The State of Natural Language Understanding

Siva Reddy
@sivareddyg

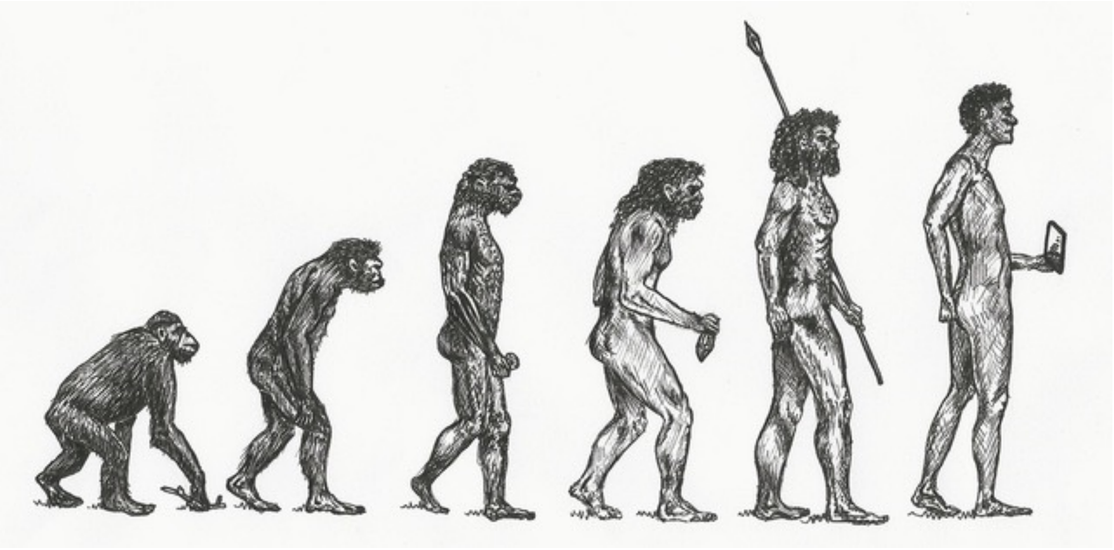
Stanford University

Focus is on question answering using

- Knowledge bases (Databases)
- Documents

Language and Intelligence

- Distinctive characteristic of human species
- An assessment of intelligence



Credits: Chris Manning

Intelligent Computers (1950)

Turing Test

Q: Please write me a sonnet on the subject of the Forth Bridge.

A: Count me out on this one. I never could write poetry.

Q: Add 34957 to 70764.

A: (Pause about 30 seconds and then give as answer) 105621.

Credits: Percy Liang

Symbolic AI (1960-1990)

- Deterministic mapping of language to symbolic representations
- Pattern matching and finite state machines
- Grammars, Syntax to Logic,
- Closed-domain language interpreters

Early AI Assistants

Question Answering Systems

- Baseball (1961), LUNAR (1971)

Chatbot

- ELIZA (1966)

Most successful AI Assistant

- SHRDLU (1972)

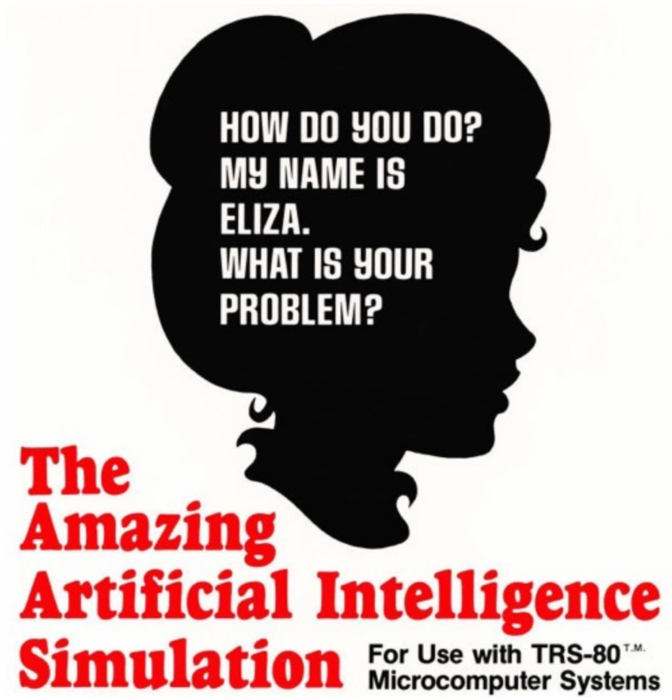
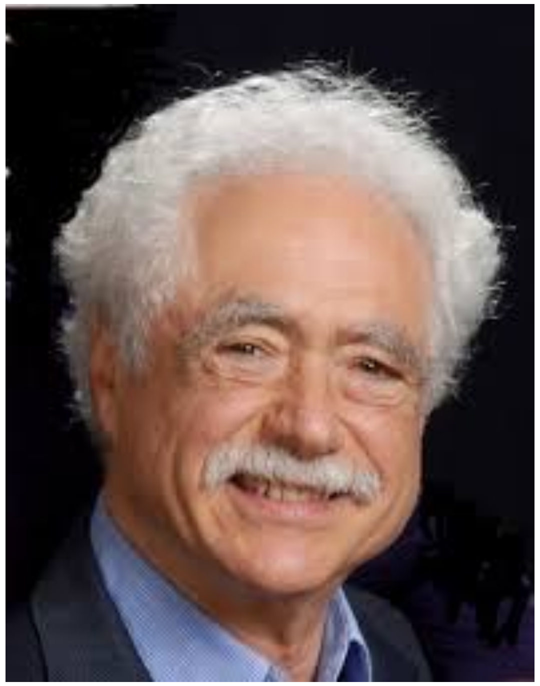


Image: analyticsindiamag.com

SHRDLU



Person: Pick up a big red block.

Computer: OK.

Person: Grasp the pyramid.

Computer: I don't understand which pyramid you mean.

Person (changing their mind): Find a block which is taller than the one you are holding and put it into the box.

Computer: By "it", I assume you mean the block which is taller than the one I am holding.

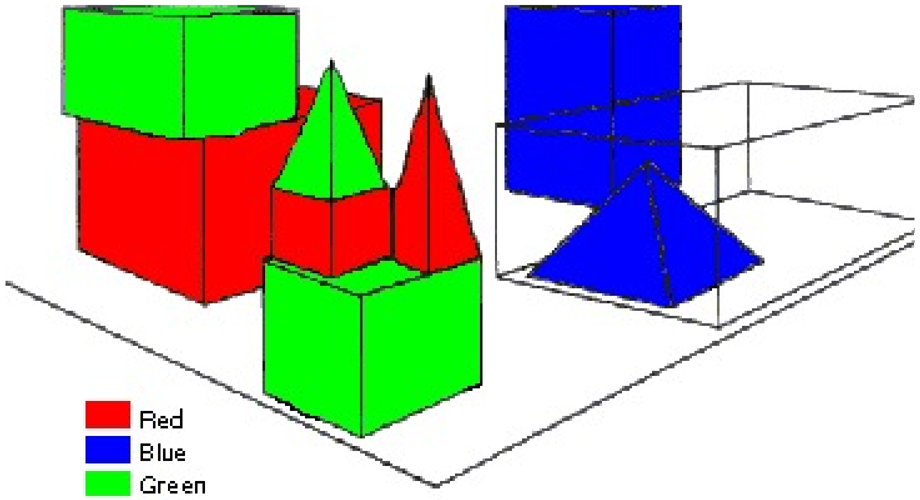
Computer: OK.

Person: What does the box contain?

Computer: The blue pyramid and the blue block.

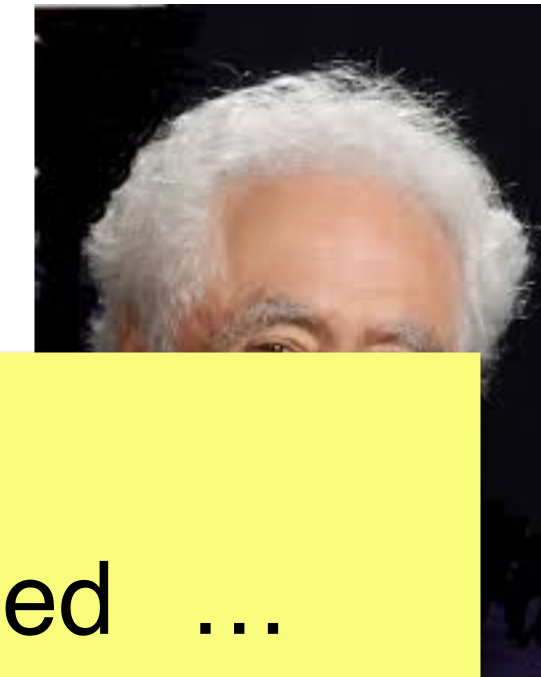
Person: What is the pyramid supported by?

Computer: The box.



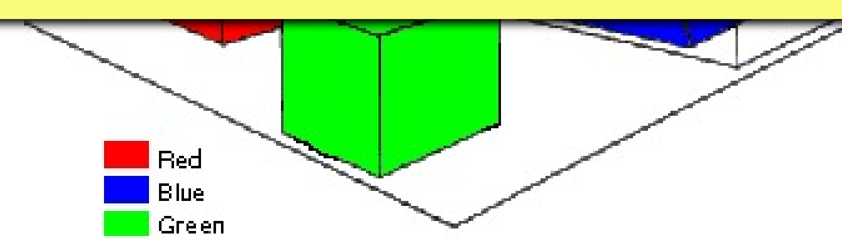
Credits: Percy Liang

SHRDLU



A number of people have suggested … SHRDLU program for understanding natural language represent a kind of dead end in AI programming. ....

Even having written the program, I find it near the limit of what I can keep in mind at once — **Terry Winograd**



Credits: Percy Liang

Statistical NLU (1990-2010)

- Probabilistic grammars from **annotated data**
- Language paired with programs (Zelle and Mooney, 1996)

Statistical NLU (1990-2010)

- Probabilistic grammars from **annotated data**
- Language paired with programs (Zelle and Mooney, 1996)

What states border Texas
$\lambda x.state(x) \wedge borders(x, texas)$

What is the largest state
$\arg\max(\lambda x.state(x), \lambda x.size(x))$

What states border the state that borders the most states
$$\lambda x.state(x) \wedge borders(x, \arg\max(\lambda y.state(y), \lambda y.count(\lambda z.state(z) \wedge borders(y, z))))$$

Credits: Zettlemoyer and Collins (2005)

Grammar Learning

[Zettlemoyer and Collins 2005, Liang et al. 2011]

Type.HumanLanguage
Type.ProgrammingLanguage

alignment

Brazil
BrazilFootballTeam

alignment

What languages do people in Brazil use

Features are the key

- languages -> Type.HumanLanguage

Train a **ML model** to identify good and bad features based on the context

Grammar Learning

[Zettlemoyer and Collins 2005, Liang et al. 2011]

Type.HumanLanguage
Type.ProgrammingLanguage

alignment

Brazil
BrazilFootballTeam

nt

What **langua** **use**

- **Expensive** to annotate training data
- Domain-specific grammars

Feat

- ype.HumanLanguage

Train a **ML model** to identify good and bad features based on the context

NLU meets IR

TREC, IBM Watson, Google

who invented the electric car

All News Images Shopping Videos More Settings Tools

About 927,000 results (0.46 seconds)

Robert Anderson

Scottish inventor **Robert Anderson** invents the first crude electric carriage powered by non-rechargeable primary cells. American **Thomas Davenport** is credited with building the first practical electric vehicle -- a small locomotive. French physicist Gaston Planté invents the rechargeable lead-acid storage battery. Oct 30, 2009

Timeline: History of the Electric Car . NOW on PBS
www.pbs.org/now/shows/223/electric-car-timeline.html

Feedback

The History of the Electric Car | Department of Energy
https://energy.gov/articles/history-electric-car
Sep 15, 2014 - And while **Robert Anderson**, a **British** inventor, developed the first crude electric carriage around this same time, it wasn't until the second half of the 19th century that French and English inventors built some of the first practical electric cars.

Train a **feature-based model** to identify the most similar sentence

Scaling NLU (2010-2016)

[Berant et al. 2013, Kwiatkowski et al. 2013, Reddy et al. 2014]

Google Knowledge Graph, WikiData

Alternate forms of supervision

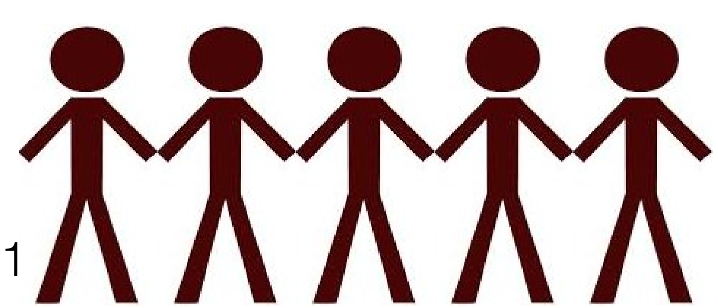
Scaling NLU (2010-2016)

[Berant et al. 2013, Kwiatkowski et al. 2013, Reddy et al. 2014]

Google Knowledge Graph, WikiData

Alternate forms of supervision

Heavy supervision	**Light supervision**
What's Bulgaria's capital?	*What's Bulgaria's capital?*
CapitalOf(Bulgaria)	`Sofia`
When was Walmart started?	*When was Walmart started?*
DateFounded(Walmart)	`1962`
What movies has Tom Cruise been in?	*What movies has Tom Cruise been in?*
Movies $\cap$ Starring(TomCruise)	`TopGun,VanillaSky,...`
...	...



1

Image Credits: Percy Liang

Training intuition

Where did Mozart tupress?

Vienna

Credits: Percy Liang

Training intuition

Where did Mozart tupress?

```
PlaceOfBirth(WolfgangMozart)
PlaceOfDeath(WolfgangMozart)
PlaceOfMarriage(WolfgangMozart)
```

Vienna

Credits: Percy Liang

Training intuition

Where did Mozart tupress?

~~`PlaceOfBirth(WolfgangMozart)` ⇒ Salzburg~~

`PlaceOfDeath(WolfgangMozart)` ⇒ Vienna

`PlaceOfMarriage(WolfgangMozart)` ⇒ Vienna

Vienna

Credits: Percy Liang

Training intuition

Where did Mozart tupress?

~~PlaceOfBirth(WolfgangMozart) ⇒ Salzburg~~

PlaceOfDeath(WolfgangMozart) ⇒ Vienna

PlaceOfMarriage(WolfgangMozart) ⇒ Vienna

Vienna

Where did Hogarth tupress?

PlaceOfBirth(WilliamHogarth) ⇒ London

PlaceOfDeath(WilliamHogarth) ⇒ London

~~PlaceOfMarriage(WilliamHogarth) ⇒ Paddington~~

London

Credits: Percy Liang

Training intuition

Where did Mozart tupress?

~~PlaceOfBirth(WolfgangMozart) ⇒ Salzburg~~

$\boxed{\texttt{PlaceOfDeath(WolfgangMozart)} \Rightarrow \text{Vienna}}$

PlaceOfMarriage(WolfgangMozart) ⇒ Vienna

Vienna

Where did Hogarth tupress?

PlaceOfBirth(WilliamHogarth) ⇒ London

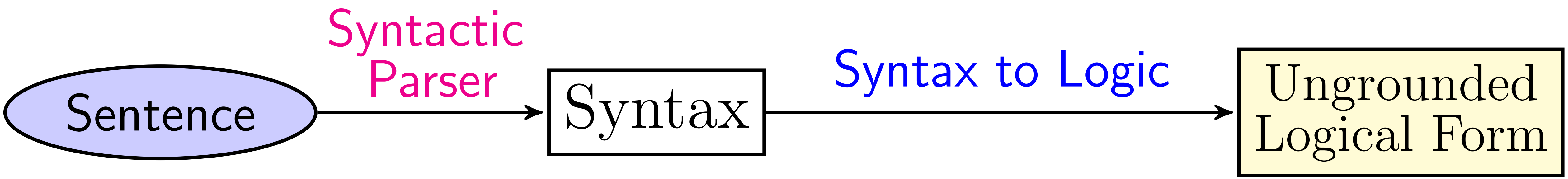
$\boxed{\texttt{PlaceOfDeath(WilliamHogarth)} \Rightarrow \text{London}}$

~~PlaceOfMarriage(WilliamHogarth) ⇒ Paddington~~

London

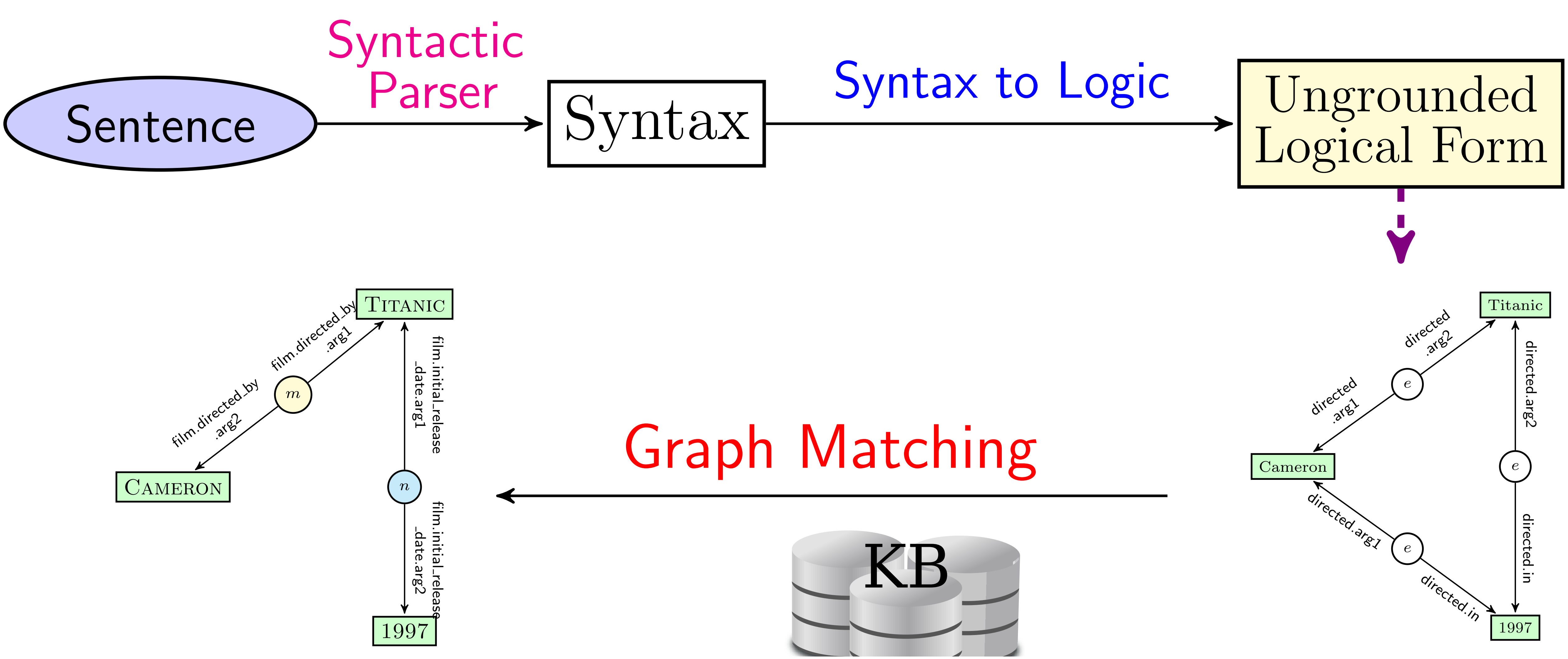
Credits: Percy Liang

Scaling using Linguistic Representations



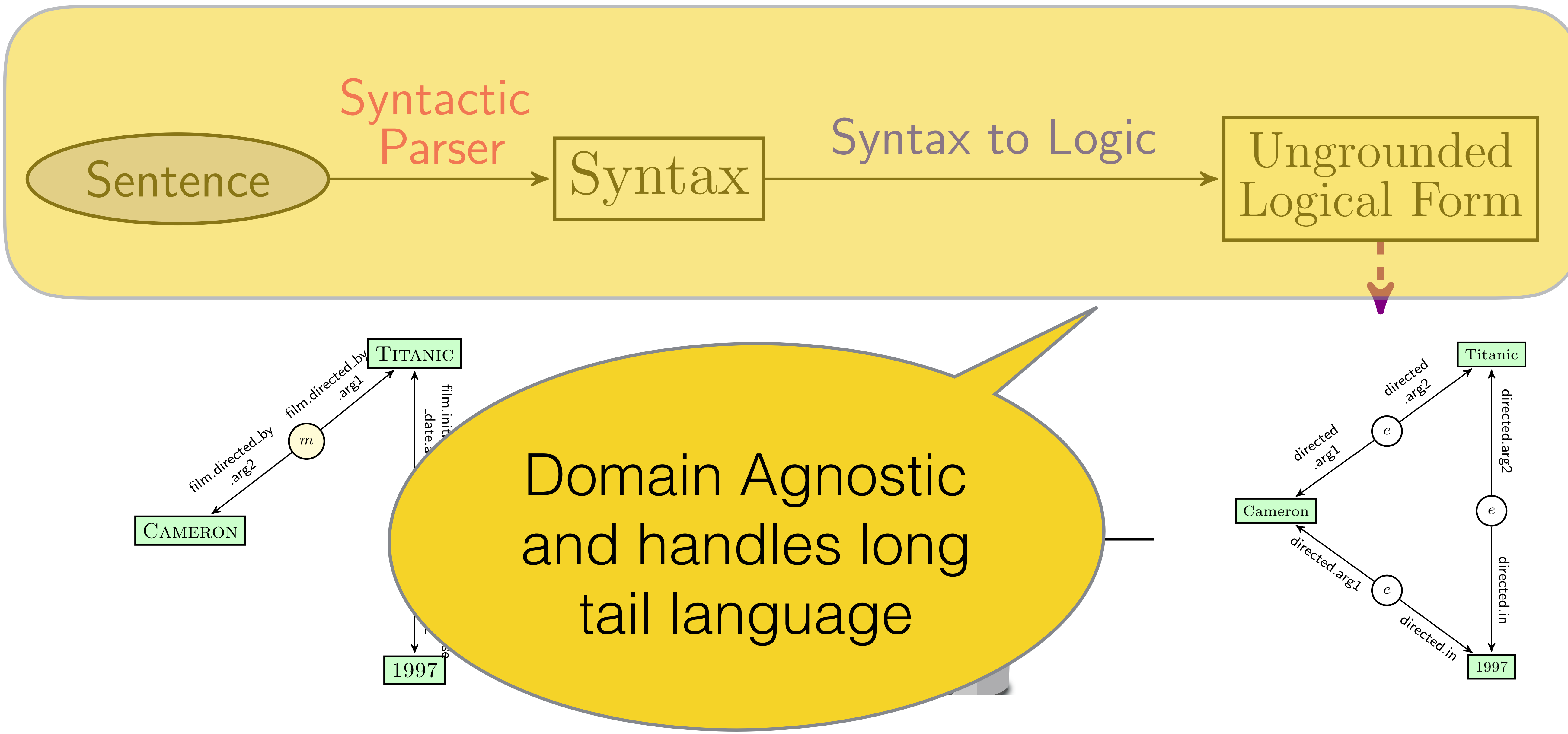
Reddy et al. (2016, 2017)

Scaling using Linguistic Representations



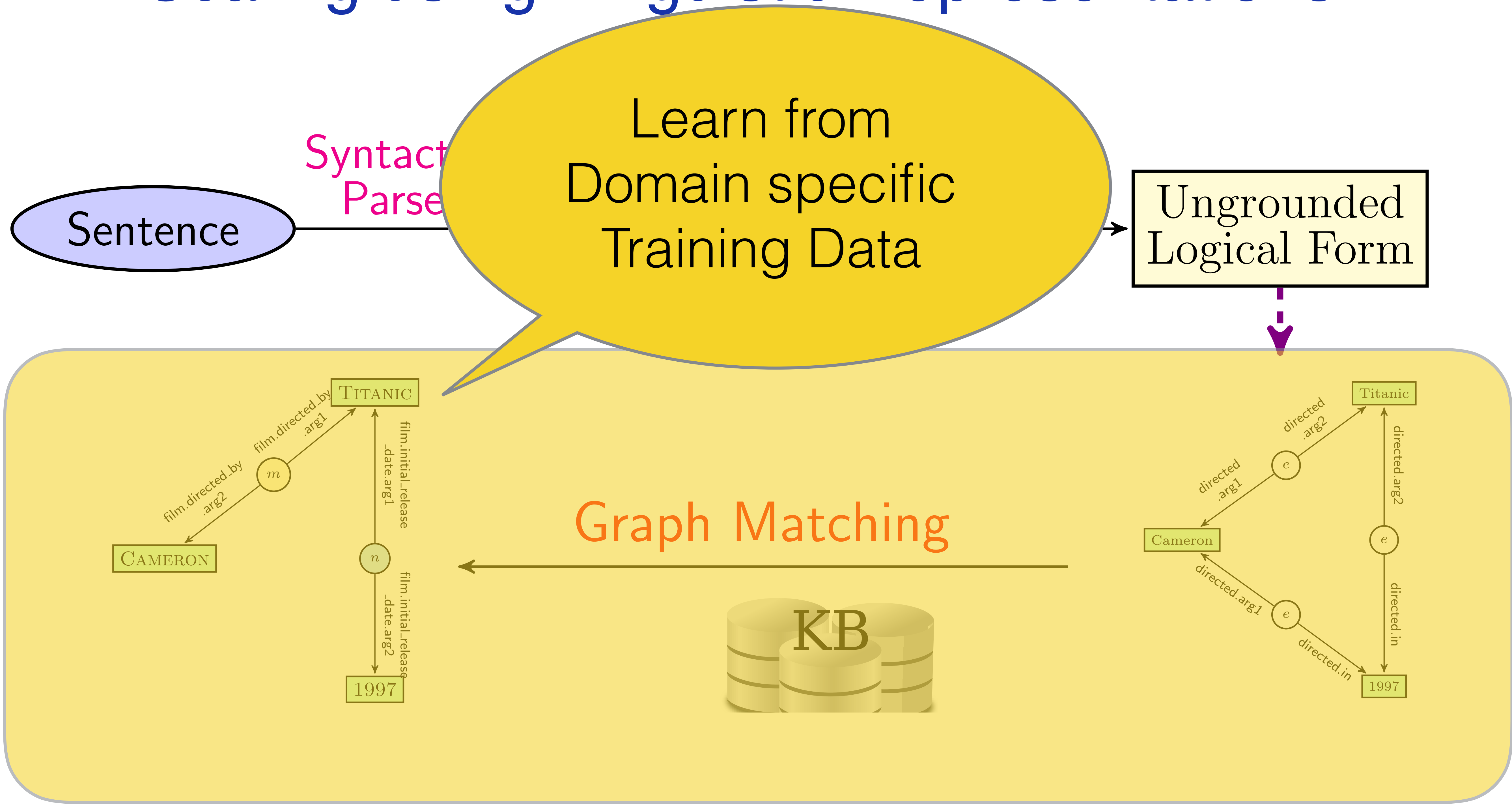
Reddy et al. (2016, 2017)

Scaling using Linguistic Representations



Reddy et al. (2016, 2017)

Scaling using Linguistic Representations



Reddy et al. (2016, 2017)

Deep Learning

The general approach to building Deep Learning systems is compelling and powerful: The researcher defines a **model architecture** and a **top-level loss function** and then both the **parameters** and the **representations** of the model **self-organize** so as to minimize this loss, in an end-to-end learning framework. — **Chris Manning (2015)**

Word Embeddings

[Mikolov et al. 2013, Pennington et al. 2014]

Symbols to numerical representations

How similar are **hotel** and **lodge**?

- **Traditional:** character overlap, dictionary
- **Embeddings**: hotel = [0.5 0.3 0.2 0.9]
  lodge = [0.9 0.2 0.4 0.1]

Entity and Relation Embeddings

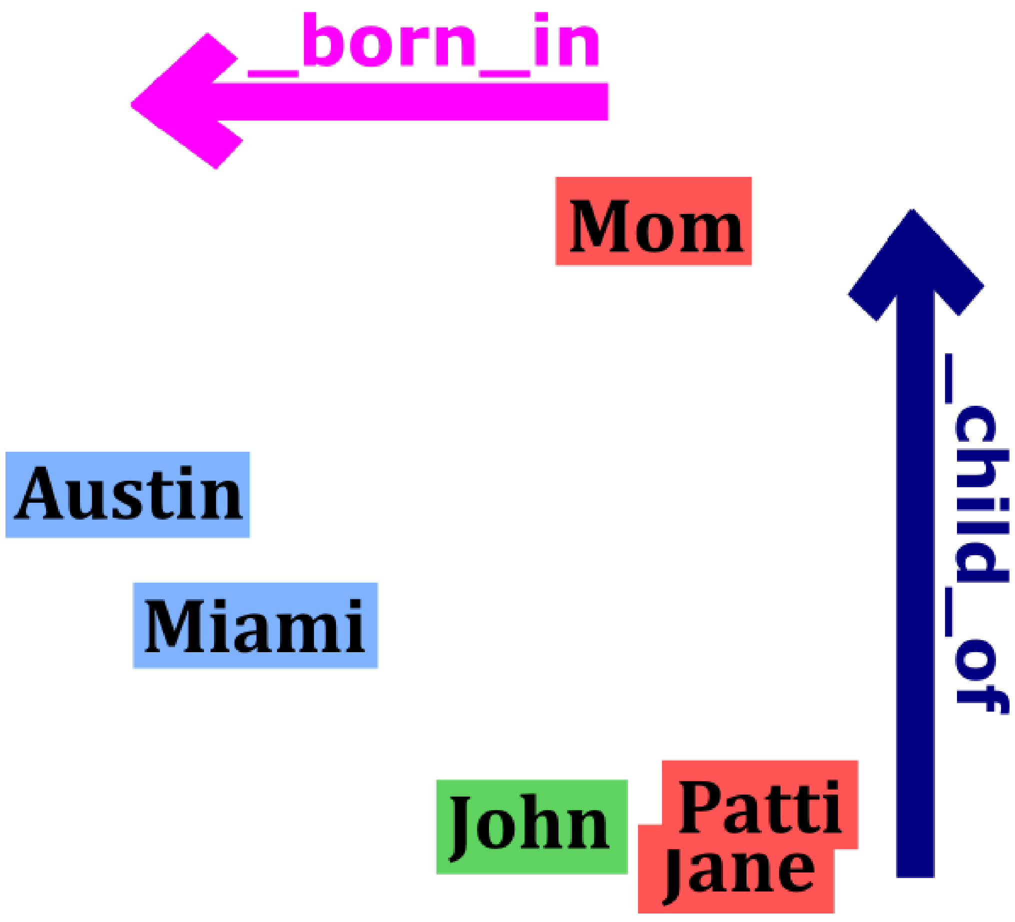
[Bordes et al. 2013]

Intuition: we want $\mathbf{s} + \mathbf{r} \approx \mathbf{o}$.

The similarity measure is defined as:

$$d(sub, rel, obj) = ||\mathbf{s} + \mathbf{r} - \mathbf{o}||_2^2$$

$\mathbf{s}$,$\mathbf{r}$ and $\mathbf{o}$ are learned to verify that.



Credits: Antoine Bordes, Jason Weston

Entity and Relation Embeddings

FRANCE	JESUS	SCRATCHED
SPAIN	CHRIST	SMASHED
ITALY	GOD	RIPPED
RUSSIA	RESURRECTION	BRUSHED
POLAND	PRAYER	HURLED
ENGLAND	YAHWEH	GRABBED
DENMARK	JOSEPHUS	TOSSED
GERMANY	MOSES	SQUEEZED
PORTUGAL	SIN	BLASTED
SWEDEN	HEAVEN	TANGLED
AUSTRIA	SALVATION	SLASHED

Credits: Antoine Bordes, Jason Weston

Entity and Relation Embeddings

WALL-E

+ movie.genre =

Credits: Antoine Bordes, Jason Weston

Entity and Relation Embeddings

WALL-E

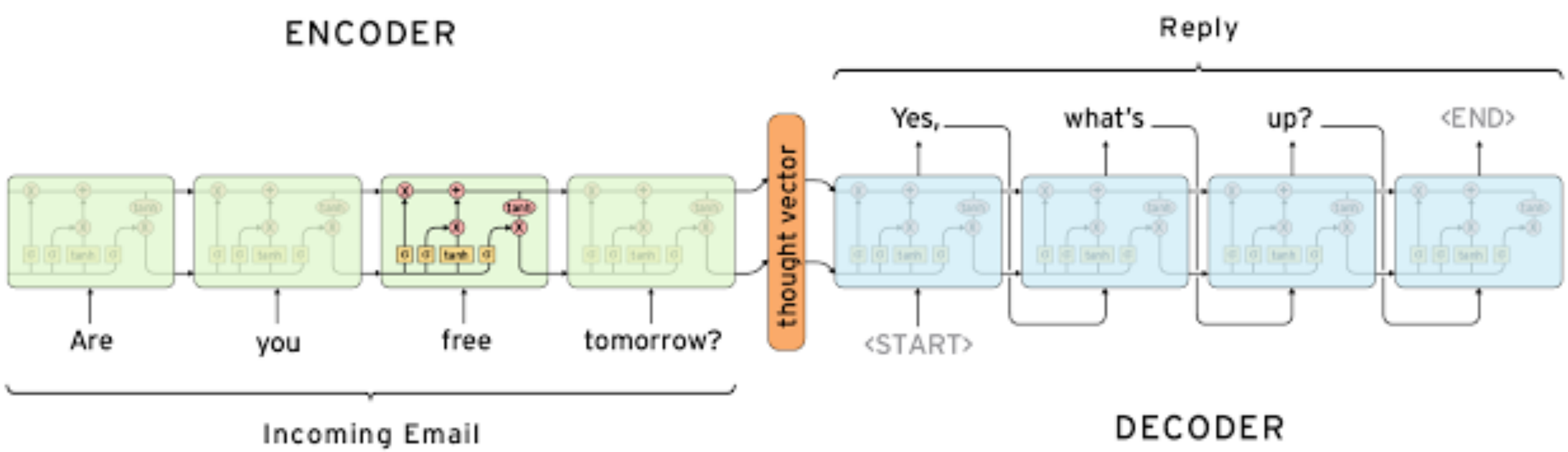
+ movie.genre =

Animation
Comedy Film
Adventure Film
Fantasy
Drama
Satire

Credits: Antoine Bordes, Jason Weston

End to End AI Assistants

- Sequence to Sequence Machine Translation (Cho et al. 2014, Sutskever et al. 2014)
- Easy to train on unstructured data



Credits: Denny Britz

Hard to control and debug

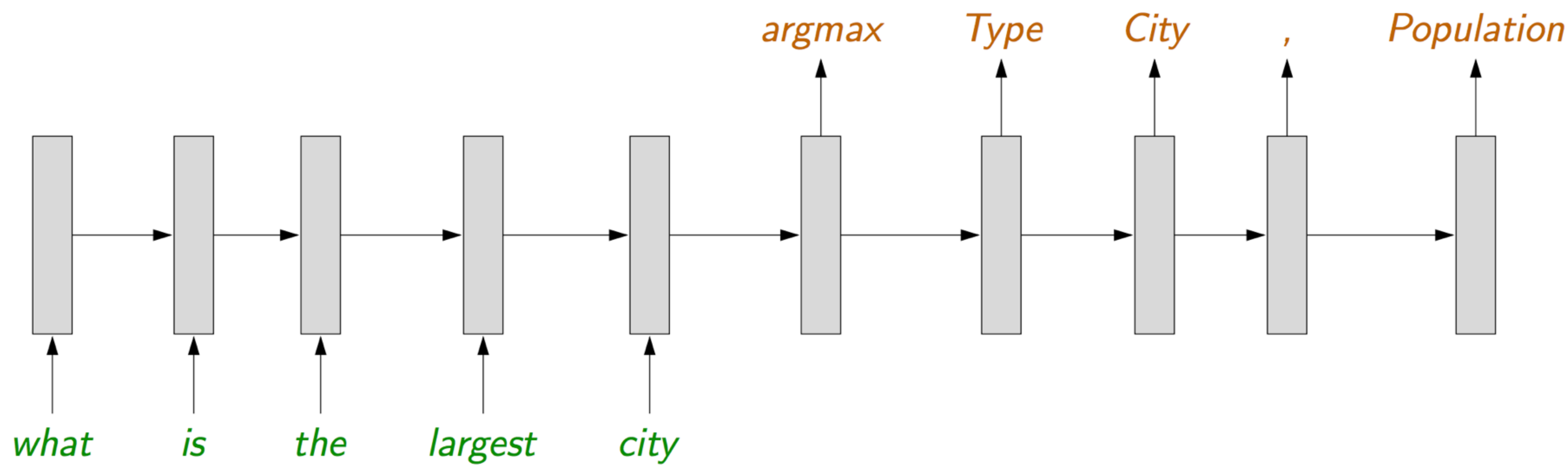
message	How old are you?
response	16 and you?
message	What's your age?
response	18.

[Li et al. 2016]

Microsoft Tay

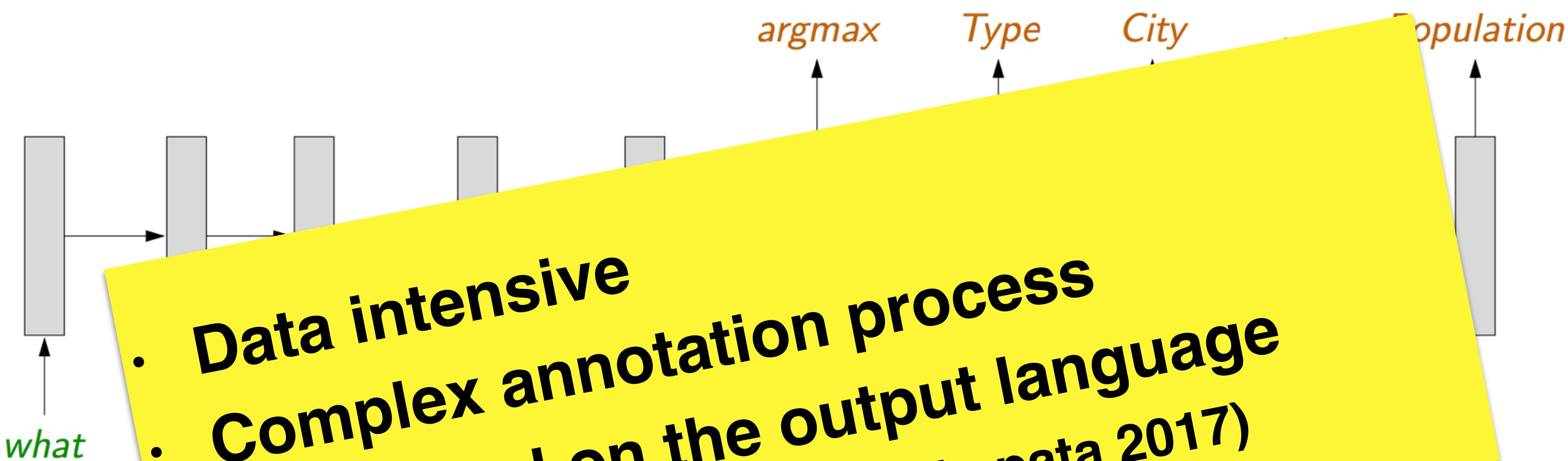
QA on Knowledge Bases

[Jia and Liang 2016, Dong and Lapata 2016]



QA on Knowledge Bases

[Jia and Liang 2016, Dong and Lapata 2016]



- Data intensive
- Complex annotation process
- No control on the output language

(See Cheng, Reddy, Saraswat, Lapata 2017)

Reading Comprehension

Question: *"The number of new Huguenot colonists declined after what year?"*

Paragraph: *"The largest portion of the Huguenots to settle in the Cape arrived between 1688 and 1689…but quite a few arrived as late as **1700**; thereafter, the numbers declined…"*

Correct Answer: *"**1700**"*

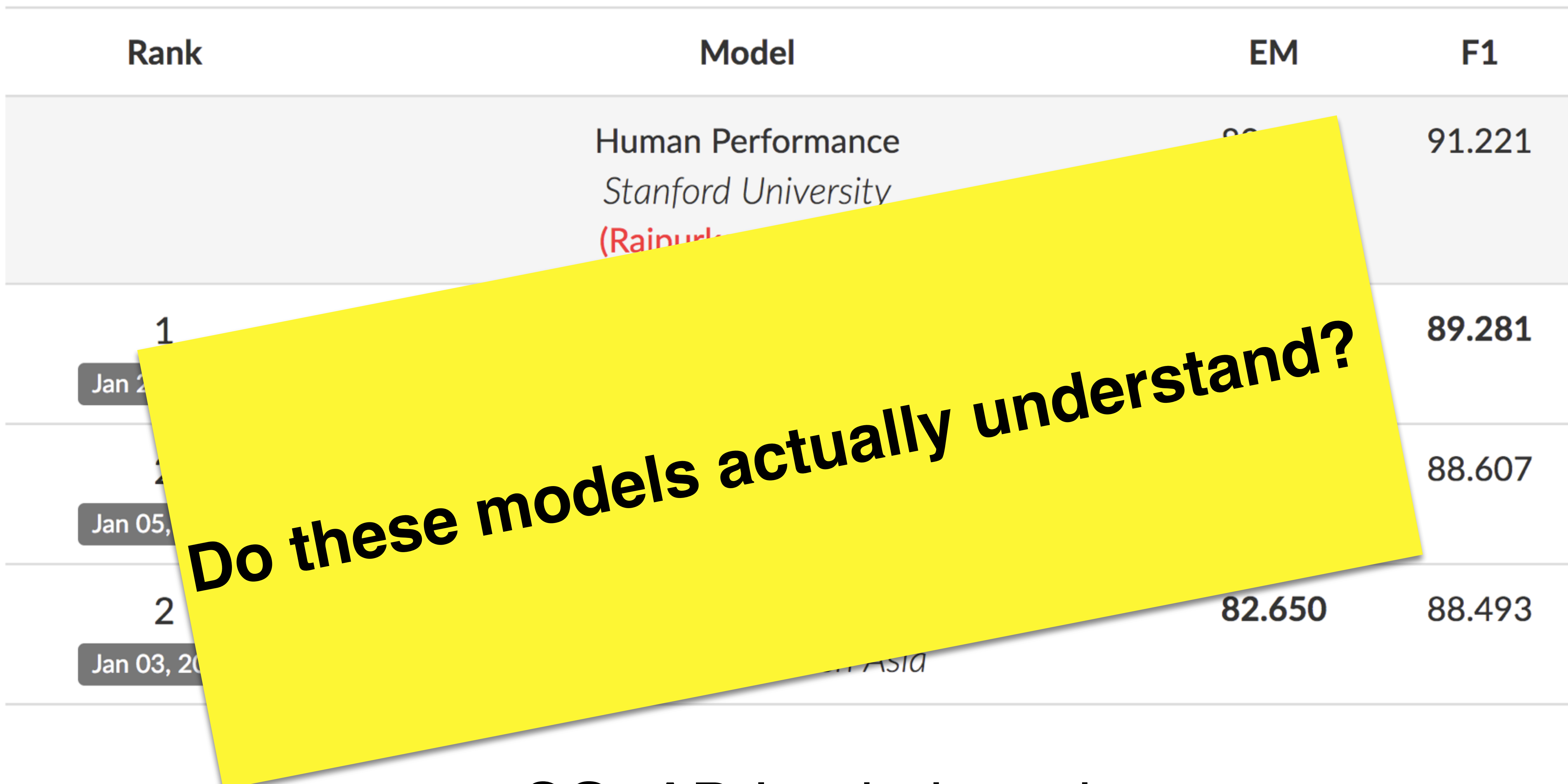
Tremendous gains using Deep Learning

Reading Comprehension

Rank	Model	EM	F1
	Human Performance *Stanford University* (Rajpurkar et al. '16)	82.304	91.221
1 Jan 22, 2018	Hybrid AoA Reader (ensemble) *Joint Laboratory of HIT and iFLYTEK Research*	82.482	**89.281**
2 Jan 05, 2018	SLQA+ (ensemble) *Alibaba iDST NLP*	82.440	88.607
2 Jan 03, 2018	r-net+ (ensemble) *Microsoft Research Asia*	**82.650**	88.493

SQuAD leaderboard

Reading Comprehension



SQuAD leaderboard

Adversarial attacks

[Jia and Liang 2016]

Question: *"The number of new Huguenot colonists declined after what year?"*

Paragraph: *"The largest portion of the Huguenots to settle in the Cape arrived between 1688 and 1689…but quite a few arrived as late as **1700**; thereafter, the numbers declined. The number of old Acadian colonists declined after the year of **1675**."*

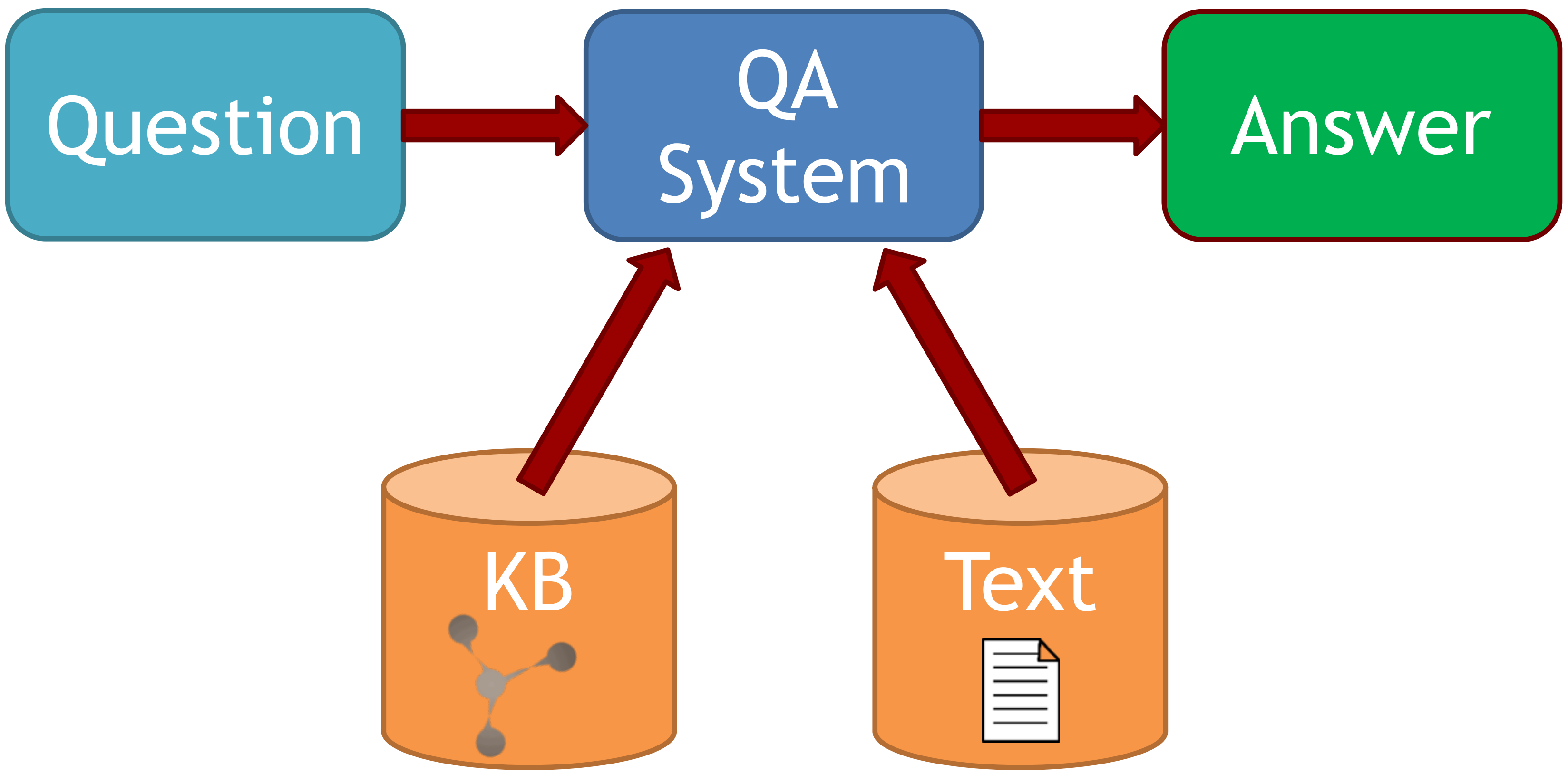
Correct Answer: *"**1700**"*

Predicted Answer: *"**1675**"*

Performance drops from 80s to 30s.

Multiple Knowledge Sources

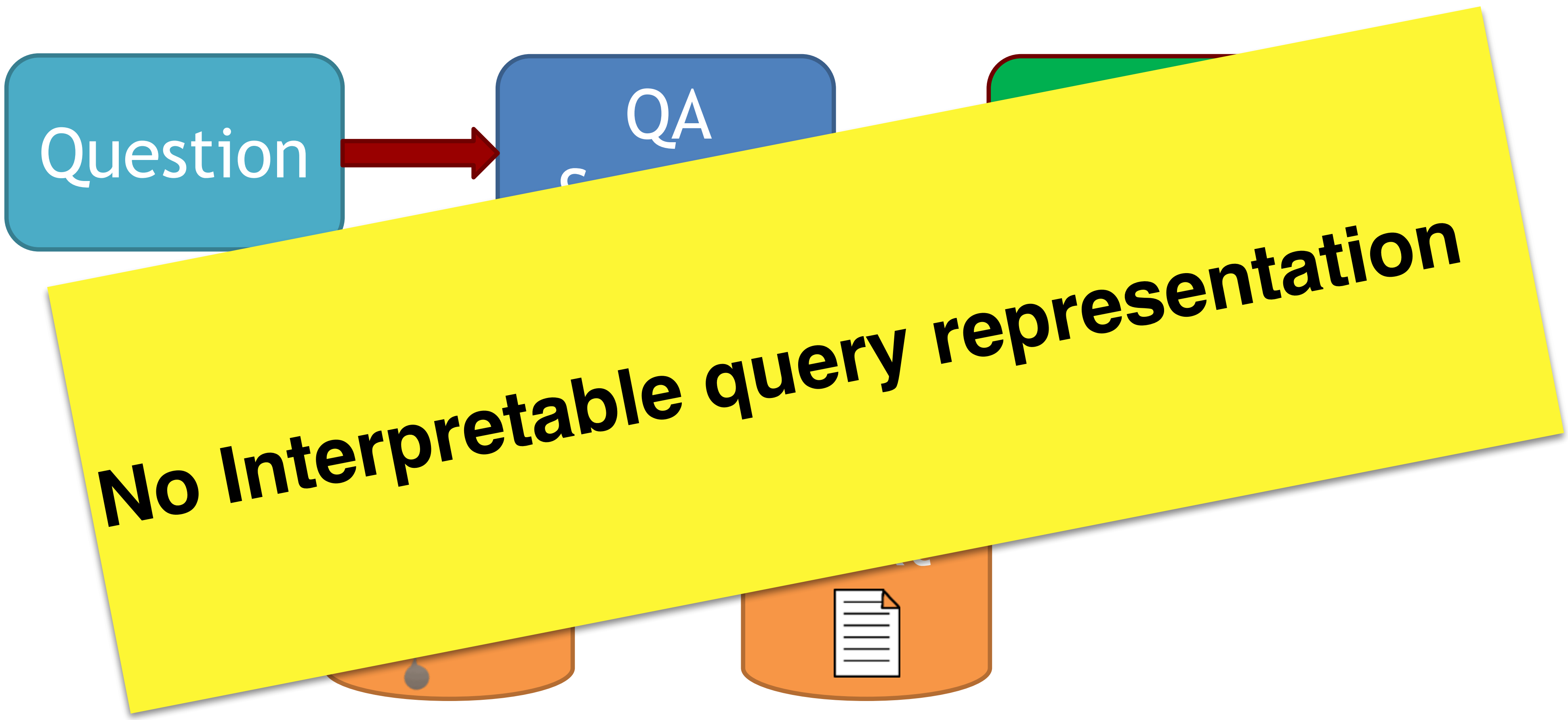
[Das, Zaheer, Reddy, McCallum 2017]



Who is the first African-american president of US?

Multiple Knowledge Sources

[Das, Zaheer, Reddy, McCallum 2017]



Who is the first African-american president of US?

Winners

Could be deployed with care

- Reading comprehension
- Simple QA on knowledge bases
- Chit chat bots
- Human assisted auto reply

Losers

Multi relational QA on knowledge bases

Goal oriented dialog

Summary

Don’t buy the Deep Learning hype for AI Assistants

NLU is harder than signal processing tasks

Data intensive and expensive

Limited control and low interpretability

Sensitive to adversarial attacks

Thank you

Special thanks to my sponsors

PhD Fellowship