Semantic Parsing on Freebase from Question-Answer Pairs

EMNLP

October 20, 2013

Jonathan Berant, Andrew Chou, Roy Frostig and Percy Liang

Semantic Parsing

Who did Humphrey Bogart marry in 1928?

Semantic Parsing

Who did Humphrey Bogart marry in 1928?

$\Big\downarrow$ semantic parsing

$\text{Type.Person} \sqcap \text{Marriage.}(\text{Spouse.HumphreyBogart} \sqcap \text{StartDate.1928})$

Semantic Parsing

Who did Humphrey Bogart marry in 1928?

semantic parsing

$\text{Type.Person} \sqcap \text{Marriage.(Spouse.HumphreyBogart} \sqcap \text{StartDate.1928)}$

execute logical form

Mary Philips

Semantic Parsing

Who did Humphrey Bogart marry in 1928?

$\Big\downarrow$ semantic parsing

Type.Person $\sqcap$ Marriage.(Spouse.HumphreyBogart $\sqcap$ StartDate.1928)

$\Big\downarrow$ execute logical form

Mary Philips

Motivation: Natural language interface to large structured knowledge-bases (Freebase, DBPedia, Yelp, ...)

[Zelle & Mooney, 1996; Zettlemoyer & Collins, 2005; ...]

Statistical semantic parsing

Supervision: manually annotated logical forms

What's California's capital?	Capital.California
How long is the Mississippi river?	RiverLength.Mississippi
...	...

[Zelle & Mooney, 1996; Zettlemoyer & Collins, 2005; ...]

Statistical semantic parsing

Supervision: manually annotated logical forms

What's California's capital?	Capital.California
How long is the Mississippi river?	RiverLength.Mississippi
...	...

Limitations:

- Requires experts — slow, expensive, does not scale!
- Restricted to limited domains

[Clarke et al. 2010; Liang et al. 2011]

Weakly supervised parsers

Supervision: question/answers pairs

What's California's capital?	Sacramento
How long is the Mississippi river?	3,734km
...	...

Weakly supervised parsers

Supervision: question/answers pairs

What's California's capital?	Sacramento
How long is the Mississippi river?	3,734km
...	...

Advantage: obtain from non-experts!

[Clarke et al. 2010; Liang et al. 2011]

Weakly supervised parsers

Supervision: question/answers pairs

What's California's capital?	Sacramento
How long is the Mississippi river?	3,734km
...	...

Advantage: obtain from non-experts!

Dataset	# word types
`GeoQuery`	279
`ATIS`	936
`KM-NP`	158

Scaling to large knowledge-bases

Unsupervised systems with no training

- Unger et al., 2012; Yahya et al., 2012

Scaling to large knowledge-bases

Unsupervised systems with no training

- Unger et al., 2012; Yahya et al., 2012

Distant supervision (on a small set of KB predicates)

- Krishnamurthy and Mitchell, 2012

Scaling to large knowledge-bases

Unsupervised systems with no training

- Unger et al., 2012; Yahya et al., 2012

Distant supervision (on a small set of KB predicates)

- Krishnamurthy and Mitchell, 2012

Parser trained from question/logical form pairs

- Cai and Yates, 2013

Scaling to large knowledge-bases

Unsupervised systems with no training

- Unger et al., 2012; Yahya et al., 2012

Distant supervision (on a small set of KB predicates)

- Krishnamurthy and Mitchell, 2012

Parser trained from question/logical form pairs

- Cai and Yates, 2013

Our goal: Training a parser from **question/answer pairs** on a **large knowledge-base**

Challenge: mapping text to the KB

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

What languages do people in Brazil use

Challenge: mapping text to the KB

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

What languages do people in Brazil use

- Exhaustive enumeration is intractable [Liang et al. 2011]

Challenge: mapping text to the KB

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

Type.HumanLanguage
SpeechLanguagePathology
LanguageAcquisition

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

What languages do people in Brazil use

- Exhaustive enumeration is intractable [Liang et al. 2011]
- String matching is not precise [Yahya et al. 2012]

Challenge: mapping text to the KB

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

Type.HumanLanguage
SpeechLanguagePathology
LanguageAcquisition

BarackObama
TopGun
MichelleObama
Type.Country
Type.city
Profession.Lawyer
PeopleBornHere
InventorOf
CapitalOf
⋮

LanguagesSpoken

What languages do people in Brazil use

- Exhaustive enumeration is intractable [Liang et al. 2011]
- String matching is not precise [Yahya et al. 2012]
- String matching has coverage issues

Contributions

BarackObama
TopGun
Type.Country
Profession.Lawyer
PeopleBornHere
InventorOf
⋮

What languages do people in Brazil use

Contributions

BarackObama
TopGun
Type.Country
Profession.Lawyer
PeopleBornHere
InventorOf
⋮

Type.HumanLanguage | Brazil
Type.ProgrammingLanguage | BrazilFootballTeam

alignment | **alignment**

What languages do people in Brazil use

Alignment: lexicon from text phrases to KB predicates

Contributions

BarackObama
TopGun
Type.Country
Profession.Lawyer
PeopleBornHere
InventorOf
⋮

LanguagesSpoken

Type.HumanLanguage
Type.ProgrammingLanguage ⋯ **bridging** ⋯ BrazilFootballTeam
Brazil

alignment **alignment**

What languages do people in Brazil use

Alignment: lexicon from text phrases to KB predicates

Bridging: Use context to generate KB predicates

Semantic parsing

- **Setup**
- Alignment
- Bridging
- Composition
- Dataset creation
- Experiments

Setup

Input:

- Knowledge-base $\mathcal{K}$
- Training set of question-answer pairs $\{(x_i, y_i)\}_1^n$

What are the main cities in California? SF, LA, ...

Setup

Input:

- Knowledge-base $\mathcal{K}$
- Training set of question-answer pairs $\{(x_i, y_i)\}_1^n$

What are the main cities in California?   SF, LA, ...

Output:

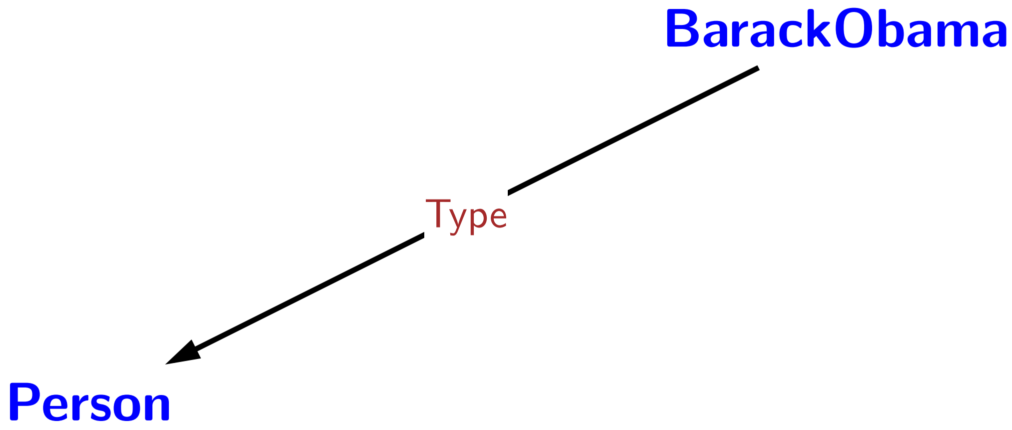
- Semantic parser that maps questions x to answers y through logical forms z

countries in Asia $\Rightarrow$ Type.Country $\sqcap$ ContainedBy.Asia

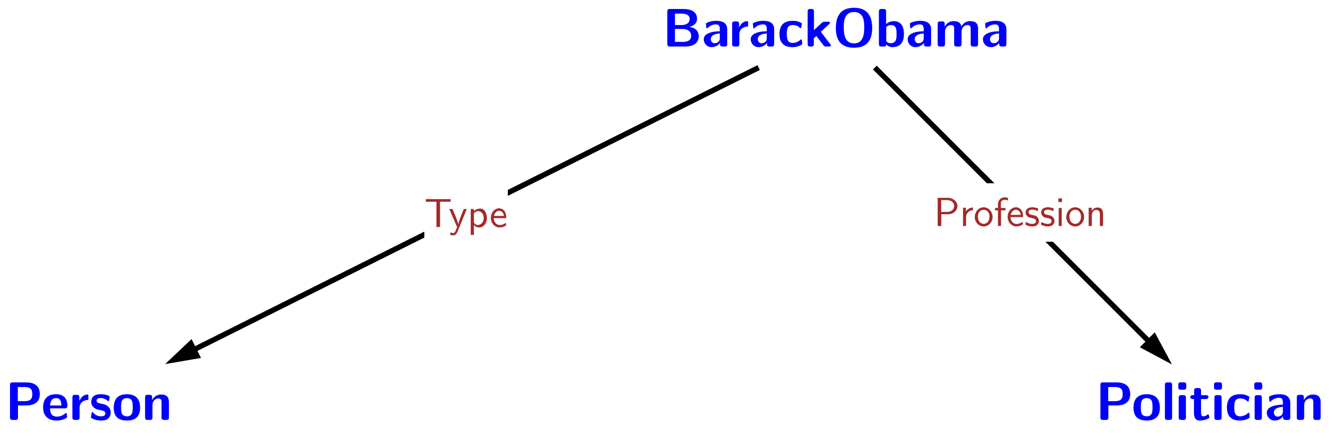
$\Rightarrow$ China, Japan, Israel, ...

Freebase knowledge graph

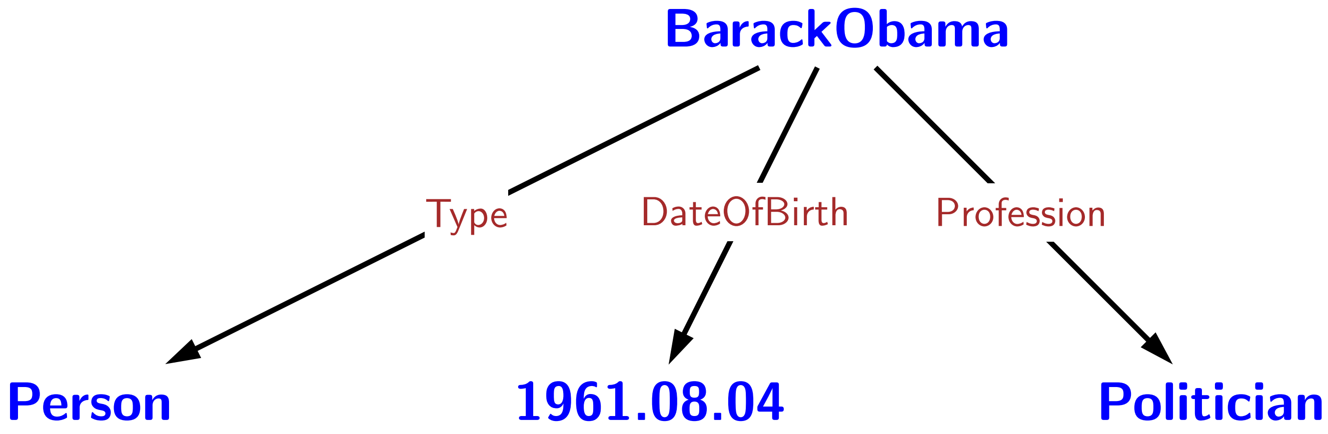
Freebase knowledge graph



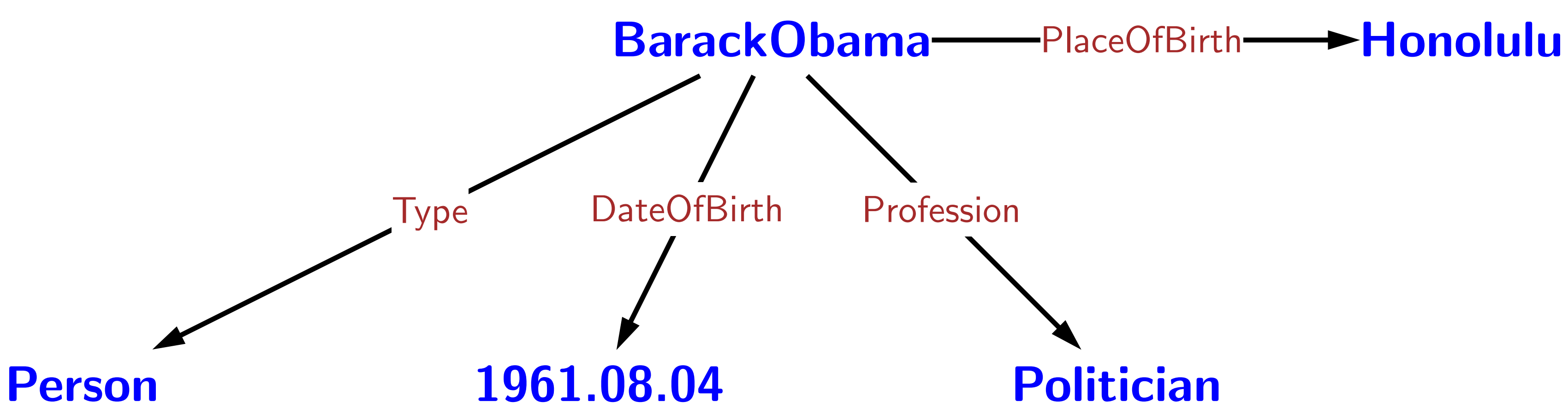
Freebase knowledge graph



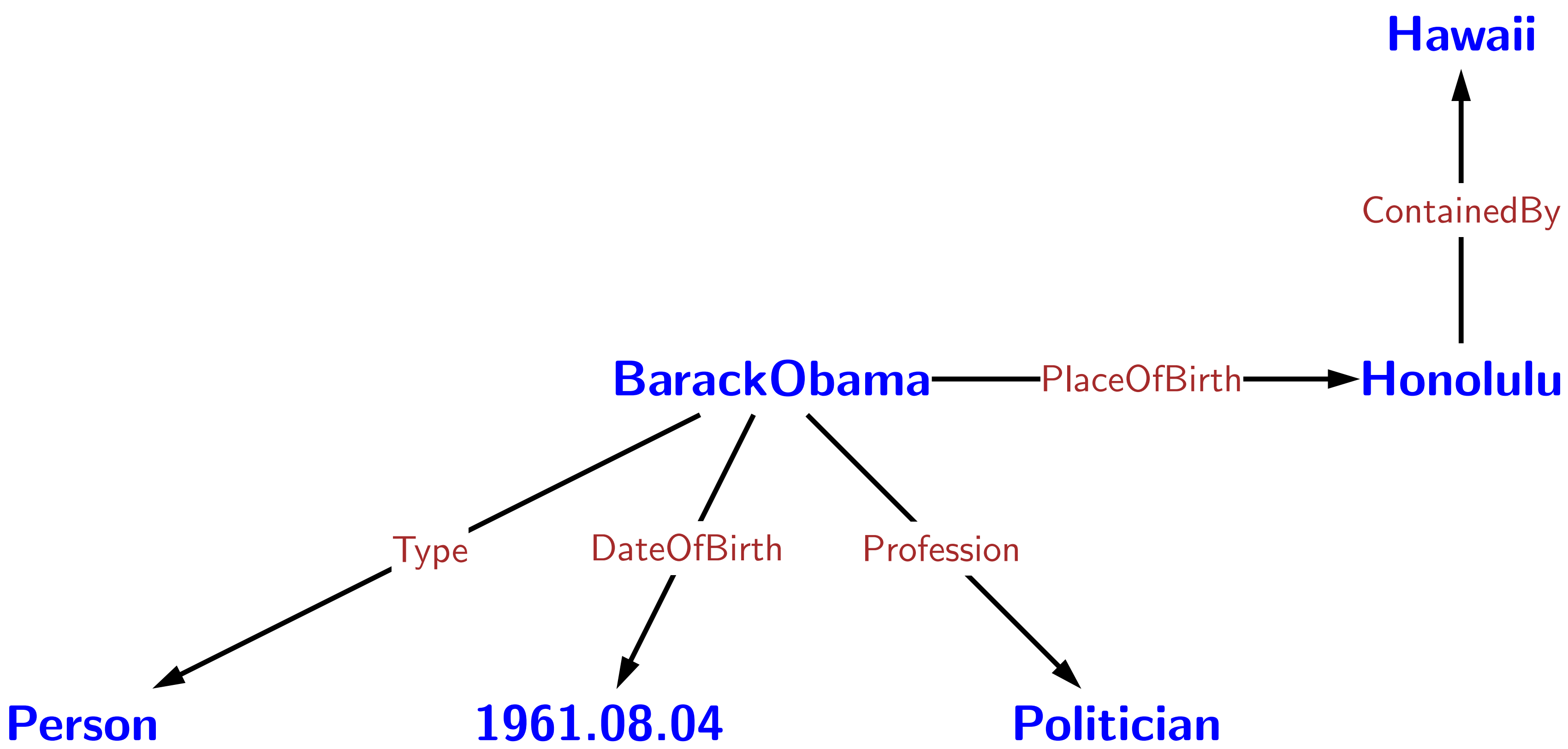
Freebase knowledge graph



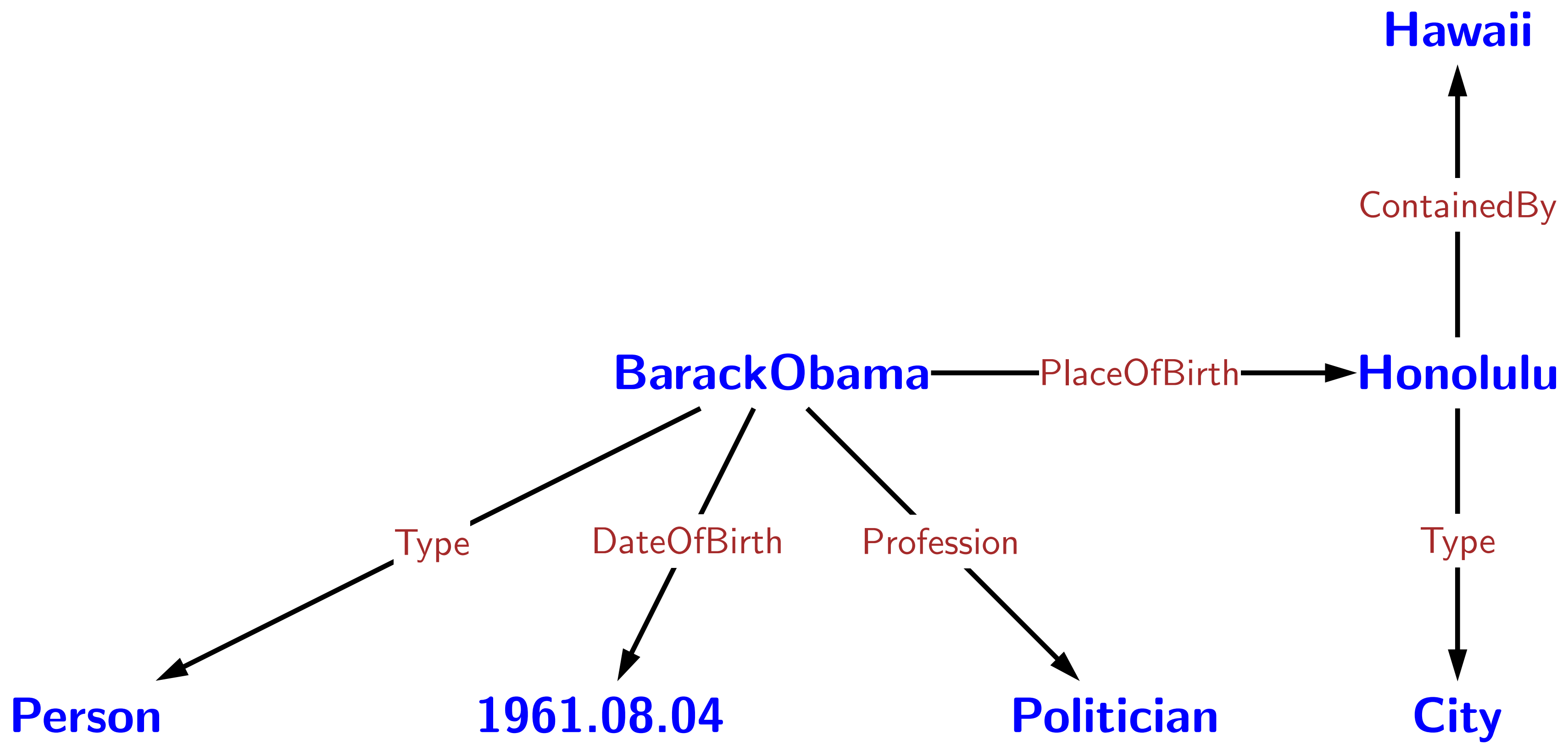
Freebase knowledge graph



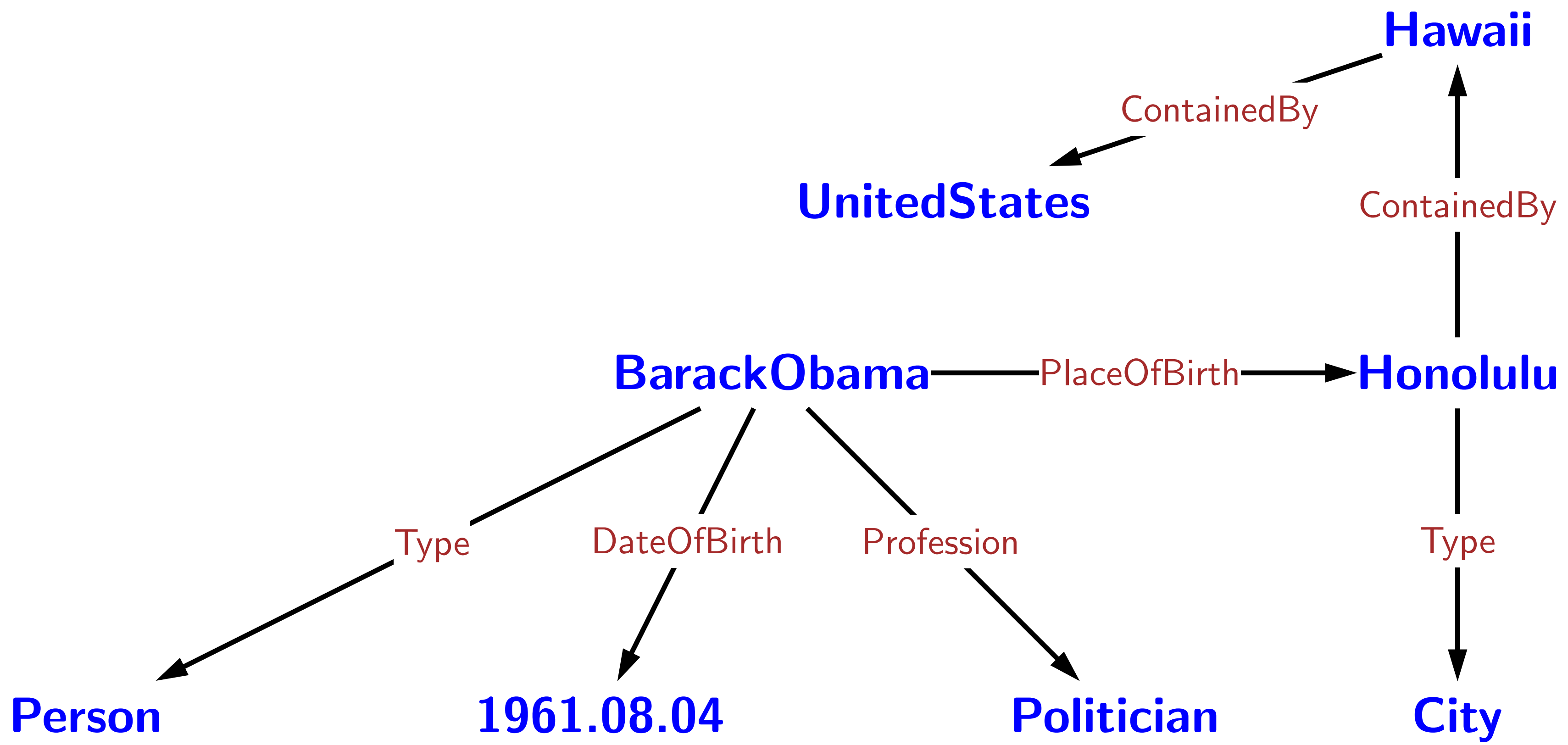
Freebase knowledge graph



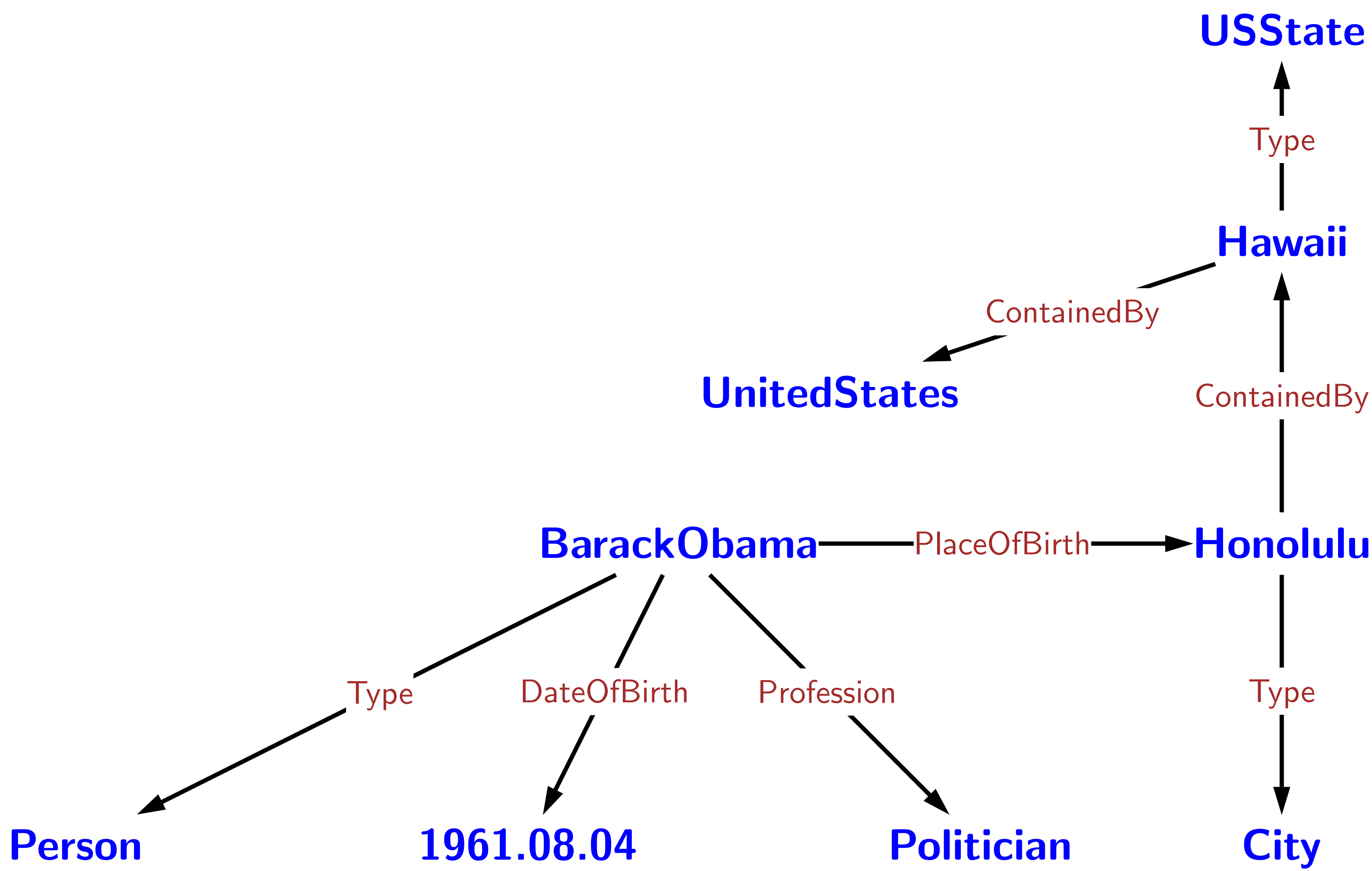
Freebase knowledge graph



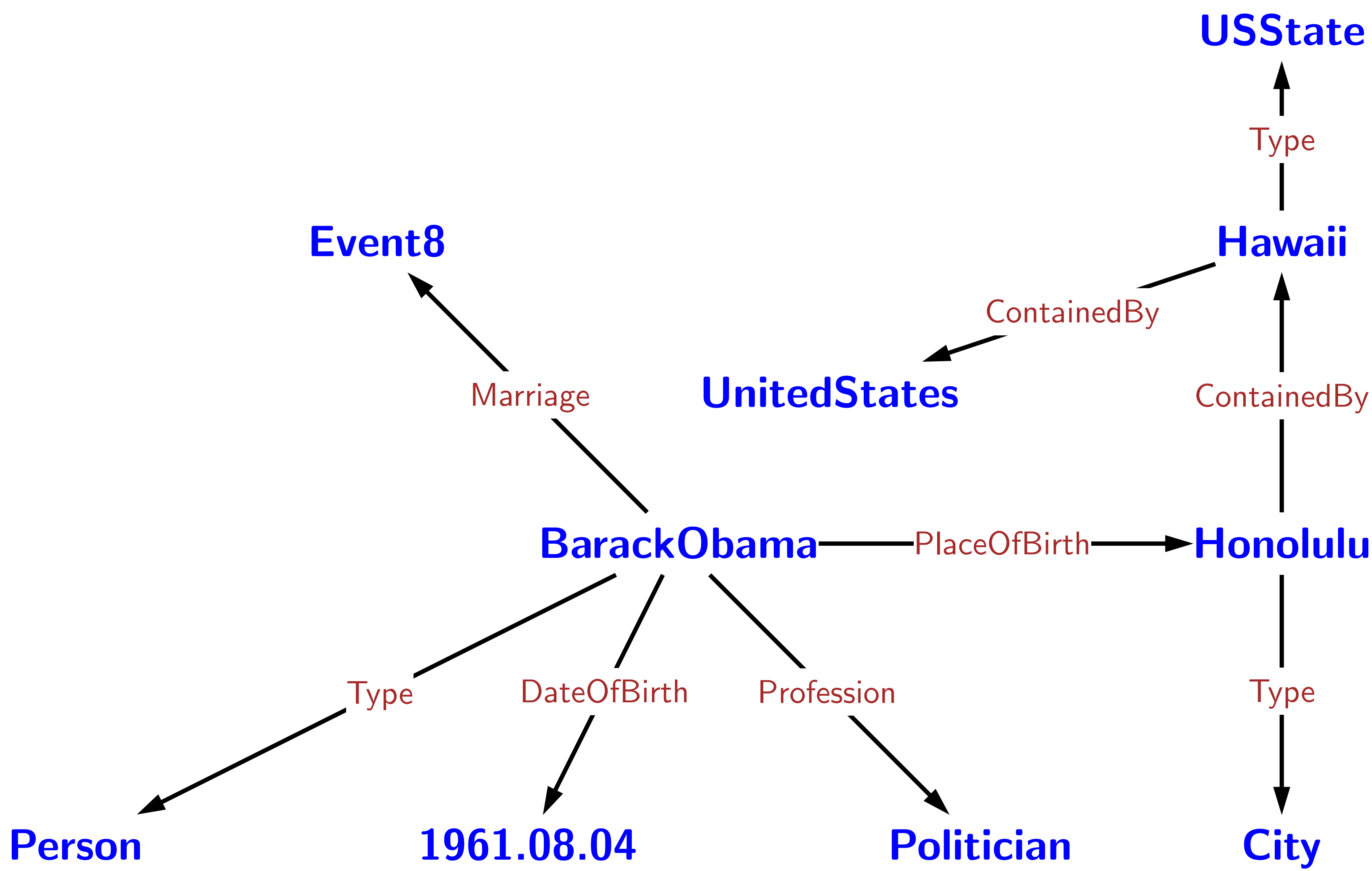
Freebase knowledge graph



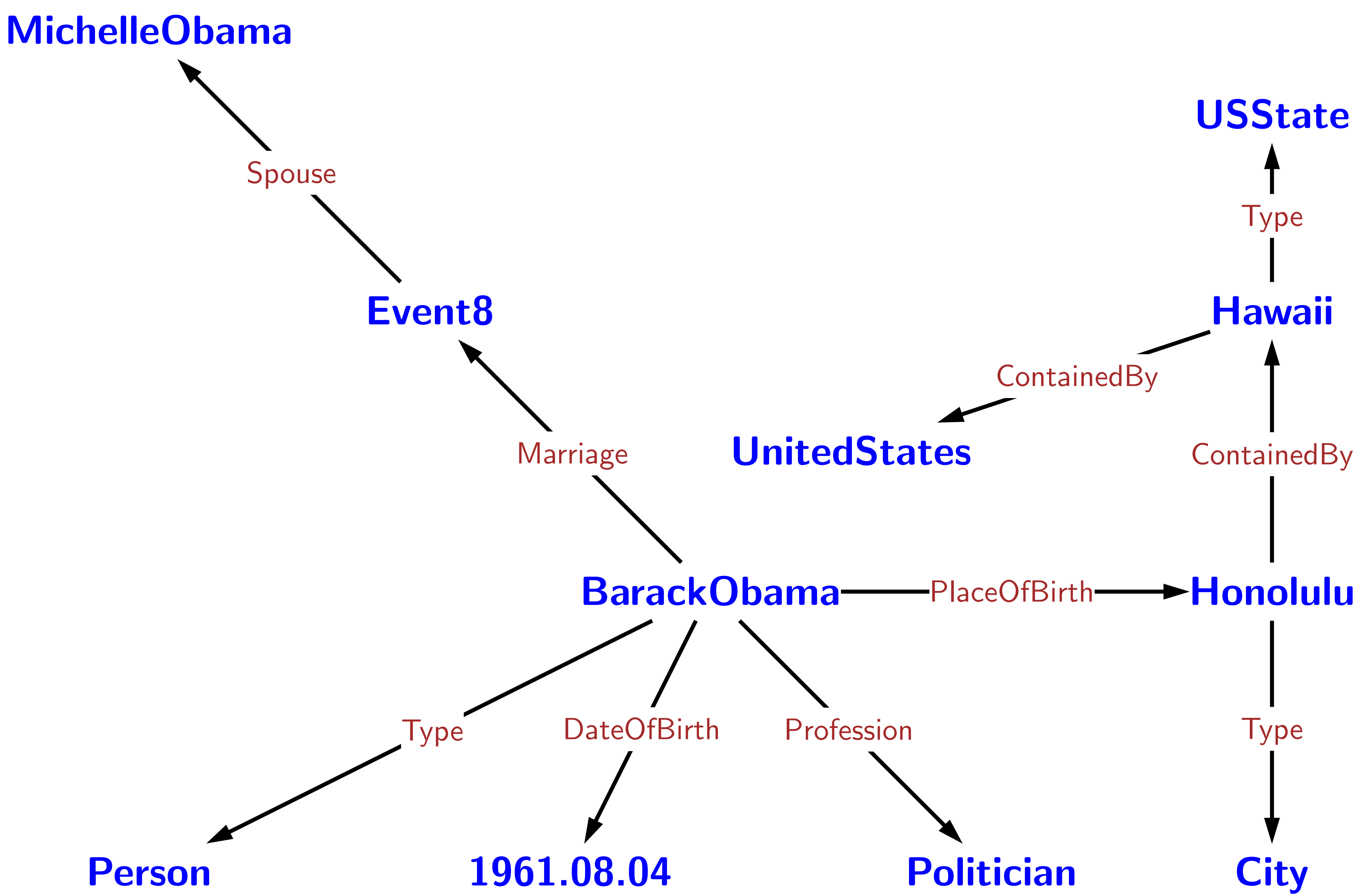
Freebase knowledge graph



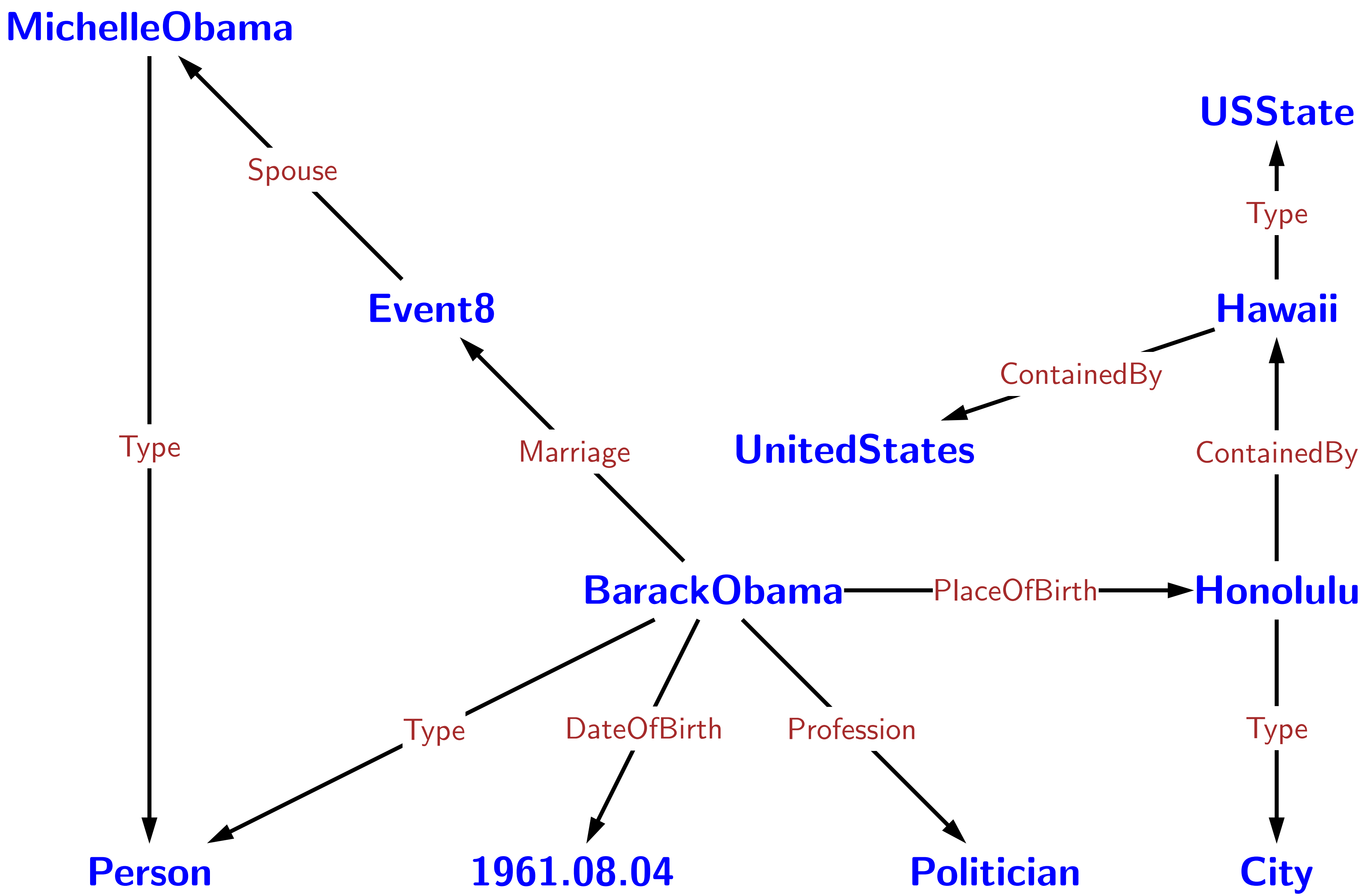
Freebase knowledge graph



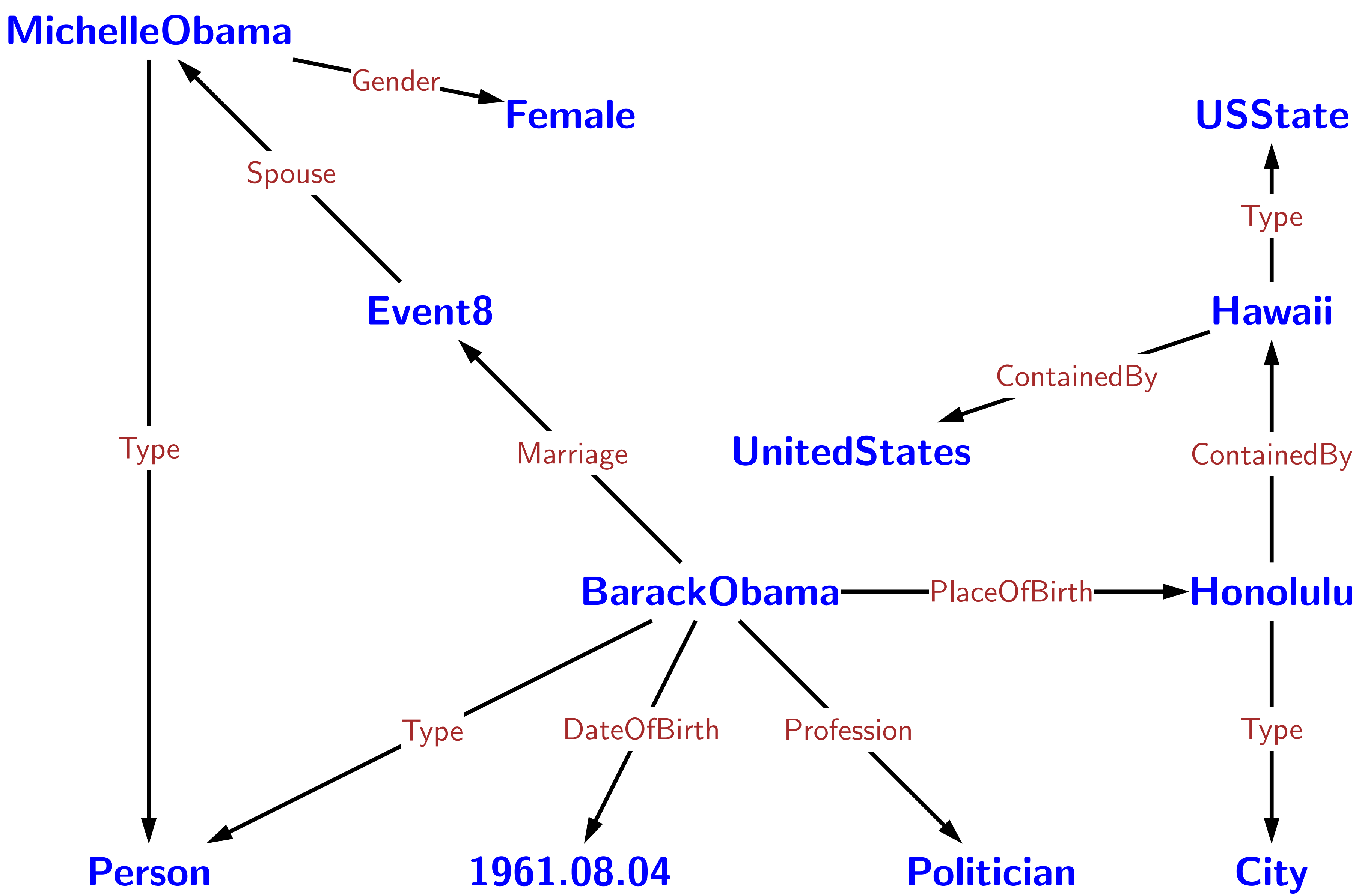
Freebase knowledge graph



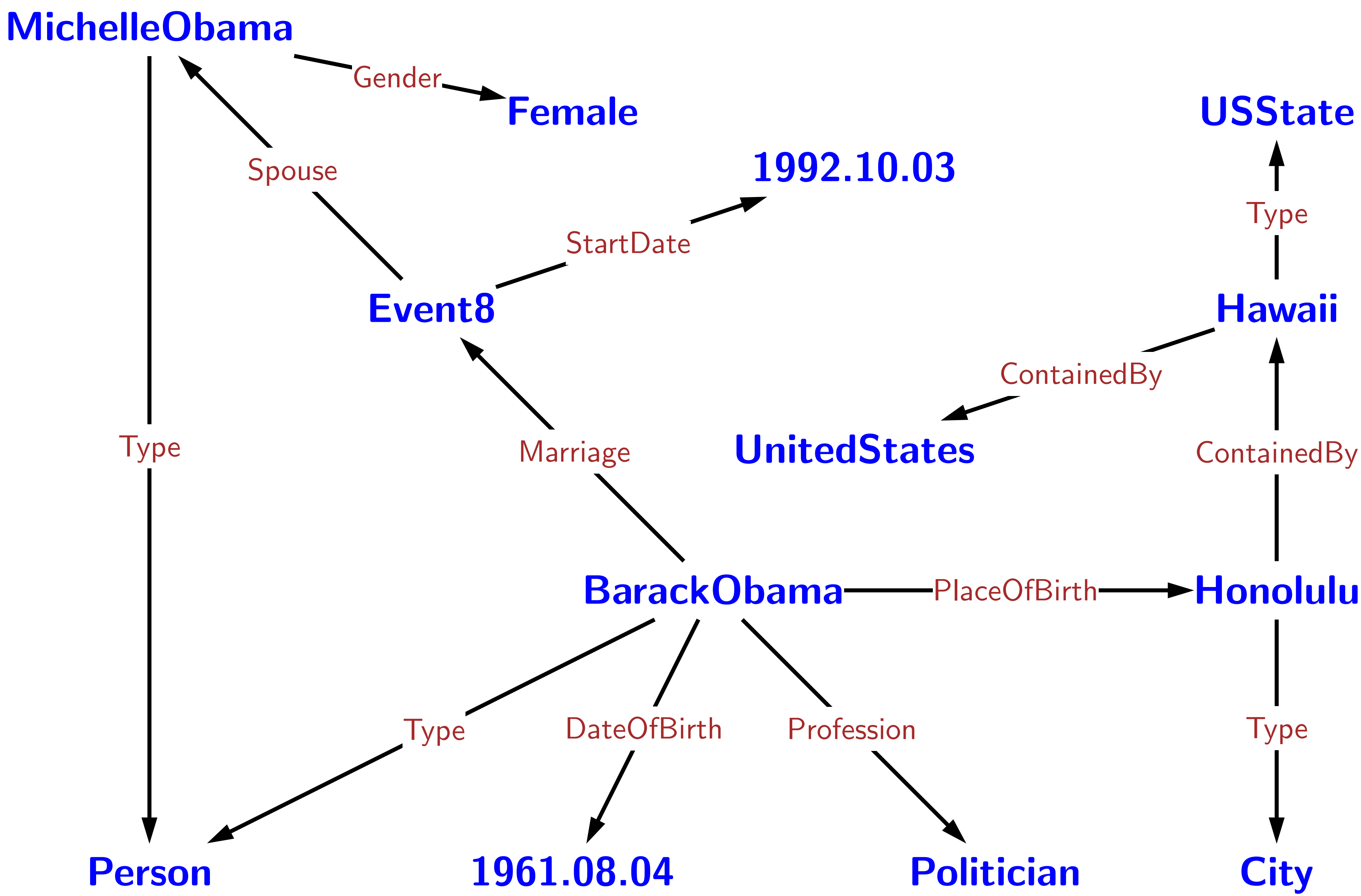
Freebase knowledge graph



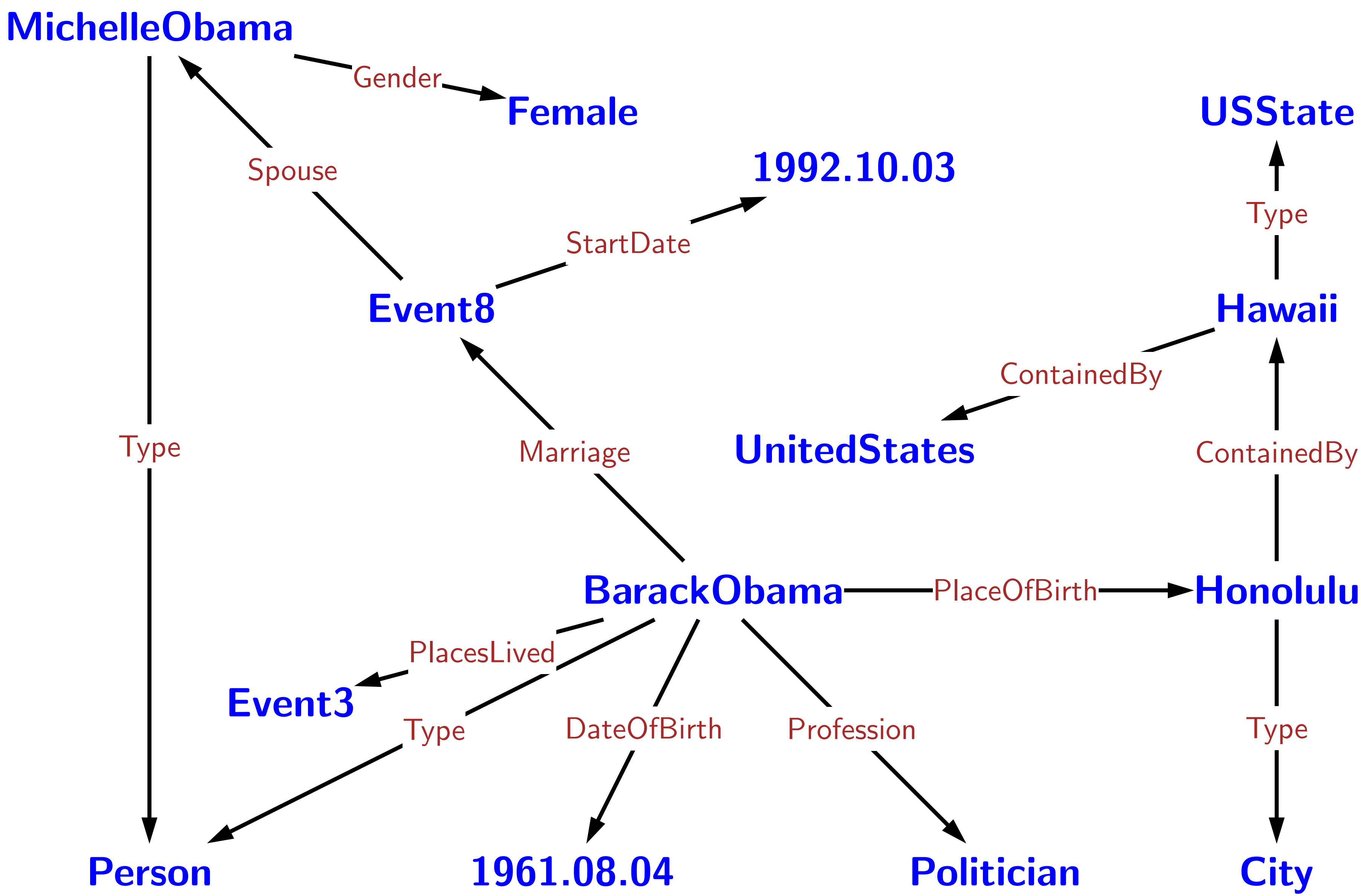
Freebase knowledge graph



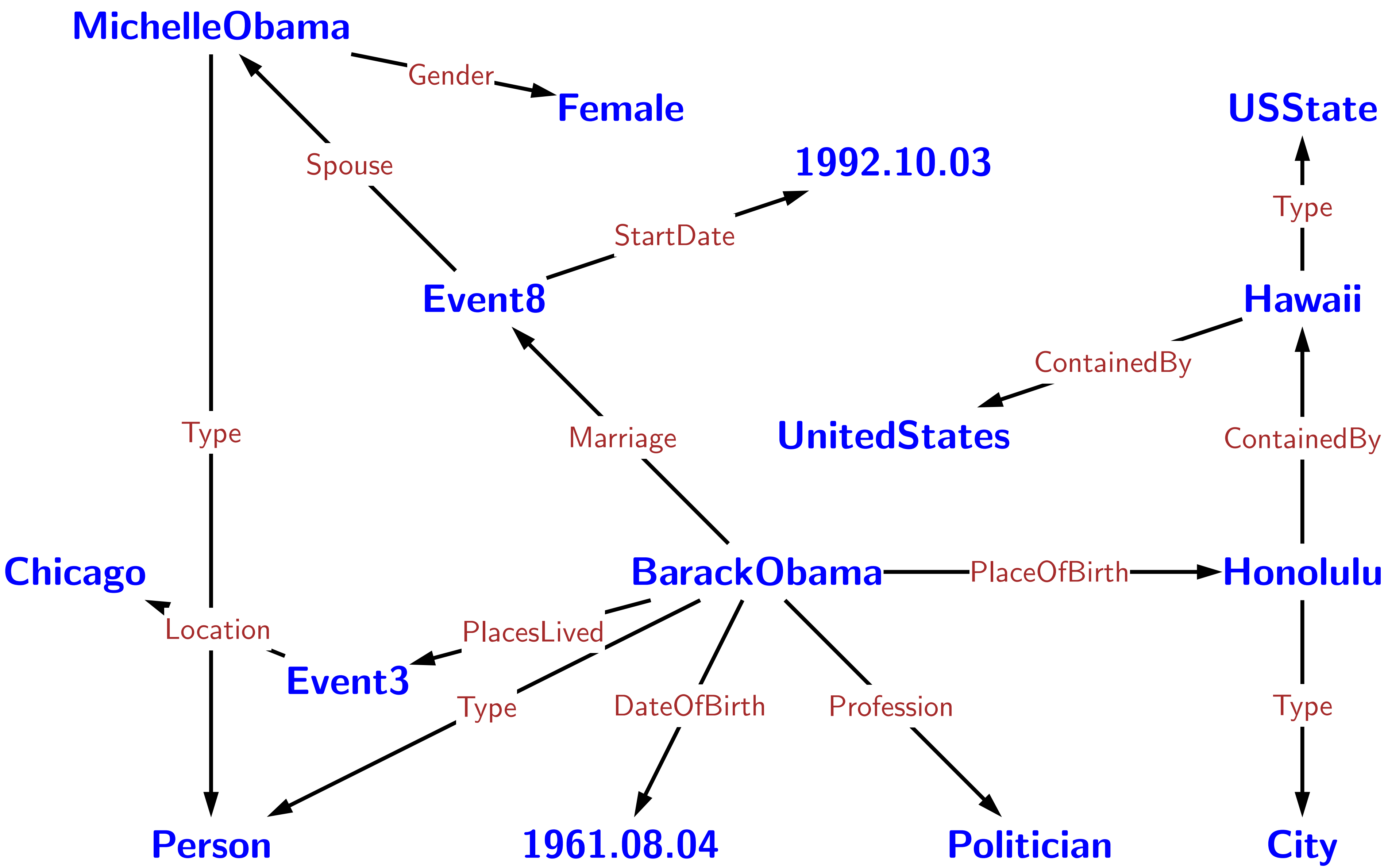
Freebase knowledge graph



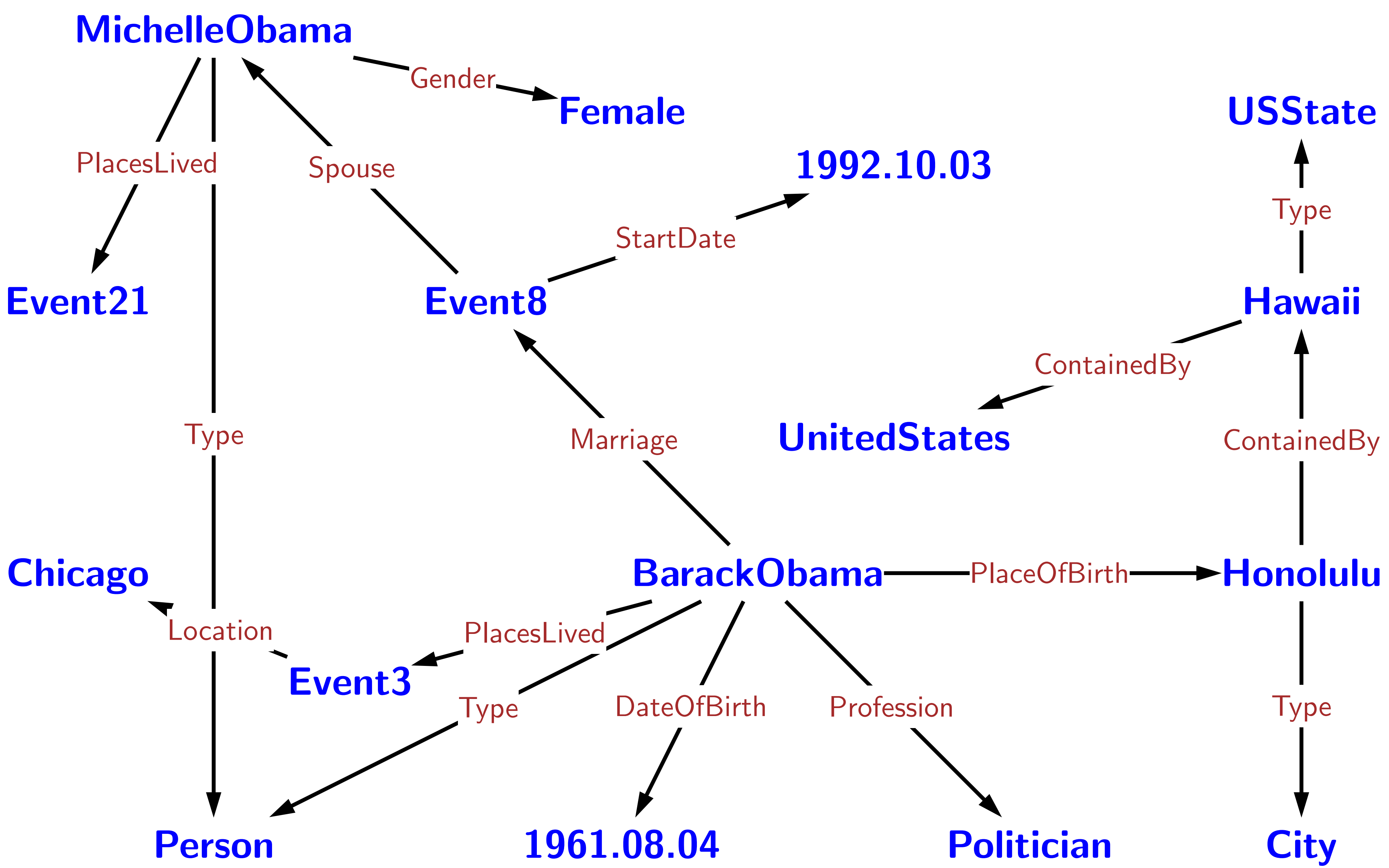
Freebase knowledge graph



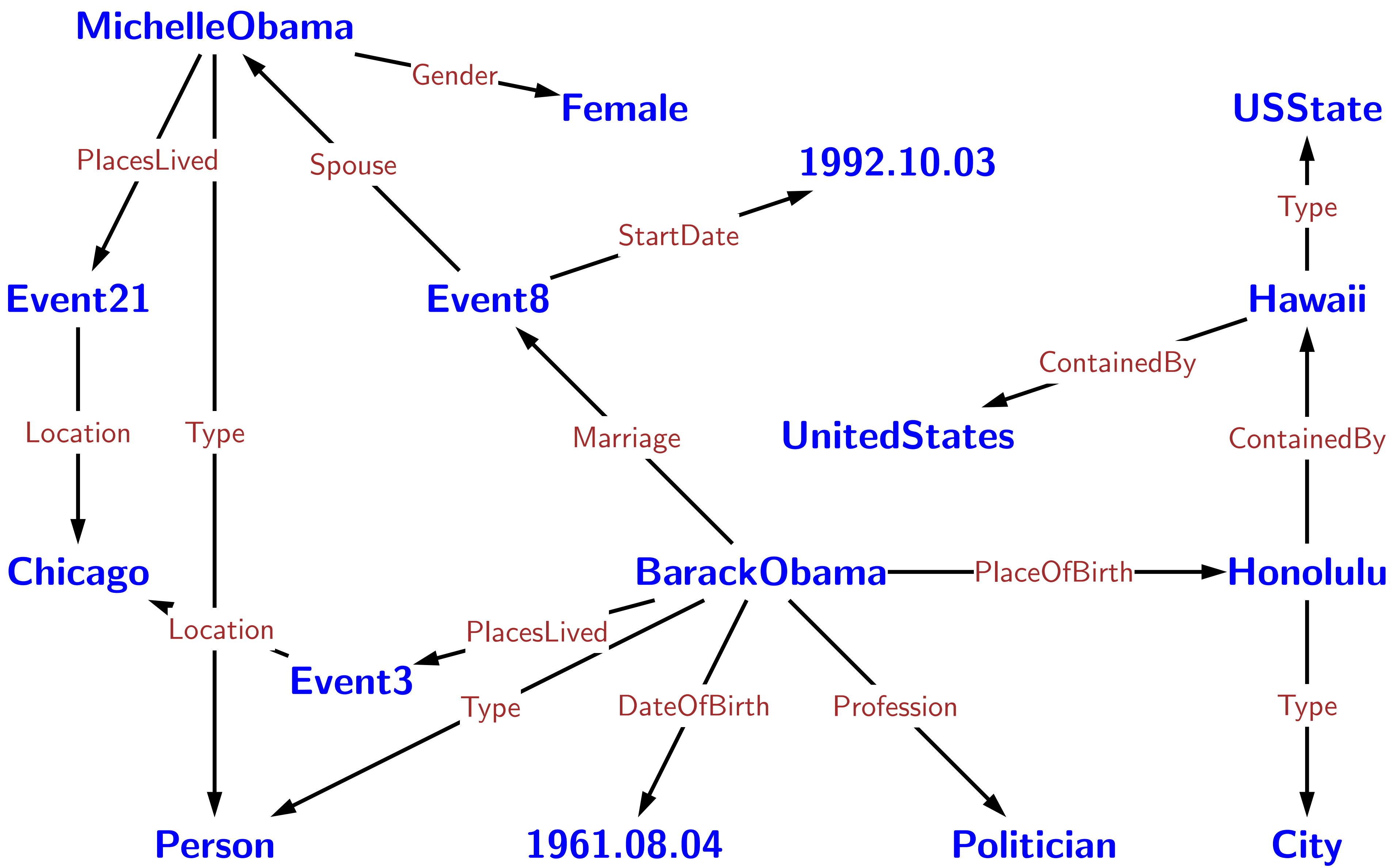
Freebase knowledge graph



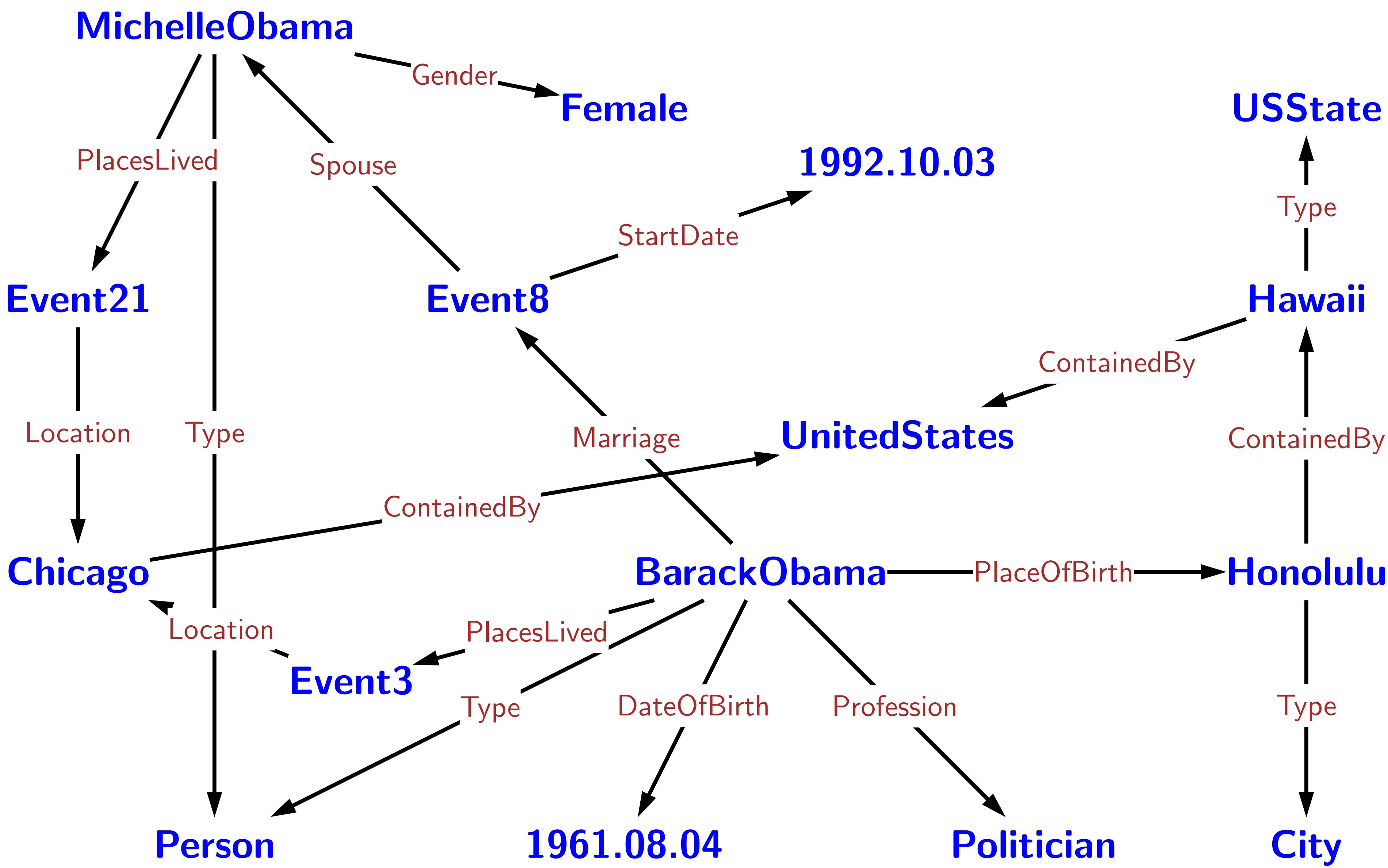
Freebase knowledge graph



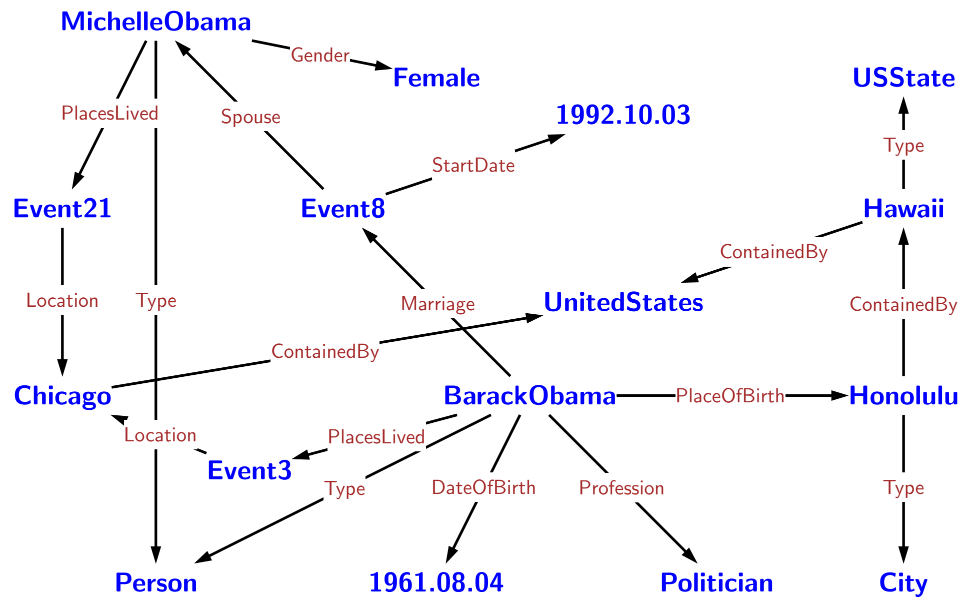
Freebase knowledge graph



Freebase knowledge graph



Freebase knowledge graph



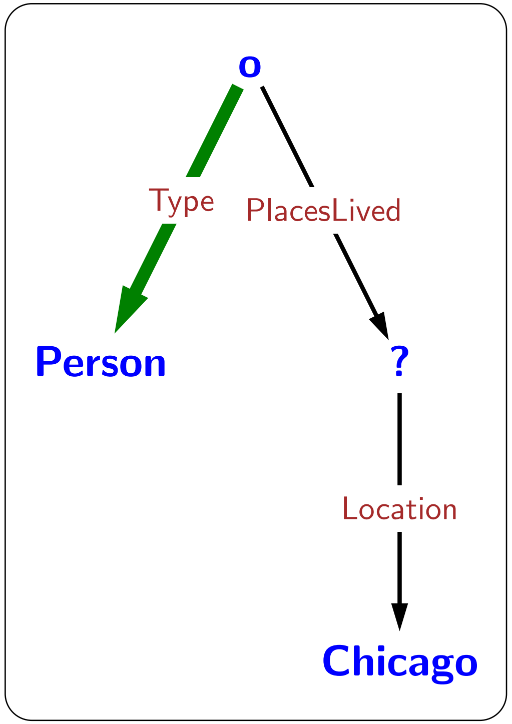
41M entities (nodes)

19K properties (edge labels)

596M assertions (edges)

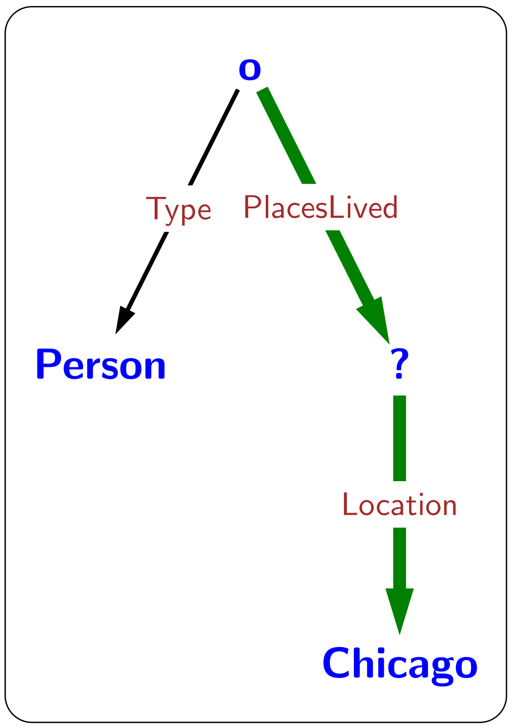
Logical forms are graph templates

Type.Person $\sqcap$ PlacesLived.Location.Chicago



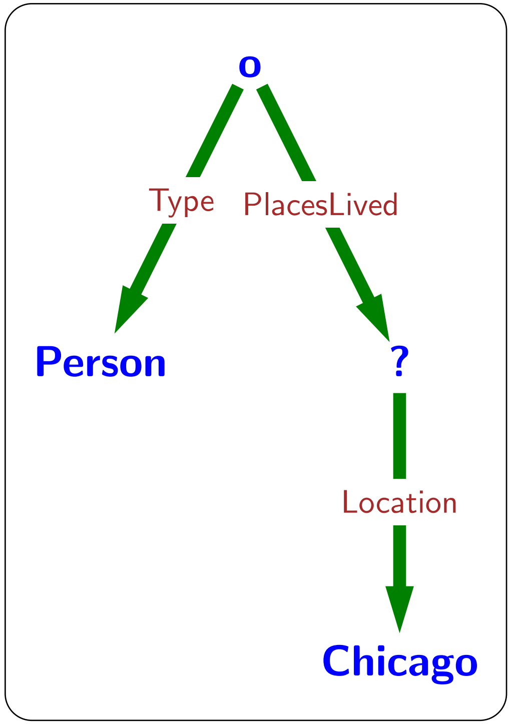
Logical forms are graph templates

$\text{Type.Person} \sqcap \text{PlacesLived.Location.Chicago}$



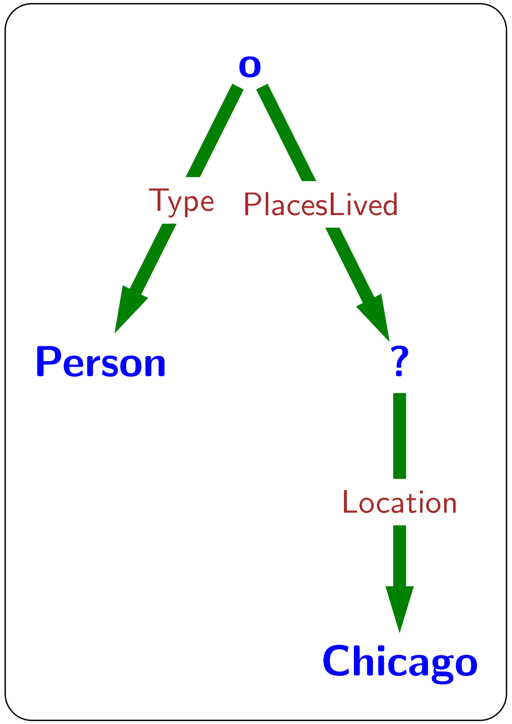
Logical forms are graph templates

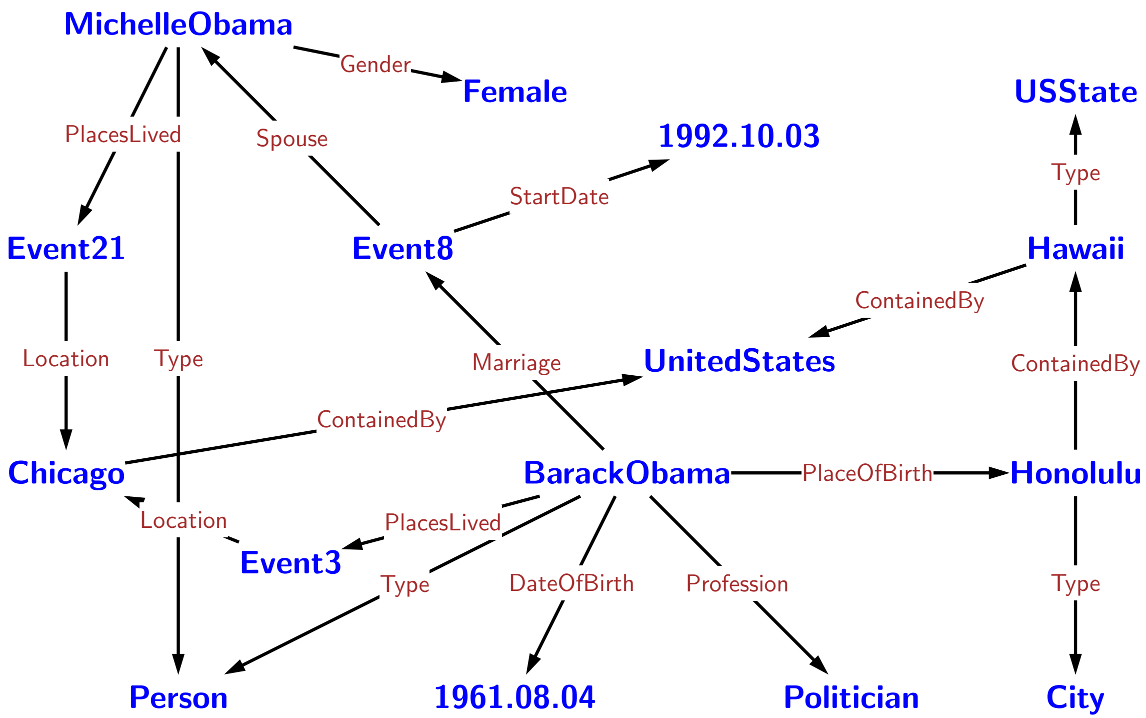
Type.Person$\sqcap$PlacesLived.Location.Chicago



Logical forms are graph templates

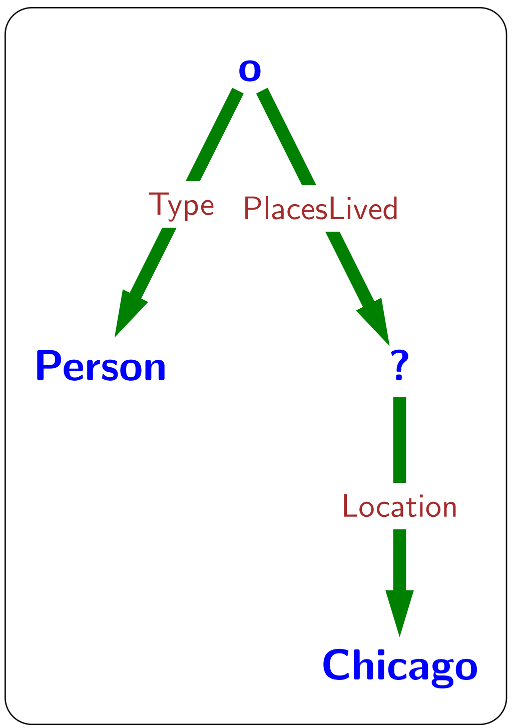
Type.Person$\sqcap$PlacesLived.Location.Chicago

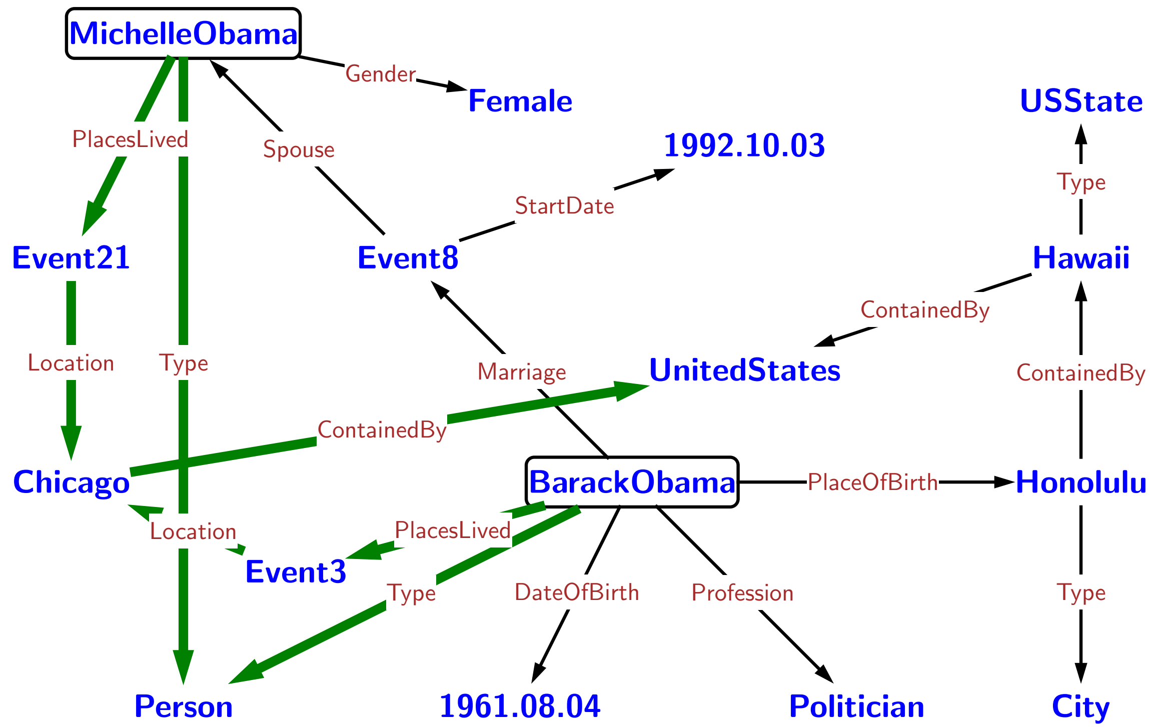




Logical forms are graph templates

Type.Person$\sqcap$PlacesLived.Location.Chicago





Semantic parsing

- Setup
- **Alignment**
- Bridging
- Composition
- Dataset creation
- Experiments

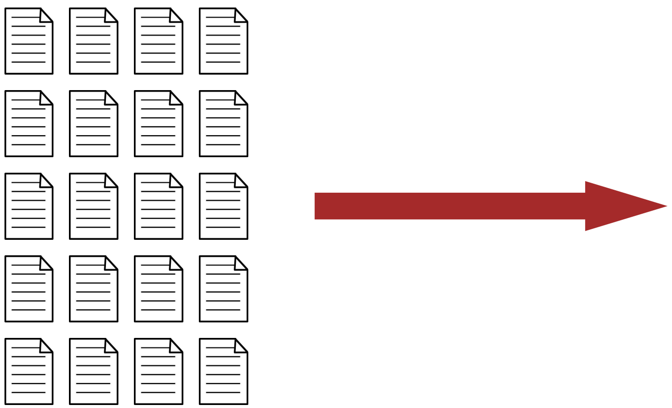
Alignment

BarackObama
TopGun
Type.Country
Profession.Lawyer
PeopleBornHere
InventorOf
⋮

	Type.HumanLanguage Type.ProgrammingLanguage				Brazil BrazilFootballTeam	
	alignment				**alignment**	
What	**languages**	**do**	**people**	**in**	**Brazil**	**use**

Alignment: text phrases

ReVerb on ClueWeb09 [Thomas Lin]:



(*Barack Obama*, *was born in*, *Honolulu*)

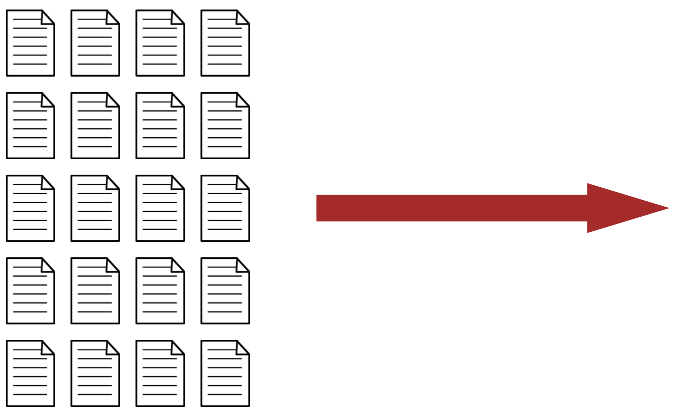
(*Albert Einstein*, *was born in*, *Ulm*)

(*Barack Obama*, *lived in*, *Chicago*)

... 15M triples ...

Alignment: text phrases

ReVerb on ClueWeb09 [Thomas Lin]:



(*Barack Obama*, *was born in*, *Honolulu*)

(*Albert Einstein*, *was born in*, *Ulm*)

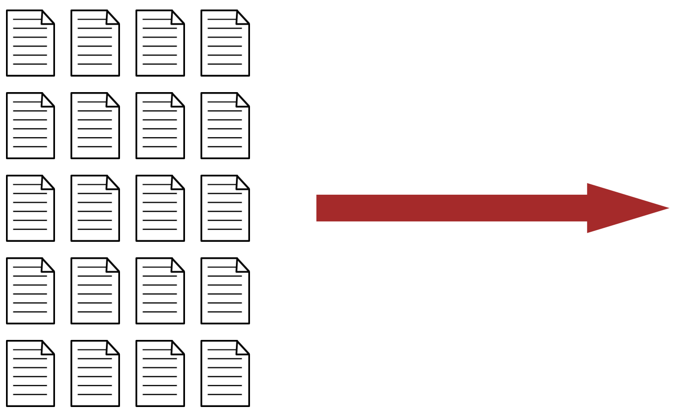
(*Barack Obama*, *lived in*, *Chicago*)

... 15M triples ...

- Entities are linked to Freebase

Alignment: text phrases

ReVerb on ClueWeb09 [Thomas Lin]:

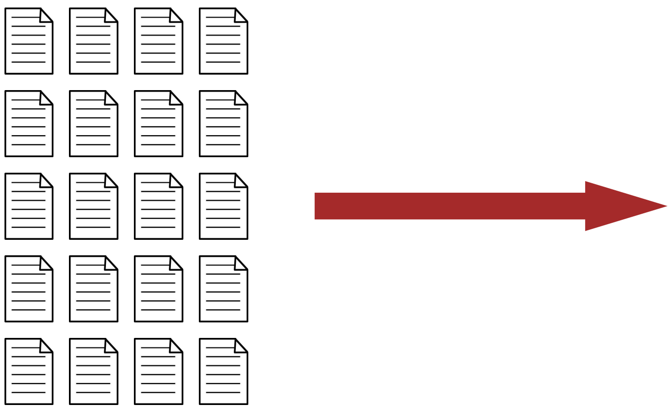


(*Barack Obama*, *was born in*, *Honolulu*)
(*Albert Einstein*, *was born in*, *Ulm*)
(*Barack Obama*, *lived in*, *Chicago*)
... 15M triples ...

- Entities are linked to Freebase
- Hearst patterns used for unaries

Alignment: text phrases

ReVerb on ClueWeb09 [Thomas Lin]:



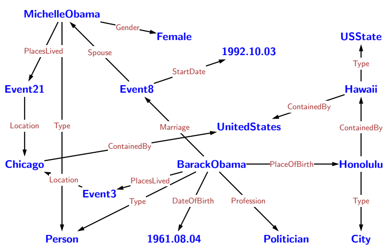
(*Barack Obama*, *was born in*, *Honolulu*)
(*Albert Einstein*, *was born in*, *Ulm*)
(*Barack Obama*, *lived in*, *Chicago*)
... 15M triples ...

- Entities are linked to Freebase
- Hearst patterns used for unaries

15,000 text phrases

Alignment: KB predicates

Freebase:



(BarackObama, PlaceOfBirth, Honolulu)

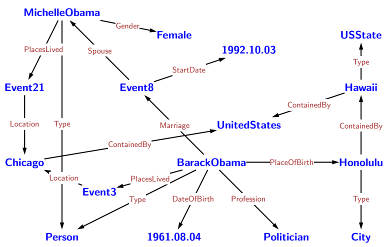
(Albert Einstein, PlaceOfBirth, Ulm)

(BarackObama, PlacesLived.Location, Chicago)

... 600M triples ...

Alignment: KB predicates

Freebase:



(BarackObama, PlaceOfBirth, Honolulu)

(Albert Einstein, PlaceOfBirth, Ulm)

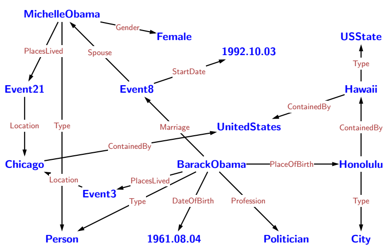
(BarackObama, PlacesLived.Location, Chicago)

... 600M triples ...

Binaries: paths of length 1 or 2 in the KB graph

Alignment: KB predicates

Freebase:



(BarackObama, PlaceOfBirth, Honolulu)

(Albert Einstein, PlaceOfBirth, Ulm)

(BarackObama, PlacesLived.Location, Chicago)

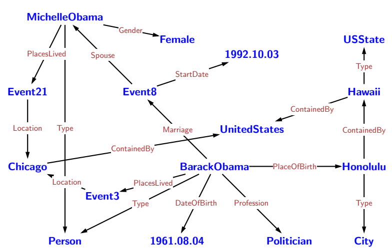
... 600M triples ...

Binaries: paths of length 1 or 2 in the KB graph

Unaries: Type.x or Profession.x

Alignment: KB predicates

Freebase:



(BarackObama, PlaceOfBirth, Honolulu)

(Albert Einstein, PlaceOfBirth, Ulm)

(BarackObama, PlacesLived.Location, Chicago)

... 600M triples ...

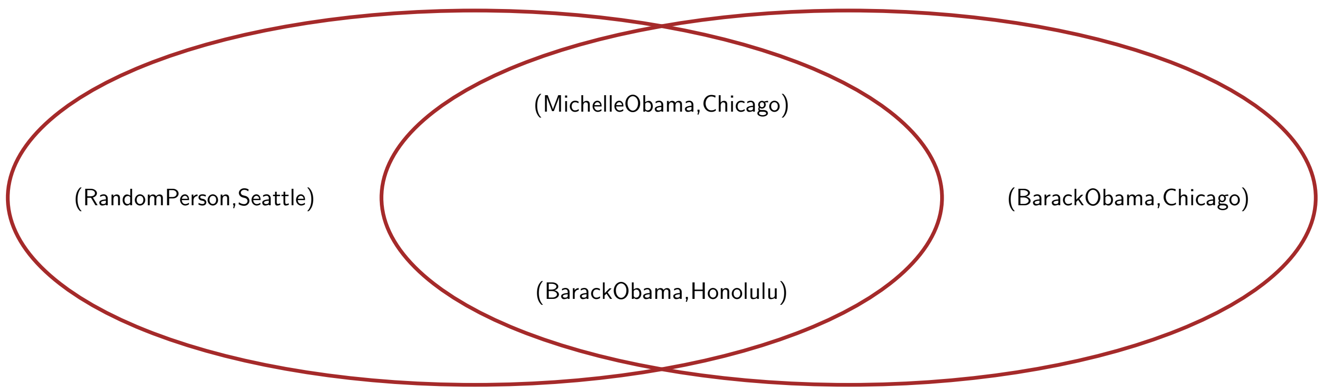
Binaries: paths of length 1 or 2 in the KB graph

Unaries: Type.x or Profession.x

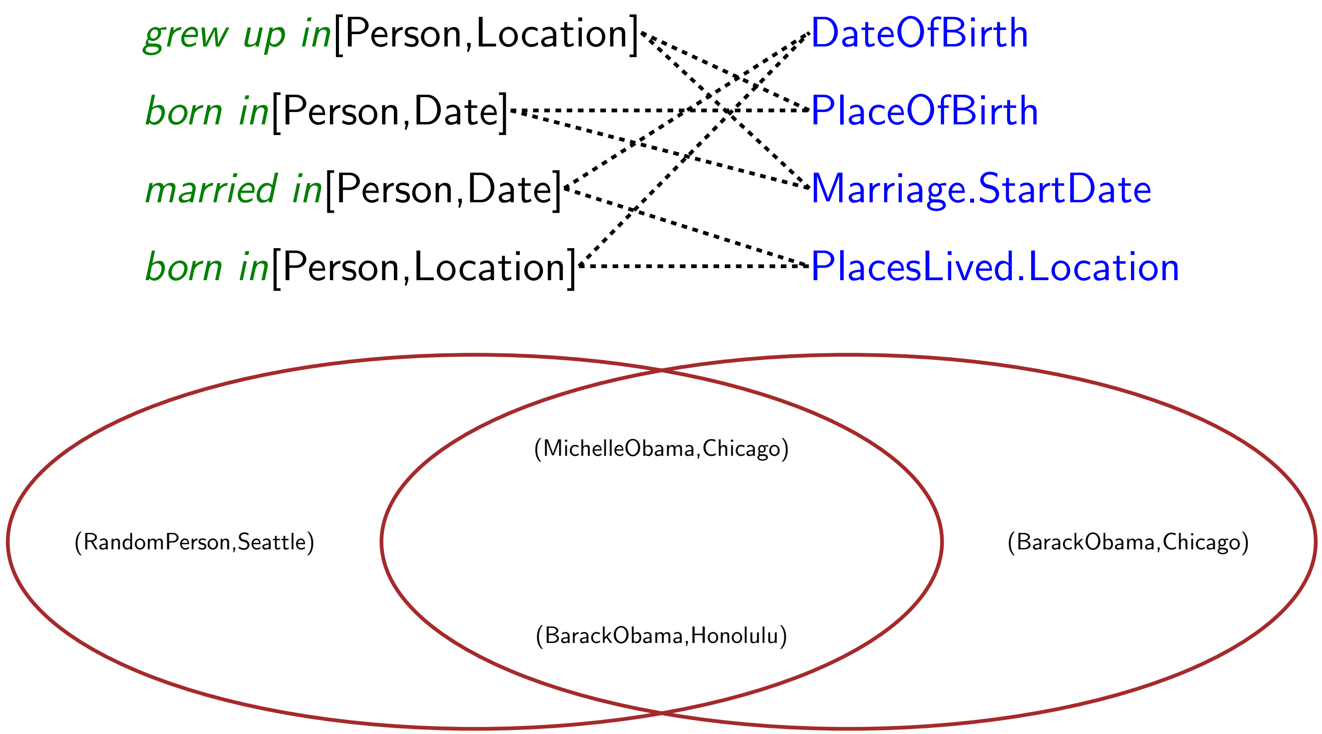
60,000 KB predicates

Alignment: match phrases and predicates

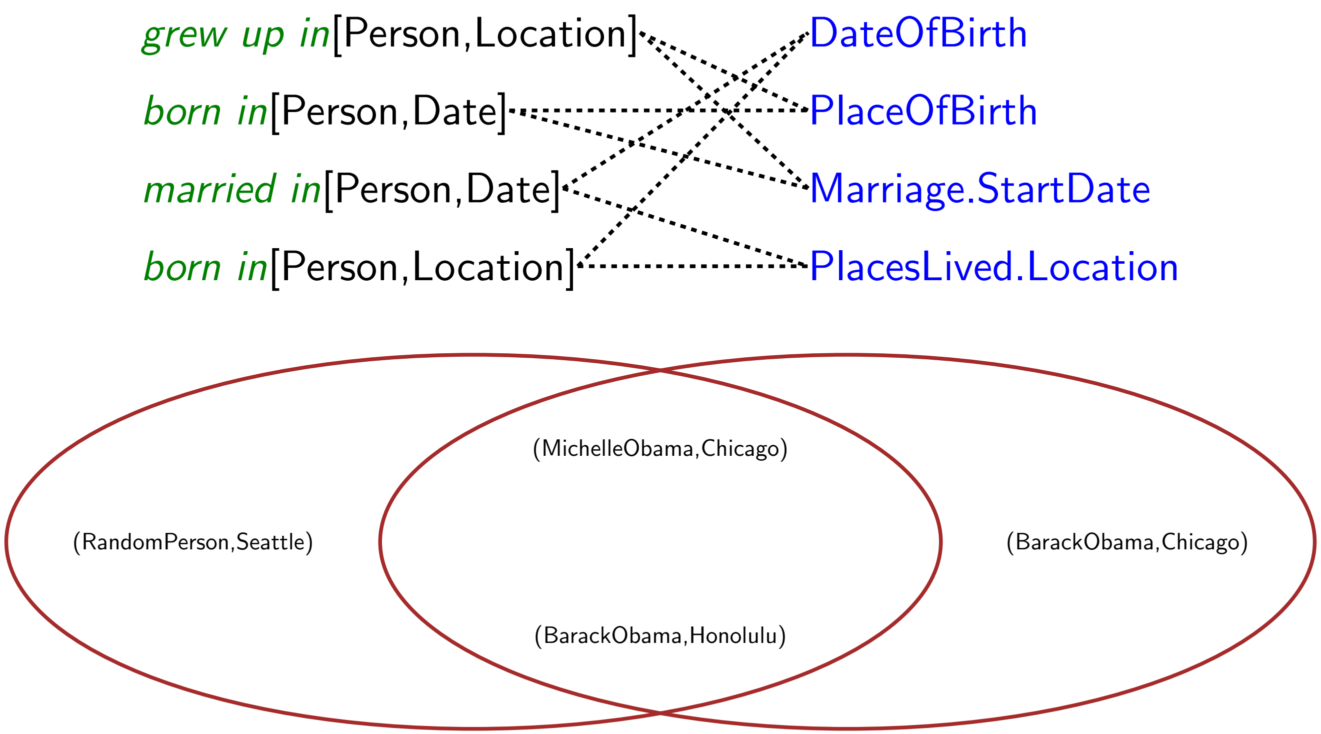
born in[Person,Location]···················PlacesLived.Location



Alignment: match phrases and predicates

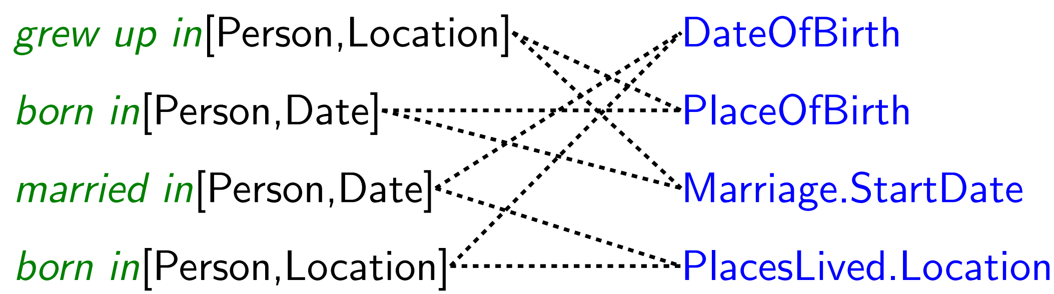


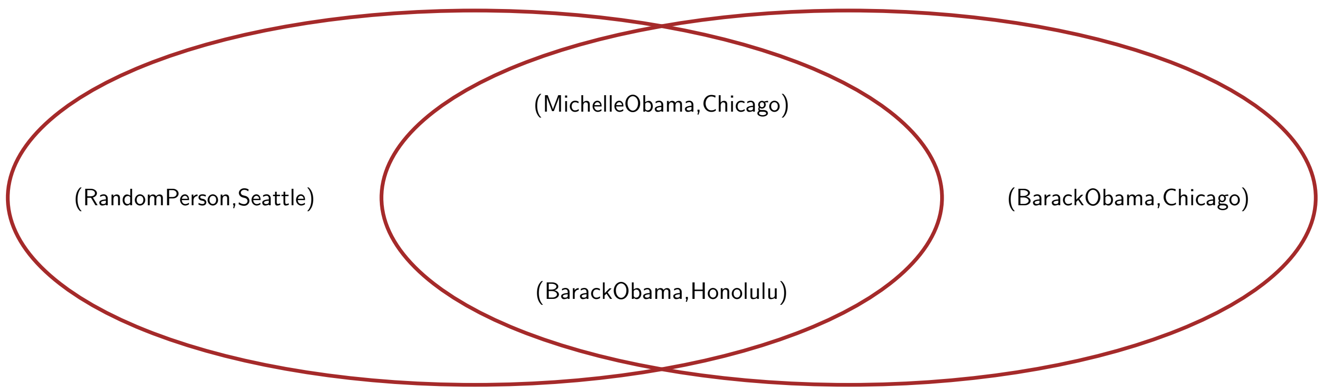
Alignment: match phrases and predicates



Lexicon: Mapping from phrases to predicates with features

Alignment: match phrases and predicates





Lexicon: Mapping from phrases to predicates with features

Alignment features

phrase-count:	15,765
predicate-count:	9,182
intersection-count:	6,048
KB-best-match:	0

Semantic parsing

- Setup
- Alignment
- **Bridging**
- Composition
- Dataset creation
- Experiments

Bridging

Often predicates are not expressed explicitly:

- *What government does Chile have?*
- *What is Italy's language?*
- *Where is Beijing?*
- *What is the cover price of X-men?*
- *Who did Humphrey Bogart marry in 1928?*

Bridging

Often predicates are not expressed explicitly:

- *What government does Chile have?*
- *What is Italy's language?*
- *Where is Beijing?*
- *What is the cover price of X-men?*
- *Who did Humphrey Bogart marry in 1928?*

Alignment: build coarse mapping from raw text

Bridging: use neighboring predicates / type constraints

Bridging 1: two unaries

Type.University

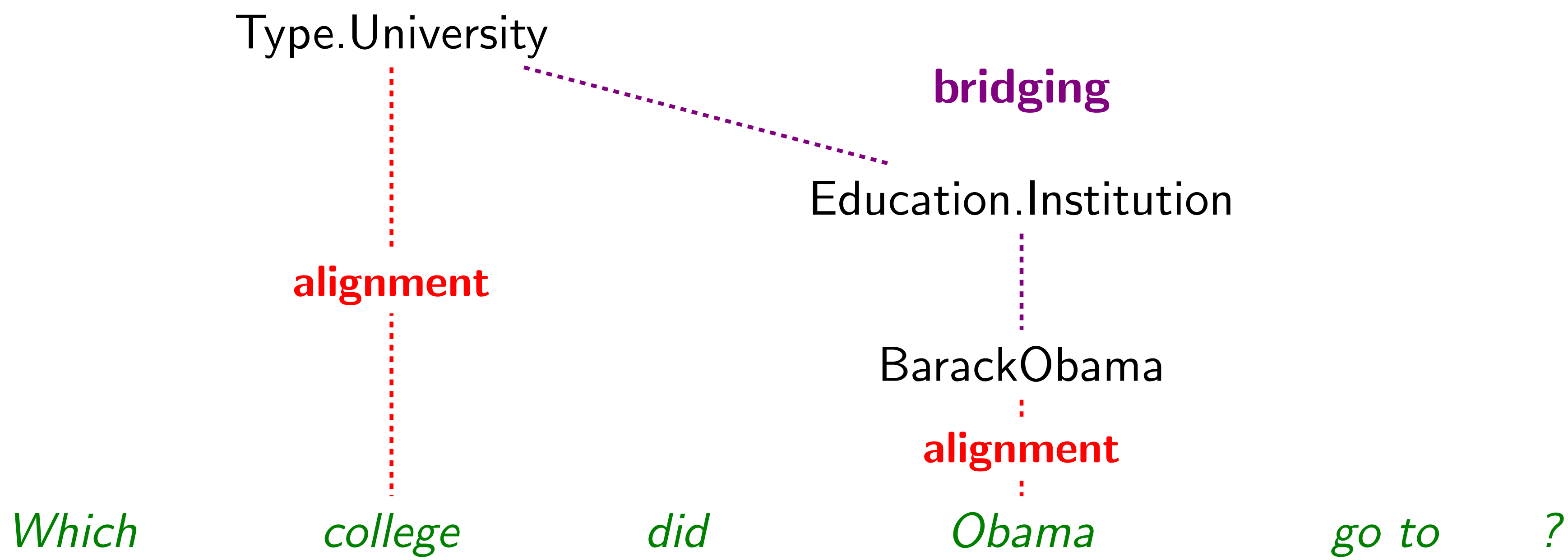
alignment

BarackObama

alignment

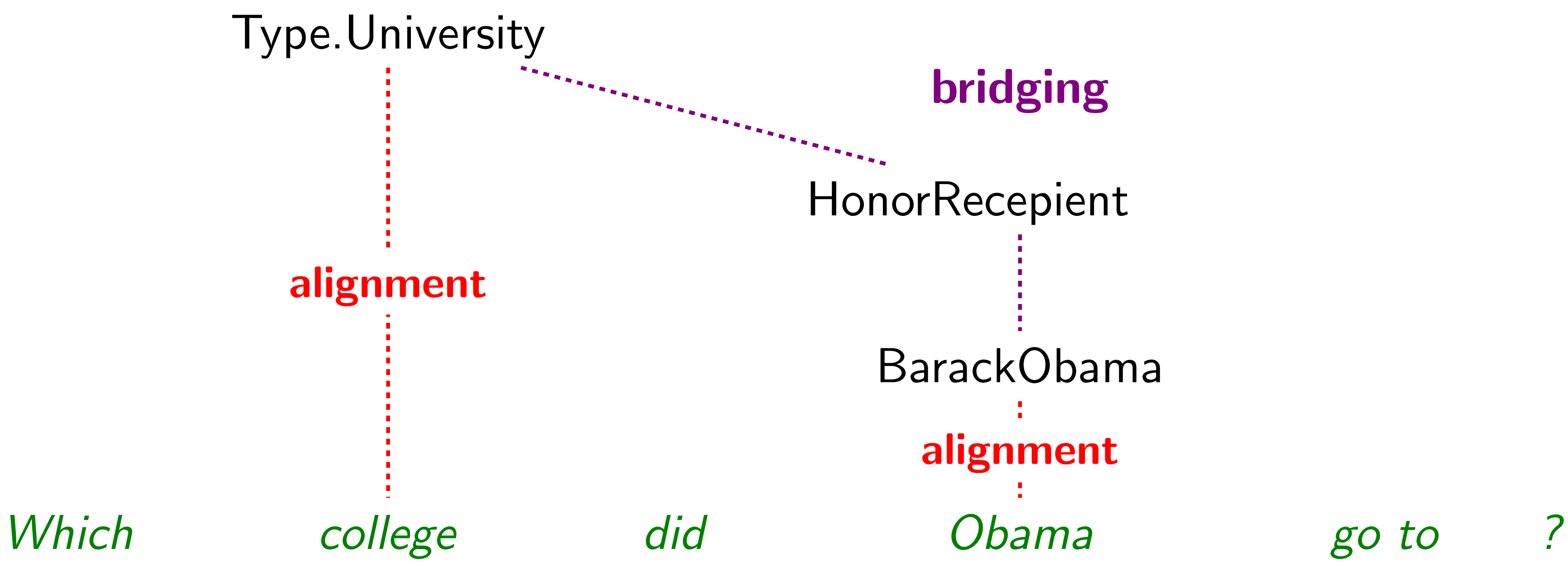
Which *college* *did* *Obama* *go to* *?*

Bridging 1: two unaries



$$\text{Type.University} \sqcap \text{Education.Institution.BarackObama}$$

Bridging 1: two unaries

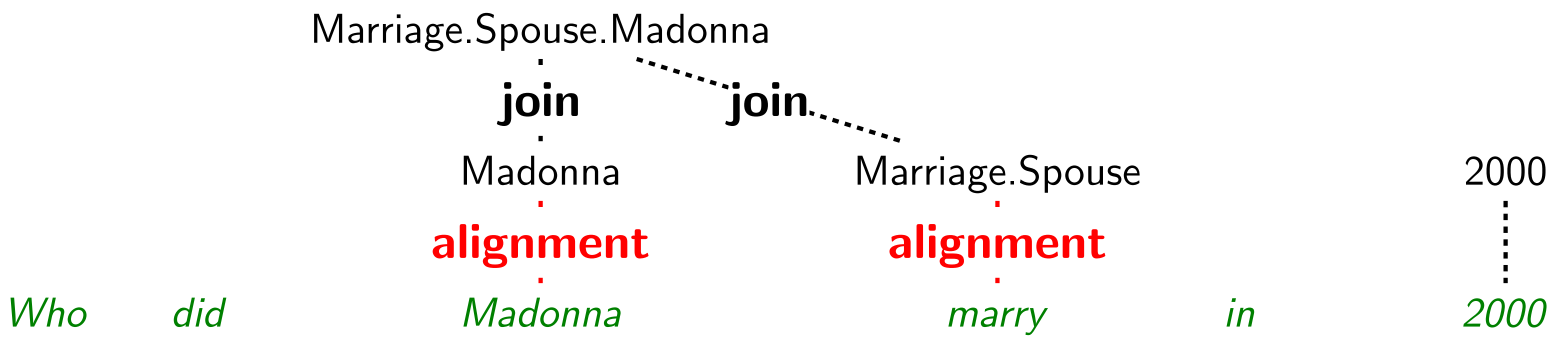


$$\text{Type.University} \sqcap \text{Education.Institution.BarackObama}$$

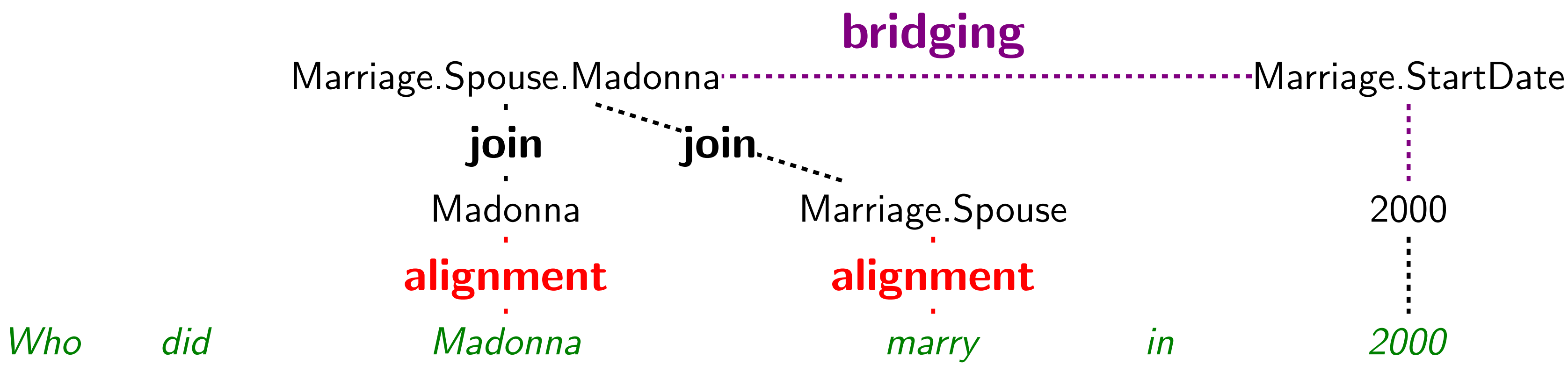
features

br-popularity	:11.37
br-two-unaries	: 1
br-education.institution:	1

Bridging 2: event modifiers



Bridging 2: event modifiers



$$\text{Marriage.(Spouse.Madonna} \sqcap \text{StartDate.2000)}$$

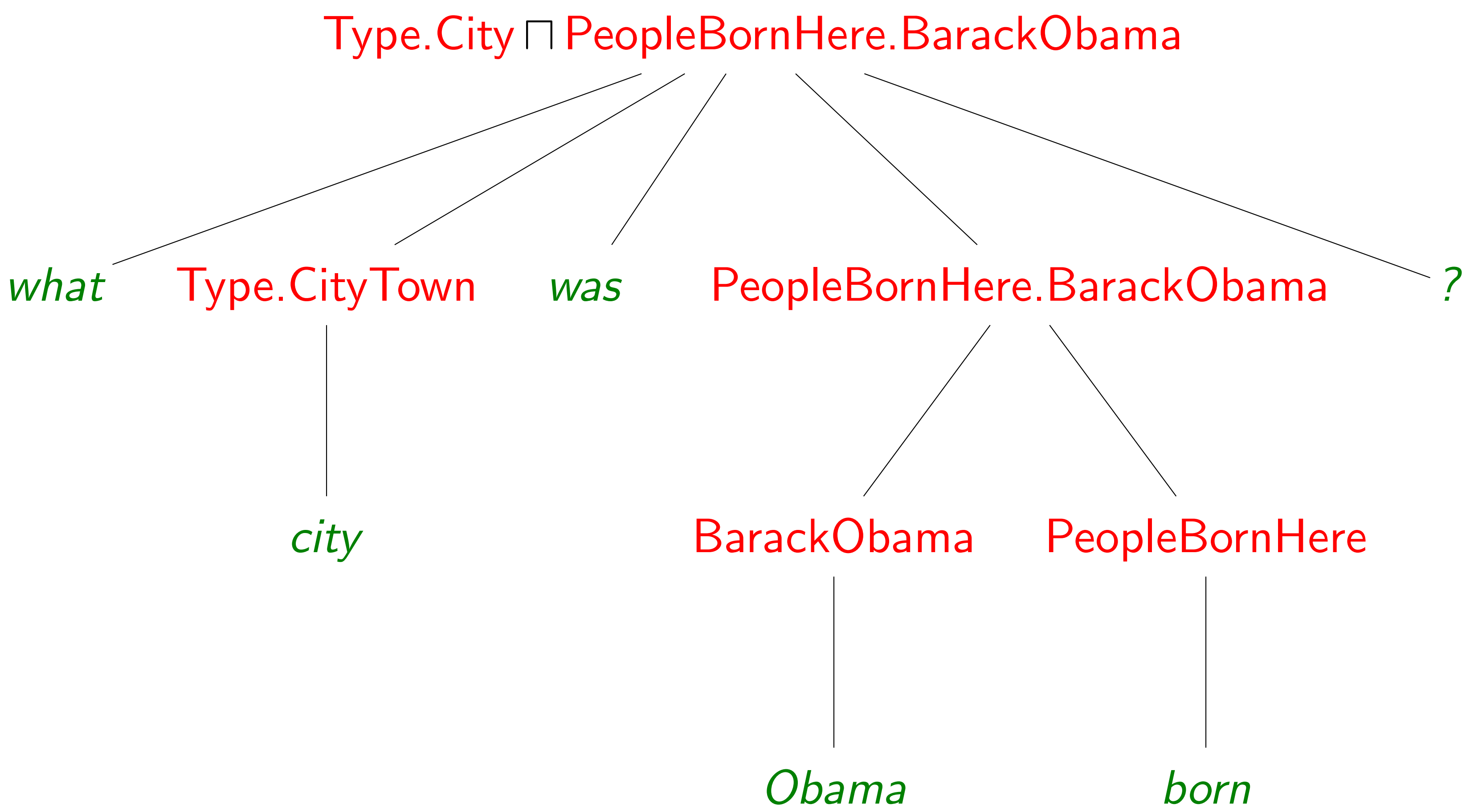
features

br-popularity	:	7.11
br-inject	:	1
br-startdate	:	1

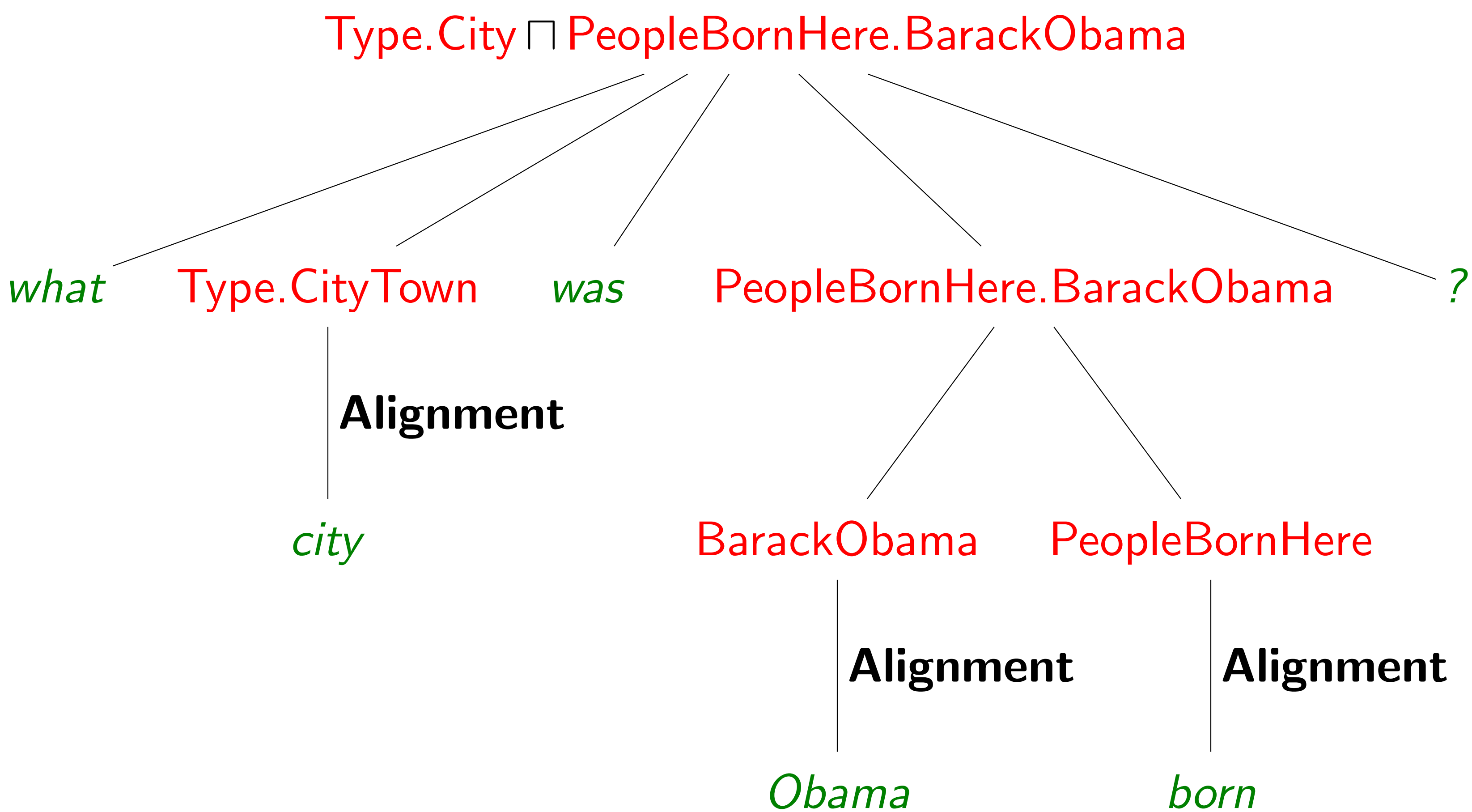
Semantic parsing

- Setup
- Alignment
- Bridging
- **Composition**
- Dataset creation
- Experiments

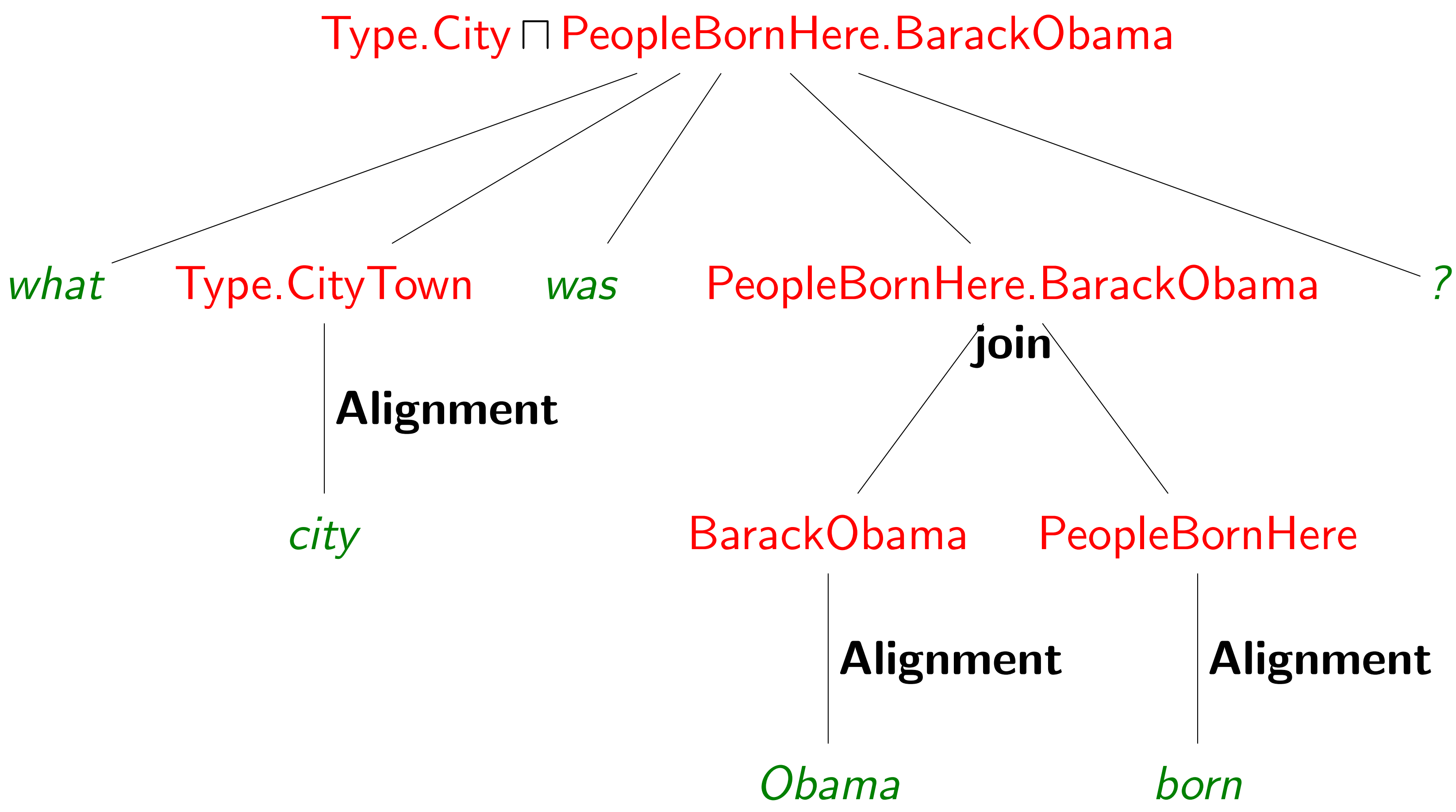
One derivation



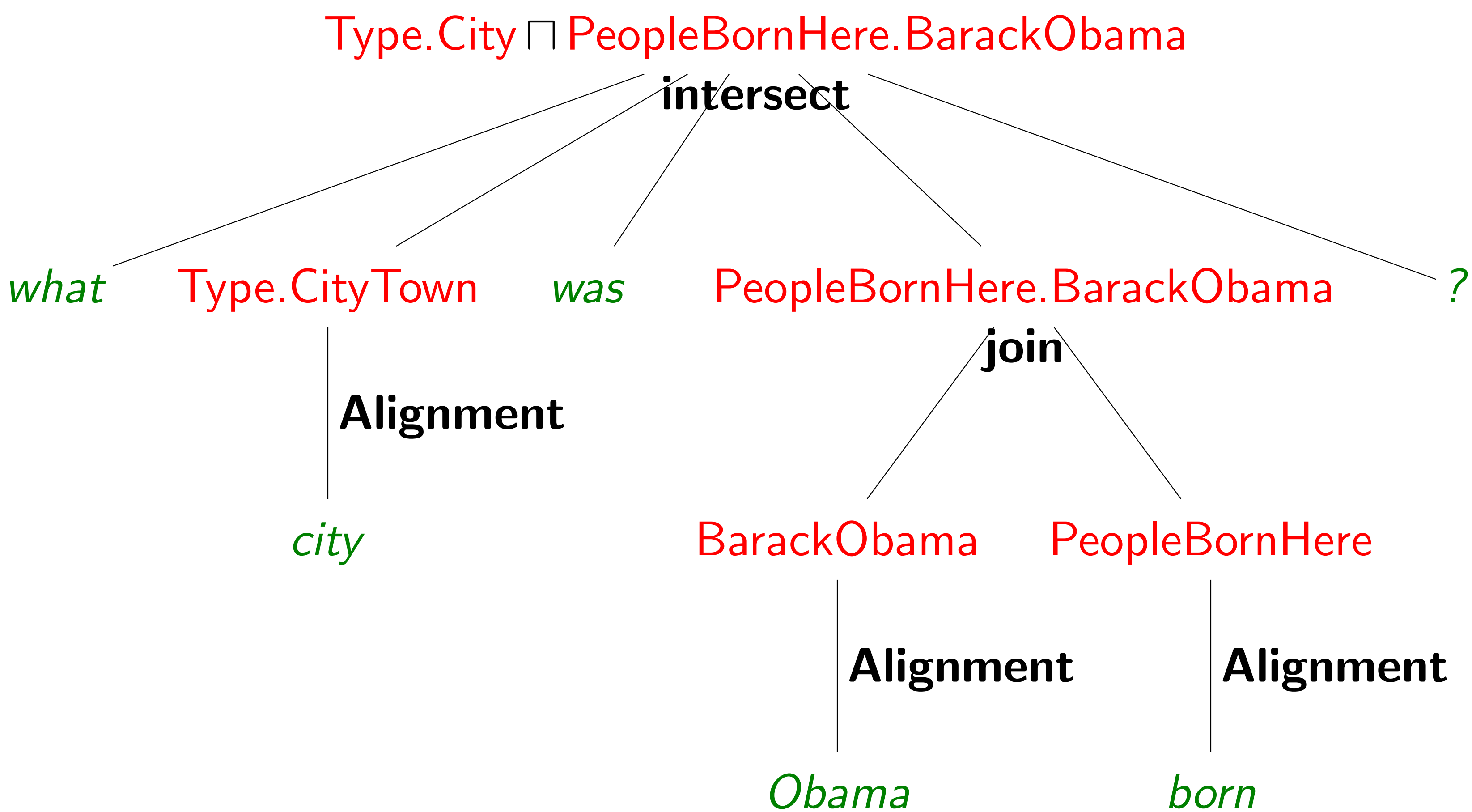
One derivation



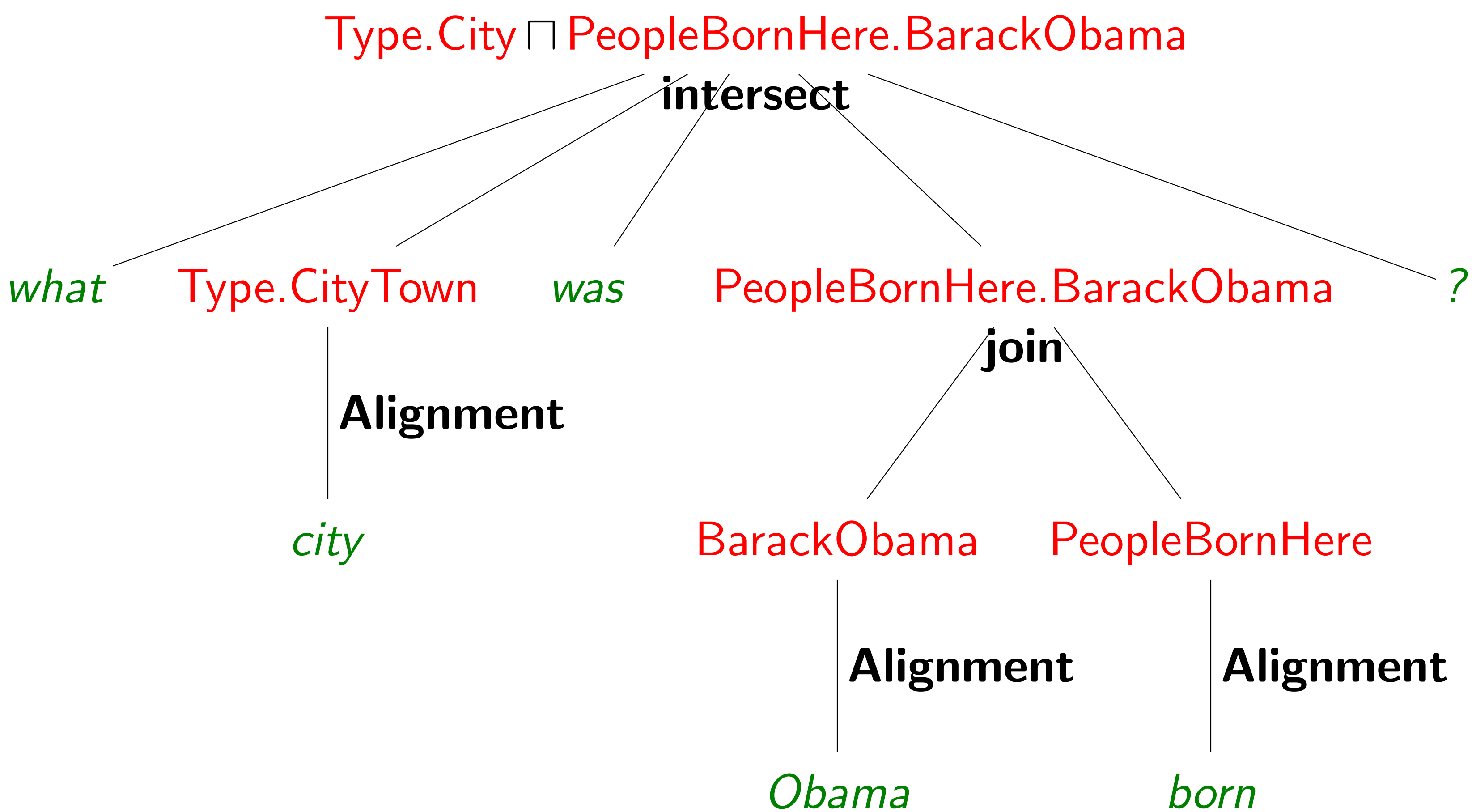
One derivation



One derivation



One derivation



Derivations are constructed using an over-general grammar

Model

Candidate derivations: $\mathcal{D}(x)$

Model

Candidate derivations: $\mathcal{D}(x)$

Model: distribution over derivations d given utterance x

$$p(d \mid x, \theta) = \frac{\exp(\phi(x,d)\cdot\theta)}{\sum_{d' \in \mathcal{D}(x)} \exp(\phi(x,d')\cdot\theta)}$$

Model

Candidate derivations: $\mathcal{D}(x)$

Model: distribution over derivations d given utterance x

$$p(d \mid x, \theta) = \frac{\exp(\phi(x,d) \cdot \theta)}{\sum_{d' \in \mathcal{D}(x)} \exp(\phi(x,d') \cdot \theta)}$$

Features:

- Alignment and bridging
- lexicalized
- syntactic
- denotation

Model

Candidate derivations: $\mathcal{D}(x)$

Model: distribution over derivations d given utterance x

$$p(d \mid x, \theta) = \frac{\exp(\phi(x,d) \cdot \theta)}{\sum_{d' \in \mathcal{D}(x)} \exp(\phi(x,d') \cdot \theta)}$$

Features:

- Alignment and bridging
- lexicalized
- syntactic
- denotation

Training (estimating θ):

- Stochastic gradient descent (AdaGrad)

Semantic parsing

- Setup
- Alignment
- Bridging
- Composition
- **Dataset creation**
- Experiments

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born? $\xrightarrow{\textbf{Google Suggest}}$ Barack Obama / Lady Gaga / Steve Jobs

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born? $\xrightarrow{\textbf{Google Suggest}}$ Barack Obama / Lady Gaga / Steve Jobs

Where was Steve Jobs born?

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born? $\xrightarrow{\textbf{Google Suggest}}$ Barack Obama, Lady Gaga, Steve Jobs

Where was Steve Jobs born?

Where was Steve Jobs _? $\xrightarrow{\textbf{Google Suggest}}$ born, raised, on the Forbes list

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born? $\xrightarrow{\textbf{Google Suggest}}$ Barack Obama / Lady Gaga / Steve Jobs

Where was Steve Jobs born?

Where was Steve Jobs _? $\xrightarrow{\textbf{Google Suggest}}$ born / raised / on the Forbes list

Where was Steve Jobs raised?

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born? $\xrightarrow{\textbf{Google Suggest}}$ Barack Obama
Lady Gaga
Steve Jobs

Where was Steve Jobs born?

Where was Steve Jobs _? $\xrightarrow{\textbf{Google Suggest}}$ born
raised
on the Forbes list

Where was Steve Jobs raised?

Result: popular web questions

WebQuestions: getting questions

Strategy: breadth-first search over Google Suggest graph

Where was Barack Obama born?

Where was _ born? $\xrightarrow{\textbf{Google Suggest}}$ Barack Obama / Lady Gaga / Steve Jobs

Where was Steve Jobs born?

Where was Steve Jobs _? $\xrightarrow{\textbf{Google Suggest}}$ born / raised / on the Forbes list

Where was Steve Jobs raised?

Result: popular web questions

Answers were obtained through crowdsourcing (AMT)

Dataset comparison

Free917 [Cai & Yates, 2013]: 917 examples, 2,036 word types

What is the engine in a 2010 Ferrari California?

What was the cover price of the X-men Issue 1?

- Generate questions based on Freebase facts

WebQuestions [our work]: 5,810 examples, 4,525 word types

What character did Natalie Portman play in Star Wars?

What kind of money to take to Bahamas?

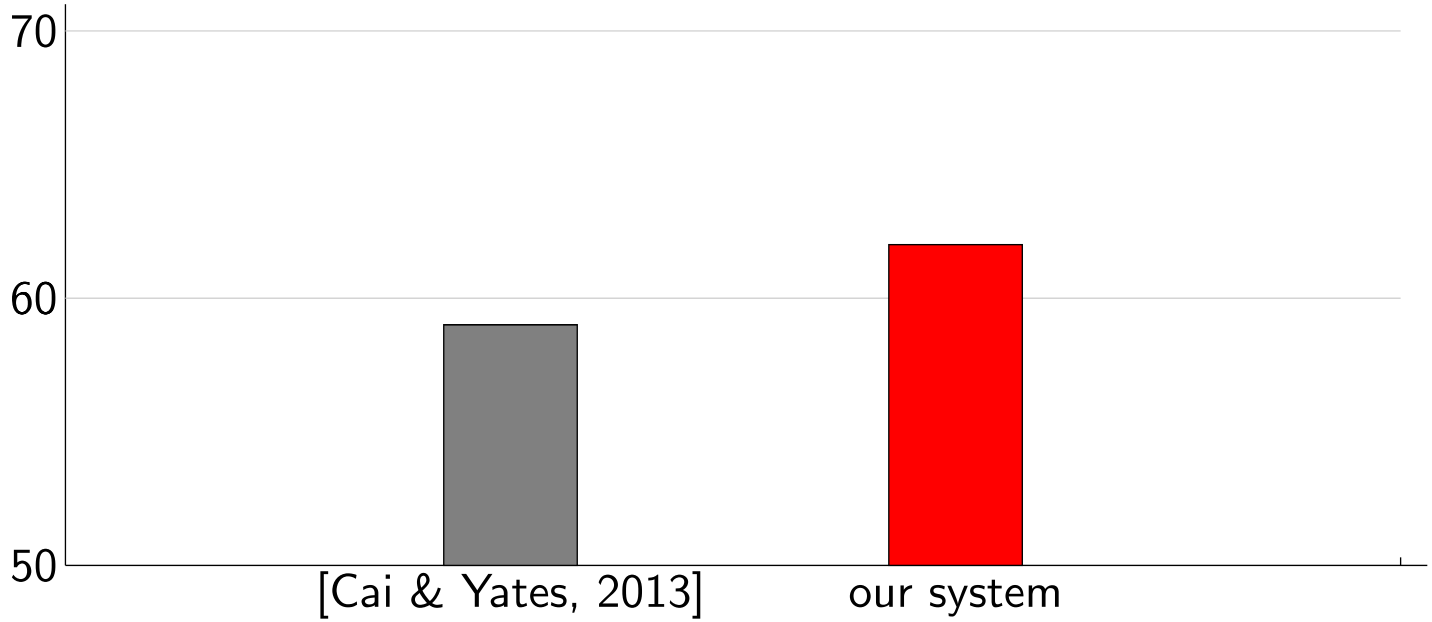
What did Edward Jenner do for a living?

- Generate questions from Google $\Rightarrow$ less formulaic

Semantic parsing

- Setup
- Alignment
- Bridging
- Composition
- Dataset creation
- **Experiments**

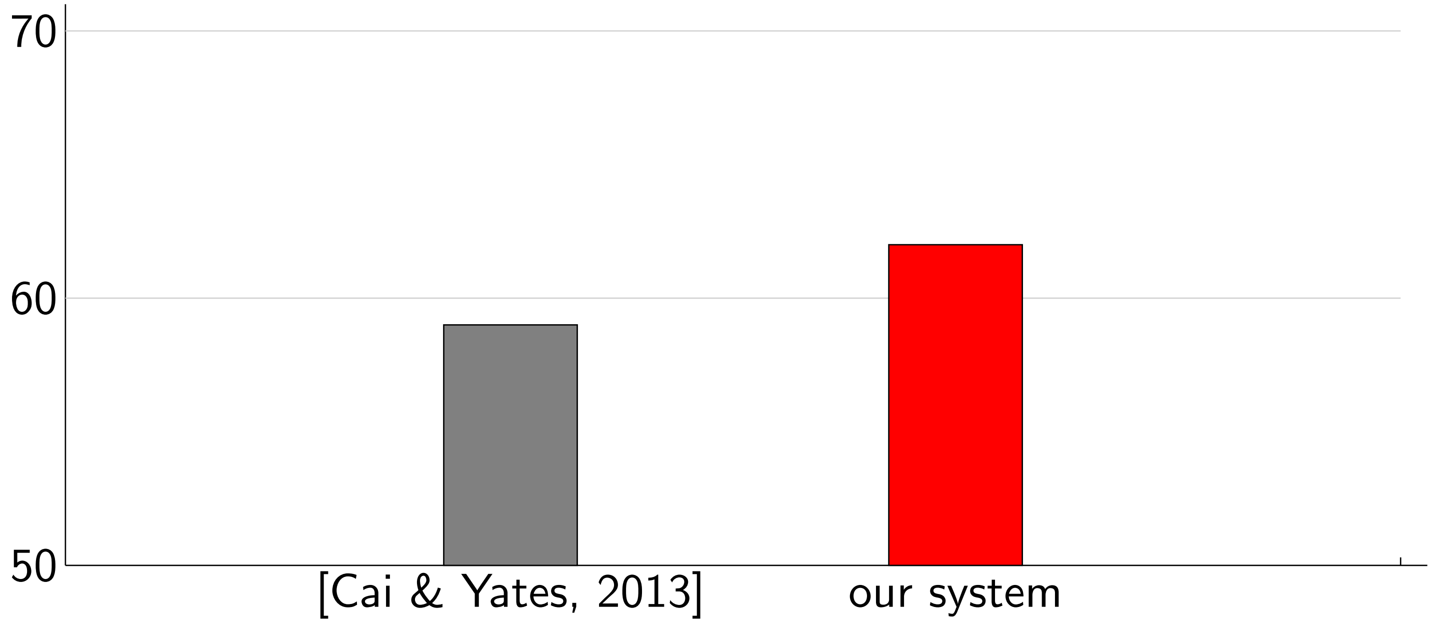
Results on Free917



Differences:

- We train from answers only, CY13 uses logical forms
- We use 12K binary predicates, CY13 used 2k binary predicates

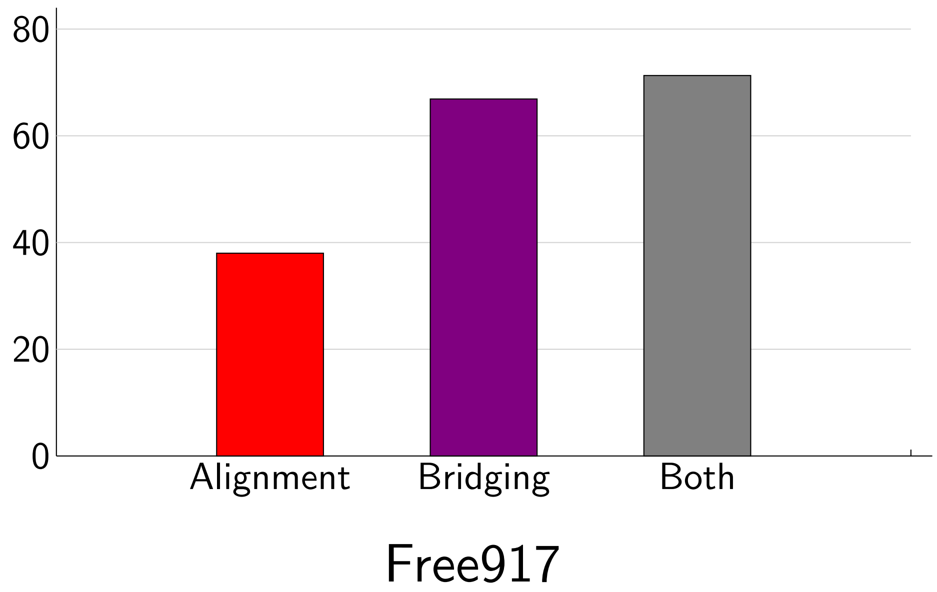
Results on Free917

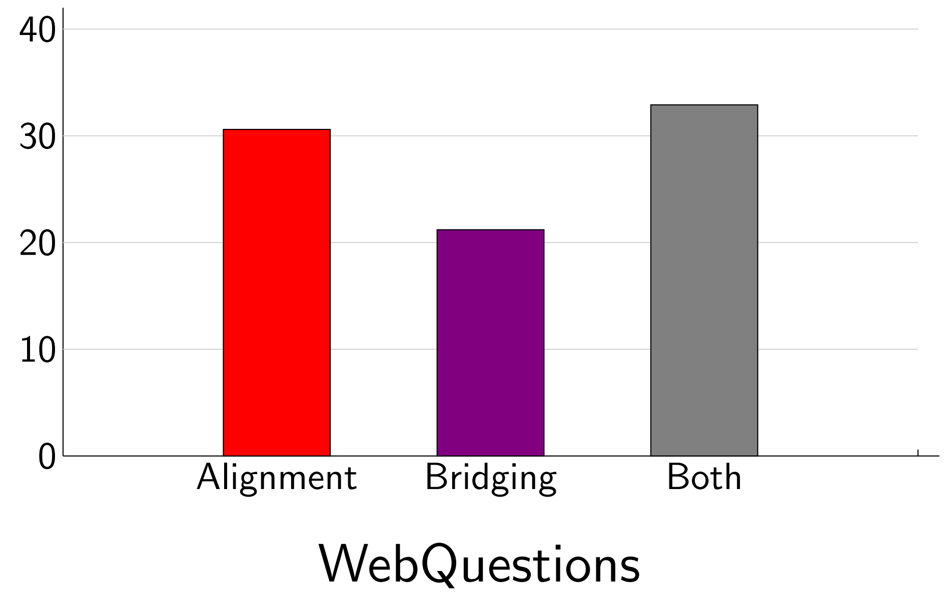


Differences:

- We train from answers only, CY13 uses logical forms
- We use 12K binary predicates, CY13 used 2k binary predicates
- Kwiatkowski et al. obtain larger improvement

Impact of alignment and bridging

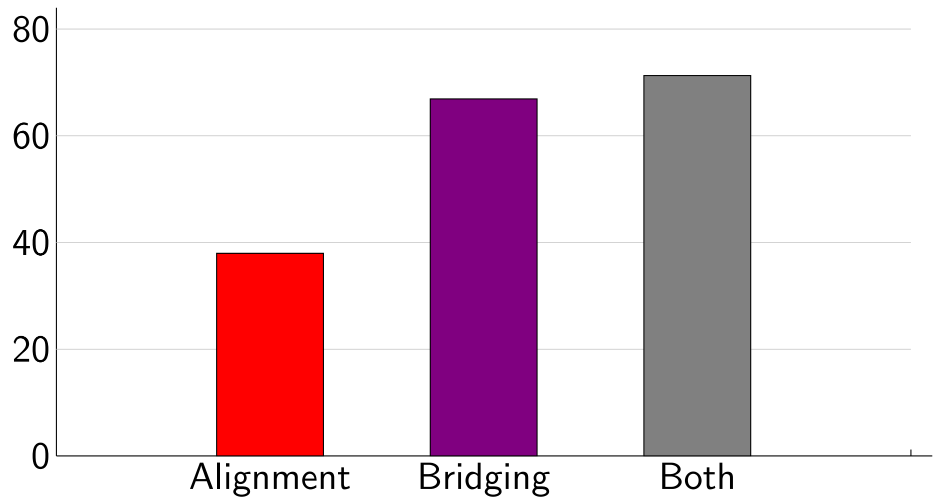




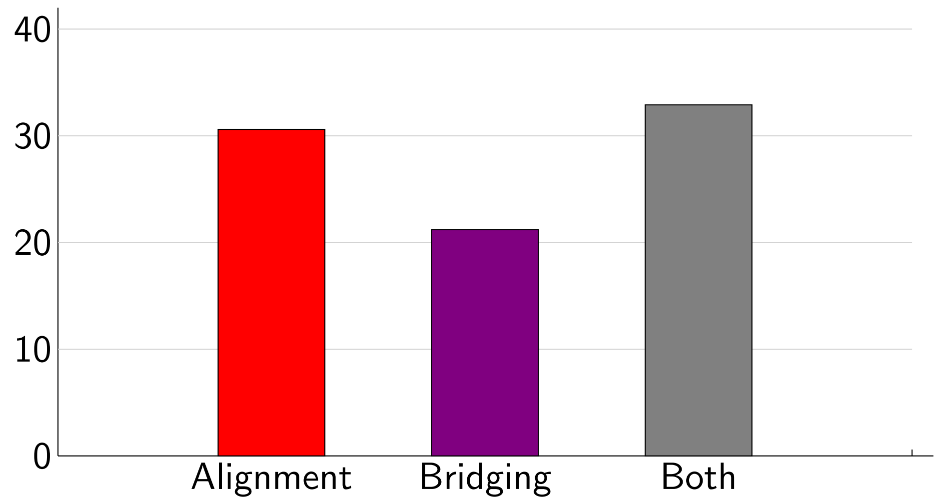
Conclusions:

- Bridging more important for Free917
- Alignment more important for WebQuestions

Impact of alignment and bridging



Free917



WebQuestions

Conclusions:

- Bridging more important for Free917
- Alignment more important for WebQuestions

Test accuracy on webQuestions: 35.7

Summary

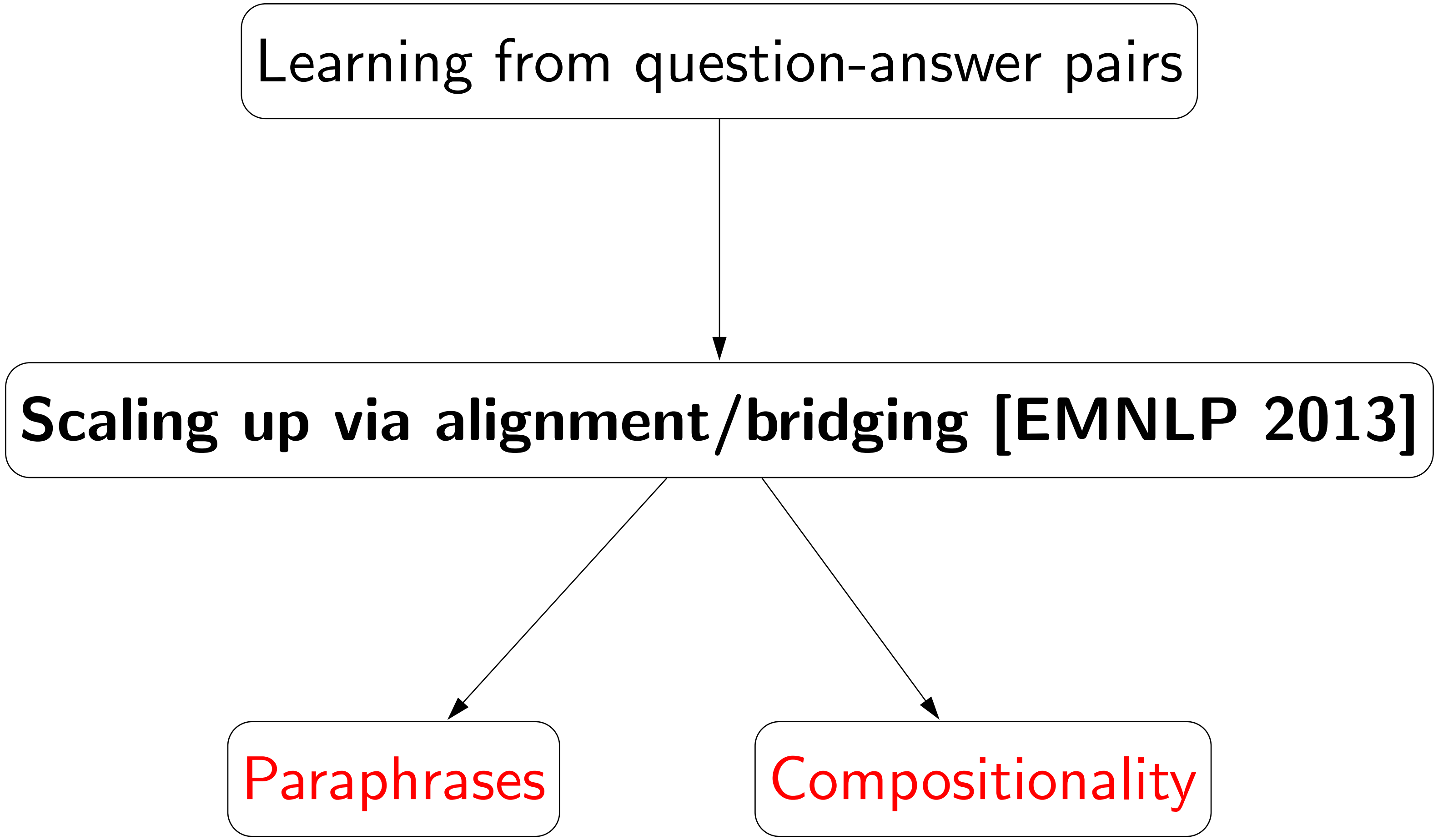
Learning from question-answer pairs

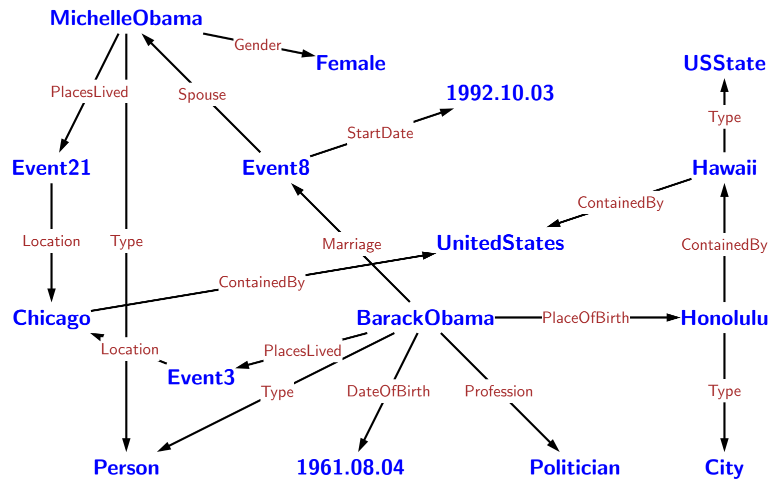
Summary

Learning from question-answer pairs

Scaling up via alignment/bridging [EMNLP 2013]

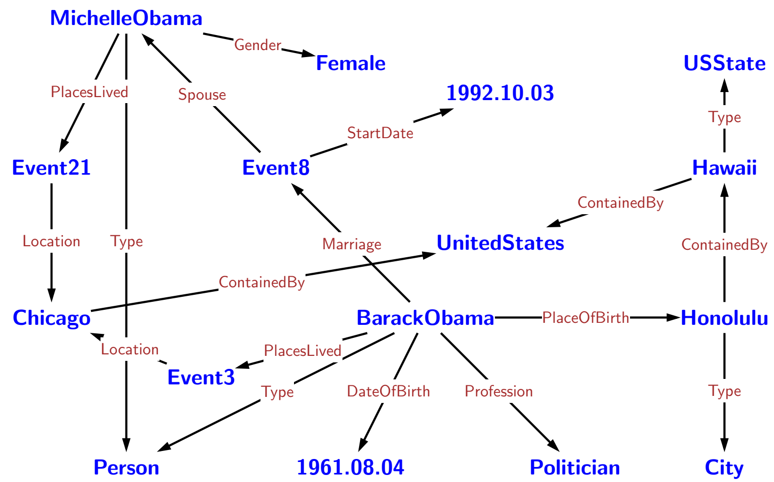
Summary


Learning from question-answer pairs
Scaling up via alignment/bridging [EMNLP 2013]
Paraphrases
Compositionality



All data and code:

http://www-nlp.stanford.edu/software/sempre/



All data and code:

http://www-nlp.stanford.edu/software/sempre/

Thank you!