

به نام خدا

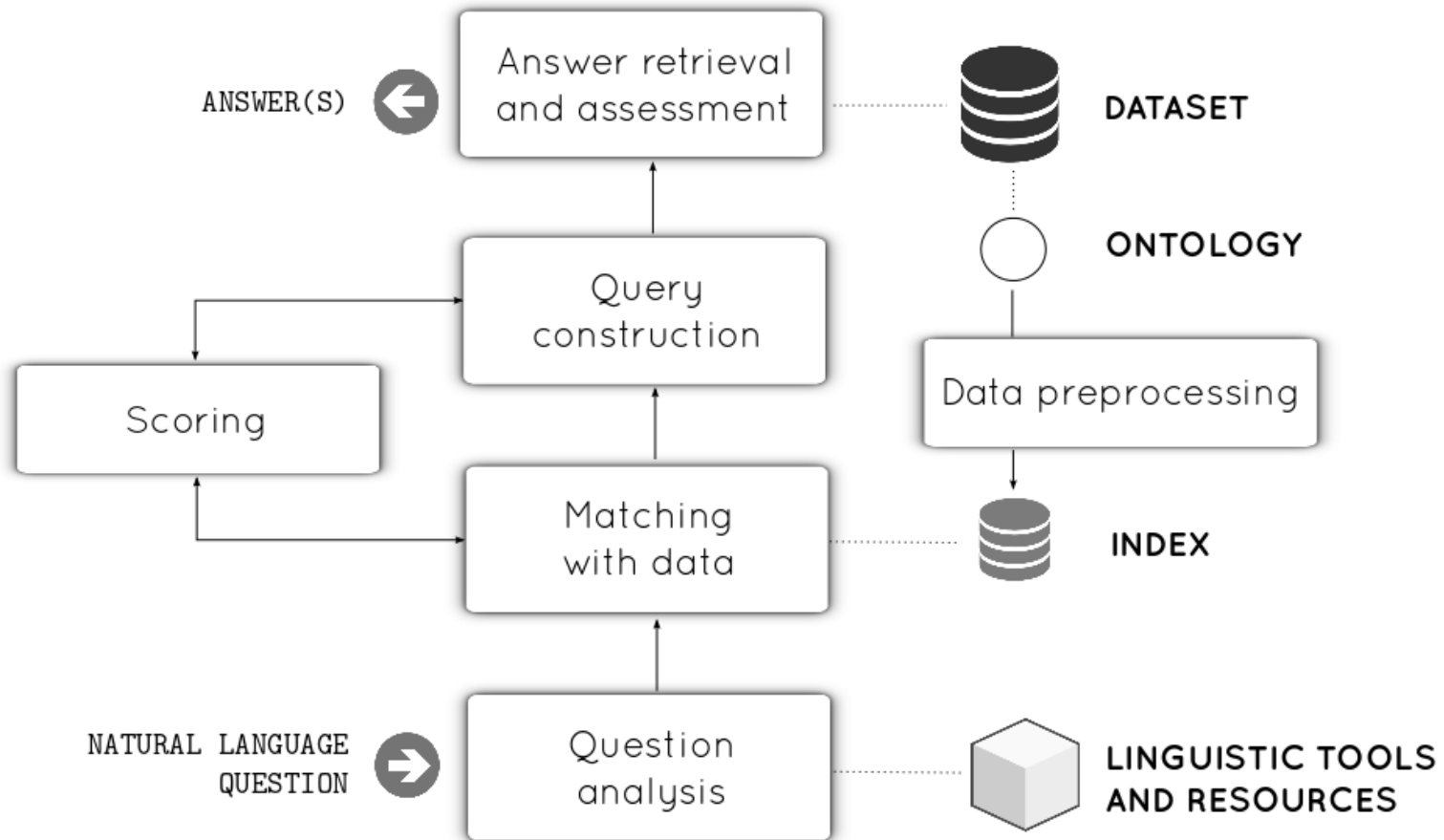
سیستم‌های پرسش و پاسخ

- اطلاعات در سطح اینترنت روز به روز در حال افزایش است.
- استفاده از موتورهای جستجو برای بازیابی اطلاعات در سطح اینترنت.
- مشکلات موتورهای جستجو
- راه حل:
- سیستم‌های پرسش و پاسخ

- تحقیقات در مورد سیستم‌های پرسش و پاسخ از سال ۱۹۶۰ آغاز شده است.
- با ظهور وب معنایی، تحول بزرگی بر روی سیستم‌های پرسش و پاسخ رخ داد.
- چالش اصلی در QA تبدیل زبان طبیعی به فرمت قابل فهم برای کامپیوتر است.

انواع سیستم‌های پرسش و پاسخ براساس محدوده کار:

- سیستم‌های باز
- سیستم‌های بسته



روش‌های مختلف تبدیل زبان طبیعی به فرمت مورد نظر:

□ تحلیل‌های زبان‌شناسی

- در این سیستم‌ها از ریخت‌شناسی و تجزیه و تحلیل نحوی استفاده می‌شود.
- استفاده از تجزیه‌گرها به ما کمک خواهد کرد تا ساختار جمله را بازنمایی کنیم.
- سیستم PYTHIA از این روش استفاده کرده است.

□ تحلیل‌های معنایی

- روابط معنایی موجود در متن سوال را استخراج می‌کنند.
- به عبارت دیگر به دنبال یافتن سه‌گانه‌ها در متن سوال هستند
- این سیستم‌ها سوال را به صورت سه‌گانه نمایش می‌دهند و بعد با کمک متریک‌های شباهت، زیرگرافی از داده RDF را که با این سه‌گانه‌ها تطبیق دارد، به عنوان جواب برمی‌گردانند.

□ روش‌های مبتنی بر قالب

سیستم پرسش و پاسخ QASYO

- یک سیستم پرسش و پاسخ مبتنی بر آنتولوژی است.
- از آنتولوژی YAGO برای پاسخ‌دهی به پرسش‌های کاربران استفاده می‌کند.
- ورودی پرسش به زبان طبیعی و تبدیل آن به کوئری T-SQL و نمایش پاسخ به کاربر.

ساختار سلسه مراتب WordNet را با میزان پوشش بالای ویکی پدیا ترکیب می کند.

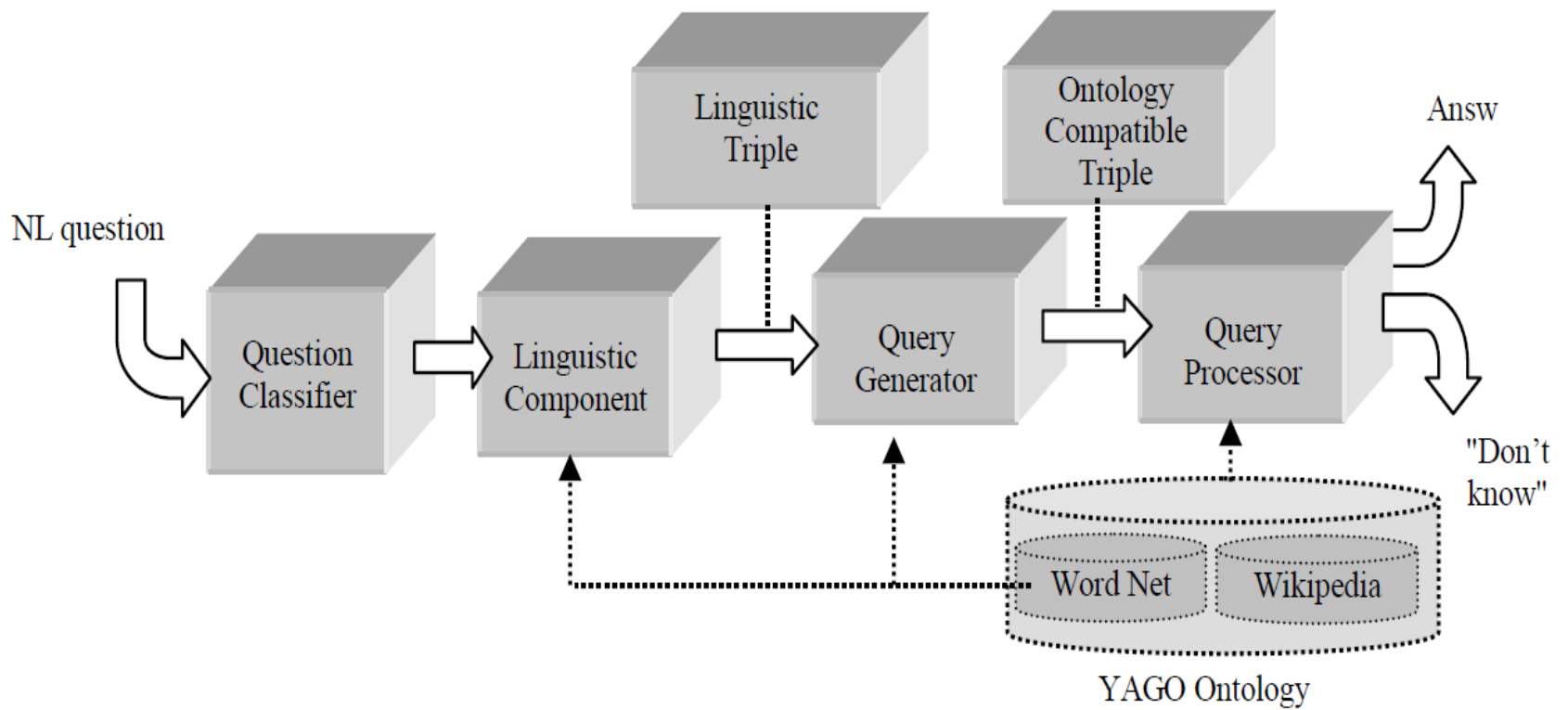
با دقت ۹۵ درصد این ترکیب را انجام می دهد.

Relation 92

Classes 224,391

Individuals 1,531,558

معماری سیستم QASYO



Question Classifier

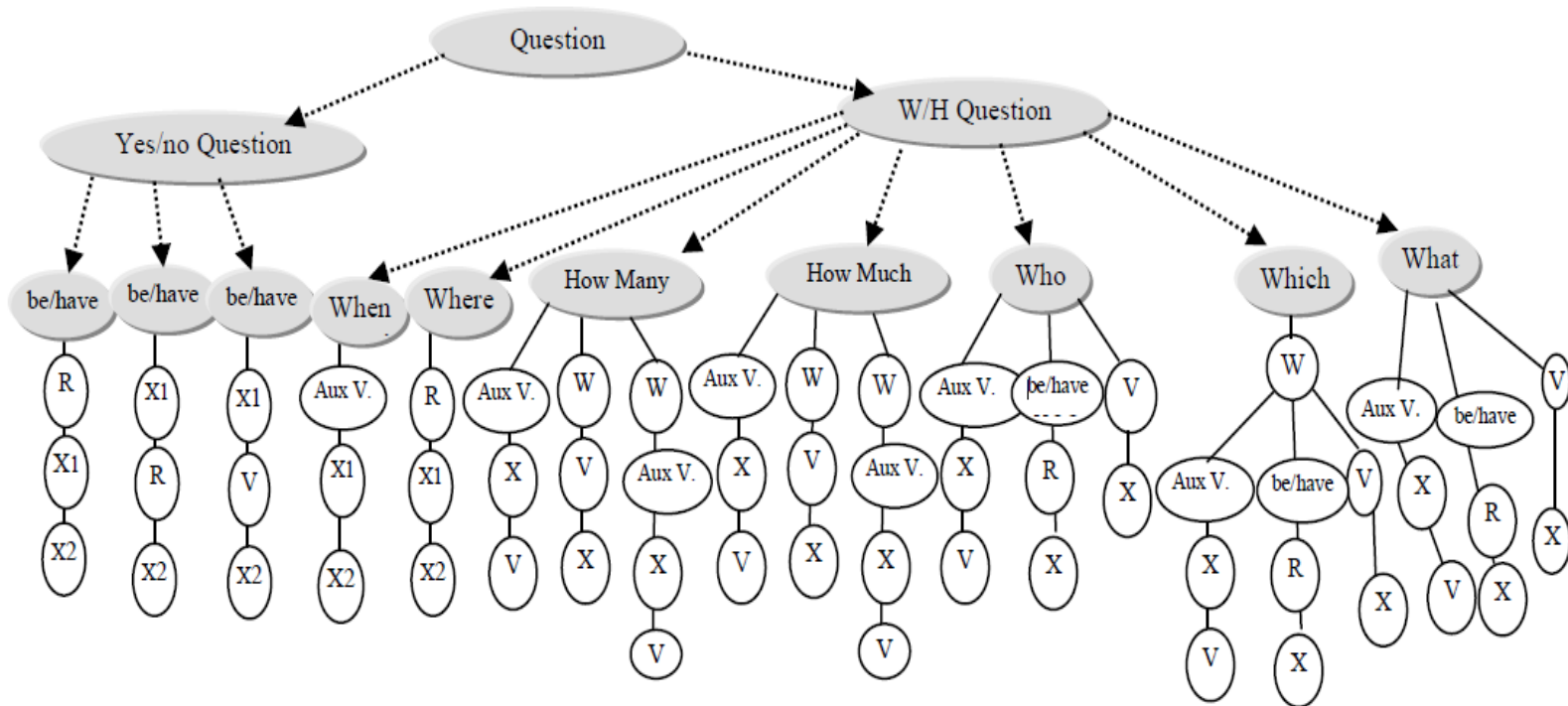
بر اساس نوع پرسش و جواب مورد انتظار، پرسش‌ها را طبقه‌بندی می‌کند.
QASYO فقط به سوالات YES/No و wh پاسخ می‌دهد.
به عنوان مثال :

اگر پرسش با کلمه who شروع شود، جواب مورد انتظار Person/Org
when: به time/date
Where: به مکان

<adjective> How: از روی adjective مشخص می‌شود. مثلا How many به تعداد
اشاره دارد.

Linguistic Component

برای تبدیل NL به سه گانه <subject,relation,object> استفاده می شود.
 کامپوننت زبانی QASYO ۲۰ گروه برای را مشخص می کند.



R شامل nounها هستند که در جدول WR در YAGO، relationها را نشان می‌دهد.

Table 5. Example of the WR Table

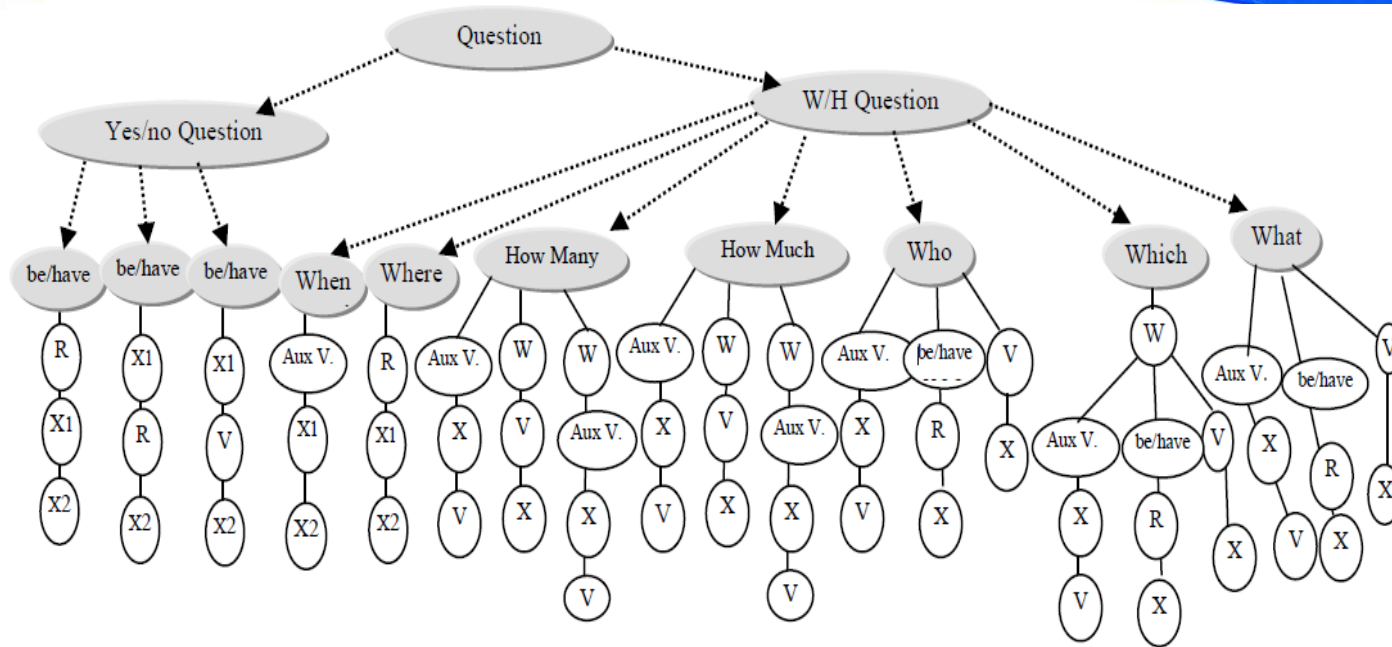
| ID | Word | Related Relation |
|----|----------|------------------|
| 1 | Currency | HASCURRENCY |
| 2 | motto | HASMOTTO |

V شامل Verbها که در جدول Verb در YAGO، relationها را نشان می‌دهد.

Table 6. Example of the Verb Table

| ID | Verb | Related Relation |
|----|------|------------------|
| 1 | Act | ACTEDIN |
| 2 | Die | DIEDONDATE |

W، شامل کلماتی مانند: Country, Scientist و ... است.
X هر کلمه‌ای خارج از اینها.



1- What is currency of brazil?

2- which scientist works at university of cambridge?

برای اینکه سیستم بتواند به سوال پاسخ دهد، باید شرایط زیر برقرار باشد:

| ARG1 | ARG2 | Case |
|---------|---------|--|
| known | known | This is a known fact [used only in YES/NO questions] |
| unknown | known | Valid question |
| known | unknown | Valid question |
| unknown | unknown | Invalid question |

بعد از آنکه رابطه از جدول WR و یا Verb استخراج شد، باید در جدول DRT، domain و range آن رابطه نیز استخراج شود.

| Domain | Relation | Range |
|--------|----------|----------|
| Person | Born in | Location |

حال با استفاده از رابطه و S/O استخراج شده و همچنین دامنه و برد استخراج شده می توان کوئری زیر را تولید کرد.

```
SELECT [ARG1 or ARG2] from FACTS where RELATIONS="retrieved relation" AND  
[ARG2 or ARG1=" the intended S/O"]
```

What is the currency of Brazil?

[Country] -> [HasCurrency] -> [???].

SELECT ARG2 from FACTS where RELATION=[HasCurrency] AND ARG1=[country].

SELECT ARG2 from FACTS where RELATION=[HasCurrency] AND ARG1=[Brazil].

Which scientist works at University of Cambridge?

[???] -> [worksAt] -> [University].

SELECT ARG1 from FACTS where RELATION=[worksAt] AND ARG2=[University].

SELECT ARG1 from FACTS where RELATION=[worksAt] AND ARG2=[Cambridge University].

Where was Max Planck born?

[Person] -> [bornIn] -> [???].

SELECT ARG2 from FACTS where RELATION=[bornIn] AND ARG1=[Person].

SELECT ARG2 from FACTS where RELATION=[bornIn] AND ARG1=[Max Planck].

یک سیستم Tempalte based بر روی داده‌های RDF

□ Tempalte ها به صورت خودکار ایجاد می‌شوند.

□ توسط مقایسه مجموعه بزرگ کوئری‌های sparql و مجموعه بزرگ پرسش‌های زبان طبیعی ایجاد می‌شوند.

□ پرسش زبان طبیعی به دلیل وجود ابهام در زبان طبیعی به uncertain graph تبدیل می‌شوند.

□ کوئری‌های اسپارکل نیز به certain graph تبدیل می‌شوند.

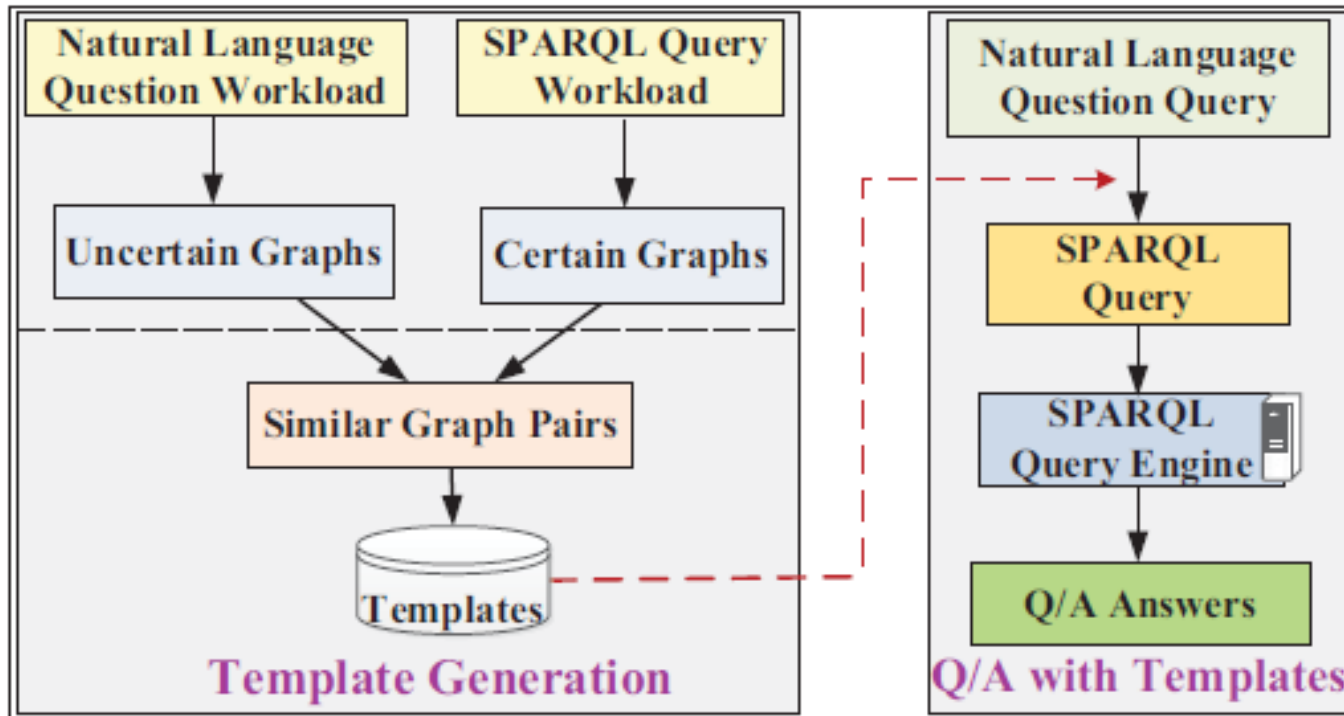
چالش :

چگونه Template ها را ایجاد کنیم؟

هدف: (q,n) پرسش زبان طبیعی n : کوئری اسپارکل q :

چگونه Template مربوطه را برای پرسش جدید پیدا کنیم؟

معماری کلی



تولید Template:

تولید semantic query graph: سه گانه‌هایی به شکل $\langle \text{rel}, \text{arg1}, \text{arg2} \rangle$
مثال:

Which actor from USA is married to Michael Jordan born in city of NY ?

$\langle \text{from}, \text{which actor}, \text{USA} \rangle$

$\langle \text{be married to}, \text{which actor}, \text{Michael Jordan} \rangle$

$\langle \text{born in}, \text{Michael Jordan}, \text{city} \rangle$

$\langle \text{located}, \text{city}, \text{NY} \rangle$

ابهام:

Michael Jordan: NBA-Player

Professor

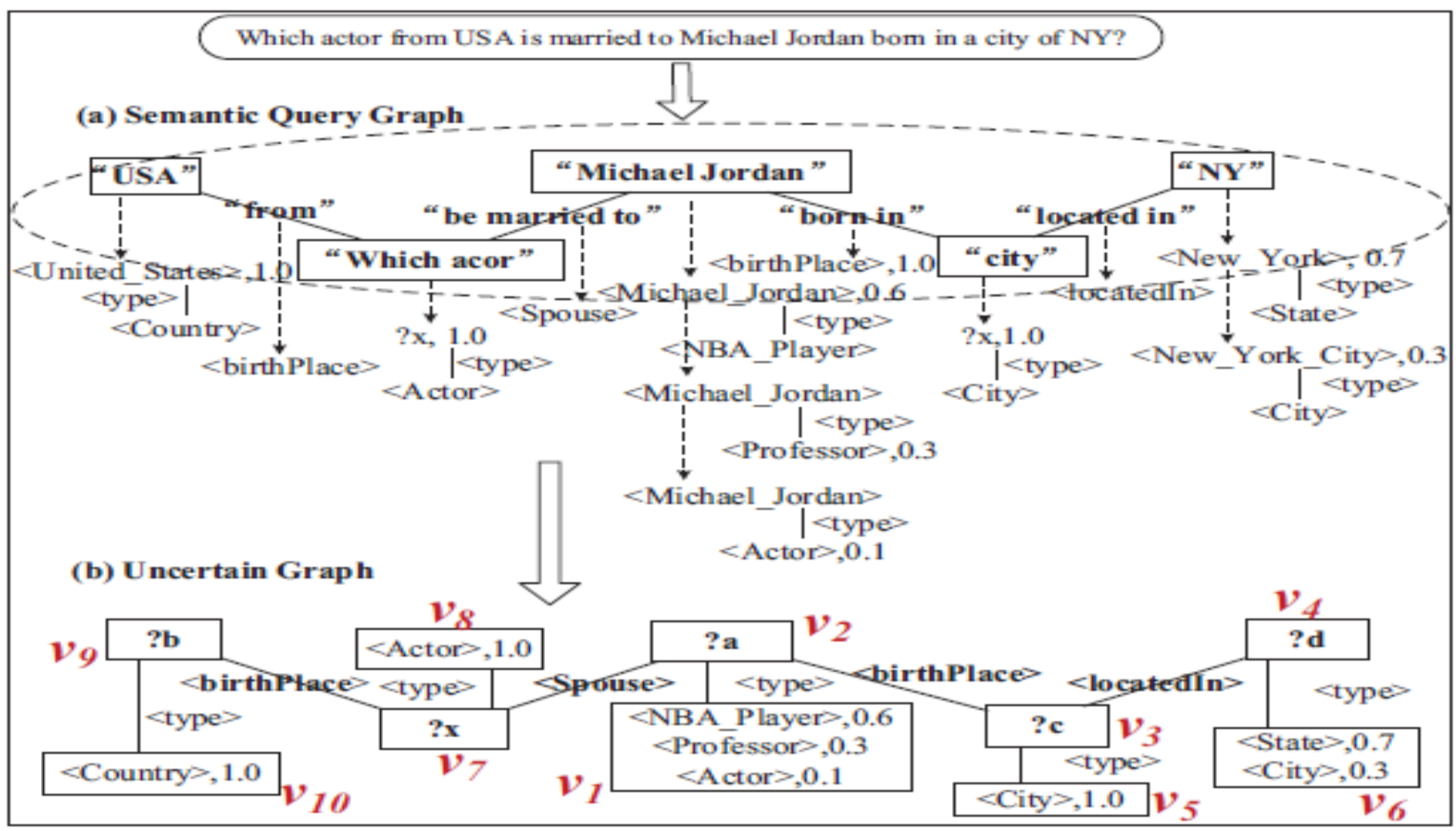
Actor

NY: city

state

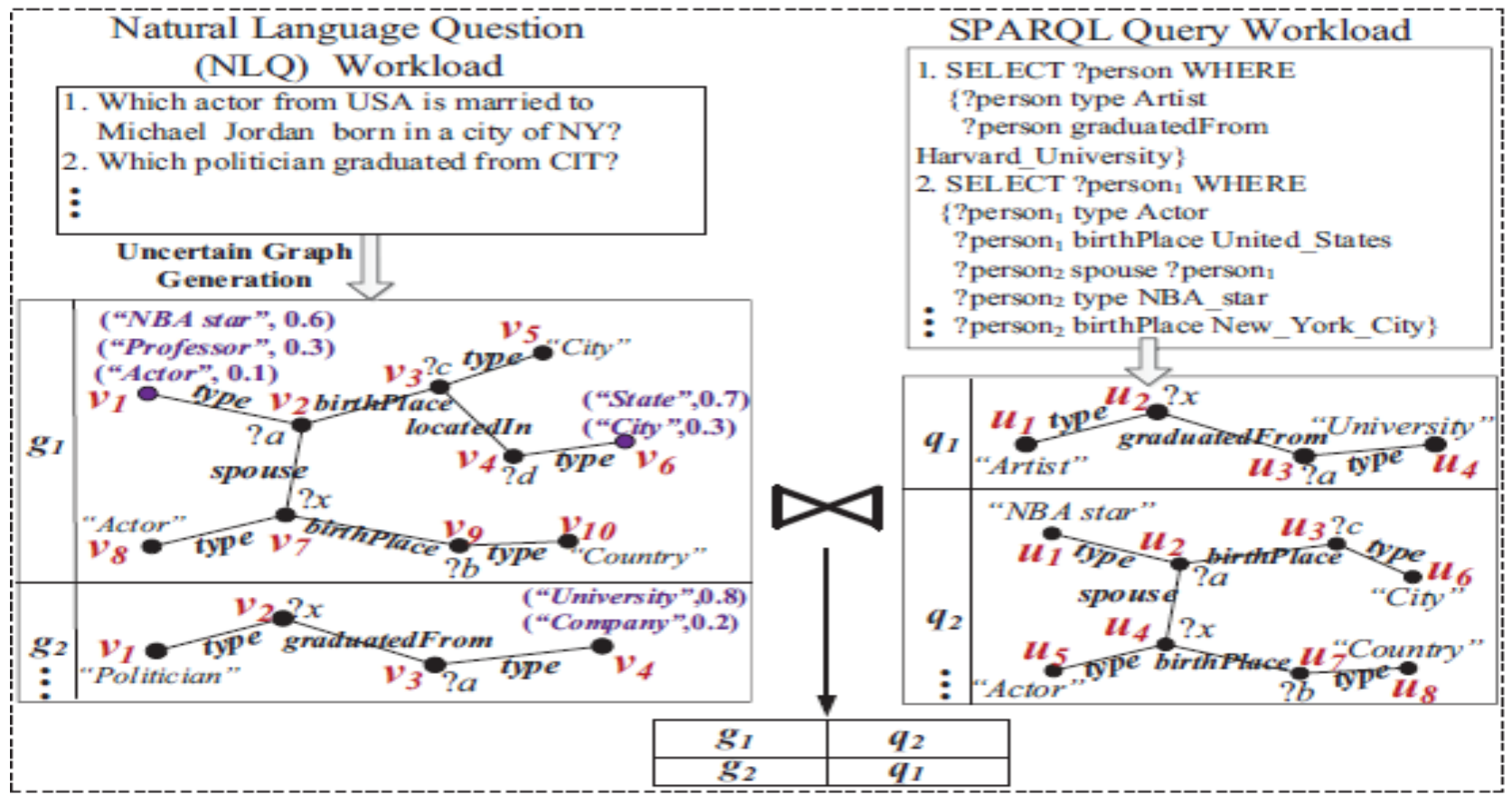
تولید Template

Which actor from USA is married to Michael Jordan born in city of NY ?

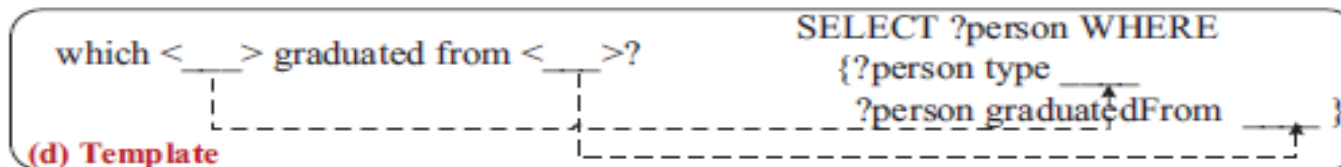
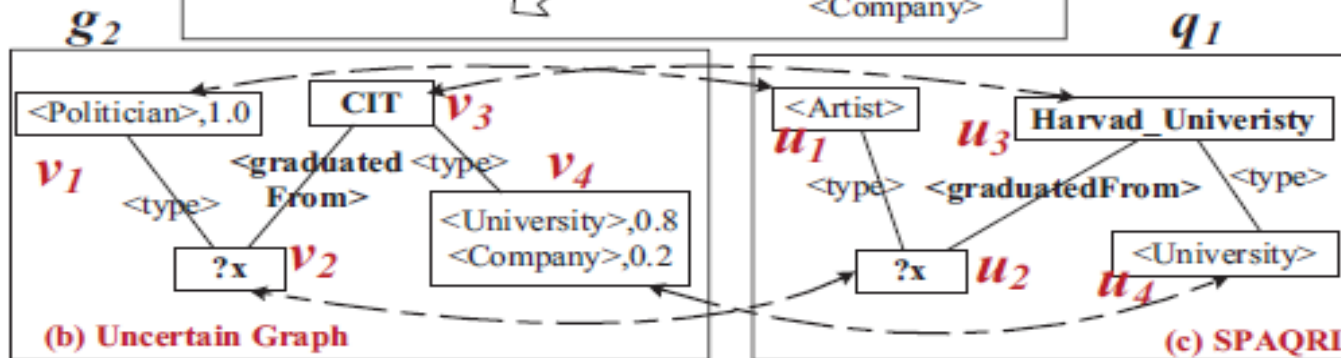
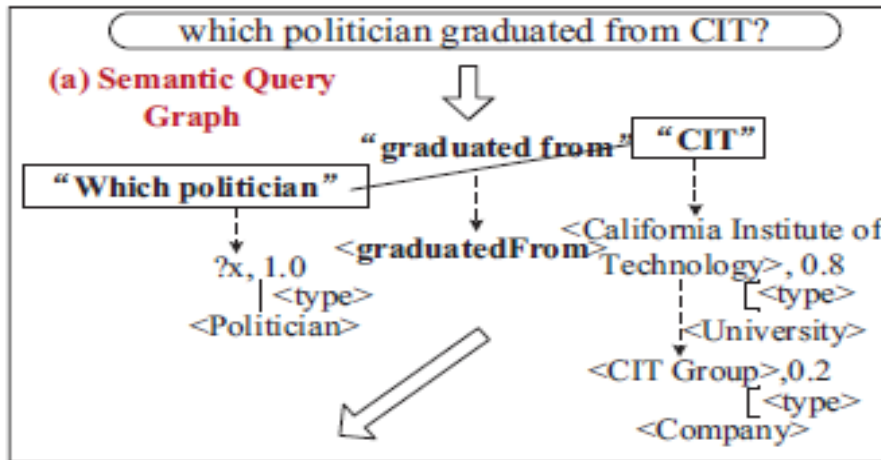


تولید Template:

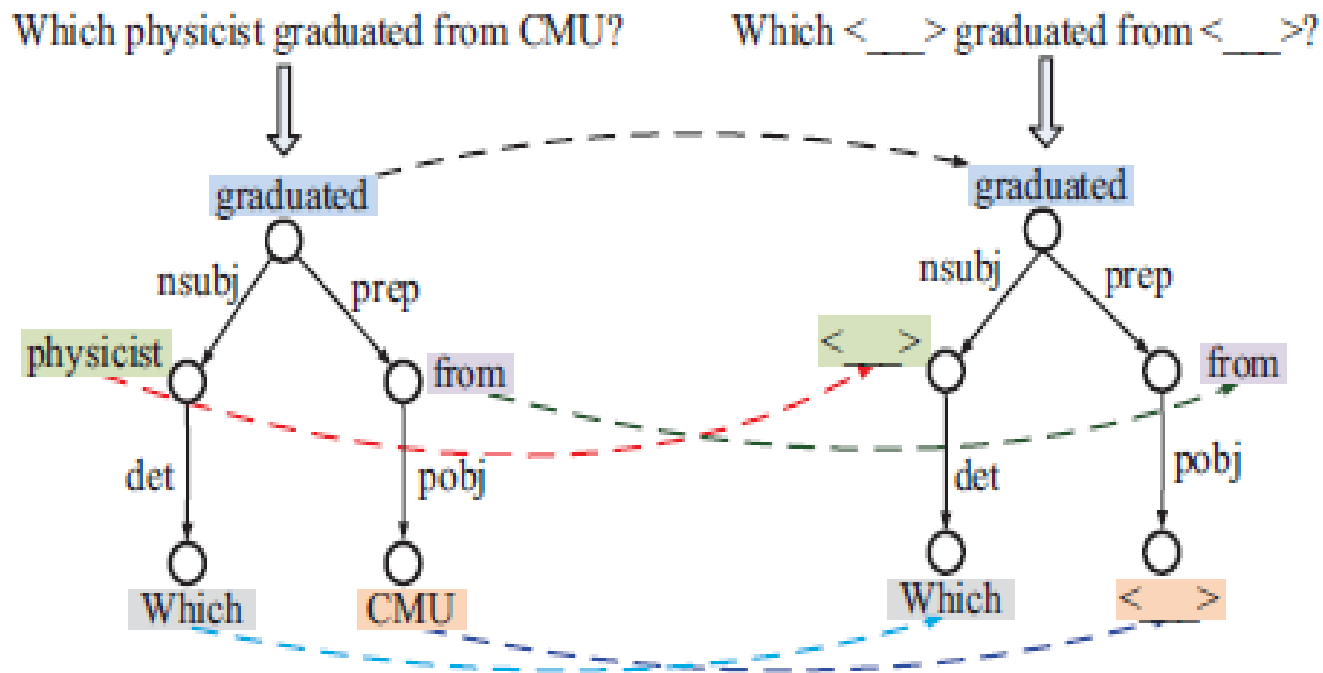
پیدا کردن شباهت گرافها



تولید Template



پیدا کردن Template برای پرسش جدید:



- Zheng, Weiguo, et al. "How to build templates for RDF question/answering: An uncertain graph similarity join approach." *Proceedings of the 2015 ACM SIGMOD International Conference on Management of Data*. ACM, 2015.
- Moussa, Abdullah M., and Rehab F. Abdel-Kader. "Qasyo: A question answering system for yago ontology." *International Journal of Database Theory and Application* 4.2 (2011): 99.
- Gerber, Daniel, and A-C. Ngonga Ngomo. "Bootstrapping the linked data web." *1st Workshop on Web Scale Knowledge Extraction@ISWC*. Vol. 2011. 2011.
- Unger, Christina, et al. "Template-based question answering over RDF data." *Proceedings of the 21st international conference on World Wide Web*. ACM, 2012.
- Unger, Christina, André Freitas, and Philipp Cimiano. "An introduction to question answering over linked data." *Reasoning Web International Summer School*. Springer International Publishing, 2014.

Question ...

