

# Scalable Semantic Parsing with Partial Ontologies

Eunsol Choi   Tom Kwiatkowski   Luke Zettlemoyer

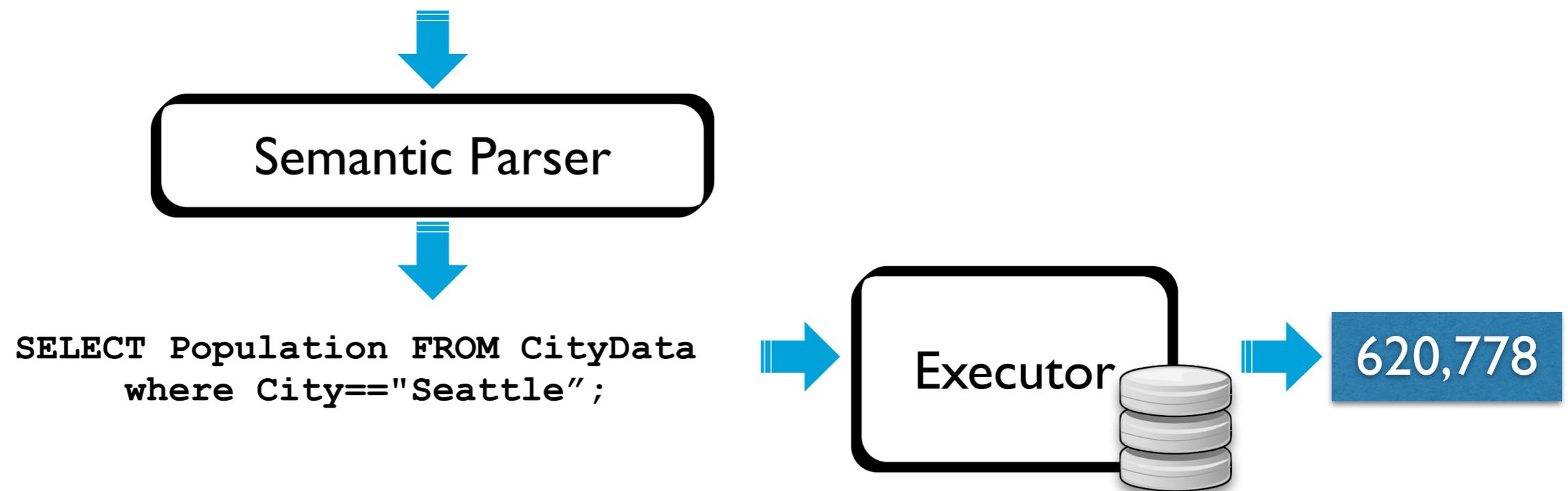


ACL 2015

# Semantic Parsing: Long-term Goal

Build meaning representations for  
open-domain texts

*How many people live in Seattle?*



(Kwiatkowski et.al, 13, Liang et.al. 11, Cai & Yates 2013, Berant et.al. 13,14, Reddy et. a. 14)

# Semantic Parsing: Large Domain

 **Freebase** is a large, community authored knowledge base with:

- 40 Million Entities
- 2 Billion Facts
- 20,000 Relations
- 10,000 Types
- 100 Domains



# Current semantic parsers

**can** parse...

*How many people live in Seattle?*

*Which college did Obama go to?*

*What party did Clay establish?*

# Current semantic parsers

**can** parse...

*How many people live in Seattle?*

*Which college did Obama go to?*

*What party did Clay establish?*

**cannot** parse...

*How many people live in Anyang?*

*Which college did Eunsol go to?*

*Who are Russian short story writers in 19th century?*

*What is a popular seaside resort city in Italy?*

# Remaining Challenges

## Fact Incompleteness

*How many people live in Anyang?*

*Which college did Eunsol go to?*

## Schema Incompleteness

*Who are Russian short story writers in 19th century?*

*What is a popular seaside resort city in Italy?*

# Remaining Challenges

## Fact Incompleteness

*How many people live in Anyang?*

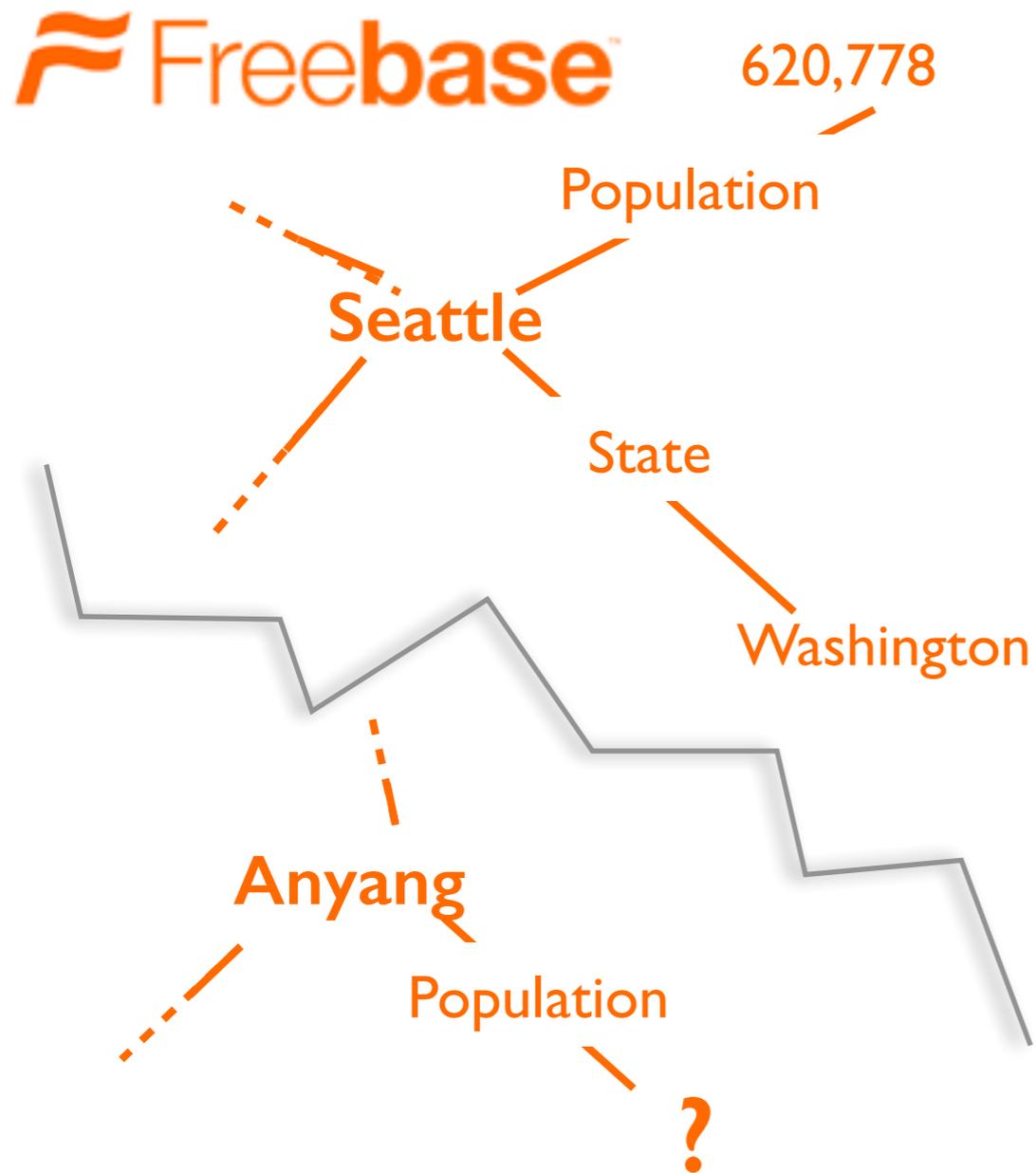
*Which college did Eunsol go to?*

## Schema Incompleteness

*Who are Russian short story writers in 19th century?*

*What is a popular seaside resort city in Italy?*

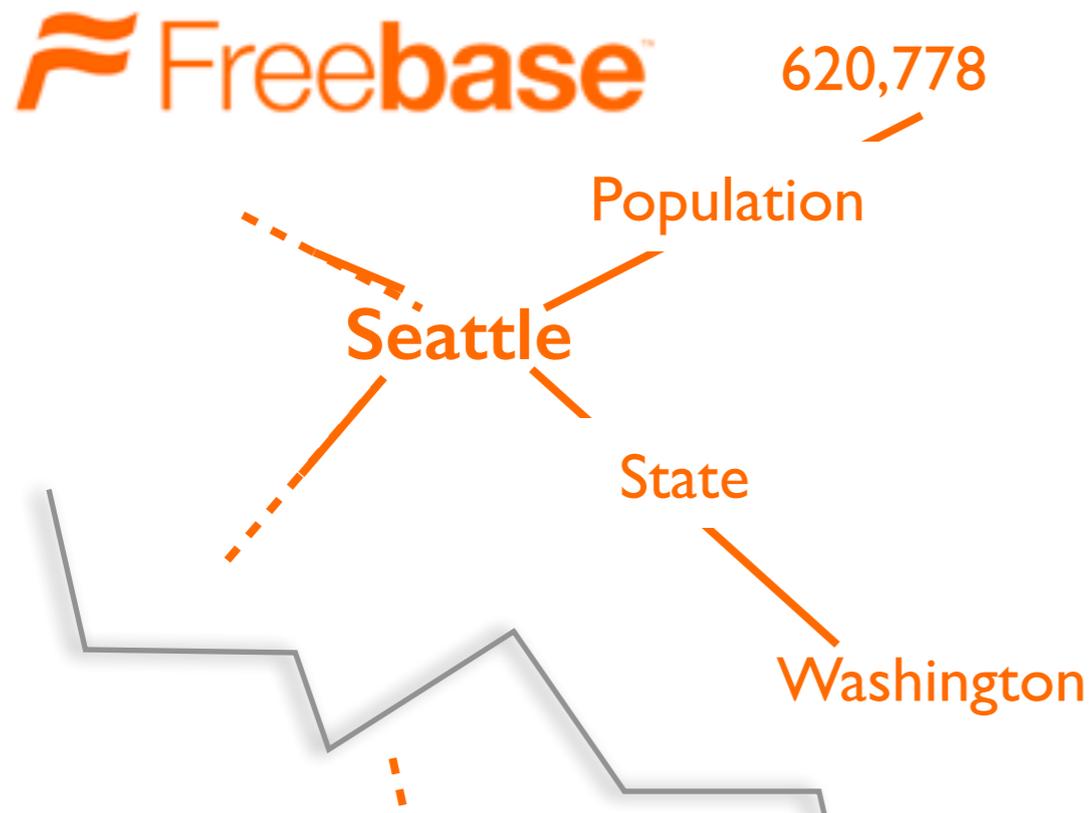
# Remaining Challenges: Fact Incompleteness



*How many people live in Anyang?  
Which college did Eunsol go to?*

Unable to handle sentences  
with facts not in Freebase

# Remaining Challenges: Fact Incompleteness



*How many people live in Anyang?*  
*Which college did Eunsol go to?*

**70%** of people in FB have no birth place (West' 14)  
**66%** of facts missing in pilot study for our dataset

# Remaining Challenges

Fact Incompleteness

*How many people live in Anyang*

*Which college did Eunsol*

Schema Incompleteness

*Who are Russian short story writers in 19th century?*

*What is a popular seaside resort city in Italy?*

# Remaining Challenges: Schema Incompleteness



short story writer|

Writer	Profession
--------	------------

Short story	Literary Genre
-------------	----------------

Author	Profession
--------	------------

*Who are Russian **short story** writers in 19th century?*

*What is **popular seaside resort city** in Italy?*

Unable to handle concepts outside existing schema

# Remaining Challenges: Schema Incompleteness

Freebase

short story writer|

Writer Profession

Short story Literary Genre

Author Profession

Who are Russian *short story writers* in 19th century?

What is *popular seaside resort city* in Italy?

In a pilot study on our dataset: **27.2%** of sentences describe concepts not in Freebase

# Previous Approach?

Existing data is filtered to ensure completeness:

- FB917 dataset is created from Freebase  
(Cai and Yates' 13)
- **93%** of originally gathered questions cannot be answered with FB (WebQuestions, Berant' 13)

# Remaining Challenges

Fact Incompleteness

New learning approach  
with broad coverage lexical statistics

Schema Incompleteness

*Who are Russian short story writers*

Semantic parser  
with partial groundings

*city in Italy?*

# This Work

Build meaning representations with both  
Freebase concepts and open concepts

*British playwright, novelist and short story writer*



Semantic Parser

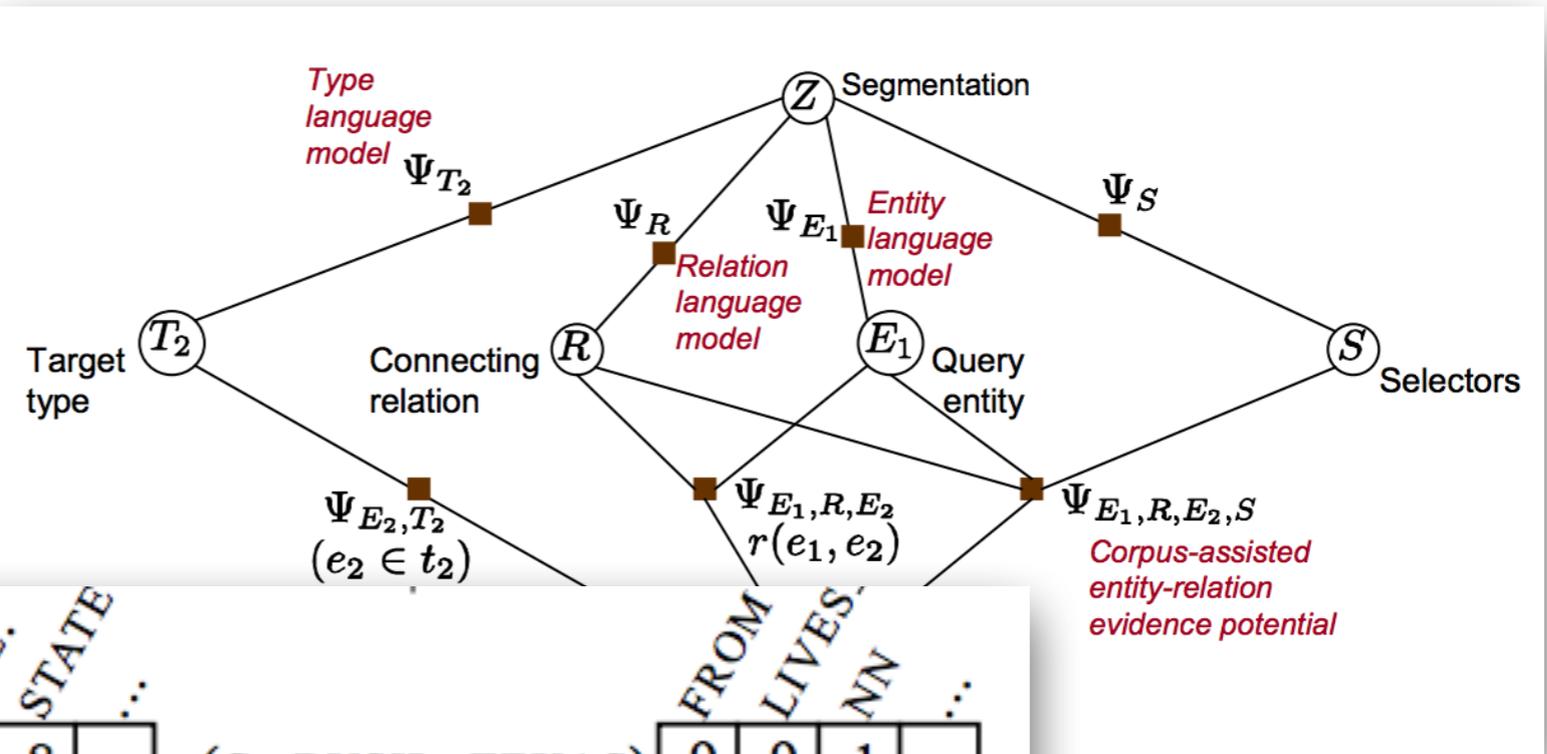


```
 $\lambda x.$ OpenRel_writer( $x$ , short_story)  
 $\wedge$  person.profession( $x$ , novelist)  
   $\wedge$  person.profession( $x$ , playwright)  
 $\wedge$  person.nationality( $x$ , united_kingdom)
```

# Parsing with Incompleteness

1. Open Information Extraction (Banko et al., 07; Fader et al., 11)
2. Matrix Factorization (Riedel, 13; Krishnamurthy, 15)
3. Web Search Queries (Joshi, 14)

(<proper noun>, acquired, <proper noun>)  
 (<proper noun>, graduated from, <proper noun>)  
 (<proper noun>, is author of, <proper noun>)  
 (<proper noun>, is based in, <proper noun>)  
 (<proper noun>, studied, <noun phrase>)  
 (<proper noun>, studied at, <proper noun>)  
 (<proper noun>, was developed by, <proper noun>)  
 (<proper noun>, was formed in, <year>)  
 (<proper noun>, was founded by, <proper noun>)  
 (<proper noun>, worked with, <proper noun>)



	F.-RUN	POL.	STATE	...
TEXAS	.1	.1	.8	
REPub.	.1	.1	.1	
G._BUSH	.9	.9	.1	
...				

	FROM	LIVES	NN	...
(G._BUSH, TEXAS)	.9	.9	.1	
(G._BUSH, REPUB.)	.1	.1	.1	
(REPub., G._BUSH)	.1	.1	.7	
...				

# Outline

*1. Task and Applications*

2. Data

3. Semantic Parser with Partial Ontology

4. Learning

5. Evaluation

# Task

Build a meaning representation with  
**Freebase** concepts and **concepts**  
**outside Freebase**

*British playwright, novelist and short story writer*

Semantic Parser

```
 $\lambda x.$ OpenRel_writer( $x$ , short_story)  
 $\wedge$  person.profession( $x$ , novelist)  
 $\wedge$  person.profession( $x$ , playwright)  
 $\wedge$  person.nationality( $x$ , united_kingdom)
```

# Focus: Noun Phrases



**William Somerset Maugham** CH ([/ˈmɔːm/](#) *MAWM*; 25 January 1874 – 16 Dec 1965) was a British **playwright, novelist and short story** writer. He was among

Alyse Tadajewski, **a spokeswoman for Fiat Chrysler,** said that the company did not believe it was responsible for the researchers to

- Interesting noun-noun modifier, implicit relations.
- Itself is a referring expression, resembling queries.

# Focus: Noun Phrases



**William Somerset Maugham** CH ([/ˈmɔːm/ MAWM](#); 25 January 1874 – 16 Dec 1965) was a British **playwright**, **novelist** and **short story** writer. He was among

**Alyse Tadajewski** a spokeswoman for Fiat Chrysler, said that the company did not believe it was responsible for the researchers to

- Interesting noun-noun modifier, implicit relations.
- Useful for information extraction, when paired with an entity.

# Applications

## Referring Expression Resolution (QA)

*Input:* Noun Phrase

*British playwright, novelist and short  
story writer*

*Output:* **Sommerset Maugham**

## Entity Attribute Extraction (IE)

(Entity, Noun Phrase)

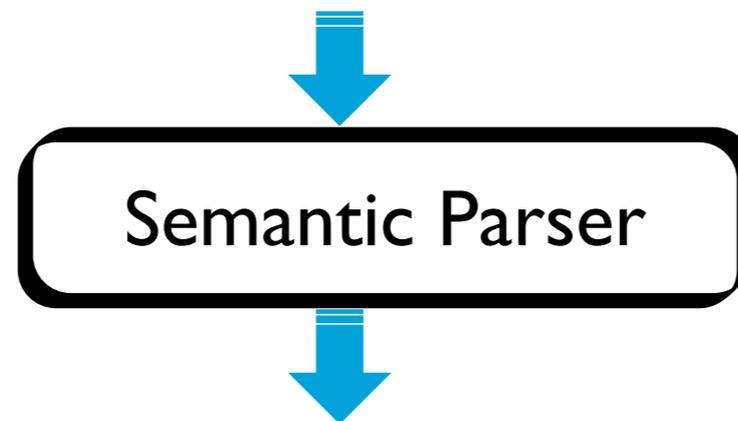
**Somerset** *British playwright, novelist*  
**Maugham** ' *and short story writer*



(S. Maugham, Nationality, U.K)  
(S. Maugham, Profession, Novelist)  
(S. Maugham, Profession, Playwright)

# Overview: Approach

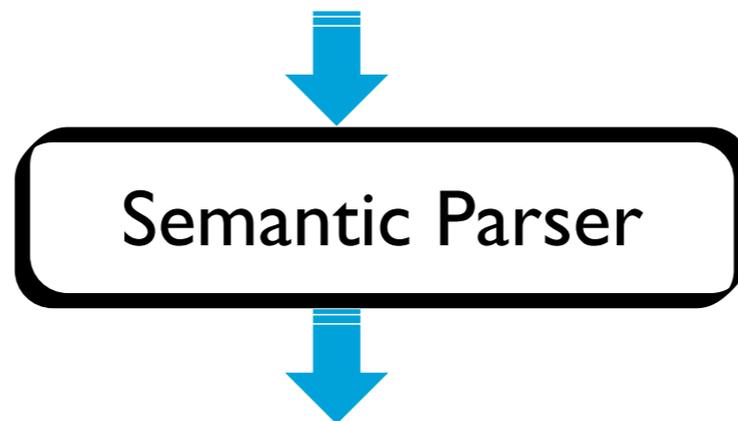
*British playwright, novelist and short story writer*



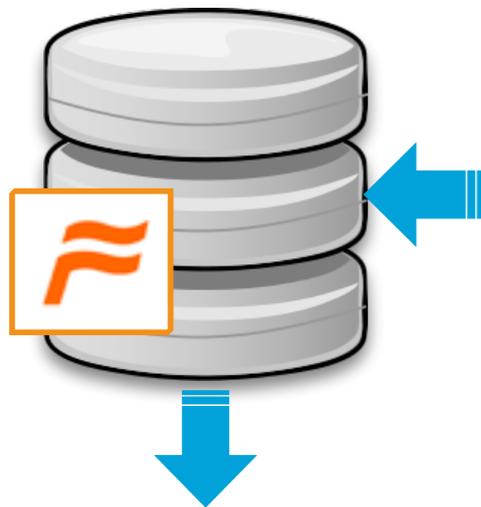
```
 $\lambda x.$ OpenRel_writer( $x$ , short_story)  
 $\wedge$  person.profession( $x$ , novelist)  
 $\wedge$  person.profession( $x$ , playwright)  
 $\wedge$  person.nationality( $x$ , united_kingdom)
```

# Referring Expression Resolution (QA)

*British playwright, novelist and short story writer*



```
 $\lambda x.$ OpenRel_writer( $x$ , short_story)  
 $\wedge$  person.profession( $x$ , novelist)  
 $\wedge$  person.profession( $x$ , playwright)  
 $\wedge$  person.nationality( $x$ , united_kingdom)
```

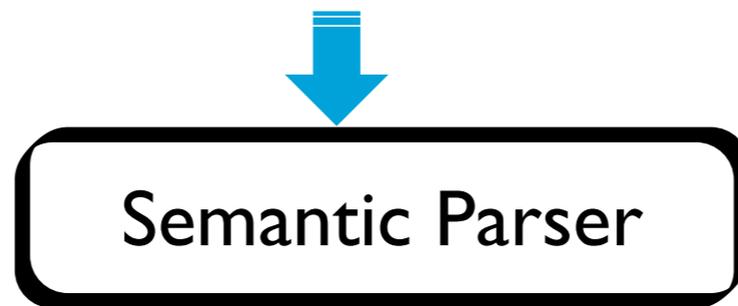


Somerset Maugham

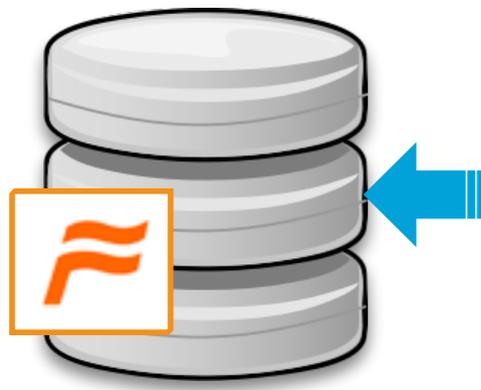
# Referring Expression Resolution (QA)

# Entity Attribute Extraction (IE)

*British playwright, novelist and short story writer*



$\lambda x.$ OpenRel\_writer( $x$ , short\_story)  
 $\wedge$  person.profession( $x$ , novelist)  
 $\wedge$  person.profession( $x$ , playwright)  
 $\wedge$  person.nationality( $x$ , united\_kingdom)



Somerset Maugham



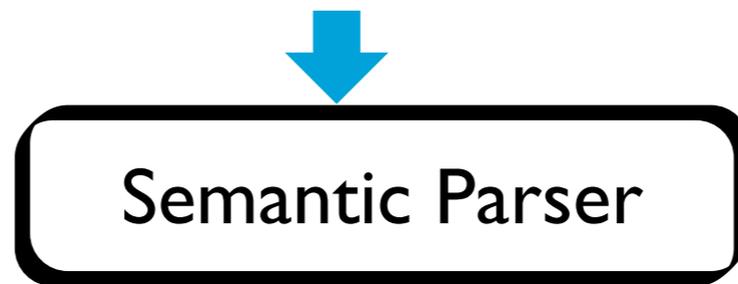
( $x$ , Nationality, U.K)  
( $x$ , Profession, Novelist)  
( $x$ , Profession, Playwright)

# Referring Expression Resolution (QA)

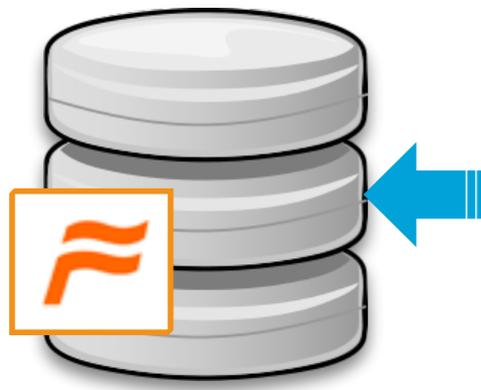
*Somerset  
Maugham*

*, British playwright, novelist and short story writer*

# Entity Attribute Extraction (IE)



```
 $\lambda x.$ OpenRel_writer( $x$ , short_story)  
 $\wedge$  person.profession( $x$ , novelist)  
 $\wedge$  person.profession( $x$ , playwright)  
 $\wedge$  person.nationality( $x$ , united_kingdom)
```



```
(S.Maugham, Nationality, U.K)  
(S.Maugham, Profession, Novelist)  
(S.Maugham, Profession, Playwright)
```

Sommerset Maugham

# Outline

1. Task and Applications

**2. *Data***

3. Semantic Parser with Partial Ontology

4. Learning

5. Evaluation

# Wikipedia Category

## Philip Seymour Hoffman

From Wikipedia, the free encyclopedia

**Philip Seymour Hoffman** (July 23, 1967 – February 2, 2014) was an American actor and director. He was prolific in both film and theater from the early 1990s until his death at age 46, after which *The New York Times* declared him "perhaps the most ambitious and widely admired American actor of his generation".<sup>[1]</sup>

Raised in [Fairport, New York](#), Hoffman studied acting at the [New York State Summer School of the Arts](#) and the [Tisch School of the Arts](#). He began his career in 1991 playing a defendant in a rape case in an episode of *Law &*

**Philip Seymour Hoffman**



# Wikipedia Category

## Philip Seymour Hoffman

From Wikipedia, the free encyclopedia

**Philip Seymour Hoffman** (July 23, 1967 – February 2, 2014) was an American actor and director. He was prolific in both film and theater from the early 1990s until his death at age 46, after which *The New York Times*

Philip Seymour Hoffman



dec  
act  
Rai  
Sun  
care

Categories: [1967 births](#) | [2014 deaths](#) | [20th-century American male actors](#)  
| [21st-century American male actors](#) | [American film producers](#) | [American film directors](#)  
| [American male film actors](#) | [American male stage actors](#) | [American male television actors](#)  
| [American male voice actors](#) | [American people of Dutch descent](#)  
| [American people of English descent](#) | [American people of German descent](#)  
| [American people of Irish descent](#) | [American theatre directors](#) | [BAFTA winners \(people\)](#)  
| [Best Actor Academy Award winners](#) | [Best Actor BAFTA Award winners](#)  
| [Best Drama Actor Golden Globe \(film\) winners](#) | [Circle in the Square Theatre School alumni](#)  
| [Cocaine-related deaths in New York](#) | [Drug-related deaths in New York](#)  
| [Film directors from New York](#) | [Film directors from New York City](#)

# Wikipedia Category

## Philip Seymour Hoffman

From Wikipedia, the free encyclopedia

**Philip Seymour Hoffman** (July 23, 1967 – February 2, 2014) was an American actor and director. He was prolific in both film and theater from the early 1990s until his death at age 46, after which *The New York Times*

Philip Seymour Hoffman



dec  
act  
Rai  
Sun  
care

Categories: [1967 births](#) | [2014 deaths](#) | [20th-century American male actors](#)  
[21st-century American male actors](#) | [American film producers](#) | [American film directors](#)  
[American male film actors](#) | [American male stage actors](#) | [American male television actors](#)  
[American male voice actors](#) | [American people of Dutch descent](#)  
[American people of English descent](#) | [American people of German descent](#)  
[American people of Irish descent](#) | [American theatre directors](#) | [BAFTA winners \(people\)](#)  
[Best Actor Academy Award winners](#) | [Best Actor BAFTA Award winners](#)  
[Best Drama Actor Golden Globe \(film\) winners](#) | [Circle in the Square Theatre School alumni](#)  
[Cocaine-related deaths in New York](#) | [Drug-related deaths in New York](#)

[Film directors from New York](#)

[from New York City](#)

# Wikipedia Category

## Philip Seymour Hoffman

From Wikipedia, the free encyclopedia

**Philip Seymour Hoffman**  
American actor and  
the early 1990s until

dec  
act  
Rai  
Sun  
care

### Categories

- 21st-cent
- American
- American
- American
- American
- Best Acto
- Best Drar
- Cocaine-

## Pages in category "Film directors from New York City"

The following 196 pages are in this category, out of 196 total. This list may not reflect recent c

### A

- George Abbott
- J. J. Abrams
- Ivan Abramson
- Jus Addiss
- John G. Adolfi
- Irwin Allen
- Woody Allen
- Roy Lee
- Spike Lee
- Marc Levin
- Jonathan Levine
- Joseph H. Lewis
- Doug Liman
- Victoria Linchong
- Fred J. Lincoln
- Steven Lisberger

[Film directors from New York](#)

from New York City

# Wikipedia Category

## Philip Seymour Hoffman

From Wikipedia, the free encyclopedia

Philip Seymour Hoffman

American

the early

dec

act

Rais

Sun

care

On average, **15%** entity overlap with Freebase  
Exciting opportunity for information extraction  
Challenge for existing learning techniques

[Film directors from New York](#)

from New York City

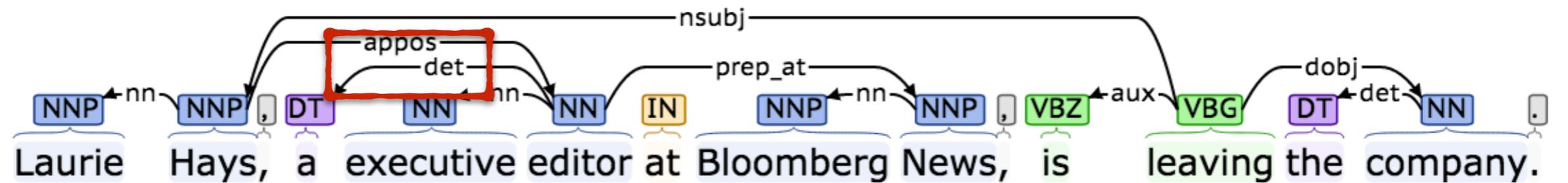
# Wikipedia Category: Data statistics

	Entire Set
Number of Category	<b>365 K</b>
Number of words per category	<b>4.1</b>
Number of entity-category pair	<b>7 million</b>

# Appositives

Relation between a named entity and a nominal.

- ▶ Laurie Hays, a executive editor at Bloomberg News, is leaving the company.



Laurie Hays ,

a executive editor at Bloomberg  
News

# Appositives

Extracted from open texts such as news articles

- Malta, an EU outpost in the Mediterranean, decided today...
- Richard Nixon, a former president of the United States...
- Maputo, the relaxed seaside capital of Mozambique, ...

# Appositives

Extracted from open texts such as news articles

- Malta, **an EU outpost in the Mediterranean**, decided today...
- Richard Nixon, a **former president of the United States**...
- Maputo, **the relaxed seaside capital of Mozambique**, ...

# Appositives: Data statistics

	Entire Set
Number of apposition	67 K
vocab	25 K
Number of words per apposition	5.73

# Outline

1. Task and Applications

2. Data

***3. Semantic Parser with Partial Ontology***

4. Learning

5. Evaluation

# Two Stage Semantic Parsing (EMNLP 13)

British playwright, novelist and short story writer

Domain Independent Parse

Ontology Match

# Two Stage Semantic Parsing (EMNLP 13)

British playwright, novelist and short story writer

Domain Independent Parse

$\lambda x.$        $\text{British}(x) \wedge \text{Playwright}(x)$   
              $\text{Novelist}(x) \wedge \text{Writer}(x, \text{Short Story})$

Ontology Match

# Two Stage Semantic Parsing (EMNLP 13)

British playwright, novelist and short story writer

Domain Independent Parse

$$\lambda x. \quad \text{British}(x) \wedge \text{Playwright}(x) \\ \text{Novelist}(x) \wedge \text{Writer}(x, \text{Short Story})$$

Ontology Match

$$\lambda x. \quad \text{person.profession}(x, \text{novelist}) \\ \wedge \text{person.profession}(x, \text{playwright}) \\ \wedge \text{person.nationality}(x, \text{united\_kingdom})$$

# Two Stage Semantic Parsing with Partial Grounding

British playwright, novelist and short story writer

Domain Independent Parse

$$\lambda x. \quad \text{British}(x) \wedge \text{Playwright}(x) \\ \text{Novelist}(x) \wedge \text{Writer}(x, \text{Short Story})$$

Ontology Match

$$\lambda x. \text{OpenRel\_writer}(x, \text{short\_story}) \\ \wedge \text{person.profession}(x, \text{novelist}) \\ \wedge \text{person.profession}(x, \text{playwright}) \\ \wedge \text{person.nationality}(x, \text{united\_kingdom})$$

# Partial Grounding: Open Schema

Explicitly model concepts not in Knowledge base  
as **OpenRel** and **OpenType**

Plants described in 1891

```
λx. Lower_classification(plant, x)  
OpenRel_described_in(x, 1891)
```

Former municipalities in Brandenburg

```
λx. OpenType_Former(x)  
OpenRel(x, Municipality)  
Located_In(x, Brandenburg)
```

# Partial Grounding: Open Schema

Benefits of open schema:

- Help learn String-Freebase concepts
- Allow partial execution
- Capture useful information, although not grounded

(**x**, **Located\_In**, **Brandenburg**)

# Two Stage Semantic Parsing with Partial Grounding

## Domain Independent Parse

$$\begin{array}{c}
 \begin{array}{ccc}
 \text{Former} & \text{municipalities} & \text{in} & \text{Brandenburgh} \\
 \hline
 N/N & N & N \setminus N/NP & NP \\
 \lambda f \lambda x. f(x) \wedge \text{former}(x) & \lambda x. \text{municipalities}(x) & \lambda f \lambda x \lambda y. f(y) \wedge \text{in}(y, x) & \text{Brandenburg} \\
 \hline
 & N & N \setminus N & \\
 \lambda x. \text{former}(x) \wedge \text{municipalities}(x) & & \lambda f \lambda y. f(y) \wedge \text{in}(y, \text{Brandenburg}) & \\
 \hline
 & N & & \\
 l_0 = \lambda x. \text{former}(x) \wedge \text{municipalities}(x) \wedge \text{in}(x, \text{Brandenburg}) & & & 
 \end{array}
 \end{array}$$

## Ontology Match

Structure Match    Constant Matches for .    OPEN Constant Matches

$$\lambda x. \text{OpenType}(x) \wedge \text{OpenRel}(x, \text{Municipality}) \wedge \text{Location.ContainedBy}(x, \text{Brandenburg})$$

# Outline

1. Task and Applications
2. Data
3. Semantic Parser with Partial Ontology
- 4. *Learning***
5. Evaluation

# Previous Work: Direct Supervision

How many people live in Seattle?

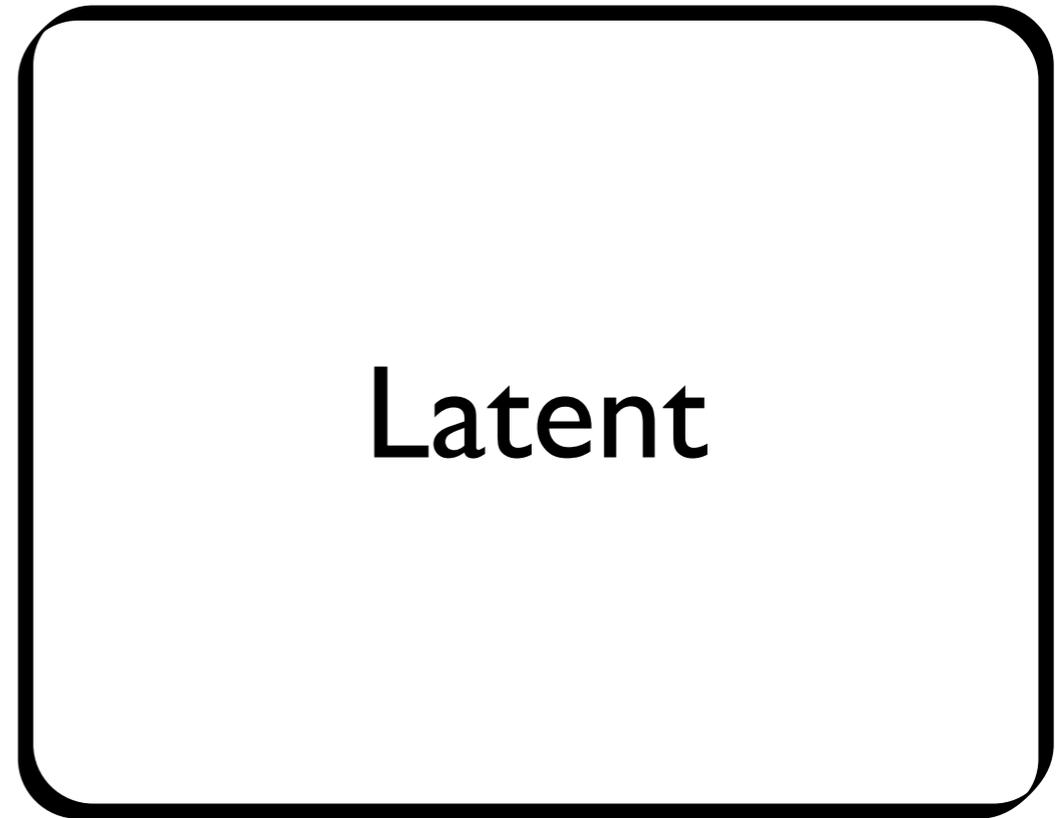


```
SELECT Population FROM CityData  
where City=="Seattle";
```



620,778

How many people live in Seattle?



620,778

# Supervision from Unfiltered Data

## Pages in category "Social democratic parties in Greece"

The following 23 pages are in this category, out of 23 total. This list may not reflect recent changes ([learn more](#)).

### A

- [Agreement for the New Greece](#)
- [Agricultural and Labour Party](#)

### D

- [Democratic Left \(Greece\)](#)
- [Democratic Regional Union](#)
- [Democratic Socialist Party of Greece](#)

### F

- [Free Citizens](#)

### G

- [Georgios Papandreou Party](#)

### M

- [Movement of Democratic Socialists](#)

### O

- [Olive Tree \(Greece\)](#)

### P

- [Panhellenic Citizen Chariot](#)
- [Panhellenic Socialist Movement](#)
- [Party of Democratic Socialism \(Greece\)](#)
- [Popular Unions of Bipartisan Social Groups](#)

### R

- [Radical Movement of Social Democratic Allia](#)

# Supervision from Unfiltered Data

Social democratic parties in Greece



$\lambda x.\text{political\_party.ideology}(x, \text{social\_democratic})$   
 $\wedge \text{political\_party.country}(x, \text{greece})$



~~{Agreement for the New Greece, Agricultural and Labour Party, Free Citizens, Democratic Social Movement....}~~

4 entities

# Supervision from Unfiltered Data

Social democratic parties in Greece



Semantic Parser

Missing facts make direct supervision difficult.

Executor



~~{Agreement with the New Greece, Agricultural and Labour Party, Free Citizens, Democratic Social Movement....}~~

4 entities

# Supervision from Unfiltered Data

Social democratic parties in Greece



Semantic Parser



```
 $\lambda x.$ political_party.ideology( $x$ , social_democratic)  
 $\wedge$ political_party.country( $x$ , greece)
```

# Supervision from Unfiltered Data

Social democratic parties in Greece



Semantic Parser



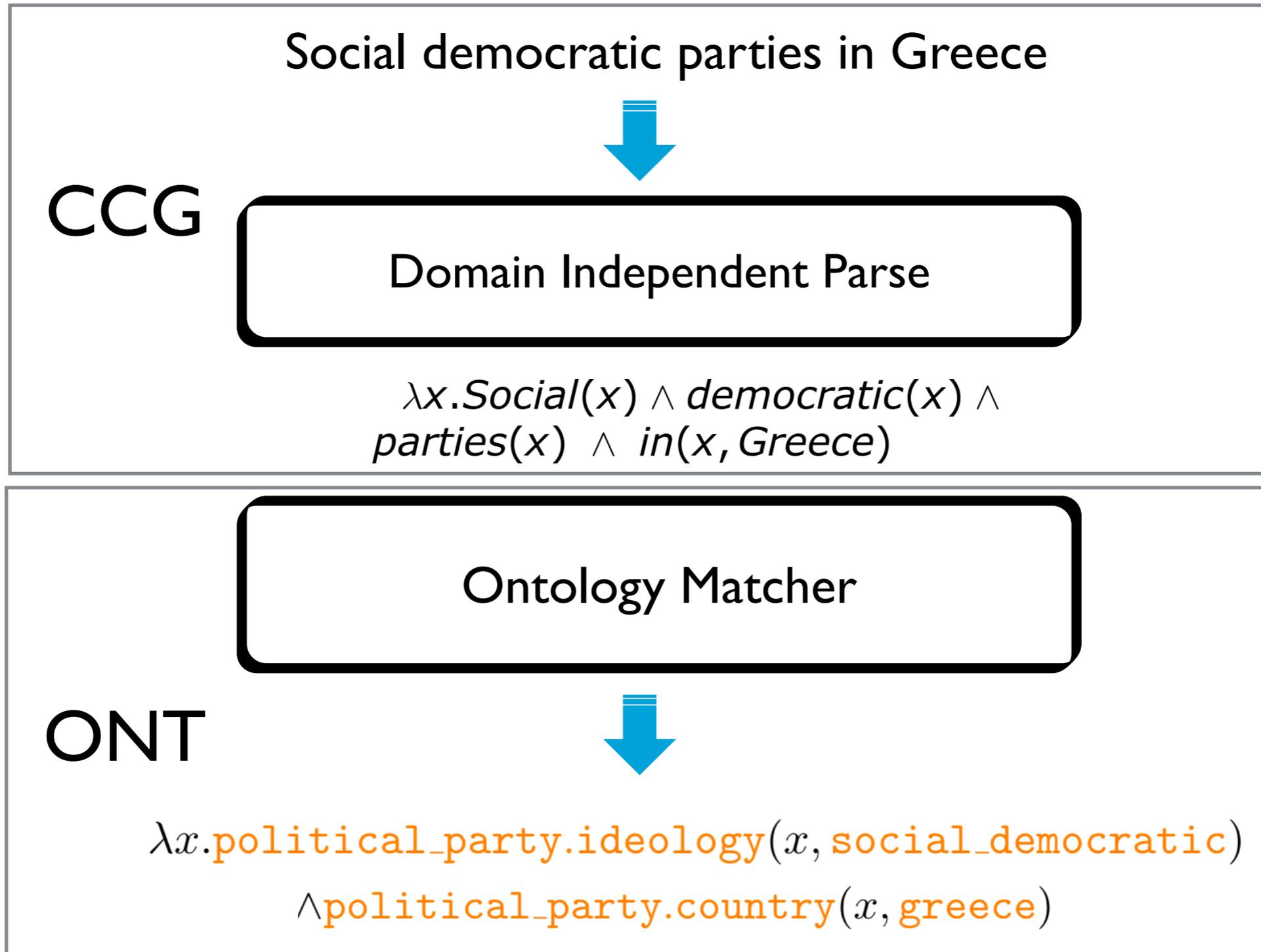
```
 $\lambda x.$ political_party.ideology( $x$ , social_democratic)  
 $\wedge$ political_party.country( $x$ , greece)
```

Gold mapping is expensive to gather large scale.

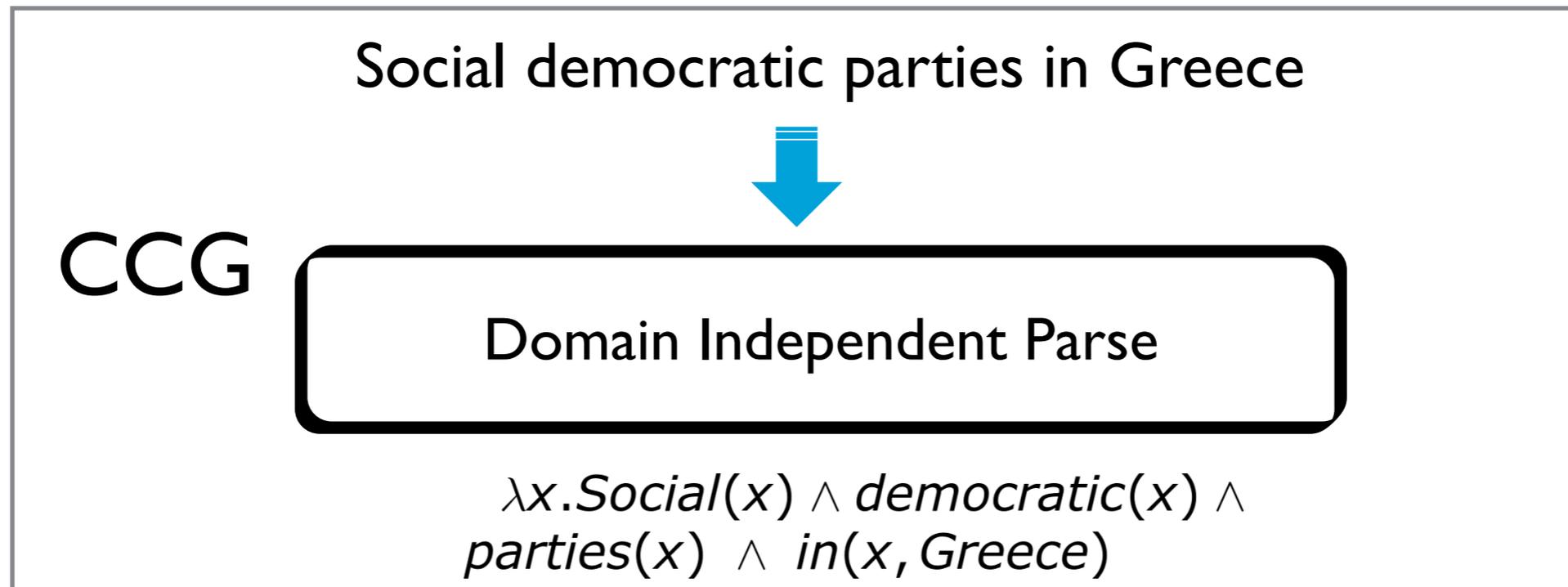
# Learning with Fact Incompleteness

1. Two-Stage Learning
2. Two kinds of data
  - i. Small Annotated Dataset
  - ii. Broad Coverage Lexical Statistics

# Two-Stage Learning



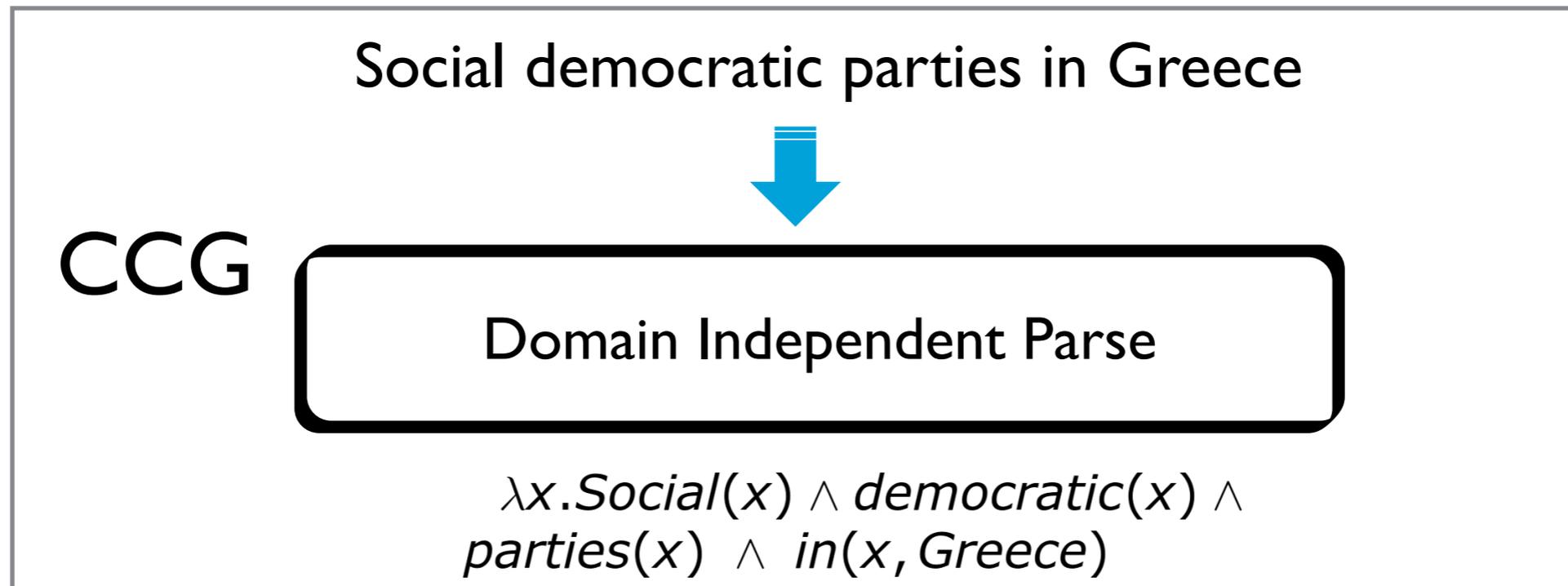
# Two-Stage Learning



Small training set with logical form

( British playwright, novelist and short story writer ,  $\lambda x. \text{OpenRel\_writer}(x, \text{short\_story}) \wedge \text{person.profession}(x, \text{novelist}) \wedge \text{person.profession}(x, \text{playwright}) \wedge \text{person.nationality}(x, \text{united\_kingdom})$  )  $\times 500$

# Two-Stage Learning



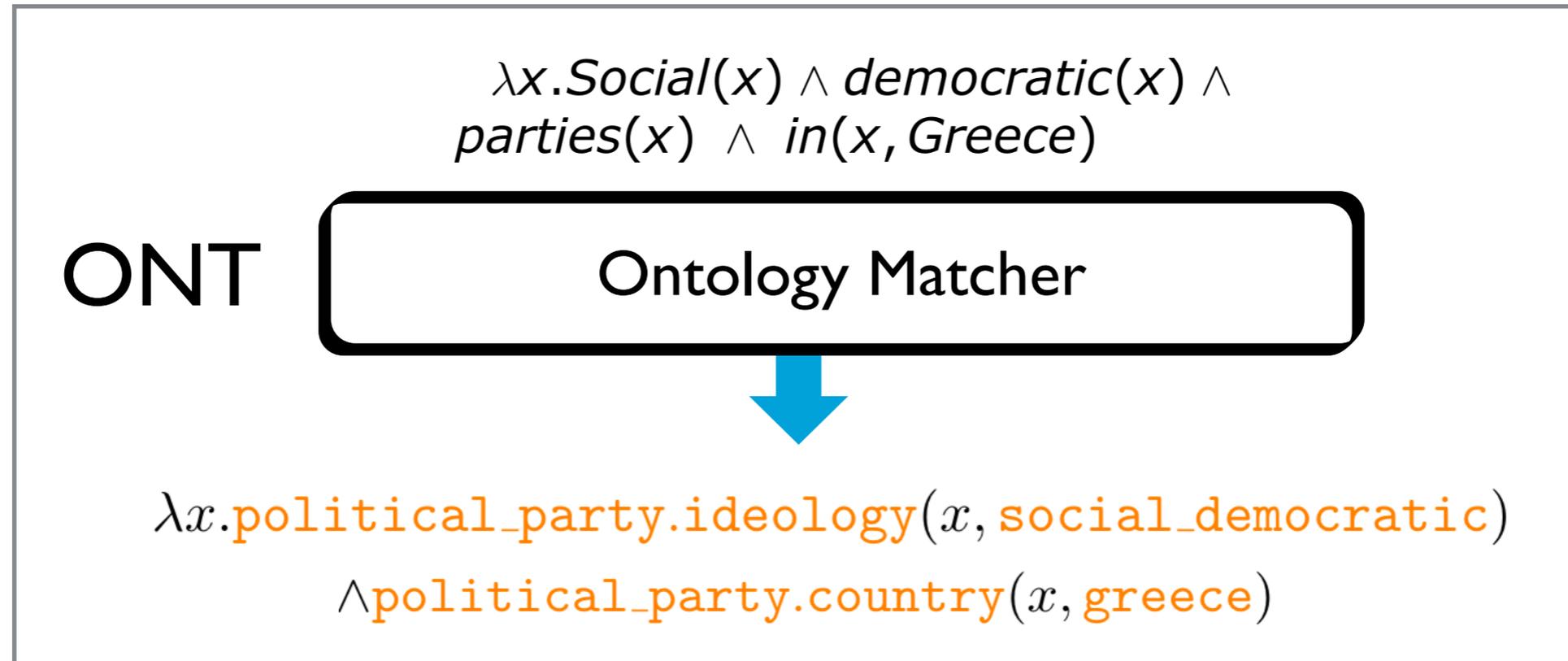
Derivations  $l$  are scored using a linear model

Highest scoring logical form ( $l^*$ ) is passed to the second stage

$$\text{score}(l) = \phi_{ccg}(l) \cdot \theta$$

$$l^* = \arg \max_{l \in \text{CCG}(x)} (\theta_p \cdot \phi_p(l))$$

# Two-Stage Learning



 **Freebase** is a large, community authored knowledge base with:

- 20,000 Relations
- 10,000 Types
- 100 Domains

# Broad Coverage Lexical Statistics : Wikipedia Category dataset

Philip Seymour Hoffman

From Wikipedia, the free encyclopedia

Philip

Ameri

the ea

dec

acto

Rais

Sun

care

85 K vocabulary

365 K category

2.5 M entity

7 M category-entity pair

an

rs

actors

ple)

alumni

[Cocaine-related deaths in New York](#) | [Drug-related deaths in New York](#)

[Film directors from New York](#) | [Film directors from New York City](#)

# Mapping words to Freebase Attribute

Pixar Feature Films  
Animation Films from Pixar  
Pixar songs

# Mapping words to Freebase Attribute

Pixar Feature Films

Animation Films from Pixar

Pixar songs

Ratatouille

Wall-E

Finding Nemo

Toy Story

Monster Inc.

Just keep swimming

# Mapping words to Freebase Attribute

Pixar Feature Films

Animation Films from Pixar

Pixar songs

Ratatouille

Wall-E

Finding Nemo

Toy Story

Monster Inc.

Just keep swimming

**(film.production\_companies, Pixar)**

(film.film.produced\_by, John Lasseter)

(film.film.directed\_by, John Lasseter)

(film.film.film\_festivals, 2011 Anima Mundi )

(film.film.starring.actor, Bob Peterson)

# Mapping words to Freebase Attribute

## Pixar Feature Films

A

Large amount of information  
for String - Entity - Freebase attribute alignment

Toy Story	(film.directed_by, John Lasseter)
Monster Inc.	(film.film_festivals, 2011 Anima Mundi)
Just keep swimming	(film.starring_actor, Bob Peterson)

# Mapping words to Freebase Attribute

Pointwise Mutual Information(PMI)=

$$\log\left(\frac{P(\text{String, Freebase Attribute})}{P(\text{String}) P(\text{Freebase Attribute})}\right)$$

as a feature

# Features

Domain  
Independent Parse

Parse Features:  
CCG Lexicon, Capitalization

Ontology Match

String -> Freebase features  
Wikipedia Lexical Statistics

Surface Lexical Features  
String Match, Stem Match

KnowledgeBase Features

# Outline

1. Task and Applications
2. Data
3. Semantic Parser with Partial Ontology
4. Learning
- 5. Evaluation***

# Experimental Setup

Training Set:

500 annotated Wikipedia Category

Test Set (Manual Evaluation):

500 unseen Wikipedia Category

300 appositives

Baseline:

SVM Classifier trained with annotated logical forms

# Applications

## Referring Expression Resolution (QA)

*Input:* Noun Phrase

*British playwright, novelist and short  
story writer*

*Output:*  $\lambda x. \text{OpenRel\_writer}(x, \text{short\_story})$   
 $\wedge \text{person.profession}(x, \text{novelist})$   
 $\wedge \text{person.profession}(x, \text{playwright})$   
 $\wedge \text{person.nationality}(x, \text{united\_kingdom})$

**Sommerset Maugham**

## Entity Attribute Extraction (IE)

(Entity, Noun Phrase)

*Somerset British playwright, novelist  
Maugham ' and short story writer*

## Entity attributes for Freebase

S. Maugham	Nationality	U.K
S. Maugham	Profession	Novelist
S. Maugham	Profession	Playwright

# Evaluation Metric: Referring Expression Resolution

Alternative Rock Groups from Nevada

Gold

music.group(x)  $\wedge$   
music.artist.origin(x, NEVADA)  $\wedge$   
music.genre.artist(alternative rock, x)

✗

✓

✗

Output

music.artist.origin(x, NEVADA)  $\wedge$   
music.genre.artist(hard rock, x)

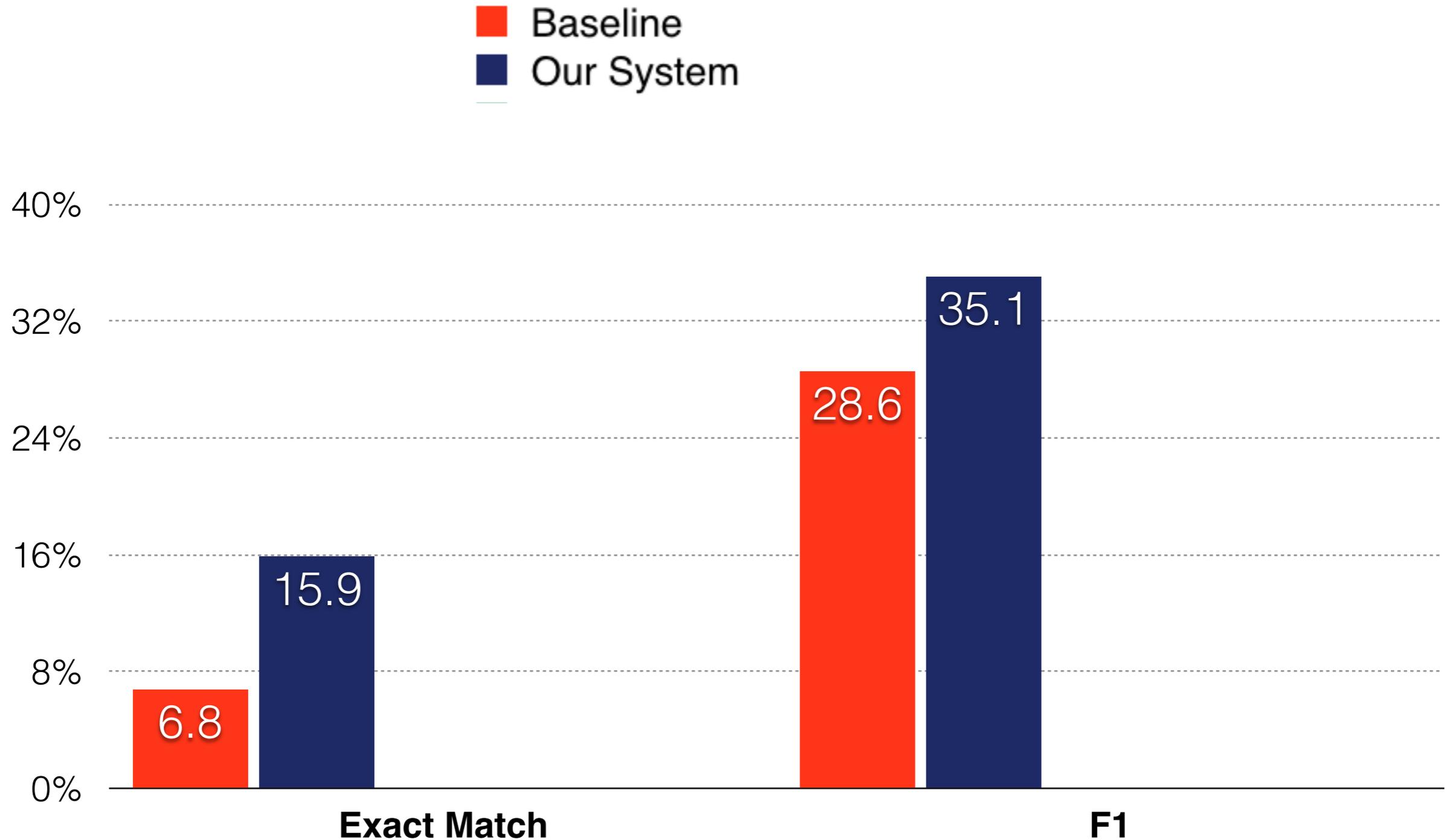
✓

✗

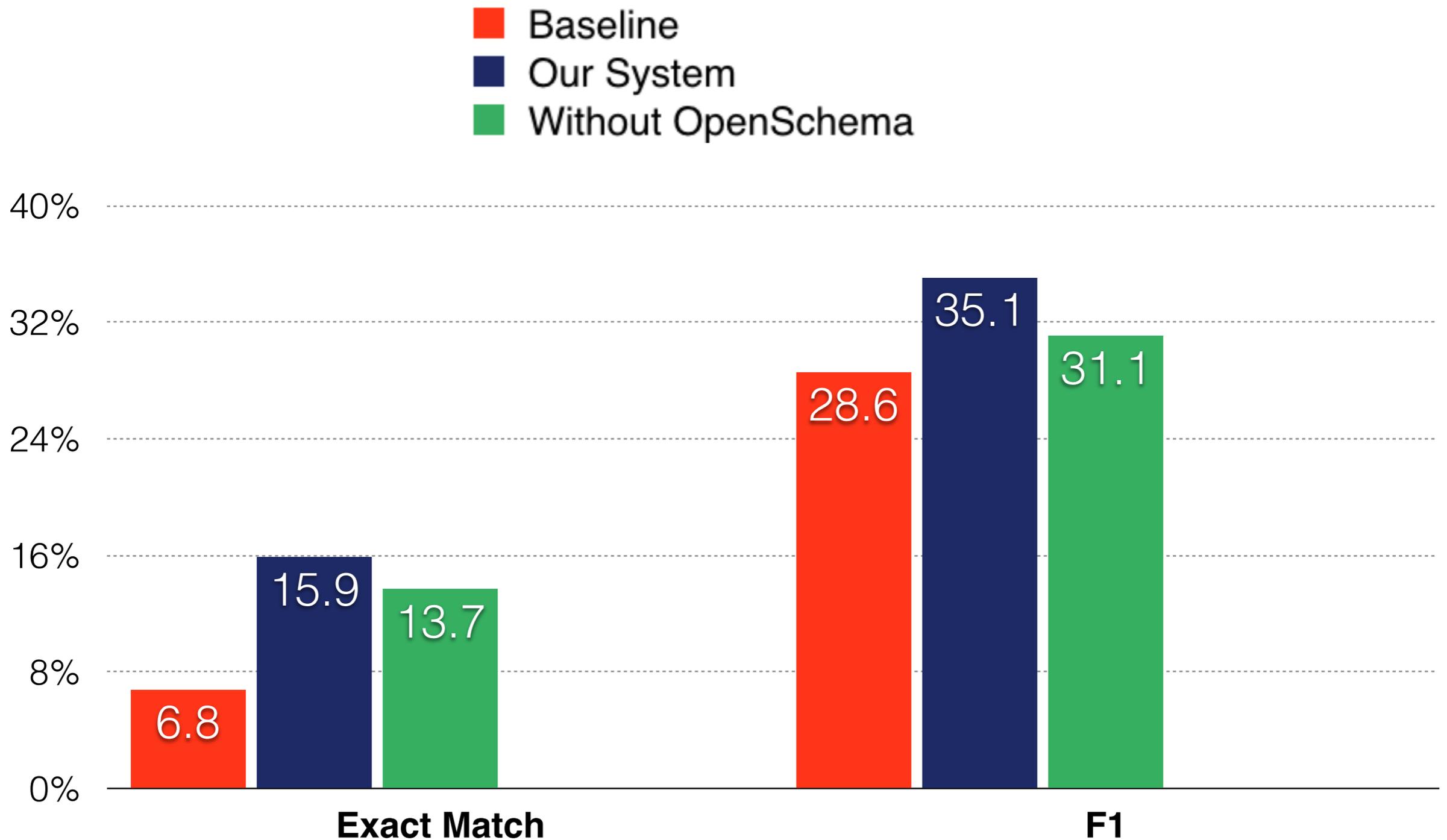
Precision: 0.5      Recall: 0.3      F1:0.375

Exact Match: **False**

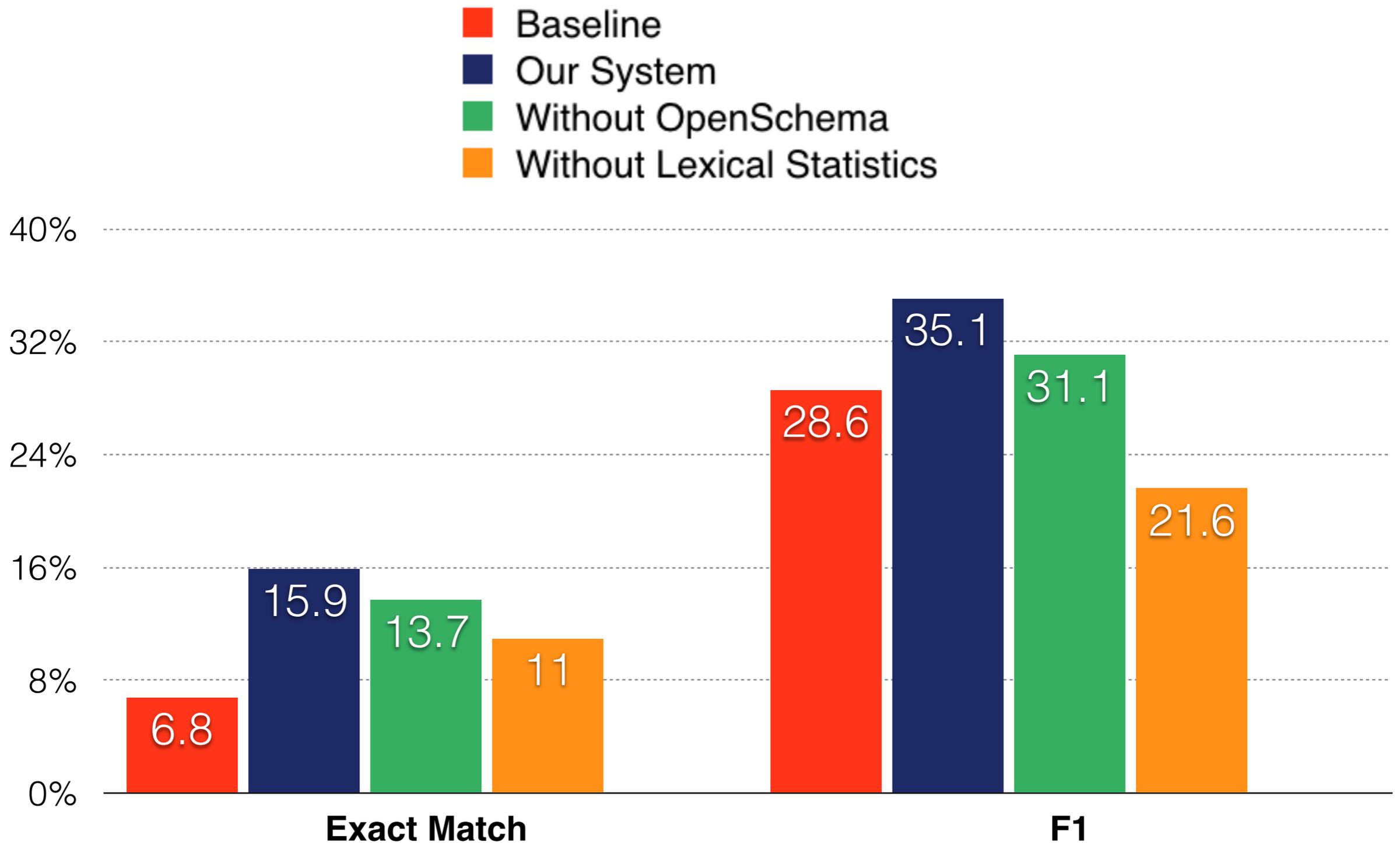
# Referring Expression Resolution: (5 fold cross validation on Training Set)



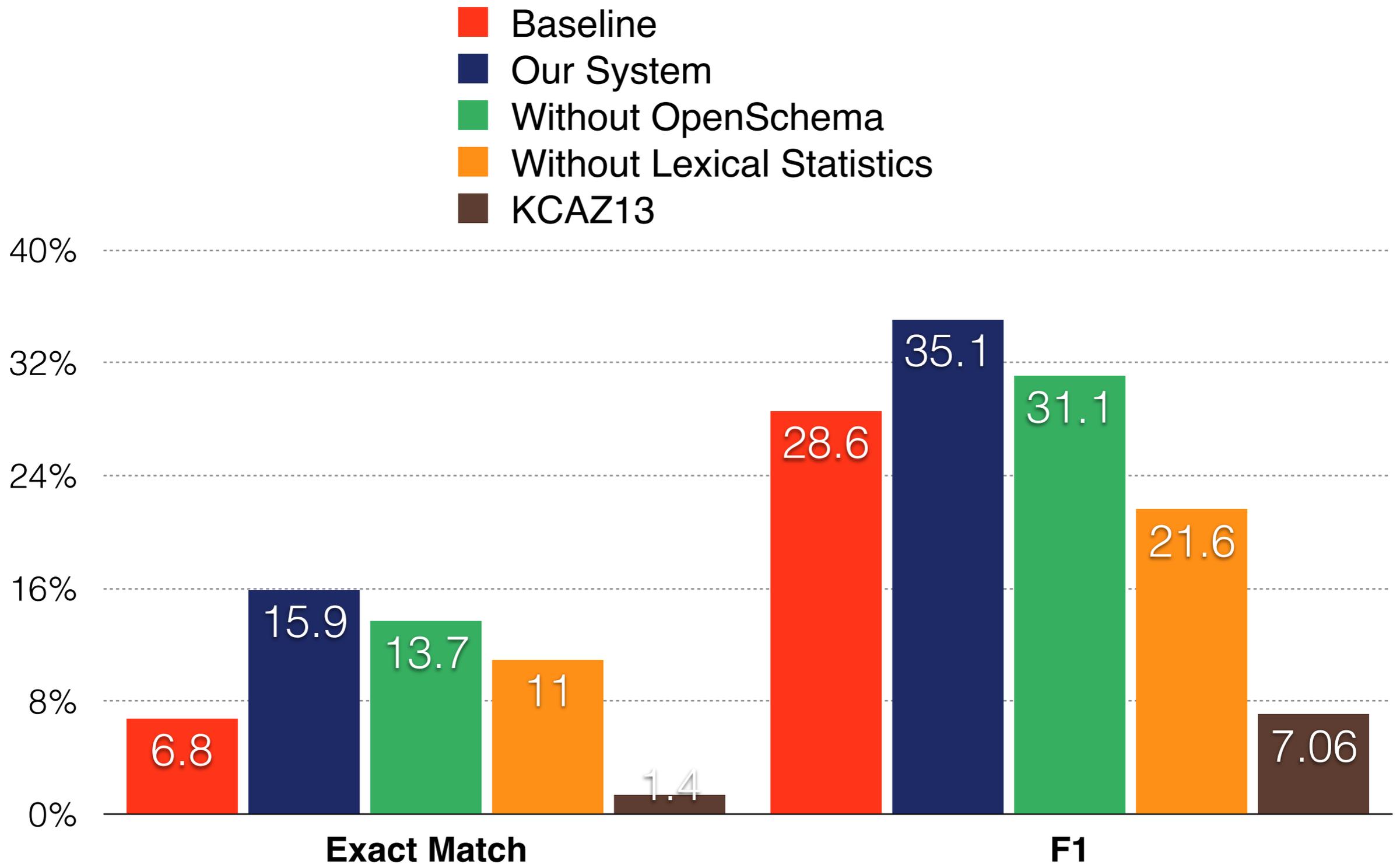
# Referring Expression Resolution: (5 fold cross validation on Training Set)



# Referring Expression Resolution: (5 fold cross validation on Training Set)



# Referring Expression Resolution: (5 fold cross validation on Training Set)



# Applications

## Referring Expression Resolution (QA)

*Input:* Noun Phrase

*British playwright, novelist and short  
story writer*

*Output:*  $\lambda x. \text{OpenRel\_writer}(x, \text{short\_story})$   
 $\wedge \text{person.profession}(x, \text{novelist})$   
 $\wedge \text{person.profession}(x, \text{playwright})$   
 $\wedge \text{person.nationality}(x, \text{united\_kingdom})$

**Sommerset Maugham**

## Entity Attribute Extraction (IE)

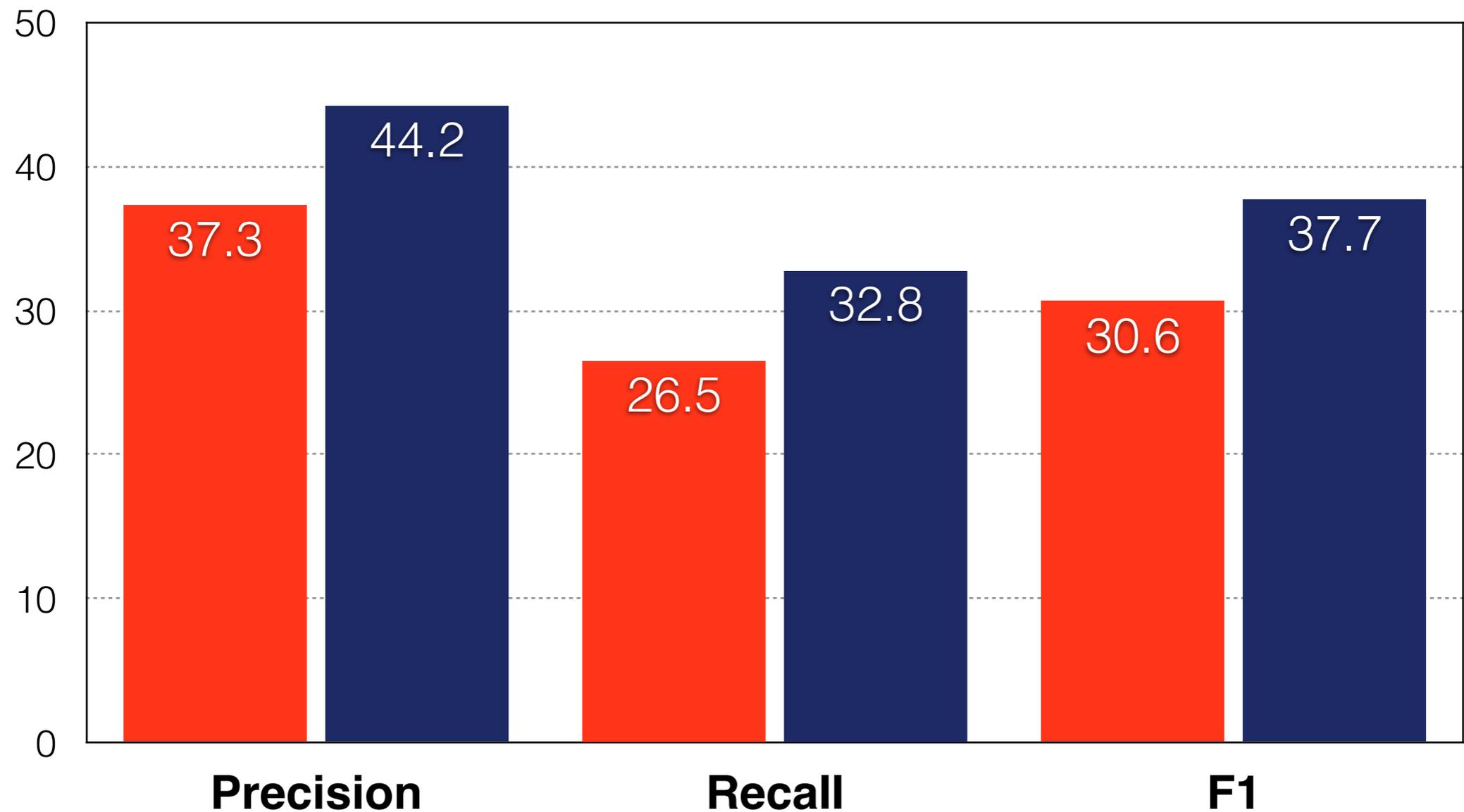
(Entity, Noun Phrase)

*Somerset British playwright, novelist  
Maugham ' and short story writer*

Entity attributes for Freebase  
**(S. Maugham, Nationality, U.K)**  
**(S. Maugham, Profession, Novelist)**  
**(S. Maugham, Profession, Playwright)**

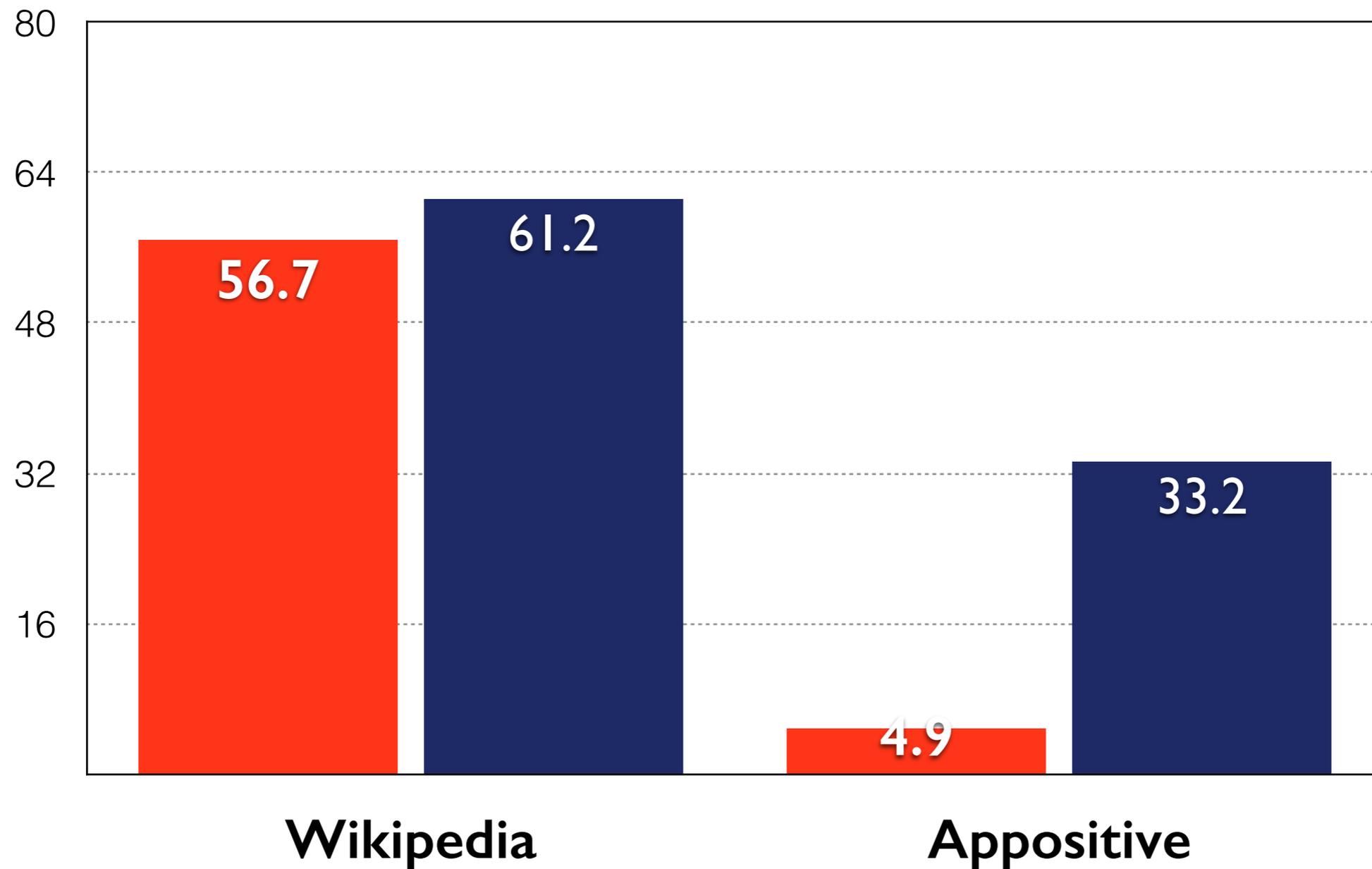
# Entity Attribute Extraction: (5 fold cross validation on Training Set)

■ Baseline      ■ Our System



# Entity Attribute Extraction: Test Result

■ Baseline      ■ Our System



# Error analysis

10% : named entity retrieving failure

10% : spurious lexical match

10% : looking at different domain

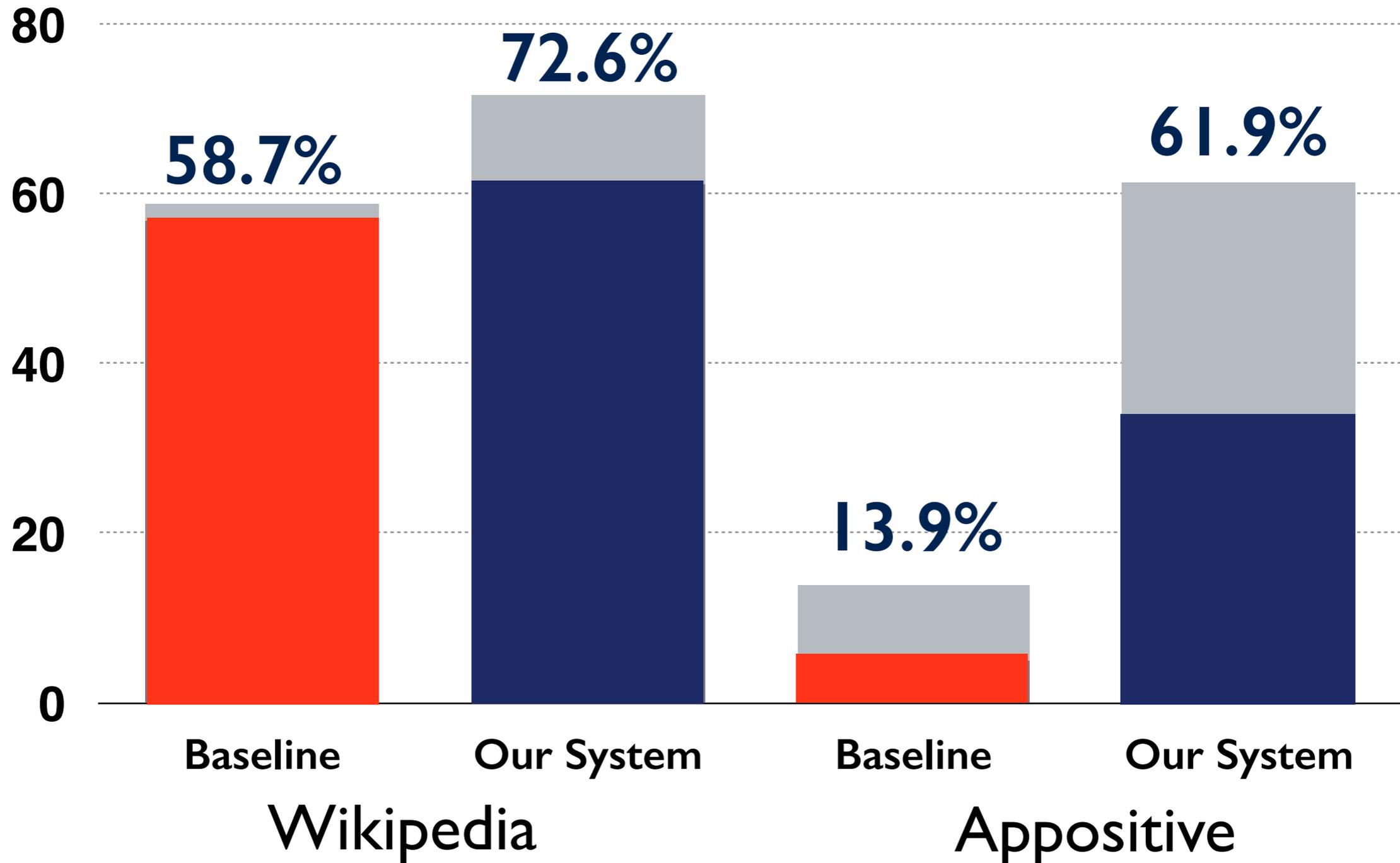
e.g: “stage actor” to “**film.actor**”

15% : wrong underspecified logical form

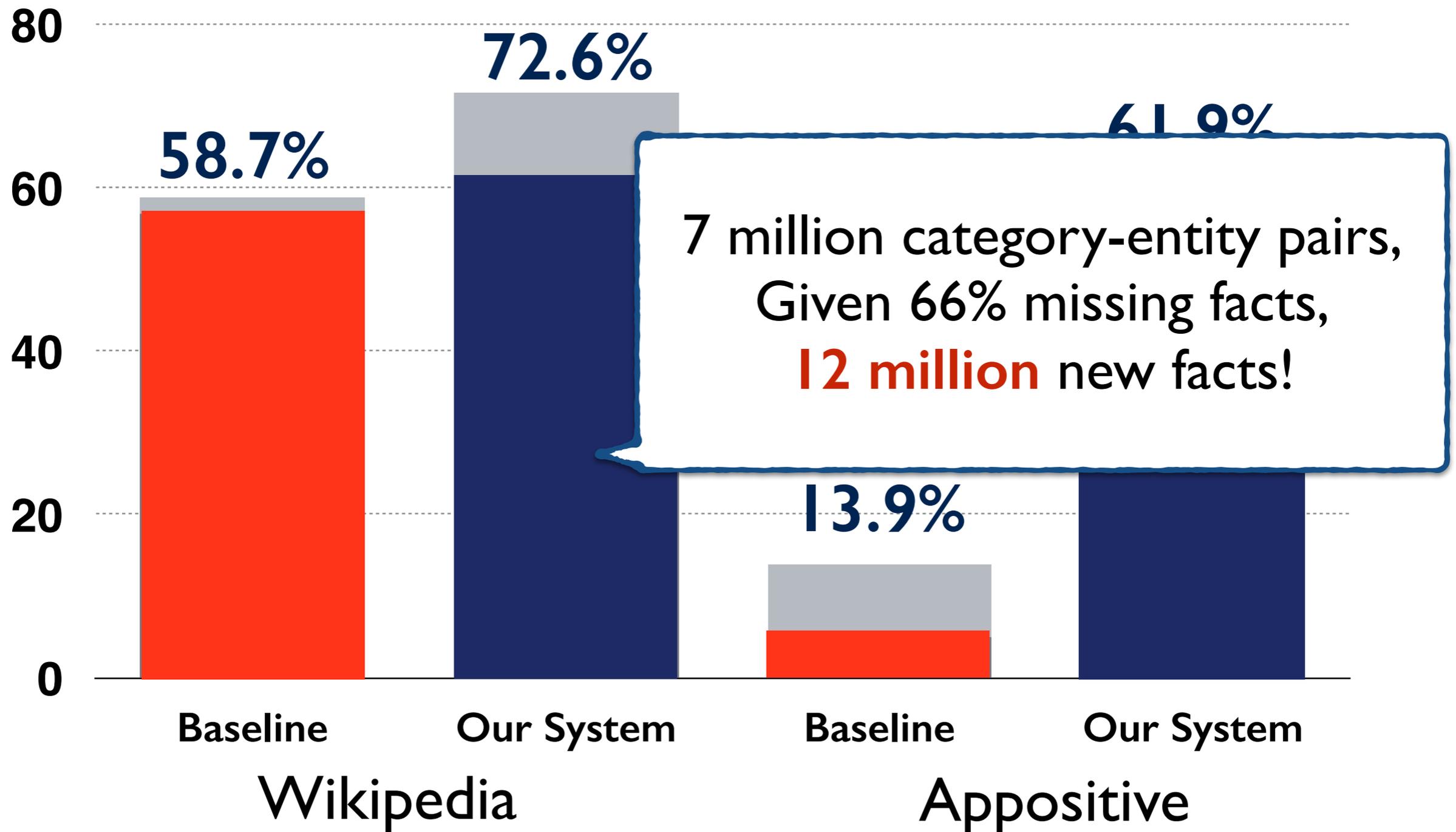
30% : mapping to superset or subset

e.g: “novel” to “**book**”

# Entity Attribute Extraction: Test Result



# Entity Attribute Extraction: Test Result



# Contributions

- Introduce large-scale semantic parsing datasets
- Partial grounding to large knowledge base
- Learn from two kinds of supervision: large-scale co-occurrence statistics and small labeled tuning set

# Future work



WIKIPEDIA  
The Free Encyclopedia

**Barack Hussein Obama** is the [44th](#) and [current](#) President of the United States, the [first African American](#) to hold the office.

```
president_number(B0, 44) ∧ ∃e.position_held(B0, e) ∧ office(e, PotUS) ∧  
∃t.from(e, t) ∧ (t < NOW) ∧ to(e, NULL) ∧ african_american(B0) ∧  
¬∃x∃e'.position_held(x, e') ∧ office(e', PotUS) ∧ african_american(x) ∧  
∃t'.from(e, t') ∧ (t' < t)
```

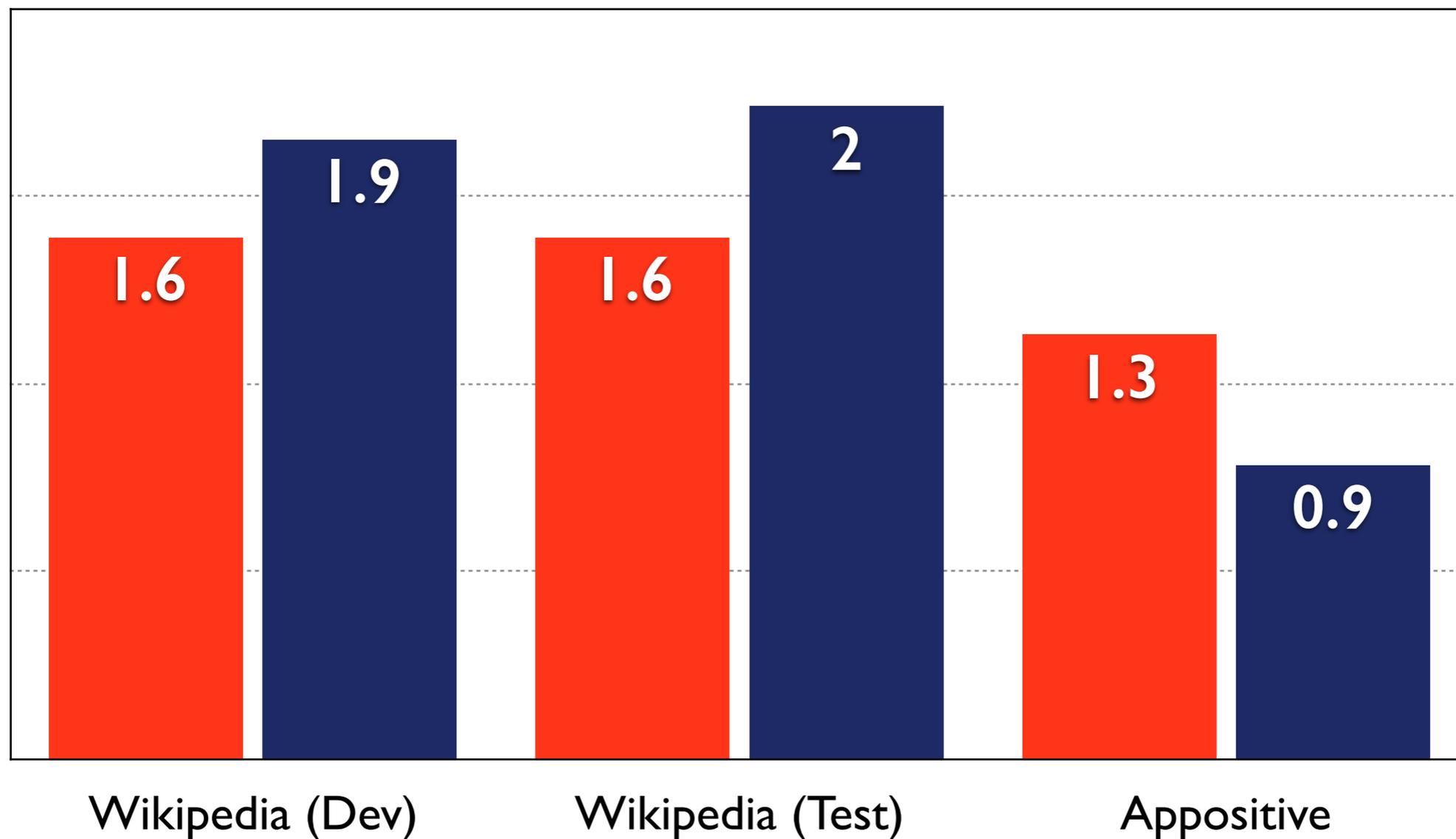
 Freebase

- More compositional and complicated structures: orders, comparison, min, max, range
- Extending to declarative sentences

# Questions?

# The Number of Extracted Facts

■ Baseline      ■ Our System



# Referring Expression Resolution: Test Set

