Natural Language Understanding with Indirect Supervision

Daniel Khashabi

Age of Big Data

Age of Big Data

- Big data:
  - Over 56 billion pages
  - Over 500 million tweets are sent every day.
  - Over 4 million blog posts are published on the Internet every day.
- Deep learning:
  - 1.5 billion parameters [Radford et al. 2019]
  - Super-human performance [Devlin et al. 2018]

Troubling Observations

Troubling Observations

- Brittleness with respect to small changes [K at al. 2016; Jia et al. 2017; Ribeiro et al. 2018; others]

Troubling Observations

- Brittleness with respect to small changes

[K at al. 2016; Jia et al. 2017; Ribeiro et al. 2018; others]

Context: *In the United States especially, several high-profile cases such as Debra LaFave, Pamela Rogers, and Mary Kay Letourneau have caused increased scrutiny on teacher misconduct.*

Question: *What has been the result of this publicity?*

Troubling Observations

- Brittleness with respect to small changes

[K at al. 2016; Jia et al. 2017; Ribeiro et al. 2018; others]

Context: *In the United States especially, several high-profile cases such as Debra LaFave, Pamela Rogers, and Mary Kay Letourneau have caused increased scrutiny on teacher misconduct.*

Question: *What has been the result of this publicity?*

→ *"increased scrutiny on teacher misconduct"*

Troubling Observations

- Brittleness with respect to small changes [K at al. 2016; Jia et al. 2017; Ribeiro et al. 2018; others]

Context: *In the United States especially, several high-profile cases such as Debra LaFave, Pamela Rogers, and Mary Kay Letourneau have caused increased scrutiny on teacher misconduct.*

Question: *~~What has~~ been the result of this publicity?* → *"increased scrutiny on teacher misconduct"*

What's

Troubling Observations

- Brittleness with respect to small changes [K at al. 2016; Jia et al. 2017; Ribeiro et al. 2018; others]

Context: *In the United States especially, several high-profile cases such as Debra LaFave, Pamela Rogers, and Mary Kay Letourneau have caused increased scrutiny on teacher misconduct.*

Question: *~~What has~~ been the result of this publicity?* → *"increased scrutiny on teacher misconduct"*

What's → *"teacher misconduct"*

Scenarios with Little (no?) Supervision

Scenarios with Little (no?) Supervision

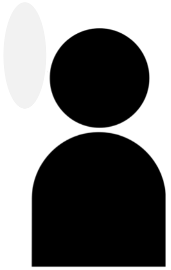
- Majority of our success has been on tasks w/ abundant **annotations.**
  - And tasks with little annotated data get the least attention.

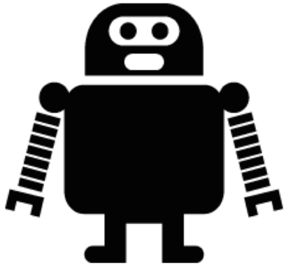
- There will be settings where there is not “enough” direct supervision.
  - Unseen/unexpected scenarios.
  - Change of style, context, domain, etc.
  - These all result in vast space of possibilities for meanings.

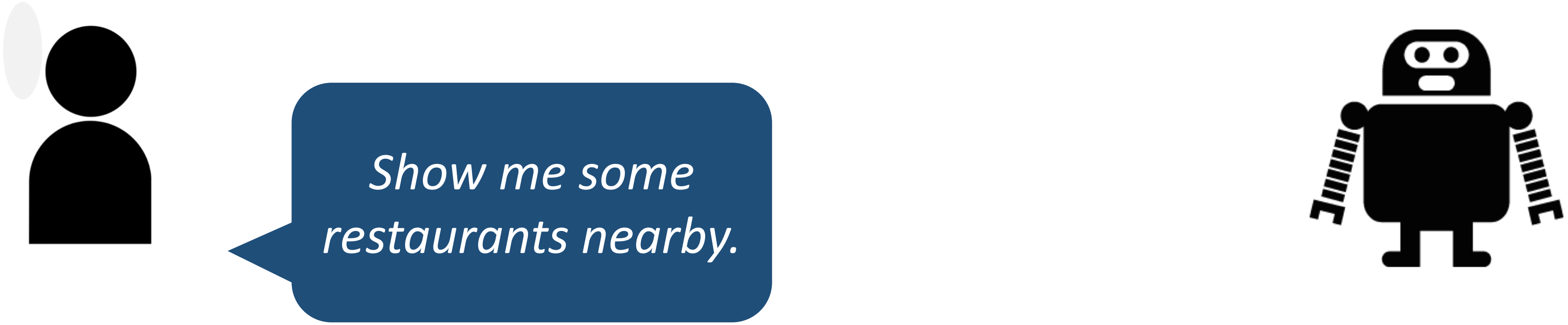
Scenarios with Little (no?) Supervision

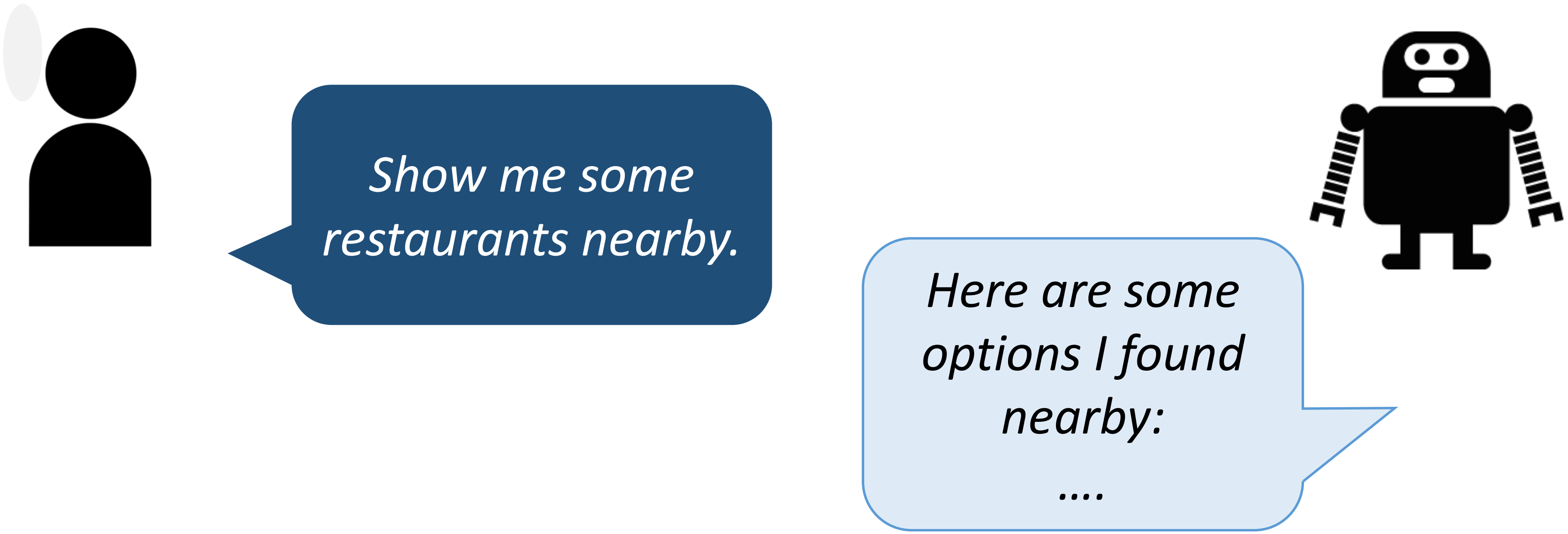
- Majority of our success has been on tasks w/ abundant **annotations.**
  - And tasks with little annotated data get the least attention.

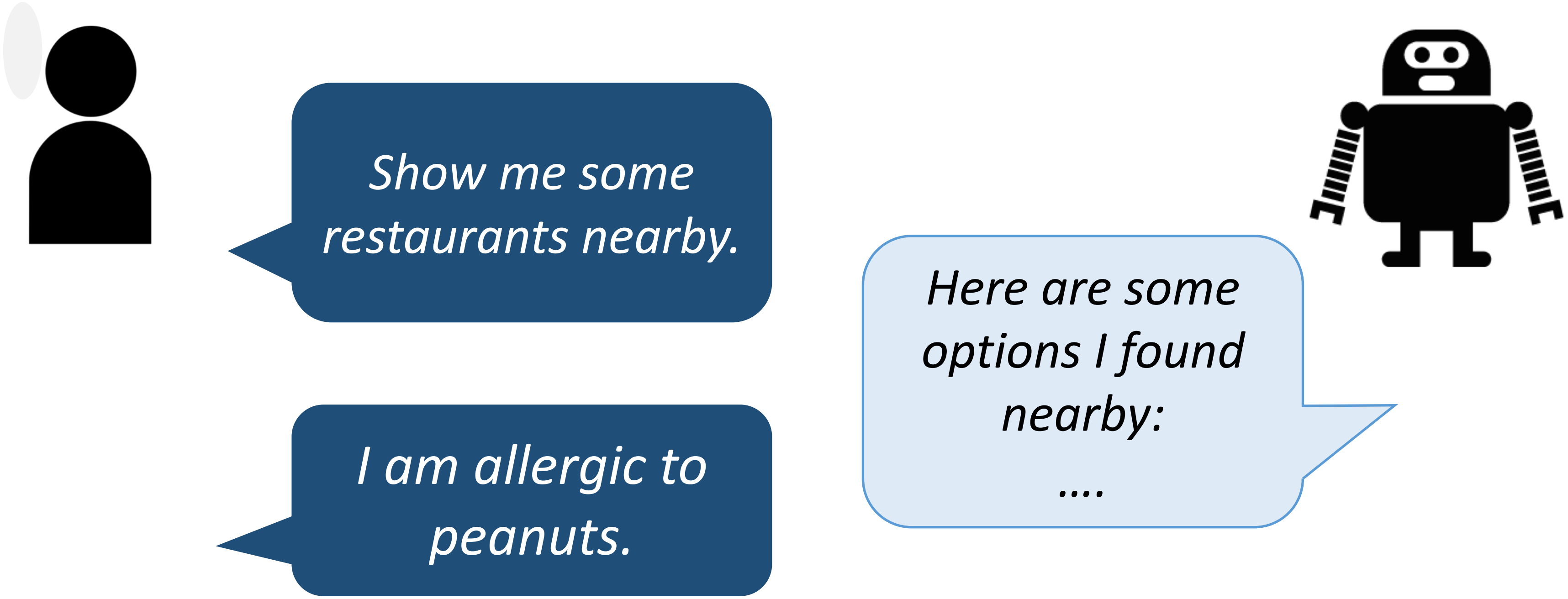
- There will be settings where there is not “enough” direct supervision.
  - Unseen/unexpected scenarios.
  - Change of style, context, domain, etc.
  - These all result in vast space of possibilities for meanings.

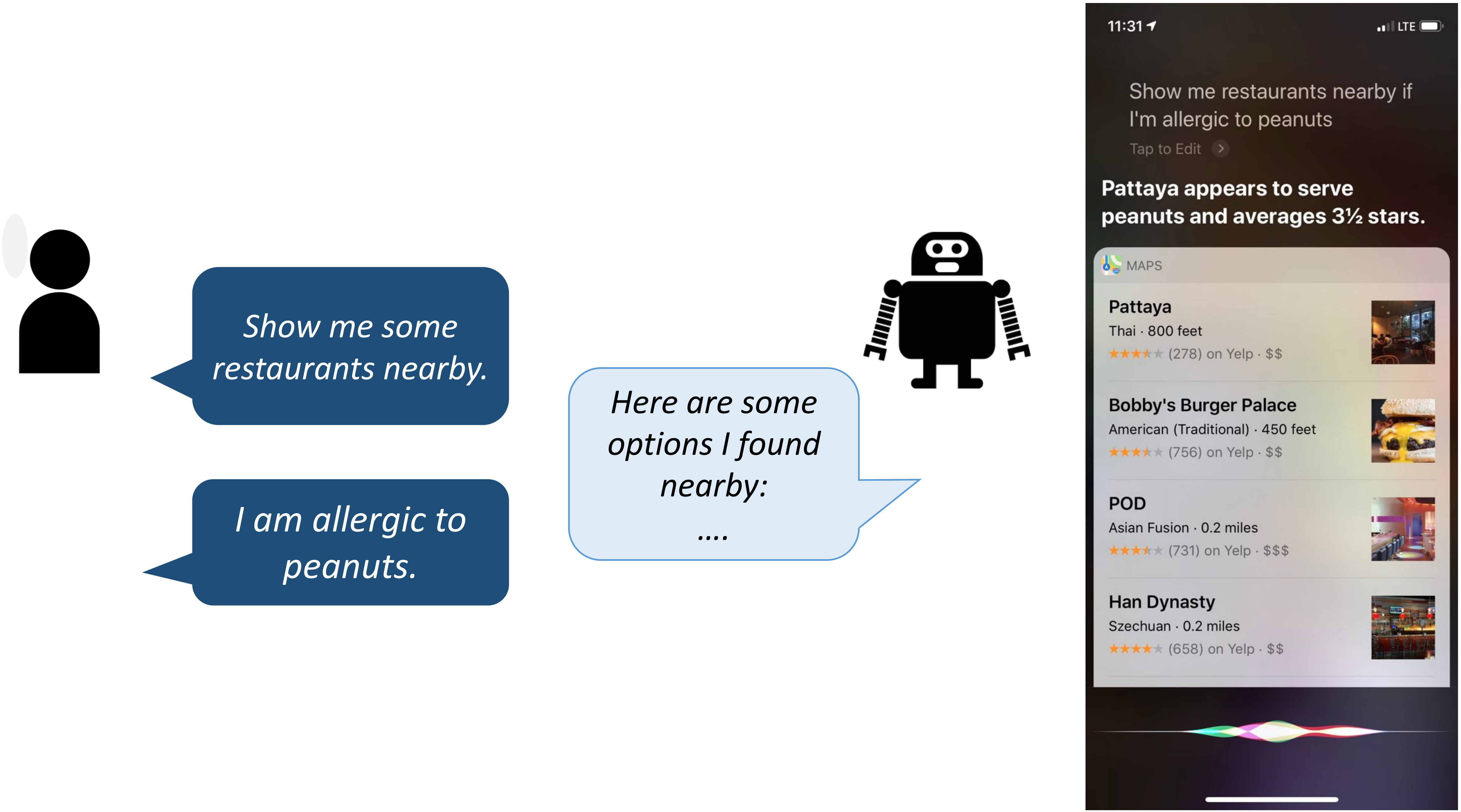




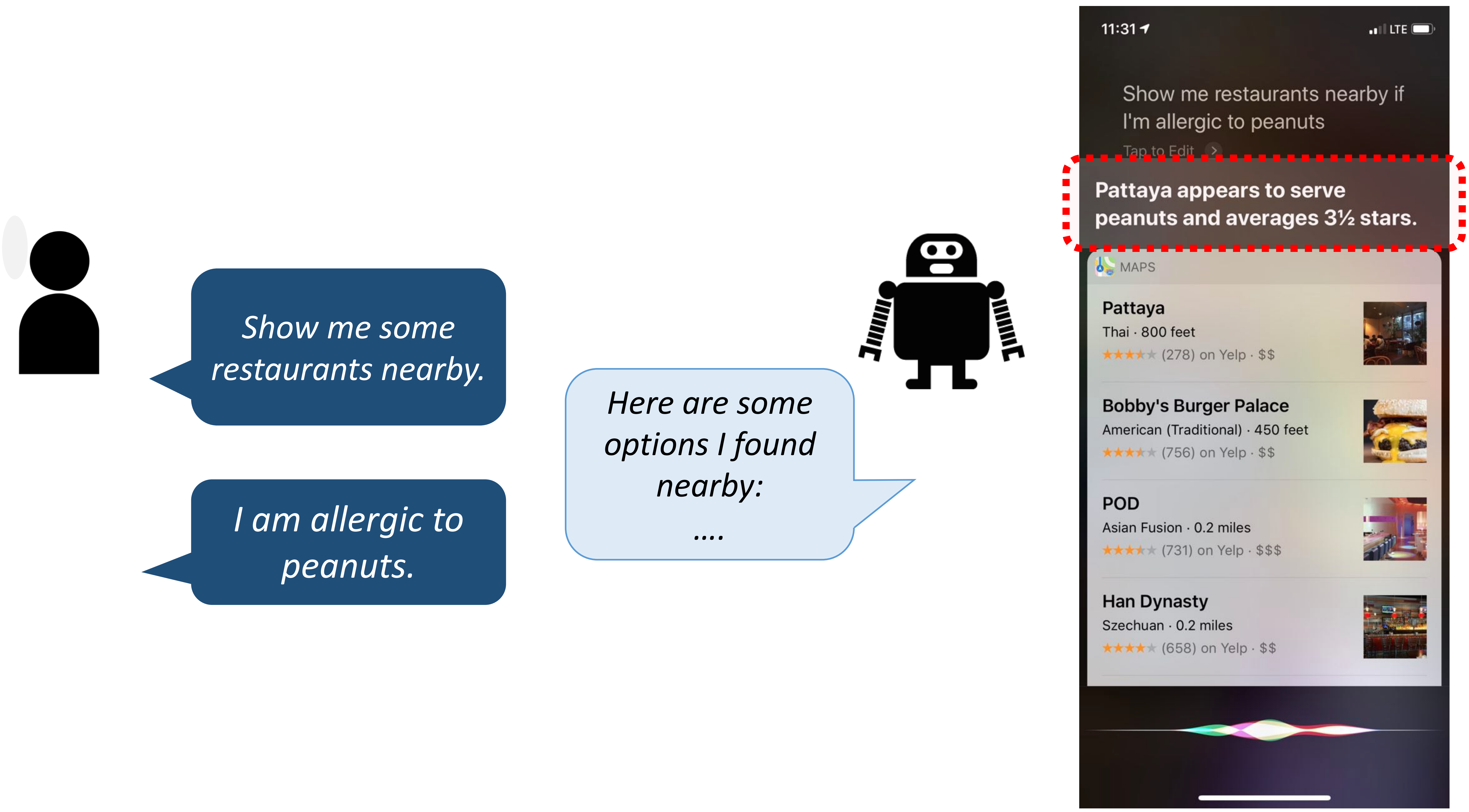

Show me some restaurants nearby.

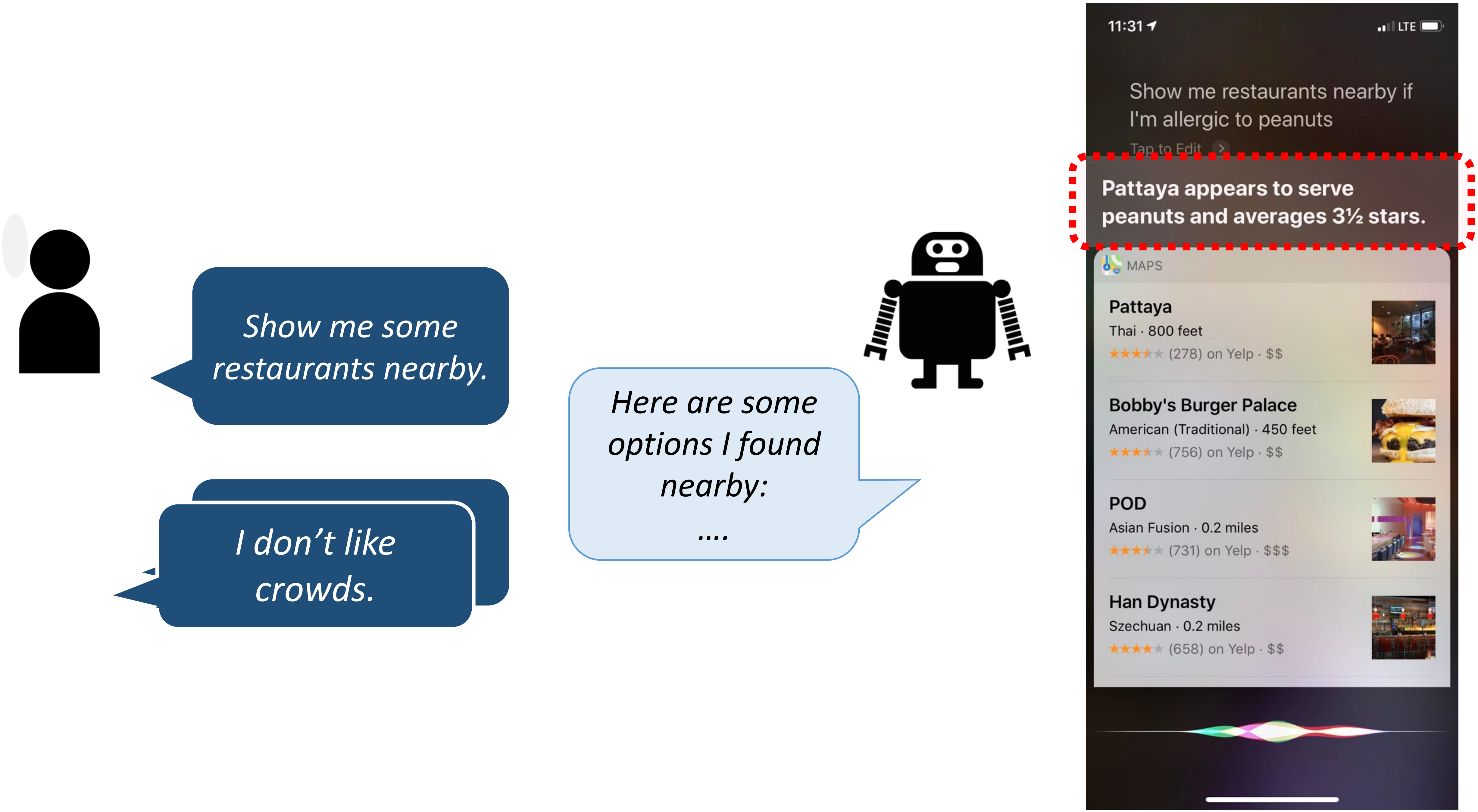

Show me some restaurants nearby.
Here are some options I found nearby:
....


Show me some restaurants nearby.
I am allergic to peanuts.
Here are some options I found nearby:
....

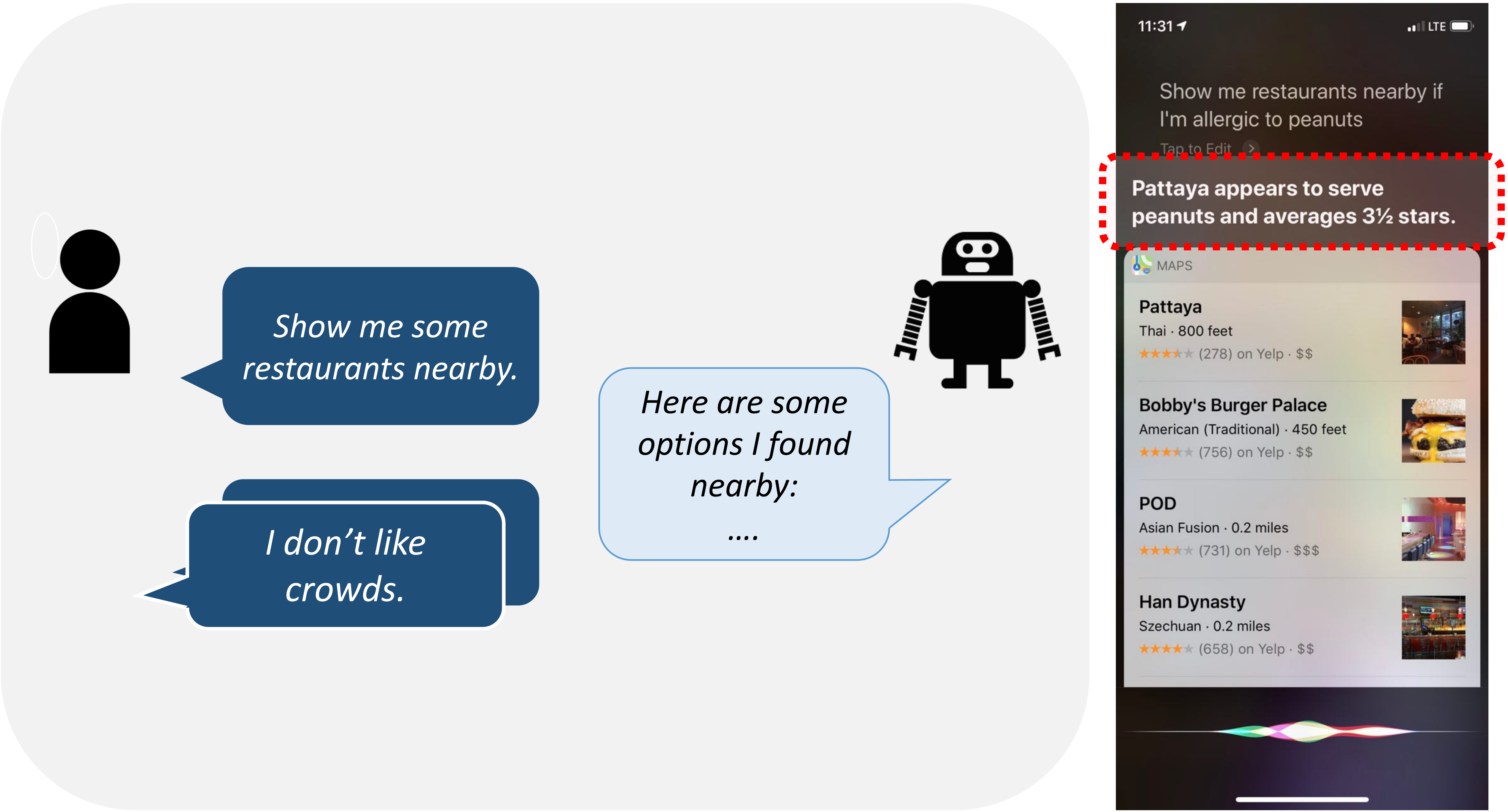


Show me some restaurants nearby.
I am allergic to peanuts.
Here are some options I found nearby:
....
11:31
LTE
Show me restaurants nearby if I'm allergic to peanuts
Tap to Edit
Pattaya appears to serve peanuts and averages 3½ stars.
MAPS
Pattaya
Thai · 800 feet
(278) on Yelp · $$
Bobby's Burger Palace
American (Traditional) · 450 feet
(756) on Yelp · $$
POD
Asian Fusion · 0.2 miles
(731) on Yelp · $$$
Han Dynasty
Szechuan · 0.2 miles
(658) on Yelp · $$


Show me some restaurants nearby.
I am allergic to peanuts.
Here are some options I found nearby:
....
11:31
LTE
Show me restaurants nearby if I'm allergic to peanuts
Tap to Edit
Pattaya appears to serve peanuts and averages 3½ stars.
MAPS
Pattaya
Thai · 800 feet
(278) on Yelp · $$
Bobby's Burger Palace
American (Traditional) · 450 feet
(756) on Yelp · $$
POD
Asian Fusion · 0.2 miles
(731) on Yelp · $$$
Han Dynasty
Szechuan · 0.2 miles
(658) on Yelp · $$



Show me some restaurants nearby.
I don't like crowds.
Here are some options I found nearby:
....
11:31
LTE
Show me restaurants nearby if I'm allergic to peanuts
Tap to Edit
Pattaya appears to serve peanuts and averages 3½ stars.
MAPS
Pattaya
Thai · 800 feet
(278) on Yelp · $$
Bobby's Burger Palace
American (Traditional) · 450 feet
(756) on Yelp · $$
POD
Asian Fusion · 0.2 miles
(731) on Yelp · $$$
Han Dynasty
Szechuan · 0.2 miles
(658) on Yelp · $$


Show me some restaurants nearby.
I don't like crowds.
Here are some options I found nearby:
....
11:31
LTE
Show me restaurants nearby if I'm allergic to peanuts
Tap to Edit
Pattaya appears to serve peanuts and averages 3½ stars.
MAPS
Pattaya
Thai · 800 feet
(278) on Yelp · $$
Bobby's Burger Palace
American (Traditional) · 450 feet
(756) on Yelp · $$
POD
Asian Fusion · 0.2 miles
(731) on Yelp · $$$
Han Dynasty
Szechuan · 0.2 miles
(658) on Yelp · $$

Talk Statement

Talk Statement

- It's unlikely that we will have directly "annotated" data that cover **all** aspects of natural language understanding.

- Data provides "hints" that exist independently of the task at hand.

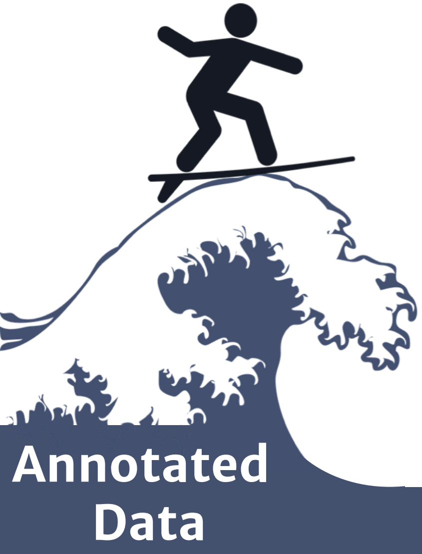
- Weak signals can be amplified to produce higher quality signals.
    - Requires effective use of representation, knowledge and putting them together.

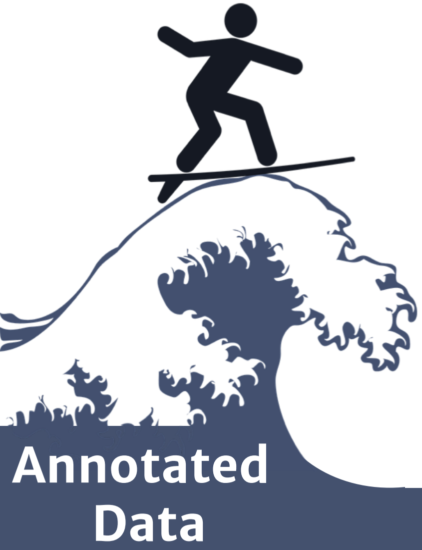
Talk Statement

- It's unlikely that we will have directly "annotated" data that cover **all** aspects of natural language understanding.

- Data provides "hints" that exist independently of the task at hand.

- Weak signals can be amplified to produce higher quality signals.
  - Requires effective use of representation, knowledge and putting them together.


Annotated
Data



This talk

“Supervision” vs “Data”

“Supervision” vs “Data”

$$\{(\mathbf{x},\mathbf{y})\} \Rightarrow$$

( ,"spam")

( ,"ham")

( ,"spam")

⋮

Supervision
(direct annotations)

“Supervision” vs “Data”

$\{(\mathbf{x},\mathbf{y})\}$ →

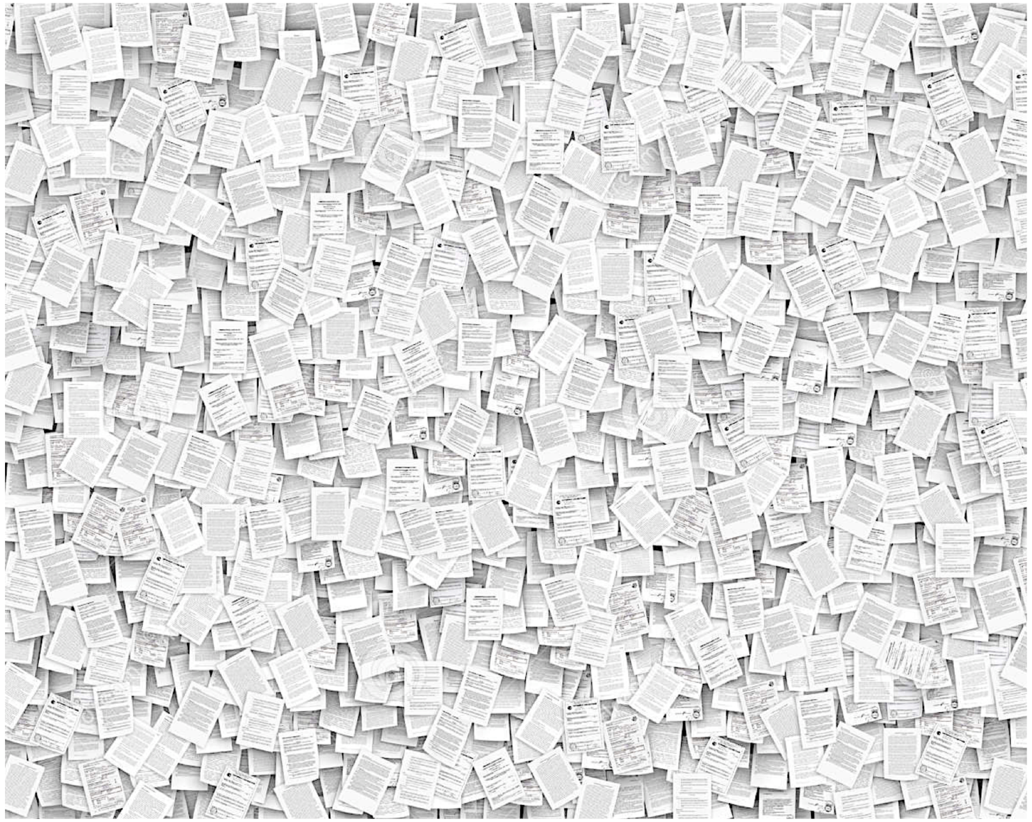
( , "spam")

( , "ham")

( , "spam")

⋮

Supervision
(direct annotations)

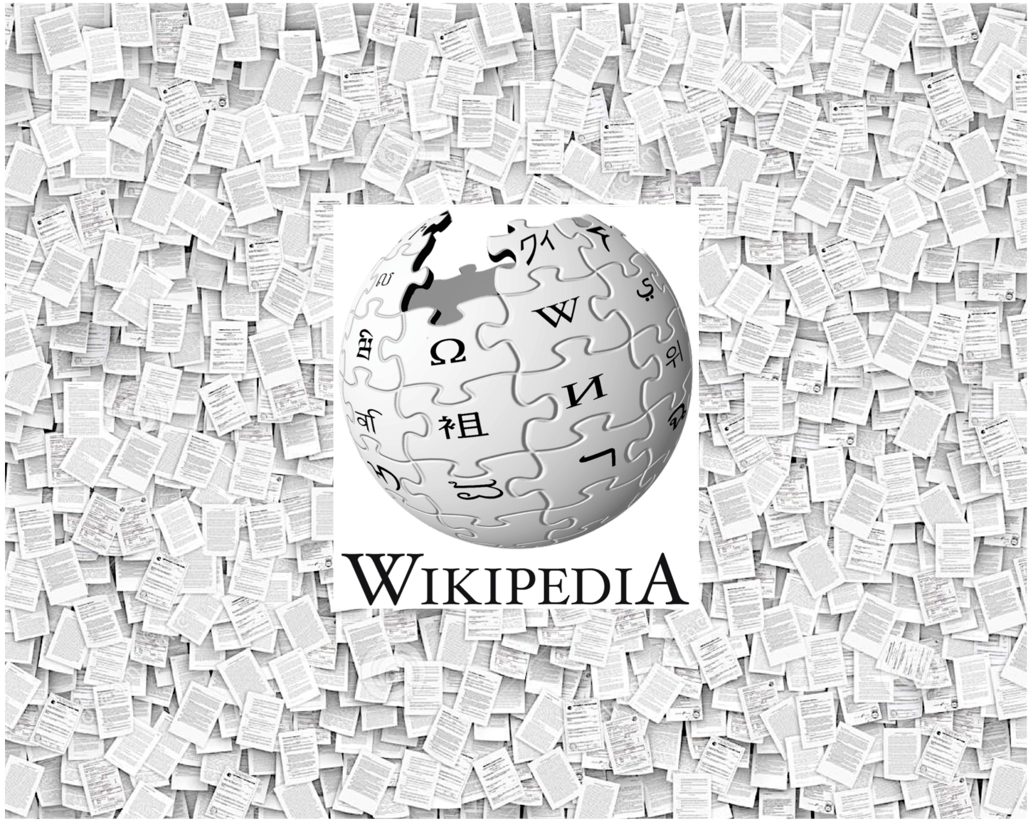


data

“Supervision” vs “Data”

$\{(\mathbf{x},\mathbf{y})\}$ → (,"spam")
(,"ham")
(,"spam")
⋮

Supervision
(direct annotations)



data

“Supervision” vs “Data”

Minimal

$\{(\mathbf{x},\mathbf{y})\}$ →

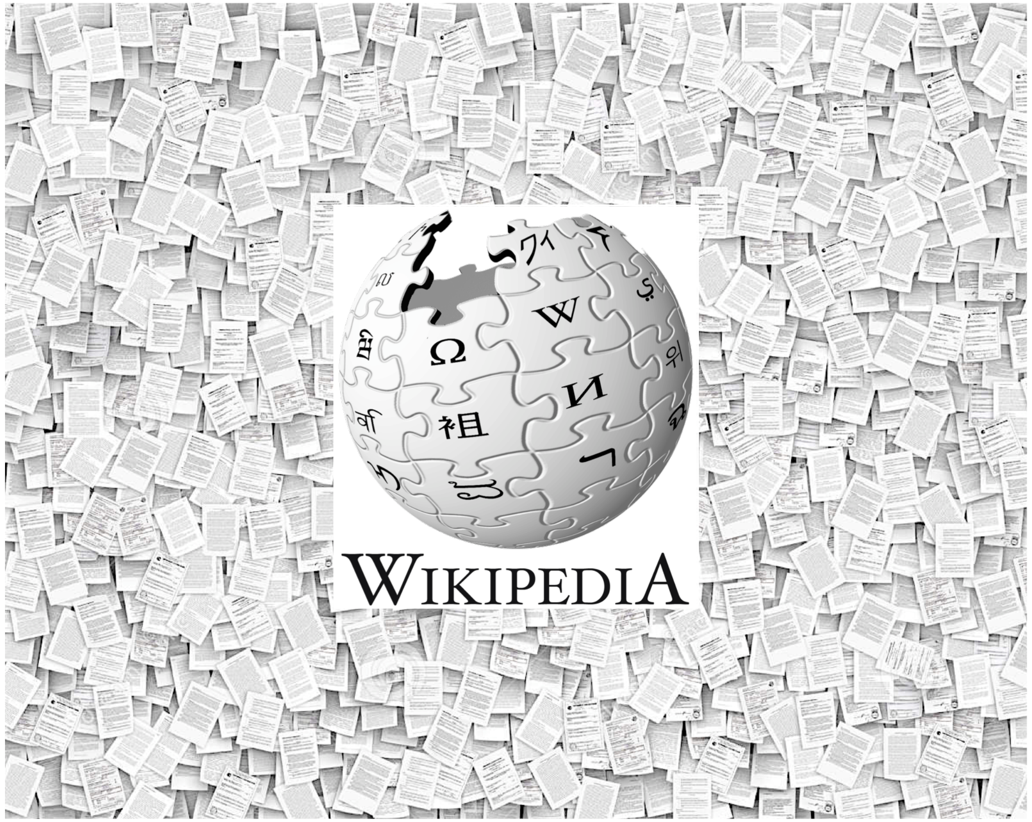
( ,"spam")

( , "ham")

( ,"spam")

⋮

Supervision
(direct annotations)



data

“Supervision” vs “Data”

Minimal

$\{(\mathbf{x},\mathbf{y})\} \rightarrow$

( ,"spam")

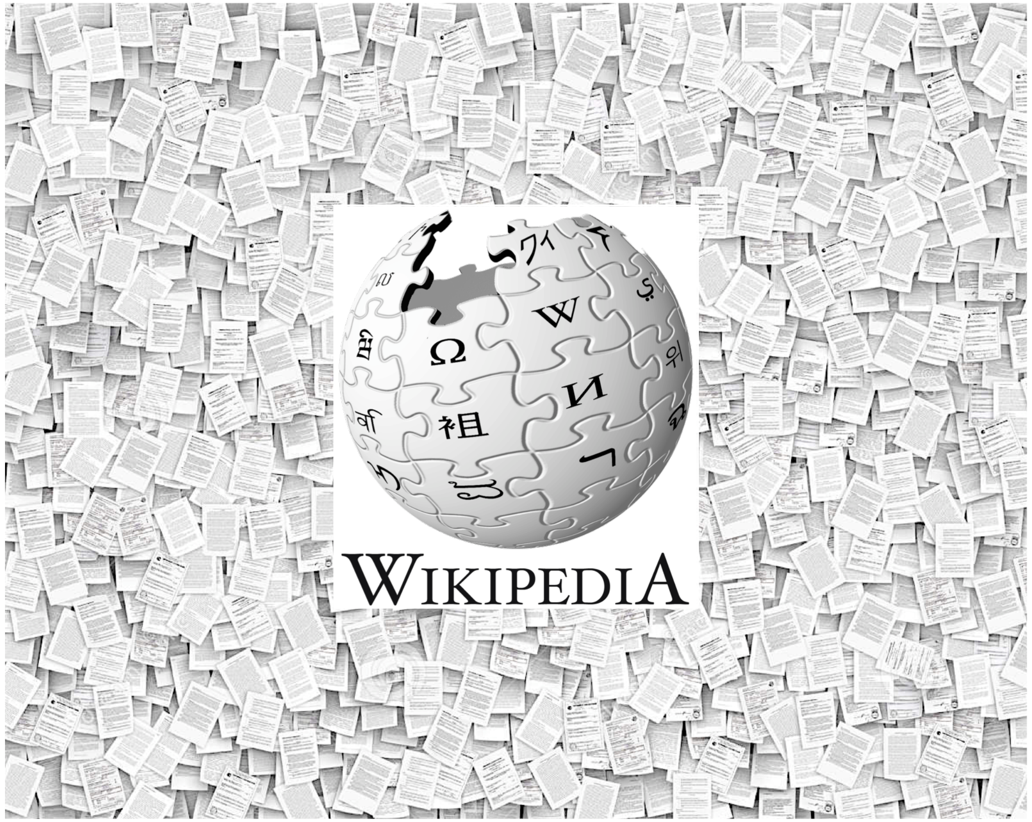
( ,"ham")

( ,"spam")

⋮

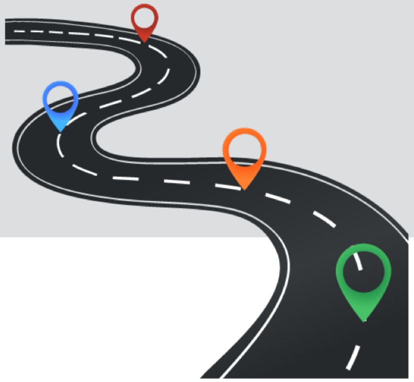
Supervision
(direct annotations)

Abundant

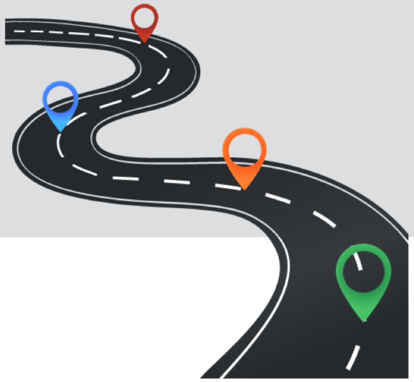


data

This talk

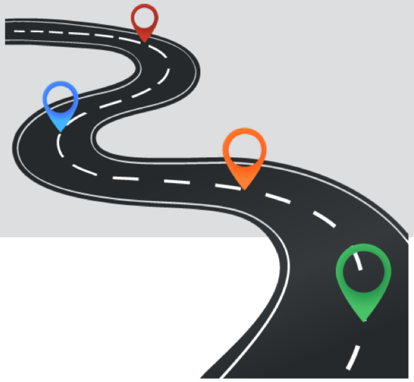


This talk



with minimal supervision

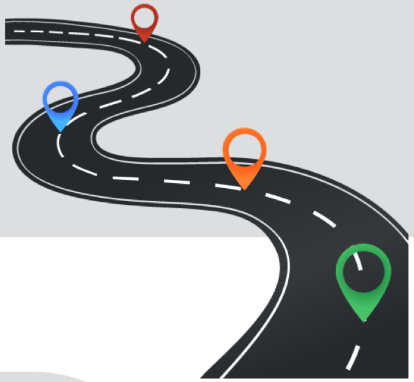
This talk



with minimal supervision

- Representations
- Wikipedia
- Structure of the problem
- Compositionality
- Other learned models
- …

This talk



▣ Introduction

▢ Answering Questions

▢ Semantic Typing of Entities

▢ Future Work

with minimal supervision

- Representations
- Wikipedia
- Structure of the problem
- Compositionality
- Other learned models
- ...

ANSWERING QUESTIONS *with minimal supervision*

K et al. Question Answering as Global Reasoning over Semantic Abstractions. AAAI 18.

K et al. Question Answering via Integer Programming over Semi-Structured Knowledge. IJCAI 16.

Clark, EKSTT**K**. Combining Retrieval, Statistics, and Inference to Answer Elementary Science Questions. AAAI 16.

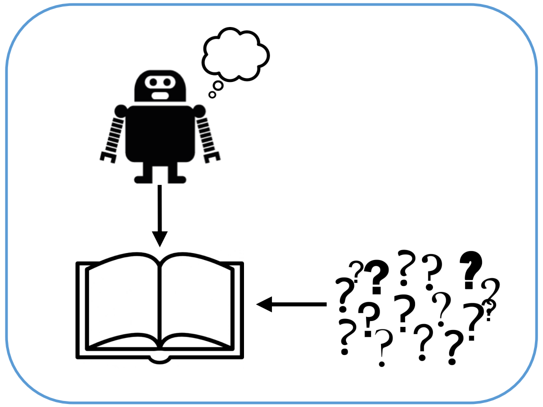
Why Answering Questions?

Why Answering Questions?

- The grand goal: *Natural Language Understanding (NLU).*
- Measuring progress by answering questions.
  - A system that is better at understanding language should have a higher chance of answering questions.
  - This has been used in the field for many years.
    [Winograd, 1972; Lehnert, 1977b; others]
  - Question Answering (QA), Reading Comprehension (RC), Textual Entailment (TE), etc.

Why Answering Questions?

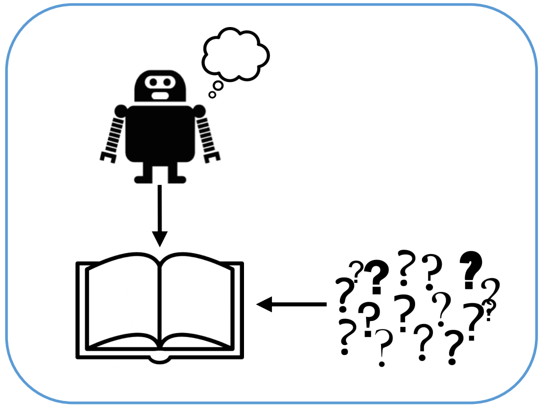
- The grand goal: *Natural Language Understanding (NLU).*
- Measuring progress by answering questions.
  - A system that is better at understanding language should have a higher chance of answering questions.



  - This has been used in the field for many years.
    [Winograd, 1972; Lehnert, 1977b; others]
  - Question Answering (QA), Reading Comprehension (RC), Textual Entailment (TE), etc.

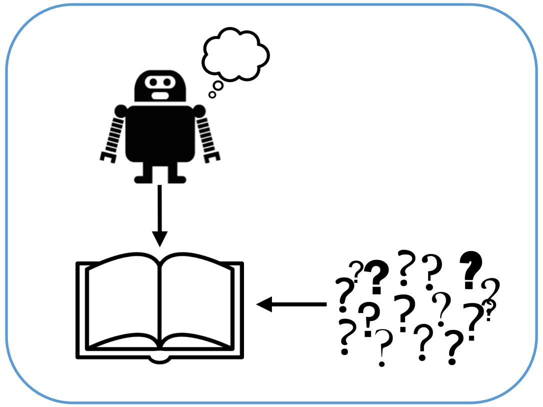
Why Answering Questions?

- The grand goal: *Natural Language Understanding (NLU).*
- Measuring progress by answering questions.
  - A system that is better at understanding language should have a higher chance of answering questions.
  - This has been used in the field for many years.
    [Winograd, 1972; Lehnert, 1977b; others]
  - Question Answering (QA), Reading Comprehension (RC), Textual Entailment (TE), etc.



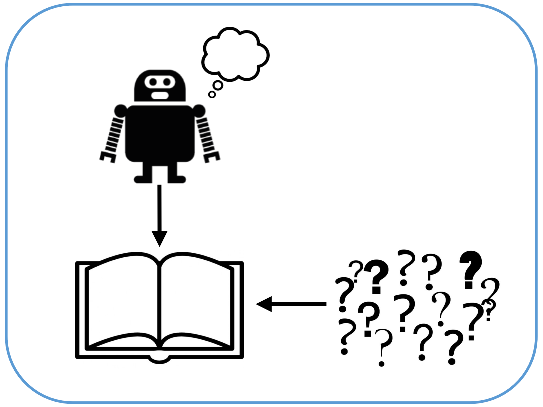
Why Answering Questions?

- The grand goal: *Natural Language Understanding (NLU).*
- Measuring progress by answering questions.
  - A system that is better at understanding language should have a higher chance of answering questions.
  - This has been used in the field for many years.
    [Winograd, 1972; Lehnert, 1977b; others]
  - Question Answering (QA), Reading Comprehension (RC), Textual Entailment (TE), etc.



Why Answering Questions?

- The grand goal: *Natural Language Understanding (NLU).*
- Measuring progress by answering questions.
    - A system that is better at understanding language should have a higher chance of answering questions.
    - This has been used in the field for many years. [Winograd, 1972; Lehnert, 1977b; others]
    - Question Answering (QA), Reading Comprehension (RC), Textual Entailment (TE), etc.



Answering Questions: The Setting

Answering Questions: The Setting

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Answering Questions: The Setting

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

\+

Attached to each question is an **evidence paragraph**, potentially with the answer to the question.

Answering Questions: The Setting

- Standardized science exams. [Clark et al. 2015]
- Simple language; machines require the ability to use the knowledge and abstract over it.

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

+ 

Attached to each question is an **evidence paragraph**, potentially with the answer to the question.

Linguistic Variability

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Linguistic Variability

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

… and a bear survives winters using its thick fur …

Linguistic Variability

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

... and a bear survives winters using its thick fur ...

Linguistic Variability

Question: *A bear survives winters with what structure?* **+**

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

… and a bear survives winters using its thick fur …

Linguistic Variability

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Linguistic Variability

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

… Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities. …

Linguistic Variability

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

... Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities. ...

A given "meaning" can be phrased in many surface forms!

Linguistic Variability

Question: *A bear survives winters with what structure?* +

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

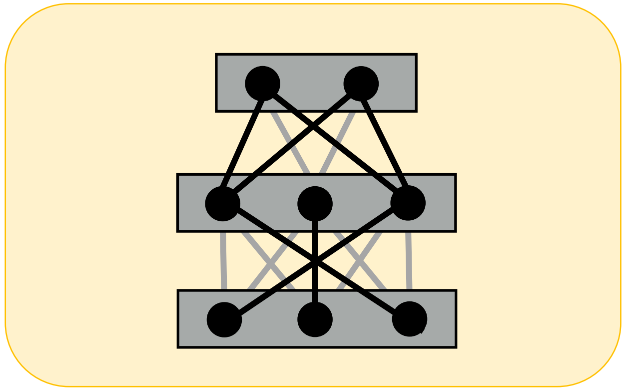
Evidence paragraph

Polar bears have white fur so that they can camouflage into their environment. Their coat is so well camouflaged in Arctic environments that it can sometimes pass as a snow drift. They have a thick layer of body fat, which keeps them warm while swimming, and a double-layered coat that insulates them from the cold Arctic air.

Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities. *Polar bears' lives depend wholly on the sea, their main source of food, and the place they spend most of their lives. But as the climate warms, that ice is melting, threatening polar bears. A common method of hunting by polar bears involves the bear keeping perfectly still by a seal's breathing hole, waiting for hours—or even days—for a seal to pop up for air.*

A Common Approach: Supervised Learning

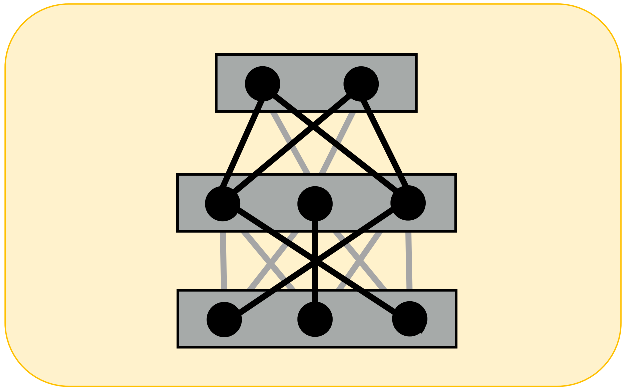
A Common Approach: Supervised Learning



A Common Approach: Supervised Learning

Question: *A bear survives … ?*

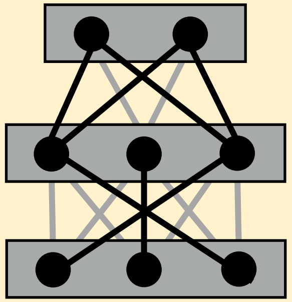
Evidence paragraph



A Common Approach: Supervised Learning

Question: *A bear survives … ?*

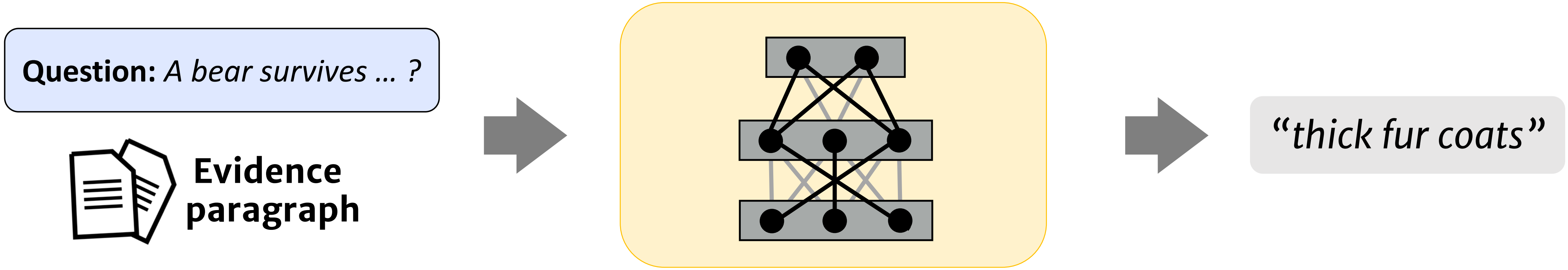
Evidence paragraph



“thick fur coats”

A Common Approach: Supervised Learning

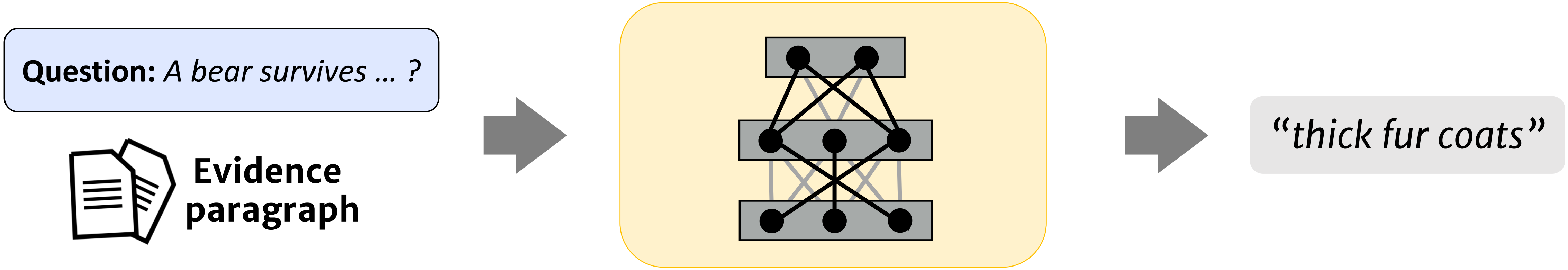
- **Input:** question, an evidence paragraph.
- **Output:** predicted answer.



- Much success: Mostly with abundantly annotated data.
- Things can break down!

A Common Approach: Supervised Learning

- **Input:** question, an evidence paragraph.
- **Output:** predicted answer.



- Much success: Mostly with abundantly annotated data.
- Things can break down!

Question: *A bear survives winters with what structure?*

Polar bears have white fur so that they can camouflage into their environment. Their coat is so well camouflaged in Arctic environments that it can sometimes pass as a snow drift. They have a thick layer of body fat, which keeps them warm while swimming, and a double-layered coat that insulates them from the cold Arctic air.

Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities. *Polar bears' lives depend wholly on the sea, their main source of food, and the place they spend most of their lives. But as the climate warms, that ice is melting, threatening polar bears. A common method of hunting by polar bears involves the bear keeping perfectly still by a seal's breathing hole, waiting for hours—or even days—for a seal to pop up for air.*

[Fetched on March 26, 2019]

https://demo.allennlp.org

[Seo et al, 17, Gardner et al, 18]

Question: *A bear survives winters with what structure?*

Polar bears have white fur so that they can camouflage into their environment. Their coat is so well camouflaged in Arctic environments that it can sometimes pass as a snow drift. They have a thick layer of body fat, which keeps them warm while swimming, and a double-layered coat that insulates them from the cold Arctic air.

Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities. *Polar bears' lives depend wholly on the sea, their main source of food, and the place they spend most of their lives. But as the climate warms, that ice is melting, threatening polar bears. A common method of hunting by polar bears involves the bear keeping perfectly still by a seal's breathing hole, waiting for hours—or even days—for a seal to pop up for air.*

Predicted Answer

[Fetched on March 26, 2019]

https://demo.allennlp.org

[Seo et al, 17, Gardner et al, 18]

Question: *A bear survives winters with what structure?*

Polar bears have white fur so that they can camouflage into their environment. Their coat is so well camouflaged in Arctic environments that it can sometimes pass as a snow drift. They have a thick layer of body fat, which keeps them warm while swimming, and a double-layered coat that insulates them from the cold Arctic air.

Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities. *Polar bears' lives depend wholly on the sea, their main source of food, and the place they spend most of their lives. But as the climate warms, that ice is melting, threatening polar bears. A common method of hunting by polar bears involves the bear keeping perfectly still by a seal's breathing hole, waiting for hours—or even days—for a seal to pop up for air.*

Predicted Answer

- Can we "explain" the decision?
- Can we "fix" such behaviors?

[Fetched on March 26, 2019]
https://demo.allennlp.org
[Seo et al, 17, Gardner et al, 18]

Semi-Structured Inference: High-level View

Question Answering

as Global Reasoning

over Semi-Structured Knowledge

Question Answering

Semi-Structured Inference: High-level View

Question Answering

as Global Reasoning

over Semi-Structured Knowledge

Question Answering

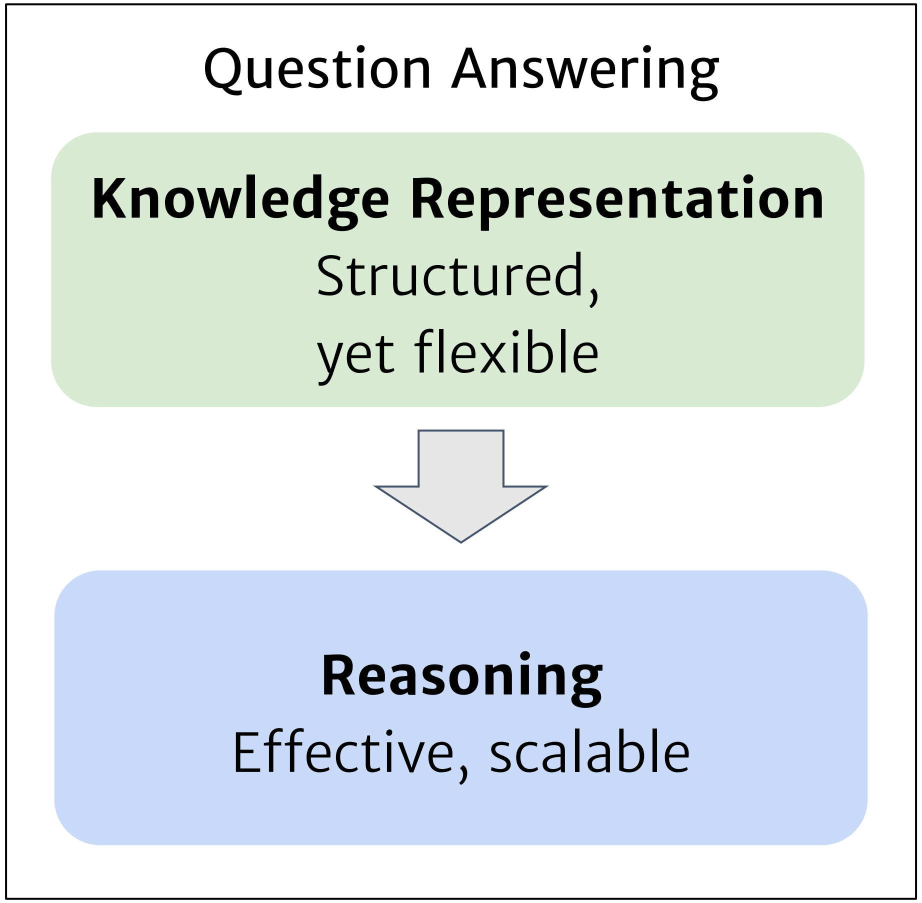
Knowledge Representation
Structured,
yet flexible

Semi-Structured Inference: High-level View

Question Answering

as Global Reasoning

over Semi-Structured Knowledge



Language Understanding Phenomena

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

*Polar bears, saved from the bitter cold **by** their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.*

Language Understanding Phenomena

verb

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

*Polar bears, saved from the bitter cold **by** their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.*

Language Understanding Phenomena

verb

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

Evidence paragraph

*Polar bears, saved from the bitter cold **by** their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.*

Language Understanding Phenomena

verb

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

preposition

Evidence paragraph

*Polar bears, saved from the bitter cold **by** their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.*

Language Understanding Phenomena

verb

Question: *A bear <u>survives</u> winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

comma

preposition

Evidence paragraph

Polar bears, <u>saved from</u> the bitter cold ***by*** *their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.*

Language Understanding Phenomena

verb

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

comma

preposition

Evidence paragraph

Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

QA is fundamentally a NLU problem

"Lifting" Meaning as Semantic Graphs

“Lifting” Meaning as Semantic Graphs

Evidence paragraph

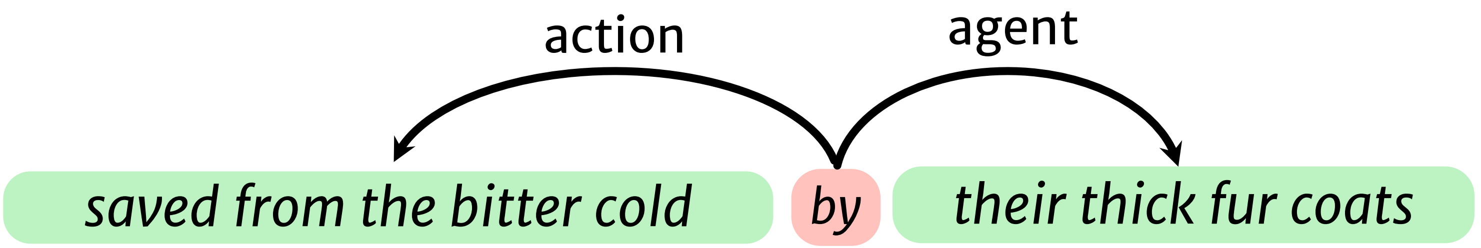
... Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

“Lifting” Meaning as Semantic Graphs

Evidence paragraph

... Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

"Lifting" Meaning as Semantic Graphs



Evidence paragraph

… Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

"Lifting" Meaning as Semantic Graphs

action

agent

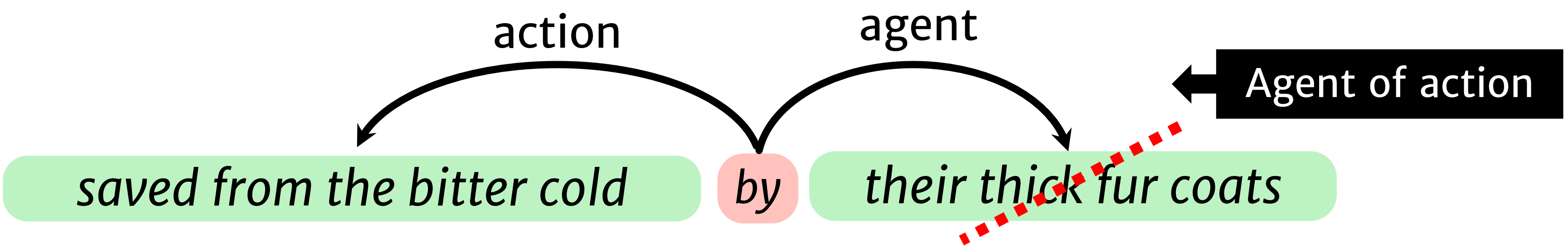
Agent of action

saved from the bitter cold *by* *their thick fur coats*

Evidence paragraph

… Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

"Lifting" Meaning as Semantic Graphs

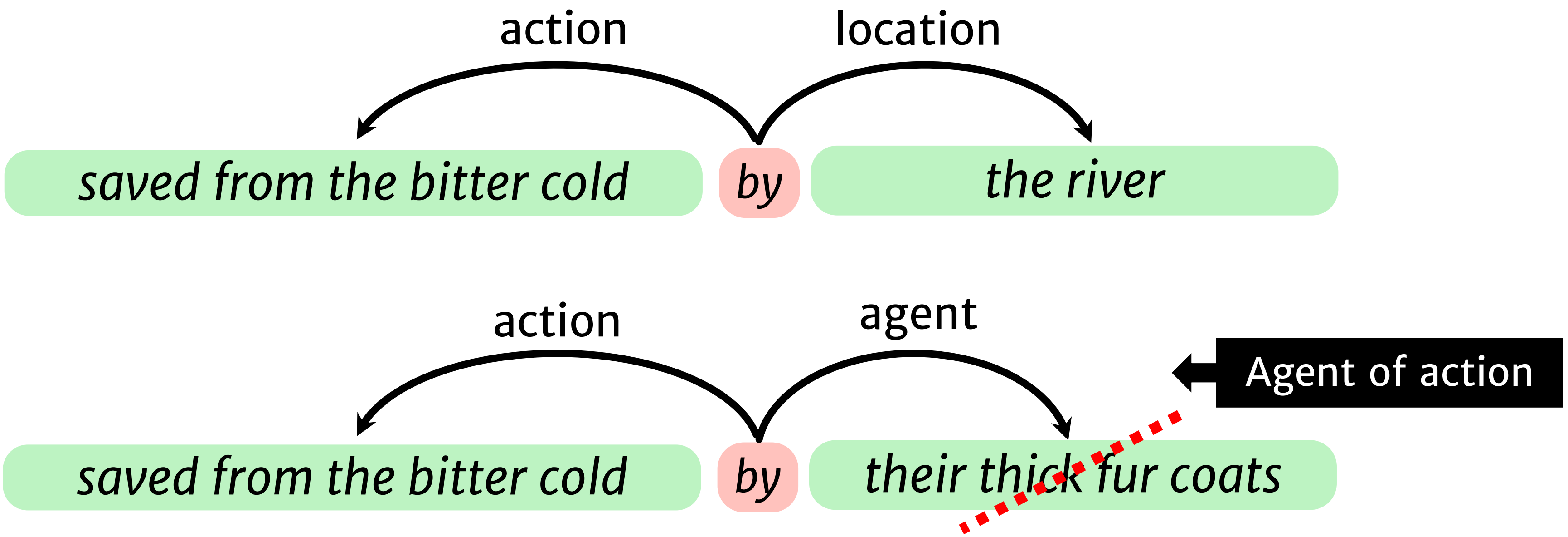


Evidence paragraph

... Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

Preposition Semantic Roles [Srikumar & Roth, 2013]

"Lifting" Meaning as Semantic Graphs

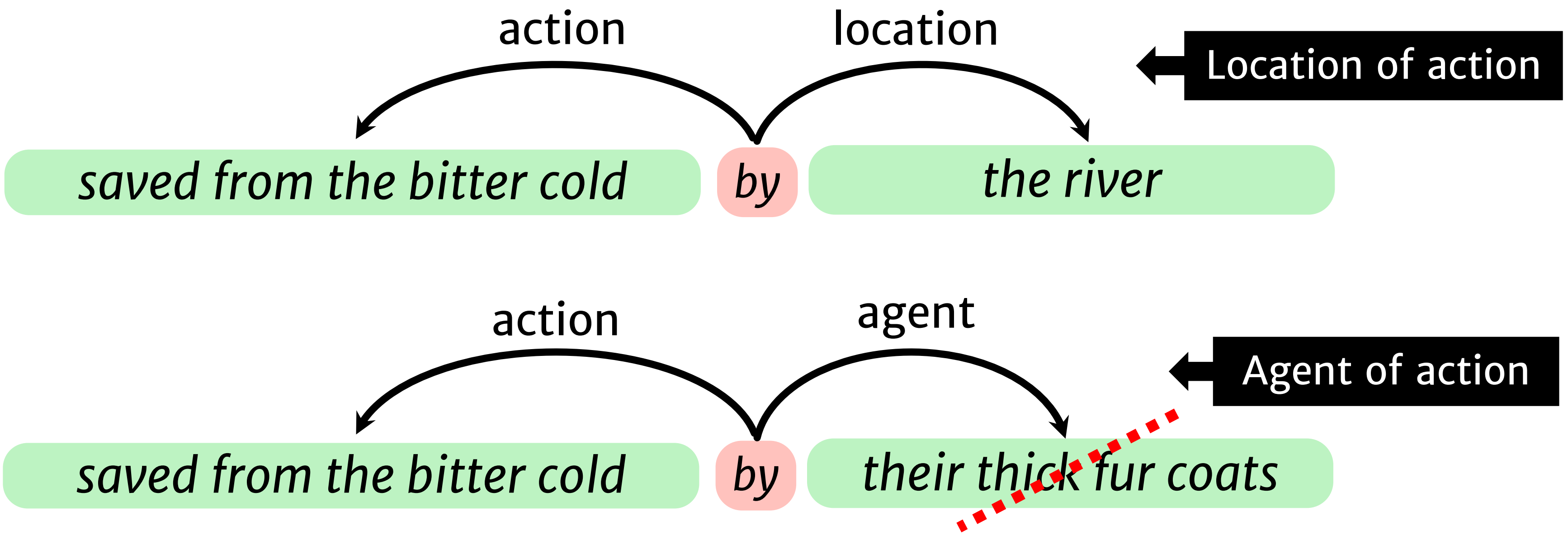


Evidence paragraph

… Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

Preposition Semantic Roles [Srikumar & Roth, 2013]

"Lifting" Meaning as Semantic Graphs



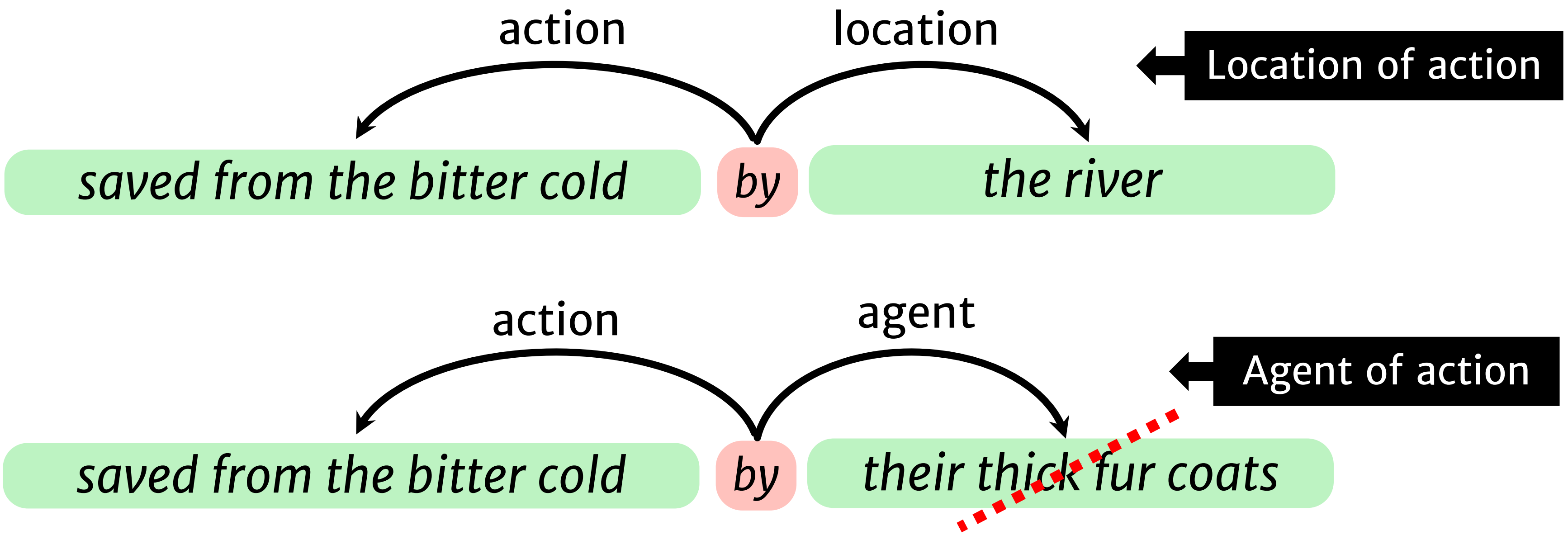
Evidence paragraph

... Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

Preposition Semantic Roles [Srikumar & Roth, 2013]

"Lifting" Meaning as Semantic Graphs

Oxford English Dictionary lists **<u>8</u>** primary meanings for "by".



Evidence paragraph

… Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

"Lifting" Meaning as Semantic Graphs

Oxford English Dictionary lists <u>**8**</u> primary meanings for "by".

Disambiguation!

action — location

saved from the bitter cold | *by* | *the river*

Location of action

action — agent

saved from the bitter cold | *by* | *their thick fur coats*

Agent of action

Evidence paragraph

... Polar bears, saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

Preposition Semantic Roles [Srikumar & Roth, 2013]

“Lifting” Meaning as Semantic Graphs

“Lifting” Meaning as Semantic Graphs

Evidence paragraph

... Polar bears , saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

"Lifting" Meaning as Semantic Graphs

Evidence paragraph

... Polar bears , saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

“Lifting” Meaning as Semantic Graphs

mention description

Polar bears , *saved from the bitter cold*

Evidence paragraph

... Polar bears , saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

"Lifting" Meaning as Semantic Graphs

mention   description

Polar bears , *saved from the bitter cold*

apposition structure

Evidence paragraph

... Polar bears , saved from the bitter cold by their thick fur coats, are among the animals in danger of extinction because of global warming and human activities.

“Lifting” Meaning as Semantic Graphs

“Lifting” Meaning as Semantic Graphs

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

"Lifting" Meaning as Semantic Graphs

Question: *A bear survives winters with what structure?*

+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

“Lifting” Meaning as Semantic Graphs

Question: *A bear survives winters with what structure?*

\+

(A) big ears
(B) black nose
(C) thick fur
(D) brown eyes

a bear — survivor

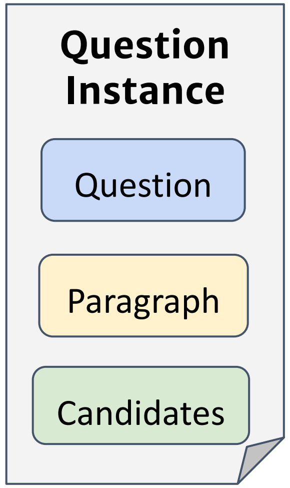
survive

a winter — adverse circumstance

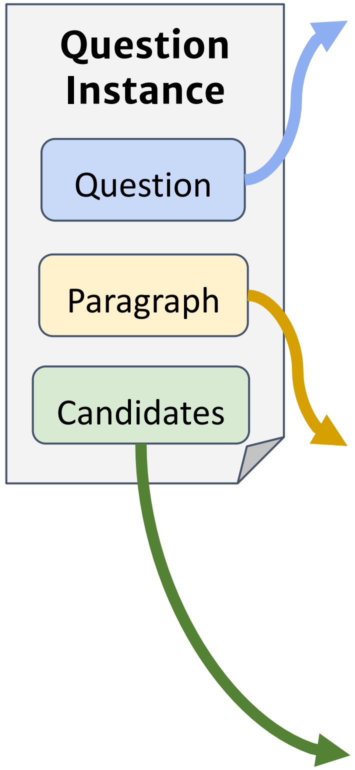
what structure

live despite circumstance

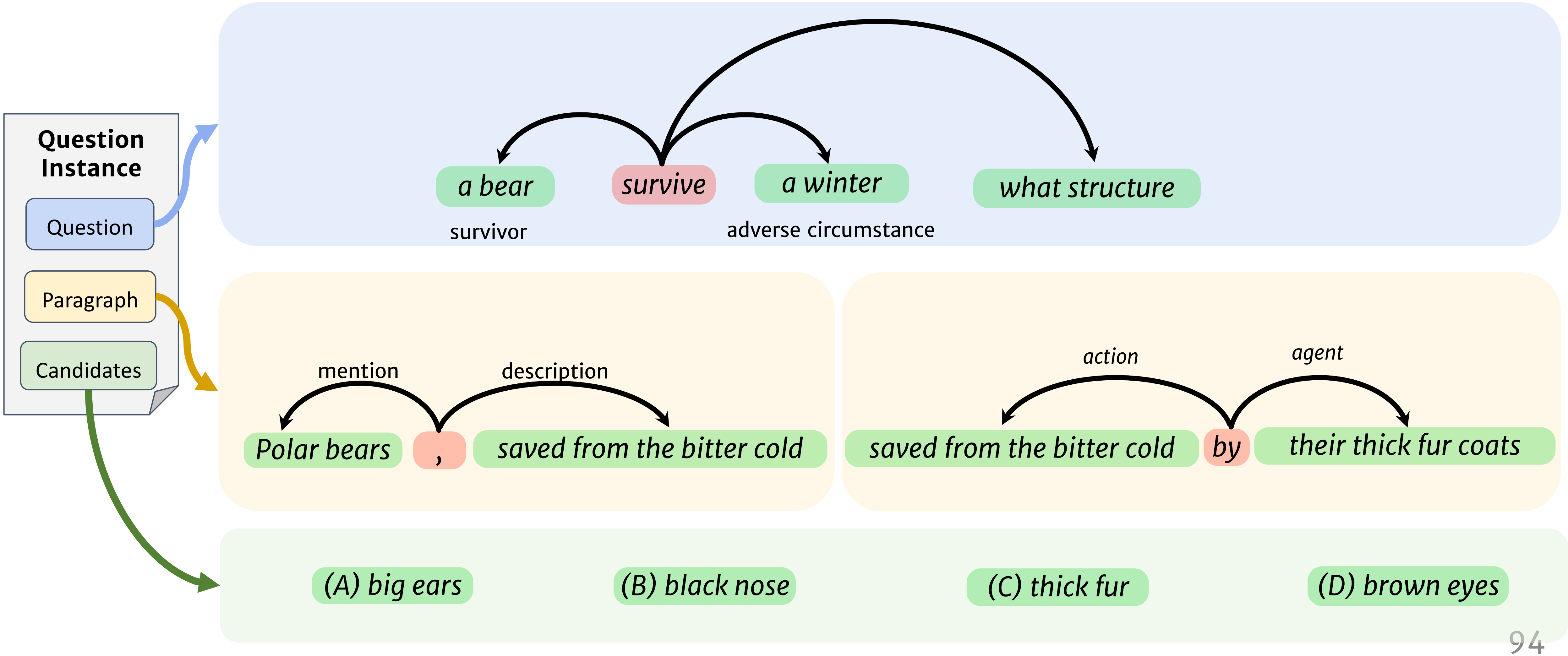
Semantic Representations Altogether



Semantic Representations Altogether

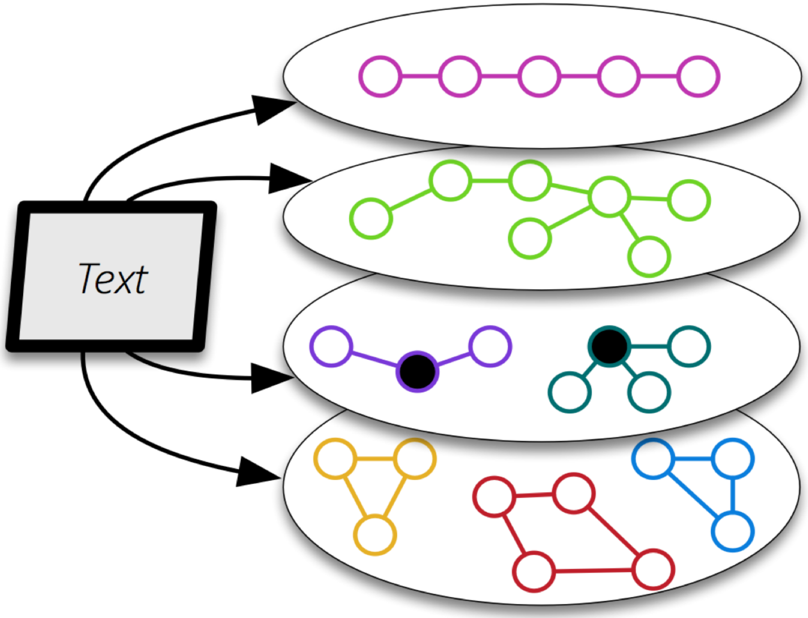


Semantic Representations Altogether



Collections of Semantic Graphs

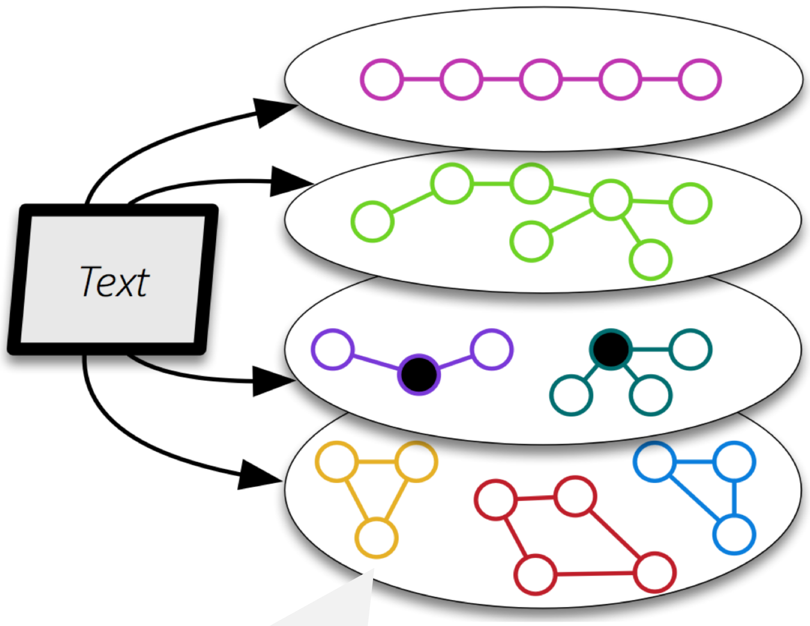
- Create a unified representation of families of graphs



- Verb Semantic Roles [Punyakanok et al. 2008]
- Preposition Semantic Roles [Srikumar & Roth 2013]
- Comma Semantic Roles [Arivazhagan et al. 2016]
- Coreference [Chang et al. 2012]
- …

Collections of Semantic Graphs

- Create a unified representation of families of graphs



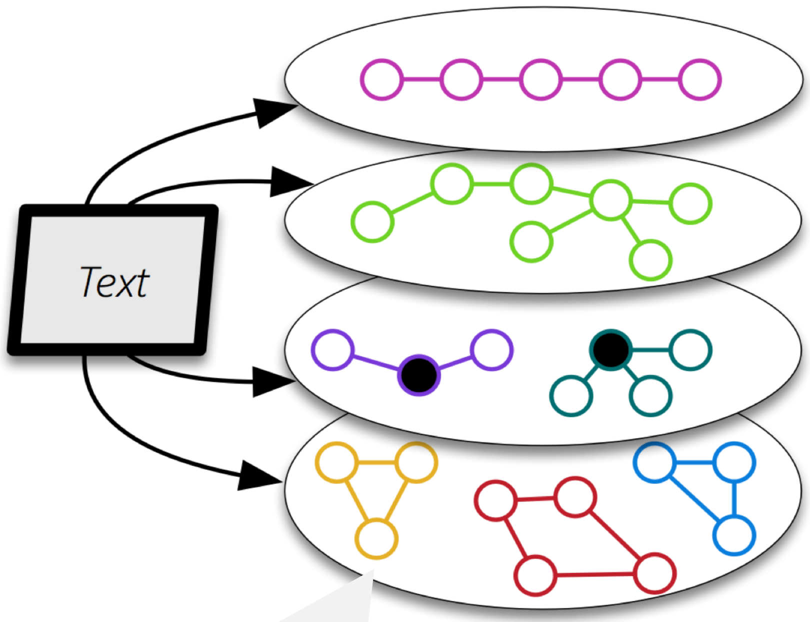
- Surface word
- Semantic labels
- Other representation
- ...

- Verb Semantic Roles [Punyakanok et al. 2008]
- Preposition Semantic Roles [Srikumar & Roth 2013]
- Comma Semantic Roles [Arivazhagan et al. 2016]
- Coreference [Chang et al. 2012]
- ...

Collections of Semantic Graphs

available in our software pipeline.
K et al. LREC'18

- Create a unified representation of families of graphs



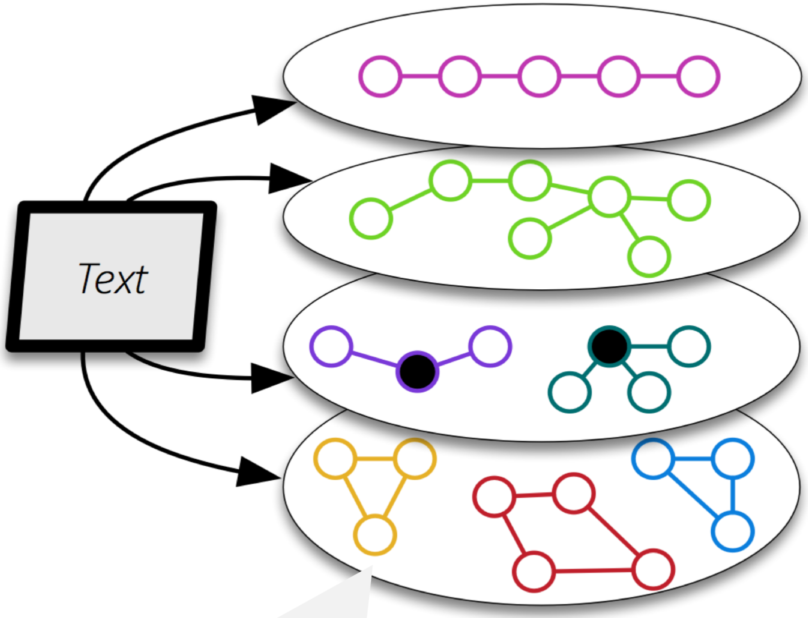
- Surface word
- Semantic labels
- Other representation
- ...

- Verb Semantic Roles [Punyakanok et al. 2008]
- Preposition Semantic Roles [Srikumar & Roth 2013]
- Comma Semantic Roles [Arivazhagan et al. 2016]
- Coreference [Chang et al. 2012]
- ...

Collections of Semantic Graphs

- Create a unified representation of families of graphs

available in our software pipeline.
K et al. LREC'18



- Surface word
- Semantic labels
- Other representation
- ...

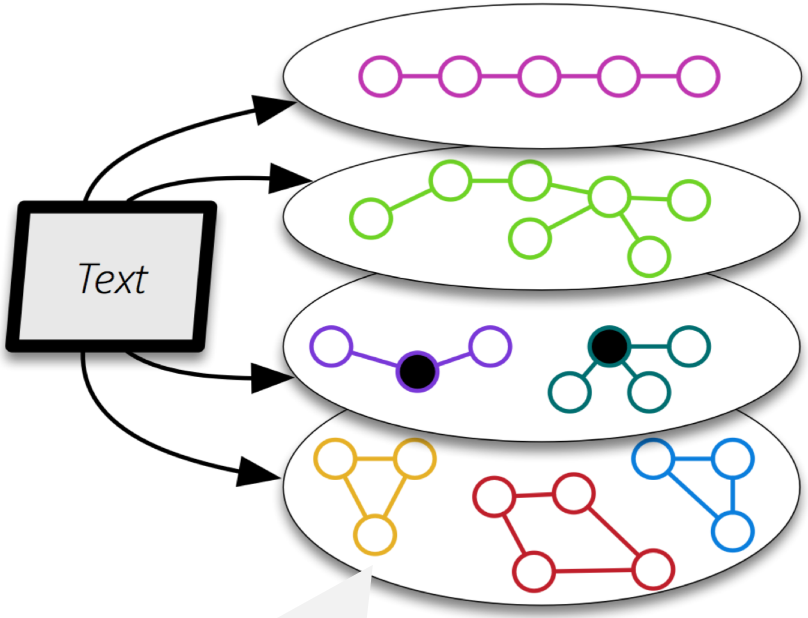
- Verb Semantic Roles [Punyakanok et al. 2008]
- Preposition Semantic Roles [Srikumar & Roth 2013]
- Comma Semantic Roles [Arivazhagan et al. 2016]
- Coreference [Chang et al. 2012]
- ...

Our representation is **not** QA-specific.
It reflects our understanding of the language

Collections of Semantic Graphs

- Create a unified representation of families of graphs

available in our software pipeline.
K et al. LREC'18



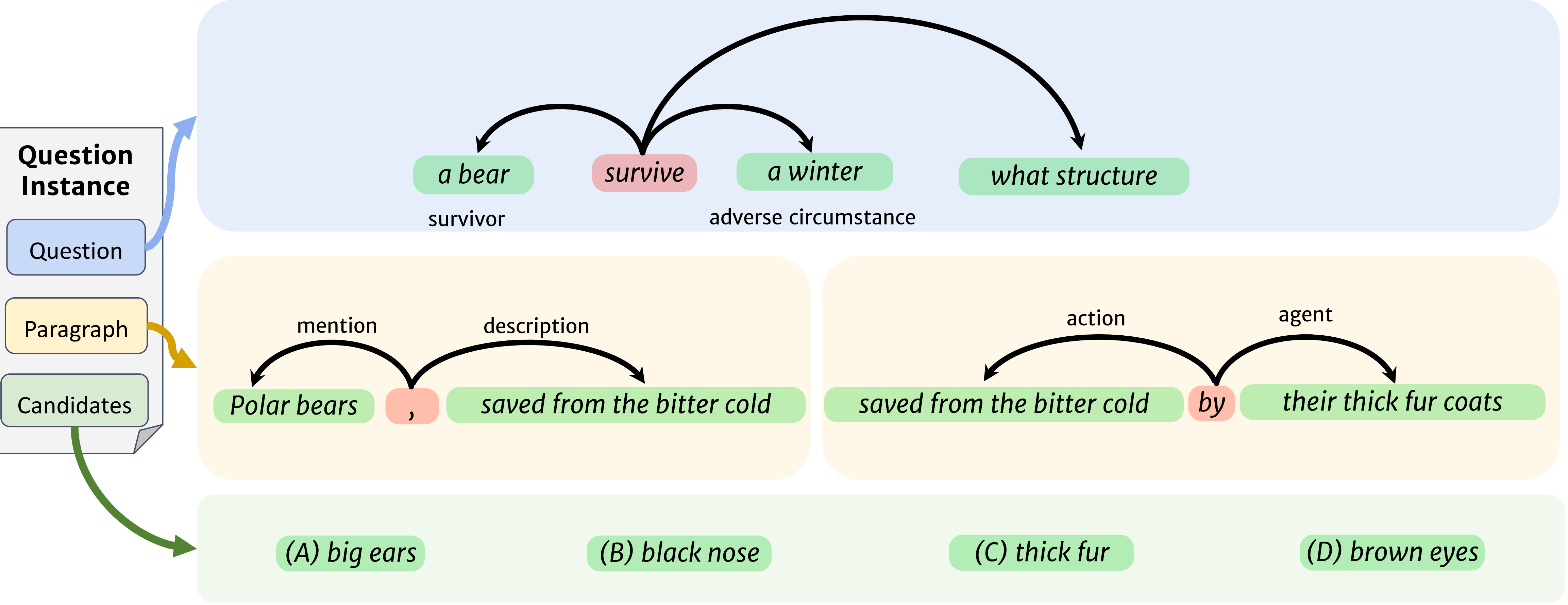
- Surface word
- Semantic labels
- Other representation
- ...

- Verb Semantic Roles [Punyakanok et al. 2008]
- Preposition Semantic Roles [Srikumar & Roth 2013]
- Comma Semantic Roles [Arivazhagan et al. 2016]
- Coreference [Chang et al. 2012]
- ...

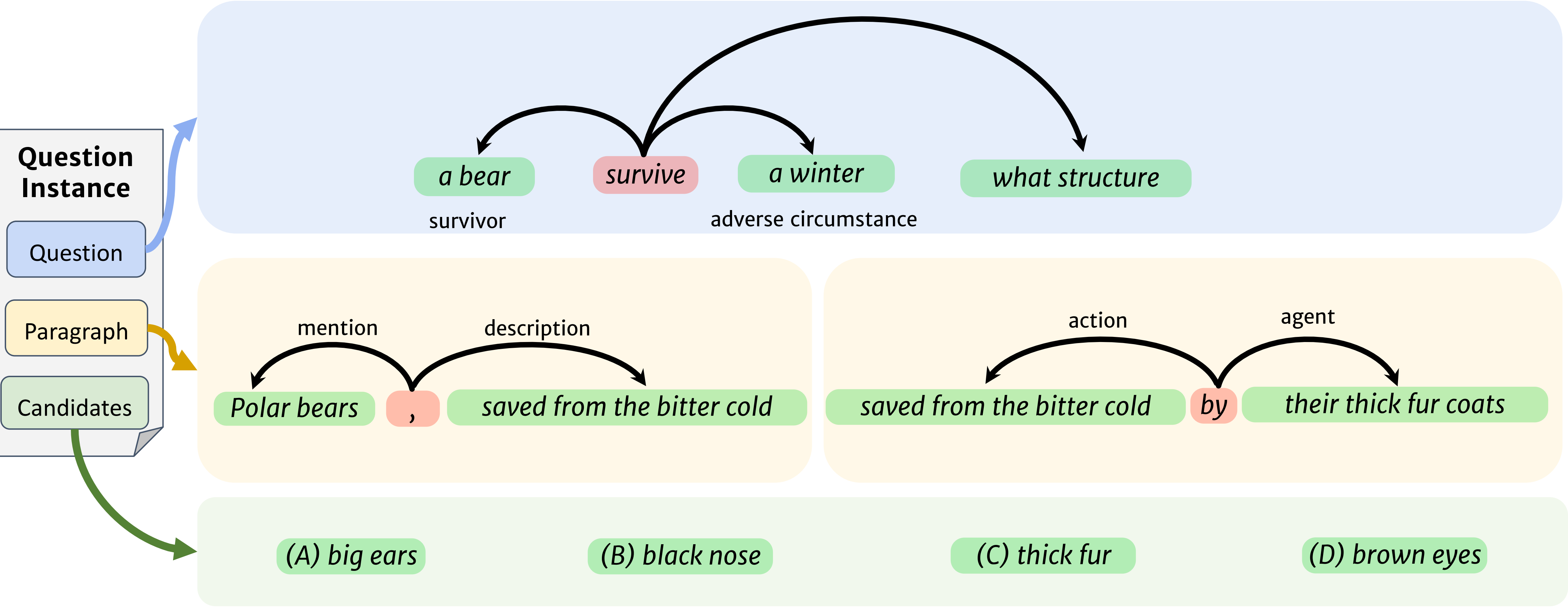
Our representation is **not** QA-specific.
It reflects our understanding of the language

Consequently, we expect these representations to be useful for a range of tasks

Support Graph

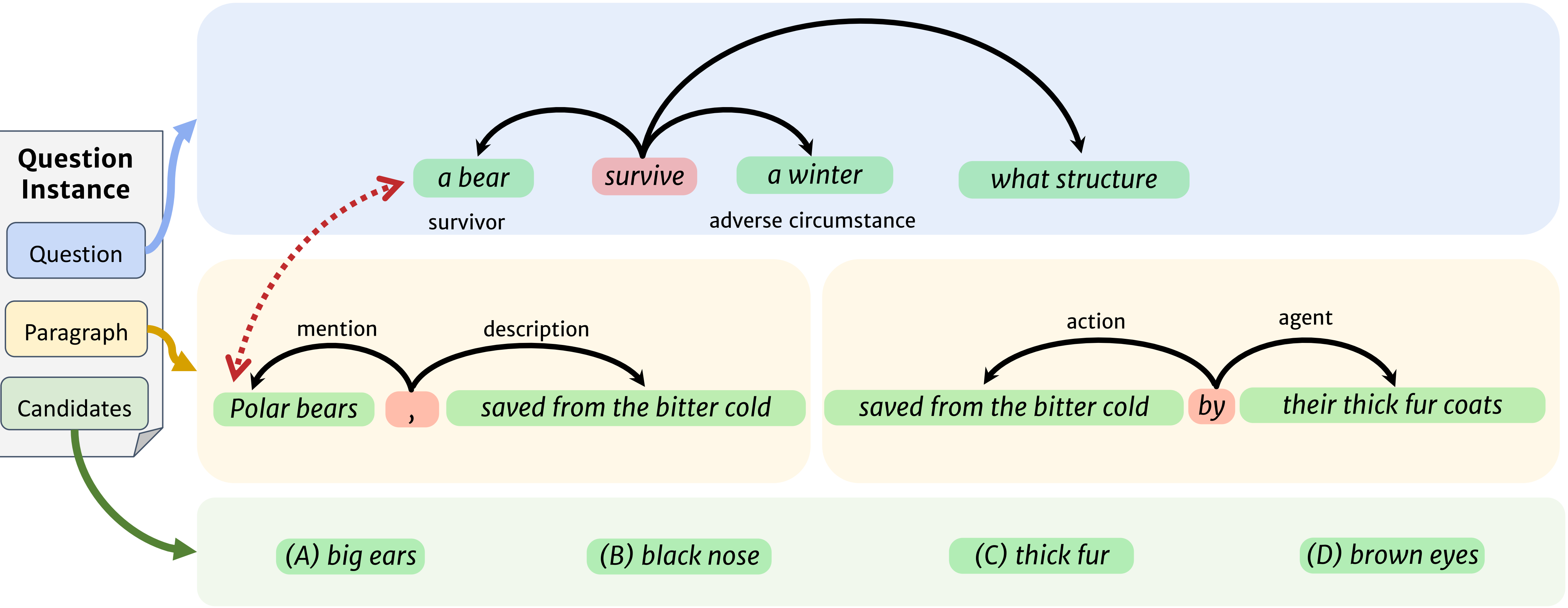


Support Graph



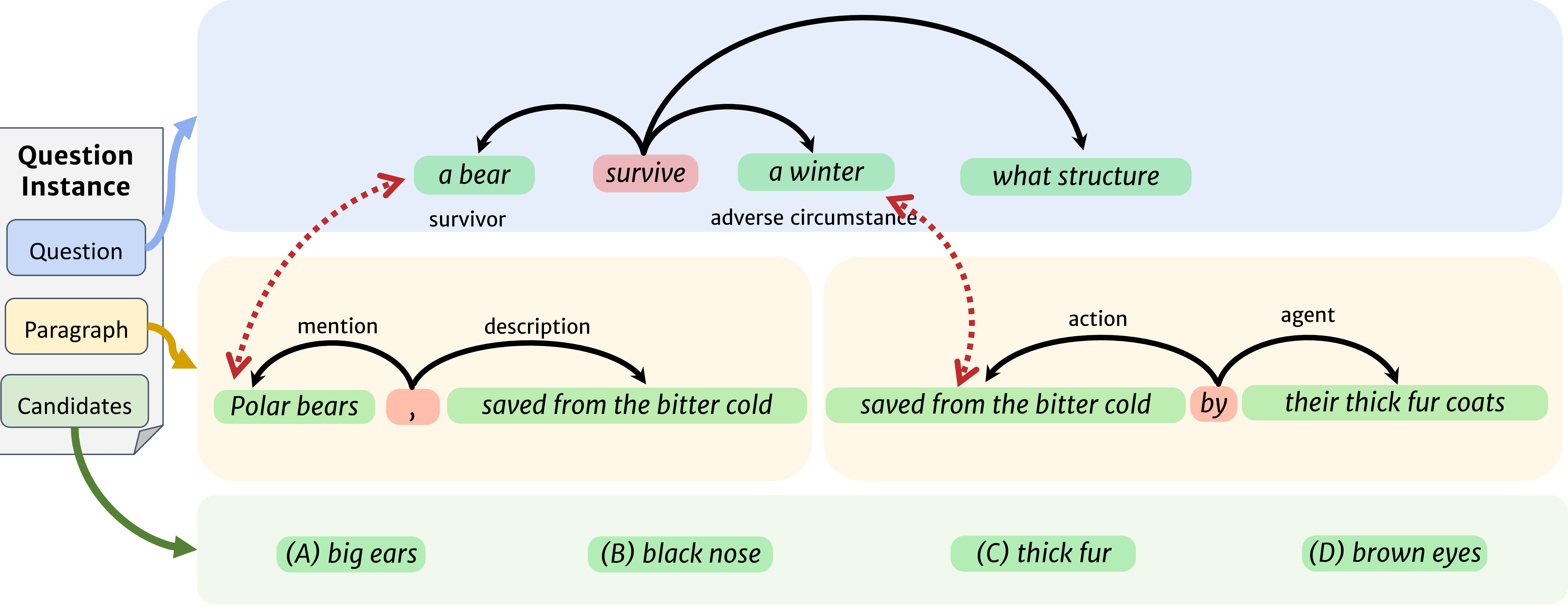
Search for the best **Support Graph** connecting the Question to an Answer through the knowledge graph.

Support Graph



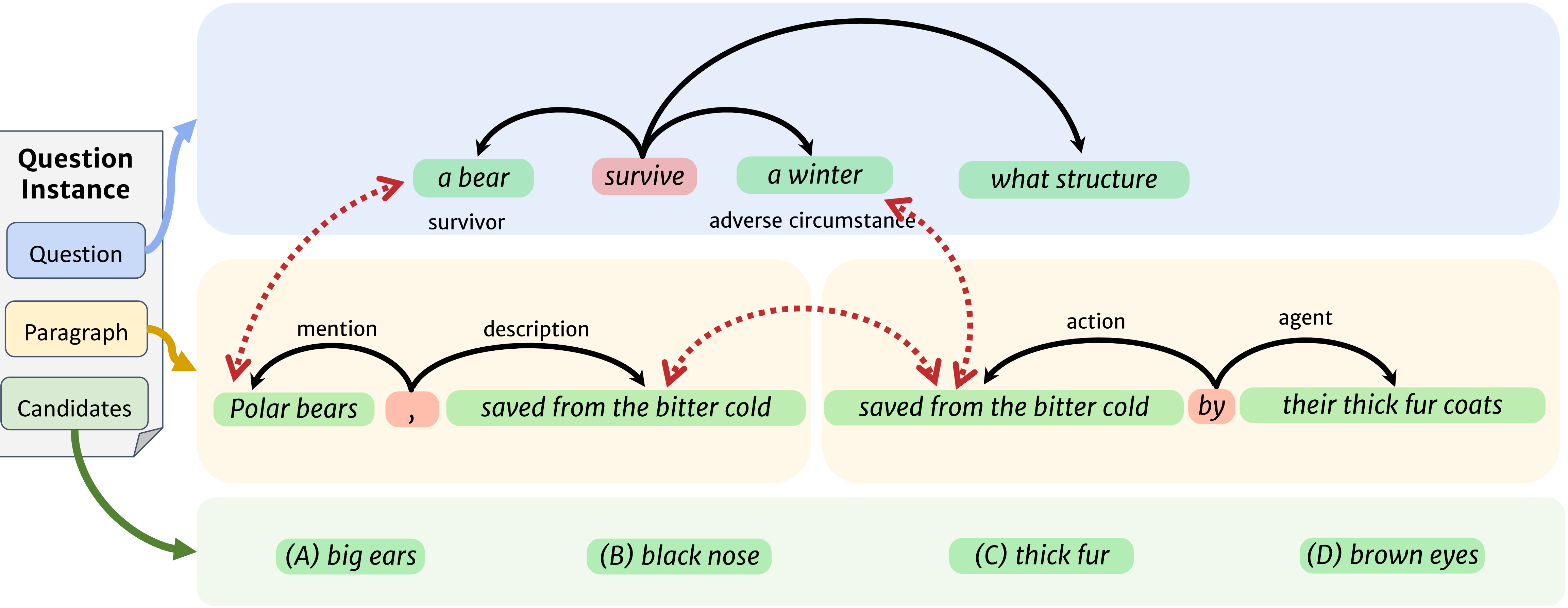
Search for the best **Support Graph** connecting the Question to an Answer through the knowledge graph.

Support Graph



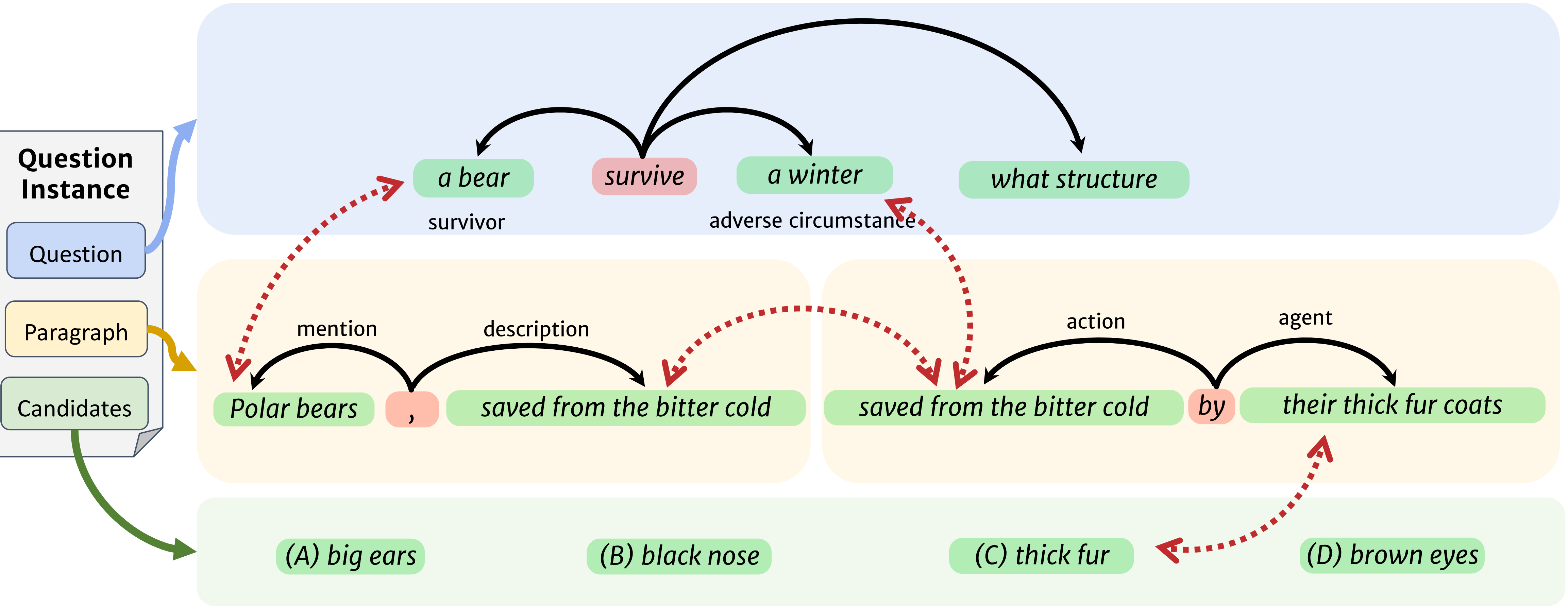
Search for the best **Support Graph** connecting the Question to an Answer through the knowledge graph.

Support Graph



Search for the best **Support Graph** connecting the Question to an Answer through the knowledge graph.

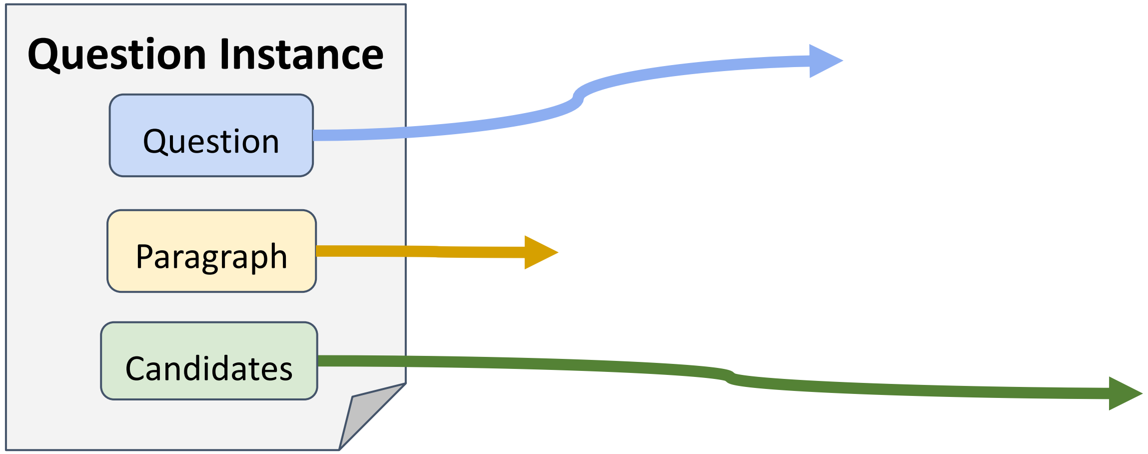
Support Graph



Search for the best **Support Graph** connecting the Question to an Answer through the knowledge graph.

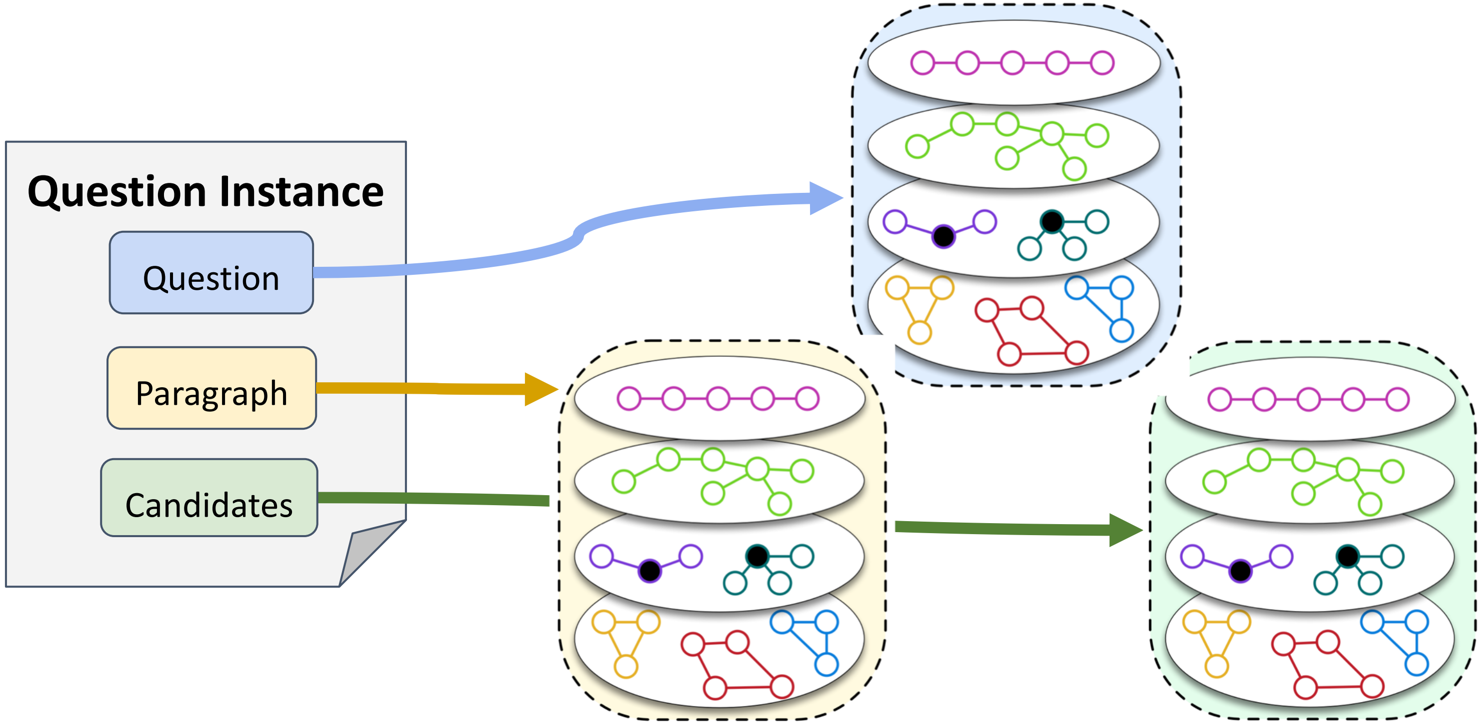
Reasoning With a Meaning Representation

- **Support Graph** creates potential alignments between various semantic abstractions.



Reasoning With a Meaning Representation

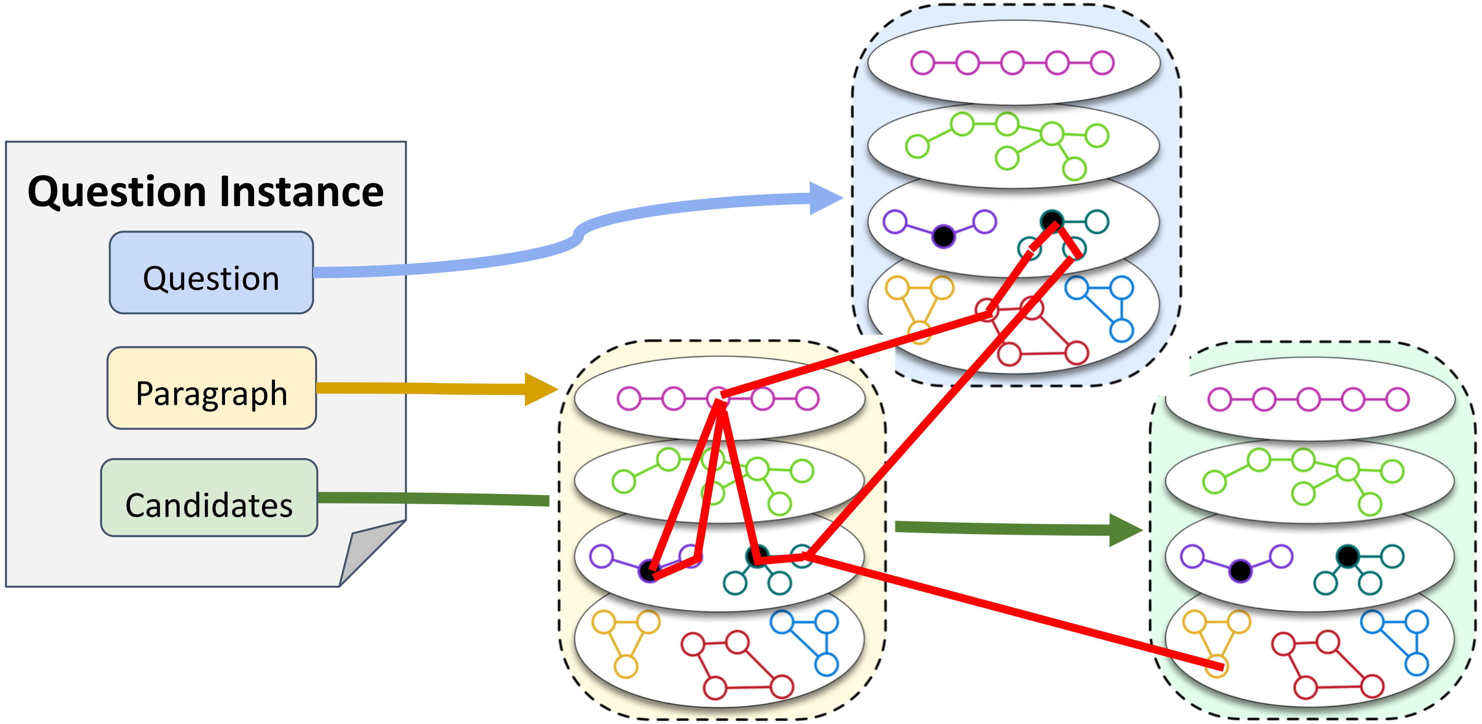
- **Support Graph** creates potential alignments between various semantic abstractions.



Edges reflect similarity / entailment

Reasoning With a Meaning Representation

- **Support Graph** creates potential alignments between various semantic abstractions.

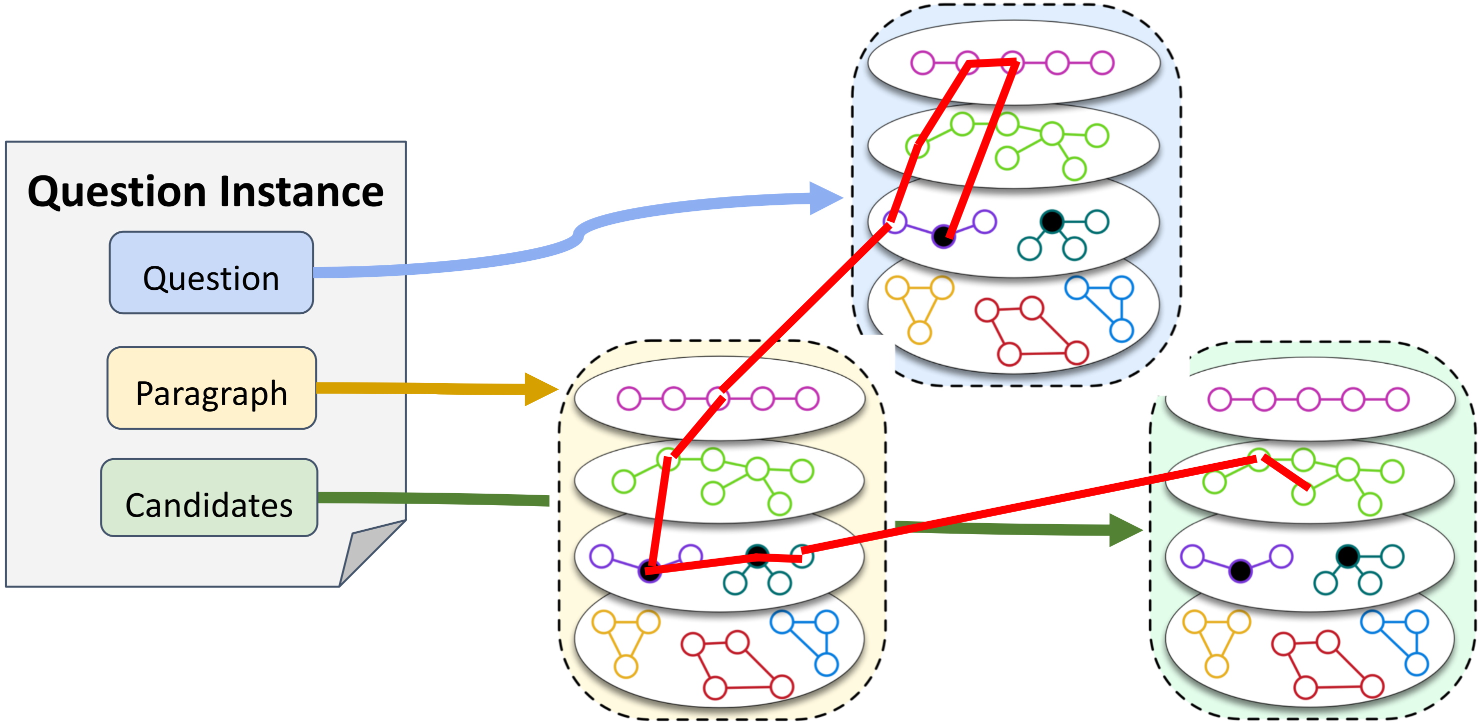


Edges reflect similarity / entailment

QA Reasoning formulated as finding "best" explanation – subgraph connecting the Question to the Answers via the Knowledge

Reasoning With a Meaning Representation

- **Support Graph** creates potential alignments between various semantic abstractions.

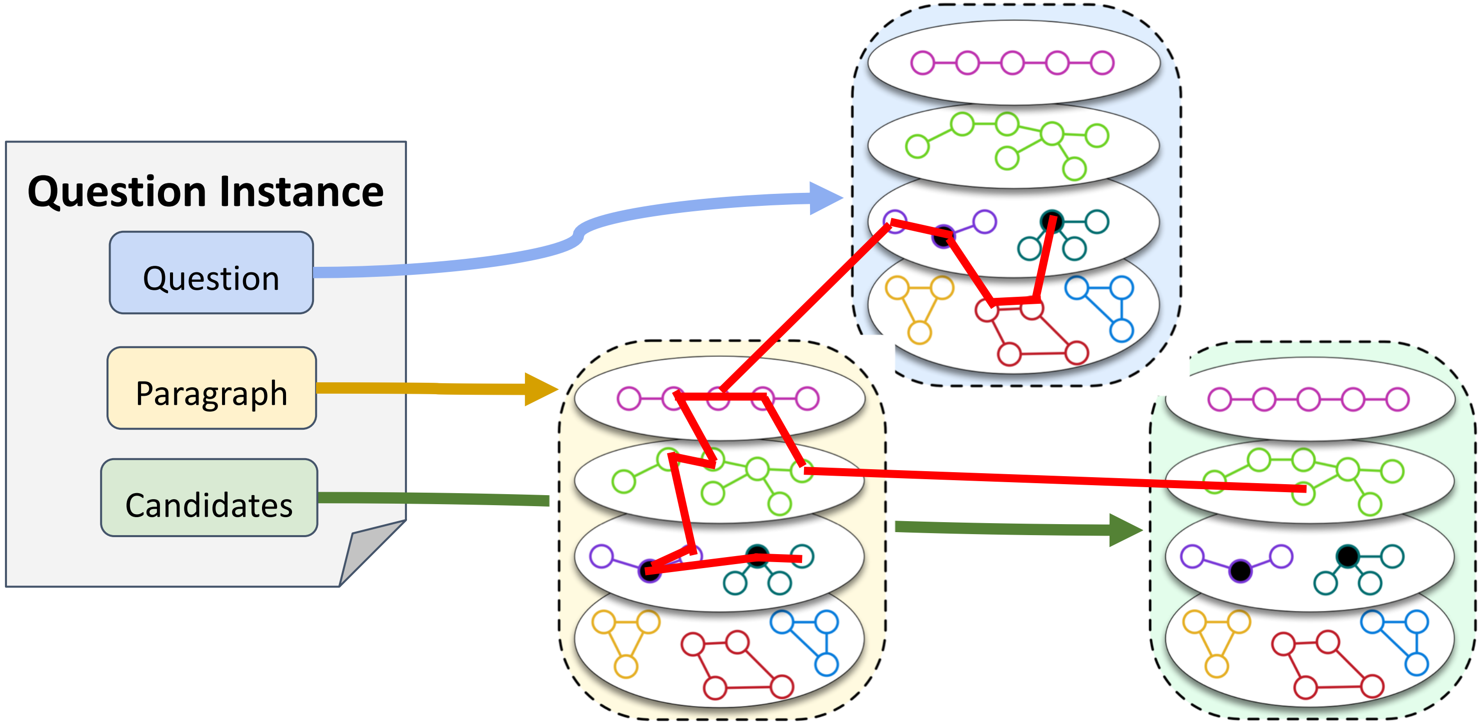


Edges reflect similarity / entailment

QA Reasoning formulated as finding "best" explanation – subgraph connecting the Question to the Answers via the Knowledge

Reasoning With a Meaning Representation

- **Support Graph** creates potential alignments between various semantic abstractions.

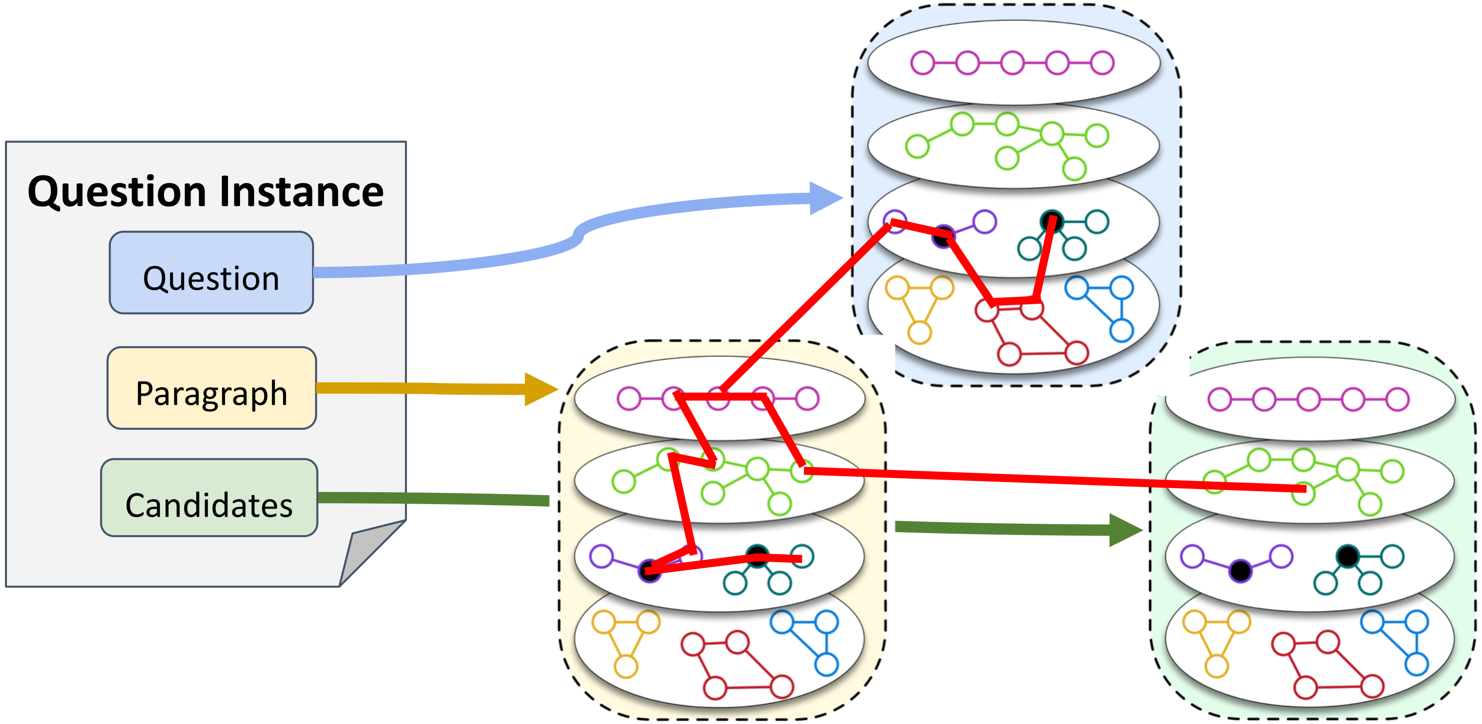


Edges reflect similarity / entailment

QA Reasoning formulated as finding "best" explanation – subgraph connecting the Question to the Answers via the Knowledge

Reasoning With a Meaning Representation

- **Support Graph** creates potential alignments between various semantic abstractions.



Edges reflect similarity / entailment

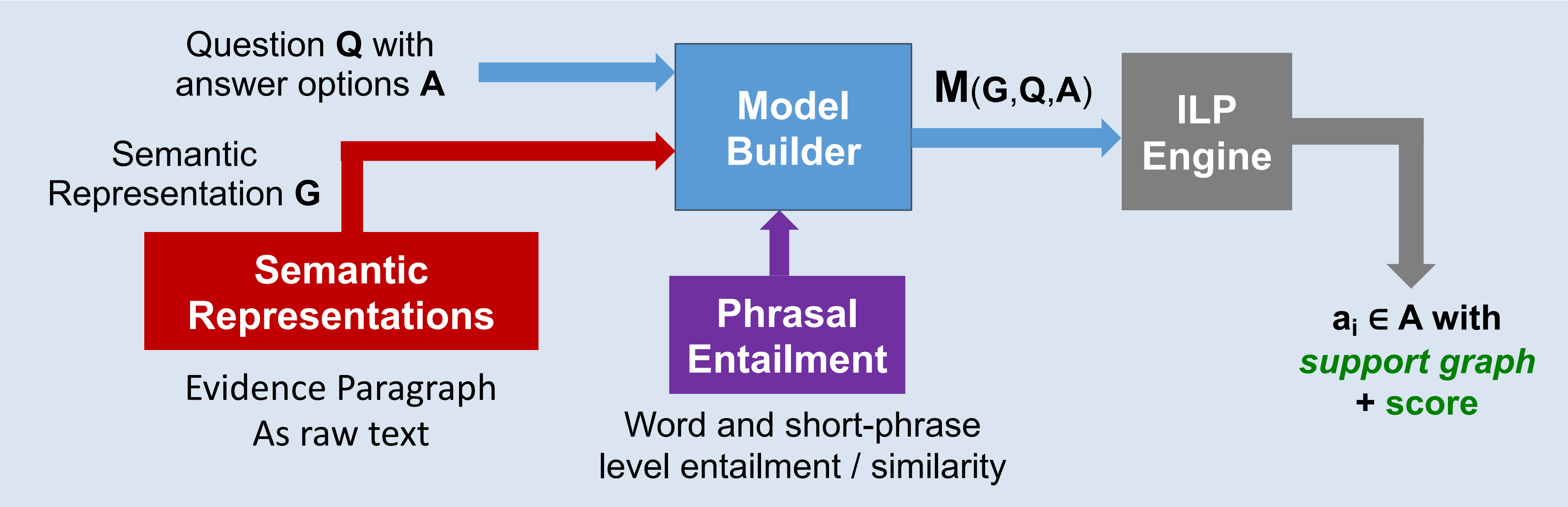
Abductive reasoning

[Peirce, 1883]

QA Reasoning formulated as finding "best" explanation – subgraph connecting the Question to the Answers via the Knowledge

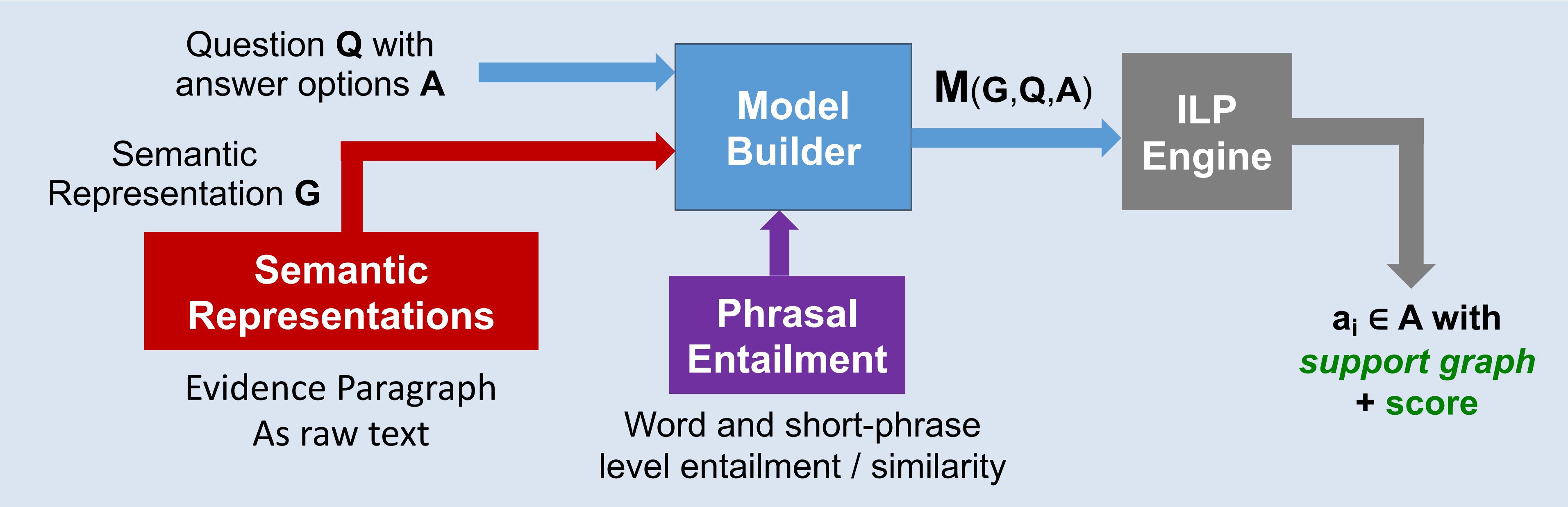
Framework Overview

- A discrete **optimization** approach to QA for multiple-choice questions



Framework Overview

- A discrete **optimization** approach to QA for multiple-choice questions



M(**G**,**Q**,**A**) ⟶

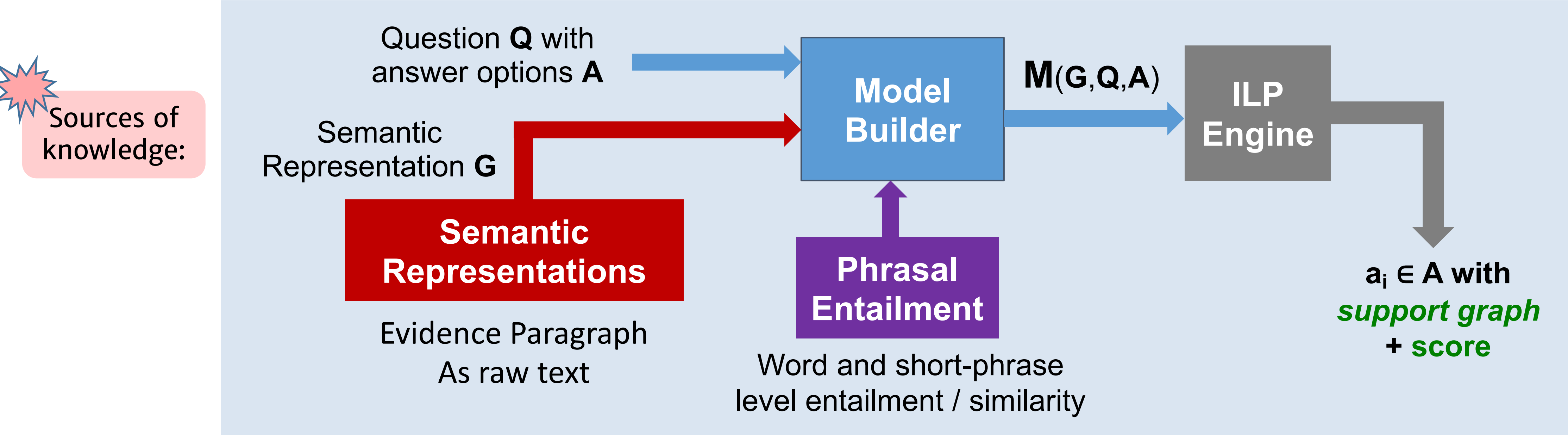
$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

Optimization using Integer Linear Program (**ILP**) formalism

Framework Overview

- A discrete **optimization** approach to QA for multiple-choice questions



M(**G**,**Q**,**A**) ⟶

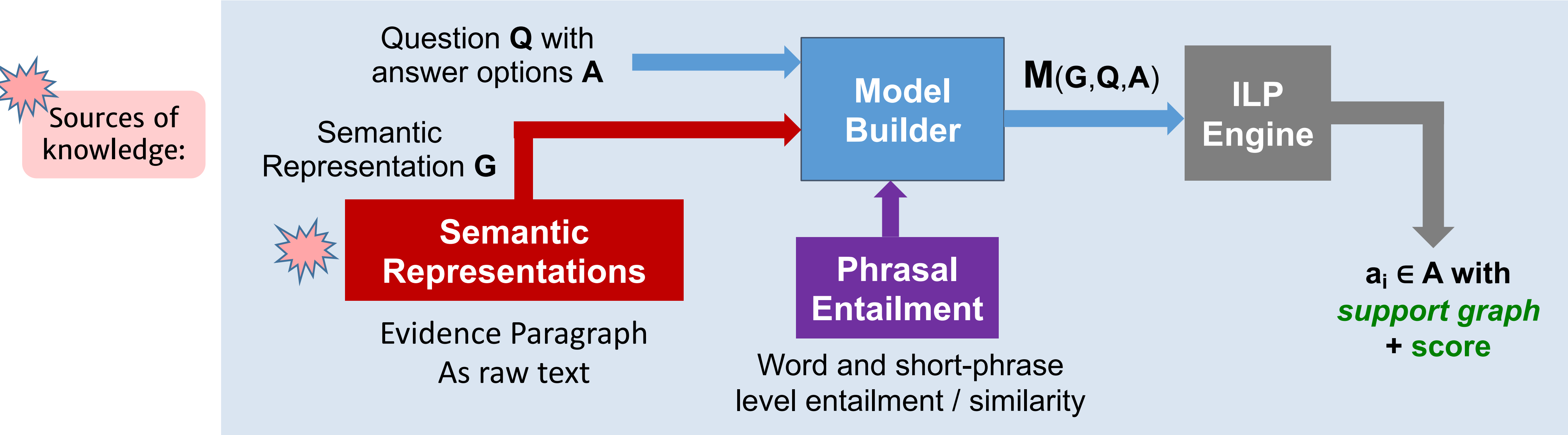
$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \dots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

Optimization using Integer Linear Program (**ILP**) formalism

Framework Overview

- A discrete **optimization** approach to QA for multiple-choice questions



M(**G**,**Q**,**A**) →

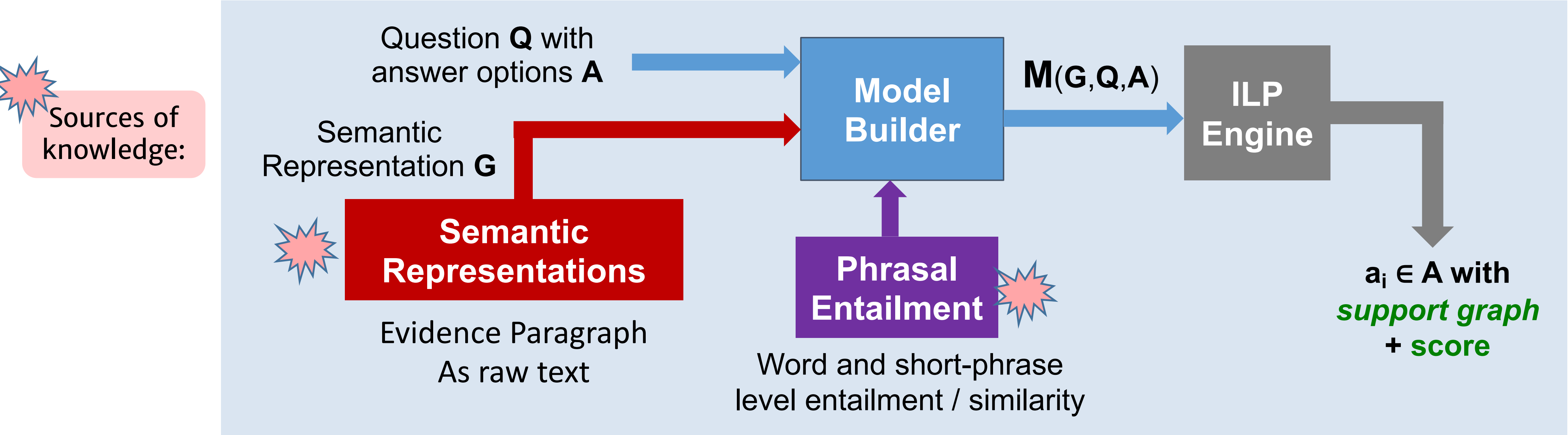
$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

Optimization using Integer Linear Program (**ILP**) formalism

Framework Overview

- A discrete **optimization** approach to QA for multiple-choice questions



M(**G**,**Q**,**A**) $\rightarrow$

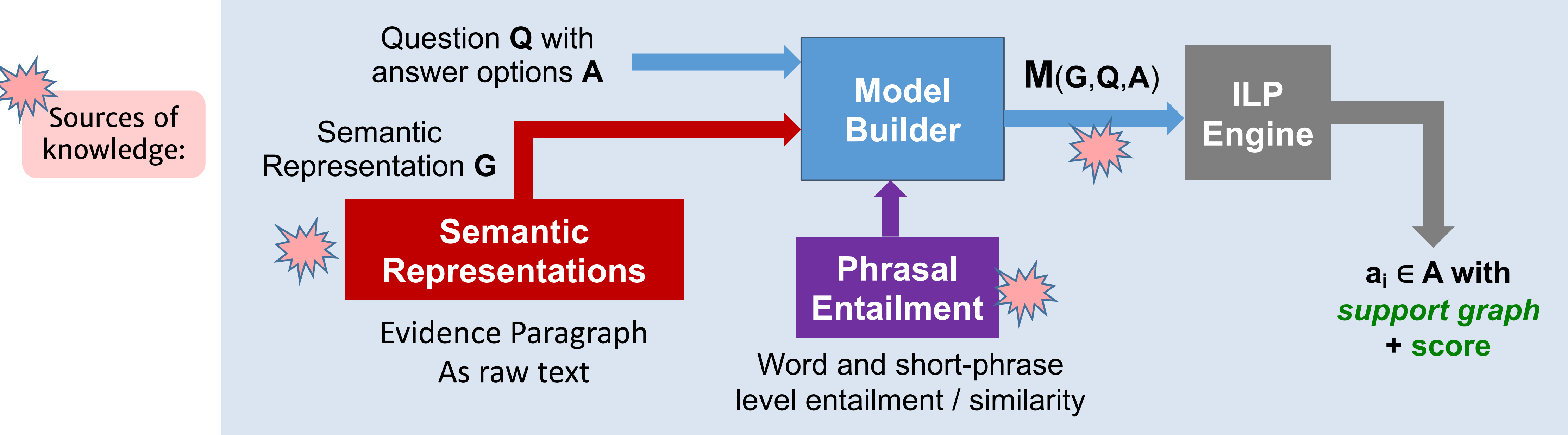
$$\max \sum_i c_i x_i \qquad \forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$

Optimization using Integer Linear Program (**ILP**) formalism

Framework Overview

- A discrete **optimization** approach to QA for multiple-choice questions



M(**G**,**Q**,**A**) ➡

$$\max \sum_i c_i x_i \qquad \forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \quad \ldots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

Optimization using Integer Linear Program (**ILP**) formalism

ILP Model: Design Challenges

Goal: *Design ILP objective function, s.t. maximizing it subject to the constraints yields a "desirable" support graph*

Not so straightforward!

$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \quad \ldots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

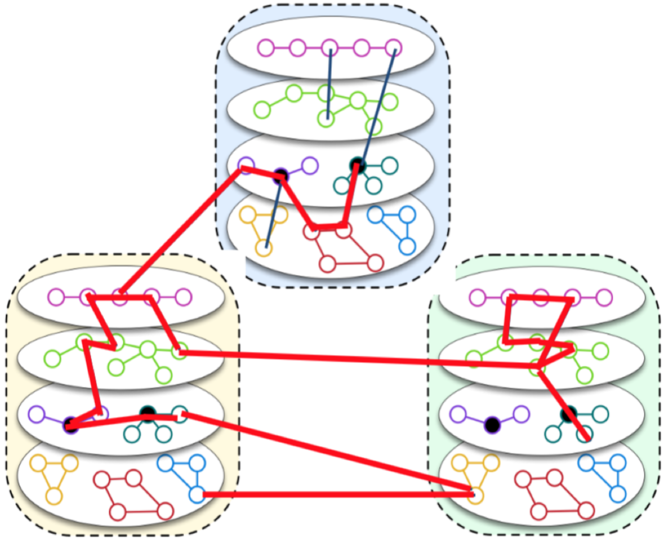
$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

- Many possible "proof structures"
- Imperfect lexical "similarity" blackbox
- Partial or missing knowledge
- Question logic (negation, conjunction, comparison)
- Scalability of ILP solvers
- …

ILP Model: Some Details

$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$

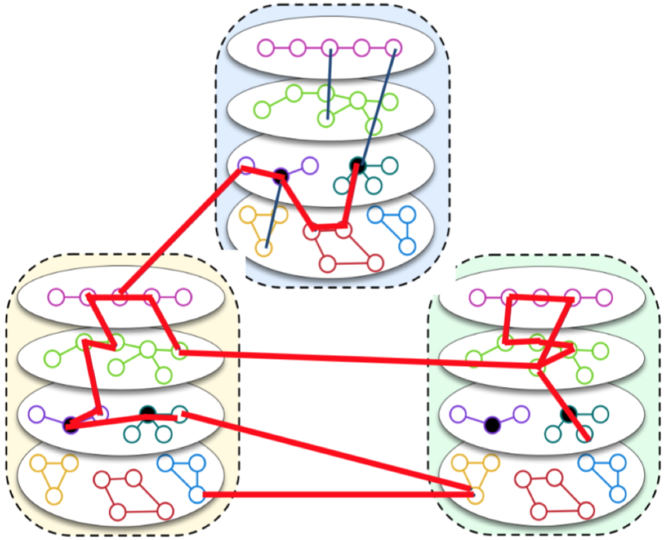
$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$



ILP Model: Some Details

$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$

$$\boxed{\forall\, x_i \in \mathbb{N} \cup \{0\}}$$

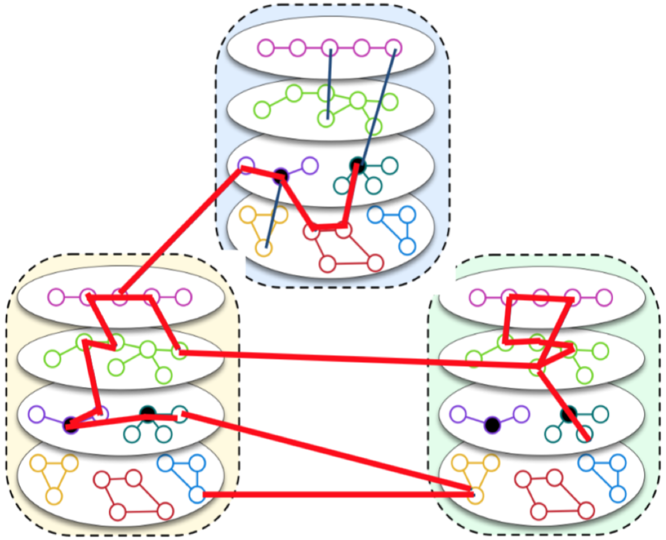


ILP Model: Some Details

$$\boxed{\max \sum_i c_i x_i}$$

$$\boxed{\forall\, x_i \in \mathbb{N} \cup \{0\}}$$

$$\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$

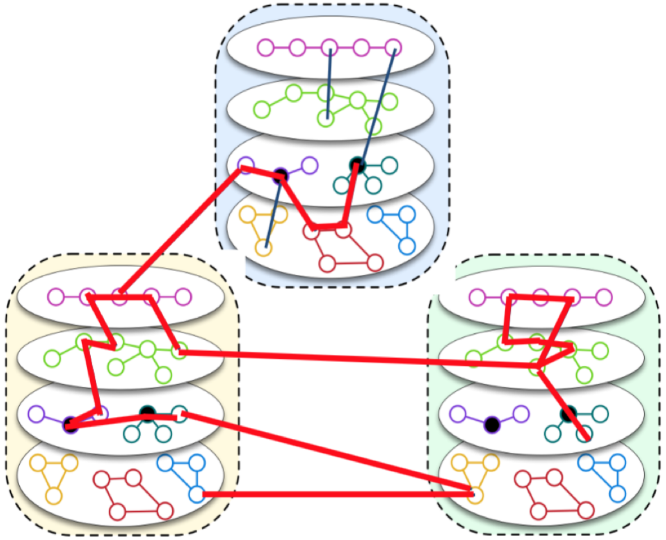


ILP Model: Some Details

$$\max \sum_i c_i x_i$$

$$\forall\ x_i \in \mathbb{N} \cup \{0\}$$

$$\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$



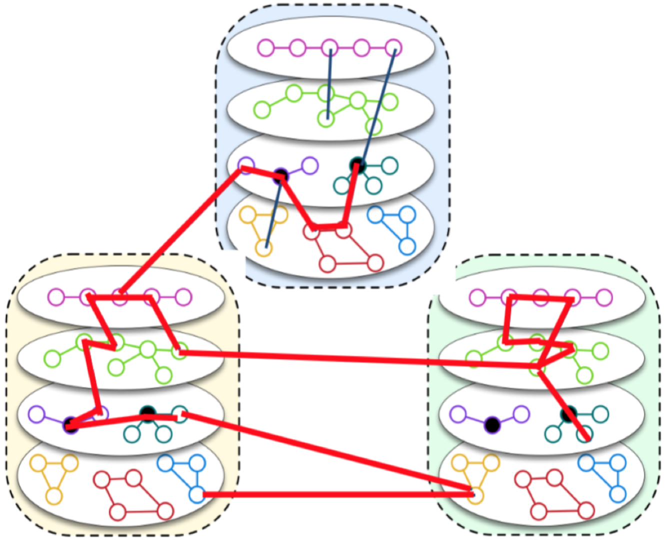
ILP Model: Some Details

Variables define the space of “support graphs”:

- Each variable corresponds to to a node or edge.
- *x=1* iff nodes / edges are part of the semantic graph.

$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \dots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$



ILP Model: Some Details

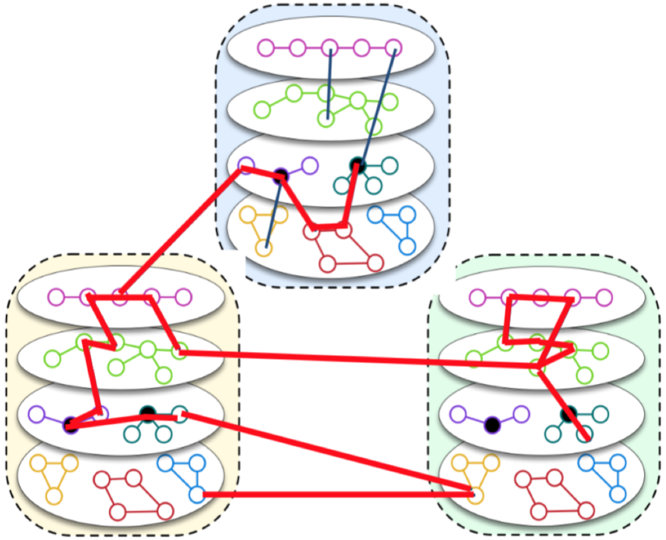
Variables define the space of "support graphs":

- Each variable corresponds to to a node or edge.
- $x=1$ iff nodes / edges are part of the semantic graph.

$$\max \sum_i c_i x_i$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$



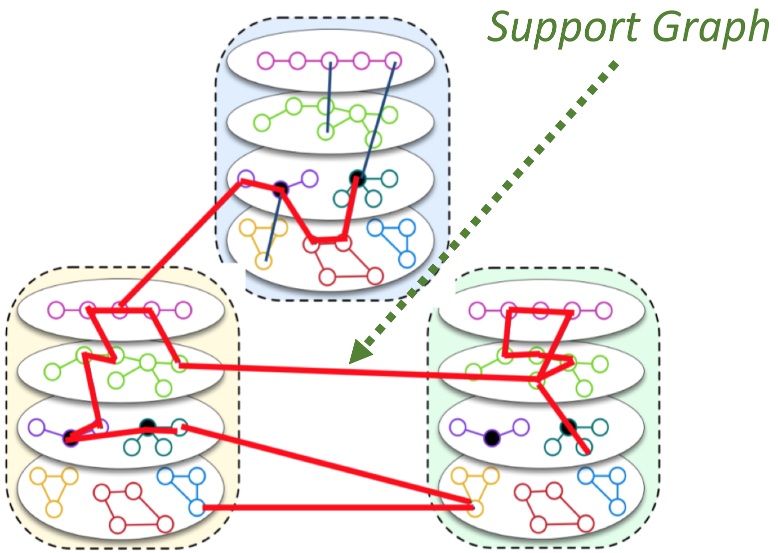
ILP Model: Some Details

Variables define the space of "support graphs":

- Each variable corresponds to to a node or edge.
- $x=1$ iff nodes / edges are part of the semantic graph.

$$\max \sum_i c_i x_i \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

$$\boxed{\forall\, x_i \in \mathbb{N} \cup \{0\}}$$

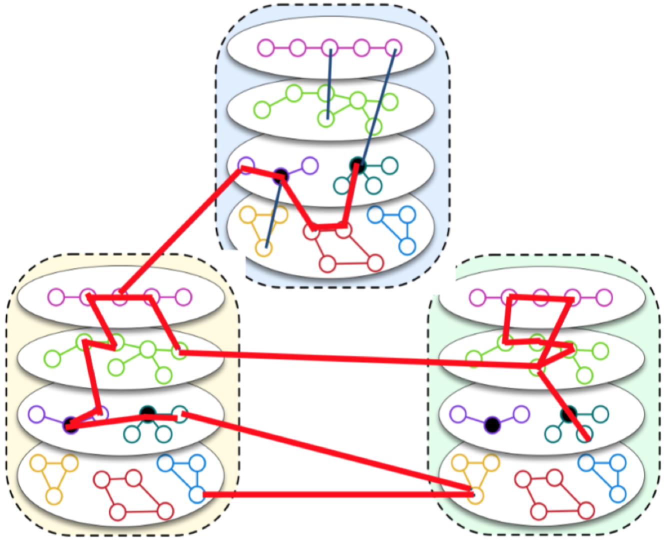


ILP Model: Some Details

$$\boxed{\max \sum_i c_i x_i}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$



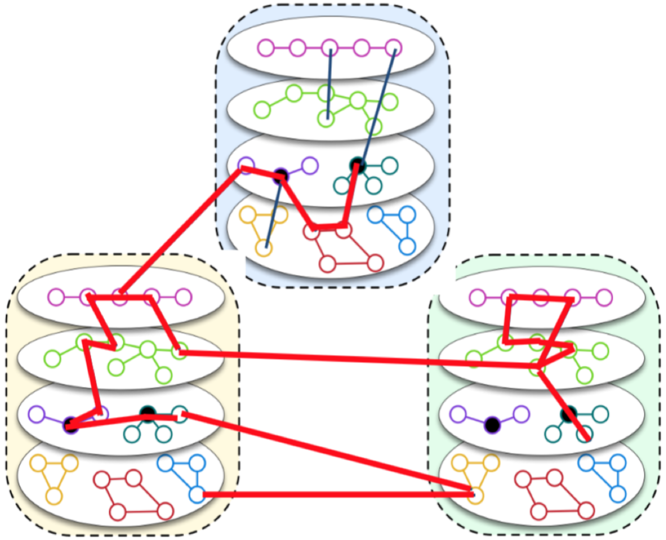
ILP Model: Some Details

Objective Function: "better" support graphs = higher objective value

- Reward good behavior:
  - High lexical match links, nearby alignments, using the subject if using a predicate-argument structure, WH-terms ("*which of energy …*"), etc.
- Penalize spurious overuse of frequently occurring terms

$$\boxed{\max \sum_i c_i x_i} \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \quad \ldots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$



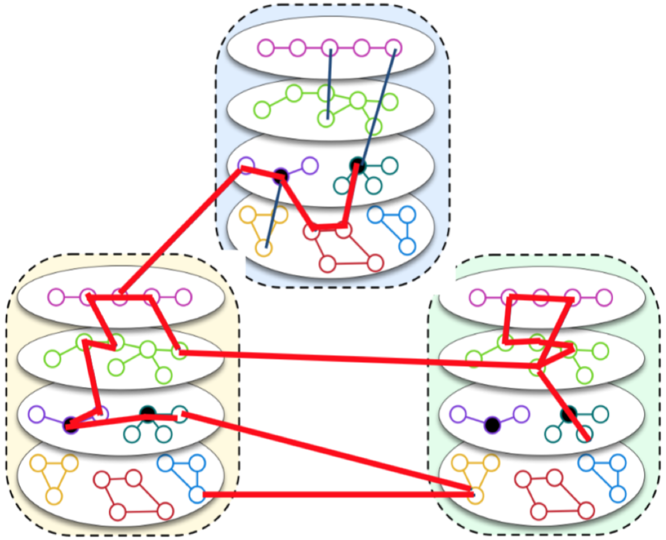
ILP Model: Some Details

Objective Function: "better" support graphs = higher objective value

- Reward good behavior:
  - High lexical match links, nearby alignments, using the subject if using a predicate-argument structure, WH-terms ("*which of energy …*"), etc.
- Penalize spurious overuse of frequently occurring terms

$$\boxed{\max \sum_i c_i x_i} \qquad \begin{cases} \sum_i a_{1i} x_i \leq b_1 \\ \dots \\ \sum_i a_{ki} x_i \leq b_k \end{cases}$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$



ILP Model: Some Details

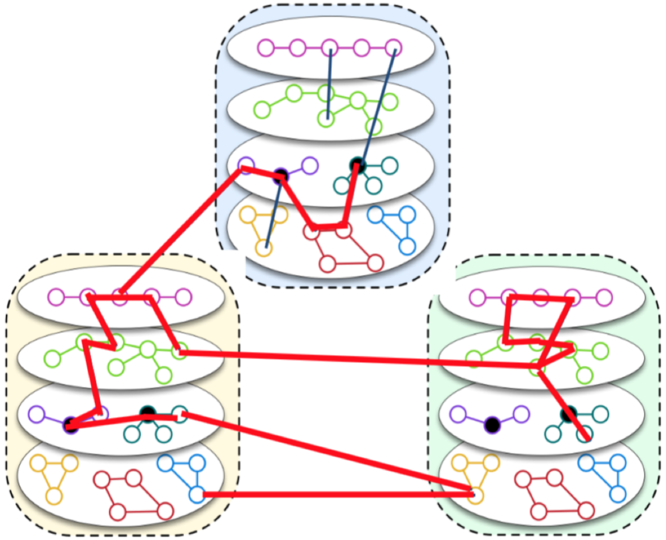
Dual goal: scalability, consider only meaningful support graphs

Incorporate global and local structure.

$$\max \sum_i c_i x_i$$

$$\forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}$$



ILP Model: Some Details

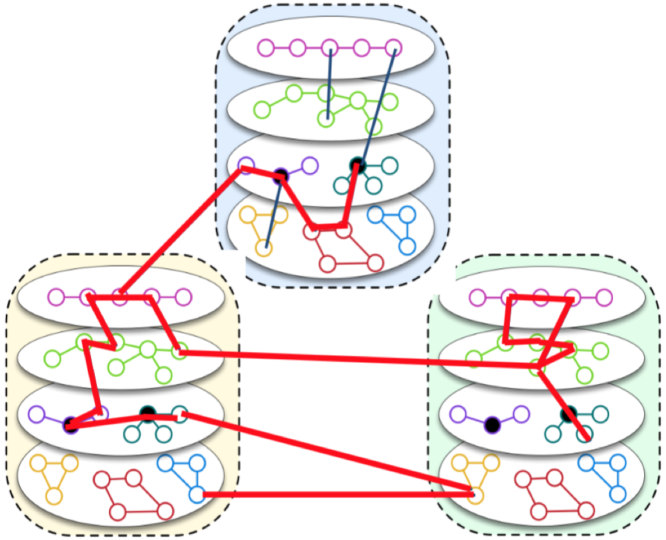
Dual goal: scalability, consider only meaningful support graphs

Incorporate global and local structure.

- **Structural Constraints**
  - Meaningful proof structures
    - connectedness, question coverage, etc.
    - single/multi-graph, etc.

$$\max \sum_i c_i x_i \qquad \forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\boxed{\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}}$$



ILP Model: Some Details

Dual goal: scalability, consider only meaningful support graphs

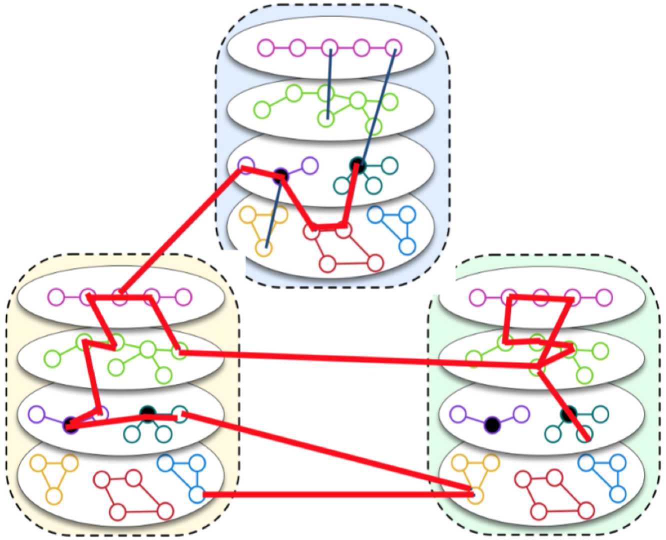
Incorporate global and local structure.

- **Structural Constraints**
  - Meaningful proof structures
    - connectedness, question coverage, etc.
    - single/multi-graph, etc.

$$\max \sum_i c_i x_i \qquad \forall\, x_i \in \mathbb{N} \cup \{0\}$$

$$\boxed{\begin{cases} \sum_i a_{1i} x_i \le b_1 \\ \quad \dots \\ \sum_i a_{ki} x_i \le b_k \end{cases}}$$

- **Semantic Constraints**
  - If using a predicate-argument graphs,
    - use at least predicate and argument



Evaluation: Notable Baselines

[Clark et al. AAAI'15]

[Khot et al. ACL'17]

[Seo et al. ICLR'16]

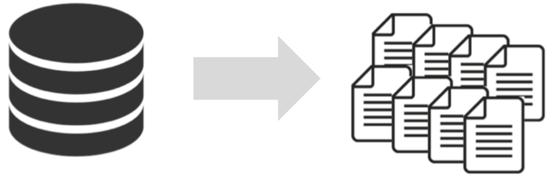
Evaluation: Notable Baselines

Information Retrieval (IR)

[Clark et al. AAAI'15]

Information retrieval baseline (Lucene)

Using 280 GB of plain text



[Khot et al. ACL'17]

[Seo et al. ICLR'16]

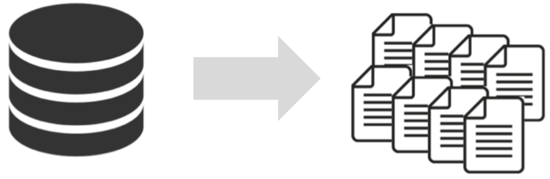
Evaluation: Notable Baselines

Information Retrieval (IR)

[Clark et al. AAAI'15]

Information retrieval baseline (Lucene)

Using 280 GB of plain text

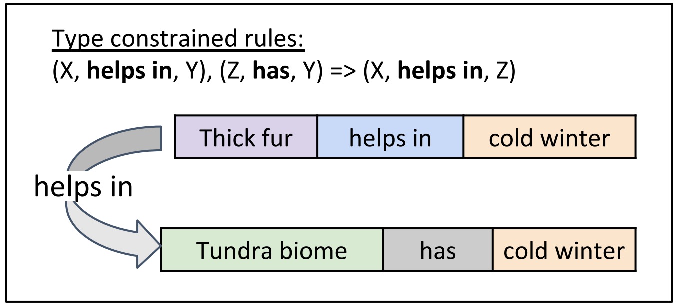


Inference over structure (TupleInf)

[Khot et al. ACL'17]

Inference over auto-generated short triples

And type-constrained rules



[Seo et al. ICLR'16]

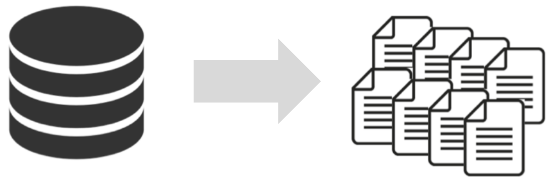
Evaluation: Notable Baselines

Information Retrieval (IR)

[Clark et al. AAAI'15]

Information retrieval baseline (Lucene)

Using 280 GB of plain text

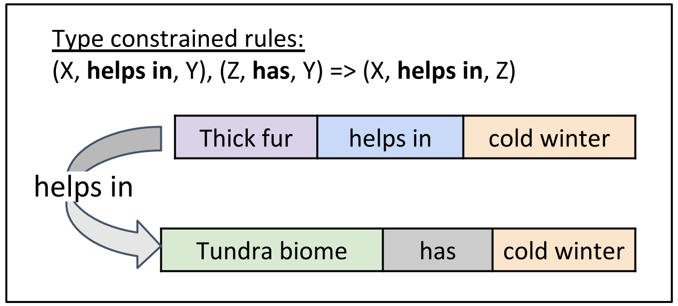


Inference over structure (TupleInf)

[Khot et al. ACL'17]

Inference over auto-generated short triples

And type-constrained rules

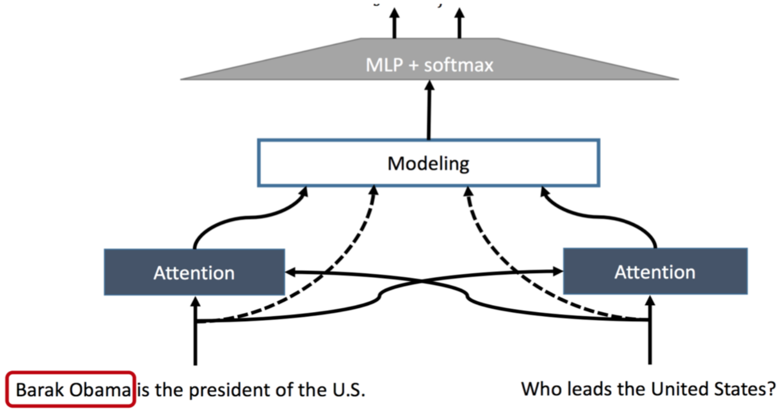


Neural Network (BiDAF)

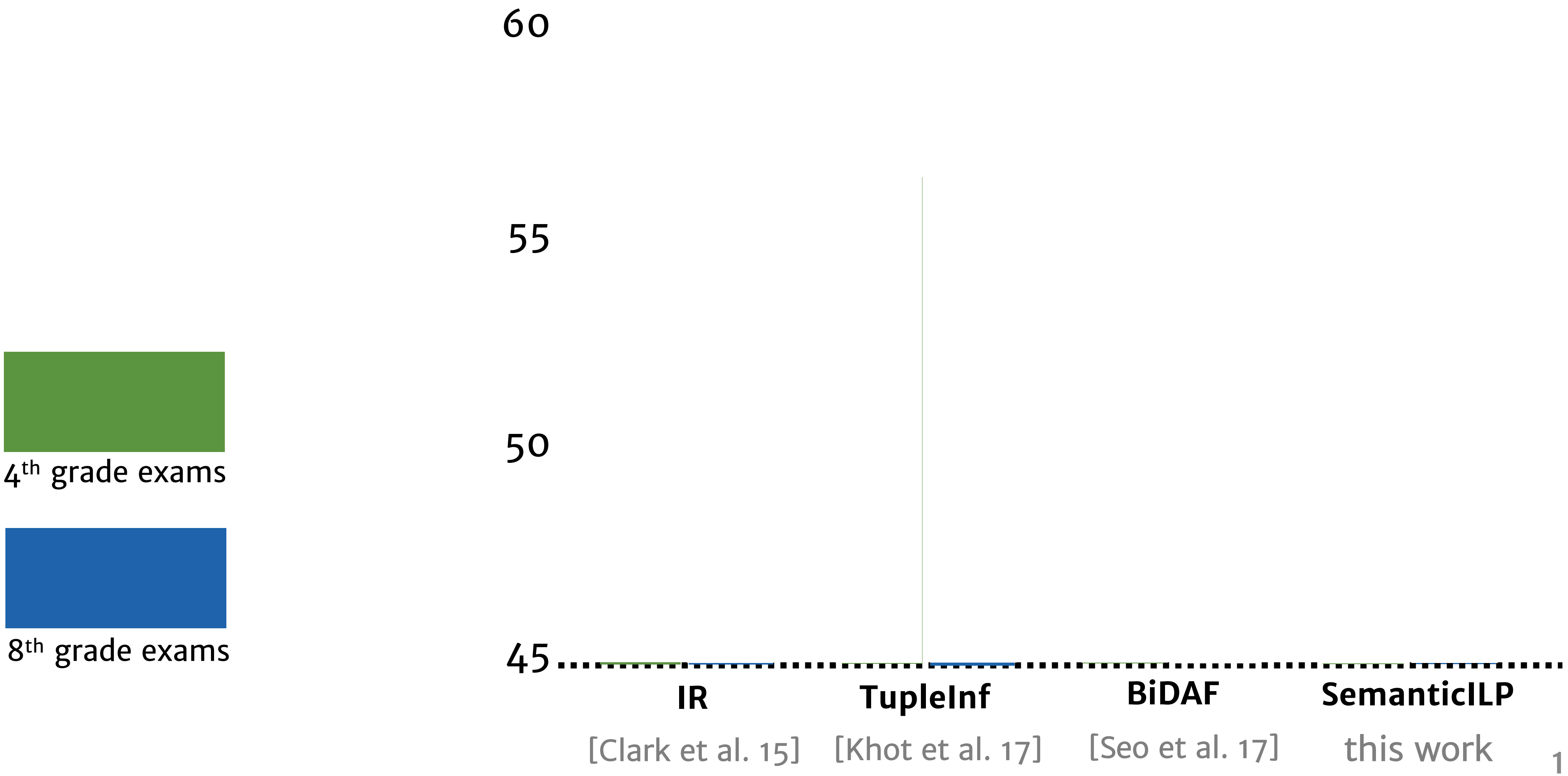
[Seo et al. ICLR'16]

Attention & LSTM

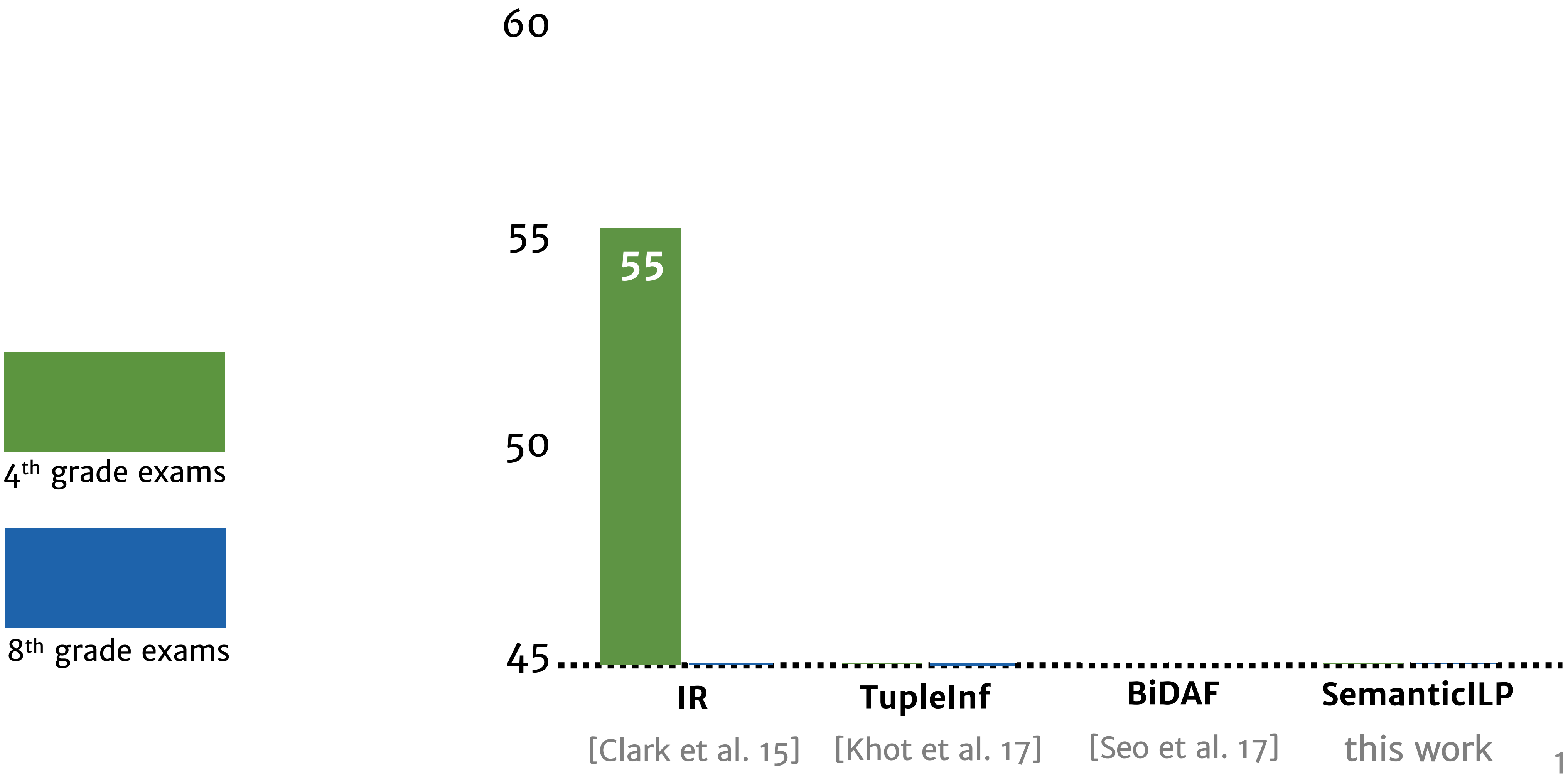
Extractive, i.e select a contiguous phrase in a given paragraph



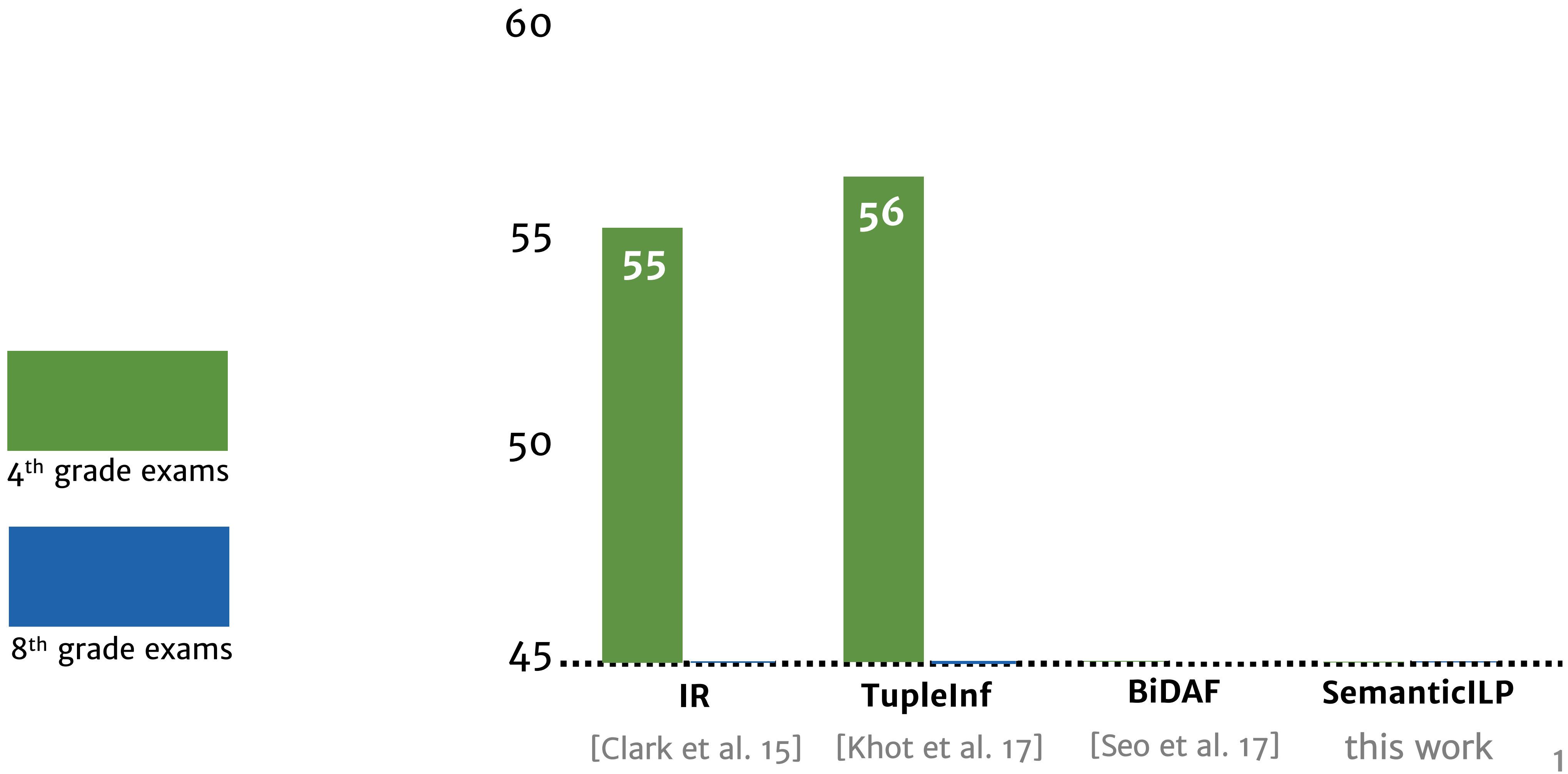
Empirical results: Science Domain [ZKTR'18]



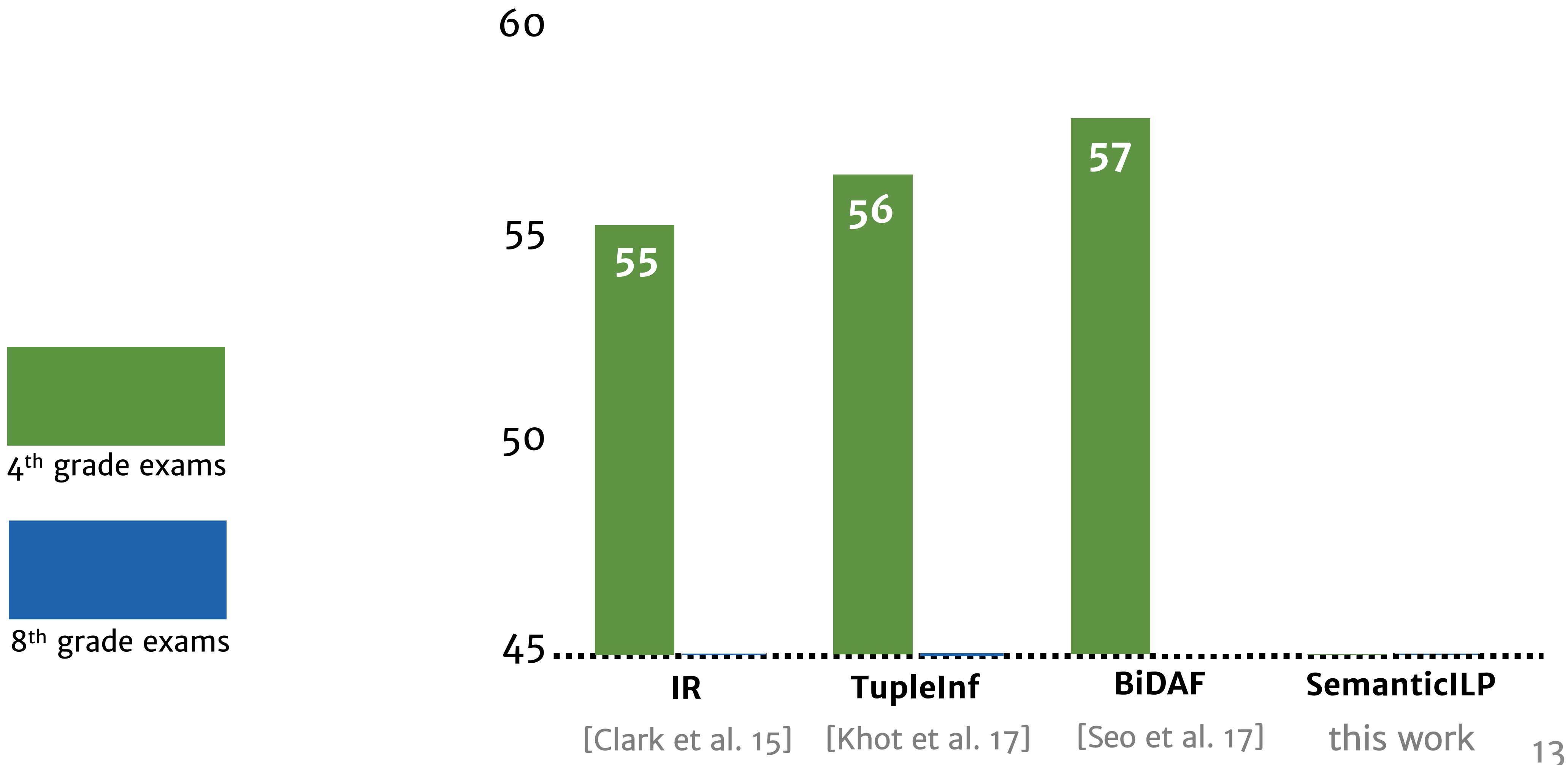
Empirical results: Science Domain [ZKTR'18]



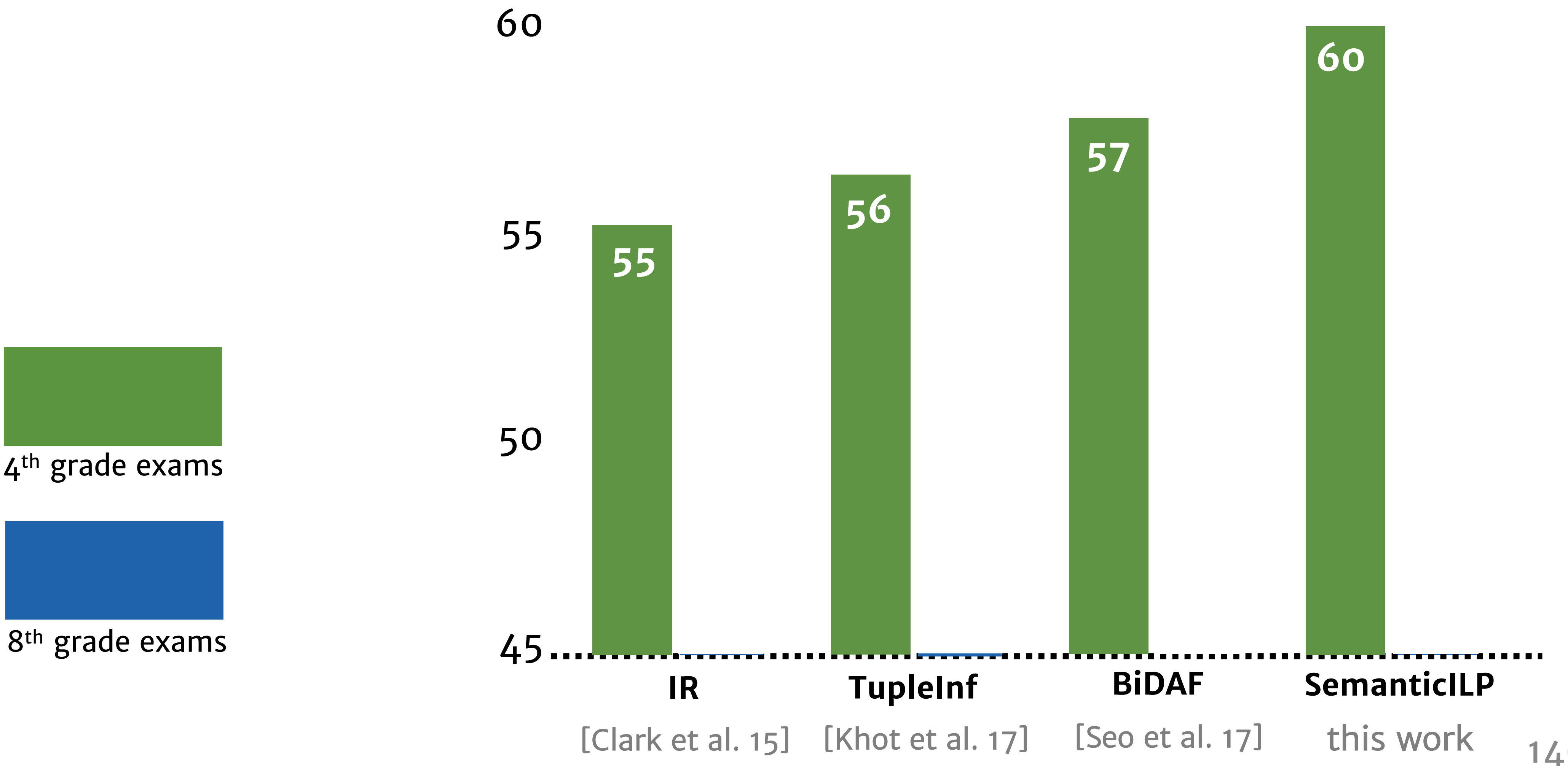
Empirical results: Science Domain [ZKTR'18]



Empirical results: Science Domain [ZKTR'18]

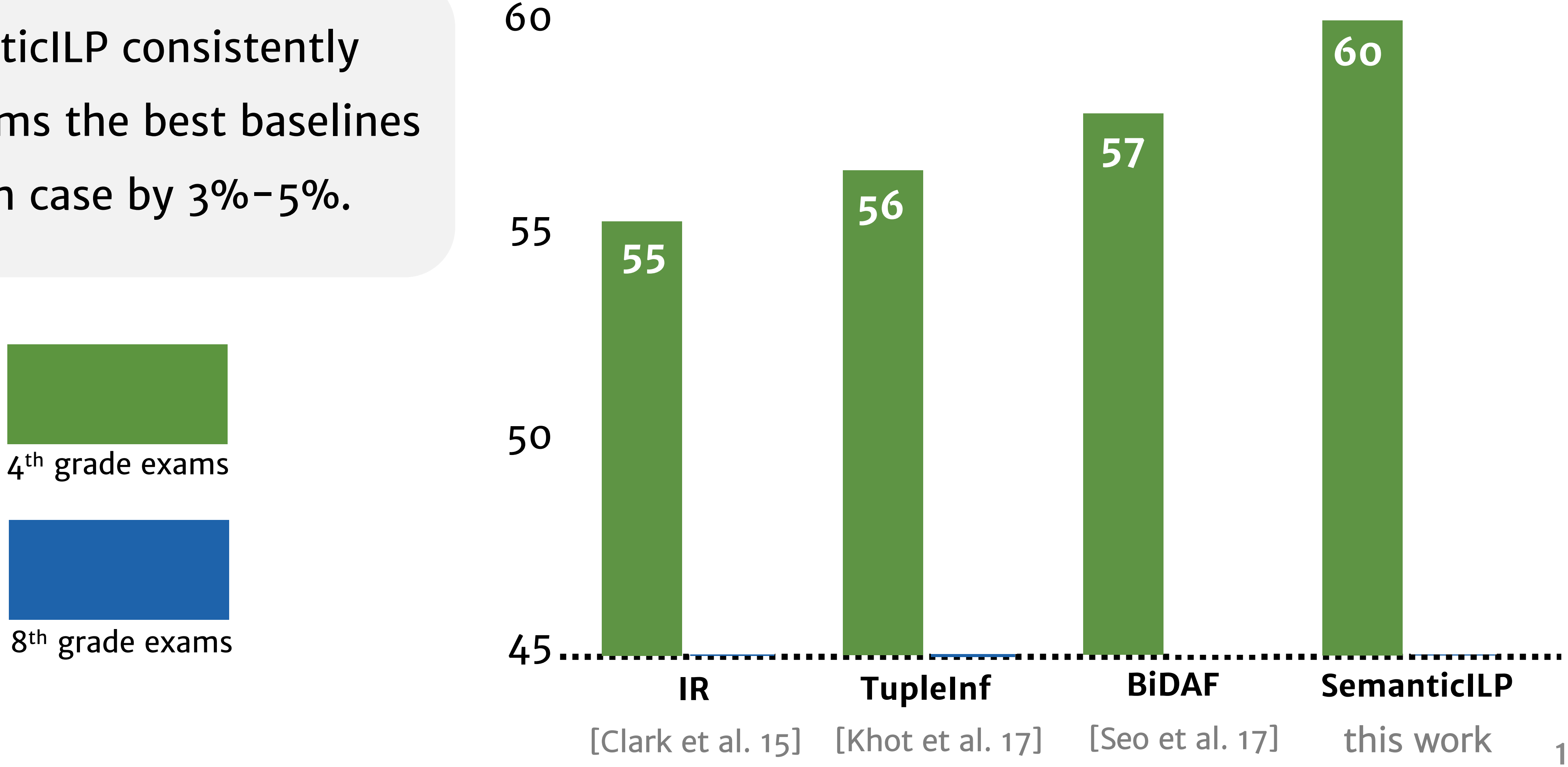


Empirical results: Science Domain [ZKTR'18]



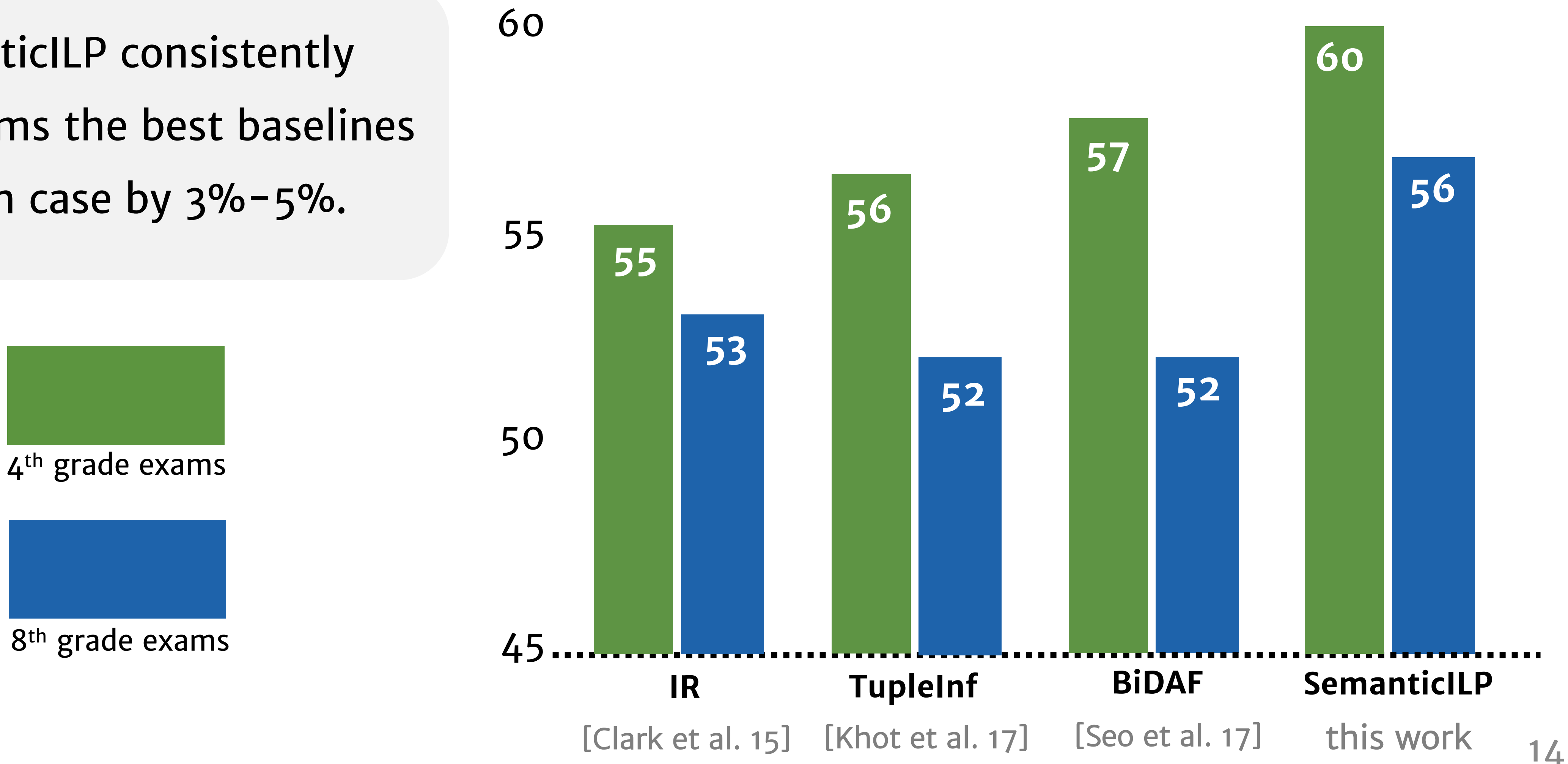
Empirical results: Science Domain [ZKTR'18]

SemanticILP consistently outperforms the best baselines in each case by 3%-5%.



Empirical results: Science Domain [ZKTR'18]

SemanticILP consistently outperforms the best baselines in each case by 3%–5%.

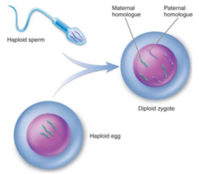


Answering Questions: Biology Exams

- **Biology exams** [Berant at al, 2014]
  - Technical terms and answer not easy to find.
  - Requires understanding complex relations.

Answering Questions: Biology Exams

- **Biology exams** [Berant at al, 2014]
  - Technical terms and answer not easy to find.
  - Requires understanding complex relations.



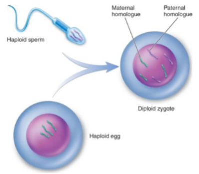
Question: *What does meiosis directly produce?*

\+

(A) Gametes
(B) Haploid cells

Answering Questions: Biology Exams

- **Biology exams** [Berant at al, 2014]
  - Technical terms and answer not easy to find.
  - Requires understanding complex relations.



Question: *What does meiosis directly produce?*

+

(A) Gametes
(B) Haploid cells

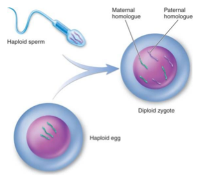
Evidence paragraph

*… **Meiosis produces not gametes but haploid cells** that then divide by mitosis and give rise to either unicellular descendants or a haploid multicellular adult organism. Subsequently, the haploid organism carries out further mitoses, producing the cells that develop into gametes….*

Answering Questions: Biology Exams

- **Biology exams** [Berant at al, 2014]
  - Technical terms and answer not easy to find.
  - Requires understanding complex relations.

We use **the same** version of our systems across our datasets.



Question: *What does meiosis directly produce?*

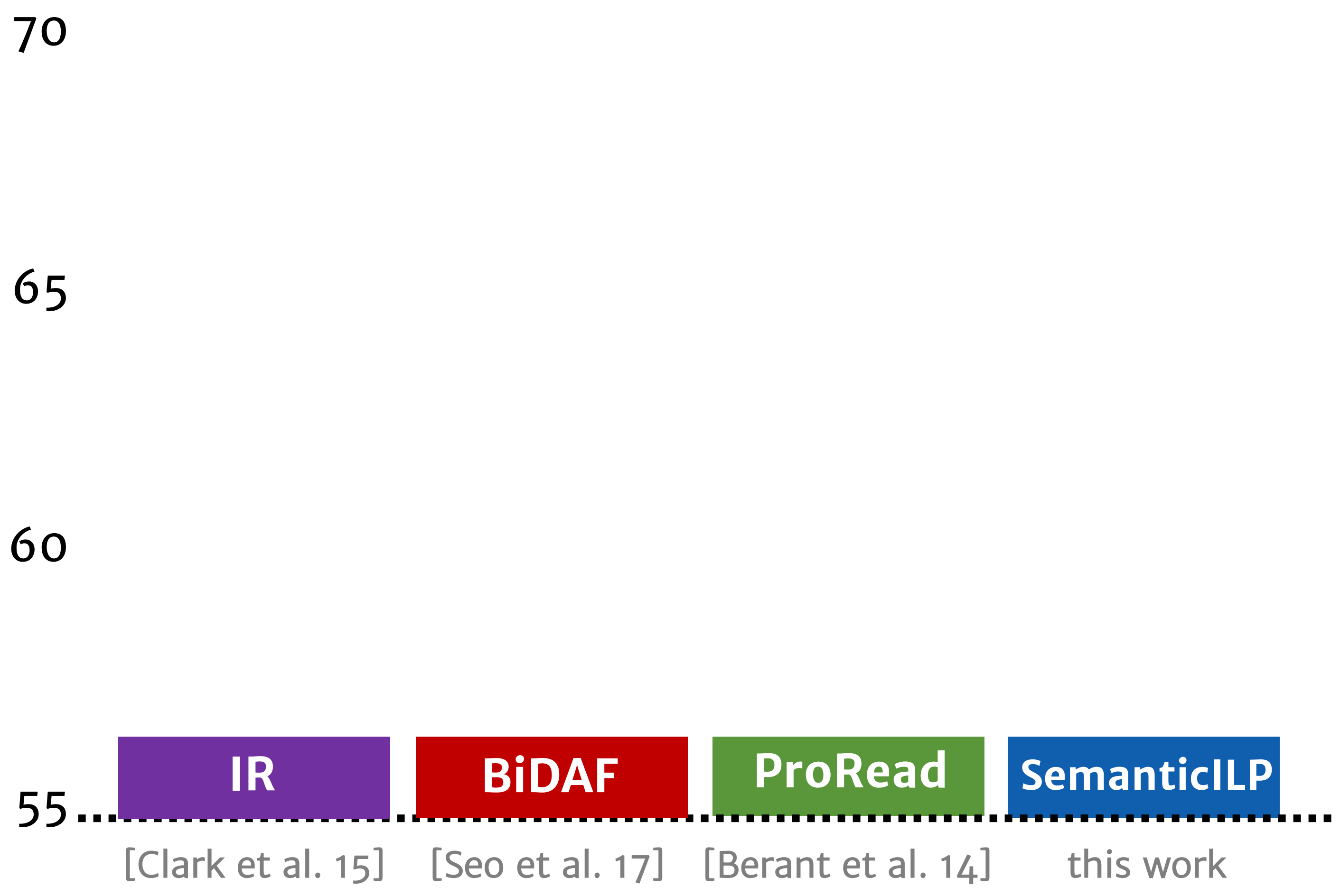
+

(A) Gametes
(B) Haploid cells

Evidence paragraph

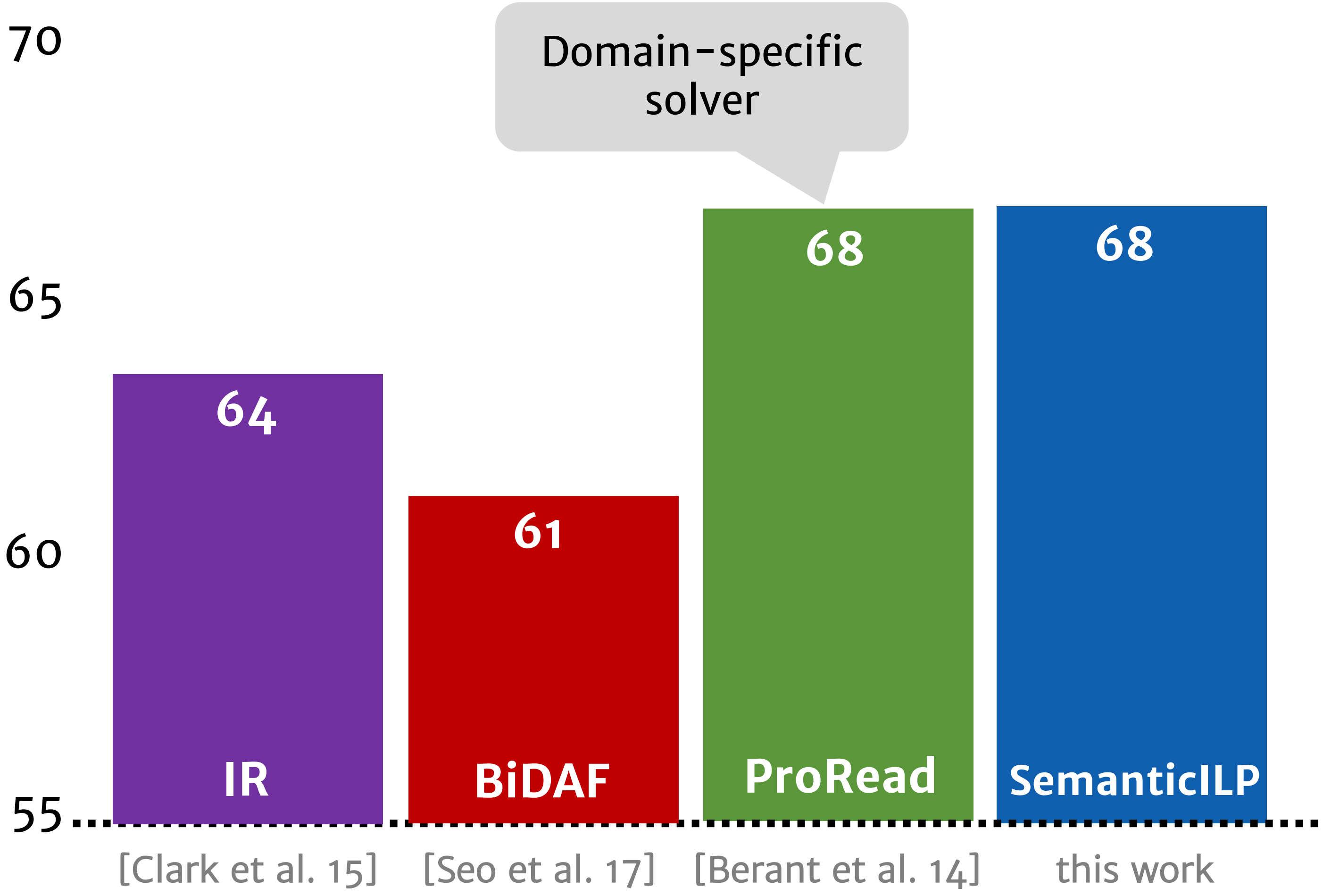
*… **Meiosis produces not gametes but haploid cells** that then divide by mitosis and give rise to either unicellular descendants or a haploid multicellular adult organism. Subsequently, the haploid organism carries out further mitoses, producing the cells that develop into gametes….*

Empirical results: Biology Domain [ZKTR'18]

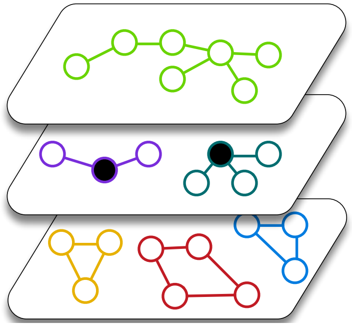


Empirical results: Biology Domain [ZKTR'18]

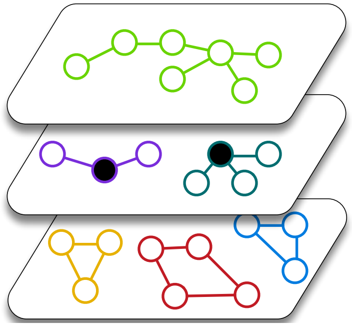
SemanticILP generalizes to a **different** domain and achieves on-par score with the best domain-specific system.



Lessons



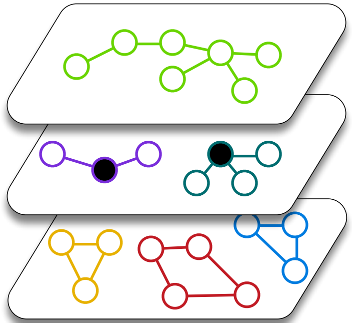
Lessons



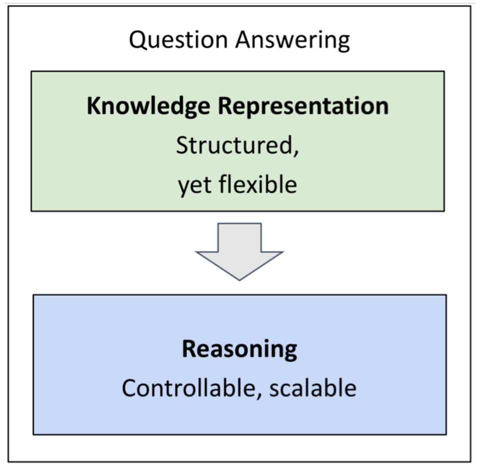
- Reasoning over language requires dealing with a diverse set of semantic phenomena.
- Collection of semantic representations of language, **independent** of the task (**indirect** supervision).
- Better generalization across two different domains.

Lessons

- Reasoning over language requires dealing with a diverse set of semantic phenomena.



- Collection of semantic representations of language, **independent** of the task (**indirect** supervision).



- Better generalization across two different domains.

ENTITY TYPING *with minimal supervision*

Zhou, **K** et al. Zero-Shot Open Entity Typing as Type-Compatible Grounding. EMNLP 18.
Fei, **K** et al. Illinois-Profiler: Knowledge Schemas at Scale. IJCAI (Cognitum) 15.

SEMANTIC TYPING OF ENTITIES

Label mentions with their semantic **types.**

*A handful of professors in the **CMU** **Department of Chemistry** are being recognized for their efforts and contributions to the scientific community.*

CMU:

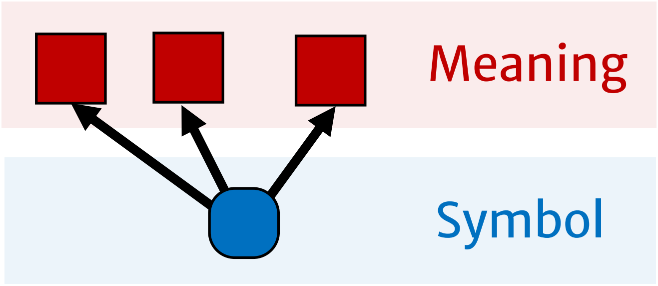
```
/organization
/organization/education_institution
```

Department of Chemistry:

```
/organization
/education
/education/department
```

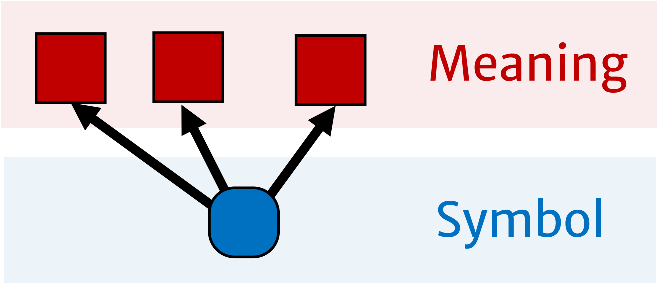
Semantic Entity Typing: The Necessity (1)

- Dealing with ambiguity



Semantic Entity Typing: The Necessity (1)

- Dealing with ambiguity



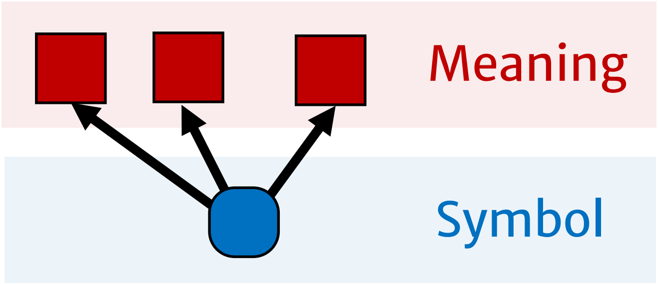
*Our break in **Paris** was quite memorable.*

*I met a girl named **Paris**.*

***Paris** issued a statement condemning the proposal.*

Semantic Entity Typing: The Necessity (1)

- Dealing with ambiguity



*Our break in **Paris** was quite memorable.*

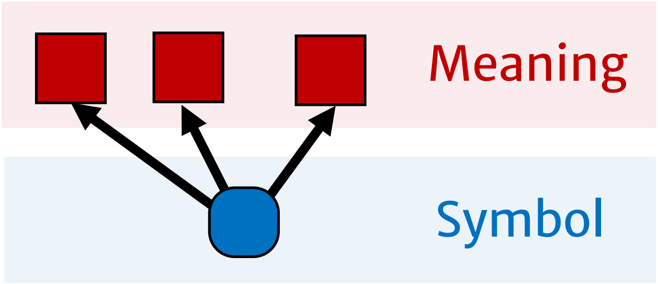
*I met a girl named **Paris**.*

***Paris** issued a statement condemning the proposal.*

city

Semantic Entity Typing: The Necessity (1)

- Dealing with ambiguity



*Our break in **Paris** was quite memorable.*

city

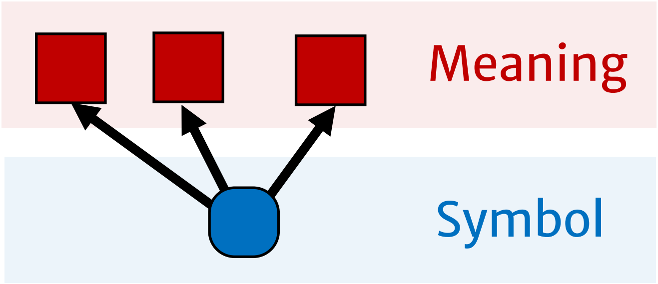
*I met a girl named **Paris**.*

person

***Paris** issued a statement condemning the proposal.*

Semantic Entity Typing: The Necessity (1)

- Dealing with ambiguity



Our break in ***Paris*** *was quite memorable.*

`city`

I met a girl named ***Paris****.*

`person`

Paris *issued a statement condemning the proposal.*

`government`

Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].

Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].

*Which month receives the most hours of daylight in **Paris**?*

“June”

Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].

Which month receives the most hours of daylight in ~~**Paris**~~*?*

Pittsburgh

“June”

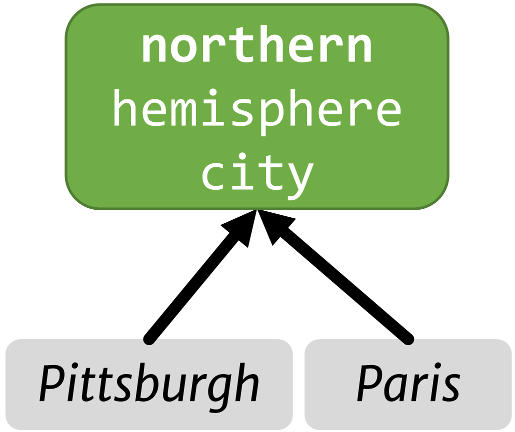
Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].

Which month receives the most hours of daylight in ~~**Paris**~~*?*

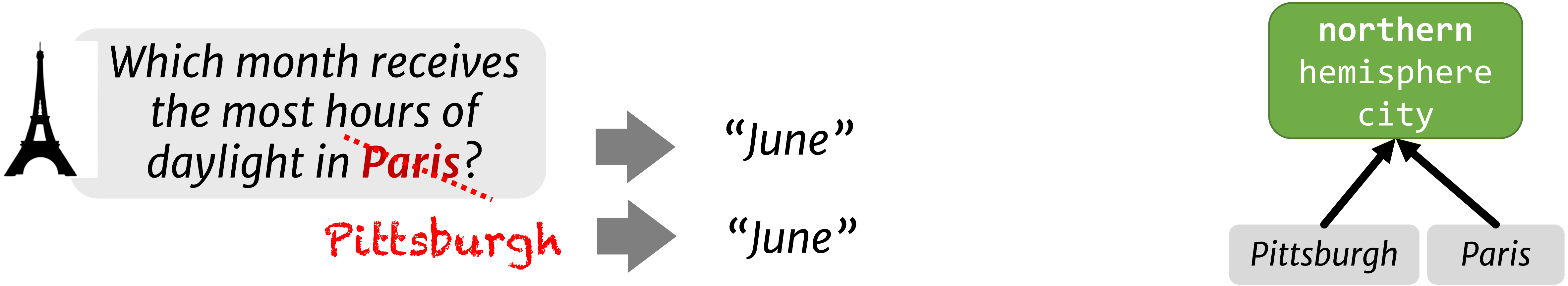
Pittsburgh

“June”



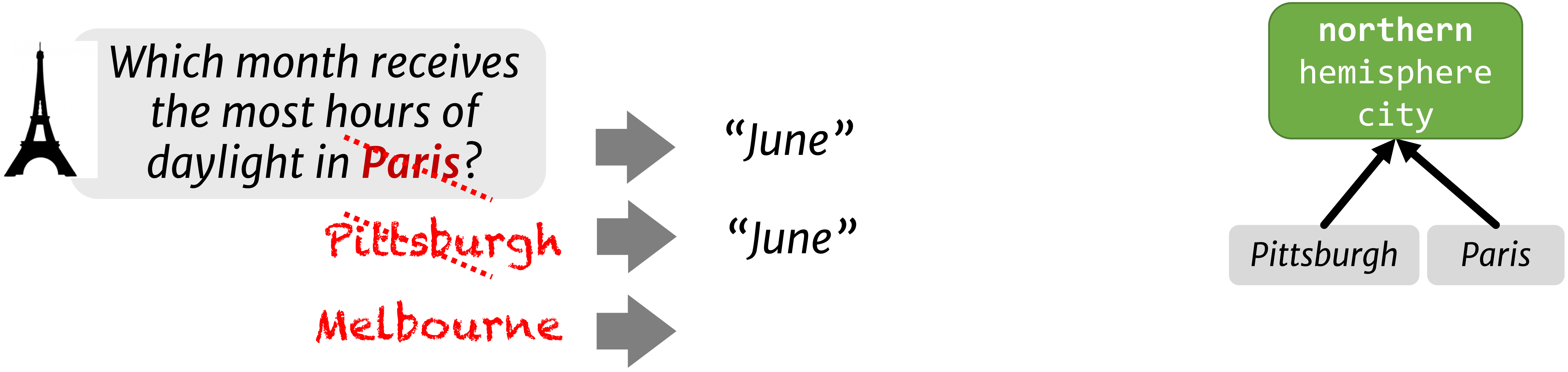
Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].



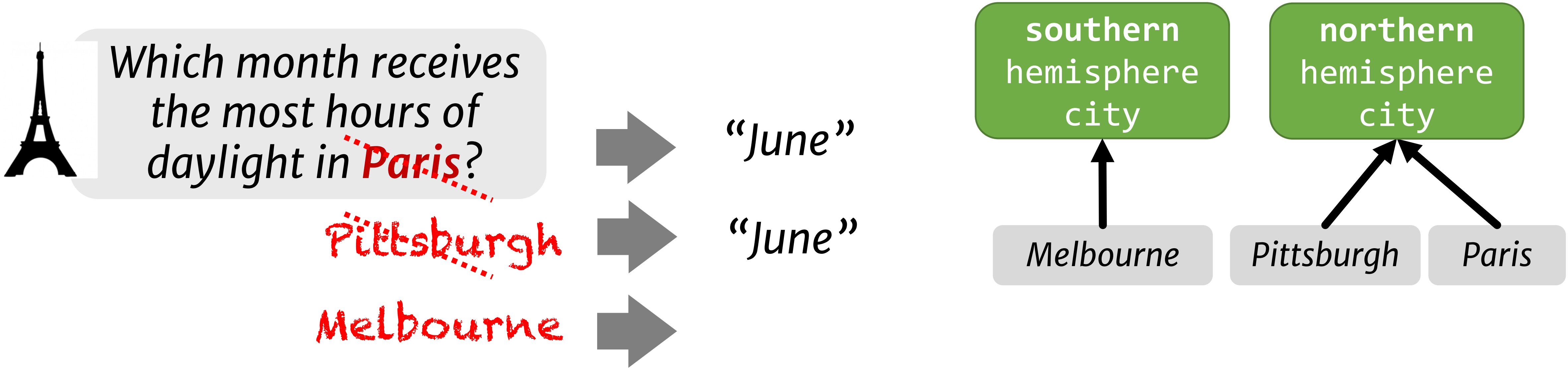
Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires "abstractions".
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].



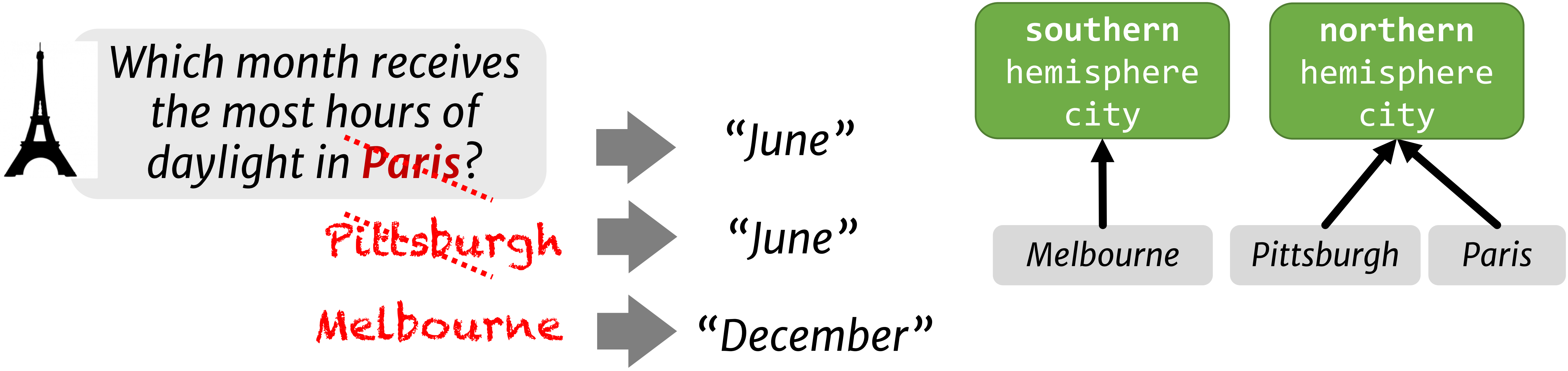
Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].



Semantic Entity Typing: The Necessity (2)

- Linguistic **generalization** requires “abstractions”.
  - Ex: Question answering [Yavuz,16]; Information Extraction [Ling,12].



Entity Typing: Existing Work

- Multiple datasets for semantic typing

Entity Typing: Existing Work

- Multiple datasets for semantic typing

granularity

Entity Typing: Existing Work

- Multiple datasets for semantic typing

granularity

Coarse Typing

Entity Typing: Existing Work

- Multiple datasets for semantic typing

[Sang&Meulder,03] [Hovy,06]

CoNLL
4 types

OntoNotes
18 types

granularity

Coarse Typing

Entity Typing: Existing Work

- Multiple datasets for semantic typing

[Sang&Meulder,03] [Hovy,06]

CoNLL
4 types

OntoNotes
18 types

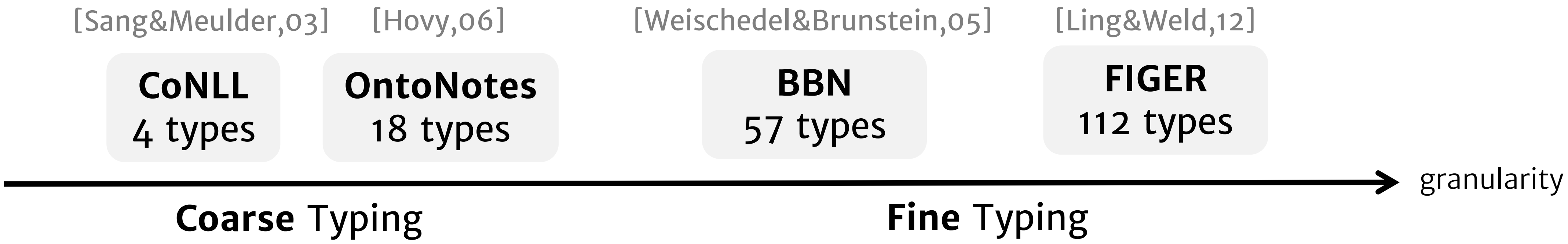
granularity

Coarse Typing

Fine Typing

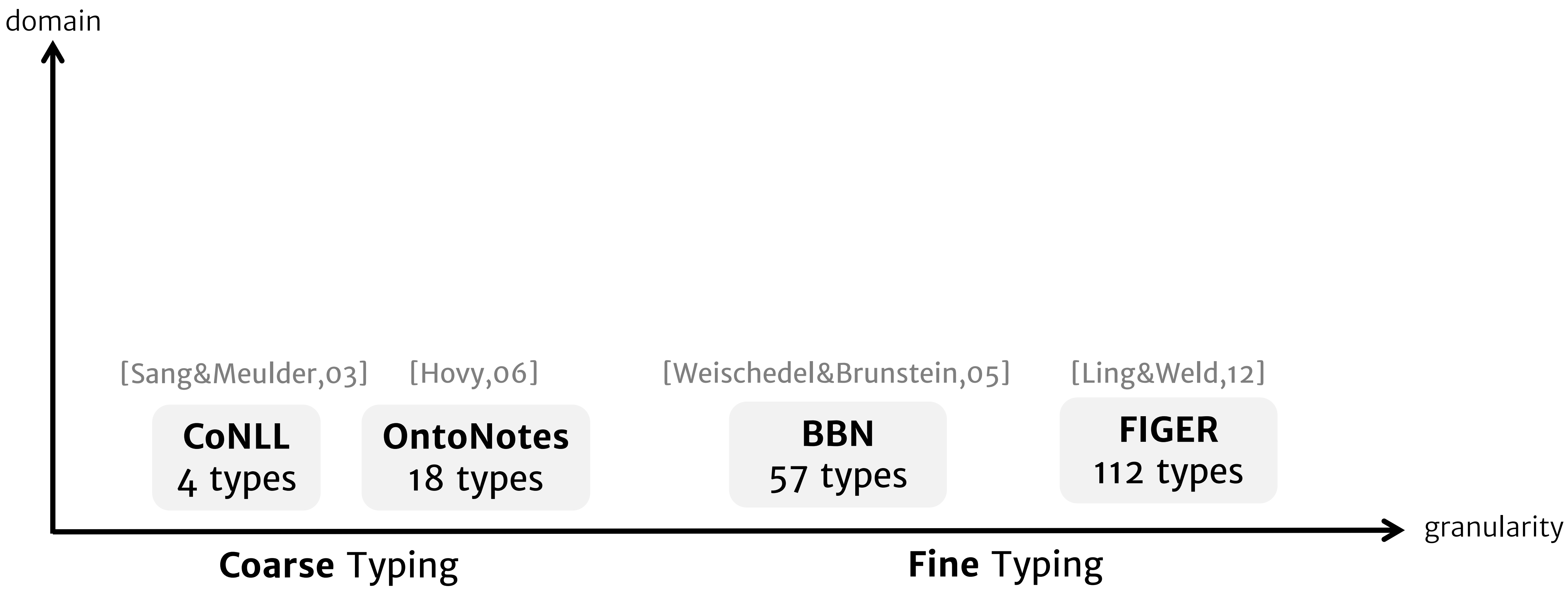
Entity Typing: Existing Work

- Multiple datasets for semantic typing



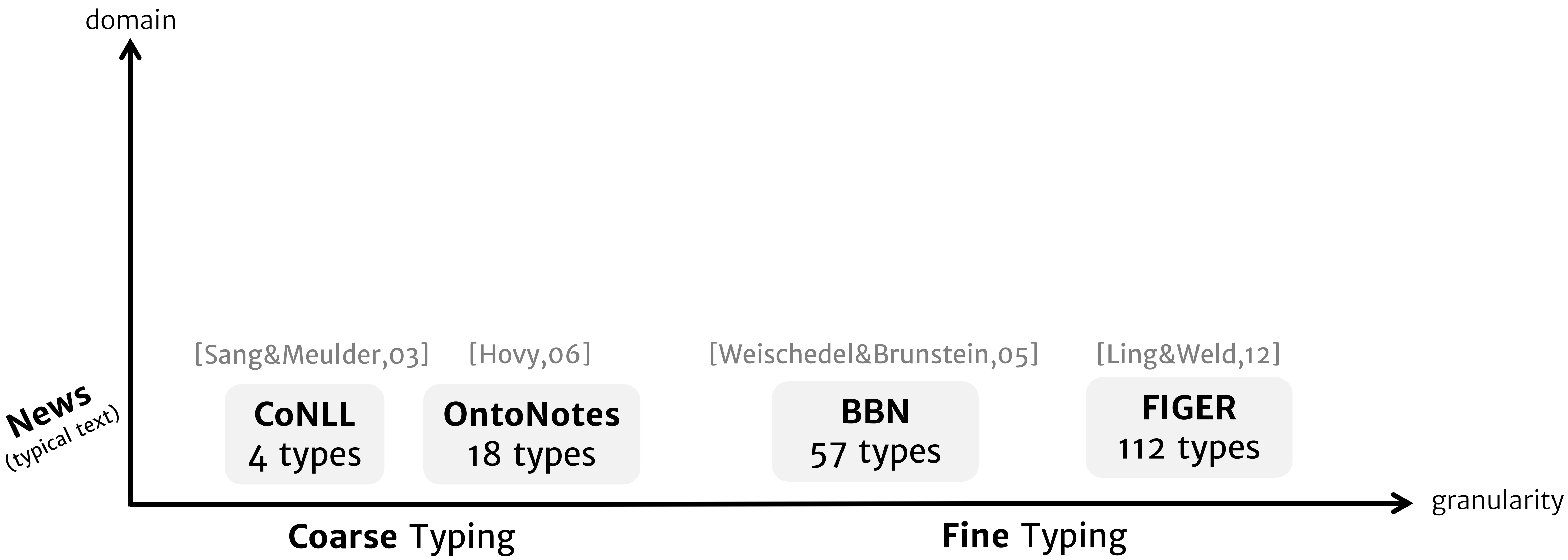
Entity Typing: Existing Work

- Multiple datasets for semantic typing



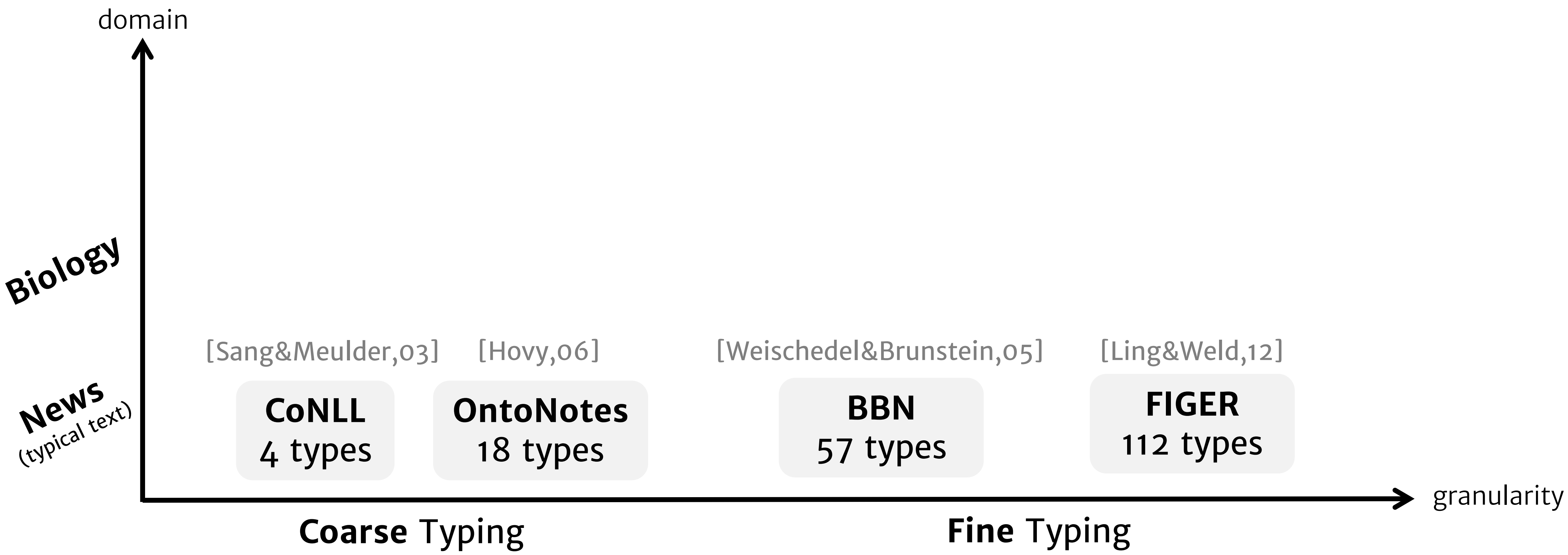
Entity Typing: Existing Work

- Multiple datasets for semantic typing



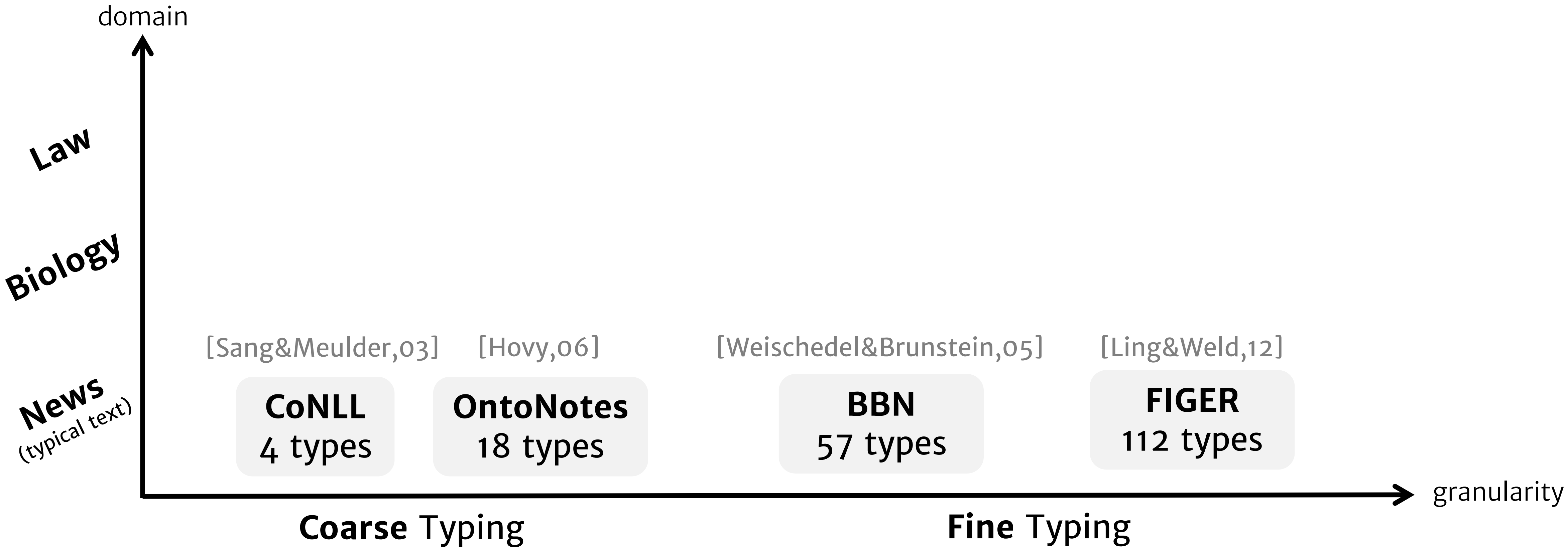
Entity Typing: Existing Work

- Multiple datasets for semantic typing



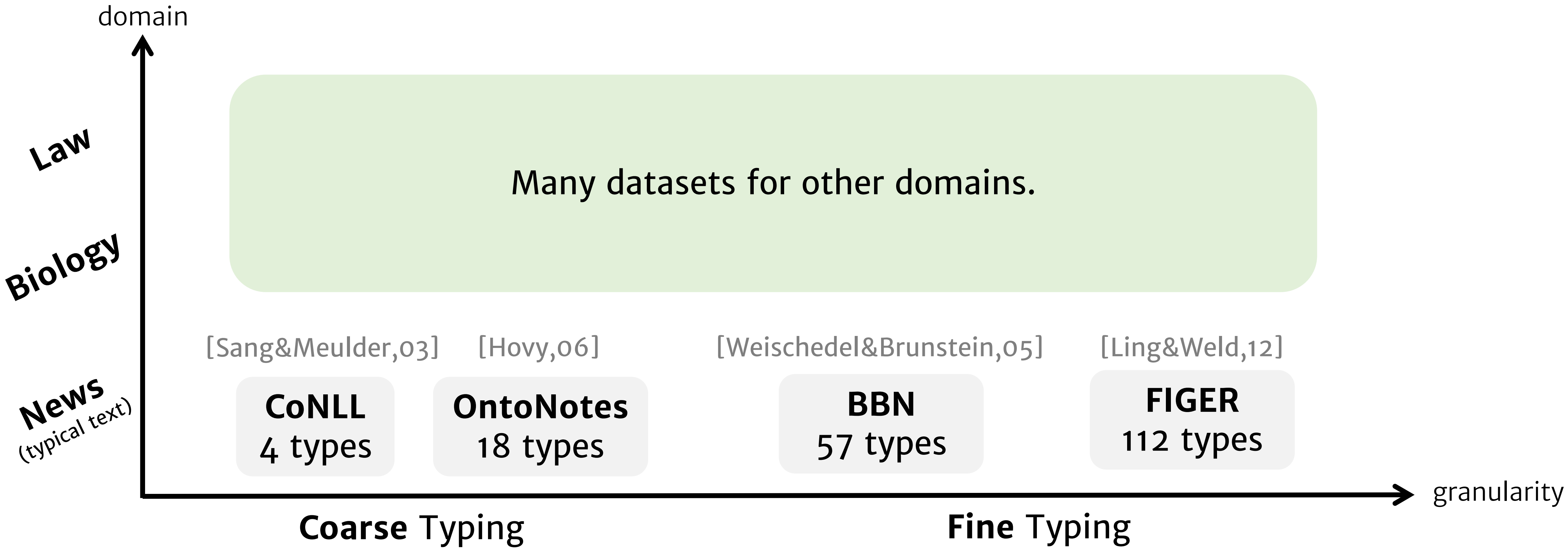
Entity Typing: Existing Work

- Multiple datasets for semantic typing



Entity Typing: Existing Work

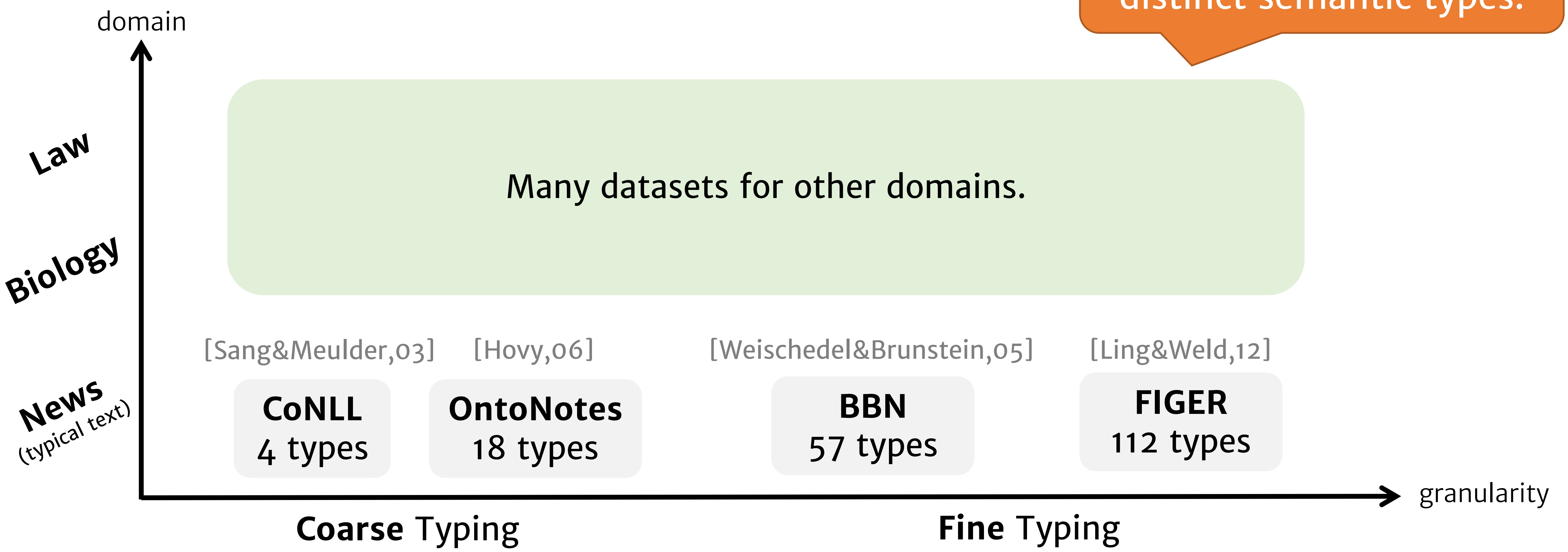
- Multiple datasets for semantic typing



Entity Typing: Existing Work

- Multiple datasets for semantic typing

Many datasets, each with distinct semantic types.



“Cheap” Typing with Wikipedia

A former Democrat, ***Bloomberg*** *switched his party registration in 2001.*

“Cheap” Typing with Wikipedia

A former Democrat, ***Bloomberg*** *switched his party registration in 2001.*

Entity Linking

[Ratinov et al. 11]

"Cheap" Typing with Wikipedia

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Entity Linking

[Ratinov et al. 11]

Freebase

[Bollacker et al. 08]

"Cheap" Typing with Wikipedia

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Entity Linking

[Ratinov et al. 11]

Freebase

[Bollacker et al. 08]

politician

businessman

philanthropist

“Cheap” Typing with Wikipedia

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Entity Linking

[Ratinov et al. 11]

WIKIPEDIA

Freebase

[Bollacker et al. 08]

politician

businessman

philanthropist

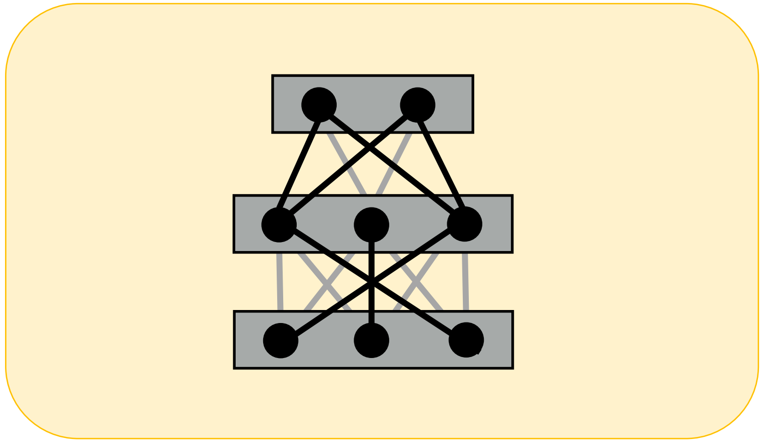
Not consistent with the context

A Common Approach: Supervised Learning

- **Input:** sentence, mention.
- **Output:** a set of types.

A Common Approach: Supervised Learning

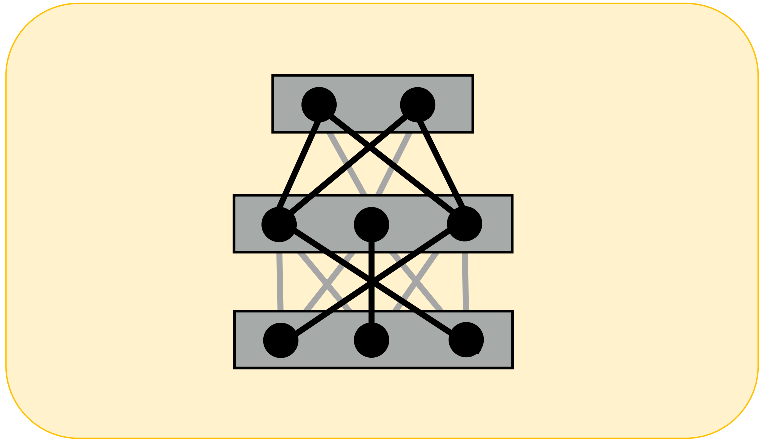
- **Input:** sentence, mention.
- **Output:** a set of types.



A Common Approach: Supervised Learning

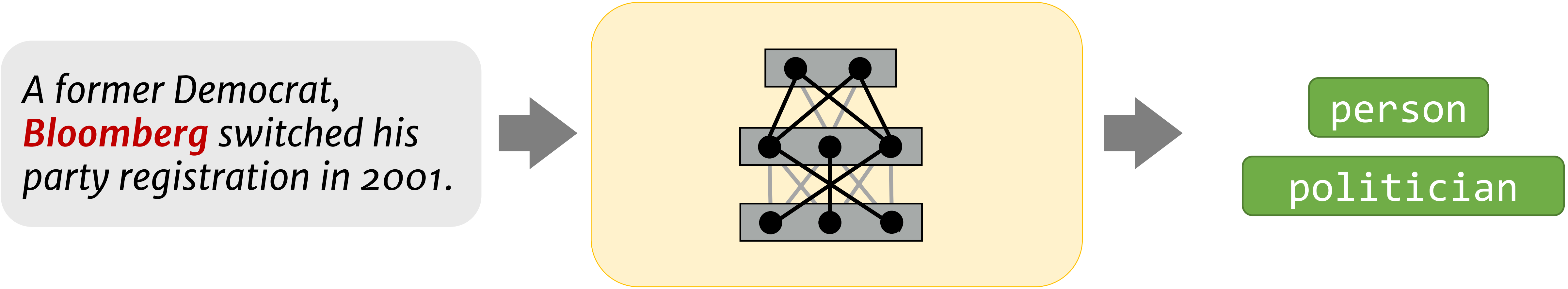
- **Input:** sentence, mention.
- **Output:** a set of types.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*



A Common Approach: Supervised Learning

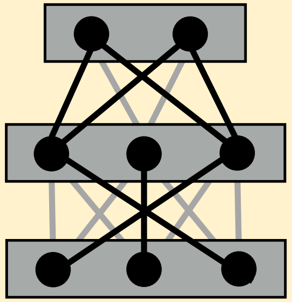
- **Input:** sentence, mention.
- **Output:** a set of types.



A Common Approach: Supervised Learning

- **Input:** sentence, mention.
- **Output:** a set of types.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*



person

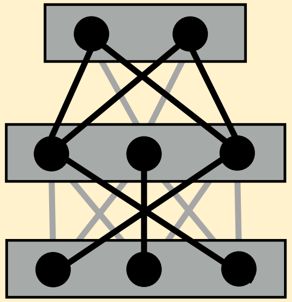
politician

Taxonomy is [indirectly] defined during the **training** time.

A Common Approach: Supervised Learning

- **Input:** sentence, mention.
- **Output:** a set of types.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*



Zero-Shot Open Entity Typing

- **Input:** sentence, mention
- **Output:** a set of types

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Zero-Shot Open Entity Typing

- **Input:** sentence, mention
- **Output:** a set of types

A former Democrat, ***Bloomberg*** *switched his party registration in 2001.*

ZOE

(Zero-shot Open Entity Typing)

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE
(Zero-shot Open Entity Typing)

Target Taxonomy

`person` `location` `organization`

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE
(Zero-shot Open Entity Typing)

Target Taxonomy

person location
organization

Classifies into a **given taxonomy**

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE
(Zero-shot Open Entity Typing)

person

Target Taxonomy

person
location
organization

Classifies into a **given taxonomy**

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE
(Zero-shot Open Entity Typing)

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE
(Zero-shot Open Entity Typing)

Target Taxonomy

- person
  - politician
  - doctor
- org
  - hospital
  - hotel

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE

(Zero-shot Open Entity Typing)

person

politician

Target Taxonomy

person → politician, doctor

org → hospital, hotel

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Open:
Classifies into a given taxonomy

ZOE
(Zero-shot Open Entity Typing)

person
politician

Target Taxonomy

person
politician
doctor
org
hospital
hotel

Zero-Shot Open Entity Typing

- **Input:** sentence, mention, **target taxonomy**.
- **Output:** a set of types (according to the target type taxonomy).

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ZOE

(Zero-shot Open Entity Typing)

person

politician

Open:
Classifies into a given taxonomy

Zero-shot:
no taxonomy-specific supervision

Target Taxonomy

- person
  - politician
  - doctor
- org
  - hospital
  - hotel

ZOE: Type-Compatible Grounding

- “Type” as conceptual container binding entities together.

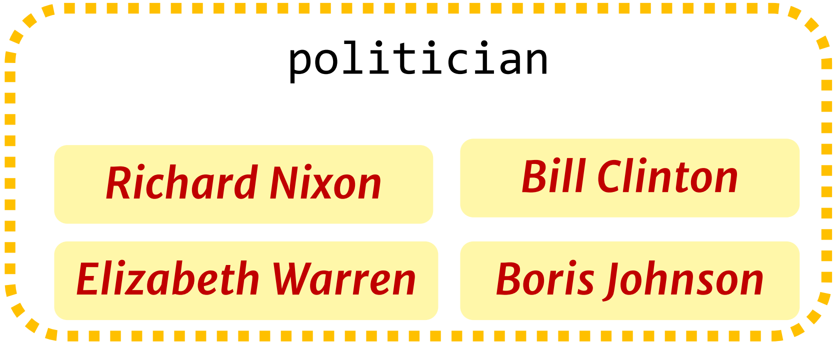
ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

politician

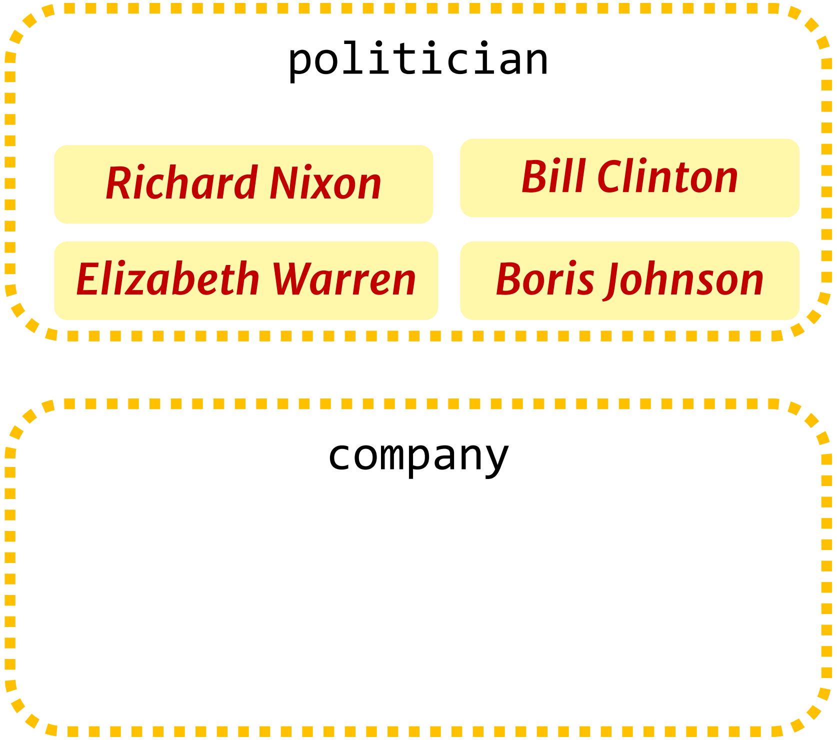
ZOE: Type-Compatible Grounding

- “Type” as conceptual container binding entities together.



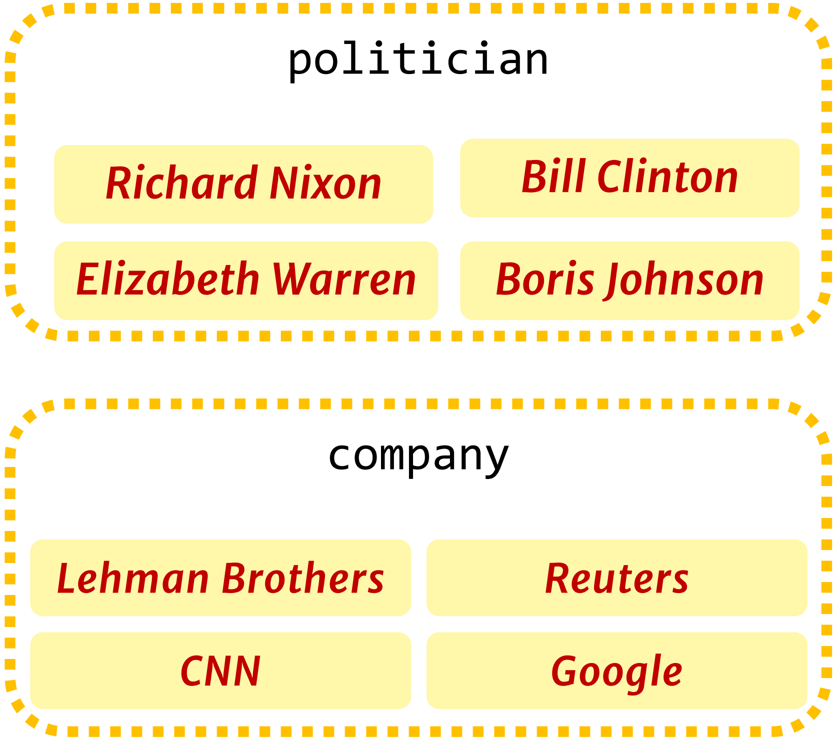
ZOE: Type-Compatible Grounding

- “Type” as conceptual container binding entities together.



ZOE: Type-Compatible Grounding

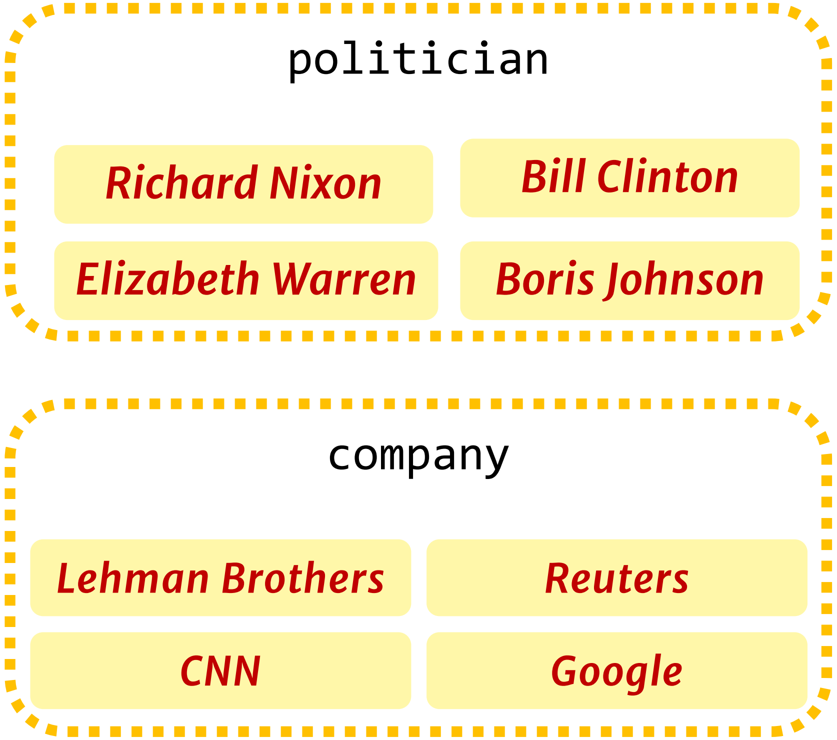
- “Type” as conceptual container binding entities together.



ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

A former Democrat, ***Bloomberg*** *switched his party registration in 2001.*



ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Key idea: Determine the **type** of an input mention by finding entities in the **type defining set** that share a similar context

politician

Richard Nixon ***Bill Clinton***

Elizabeth Warren ***Boris Johnson***

company

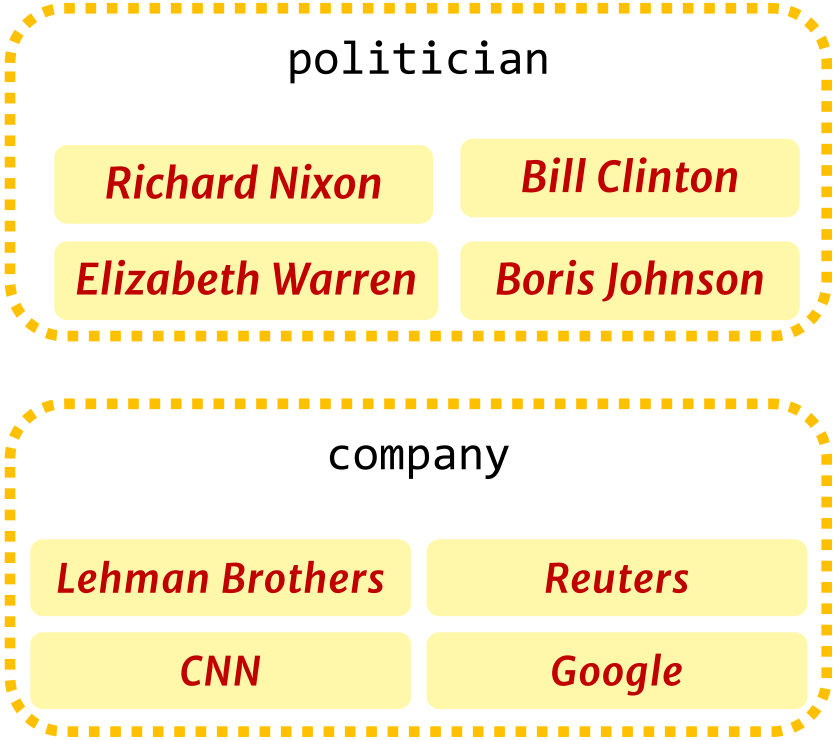
Lehman Brothers ***Reuters***

CNN ***Google***

ZOE: Type-Compatible Grounding

- “Type” as conceptual container binding entities together.

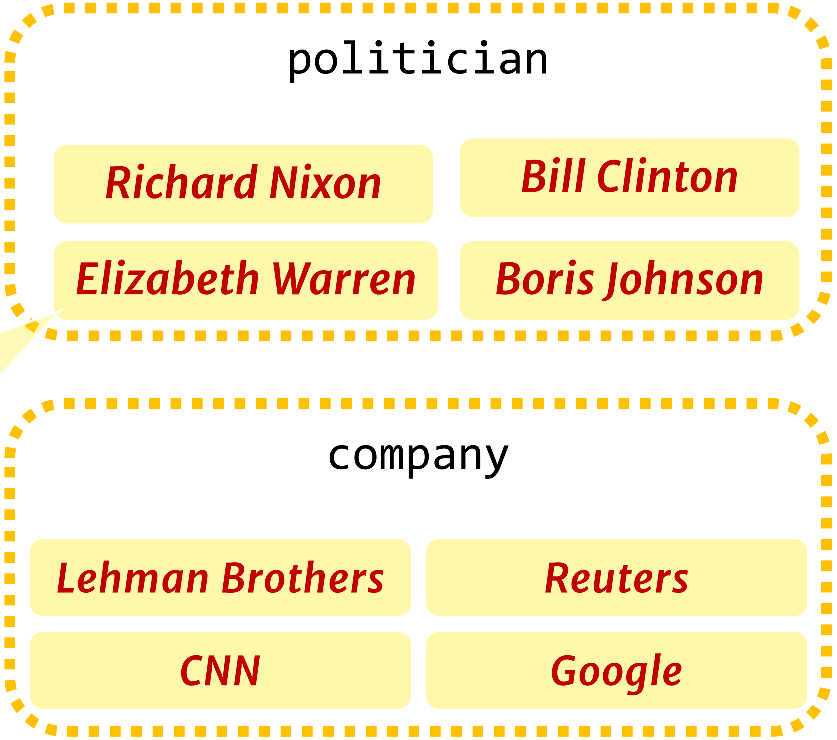
A former Democrat, ***Bloomberg*** *switched his party registration in 2001.*



ZOE: Type-Compatible Grounding

- “Type” as conceptual container binding entities together.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*



Different contexts where ***“Elizabeth Warren”*** is mentioned

WikiLinks [Singh et al. 12]

ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

***Warren** is a member of the Democratic Party, after switching party affiliation from the Republican Party in 1996.*

Different contexts where ***"Elizabeth Warren"*** is mentioned

WikiLinks [Singh et al. 12]

politician

Richard Nixon ***Bill Clinton***

Elizabeth Warren ***Boris Johnson***

company

Lehman Brothers ***Reuters***

CNN ***Google***

ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

ELMo [Peters et al. 18]

Context-consistent

***Warren** is a member of the Democratic Party, after switching party affiliation from the Republican Party in 1996.*

Different contexts where ***"Elizabeth Warren"*** is mentioned

WikiLinks [Singh et al. 12]

politician

Richard Nixon ***Bill Clinton***

Elizabeth Warren ***Boris Johnson***

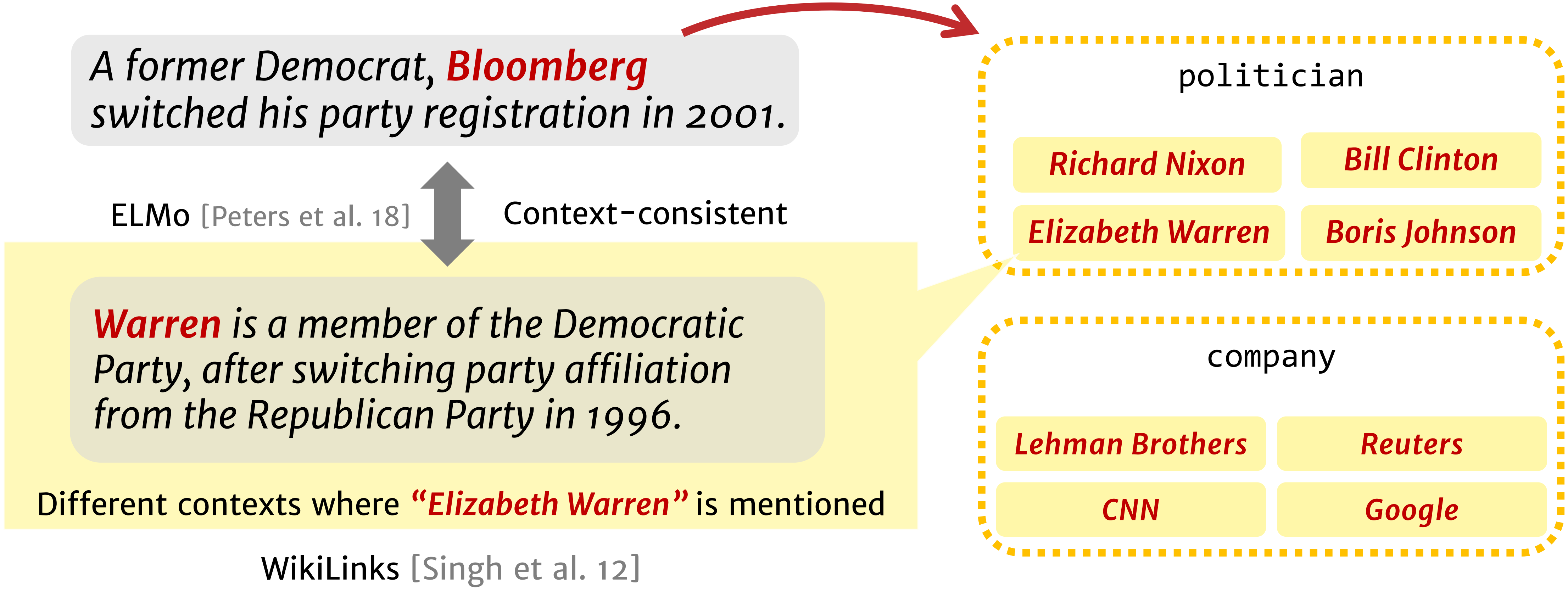
company

Lehman Brothers ***Reuters***

CNN ***Google***

ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.



ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*Since its founding, **Bloomberg** has made several acquisitions.*

politician

Richard Nixon | ***Bill Clinton***

Elizabeth Warren | ***Boris Johnson***

company

Lehman Brothers | ***Reuters***

Berkshire Hathaway | ***Enron***

ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*Since its founding, **Bloomberg** has made several acquisitions.*

Different contexts where *"Berkshire Hathaway"* is mentioned

politician

Richard Nixon ***Bill Clinton***

Elizabeth Warren ***Boris Johnson***

company

Lehman Brothers ***Reuters***

Berkshire Hathaway ***Enron***

ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*Since its founding, **Bloomberg** has made several acquisitions.*

__Berkshire Hathaway__ Berkshire Hathaway Inc has acquired a 25 billion rupees ($356 million) stake in the parent of digital payments company Paytm.

Different contexts where ***"Berkshire Hathaway"*** is mentioned

politician

Richard Nixon ***Bill Clinton***

Elizabeth Warren ***Boris Johnson***

company

Lehman Brothers ***Reuters***

Berkshire Hathaway ***Enron***

ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*Since its founding, **Bloomberg** has made several acquisitions.*

Context-consistent

Berkshire Hathaway *Berkshire Hathaway Inc has acquired a 25 billion rupees ($356 million) stake in the parent of digital payments company Paytm.*

Different contexts where ***"Berkshire Hathaway"*** is mentioned

politician

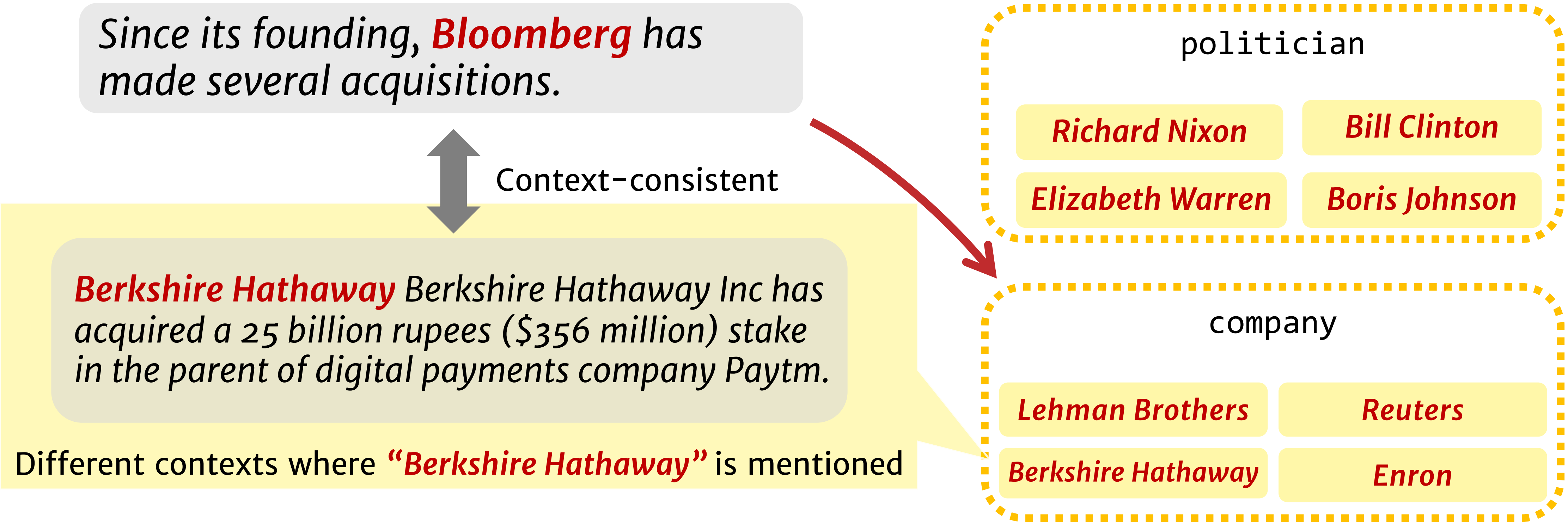
Richard Nixon ***Bill Clinton***
Elizabeth Warren ***Boris Johnson***

company

Lehman Brothers ***Reuters***
Berkshire Hathaway ***Enron***

ZOE: Type-Compatible Grounding

- “Type” as conceptual container binding entities together.



ZOE: Type-Compatible Grounding

- "Type" as conceptual container binding entities together.

*Since its founding, **Bloomberg** has made several acquisitions.*

Context-consistent

***Berkshire Hathaway** Berkshire Hathaway Inc has acquired a 25 billion rupees ($356 million) stake in the parent of digital payments company Paytm.*

Different contexts wh

politician

Richard Nixon

Bill Clinton

Elizabeth Warren

Boris Johnson

company

Lehman Brothers

Reuters

ay

Enron

Context consistency allows us to determine good candidates for **type compatibility.**

Zero-Shot Open Typing: Big Picture

A mention & its context

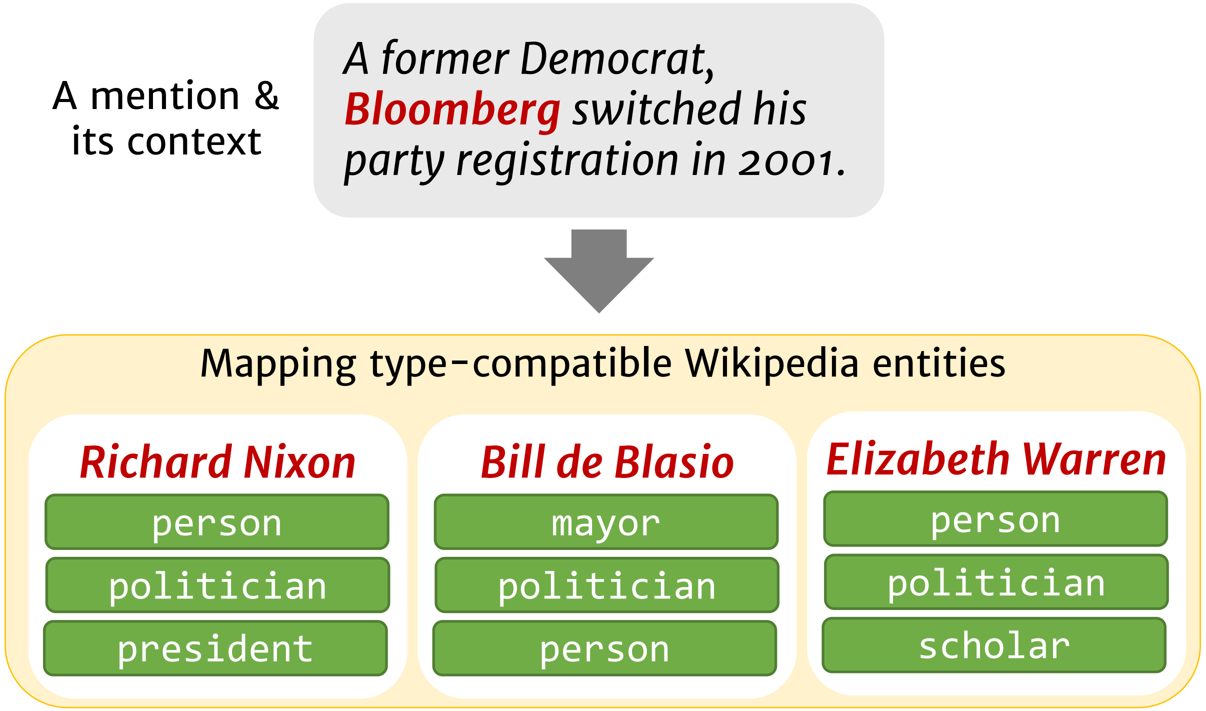
*A former Democrat, **Bloomberg** switched his party registration in 2001.*

High-level Algorithm:

1. Map the mention to **context-consistent** Wikipedia concepts

2. Rank candidate titles by **context-consistency** and infer the types according to the **type taxonomy**.

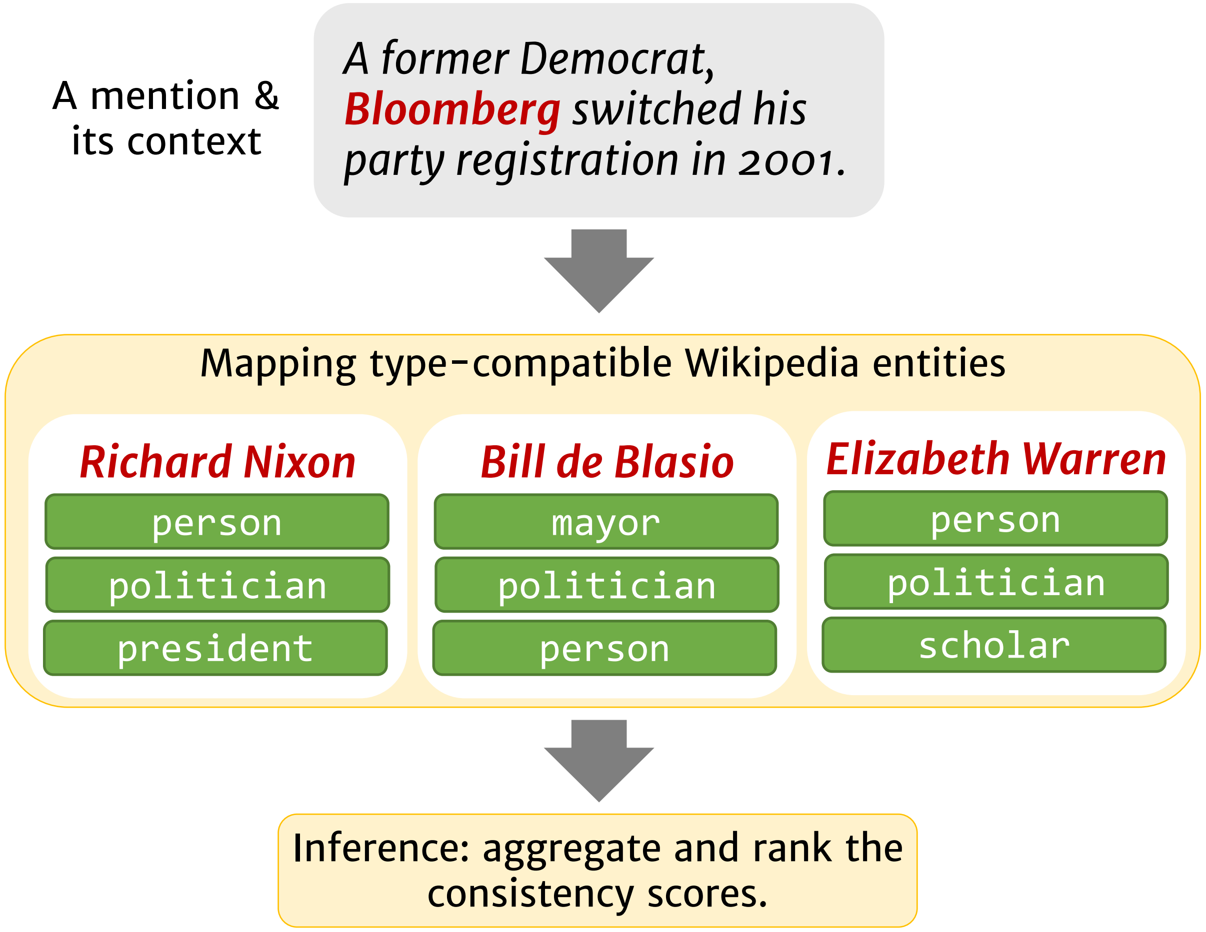
Zero-Shot Open Typing: Big Picture



High-level Algorithm:

1. Map the mention to **context-consistent** Wikipedia concepts
2. Rank candidate titles by **context-consistency** and infer the types according to the **type taxonomy**.

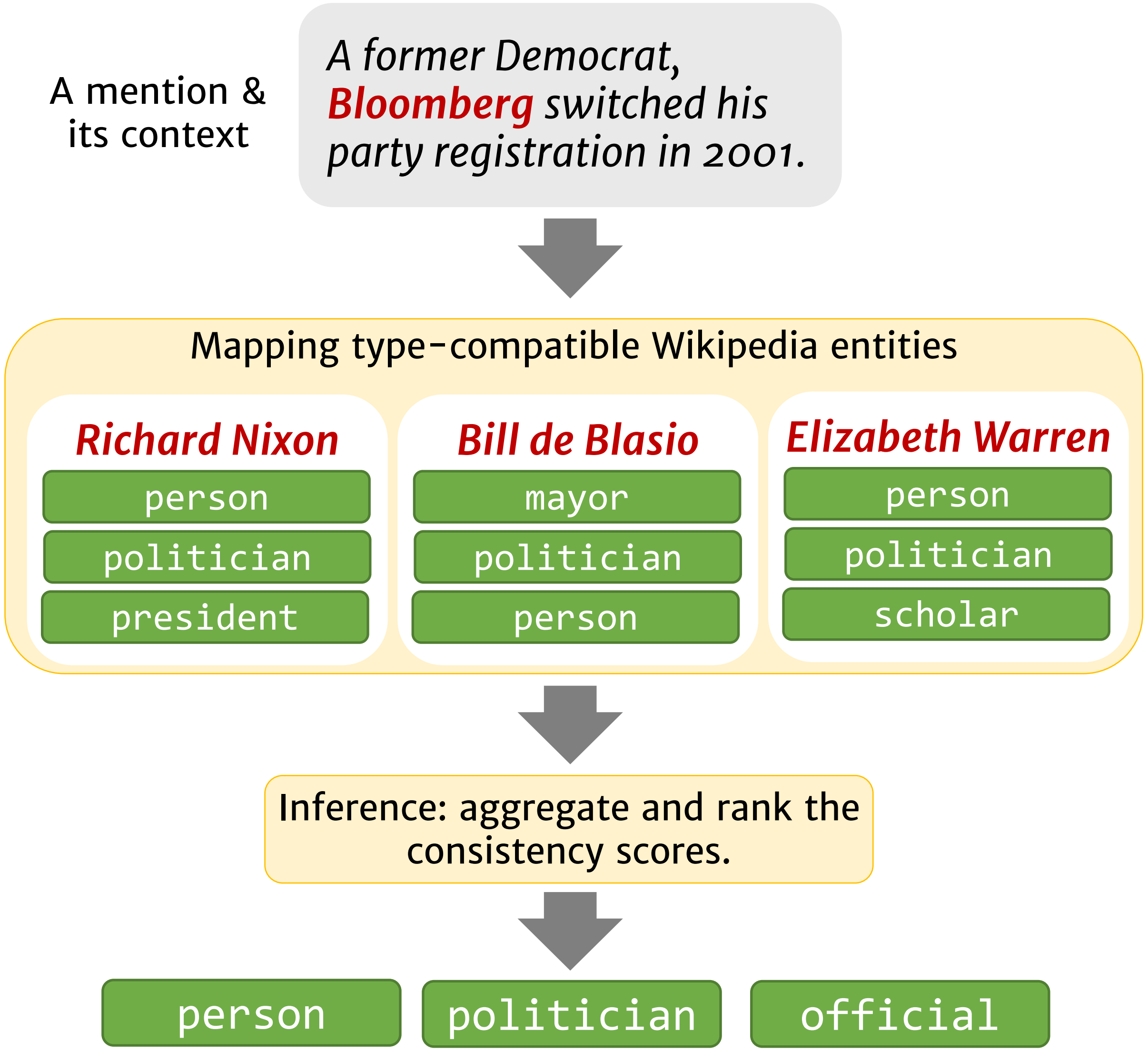
Zero-Shot Open Typing: Big Picture



High-level Algorithm:

1. Map the mention to **context-consistent** Wikipedia concepts
2. Rank candidate titles by **context-consistency** and infer the types according to the **type taxonomy**.

Zero-Shot Open Typing: Big Picture



High-level Algorithm:

1. Map the mention to **context-consistent** Wikipedia concepts
2. Rank candidate titles by **context-consistency** and infer the types according to the **type taxonomy**.

Zero-Shot Open Typing: Big Picture

A mention & its context

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Mapping type-compatible Wikipedia entities

Richard Nixon	***Bill de Blasio***	***Elizabeth Warren***
person	mayor	person
politician	politician	politician
president	person	scholar

Inference: aggregate and rank the consistency scores.

person | politician | official

Resources

Zero-Shot Open Typing: Big Picture

A mention & its context

*A former Democrat, **Bloomberg** switched his party registration in 2001.*

Mapping type-compatible Wikipedia entities

Richard Nixon	*Bill de Blasio*	*Elizabeth Warren*
person	mayor	person
politician	politician	politician
president	person	scholar

Inference: aggregate and rank the consistency scores.

person politician official

Resources

WIKIPEDIA

Freebase

[Bollacker et al. 08]

WikiLinks

[Singh et al. 12]

Contextualized Representations

[Peters et al.18]

Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.

- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	BBN	Ontonotes

Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.

- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	BBN	Ontonotes
AFET [Ren et al. 16]	FIGER			
NFETC [Xu&Barbosa 18]	FIGER			

Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.

- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	BBN	Ontonotes
AFET [Ren et al. 16]	**FIGER**	66		
NFETC [Xu&Barbosa 18]	**FIGER**	79		

Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.
- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	BBN	Ontonotes
AFET [Ren et al. 16]	**FIGER**	66	?	?
NFETC [Xu&Barbosa 18]	**FIGER**	79	?	?

Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.
- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	**BBN**	**Ontonotes**
AFET [Ren et al. 16]	**FIGER**	66	?	?
NFETC [Xu&Barbosa 18]	**FIGER**	79	?	?
AFET [Ren et al. 16]	**BBN**	?	75	?
AAA [Abishek et al. 17]	**BBN**	?	79	?

Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.
- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	BBN	Ontonotes
AFET [Ren et al. 16]	**FIGER**	66	?	?
NFETC [Xu&Barbosa 18]	**FIGER**	79	?	?
AFET [Ren et al. 16]	**BBN**	?	75	?
AAA [Abishek et al. 17]	**BBN**	?	79	?
AFET [Ren et al. 16]	**Ontonotes**	?	?	65
NFETC [Xu&Barbosa 18]	**Ontonotes**	?	?	70

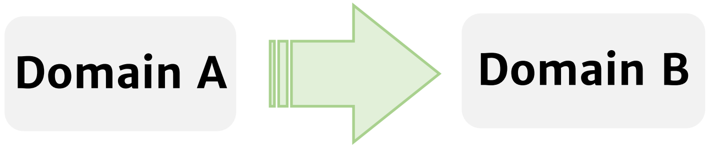
Empirical Results: Fine-Typing [ZKTR18]

- Outperforms supervised system in cross-domain.
- Comparable results with supervised systems.

System	Trained on	Evaluated on		
		FIGER	BBN	Ontonotes
AFET [Ren et al. 16]	**FIGER**	66	?	?
NFETC [Xu&Barbosa 18]	**FIGER**	79	?	?
AFET [Ren et al. 16]	**BBN**	?	75	?
AAA [Abishek et al. 17]	**BBN**	?	79	?
AFET [Ren et al. 16]	**Ontonotes**	?	?	65
NFETC [Xu&Barbosa 18]	**Ontonotes**	?	?	70
ZOE (this work)	-	71	75	61

Lessons

- Reformulating the task and using weak signals helps us reduce our dependence on direct "supervision".



- This type-aware approach leads to the ability to transfer across **domains & taxonomies.**

Beyond Supervision-rich “tasks”

- We will never have **enough annotated data** to train all the models for **all** the tasks.
    - Annotation for complex tasks is difficult, costly and sometimes impossible.

- We don’t even know what are “all the tasks”.

Beyond Supervision-rich “tasks”

- Two samples of research projects in an attempt to utilize hints in data to infer supervision signals:
  - Representation
  - Structure

- Not just two systems:
  - Initial steps towards a broader theory of using “incidental” signals.

[Roth, AAAI’17]

BIG PICTURE + LOOK AHEAD

Machine Learning,
Optimization &
applications

Natural Language Processing

Machine Learning, Optimization & applications

KSKCSSR. StartAI'18

KKCMSR. COLING'16

QK. NourIPS'15

KNJF. TIP'14

NKTNJ. SMC'11

Natural Language Processing

Machine Learning, Optimization & applications

KSKCSSR. StartAI'18
KKCMSR. COLING'16
QK. NourIPS'15
KNJF. TIP'14
NKTNJ. SMC'11

Natural Language Processing

Semantics

Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.

ZKCR. EMNLP'18
KCRUR. NAACL'18
FKPWR. Cognitum'15
PKR. NAACL'15

Machine Learning, Optimization & applications

KSKCSSR. StartAI'18

KKCMSR. COLING'16

QK. NourIPS'15

KNJF. TIP'14

NKTNJ. SMC'11

Natural Language Processing

Semantics

Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.

ZKCR. EMNLP'18

KCRUR. NAACL'18

FKPWR. Cognitum'15

PKR. NAACL'15

Learning & Inference

Question Answering, Textual Entailment, etc.

KKSR. AAAI'18

KKSR. CoNLL'17

CEKSTTK. AAAI'16

KKSR. IJCAI'16

Machine Learning, Optimization & applications

KSKCSSR. StartAI'18
KKCMSR. COLING'16
QK. NourIPS'15
KNJF. TIP'14
NKTNJ. SMC'11

Natural Language Processing

Semantics

Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.

ZKCR. EMNLP'18
KCRUR. NAACL'18
FKPWR. Cognitum'15
PKR. NAACL'15

Learning & Inference

Question Answering, Textual Entailment, etc.

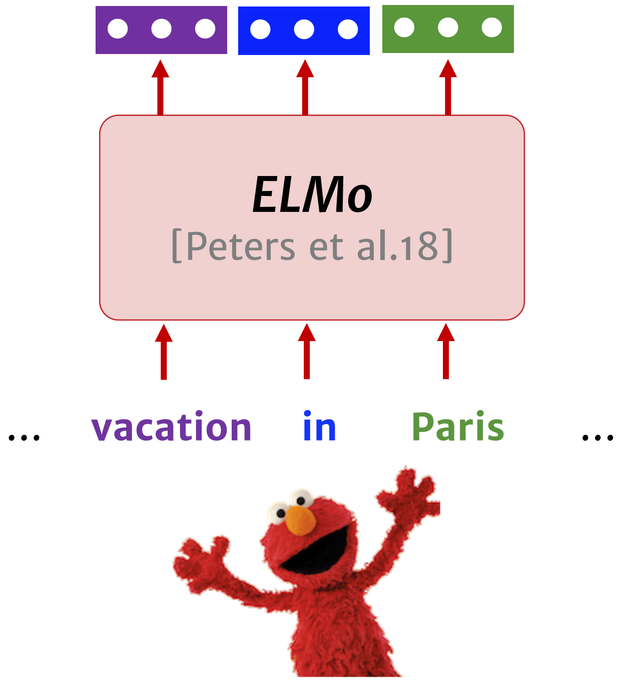
KKSR. AAAI'18
KKSR. CoNLL'17
CEKSTTK. AAAI'16
KKSR. IJCAI'16

NLP tools/software

K et al. LREC'18
SCKKSVBWR. LREC'16

Beyond Supervision-rich "tasks"

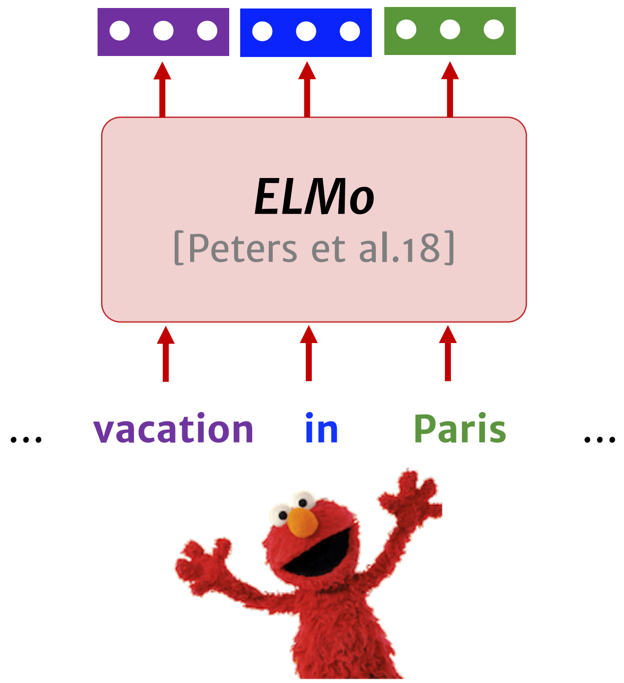
- A major shift in the field:
  - Being able to make use of massive loads of **unlabeled** data in the form of language models.
  - Compatible with the philosophy I advocated for here.



Other works:
[Devlin et al.18; Radford et al.18]

Language Models: Means to Access Knowledge

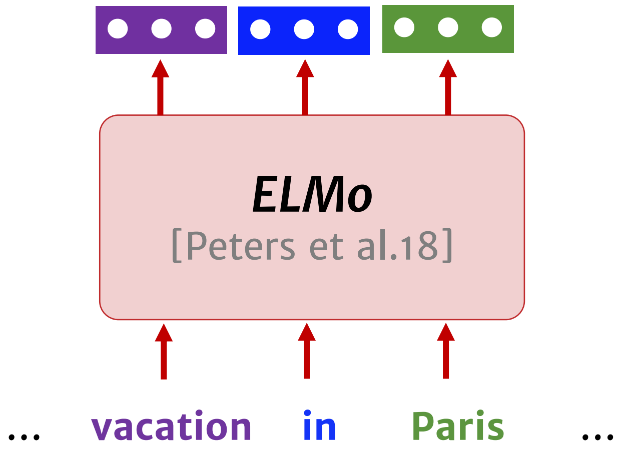
- They let you "query" for knowledge:



Other works:
[Devlin et al.18; Radford et al.18]

Language Models: Means to Access Knowledge

- They let you “query” for knowledge:

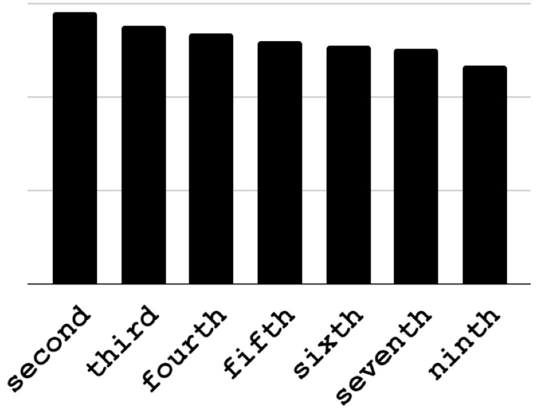


Other works:
[Devlin et al.18; Radford et al.18]

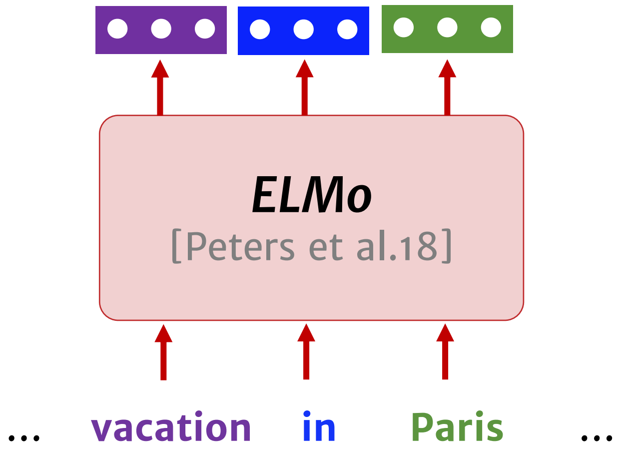
Pittsburgh is the _____ –largest populated city in Pennsylvania.

Language Models: Means to Access Knowledge

- They let you "query" for knowledge:



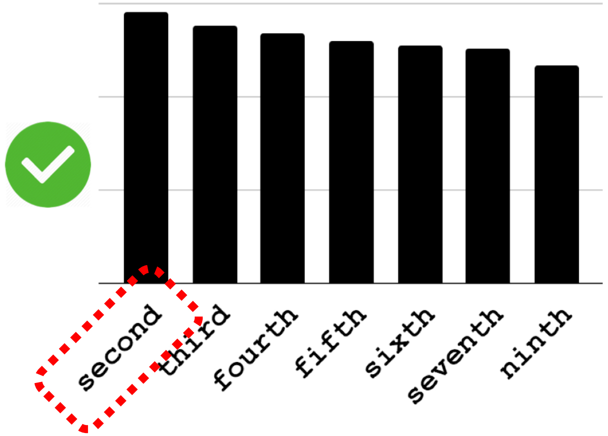
Pittsburgh is the _____ –largest populated city in Pennsylvania.



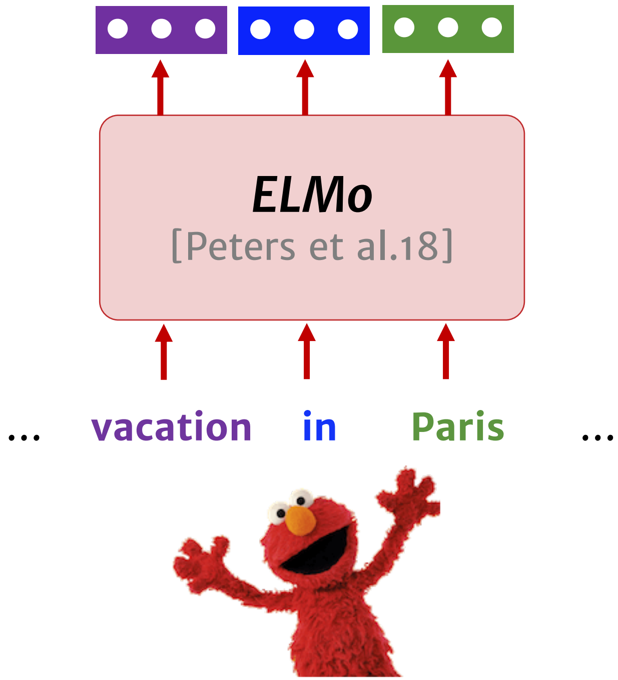
Other works:
[Devlin et al.18; Radford et al.18]

Language Models: Means to Access Knowledge

- They let you "query" for knowledge:

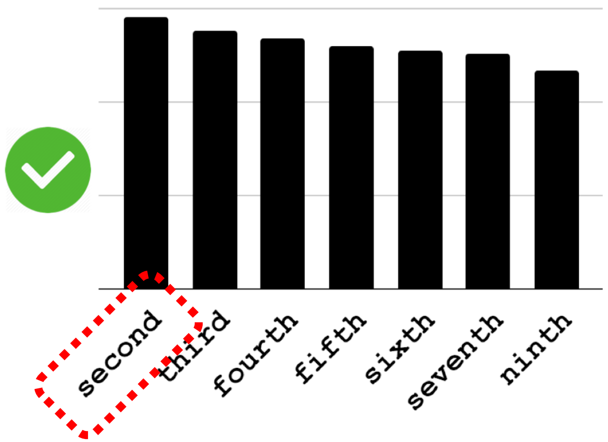


Pittsburgh is the _____ -largest populated city in Pennsylvania.



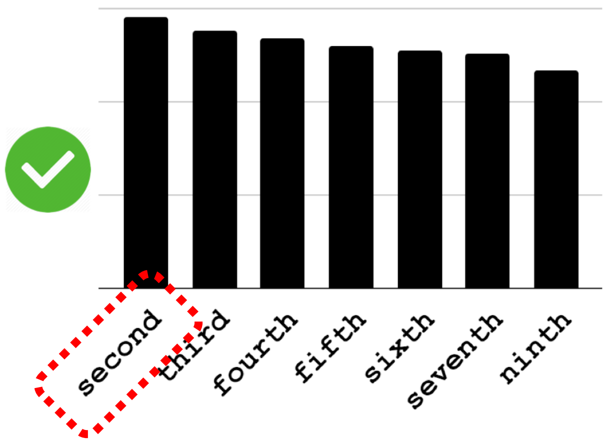
Other works:
[Devlin et al.18; Radford et al.18]

Language Models: Means to Access Knowledge



Pittsburgh is the _____ -largest populated city in Pennsylvania.

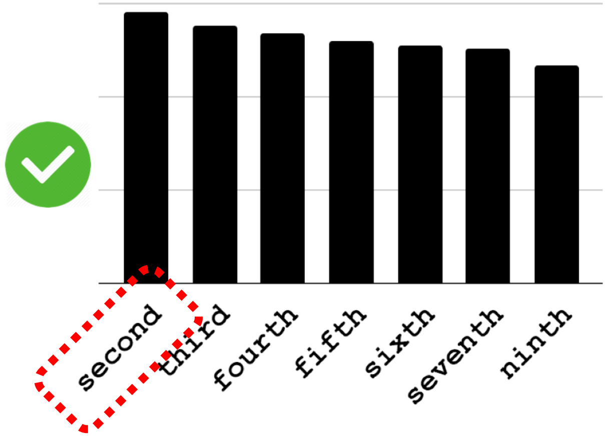
Language Models: Means to Access Knowledge



Pittsburgh is the _____ –largest populated city in Pennsylvania.

- **What is known:**
  - What is the nature the knowledge that they have internalized?
- **Know what you know:**
  - Is there a mechanism to decide whether something is [not]?
- **Inference with knowledge:**
  - Access what is known and be able to solve bigger problems.

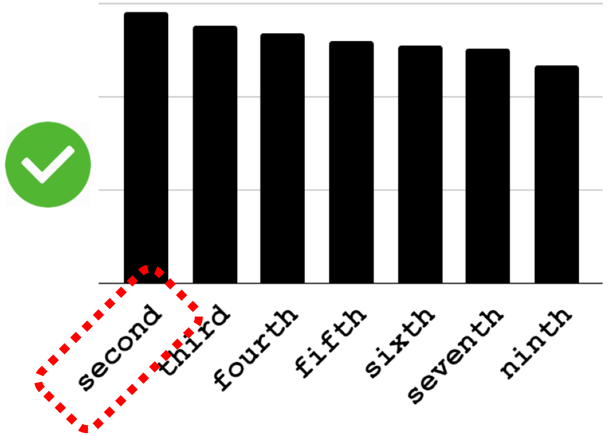
Language Models: Means to Access Knowledge



Pittsburgh is the _____ -largest populated city in Pennsylvania.

- What is known:
  - What is the nature the knowledge that they have internalized?
- Know what you know:
  - Is there a mechanism to decide whether something is [not]?
- Inference with knowledge:
  - Access what is known and be able to solve bigger problems.

Language Models: Means to Access Knowledge



Pittsburgh is the _____ -largest populated city in Pennsylvania.

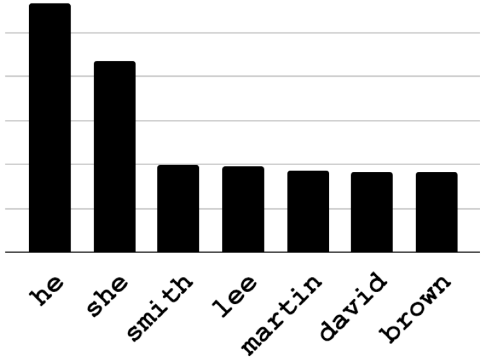
- **What is known:**
  - What is the nature the knowledge that they have internalized?
- **Know what you know:**
  - Is there a mechanism to decide whether something is [not]?
- **Inference with knowledge:**
  - Access what is known and be able to solve bigger problems.

Language Models: Biases

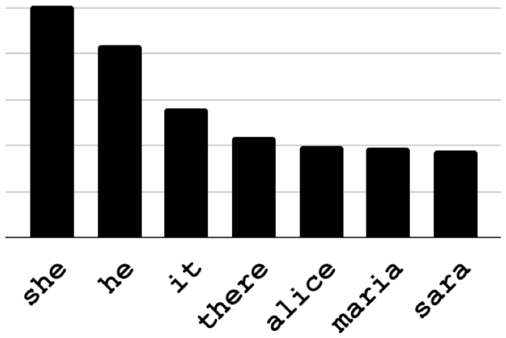
- What does this mean for the NLP systems built out of such systems?
- **Discovery:**
  - How can we automate the discovery of issues?
- **Mitigation:**
  - How can we resolve the such biases?

Language Models: Biases

____ is a lawyer.



____ is a babysitter.

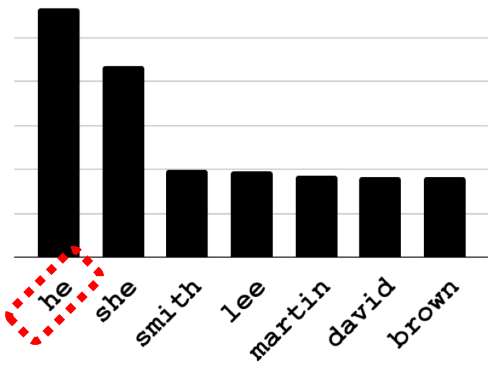


Bias
[May et al.19; Zhao et al.19]

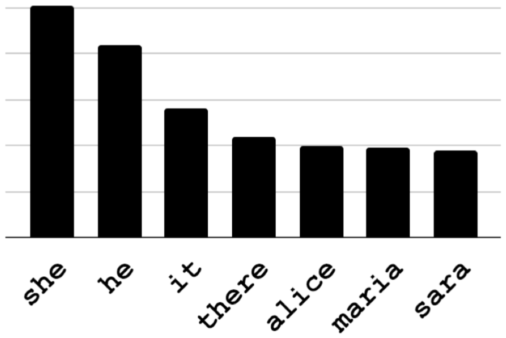
- What does this mean for the NLP systems built out of such systems?
- Discovery:
  - How can we automate the discovery of issues?
- Mitigation:
  - How can we resolve the such biases?

Language Models: Biases

____ *is a lawyer.*



____ *is a babysitter.*



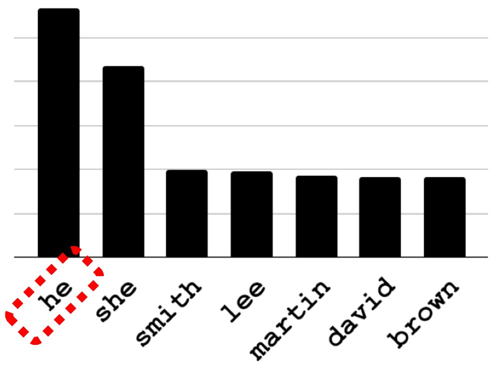
Bias

[May et al.19; Zhao et al.19]

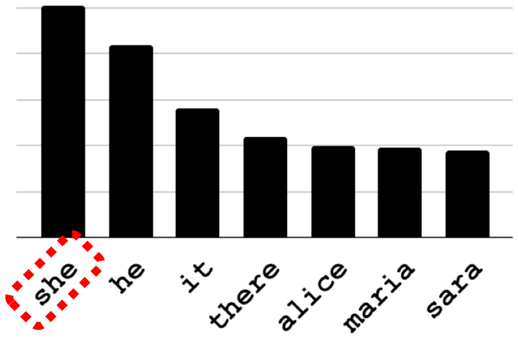
- What does this mean for the NLP systems built out of such systems?
- **Discovery:**
  - How can we automate the discovery of issues?
- **Mitigation:**
  - How can we resolve the such biases?

Language Models: Biases

____ *is a lawyer.*



____ *is a babysitter.*



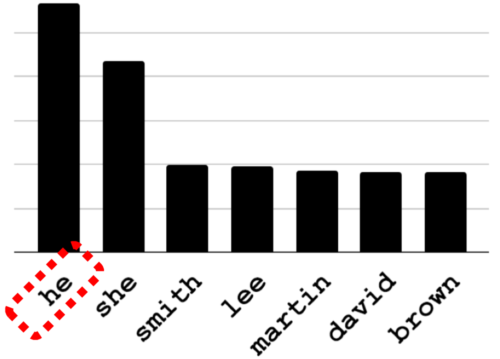
Bias

[May et al.19; Zhao et al.19]

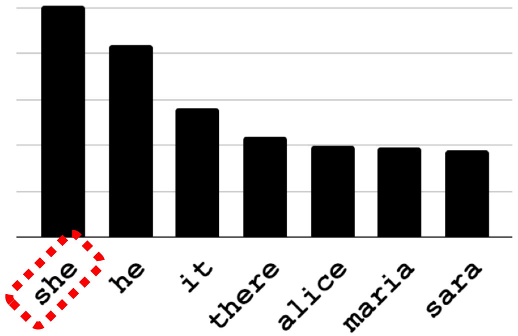
- What does this mean for the NLP systems built out of such systems?
- **Discovery:**
  - How can we automate the discovery of issues?
- **Mitigation:**
  - How can we resolve the such biases?

Language Models: Biases

____ is a lawyer.



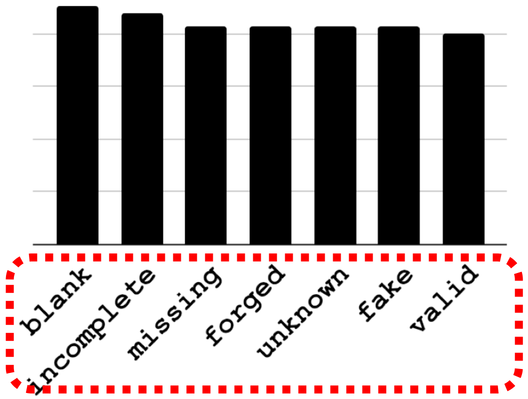
____ is a babysitter.



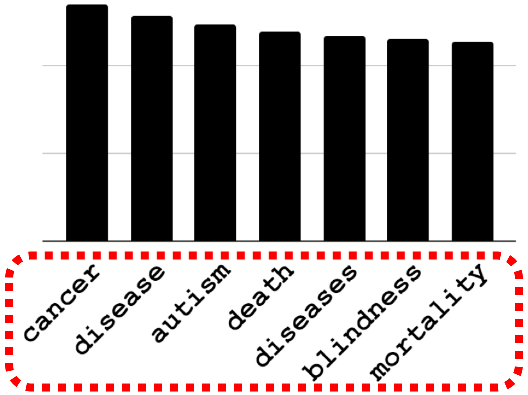
Bias

[May et al.19; Zhao et al.19]

Obama's birth certificate is ____.



Vaccines cause ____.



Conspiracy Theories

- What does this mean for the NLP systems built out of such systems?
- **Discovery:**
  - How can we automate the discovery of issues?
- **Mitigation:**
  - How can we resolve the such biases?

Machine Learning, Optimization & applications

KSKCSSR. StartAI'18
KKCMSR. COLING'16
QK. NourIPS'15
KNJF. TIP'14
NKTNJ. SMC'11

Natural Language Processing

Semantics

Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.

ZKCR. EMNLP'18
KCRUR. NAACL'18
FKPWR. Cognitum'15
PKR. NAACL'15

Learning & Inference

Question Answering, Textual Entailment, etc.

KKSR. AAAI'18
KKSR. CoNLL'17
CEKSTTK. AAAI'16
KKSR. IJCAI'16

NLP tools/software

K et al. LREC'18
SCKKSVBWR. LREC'16

Machine Learning, Optimization & applications

KS**K**CSSR. StartAI’18
K**K**CMSR. COLING’16
Q**K**. NourIPS’15
KNJF. TIP’14
N**K**TNJ. SMC’11

Natural Language Processing

Semantics

Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.

Z**K**CR. EMNLP’18
KCRUR. NAACL’18
F**K**PWR. Cognitum’15
P**K**R. NAACL’15

Learning & Inference

Question Answering, Textual Entailment, etc.

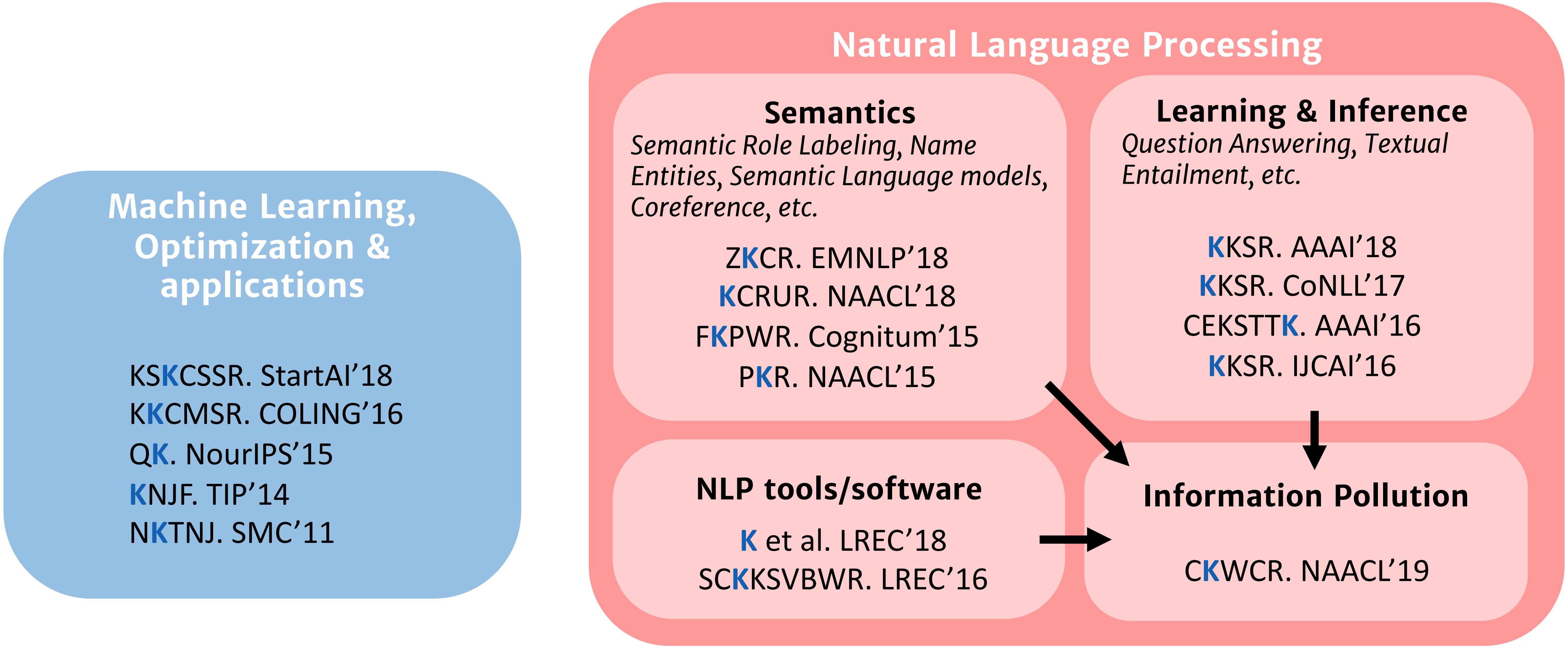
KKSR. AAAI’18
KKSR. CoNLL’17
CEKSTT**K**. AAAI’16
KKSR. IJCAI’16

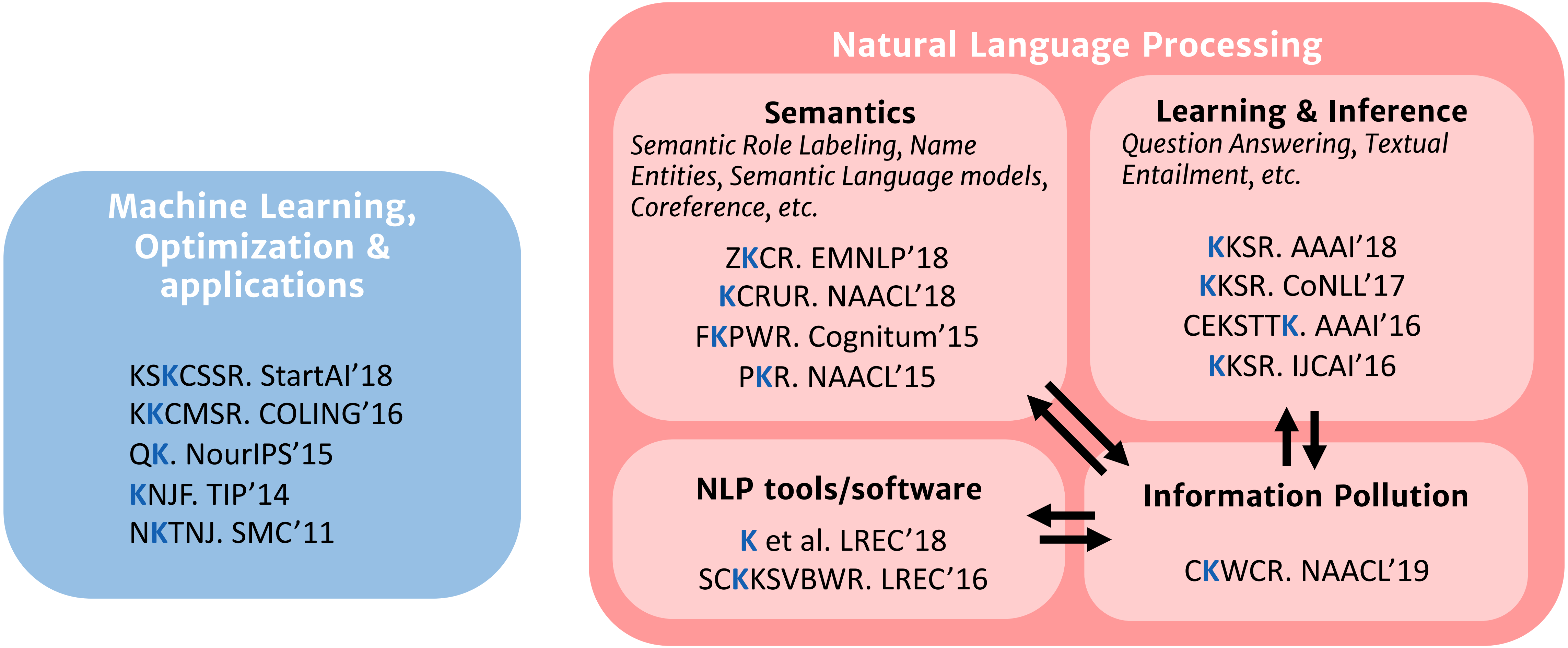
NLP tools/software

K et al. LREC’18
SC**K**KSVBWR. LREC’16

Information Pollution

C**K**WCR. NAACL’19


Machine Learning, Optimization & applications
KSKCSSR. StartAI'18
KKCMSR. COLING'16
QK. NourIPS'15
KNJF. TIP'14
NKTNJ. SMC'11
Natural Language Processing
Semantics
Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.
ZKCR. EMNLP'18
KCRUR. NAACL'18
FKPWR. Cognitum'15
PKR. NAACL'15
Learning & Inference
Question Answering, Textual Entailment, etc.
KKSR. AAAI'18
KKSR. CoNLL'17
CEKSTTK. AAAI'16
KKSR. IJCAI'16
NLP tools/software
K et al. LREC'18
SCKKSVBWR. LREC'16
Information Pollution
CKWCR. NAACL'19


Machine Learning, Optimization & applications
KSKCSSR. StartAI'18
KKCMSR. COLING'16
QK. NourIPS'15
KNJF. TIP'14
NKTNJ. SMC'11
Natural Language Processing
Semantics
Semantic Role Labeling, Name Entities, Semantic Language models, Coreference, etc.
ZKCR. EMNLP'18
KCRUR. NAACL'18
FKPWR. Cognitum'15
PKR. NAACL'15
Learning & Inference
Question Answering, Textual Entailment, etc.
KKSR. AAAI'18
KKSR. CoNLL'17
CEKSTTK. AAAI'16
KKSR. IJCAI'16
NLP tools/software
K et al. LREC'18
SCKKSVBWR. LREC'16
Information Pollution
CKWCR. NAACL'19

Information Pollution

Information Pollution

- Information Technology started with much optimism:
  - Democratizing information and greater liberties.
- Few foresaw the huge radical impact of the information revolution.
  - Massive amount of Information pollution:

Information Pollution

- Information Technology started with much optimism:
  - Democratizing information and greater liberties.
- Few foresaw the huge radical impact of the information revolution.
  - Massive amount of Information pollution:

> "The contamination of the information supply with irrelevant, redundant, unsolicited, incorrect, and otherwise low-value information."

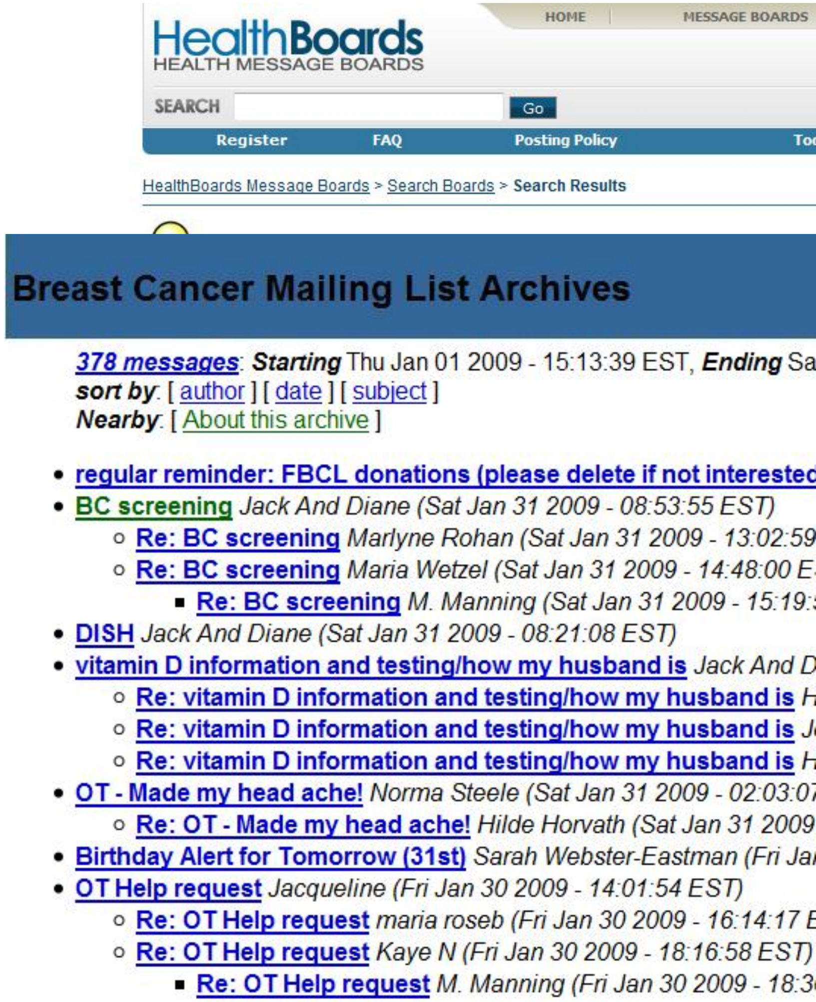
[Levent Orman'15]

Information Pollution: Not Just Politics

- Medical Domain, Education, Public Policy, etc.
  - *"Best treatment for X;" "Side effects of X ."*

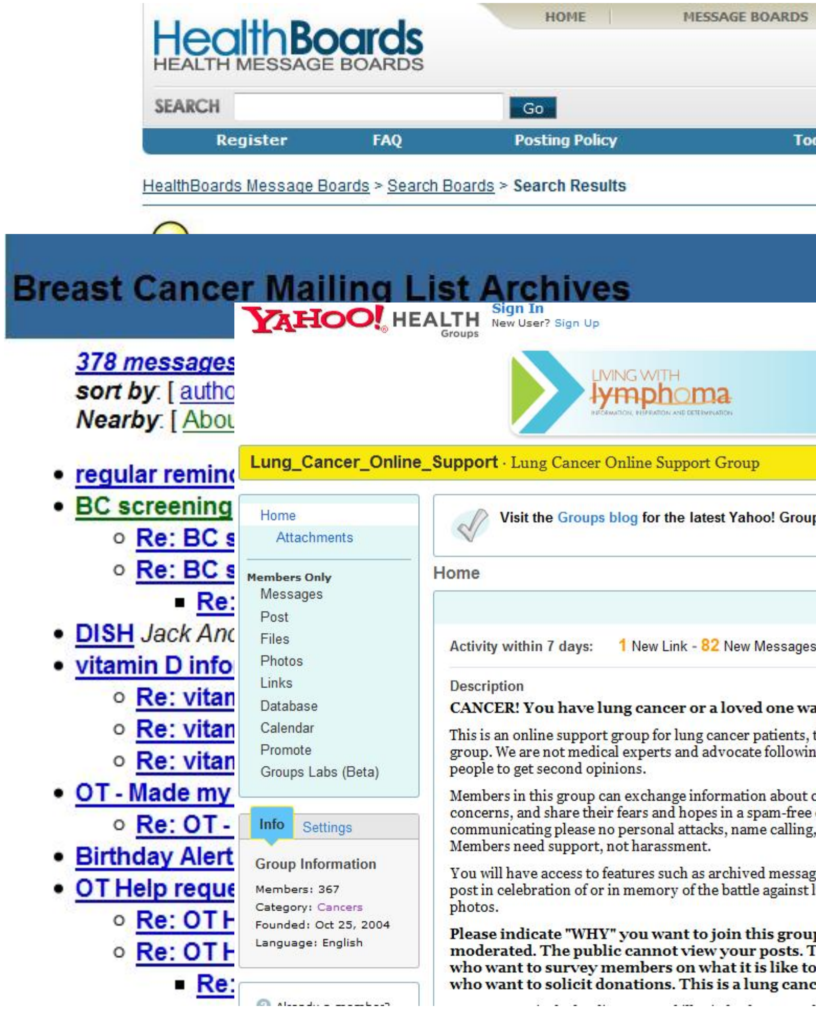
Information Pollution: Not Just Politics

- Medical Domain, Education, Public Policy, etc.
    - *“Best treatment for X;” “Side effects of X .”*



Information Pollution: Not Just Politics

- Medical Domain, Education, Public Policy, etc.
  - *“Best treatment for X;” “Side effects of X .”*



Information Pollution: Not Just Politics

- Medical Domain, Education, Public Policy, etc.
  - *“Best treatment for X;” “Side effects of X .”*

Information Pollution: Not Just Politics

- Medical Domain, Education, Public Policy, etc.
  - *“Best treatment for X;” “Side effects of X .”*

- Are they consistent?
- Are they trustworthy?
- Are they written by someone with an agenda?

Information Pollution: Not Just Fact-Checking

Information Pollution: Not Just Fact-Checking

- Many issues don't have a single "answer."
  - "*Should X be legalized?*"
    - Possible answers are subject to situations, world views or background.
    - Moral, utilitarian, libertarian, philosophy, etc.

Information Pollution: Not Just Fact-Checking

- Many issues don't have a single "answer."
  - *"Should X be legalized?"*
    - Possible answers are subject to situations, world views or background.
    - Moral, utilitarian, libertarian, philosophy, etc.

Information Pollution: Not Just Fact-Checking

- Many issues don't have a single "answer."
  - *"Should X be legalized?"*
    - Possible answers are subject to situations, world views or background.
    - Moral, utilitarian, libertarian, philosophy, etc.

Information Pollution: Not Just Fact-Checking

- Many issues don't have a single "answer."
  - *"Should X be legalized?"*
    - Possible answers are subject to situations, world views or background.
    - Moral, utilitarian, libertarian, philosophy, etc.

Information Pollution: Not Just Fact-Checking

- Many issues don't have a single "answer."
  - *"Should X be legalized?"*
    - Possible answers are subject to situations, world views or background.
    - Moral, utilitarian, libertarian, philosophy, etc.

Factual information (or lack of) is **not** really the core of the problem.

Information Pollution, as NLU Problems

Not only applications for NLP, but also drive the research in important directions.

Information Pollution, as NLU Problems

- **Understanding Sources**
- But **what should we believe**, and who should we trust?
- Sources may
  - Have their own, often hidden, motivations
  - Make different or even contradictory claims

Not only applications for NLP, but also drive the research in important directions.

Information Pollution, as NLU Problems

- **Understanding Sources**
- But **what should we believe**, and who should we trust?
- Sources may
  - Have their own, often hidden, motivations
  - Make different or even contradictory claims

- **Understanding the evidence**
- Sources may present different, but legitimate, perspectives
- Most interesting issues/questions have multiple "right" answers
  - Perspectives must be supported by evidence

Not only applications for NLP, but also drive the research in important directions.

Information Pollution, as NLU Problems

- **Understanding Sources**
- But **what should we believe**, and who should we trust?
- Sources may
  - Have their own, often hidden, motivations
  - Make different or even contradictory claims

- **Understanding the evidence**
- Sources may present different, but legitimate, perspectives
- Most interesting issues/questions have multiple "right" answers
  - Perspectives must be supported by evidence

Not only applications for NLP, but also drive the research in important directions.

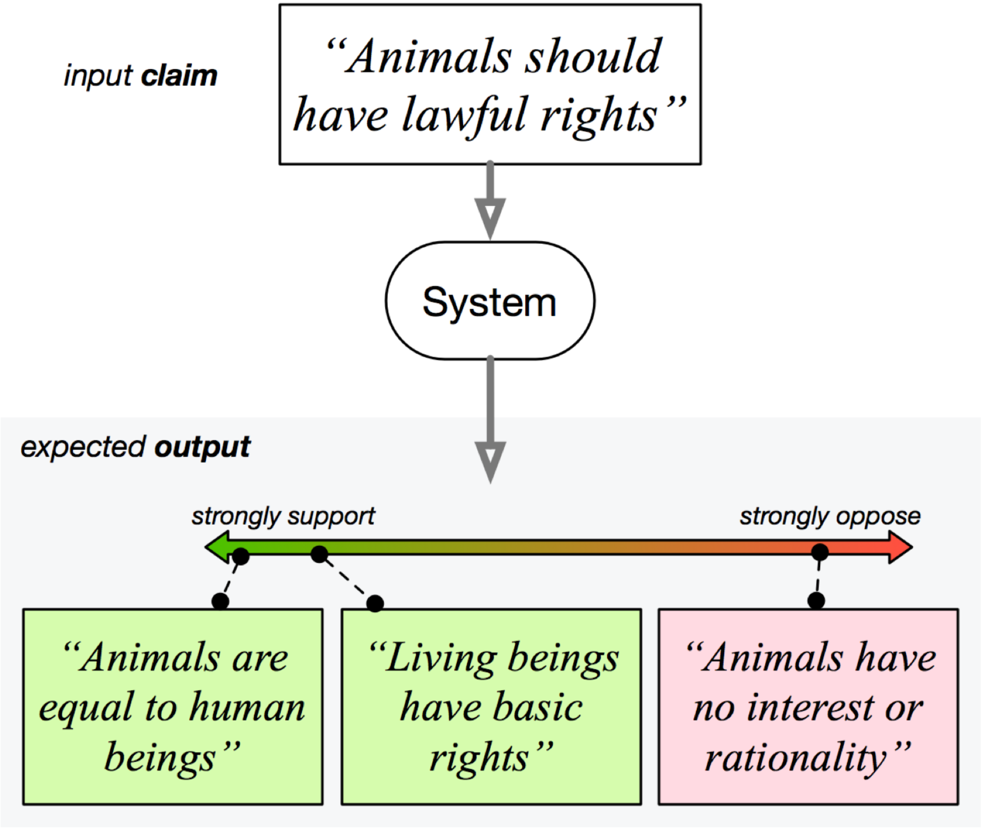
Information Pollution, as NLU Problems

- **Understanding Sources**
- But **what should we believe**, and who should we trust?
- Sources may
  - Have their own, often hidden, motivations
  - Make different or even contradictory claims

- **Understanding the evidence**
- Sources may present different, but legitimate, perspectives
- Most interesting issues/questions have multiple “right” answers
  - Perspectives must be supported by evidence

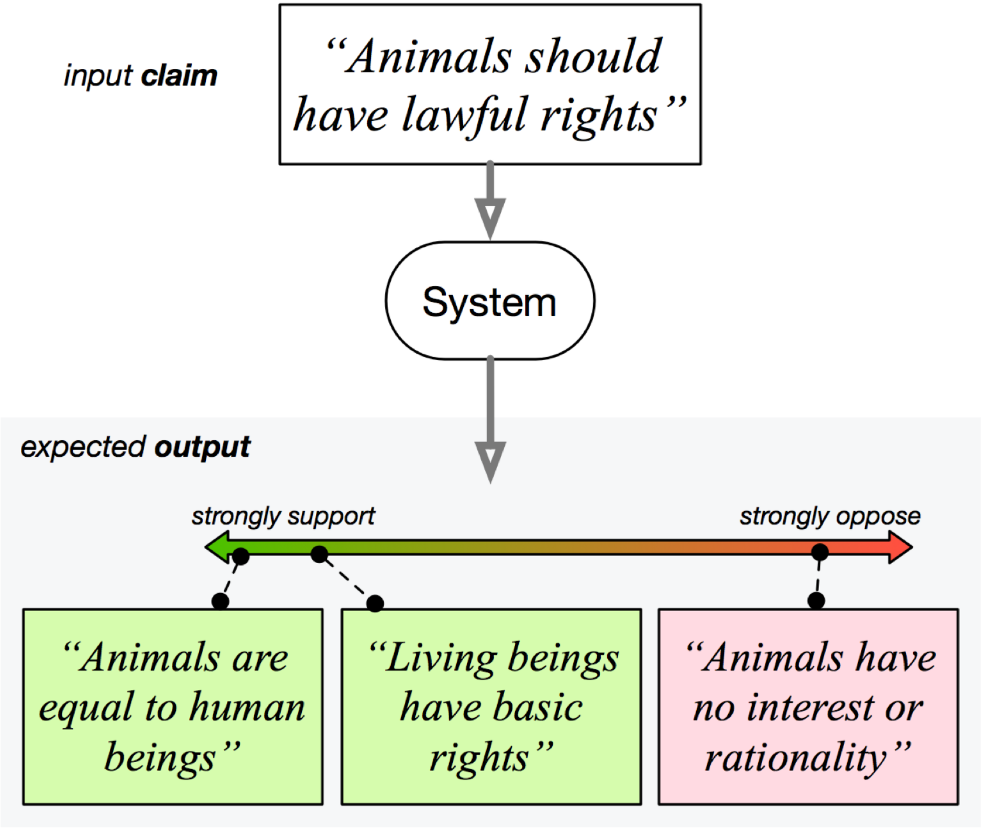
Not only applications for NLP, but also drive the research in important directions.

Discovering Diverse "Perspectives"



[Chen, **K**, et al. NAACL'19]

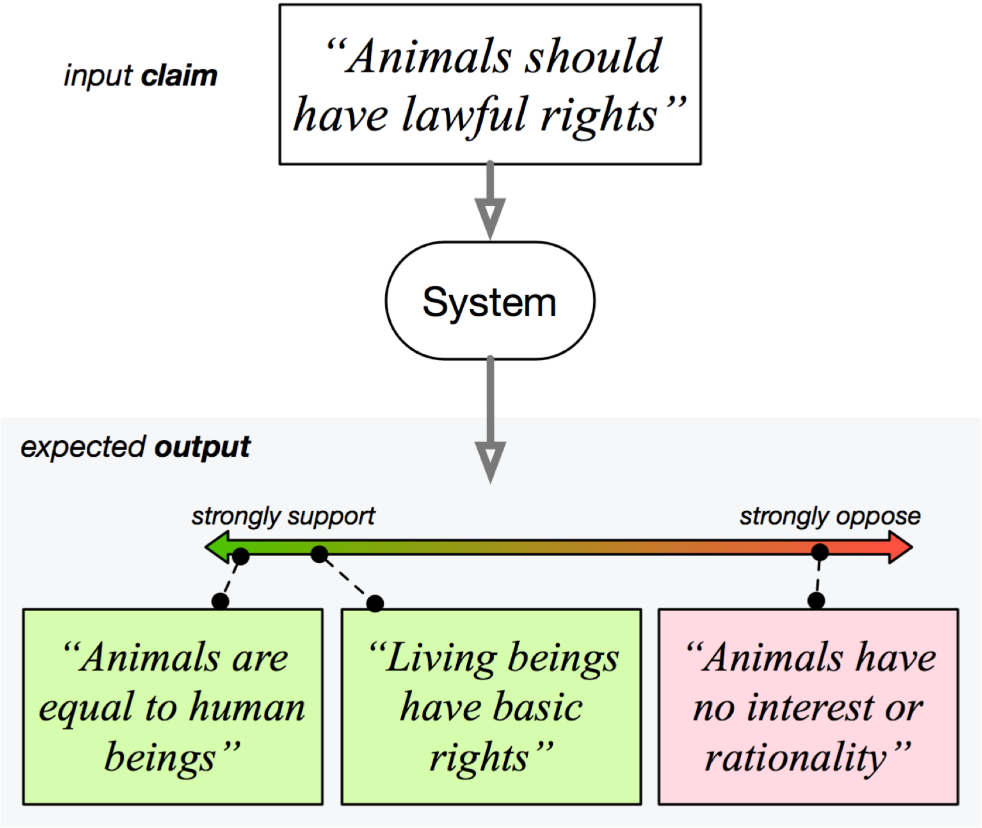
Discovering Diverse "Perspectives"



[Chen, K, et al. NAACL'19]

- **Our recent work:** provide users with the understanding that each "story" has more than one "perspective."
- **Goal:**
  - Perspectives could give a fuller understanding of an issue.
  - Make us more open-minded, less afraid & more likely to consider other views.

Discovering Diverse “Perspectives”



[Chen, K, et al. NAACL’19]

- **Our recent work:** provide users with the understanding that each “story” has more than one “perspective.”
- **Goal:**
    - Perspectives could give a fuller understanding of an issue.
    - Make us more open-minded, less afraid & more likely to consider other views.

Information Pollution: an NLU Challenge

- Suffering from this pollution is not a forgone conclusion.
- A computational model that will help us navigate the polluted world.
  - Natural Language Processing/Understanding + Algorithmic Components
  - Collaborative efforts involving experts from the social sciences, policy, and others.
- Overreliance on fully annotated data, unlikely to solve the problem.
- Interesting challenge, important, and will have societal impact.

Information Pollution: an NLU Challenge

- Suffering from this pollution is not a forgone conclusion.
- A computational model that will help us navigate the polluted world.
    - Natural Language Processing/Understanding + Algorithmic Components
    - Collaborative efforts involving experts from the social sciences, policy, and others.
- Overreliance on fully annotated data, unlikely to solve the problem.
- Interesting challenge, important, and will have societal impact.

Information Pollution: an NLU Challenge

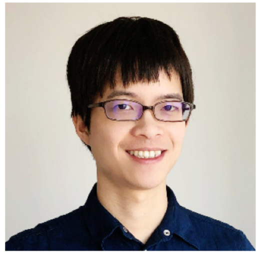
- Suffering from this pollution is not a forgone conclusion.
- A computational model that will help us navigate the polluted world.
  - Natural Language Processing/Understanding + Algorithmic Components
  - Collaborative efforts involving experts from the social sciences, policy, and others.
- Overreliance on fully annotated data, unlikely to solve the problem.
- Interesting challenge, important, and will have societal impact.

Tushar Khot
(AI2)

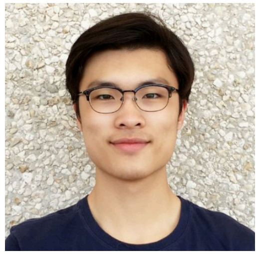
Dan Roth
(UPenn)

Ashish Sabharwal
(AI2)

Peter Clark
(AI2)



Chen-Tse Tsai
(Bloomberg)



Ben Zhou
(UIUC → UPenn)

That’s it, folks!