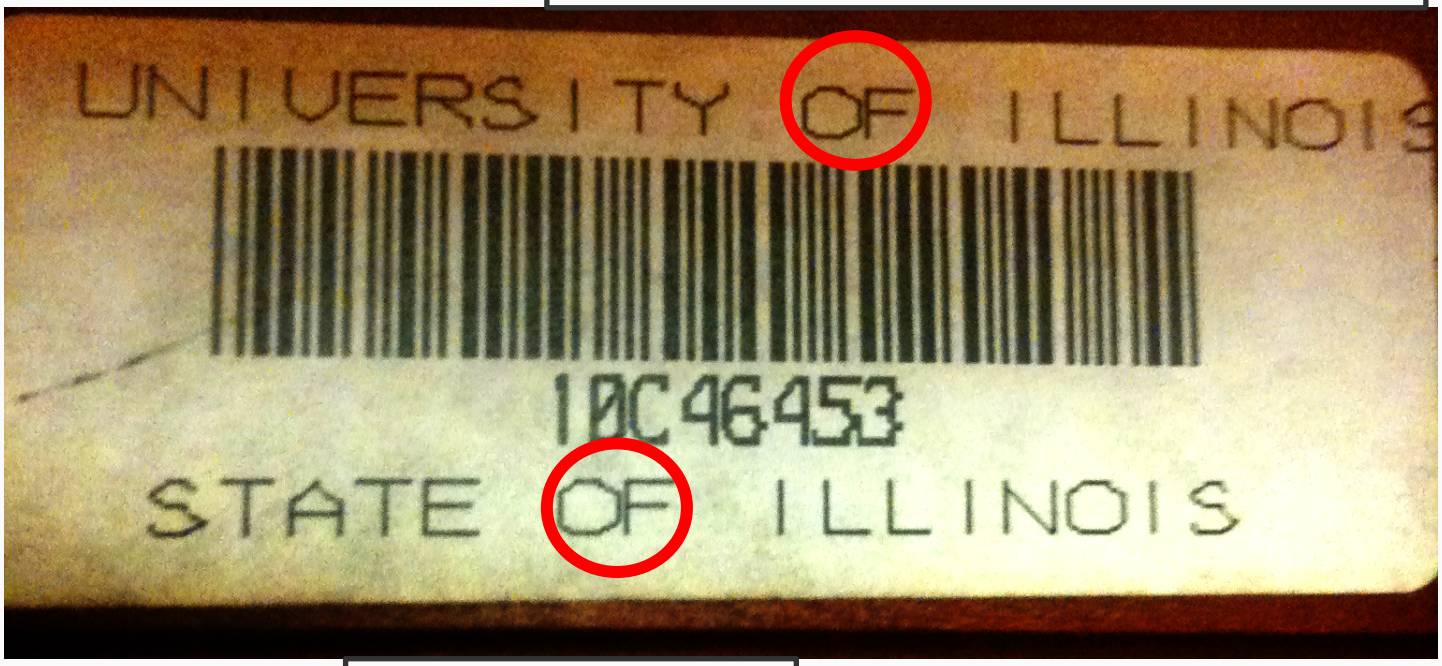
based on COCA list of 5000 most frequent English words

Prepositions trigger semantic relations



Prepositions are ambiguous

The university belongs to Illinois



Illinois is a state

Preposition Sense Disambiguation

Eg. *State* ***of*** *Illinois* vs. *University* ***of*** *Illinois*

- The Preposition Project [Litkowski and Hargraves, 2005]
  - Word sense for 34 prepositions
  - Based on preposition definitions in Oxford Dictionary of English

Preposition Senses

Merriam-Webster over

Fine grained details
Like Over 15(6)–1

Full Definition of OVER

1 —used as a function word to indicate motion or situation in a position higher than or above another <towered *over* his mother> <flew *over* the lake> <rode *over* the old Roman road>

2 **a** —used as a function word to indicate the possession of authority, power, or jurisdiction in regard to some thing or person <respected those *over* him>

b —used as a function word to indicate superiority, advantage, or preference <a big lead *over* the others>

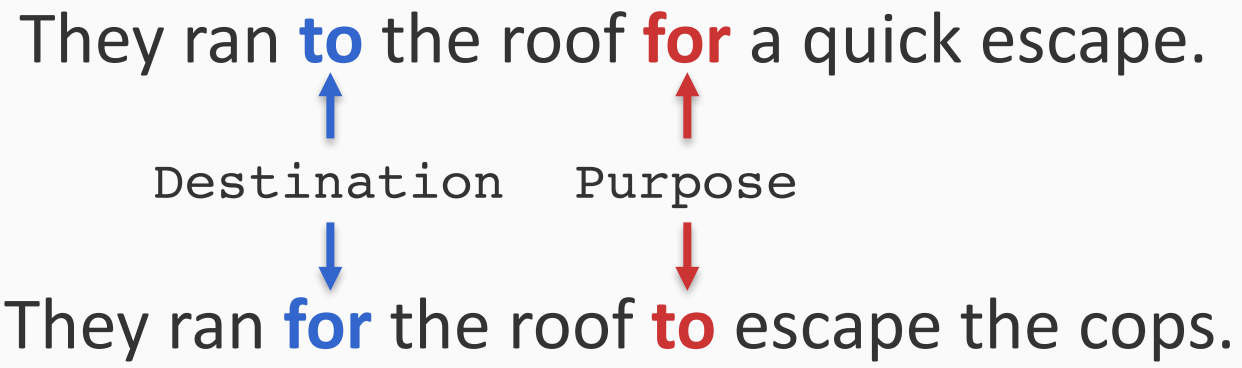
c —used as a function word to indicate one that is overcome, circumvented, or disregarded <passed *over* the governor's veto>

3 **a :** more than <cost *over* $5>

b : ABOVE 4

4 **a** —used as a function word to indicate position upon or movement down upon <laid a blanket *over* the child> <hit

Supersenses = shared functions



Adposition Supersenses

- **Early attempt: Preposition relations**
  - Mapping preposition senses to unlexical supersenses [TACL 2013]
- **English preposition supersenses**
  - Annotation scheme & lexicon [LAW 2015]
  - Annotated corpus [LAW 2016]
- **CARMLS**: making these categories multilingual [work in progress!]

English Preposition Supersenses data

Comprehensive annotation: first dataset with **all** prepositions (types+tokens) semantically annotated

- Sentences not hand-selected
- Sentences fully annotated
- Preposition types not constrained by a lexicon (labels generalize)
- All sentences seen by multiple annotators

Challenges

1. Scoping phenomena of interest: multiwords, particles, grammaticized uses of prepositions, etc.

2. Deciding categories and their boundaries

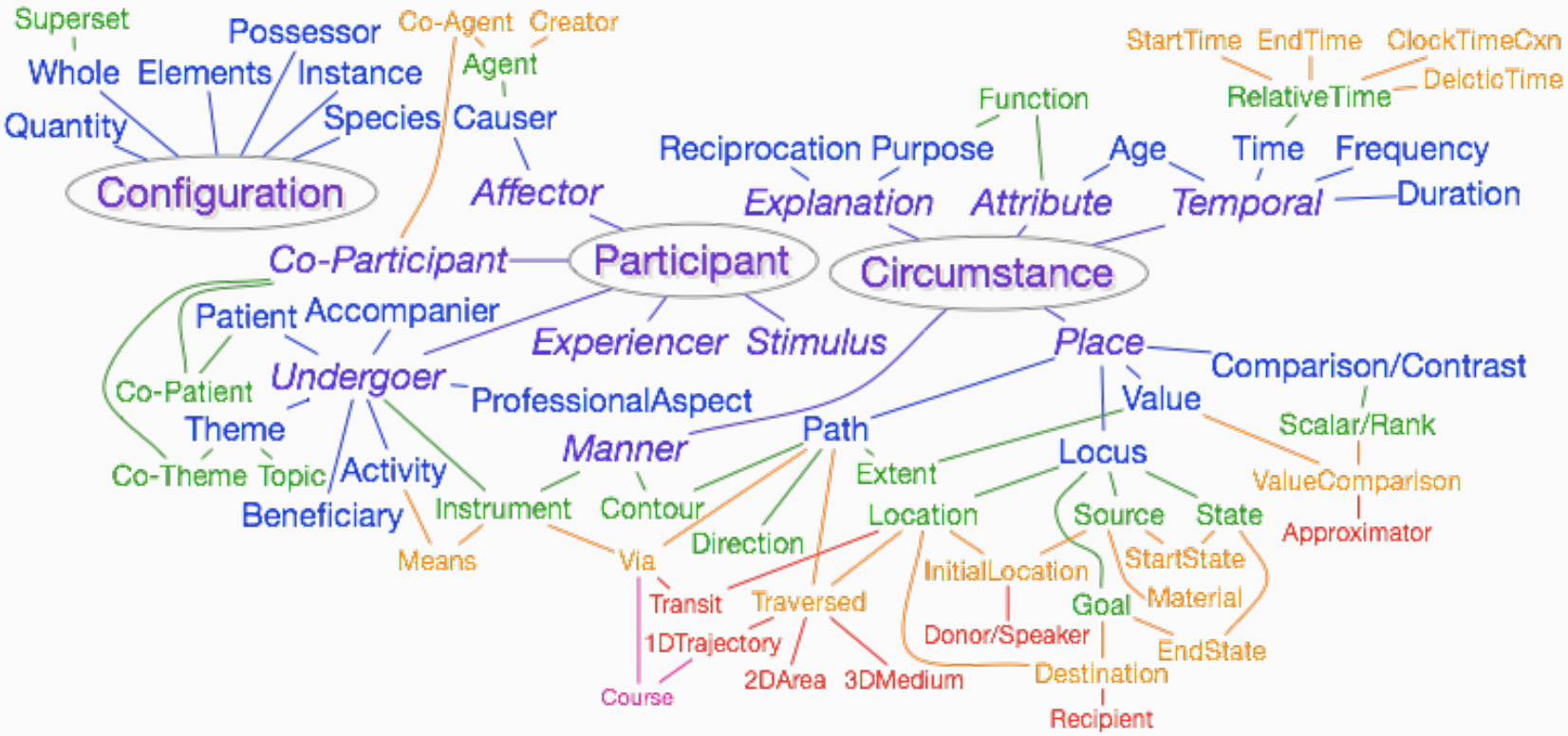
3. Thorough documentation with example usages

Defining the supersenses

- Starting point: 32 categories from S & Roth (2013) + VerbNet thematic role hierarchy

- Extensive discussion of examples and pilot annotation to identify categories
  - Multiple inheritance among categories
    E.g., VIA (*talk* over *Skype*) = INSTRUMENT + PATH

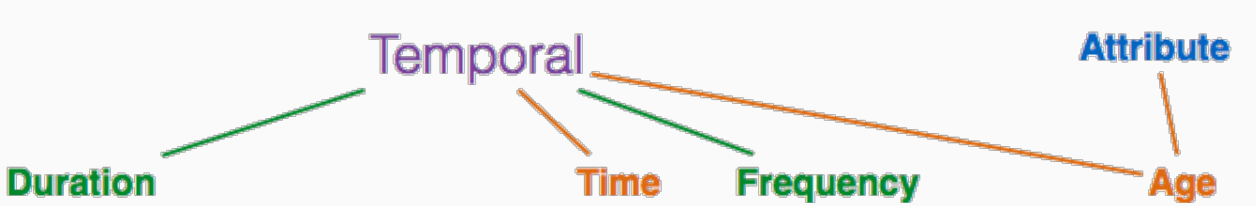
Supersense hierarchy



75 preposition supersense categories

http://tiny.cc/prepwiki

Temporal hierarchy



PrepWiki

- Because there are so many supersenses, we showed annotators **example sentences** for each prep/supersense pair
  - So we needed to create a lexical resource documenting possible usages

Category | Discussion | Read | Edit | View history | Search

Category:In

http://tiny.cc/prepwiki

Word sense

	Definition	Example	Supersense
In 1(1)	expressing the situation of something that is or appears to be enclosed or surrounded by something else:	I'm living in London; she saw the bus in the rear-view mirror.;	SST-Location
In 1(1)-1	used to introduce someone's clothing	Trouserless men looked absurd in socks.	SST
In 10(7a)	indicating the key in which a piece of music is written:	Mozart's Piano Concerto in E flat.;	SST-Attribute
In 11(8)	as an integral part of (an activity):	In announcing the program, Computershare pointed out the environmental benefits of reducing the use of valuable resources such as trees.; He was unsuccessful in running for president	SST-Circumstance
In 12(9)	expressing a value as a proportion of (a whole):	a local income tax running at six pence in the pound.; Perhaps only one in twenty of the city's adult residents had been born there.	SST-Superset,Value
In 2(1a)	expressing motion with the result that something ends up within or surrounded by something else:	don't put coal in the bath; he got in his car and drove off.;	SST-Location
In 3(2)	expressing a period of time during which an event happens or a situation	they met in 1885; at one o'clock in the morning; I hadn't seen him in years.; In my life I have never seen so many cats on a couch; the leaders met for	SST-DeicticTime,Time

examples

supersense(s) mapped to individual examples (not shown)

Corpus annotation

- Fully annotated a 55,000 word corpus of online reviews
- Several annotators at CU-Boulder
- All sentences were independently annotated twice, then adjudicated by an expert annotator
  - Original IAA rates varied considerably; mostly 60%–78%
  - IAA between experts: 88%

Preposition types

to	from	before	without	during
of	about	over	off	...
in	like	than	until	
for	by	since	through	
with	after	around	ago	
on	as	into	away	
at	back	out	within	

out of	as soon as	in front of	rather than
at least	all over	more than	as to
due to	instead of	more like	such as
because of	other than	along with	as long as
at all	next to	just about	aside from
less than	nothing but	thanks to	...

Ongoing work: CARMLS

Case and Adposition
Representation
for Multi-Lingual Semantics

Can we use supersenses for case markers and adpositions in other languages?

Focus languages (besides English)

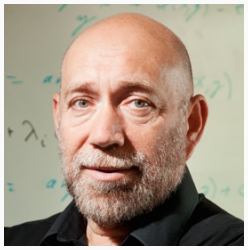
Hindi

chna Bhatia

la-Rae Han

This talk

- The Ins and Outs of Preposition Semantics
- Predicting Preposition Relations
- Learning to Read



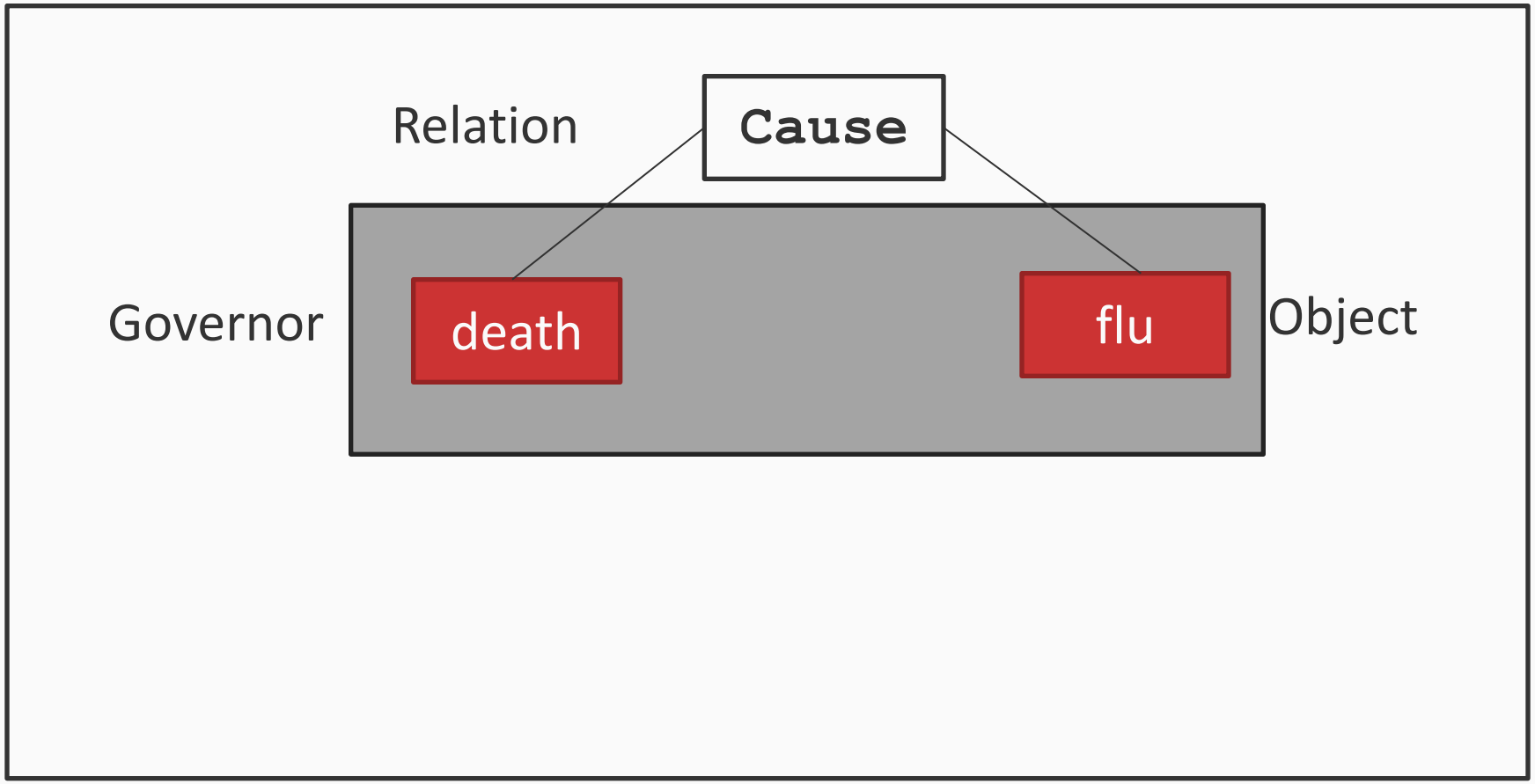
Dan Roth

"zoo in NYC" ⟶ **Location**(zoo, NYC)

Two Models for Predicting Preposition Relations [TACL 2013]

Structure of prepositions

Poor care led to her death **from** flu.



Relation depends on argument types

Poor care led to her death **from** flu.

`Cause`(death, flu)

Poor care led to her death **from** pneumonia.

How do we generalize the classifier to unseen arguments in the same “type”?

Why are types important?

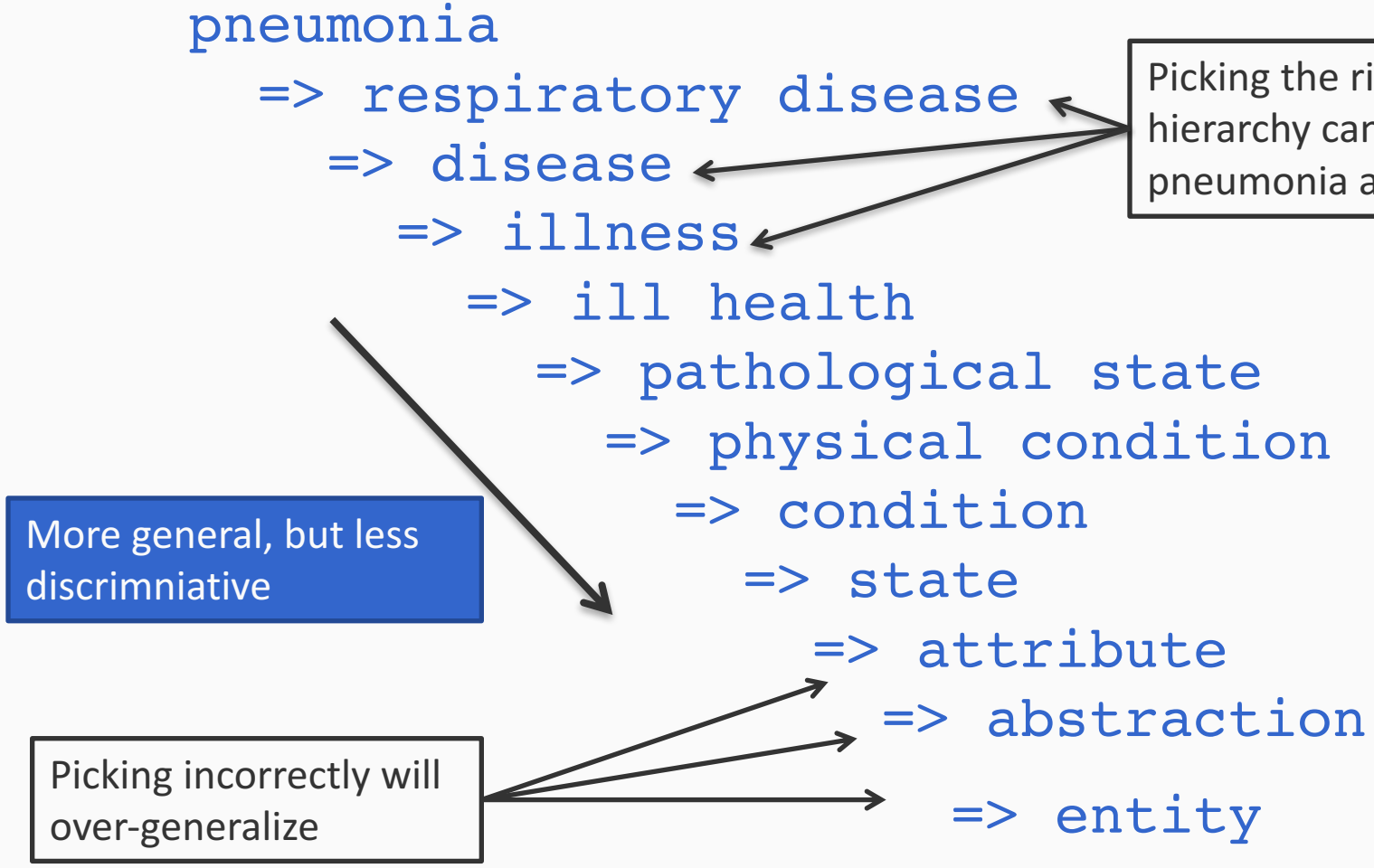
- Goes beyond words
  - Abstract *flu and pneumonia* into the same group

- Some semantic relations hold only for certain types of entities

- How do we know the types of words?

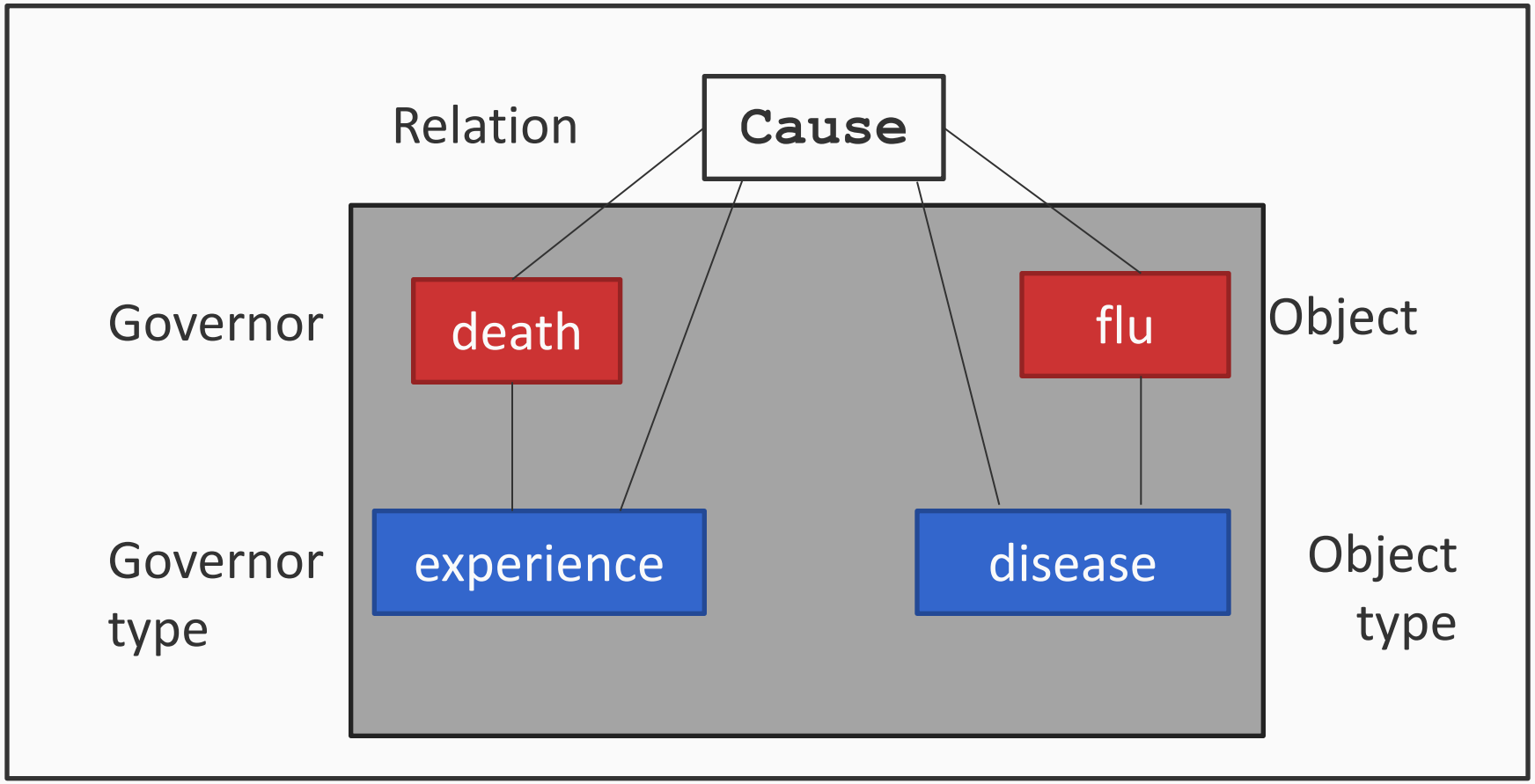
  A: Hypernymy

WordNet IS-A hierarchy



Structure of prepositions

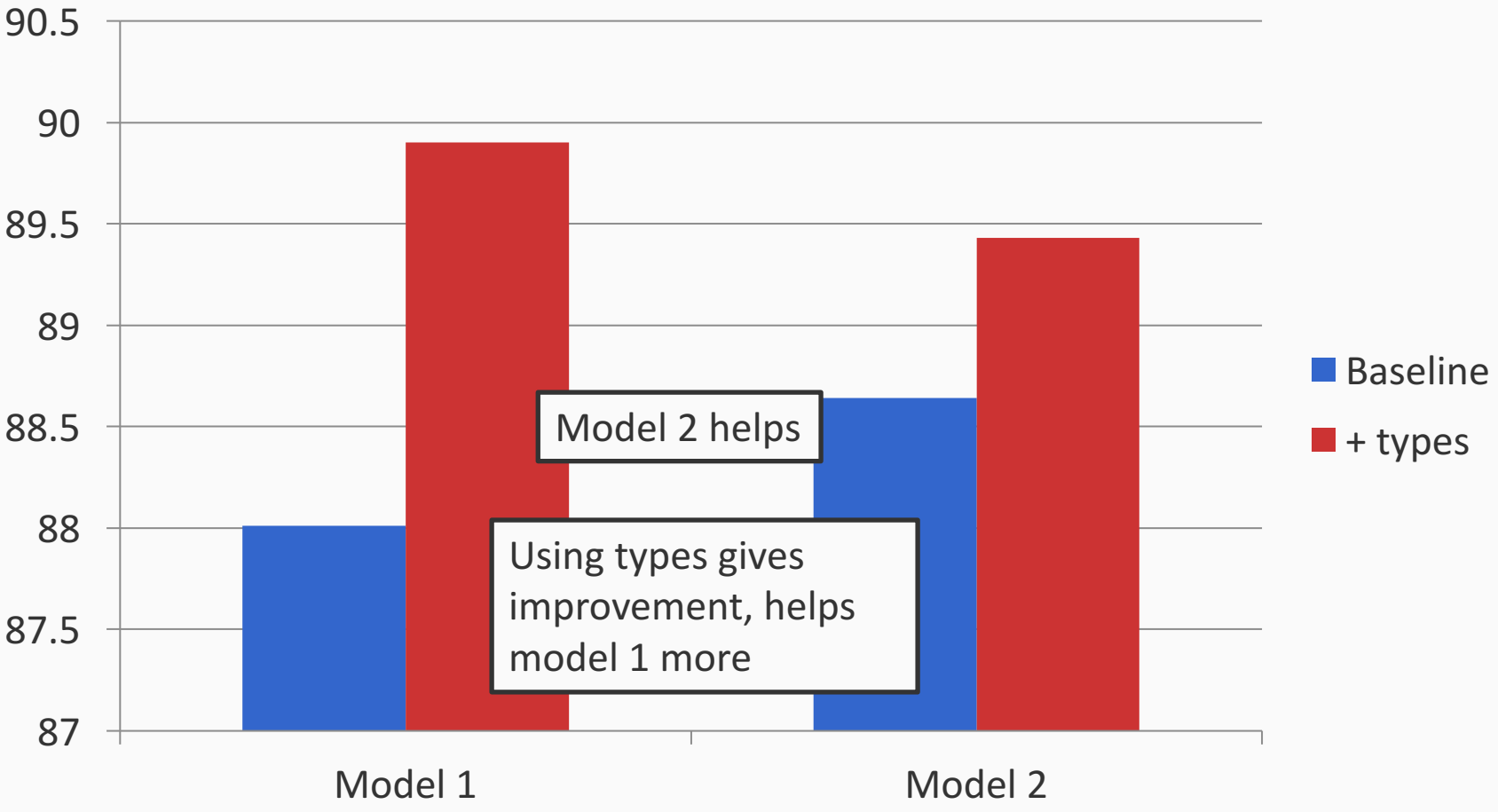
Poor care led to her death **from** flu.



Two models

- Model 1
  - Predict only relation label: Multi-class
  - Use features from all possible governor and object candidates
    - Also types
- Model 2 uses features from the structure
  - Predict full structure: relation and arguments
    - Also types

Accuracy of supersense labeling



What do we have?

Input	Relation	Governor type	Object type
Died of pneumonia	`Cause`	Experience	Disease
Suffering from flu	`Cause`	Experience	Disease
Recovered from flu	`StartState`	Change	Disease

Governor, object and their types as a certificate for the choice of relation label

Summary

- Prepositions express a diverse set of relations
  - An ontology of preposition supersenses
  - Can enrich existing PropBank/FrameNet representation

- Models for predicting preposition relations
  - Arguments and types help

This talk

- The Ins and Outs of Preposition Semantics
- Predicting Preposition Relations
- Learning to Read

Jonathan Berant

Heather Chen

Chris Manning

Modeling Biological Processes for Reading Comprehension

EMNLP 2014

What can we do with a semantic representation?

Reading comprehension is hard!

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. Light absorbed by chlorophyll drives a transfer of the electrons and hydrogen ions from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Reading comprehension is hard!

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. Light absorbed by chlorophyll drives a transfer of the electrons and hydrogen ions from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Reading comprehension is hard!

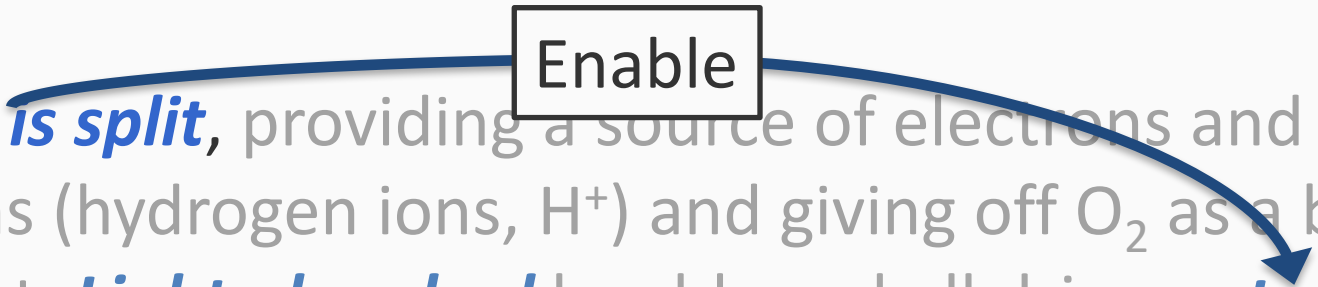
Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Reading comprehension is hard!



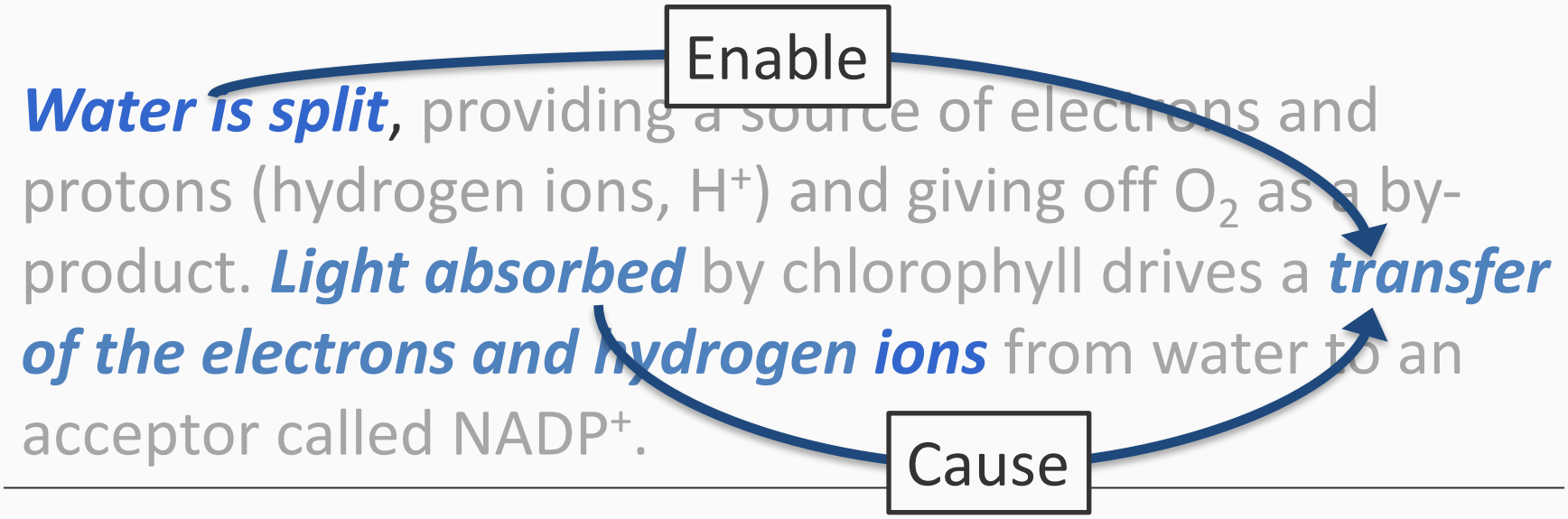
Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Reading comprehension is hard!



Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

The punchline

1. A new reading comprehension task requiring reasoning over processes
    Processes are fundamental in many domains

2. A new dataset `ProcessBank` consisting of descriptions of biological processes with
   – Rich process structure annotated, *and*
   – Multiple-choice questions

3. A new end-to-end system for reading comprehension
   – Predict structure and treat it as a knowledge base
   – Parse question as query to this knowledge base (semantic parsing)

A new dataset: `ProcessBank`

Motivation: macro vs. micro reading

- Macro reading:
  - Exploits web-scale redundancy
    [Etzioni et al., 2006, Carlson et al., 2010, Fader et al., 2011]
  - Factoid questions
    [Berant el al., 2014, Fader et al., 2014]

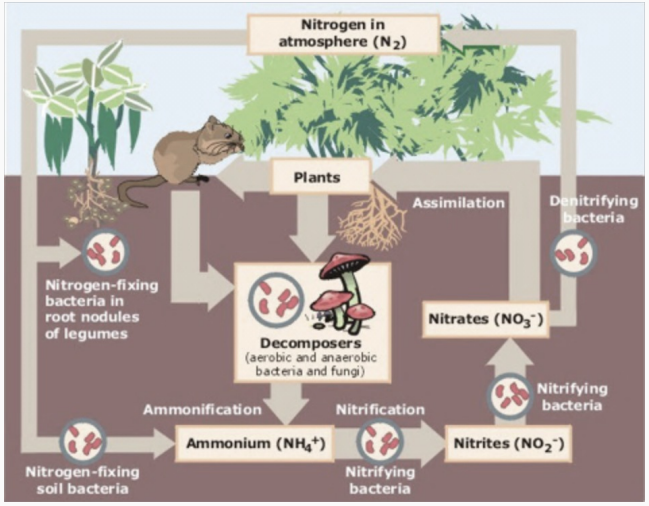
- Micro reading:
  - Single document
  - Requires reasoning
  - Non-factoid questions

  [Richardson et al., 2013, Kushman et al., 2014]

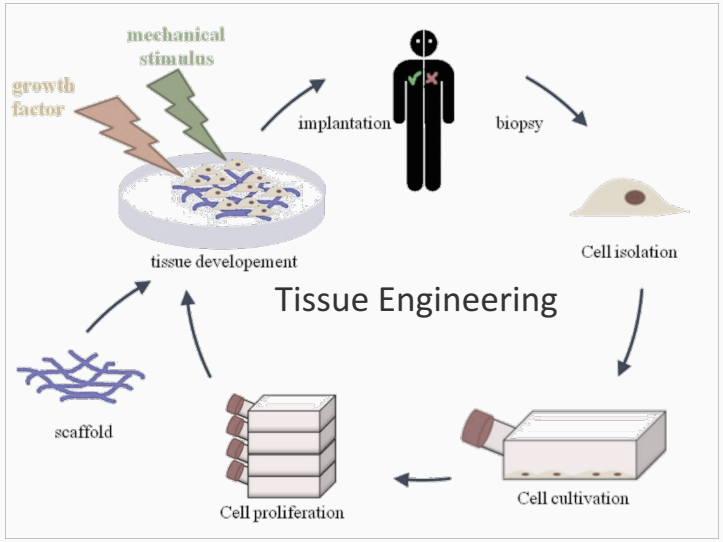
Chosen domain:

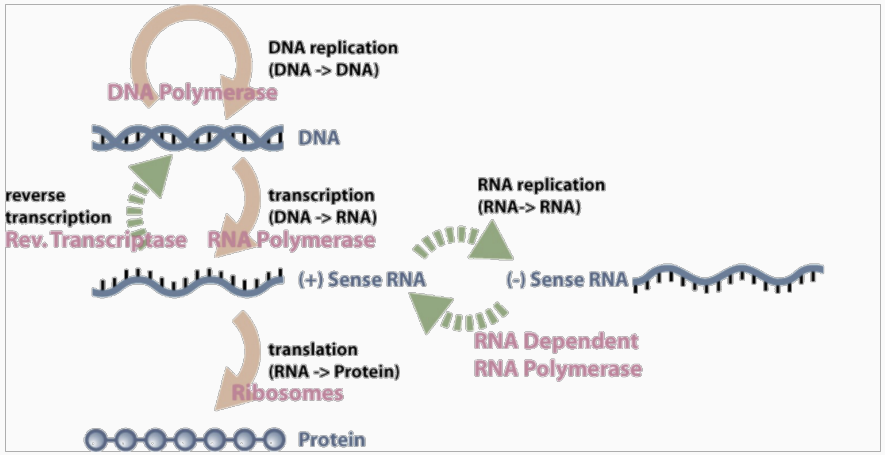
Biological process descriptions

Processes abound in biology

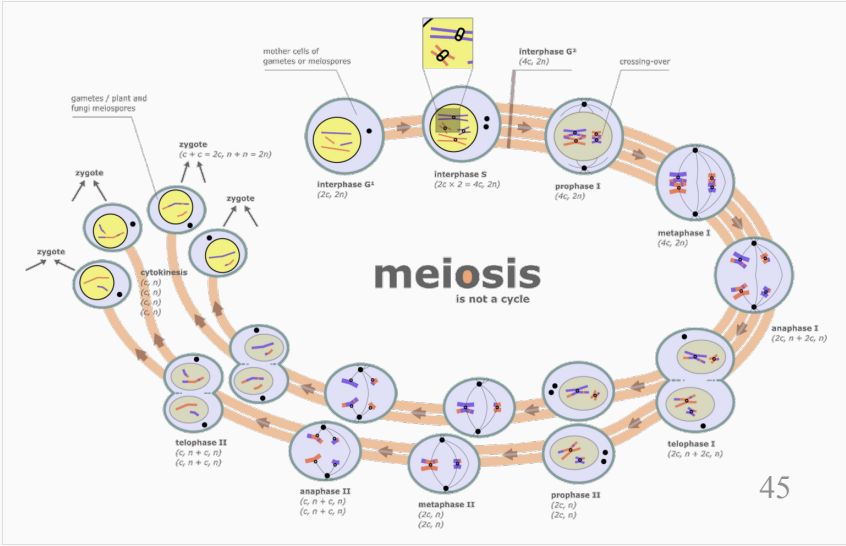


Nitrogen Cycle





Central Dogma of Molecular Biology



Creating a difficult reading comprehension task

200 paragraphs from the textbook *Biology*

[Campbell & Reese, 2005]

Desiderata

1. Test understanding of inter-relations between events and entities
2. Both answers should have similar lexical overlap:
   - Trump shallow approaches
   - Sidestep lexical variability

Reading comprehension annotation

- Annotation instructions: Ask questions about events, entities and their relationships
  - 10 examples provided
  - Two answer choices, only one unambiguously correct
- 200 paragraphs $\rightarrow$ 585 questions
- Second annotator answered the questions
  - 98.1% agreement

Examples of annotated questions

Dependencies between events/entities (70%)

Q: *What can the splitting of water lead to?*

A: Light absorption
B: Transfer of ions

Temporal ordering of events (10%)

Q: *What is the correct order of events?*

A: PDGF binds to tyrosine kinases, then cells divide, then wound healing

B: Cells divide, then PDGF binds to tyrosine kinases, then wound healing

True-False questions (20%)

Q: *Cdk associates with MPF to become cyclin*

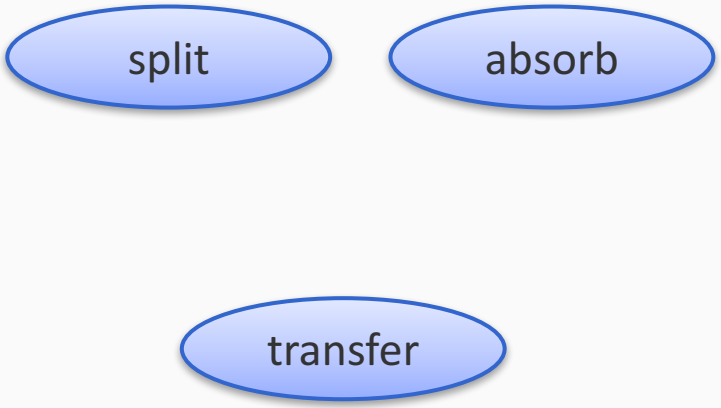
A: True
B: False

A second layer of annotation: Process structures

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

A second layer of annotation: Process structures

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

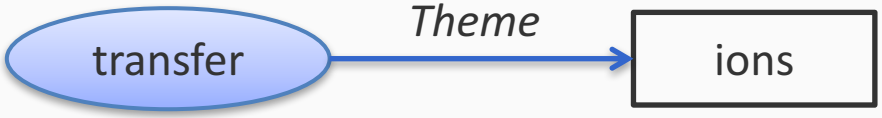


Triggers: Tokens denoting occurrence of an event

A second layer of annotation: Process structures

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

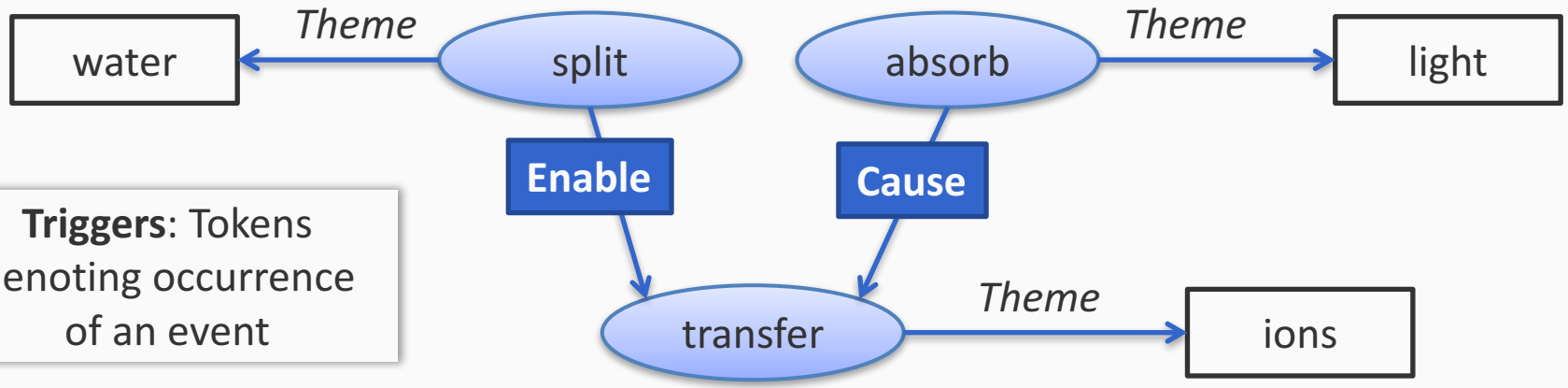
Triggers: Tokens denoting occurrence of an event



Arguments: *Agent, Theme, Source, Destination, Location, Result* and *Other*

A second layer of annotation: Process structures

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.



Triggers: Tokens denoting occurrence of an event

Arguments: *Agent, Theme, Source, Destination, Location, Result* and *Other*

E-E relations: *Cause, Enable, Prevent, Super, Same*

Process structure data

- Same 200 paragraphs from *Biology*
  - Paragraphs annotated and verified
- Three annotators
  - Biologists
  - Independent from QA annotator
    - Potentially conflicting with questions
- More nuances
  - Eg: No temporal ordering of events

What is `ProcessBank`?

- 200 paragraphs from the textbook *Biology*
  - Manually chosen to represent biological processes

- Each paragraph annotated with
  1. Non-factoid reading comprehension questions
  2. Process structures

Answering questions: Overview

System in a nutshell

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

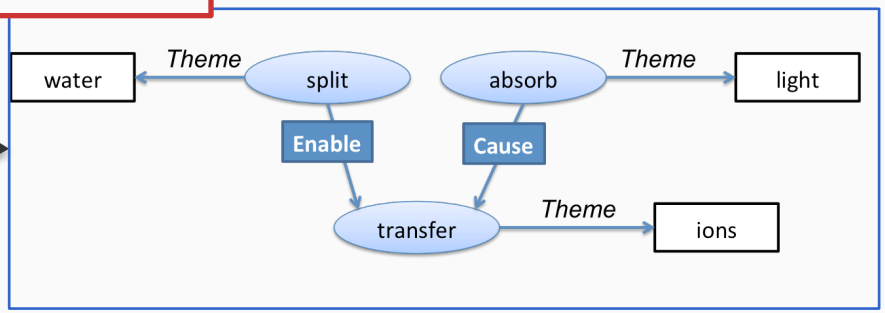
What can the splitting of water lead to?
A: Light absorption
B: Transfer of ions

System in a nutshell

Process Structure Prediction

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

Step 1



What can the splitting of water lead to?

A: Light absorption

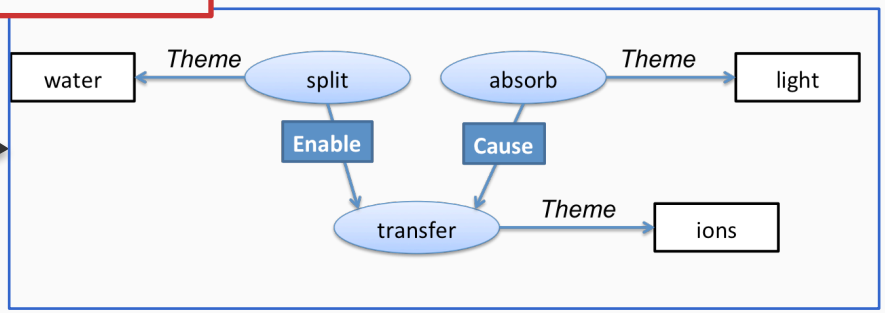
B: Transfer of ions

System in a nutshell

Process Structure Prediction

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

Step 1



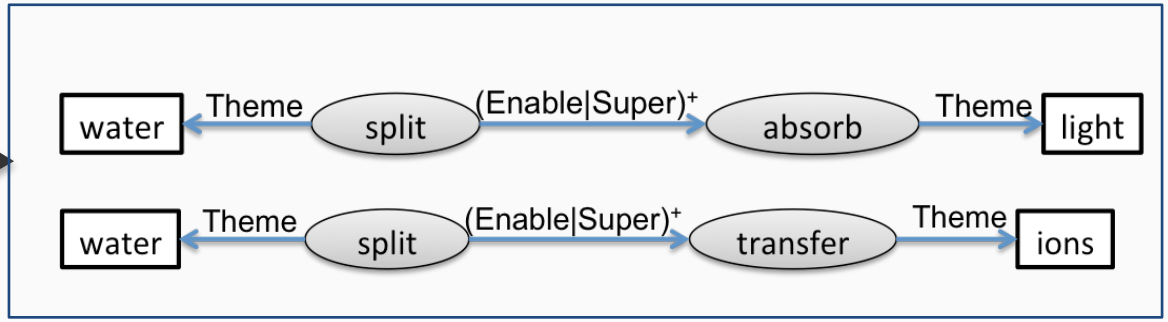
What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Question Parsing

Step 2

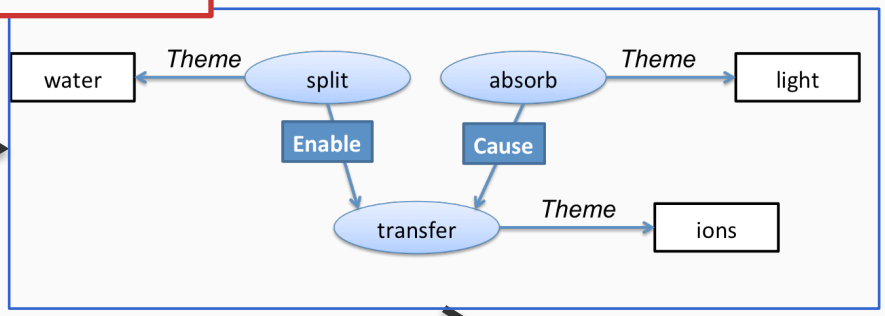


System in a nutshell

Process Structure Prediction

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

Step 1



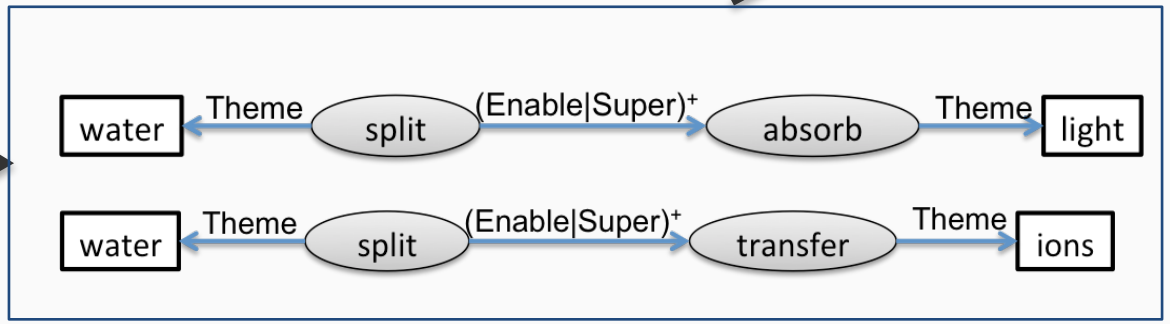
What can the splitting of water lead to?
A: Light absorption
B: Transfer of ions

Answering Question

Step 3: Answer = **B**

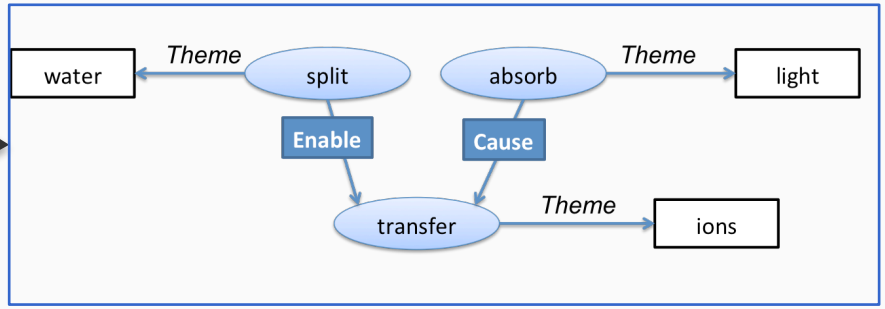
Step 2

Question Parsing



Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

Step 1



Predicting process structures

Process structure prediction

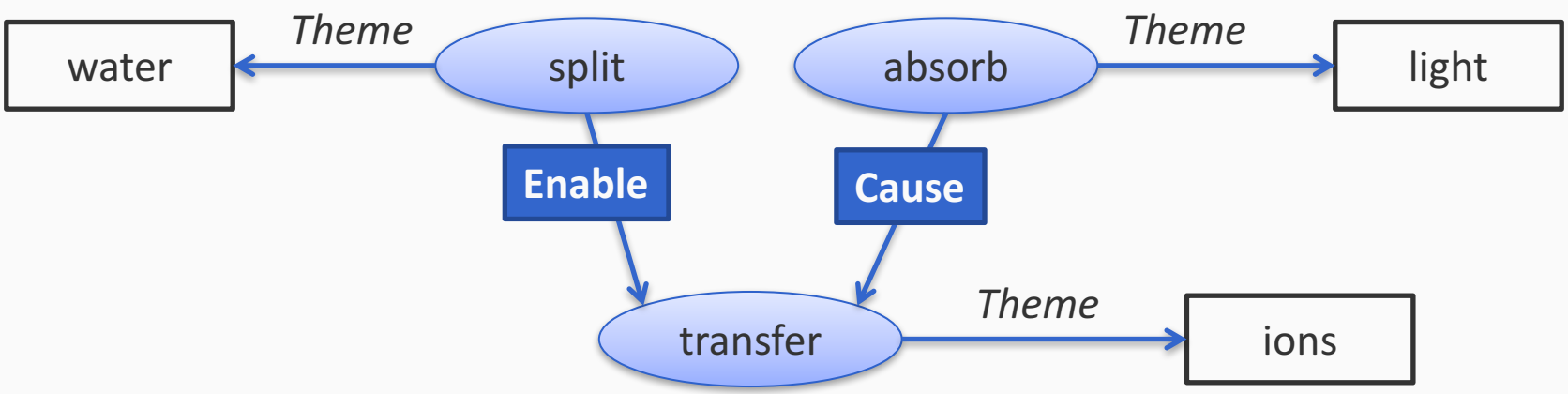
1. Train trigger identifier

   Logistic regression; features from words, lists

2. Joint learning and inference over arguments and event-event relations using predicted triggers

Event-arguments and event-event relations

$$\max \left(\text{Total event-argument score} + \text{Total event-event relation score} \right)$$



[Roth and Yih, 2004, Punyakanok et al. 2008, etc]

Event-arguments and event-event relations

$$\max \sum_{t,a,L}$$

Score that arg-candidate has label L

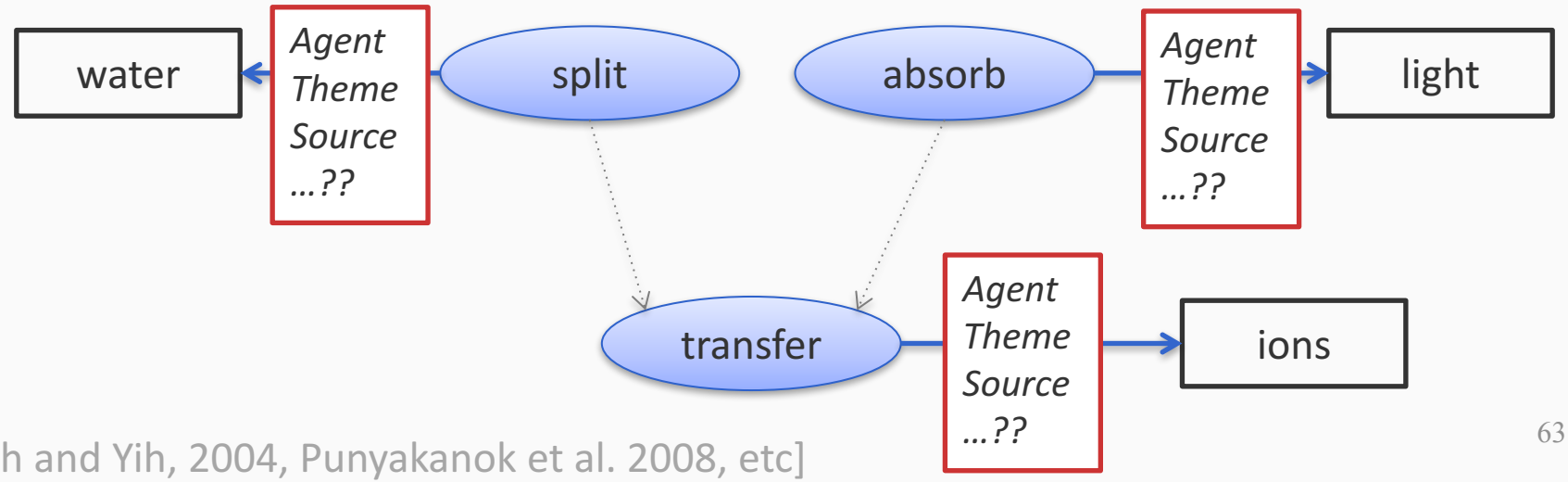
+

Total event-event relation score

Trigger *t*

Argument candidate *a*

Argument label *L*



[Roth and Yih, 2004, Punyakanok et al. 2008, etc]

Event-arguments and event-event relations

$$\max \sum_{t,a,L} \text{Score that arg-candidate has label L} + \sum_{t_1,t_2,R} \text{Score that trigger pair connected by relation R}$$

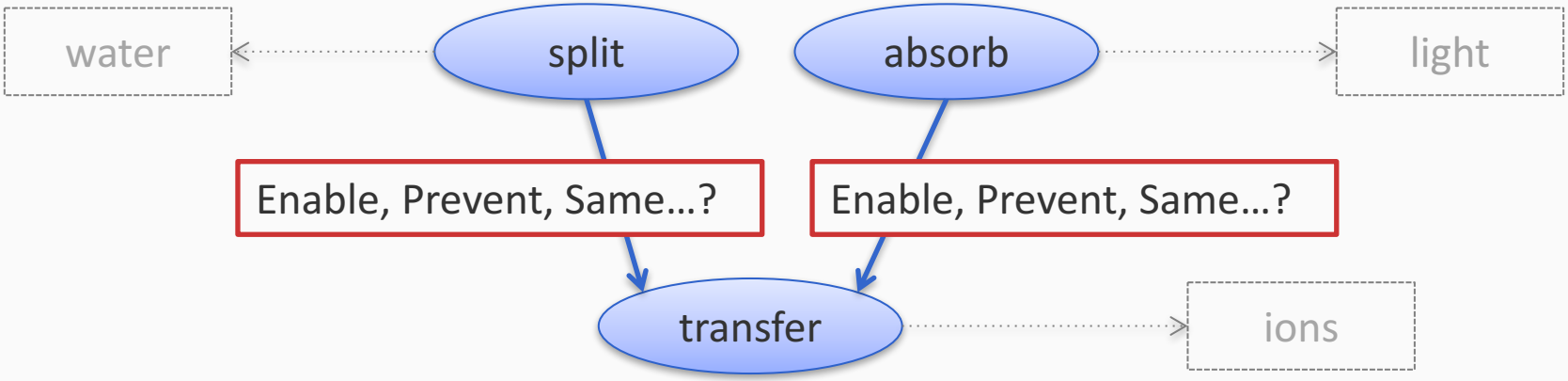
Trigger *t*

Argument candidate *a*

Argument label *L*

Pair of events (t_1, t_2)

Relation label *R*



[Roth and Yih, 2004, Punyakanok et al. 2008, etc]

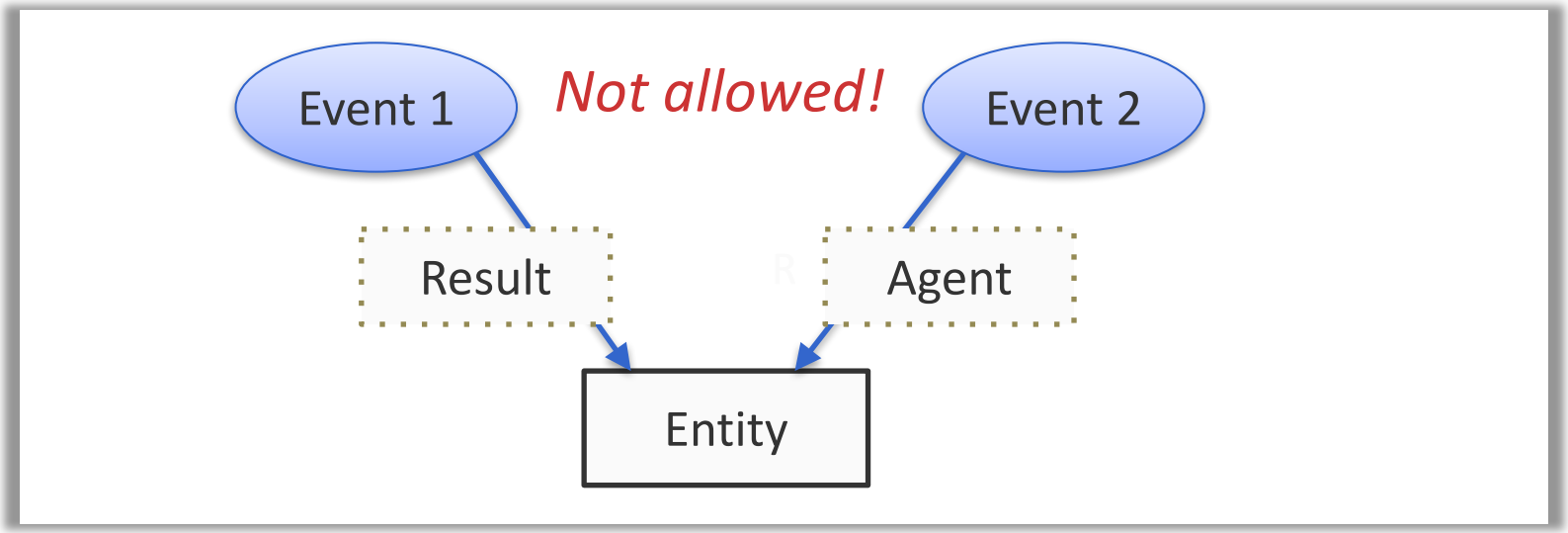
Joint inference with constraints

1. No overlapping arguments
2. Maximum number of arguments per trigger
3. Maximum number of triggers per entity
4. Connectivity
5. Events that share arguments must be related

And a few other constraints

Joint inference with constraints

1.
2.
3.
4.

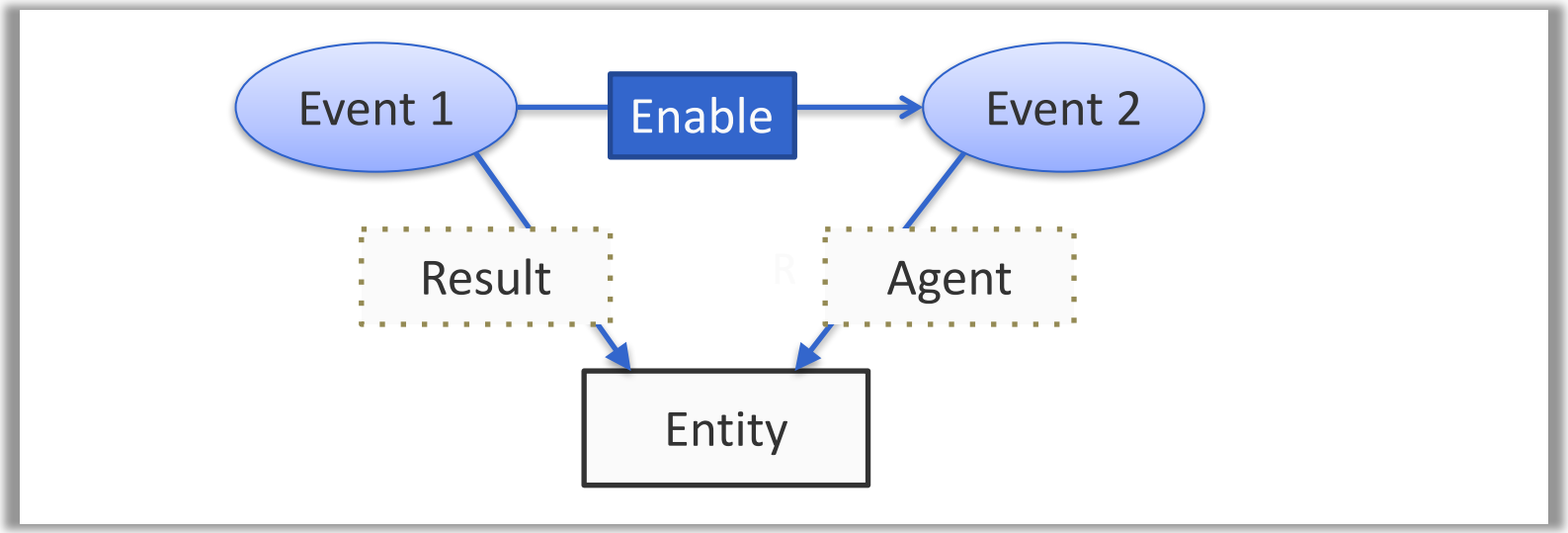


5. ***Events that share arguments must be related***

And a few other constraints

Joint inference with constraints

1.
2.
3.
4.



5. ***Events that share arguments must be related***

And a few other constraints

Learning and Inference

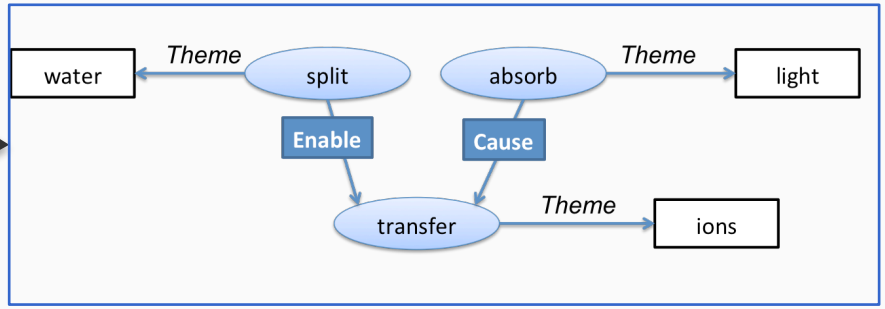
- Linear model to score argument labels and event-event relations
  - Related: Semantic role labeling, information extraction
- Structured averaged perceptron
- Gurobi ILP solver (exact solution)

Answering questions

Where are we?

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

Step 1



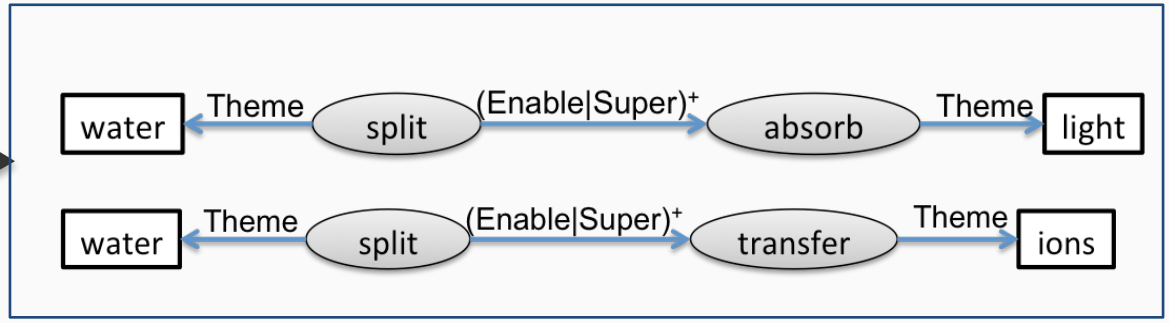
What can the splitting of water lead to?

A: Light absorption

B: Transfer of ions

Question Parsing

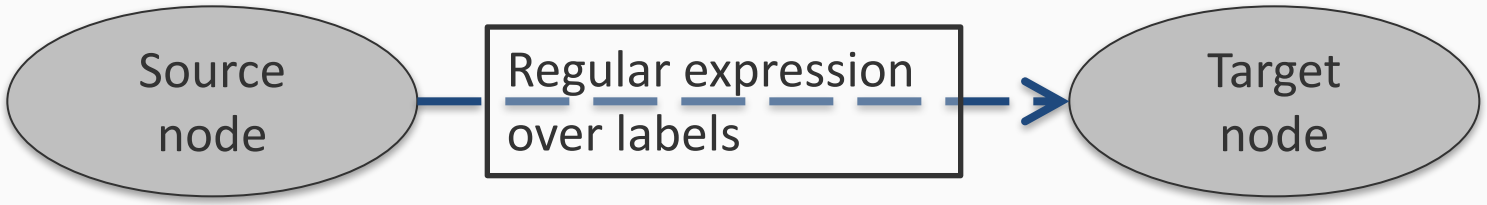
Step 2



Question parsing

- Task: Given a question and two answers, produce two queries
  - One for each answer

- Query structure

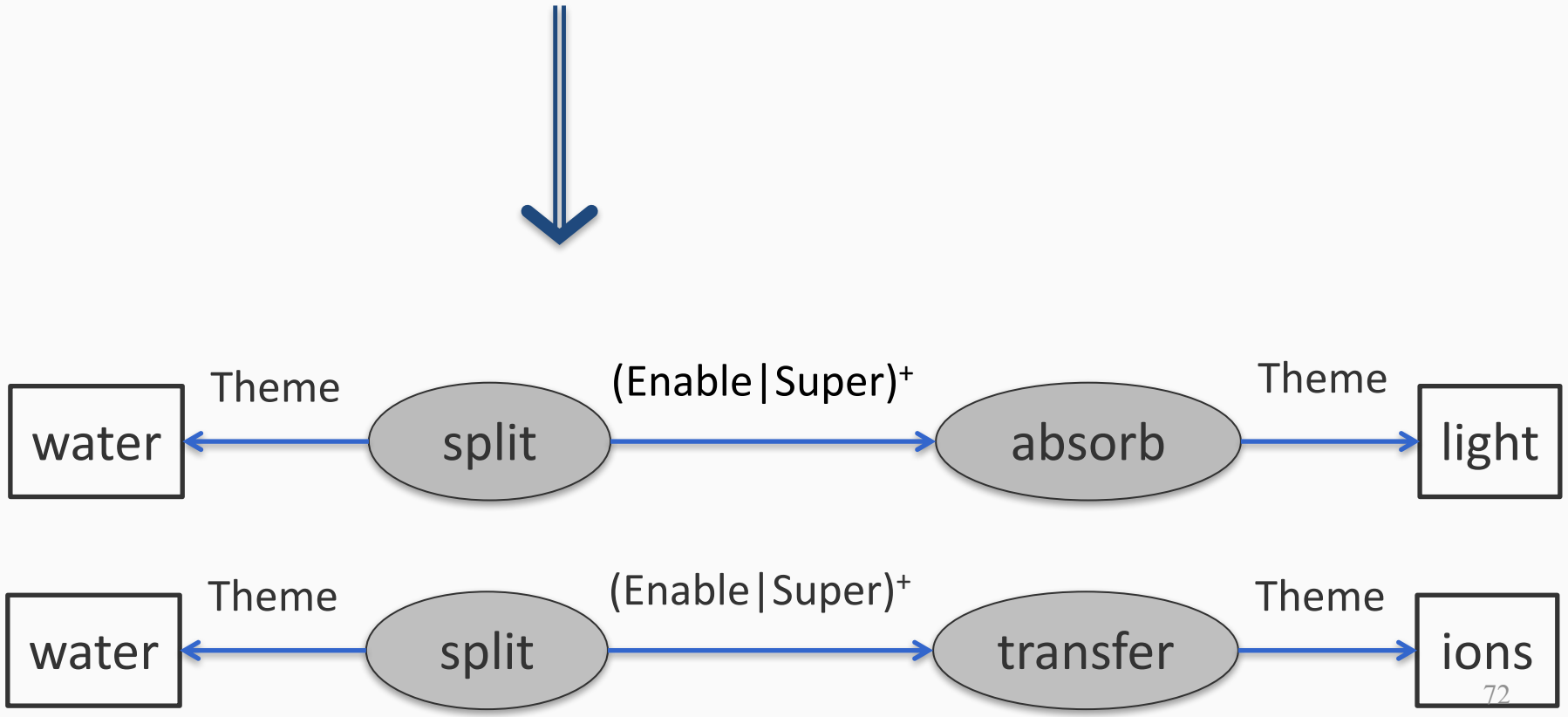


Parsing question to produce formal queries

What does the splitting of water lead to?

A: Light absorption

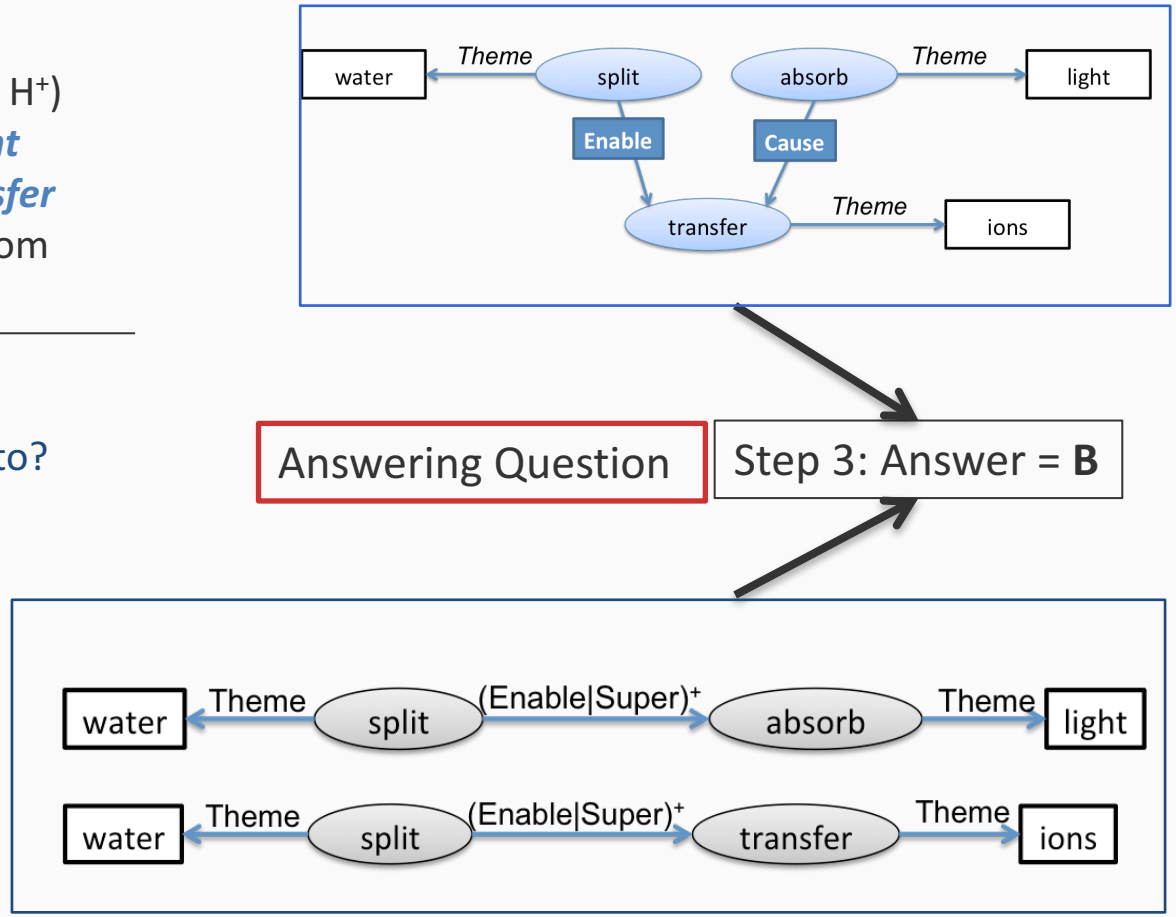
B: Transfer of ions



Where are we?

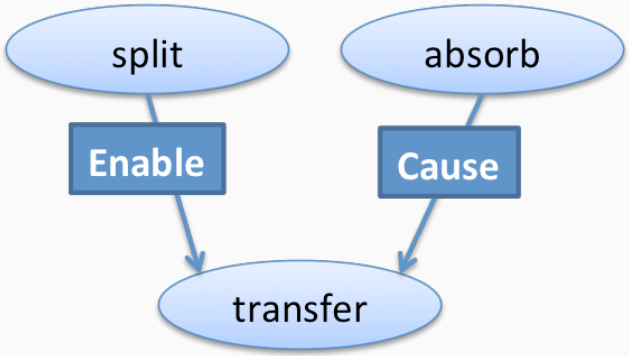
Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.

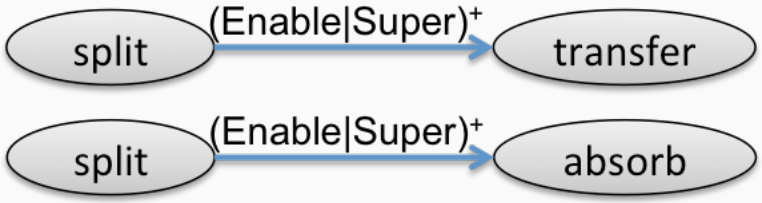
What can the splitting of water lead to?
A: Light absorption
B: Transfer of ions



Step 3: Answering questions

- Given
  - Process structure
  - Two queries

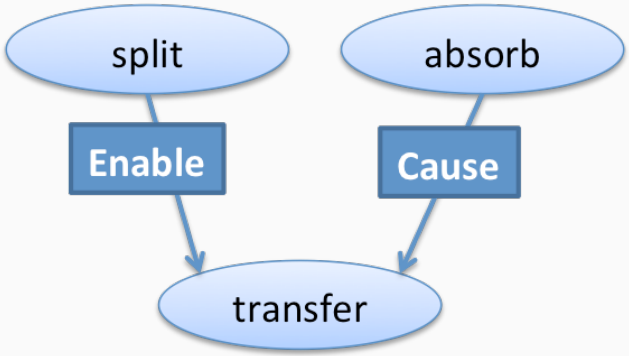


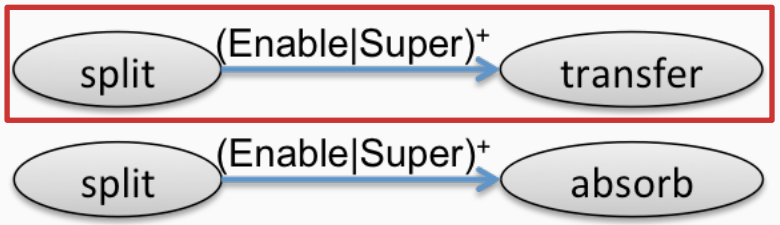


- Answering algorithm

Step 3: Answering questions

- Given
  - Process structure
  - Two queries

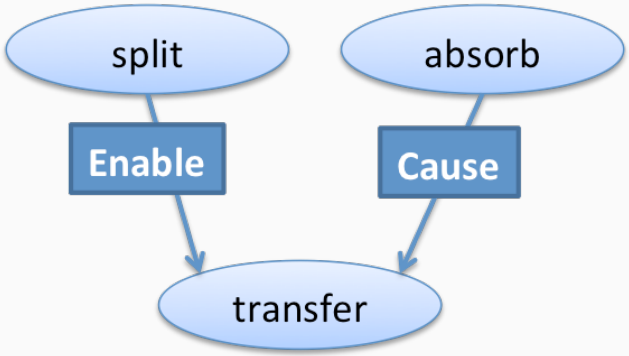


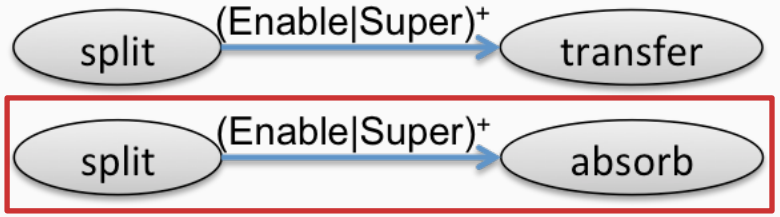


- Answering algorithm
  1. Find matching path (valid proof)

Step 3: Answering questions

- Given
  - Process structure
  - Two queries

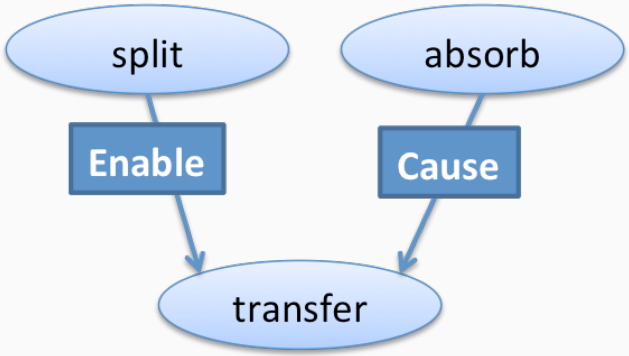


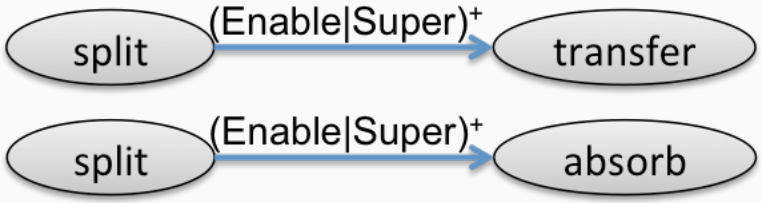


- Answering algorithm
  1. Find matching path (valid proof)
  2. Else, find contradiction of causality (refutation)

Step 3: Answering questions

- Given
  - Process structure
  - Two queries

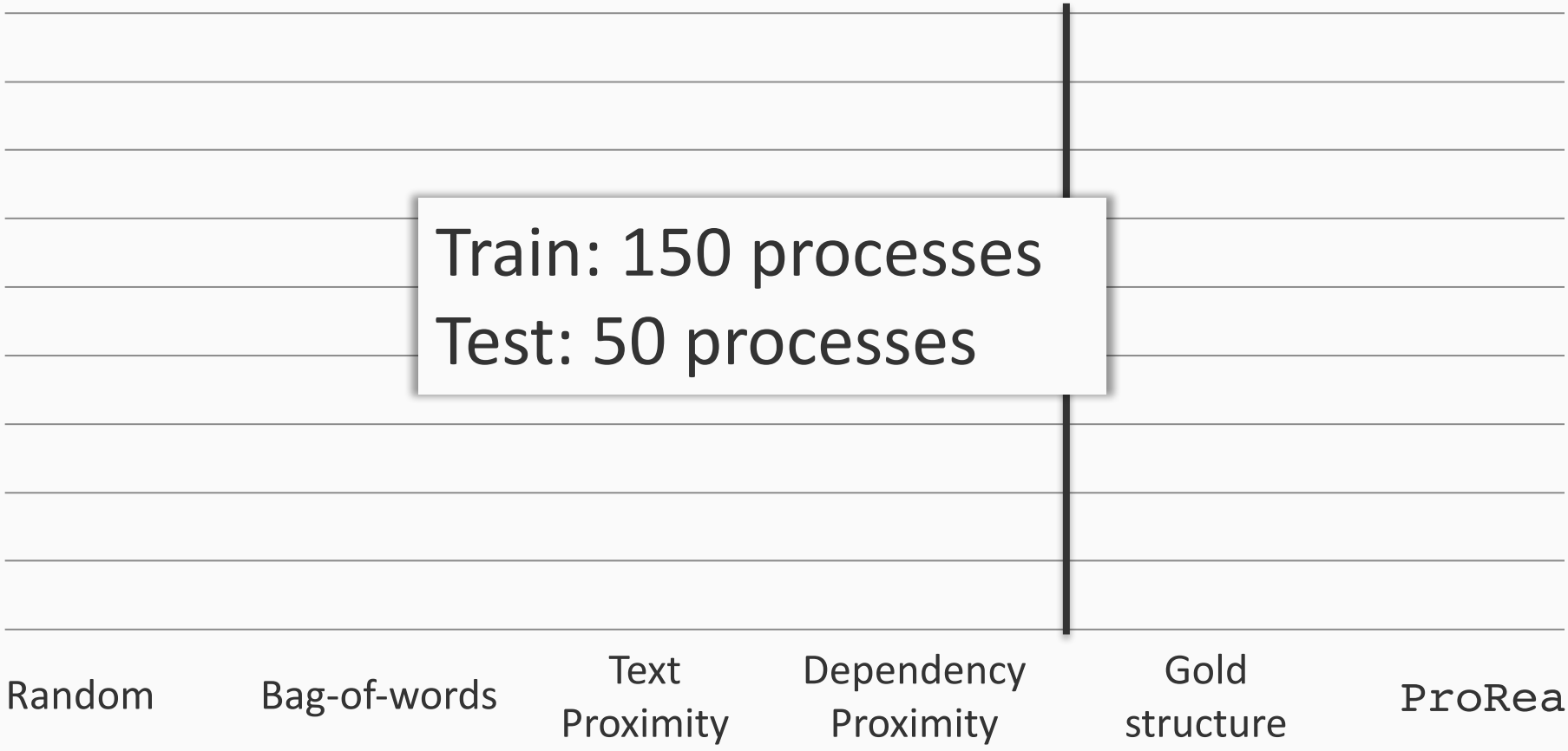




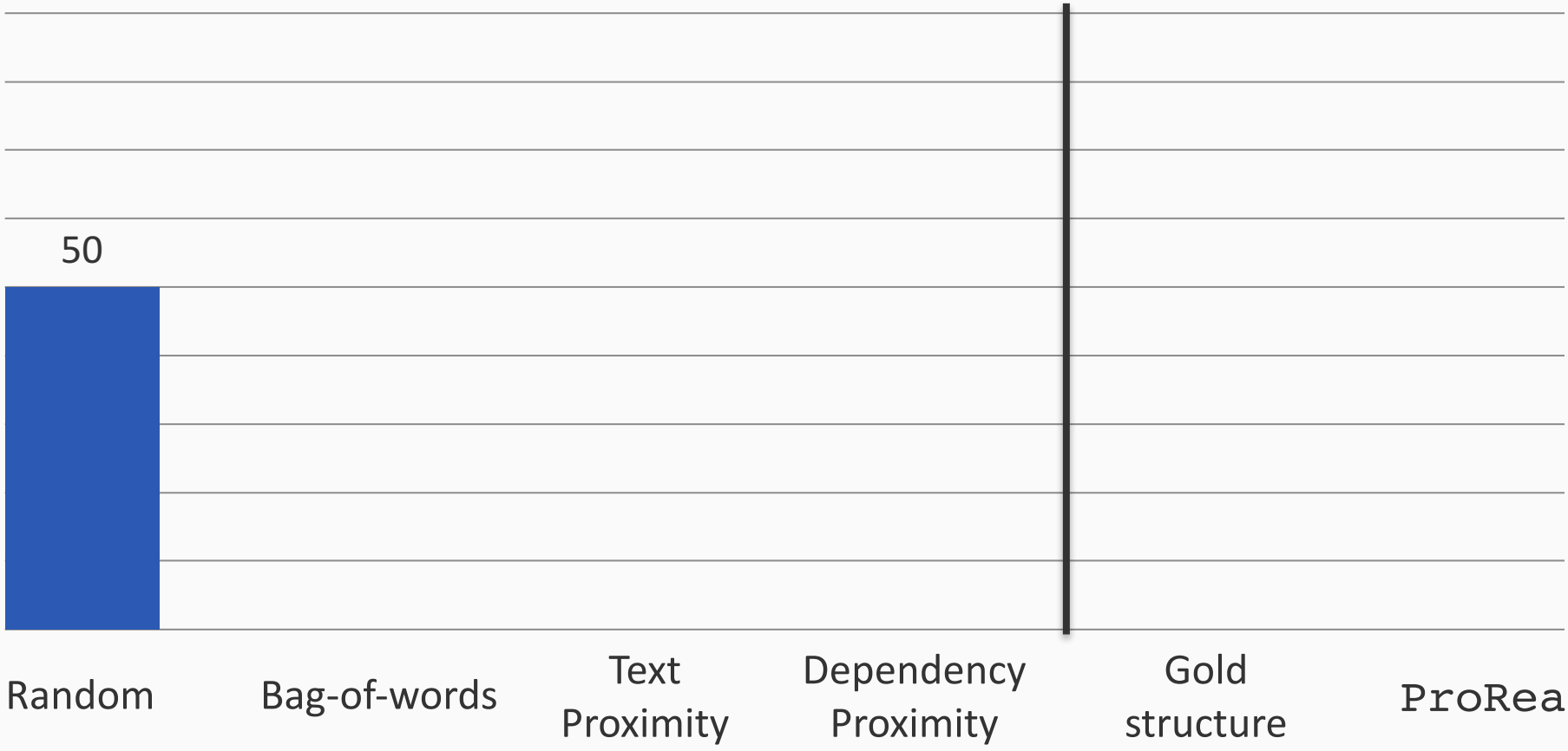
- Answering algorithm
  1. Find matching path (valid proof)
  2. Else, find contradiction of causality (refutation)
  3. Back off to baseline

Experiments and Results

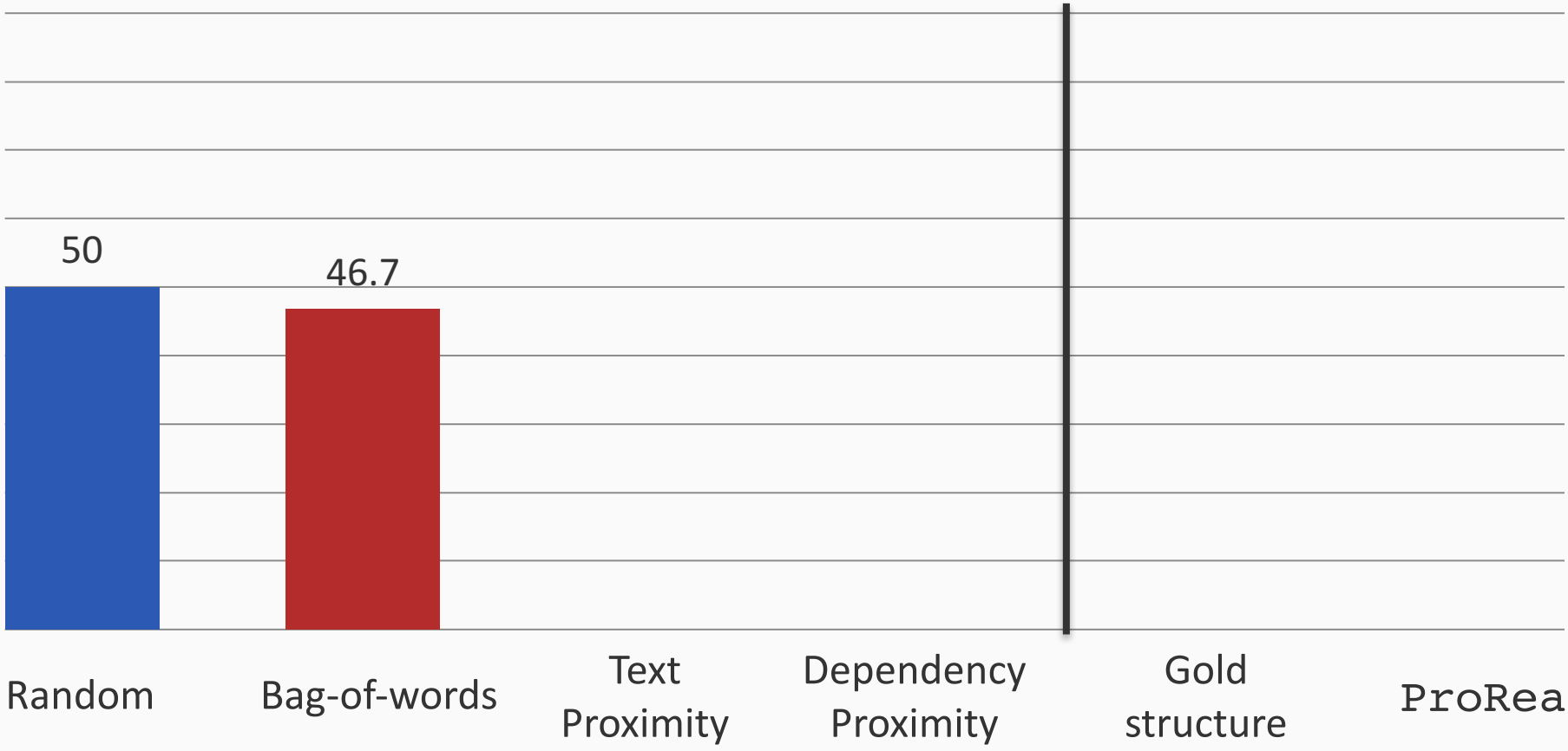
Question Answering Accuracy



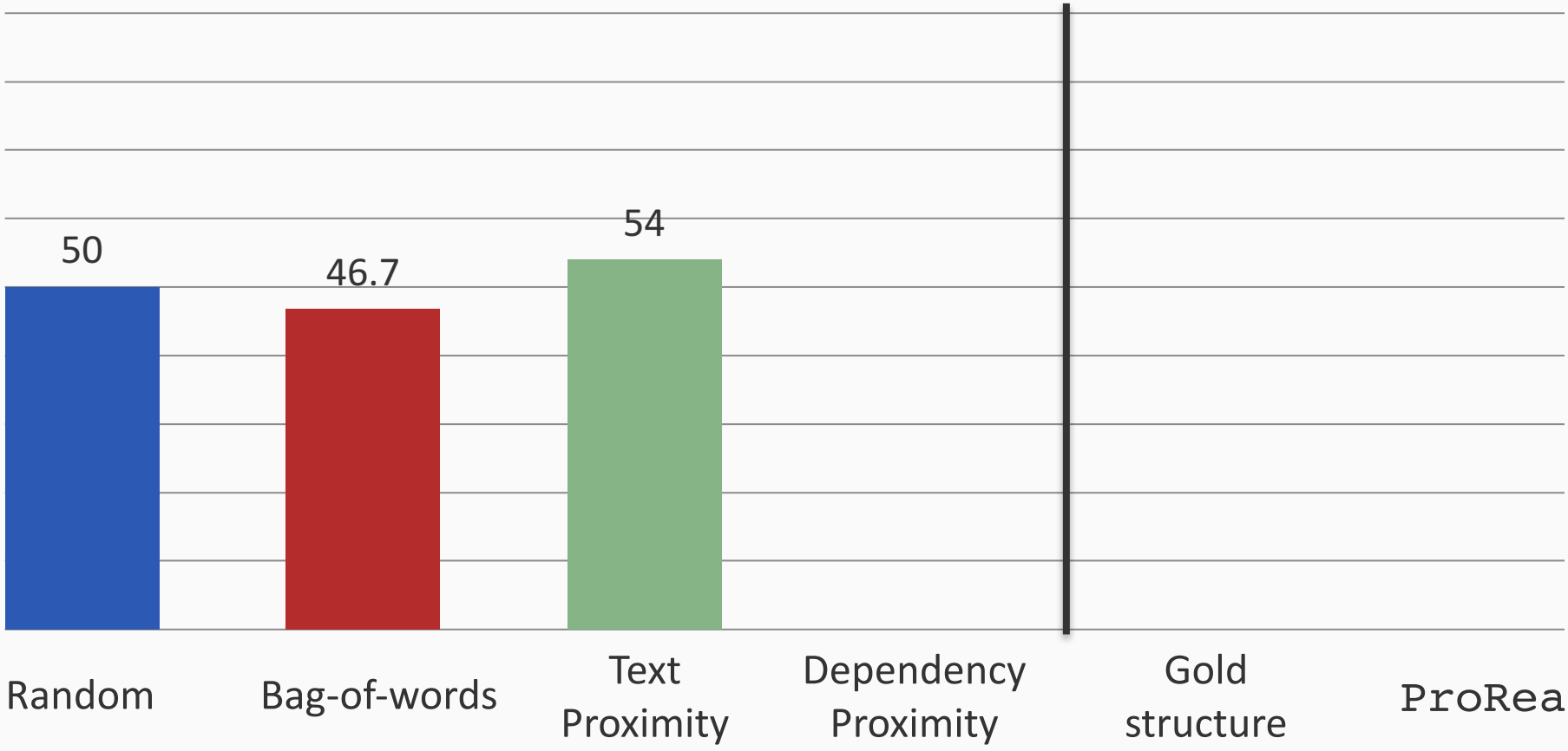
Question Answering Accuracy



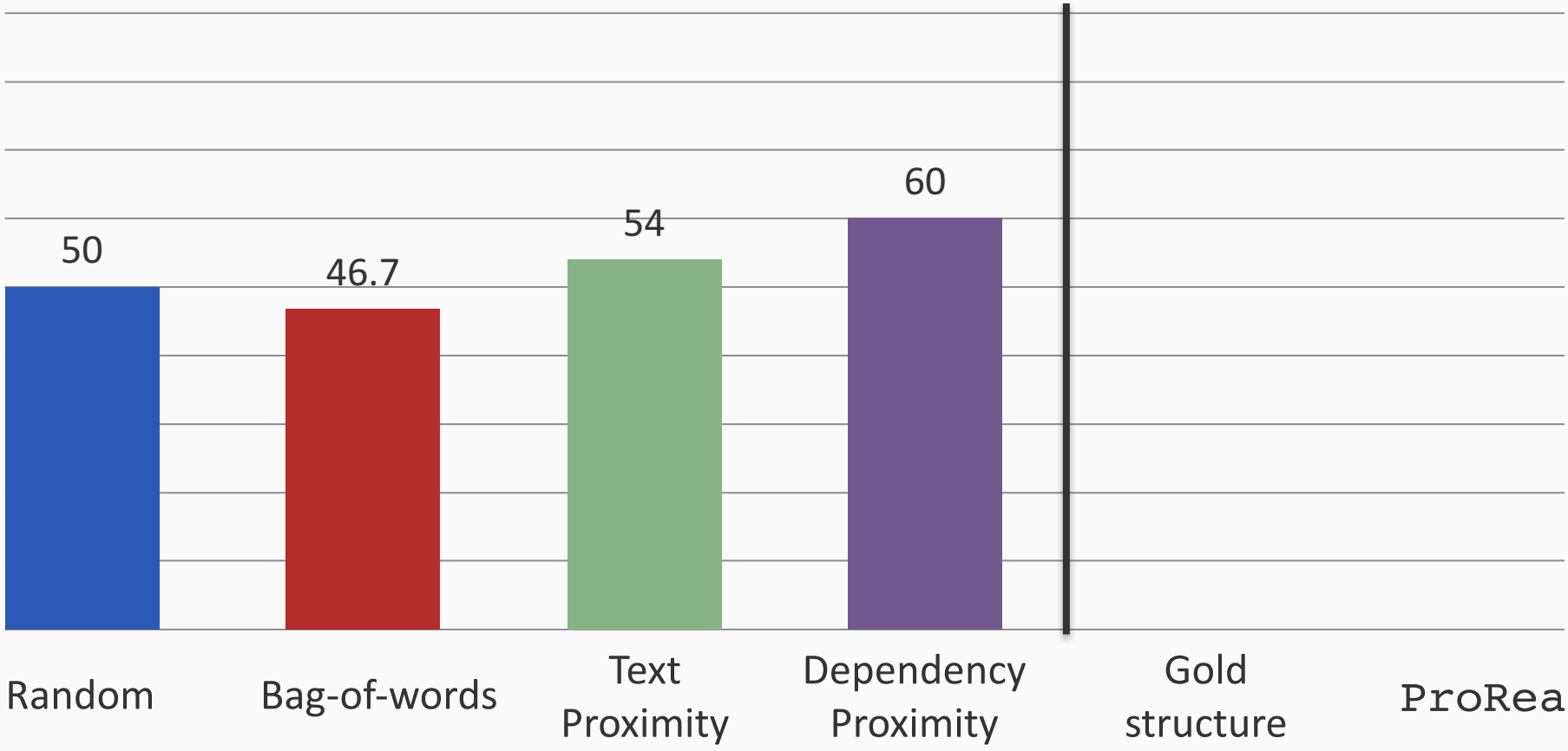
Question Answering Accuracy



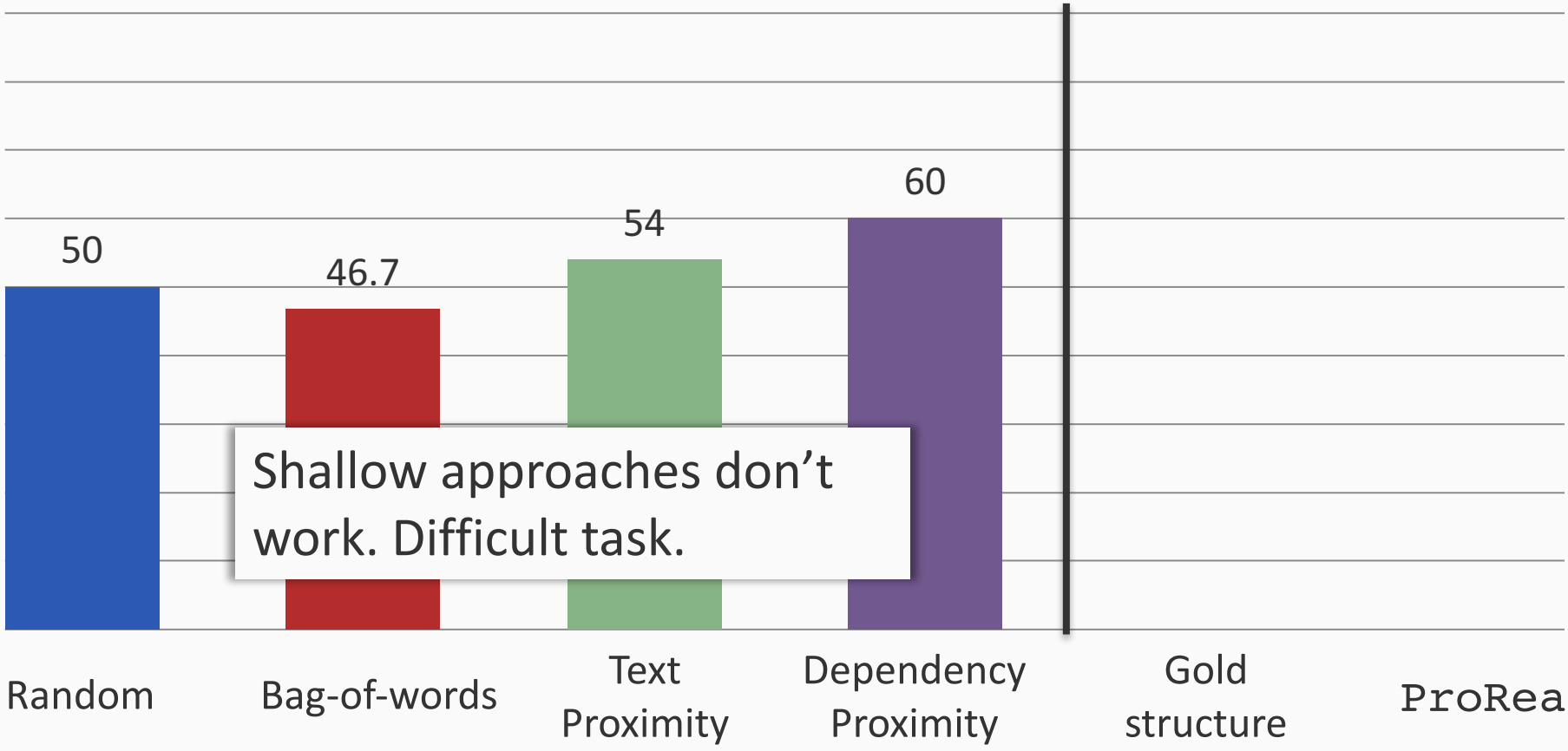
Question Answering Accuracy



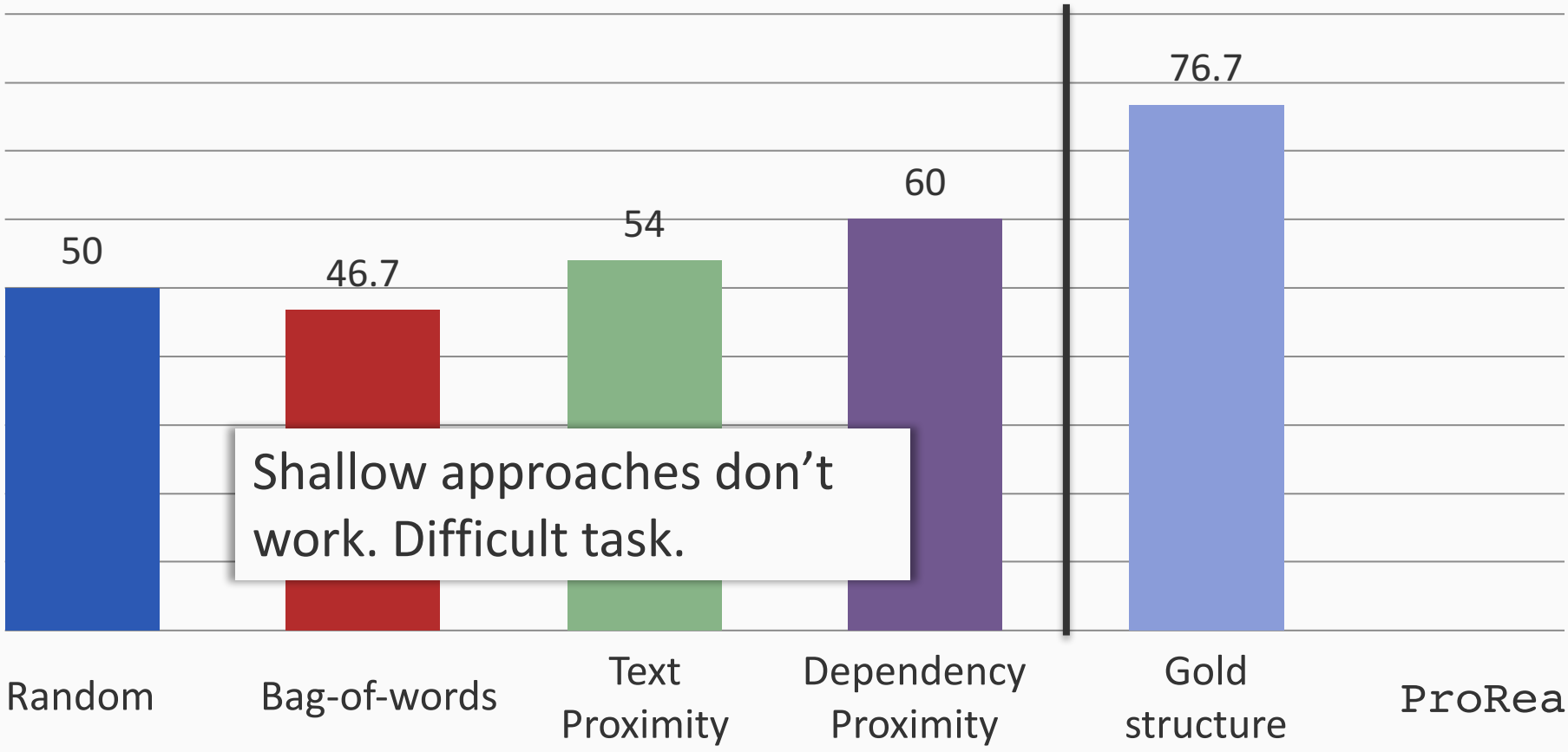
Question Answering Accuracy



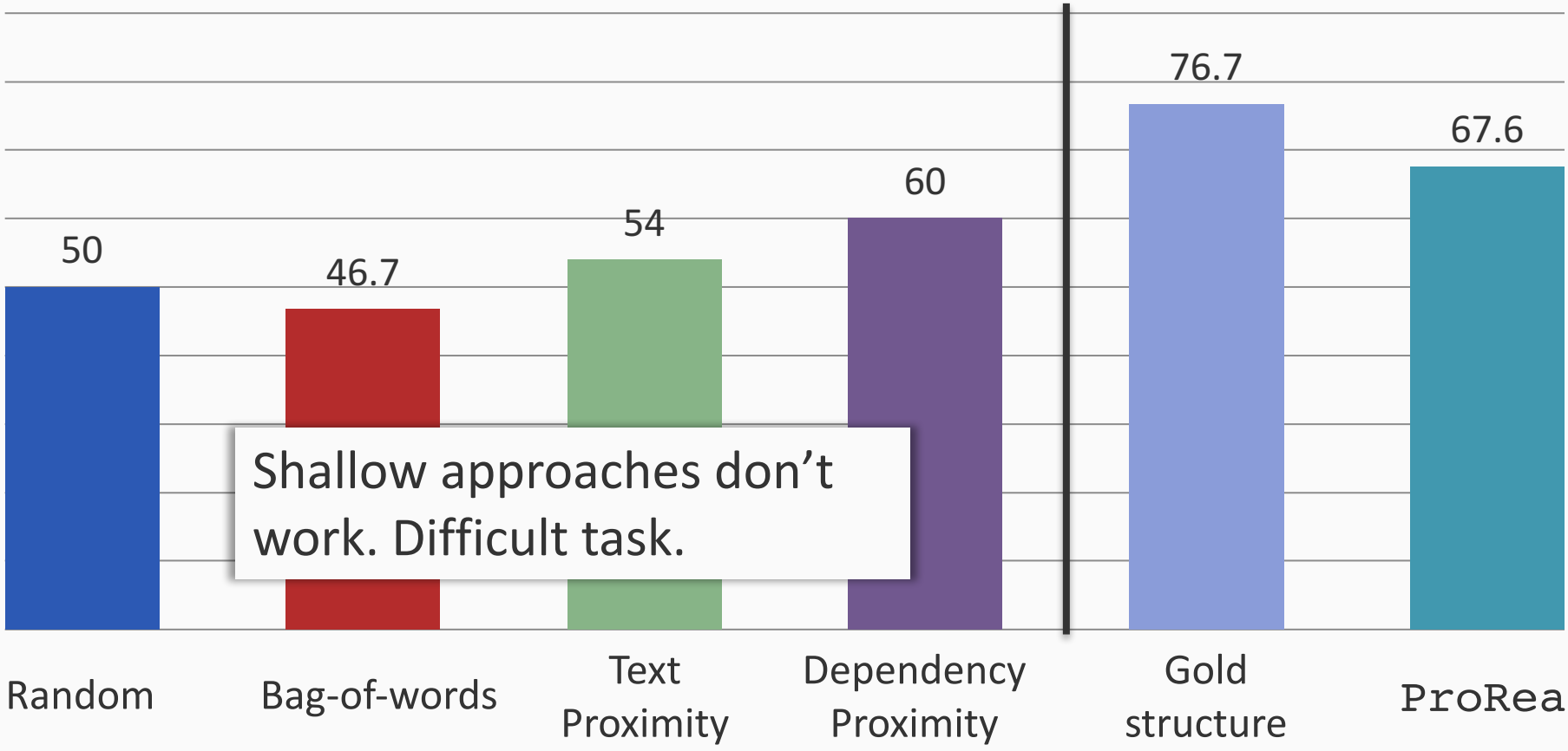
Question Answering Accuracy



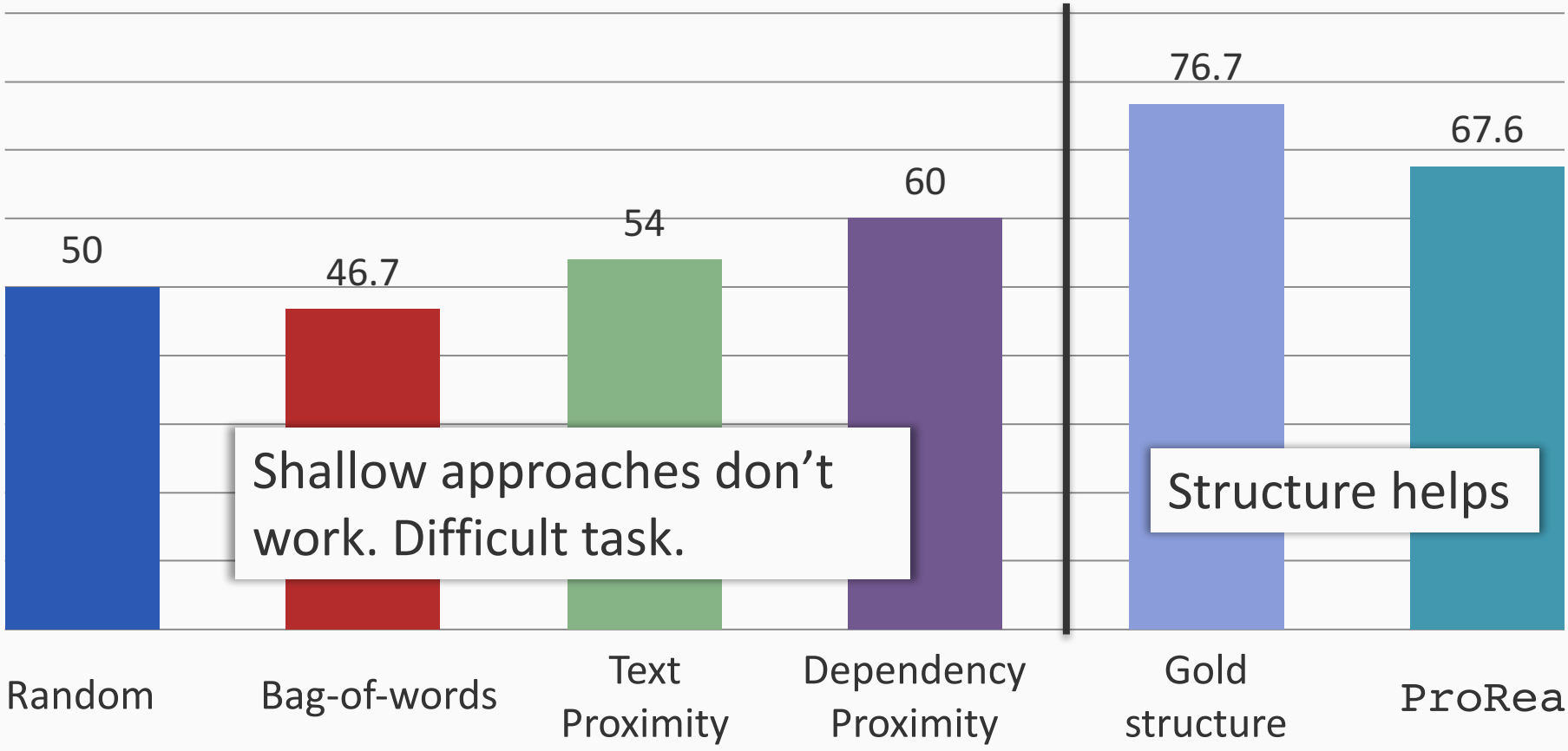
Question Answering Accuracy



Question Answering Accuracy



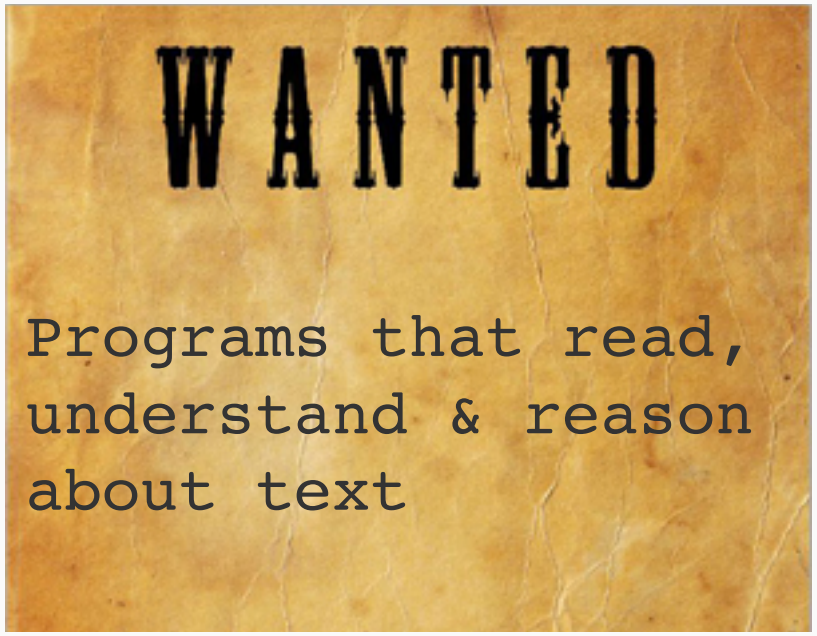
Question Answering Accuracy



Summary

- A new reading comprehension task
- A dataset of structures with Q&A: `ProcessBank`
- An end-to-end system for question answering via predicted structures
- Rich entity and event structure helps

Bigger picture



Need linguistically motivated structured representations for machine reading