Compositional Semantic Parsing on Semi-Structured Tables

Ice Pasupat and Percy Liang
Stanford University

ACL 2015
Tuesday, July 28, 2015

Task

Question answering given a knowledge source

In which city was Ada Lovelace born?

Database

Semantic Parsing

Parse questions into executable logical forms

In which city was Ada Lovelace born?

Type.City ⊓ PeopleBornHere.AdaLovelace

(Lambda DCS)

Database

Semantic Parsing

Logical forms can be executed on the knowledge source to get denotations

Type.City ⊓ PeopleBornHere.AdaLovelace

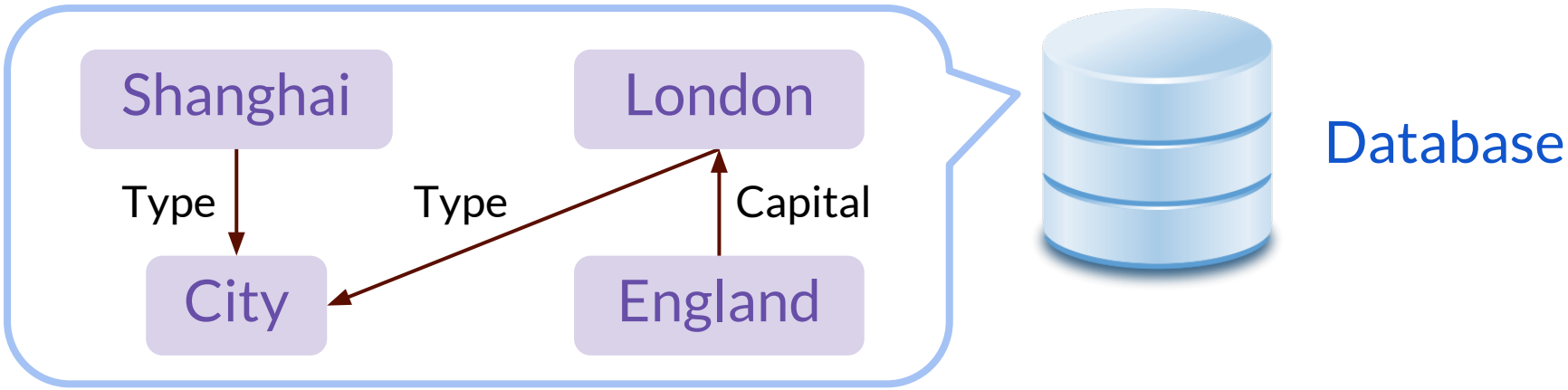
City

Database

Semantic Parsing

Logical forms can be executed on the knowledge source to get denotations

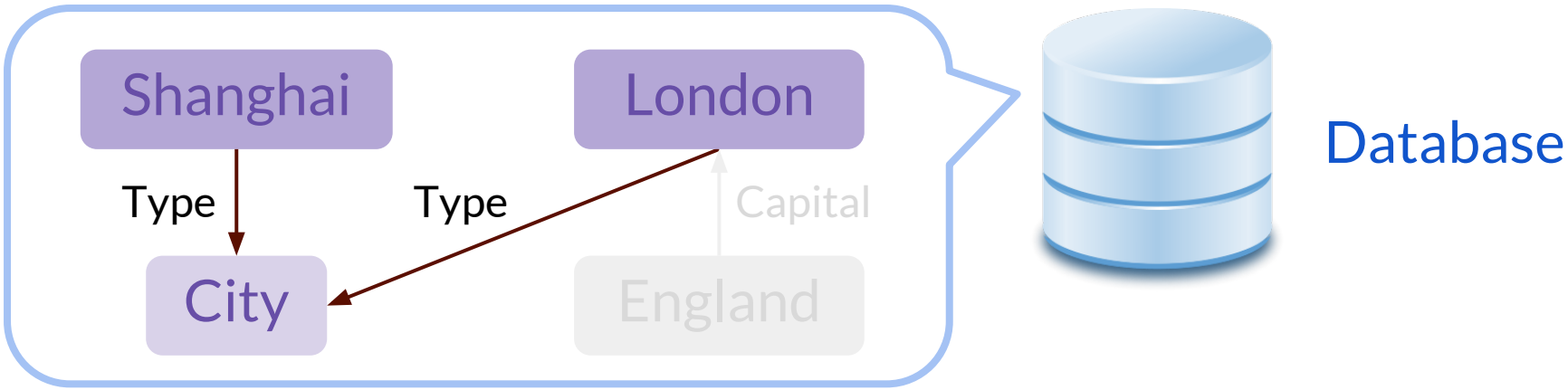
Type.City ⊓ PeopleBornHere.AdaLovelace



Semantic Parsing

Logical forms can be executed on the knowledge source to get denotations

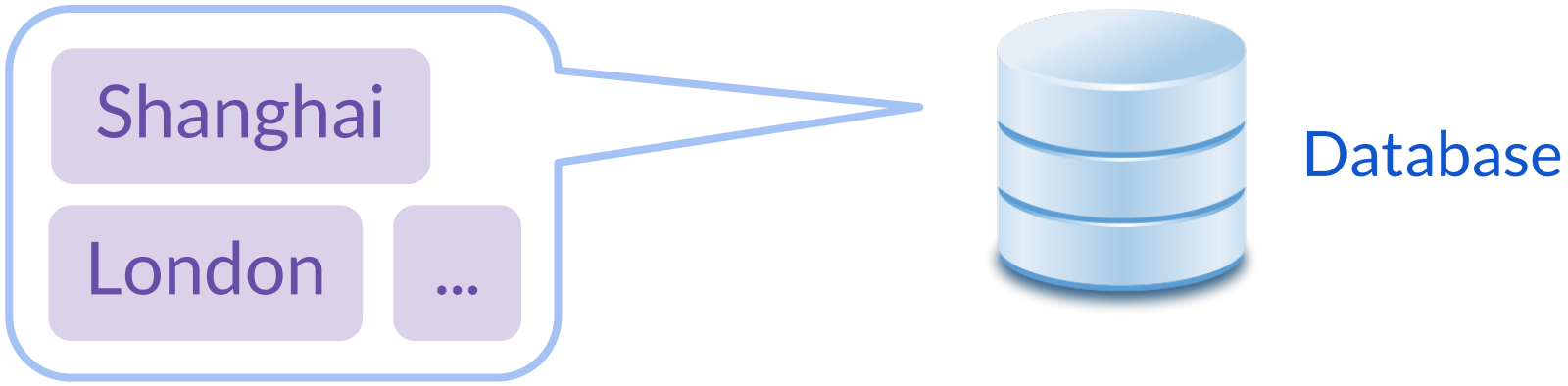
Type.City ⊓ PeopleBornHere.AdaLovelace



Semantic Parsing

Logical forms can be executed on the knowledge source to get denotations

Type.City ⊓ PeopleBornHere.AdaLovelace



Compositionality

We can compose logical forms into bigger ones with logical operations

Type.City ⊓ PeopleBornHere.AdaLovelace

Compositionality

We can compose logical forms into bigger ones with logical operations

Intersection

Type.City ⊓ PeopleBornHere.AdaLovelace

Compositionality

We can compose logical forms into bigger ones with logical operations

Intersection

Type.City ⊓ PeopleBornHere.AdaLovelace

London

Compositionality

We can compose logical forms into bigger ones with logical operations

- Type.City ⊔ Type.State — cities and / or states
- count(Type.City) — how many cities
- argmax(Type.City, Area) — largest city
- sum(AreaOf.Type.City) — total area of all cities
- AreaOf.London – AreaOf.Paris — how much bigger is London than Paris?

Related Work

Early systems: Parse very compositional questions into database queries

How many rivers are in the state with the largest population?

```
answer(A,
  count(B,
    (river(B), loc(B, C),
     largest(D, (state(C), population(C, D)))),
    A)))
```

Compositionality: High

[Zelle + Mooney, 1996 / Wong + Mooney, 2007 / Zettlemoyer + Collins, 2007 / Kwiatkowski et al., 2011 / ...]

Related Work

Early systems: Parse very compositional questions into database queries

How many rivers are in the state with the largest population?

```
answer(A,
  count(B,
    (river(B), loc(B, C),
     largest(D, (state(C), population(C, D)))),
    A)))
```

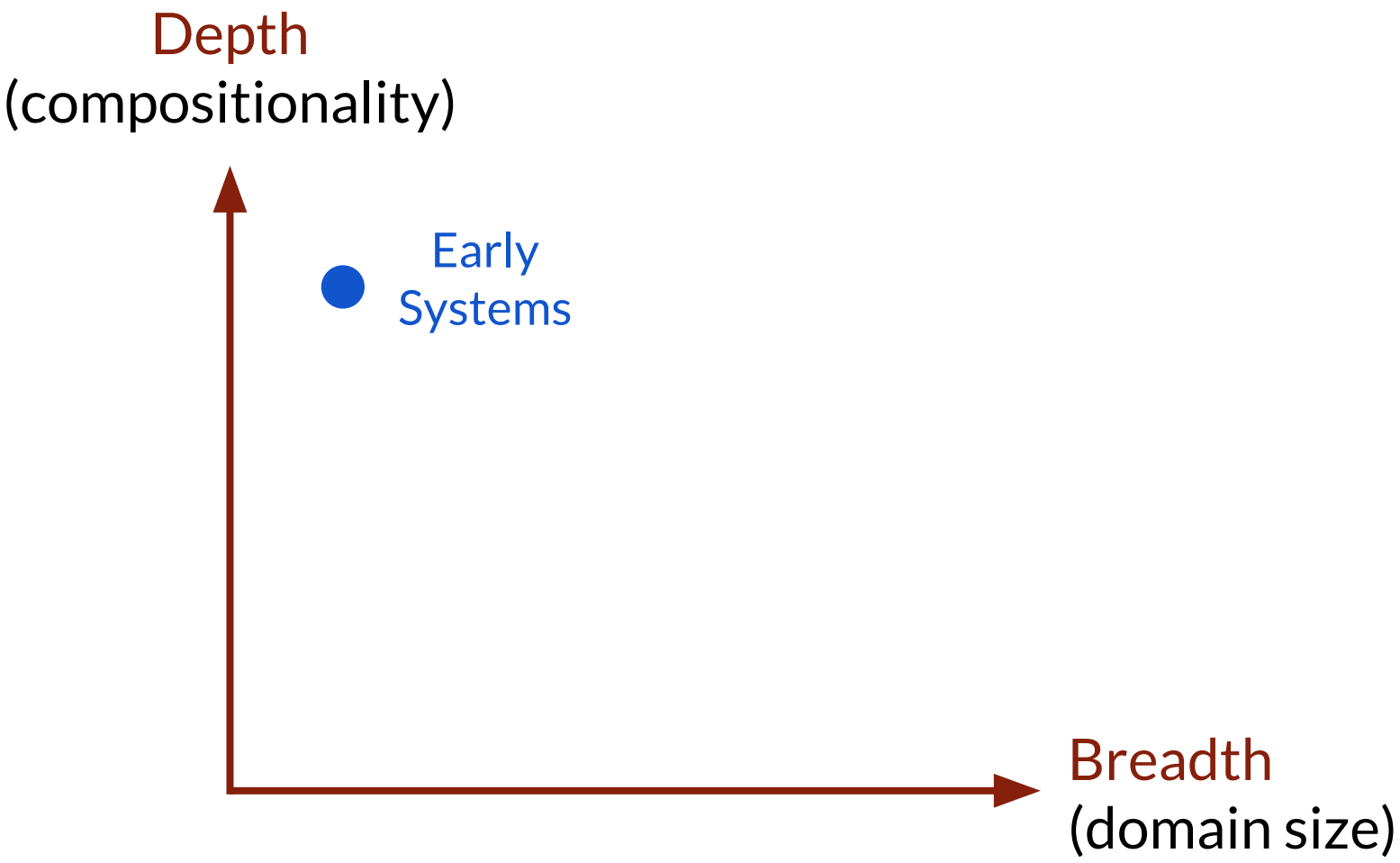
Compositionality: High

Knowledge source: Database

- few entities / relations
- fixed schema

[Zelle + Mooney, 1996 / Wong + Mooney, 2007 / Zettlemoyer + Collins, 2007 / Kwiatkowski et al., 2011 / ...]

Related Work



Related Work

Scaling to large knowledge bases (KBs): Answer open-domain questions using curated KBs

In which comic book issue did Kitty Pryde first appear?

Knowledge source: Large KBs

- lots of entities / relations
- fixed schema

[Cai + Yates, 2013 / Berant et al., 2013 + 2014 / Fader et al., 2014 / Reddy et al., 2014 / ...]

Related Work

Scaling to large knowledge bases (KBs): Answer open-domain questions using curated KBs

In which comic book issue did Kitty Pryde first appear?

R[FirstAppearance].KittyPryde

Compositionality: Lower

Knowledge source: Large KBs

- lots of entities / relations
- fixed schema

[Cai + Yates, 2013 / Berant et al., 2013 + 2014 / Fader et al., 2014 / Reddy et al., 2014 / ...]

Related Work

Scaling to large knowledge bases (KBs): Answer open-domain questions using curated KBs

In which comic book issue did Kitty Pryde first appear?

R[FirstAppearance].KittyPryde

Still, only < 10% of general questions can be answered by Freebase [Berant et al., 2013]

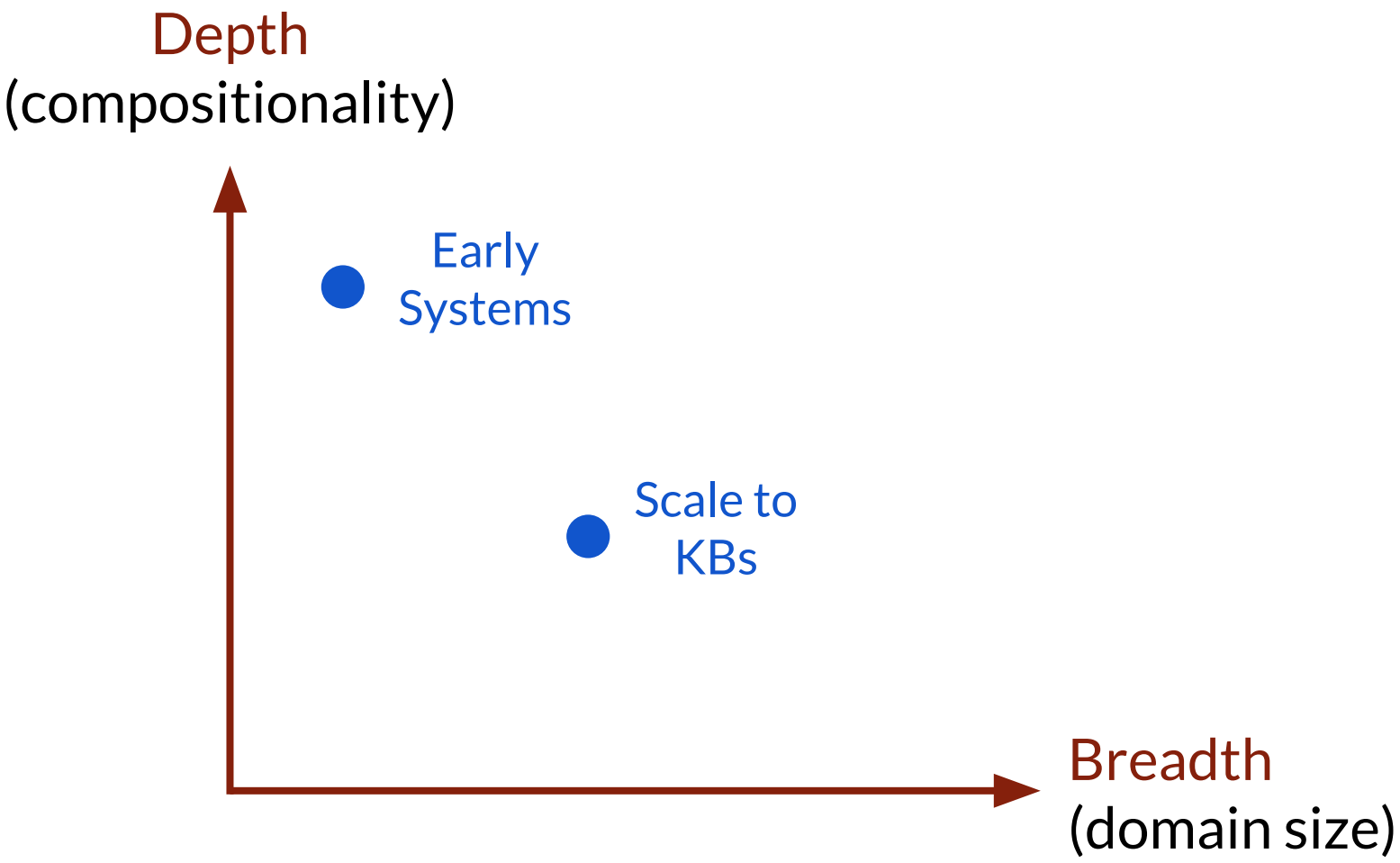
Compositionality: Lower

Knowledge source: Large KBs

- lots of entities / relations
- fixed schema

[Cai + Yates, 2013 / Berant et al., 2013 + 2014 / Fader et al., 2014 / Reddy et al., 2014 / ...]

Related Work



Related Work

Web search: Keyword search over the whole Web
(information retrieval / not semantic parsing)

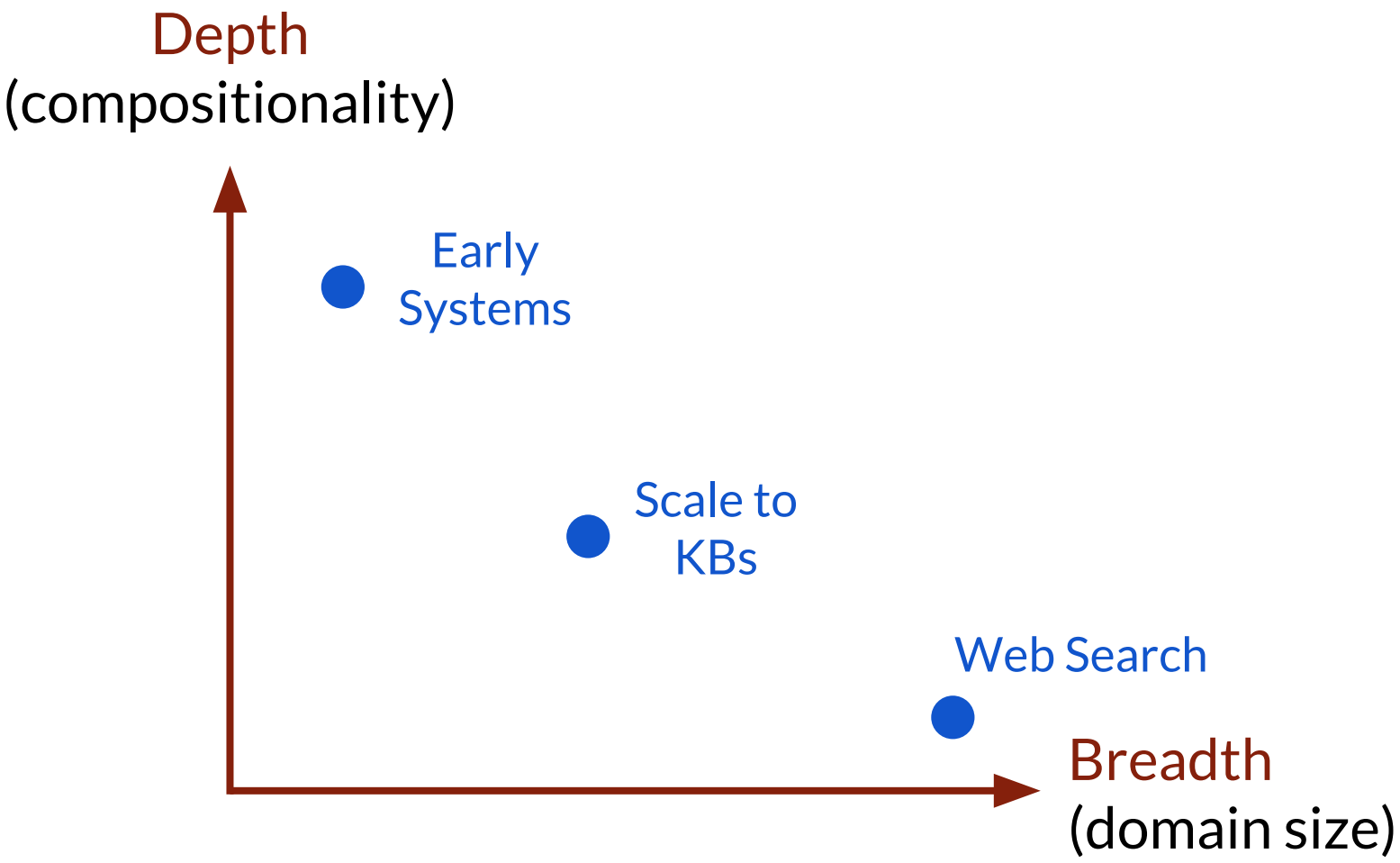
stanford cs professors

Compositionality: None

Knowledge source: Internet

- open-domain
- unstructured (no schema)

Related Work



Motivation

Web text in general is too unstructured

However, the Web also contains semi-structured data (tables, lists, repeated headings, ...)

stanford cs professors
http://cs.stanford.edu/faculty

Regular Faculty

57 people

Name	Phone	Office	email
Maneesh Agrawala		Gates 364	
Alex Aiken	5-3359	GATES 411	aiken
Peter Bailis			pbailis
Serafim Batzoglou	3-3334	Clark S266	serafim
Gill Bejerano	650 723-7666	Beckman B321	Click here
Michael Bernstein	4-1248	Gates 384	msb
Dan Boneh	5-3897	GATES 475	dabo
Moses Charikar		Gates 4B	
David Cheriton	3-1131	GATES 439	cheriton
Steve Cooper	723-9798	Gates 190	coopers
Bill Dally	5-8945	GATES 301	
David Dill	5-3642	GATES 344	dill
Ron Dror	497-8586	Gates 204	rondror
Dawson Engler	3-0762	GATES 314	engler
Stefano Ermon		Gates 158	stefano.ermon
Ron Fedkiw		GATES 207	
Hector Garcia-Molina	3-0685	GATES 434	hector
Mike Genesereth	3-0934	GATES 220	

Motivation

Web text in general is too unstructured

However, the Web also contains semi-structured data (tables, lists, repeated headings, ...)

- Open-domain: lots of information with arbitrary data schema [Cafarella et al., 2008 (WebTables)]

Motivation

Web text in general is too unstructured

However, the Web also contains semi-structured data (tables, lists, repeated headings, ...)

- Open-domain: lots of information with arbitrary data schema [Cafarella et al., 2008 (WebTables)]
- Structured enough to allow complex logical operations (~ mini knowledge base)

How many Stanford CS professors do not have offices in the Gates building?

Motivation

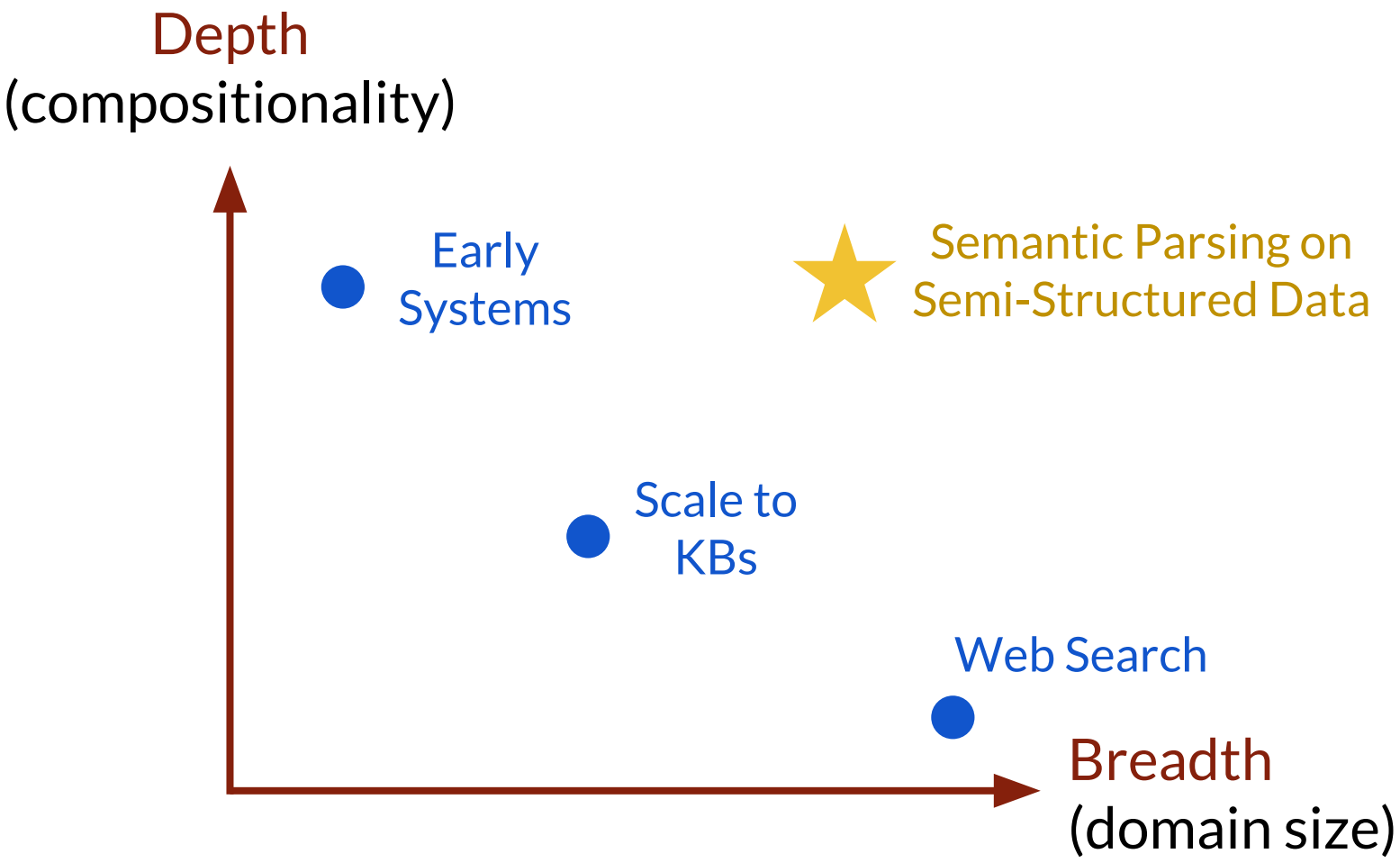
Web text in general is too unstructured

However, the Web also contains semi-structured data (tables, lists, repeated headings, ...)

- Open-domain: lots of information with arbitrary data schema [Cafarella et al., 2008 (WebTables)]
- Structured enough to allow complex logical operations (~ mini knowledge base)

Task: Answer compositional questions based on semi-structured tables from the Web

Motivation


Depth
(compositionality)
Early
Systems
Semantic Parsing on
Semi-Structured Data
Scale to
KBs
Web Search
Breadth
(domain size)

Outline

- Background and Related Work
- **Task and Dataset**
- Approach
- Experiments

Task Description

Input: utterance x and HTML table t

Output: answer y

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
...	...	...	...
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

x = Greece held its last Summer Olympics in which year?

y = 2004

Task Description

Input: utterance x and HTML table t

Output: answer y

Training data: list of (x, t, y) — no logical form

Tables in test data are **not seen** during training

- The model must generalize to unseen table schemas!

Dataset

WikiTableQuestions dataset:

- Tables t are from Wikipedia

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

https://en.wikipedia.org/wiki/Piotr_Kędzia

Dataset

WikiTableQuestions dataset:

- Tables t are from Wikipedia
- Questions x and answers y are from Mechanical Turk — Prompts are given to encourage compositionality

How many … … ___est … … last …

… above … … same … as … … difference …

… or … … his … Requires counting

etc.

Dataset

WikiTableQuestions dataset:

- Tables t are from Wikipedia
- Questions x and answers y are from Mechanical Turk — Prompts are given to encourage compositionality

Prompt: The question must contains "last" (or a synonym)

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

In what city did Piotr's last 1st place finish occur?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

How long did it take this competitor to finish the 4x400 meter relay at Universiade in 2005?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

Where was the competition held immediately before the one in Turkey?

Year	Competition	Venue	Position	Event	Notes
Representing Poland					
2001	World Youth Championships	Debrecen, Hungary	2nd	400 m	47.12
			1st	Medley relay	1:50.46
	European Junior Championships	Grosseto, Italy	1st	4x400 m relay	3:06.12
2003	European Junior Championships	Tampere, Finland	3rd	400 m	46.69
			2nd	4x400 m relay	3:08.62
2005	European U23 Championships	Erfurt, Germany	11th (sf)	400 m	46.62
			1st	4x400 m relay	3:04.41
	Universiade	Izmir, Turkey	7th	400 m	46.89
			1st	4x400 m relay	3:02.57
2006	World Indoor Championships	Moscow, Russia	2nd (h)	4x400 m relay	3:06.10
	European Championships	Gothenburg, Sweden	3rd	4x400 m relay	3:01.73
2007	European Indoor Championships	Birmingham, United Kingdom	3rd	4x400 m relay	3:08.14
	Universiade	Bangkok, Thailand	7th	400 m	46.85
			1st	4x400 m relay	3:02.05
2008	World Indoor Championships	Valencia, Spain	4th	4x400 m relay	3:08.76
	Olympic Games	Beijing, China	7th	4x400 m relay	3:00.32
2009	Universiade	Belgrade, Serbia	2nd	4x400 m relay	3:05.69

How many times has this competitor placed 5th or better in competition?

Dataset

WikiTableQuestions dataset:

- 2100 tables
  - Average: 6.3 columns / 27.5 rows
- 22000 examples

Challenges

With increased breadth (semi-structured data):

- Must generalize to arbitrary table schemas
  (as opposed to the fixed database schema)
- Test tables are **unseen** → Cannot precompute a lexicon mapping phrases to table relations

Table headers
(Year, Competition, Venue, …)

Challenges

With increased breadth (semi-structured data):

- Must generalize to arbitrary table schemas
  (as opposed to the fixed database schema)
- Test tables are **unseen** → Cannot precompute a lexicon mapping phrases to table relations

With increased depth (compositional questions):

- More operations and deeper recursion →
  Number of possible parses grows exponentially

Outline

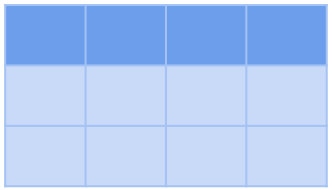
- Background and Related Work
- Task and Dataset
- **Approach**
- Experiments

Approach

Greece held its last Summer Olympics in which year?

x

t

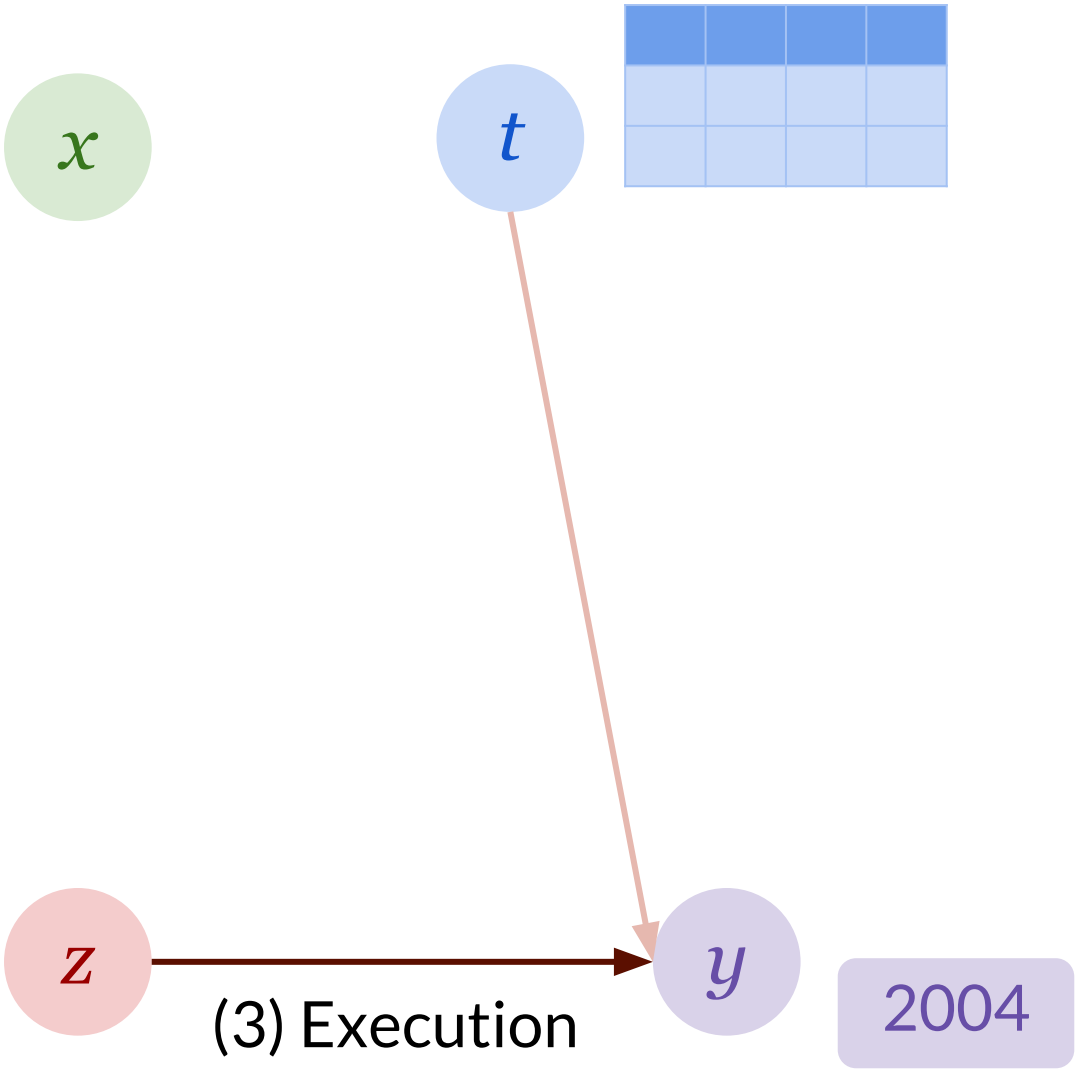


y

2004

Approach

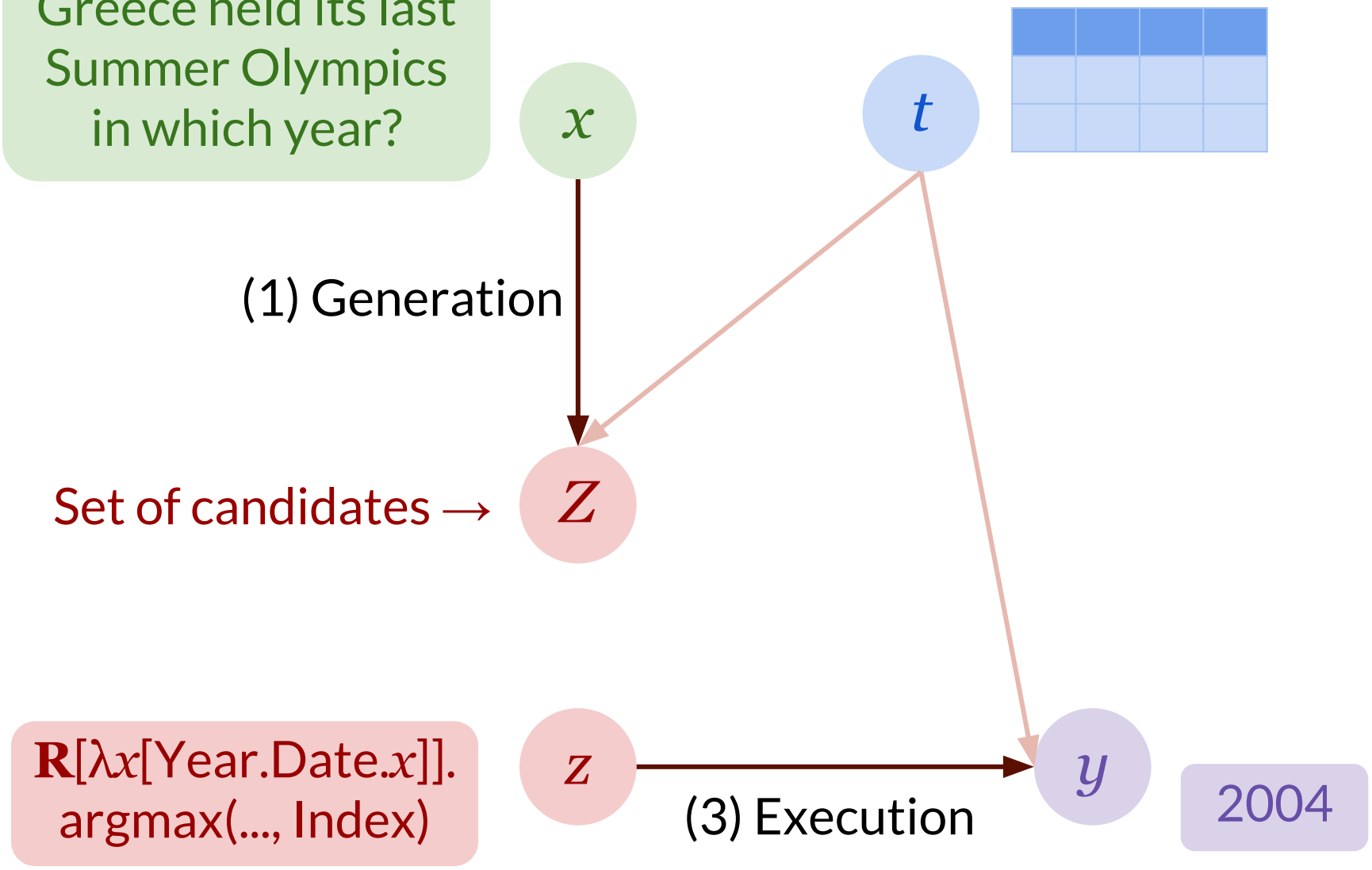
Greece held its last Summer Olympics in which year?



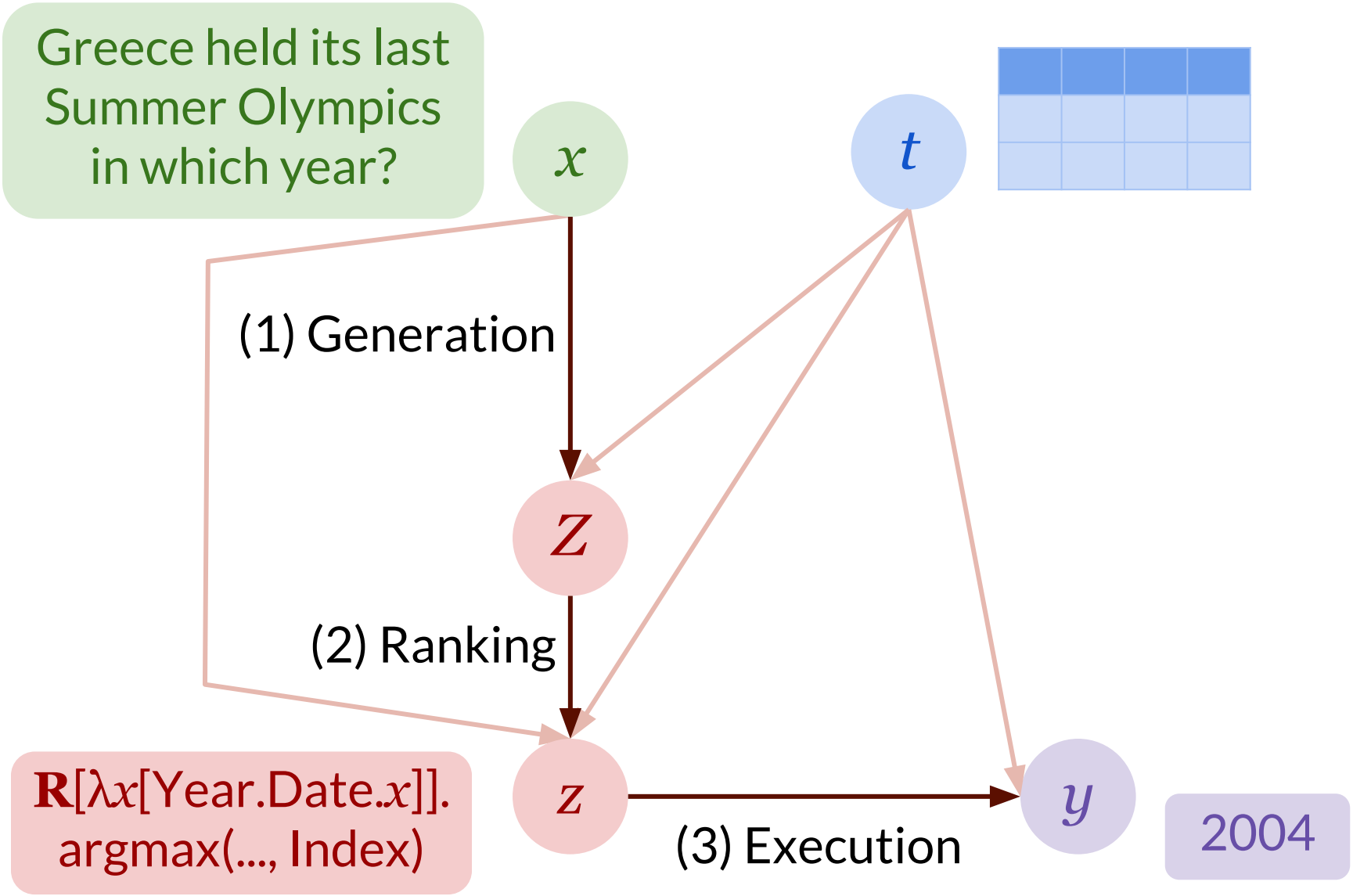
R[λ*x*[Year.Date.*x*]].
argmax(..., Index)

Approach

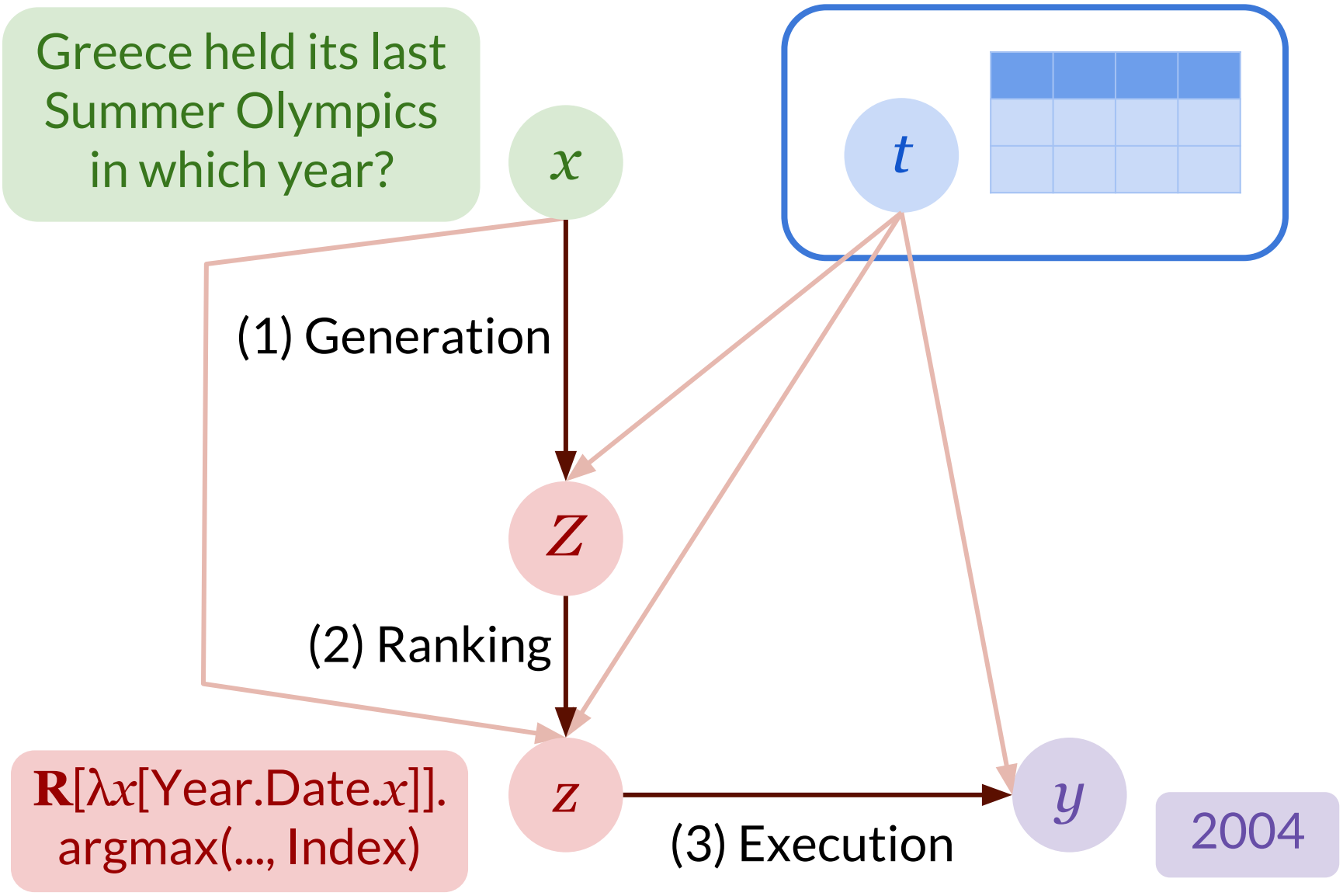
Greece held its last Summer Olympics in which year?



Approach


Greece held its last Summer Olympics in which year?
x
t
(1) Generation
z
(2) Ranking
R[λx[Year.Date.x]]. argmax(..., Index)
z
(3) Execution
y
2004

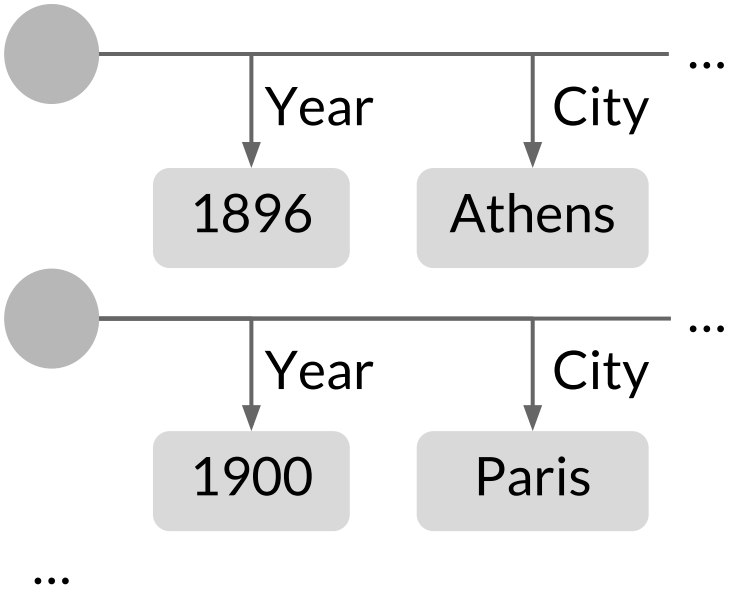
Approach


Greece held its last Summer Olympics in which year?
x
t
(1) Generation
z
(2) Ranking
R[λx[Year.Date.x]]. argmax(..., Index)
z
(3) Execution
y
2004

Representation

Convert table t to knowledge graph w

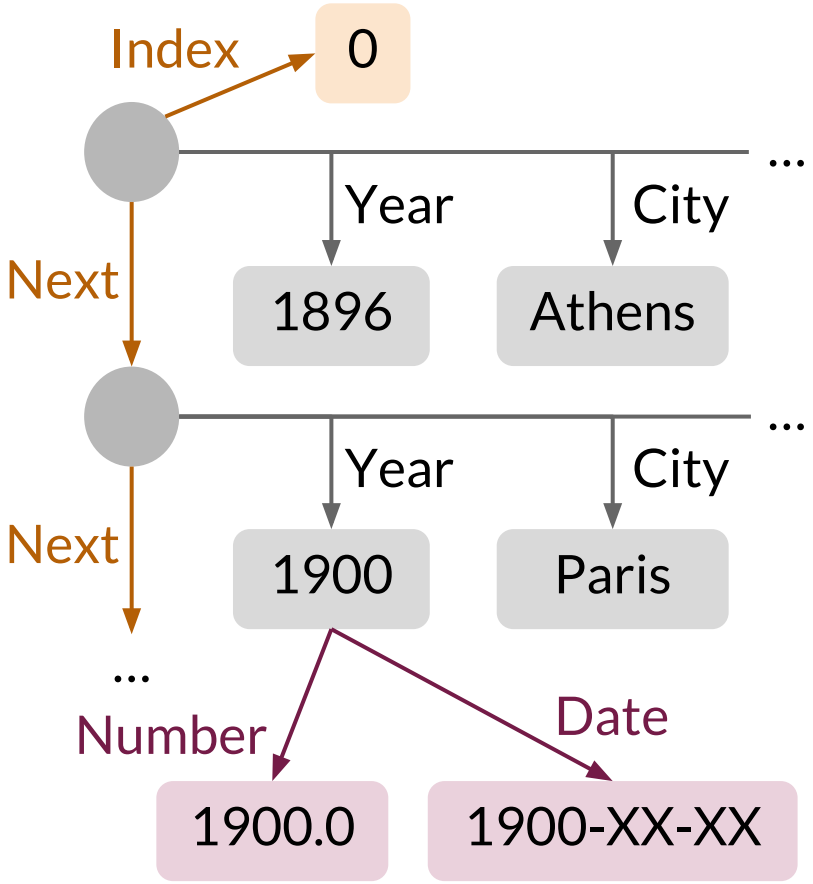
Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
…	…	…	…
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204



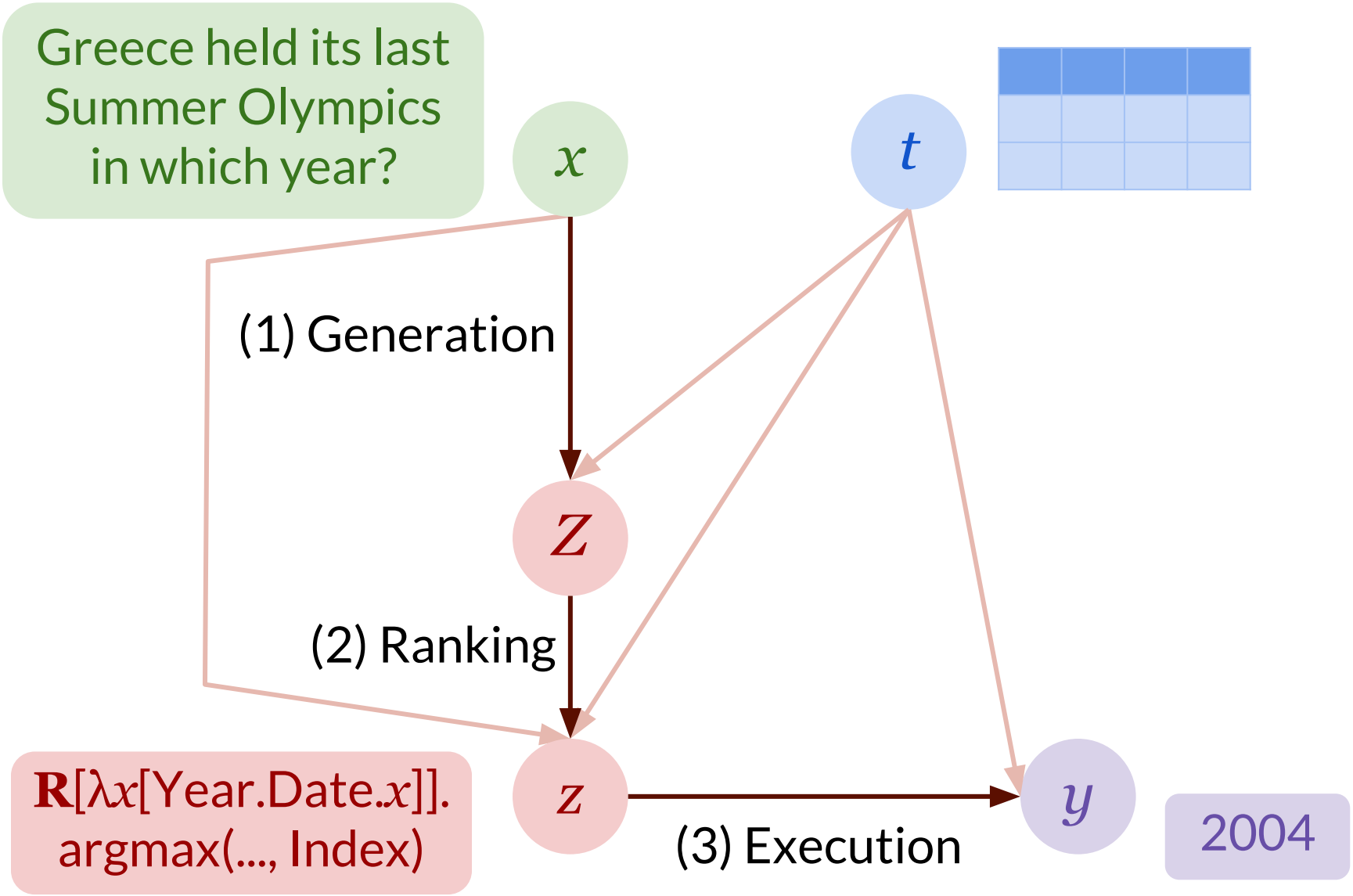
Representation

Convert table t to knowledge graph w

Year	City	Country	Nations
1896	Athens	Greece	14
1900	Paris	France	24
1904	St. Louis	USA	12
…	…	…	…
2004	Athens	Greece	201
2008	Beijing	China	204
2012	London	UK	204

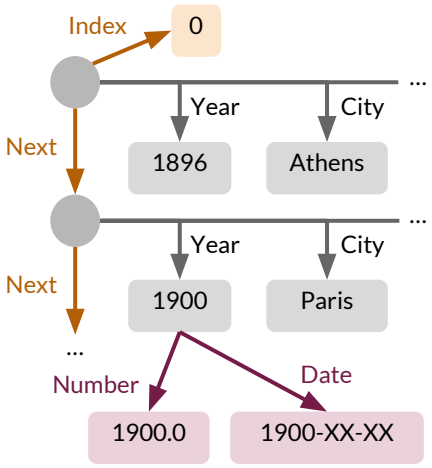


Approach


Greece held its last Summer Olympics in which year?
x
t
(1) Generation
z
(2) Ranking
R[λx[Year.Date.x]]. argmax(..., Index)
z
(3) Execution
y
2004

Approach

Greece held its last Summer Olympics in which year?



x

w

(1) Generation

z

(2) Ranking

z

(3) Execution

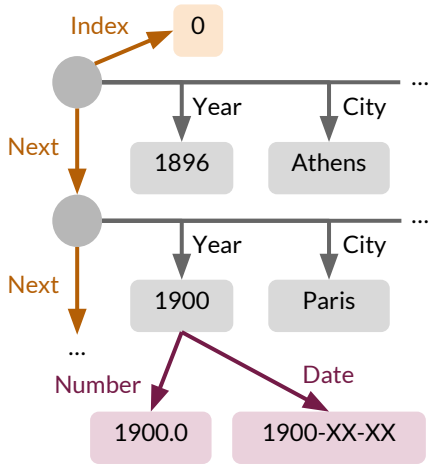
y

2004

R[λx[Year.Date.x]].
argmax(..., Index)

Approach

Greece held its last Summer Olympics in which year?



x

w

(1) Generation

z

(2) Ranking

$\mathbf{R}[\lambda x[\text{Year.Date}.x]].$ argmax(..., Index)

z

(3) Execution

y

2004

Generation

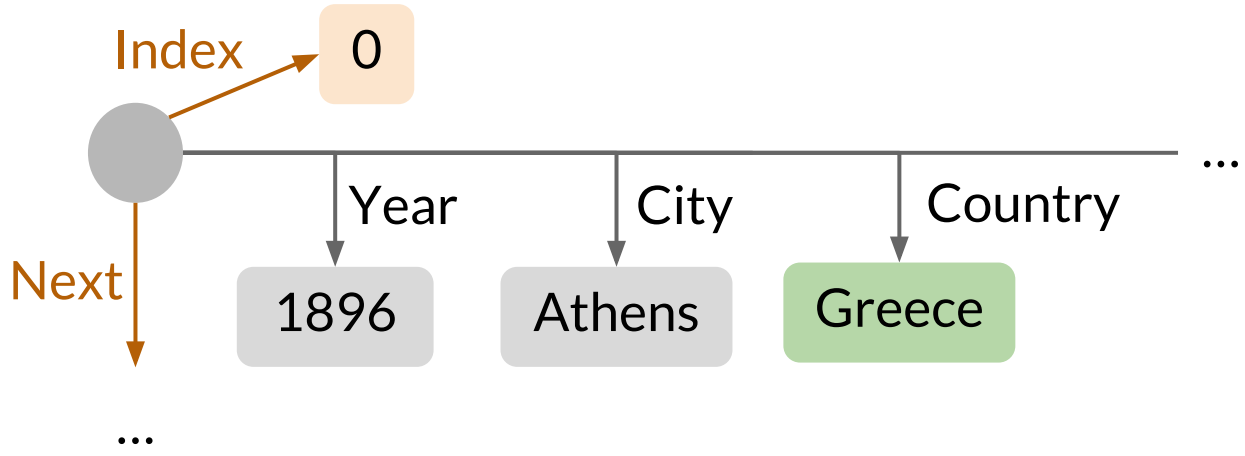
Build formulas **bottom-up** according to a set of deduction rules

$\mathbf{R}[\lambda x[\text{Year.Date.}x]]$.argmax(Country.Greece, Index)

Greece held its last Summer Olympics in which year?

Generation

Build formulas **bottom-up** according to a set of deduction rules



Greece held its last Summer Olympics in which year?

Generation

Build formulas **bottom-up** according to a set of deduction rules

Generation

Complication: Some logical predicates (e.g., relation Country) don't map to any phrase

R[λ*x*[Year.Date.*x*]].argmax(**Country**.Greece, Index)

Generation

Complication: Some logical predicates (e.g., relation Country) don't map to any phrase

R[λ*x*[Year.Date.*x*]].argmax(**Country**.Greece, Index)

Even when there is such a phrase, **we may still don't know the mapping** if we have not seen the relation in any table in the training data

Greece

Greece held its last Summer Olympics in which year?

Generation

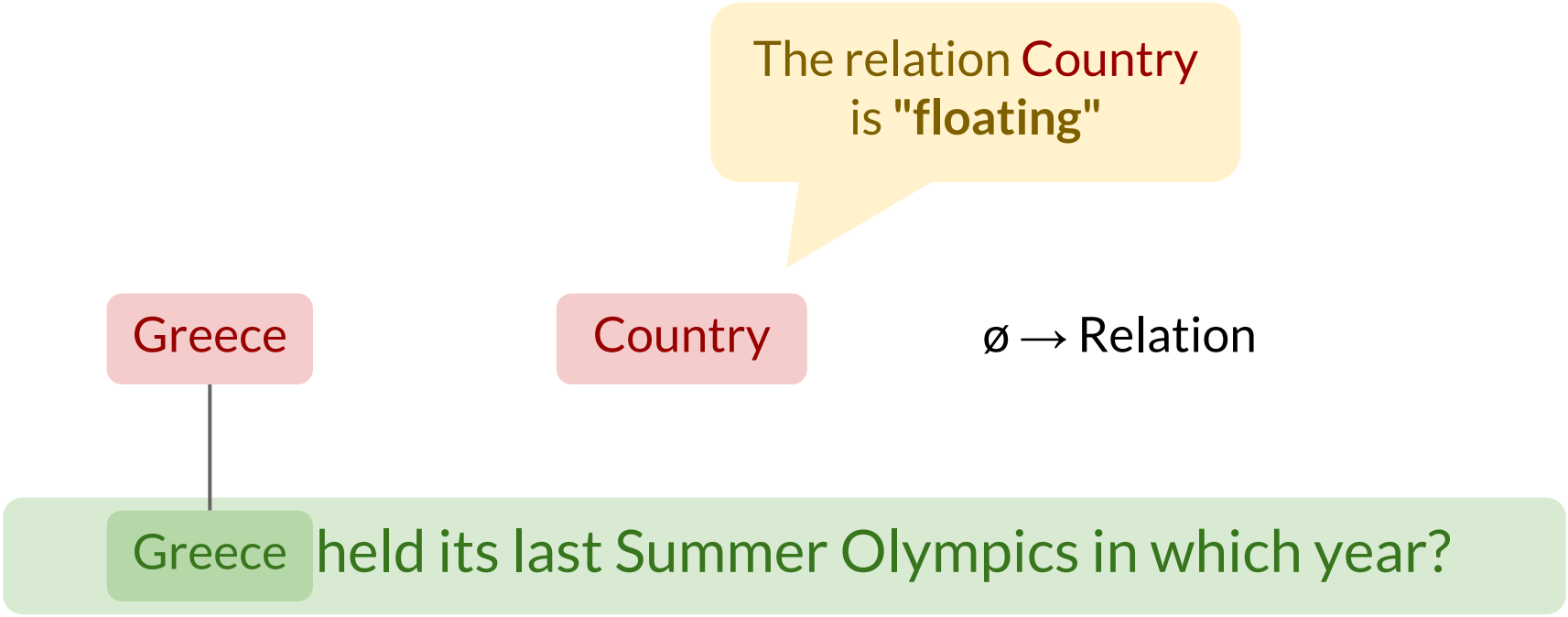
Idea: Allow formulas to be created from nothing (**"floating"**)

- Inspired by "bridging" [Berant et al., 2013]

Generation

Idea: Allow formulas to be created from nothing (**"floating"**)

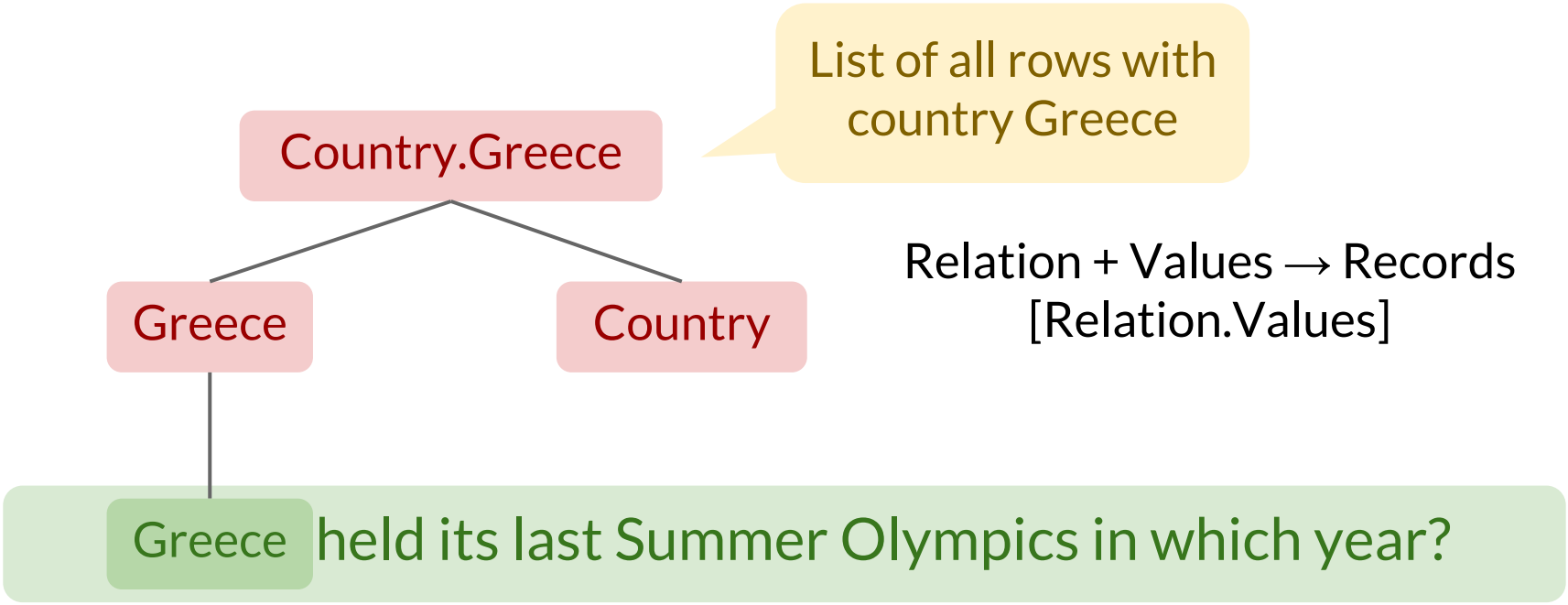
- Inspired by "bridging" [Berant et al., 2013]



Generation

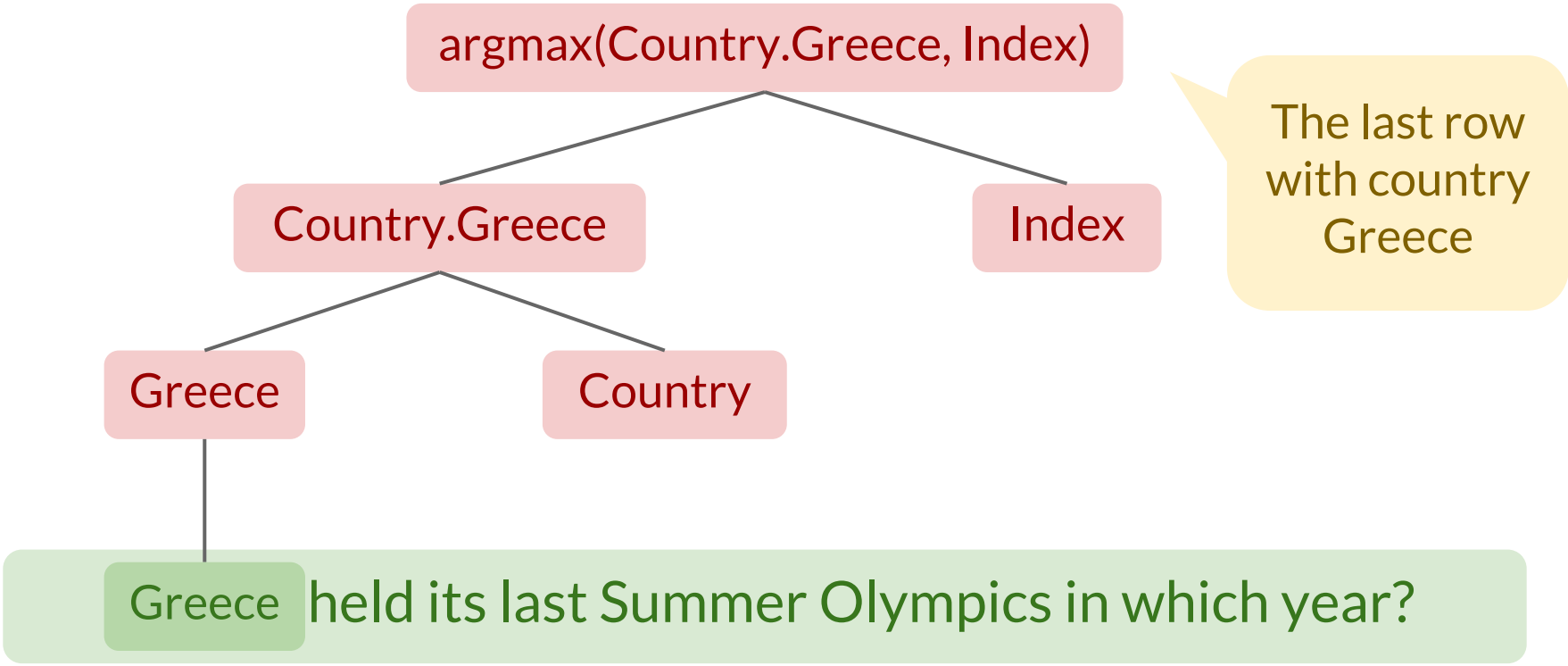
Idea: Allow formulas to be created from nothing (**"floating"**)

- Inspired by "bridging" [Berant et al., 2013]



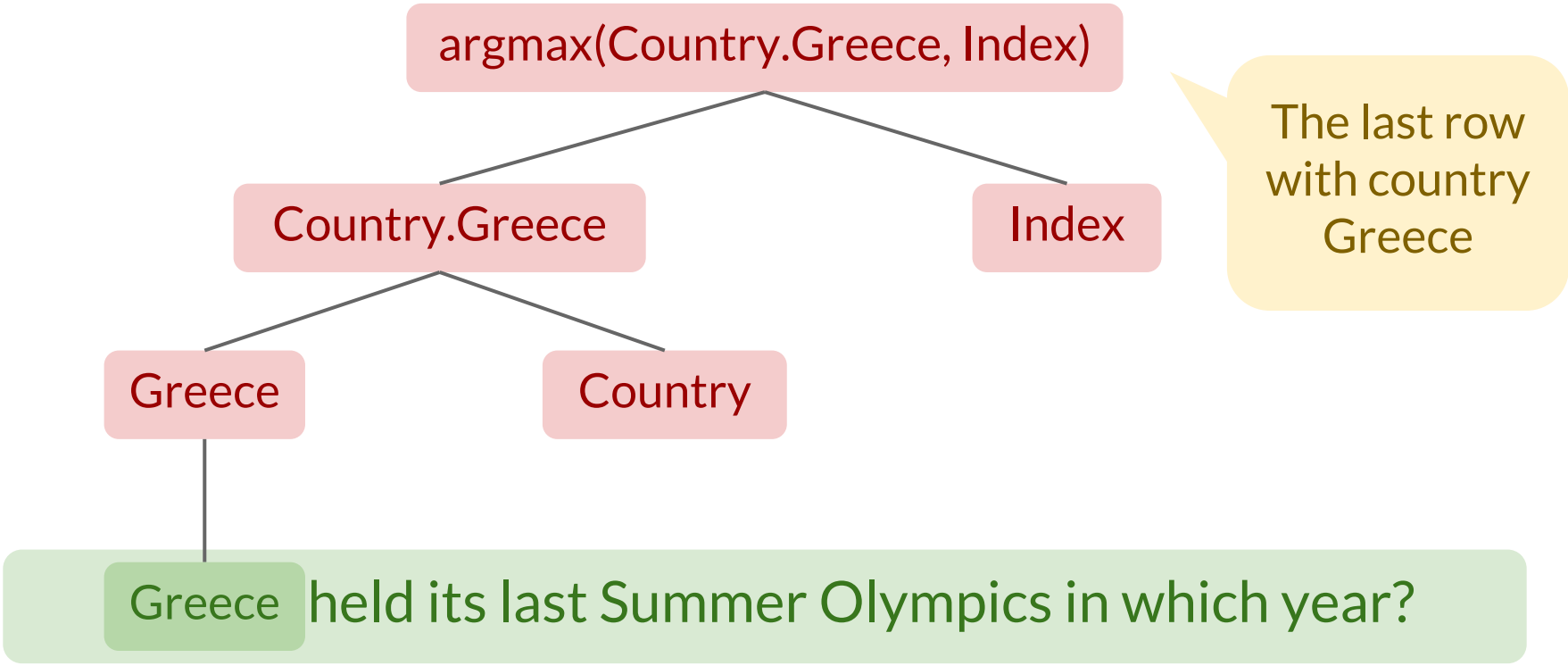
Generation

- **Entities** are anchored to token spans
- **Relations** and **Operations** are not



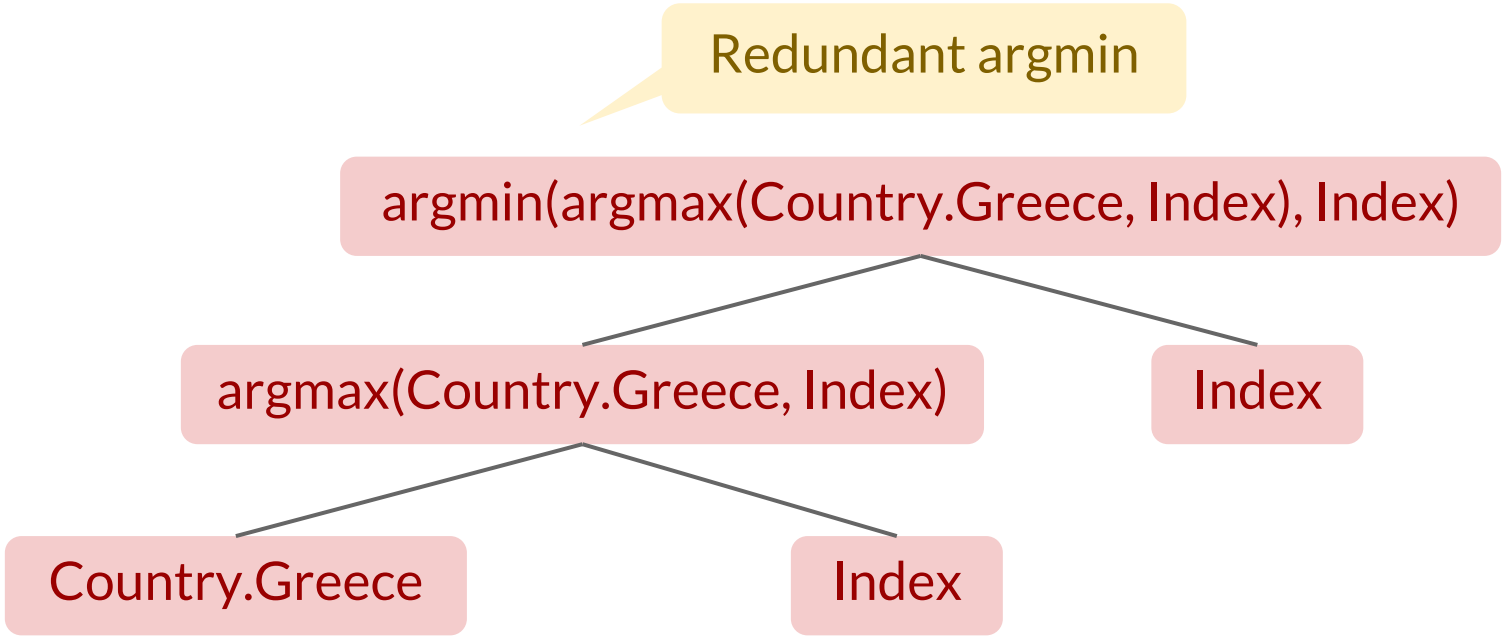
Generation

Connection between floating predicates and phrases in the question are made during **ranking**



Generation

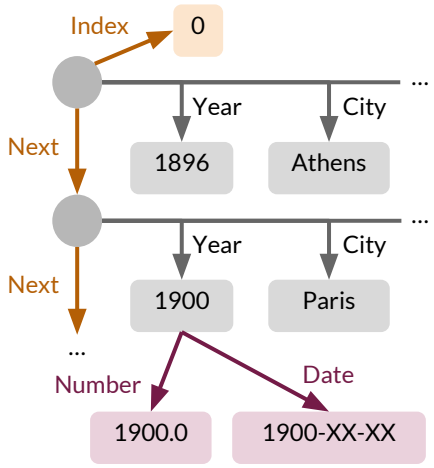
Problem: Over-generation due to high recursion



Handled by beam search and pruning heuristics

Approach

Greece held its last Summer Olympics in which year?



x

w

(1) Generation

z

(2) Ranking

$\mathbf{R}[\lambda x[\text{Year.Date}.x]].$ argmax(..., Index)

z

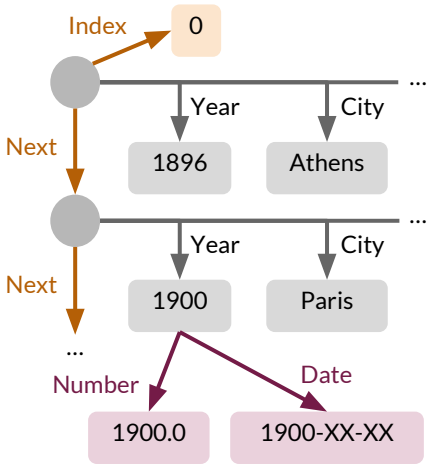
(3) Execution

y

2004

Approach

Greece held its last Summer Olympics in which year?



x

w

(1) Generation

z

(2) Ranking

z

y

2004

(3) Execution

R[λx[Year.Date.x]].
argmax(..., Index)

Ranking

Given a set Z of candidate formulas z, define a log-linear distribution:

$$p_\theta(z \mid x, w) \propto \exp\{\theta^\top \varphi(x, w, z)\}$$

where

- θ = parameter vector
- $\varphi(x, w, z)$ = feature vector

Ranking

Features:

- Relate phrases in x to predicates in z

(phrase = last, predicate = argmax)

(phrase = year, predicate = Year)

phrase == predicate

Ranking

Features:

- Relate phrases in x to predicates in z

(phrase = last, predicate = argmax)

(phrase = year, predicate = Year)

phrase == predicate

- Relate phrases in x to properties of $y = [\![z]\!]_w$

(headword = year, answer's type = NUMBER)

headword == answer's column

Learning

Given training example (x, w, y), define

$$p_\theta(y \mid x, w) = \sum_{z \in Z} p_\theta(z \mid x, w)\, \mathbf{I}(y = [\![z]\!]_w)$$

As usual, we choose θ to maximize the (L1 regularized) expectation of $\log p_\theta(y \mid x, w)$ over training data

Outline

- Background and Related Work
- Task and Dataset
- Approach
- **Experiments**

Results

- Oracle: Able to generate a candidate formula $z \in Z$ that executes to y
- Accuracy: The highest-ranked z executes to y

Results

- Oracle: Able to generate a candidate formula $z \in Z$ that executes to y
- Accuracy: The highest-ranked z executes to y

Two baselines:

- IR-inspired: Pick an answer among table cells by putting softmax over table cells
- WQ: Restrict the generation rules to the ones from Berant and Liang (2014)

Results on Test Set

	accuracy	oracle
IR-inspired	12.7	70.6
WQ		
This work		

In all settings, tables in test data are not seen during training

Results on Test Set

	accuracy	oracle
IR-inspired	12.7	70.6
WQ	24.3	35.6
This work		

In all settings, tables in test data are not seen during training

Results on Test Set

	accuracy	oracle
IR-inspired	12.7	70.6
WQ	24.3	35.6
This work	**37.1**	76.6

In all settings, tables in test data are not seen during training

Positive Examples

What is the **last** title that spicy horse produced?

How many districts have a population density of **at least** 1000?

Who finished directly **after** the driver who finished in 1:28.745?

(Information retrieval alone can't answer these questions)

Error Analysis

(1) Anchoring [18%]

how many **mexican** swimmers ranked in the top 10?

Rank	Swimmer	Country	Time	Note
		Mexico		

Error Analysis

(2) Normalization [29%]

how long does the show defcon 3 last?

ET	Days	Program	Hosts	Description
2pm–3pm				

Error Analysis

(3) Unhandled Operations [19%]

was there more gold medals won than silver?

(boolean answer)

which movies were number 1 for at least two consecutive weeks?

(consecutive count)

how many titles had the same author listed as the illustrator?

(count rows with arbitrary conditions)

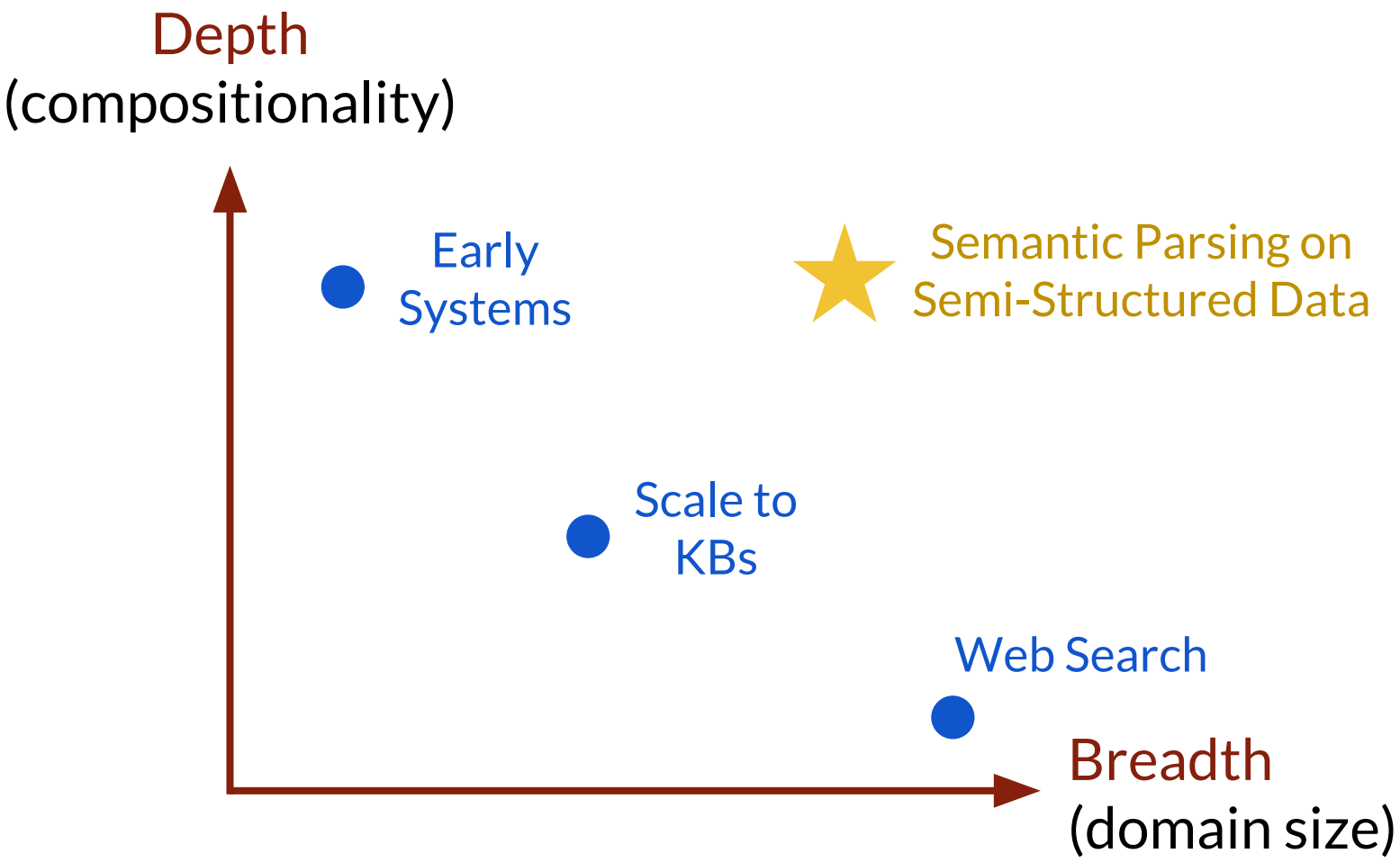
Error Analysis

(4) Ranking Errors [24%]

how many buildings on the list are **taller** than 200 feet?

Name	Street Address	Years as Tallest	Height ft (m)	Floors
			792 (241)	

Conclusion


Depth
(compositionality)
Early
Systems
Semantic Parsing on
Semi-Structured Data
Scale to
KBs
Web Search
Breadth
(domain size)

Conclusion

Dataset and reproducible experiments are available on CodaLab:

nlp.stanford.edu/software/sempre/wikitable

Thank you