Question Answering over Knowledge Graphs

Prof. Dr. Jens Lehmann
Smart Data Analytics Group
University of Bonn & Fraunhofer IAIS
The 2017 Voice Report Predicts More Than 24 Million Amazon Echo And Google Home Devices Will Be Sold This Year

Chatbots are revolutionizing customer support

WITH QUARTZ’S APP, YOU DON’T READ THE NEWS. YOU CHAT WITH IT

Chatbots Raised Over $170M+ in 8 Months

Baidu is bringing AI chatbots to healthcare

How Analytics, Big Data and AI Are Changing Call Centers Forever
Definition of QA

Definition adopted from Hirschman and Gaizauskas [2]:

1. Ask question in natural language

Example (Natural Language)

Which books are written by Dan Brown?
Definition of QA

Definition adopted from Hirschman and Gaizauskas [2]:

1. Ask question in natural language
2. People use own terminology

Example

Which books are *written by* Dan Brown?
Which books have Dan Brown as one of their *authors*?
What are *notable works* of Dan Brown?
Definition of QA

Definition adopted from Hirschman and Gaizauskas [2]:

1. Ask question in natural language
2. People use own terminology
3. Receive a concise answer

Example (Formal Language)


The Official Website of Dan Brown
www.danbrown.com/
Question Answering Timeline

1961: BASEBALL
1971: NASA LUNAR

1993: MIT’s START

2009: Wolfram Alpha
2010: SIRI

2011: IBM Watson won Jeopardy

2014: Microsoft Cortana

2012: Google uses Knowledge Graph for QA

2015-16: Google Assistant, Amazon Alexa
BASEBALL [1]:
- First QA system in 1961
- Answered questions about the US baseball league over a period of one year
LUNAR [3] 1971:

- Developed for NASA to answer questions on lunar geology
- First evaluated QA system with 78% correctly answered questions from lunar science convention visitors in 1971
- Compiles English into a “meaning representation language”
What is the fifty-second smallest country by GDP per capita?

Result:
Djibouti
($1814 per year per person)
Question Answering Timeline

- 1961: BASEBALL
- 1971: NASA LUNAR
- 1993: MIT's START
- 2011: IBM Watson won Jeopardy
- 2014: Microsoft Cortana
- 2009: Wolfram Alpha
- 2010: SIRI
- 2012: Google uses Knowledge Graph for QA
- 2015-16: Google Assistant, Amazon Alexa
Question Answering Timeline

1961: BASEBALL
1971: NASA LUNAR
1993: MIT’s START
2009: Wolfram Alpha
2010: SIRI
2011: IBM Watson won Jeopardy
2012: Google uses Knowledge Graph for QA
2014: Microsoft Cortana
2015-16: Google Assistant, Amazon Alexa

“Alexa, turn on Welcome Home”
“Alexa, turn off my Bedroom Sonos”
“Alexa, turn on my Chili Time”
“Alexa, turn on the TV”

Hi, how can I help?
how is the weather in Bonn
Don't forget your sunglasses

31°C in Bonn
Sunny - High: 33°C Low: 16°C
Precip: 10%
1 Knowledge Graphs
Knowledge Graph Example

Labeled, directed multigraph

Spock
 Leonard Nimoy
 played
 Star Trek
 starredIn
 Science Fiction
 genre
 Obi-Wan Kenobi
 played
 Alec Guinnes

Star Trek

Science Fiction

Star Wars

Obi-Wan Kenobi

Popular Knowledge Graphs

DBpedia 2007

Freebase 2007

yago 2008

NELL 2010

Knowledge Graph 2012

Microsoft Satory 2012

facebook 2013

LinkedIn 2016
Knowledge Graph in Google Search
RDF and the Linked Data Principles

RDF Triple / Fact:

Subject ➔ Predicate ➔ Object

Linked Data principles (simplified version):
1. Use RDF and URLs as identifiers
2. Include links to other datasets
RDF and the Linked Data Principles

RDF Triple / Fact:

Subject ➔ Predicate ➔ Object

Example:

\[
\begin{align*}
\text{dbr:Germany} & \quad \text{dbo:hasCapital} & \quad \text{dbr:Berlin} \\
\text{Subject} & \quad \text{Predicate} & \quad \text{Object}
\end{align*}
\]

dbo : http://dbpedia.org/ontology/
dbr : http://dbpedia.org/resource/
RDF and the Linked Data Principles

RDF Triple / Fact:

Example:

Subject: dbr:Germany
Predicate: dbo:hasCapital
Object: dbr:Berlin

dbo: http://dbpedia.org/ontology/
dbr: http://dbpedia.org/resource/

Linked Data principles (simplified version):

1. Use RDF and URLs as identifiers
2. Include links to other datasets
Linked Data / RDF Graphs

DHL

full name

DHL International GmbH

industry

Logistics

headquarters

located in

Post Tower

height

162.5 m

Bonn

Logistik

物流

RDF = Resource Description Framework

(URLs not shown; a set of triples is strictly not a graph but often labeled as such)
Linked Open Data Cloud

10000+ open datasets and 100+ billion facts
2 QA over KG Applications
Personal Assistants

- Personal assistants on smartphones with voice interface
- Users expect devices to have **encyclopedic knowledge**
- Also becoming viable for other medium and large enterprises

"The world’s knowledge in your pocket“
Entertainment & Smart Home

- Will get increasingly complex → people will ask devices more complex questions
- Example: QA on knowledge graphs plus images/videos
- Networked devices (Internet of Things)
When other devices are inconvenient . . .

- Crisis situations
- In-car QA systems (spatial and temporal QA)
- Childcare ;-)

Is this substance toxic?

When will the wave arrive here?

Integrated Overview

RDF / AMQP
Search

- QA is becoming part of main stream search engines
- Google:
  - Knowledge Graph in 2012
  - Question Intent Understanding in 2015
  - Can understand superlatives, ordered items, time e.g. "Who was the U.S. President when the Angels won the World Series?"
- Also relevant for enterprise search in medium and large enterprises
June 23, 1912
3 Challenges for QA on KGs
Goal of QA over RDF Datasets

NLQ

What does the **Rhine** cross through?

SPARQL
TRIPLE
PATTERN

KB

Rhine

crosses through

?x

Rhine

crosses through

Mainz

Bonn

Köln
Who developed Minecraft?

SELECT DISTINCT ?uri
WHERE {
}
How **tall** is Michael Jordan?

```sparql
SELECT DISTINCT ?num
WHERE {
}
```
Give me all taikonauts.

SELECT DISTINCT ?uri
WHERE {
}
Lexical Knowledge

- Who was the last princess of Joseon?
  → !BOUND :successor
- Which of the Beatles is still alive?
  → !BOUND :deathDate
How many **banks** are there in London?

```
SELECT DISTINCT count(?bank)
WHERE {
  ?bank a dbo:Bank.
}
```

```
SELECT DISTINCT count(?bank)
WHERE {
  ?bank a lgdo:Riverbank.
}
```
Give me all communist countries.

SELECT DISTINCT ?uri
WHERE {
?uri dbpedia-owl:governmentType dbpedia:Communism .
}"
Who are the parents of the wife of Juan Carlos I?

```sql
SELECT DISTINCT ?uri
WHERE {
  dbpedia:Juan_Carlos_I_of_Spain dbpedia-owl:spouse ?x .
  ?x dbpedia-owl:parent ?uri .
}
```
How many companies were founded in the same year as Google?

SELECT COUNT(DISTINCT ?c)
WHERE {
  ?c rdf:type dbo:Company .
  ?c dbo:foundingYear ?year .
  res:Google dbo:foundingYear ?year .
}
Scalability & Real World Problems

- Distributed data
  → aggregating information from different datasets
  (Question Answering over linked data)
- Multilingual data
- Conflicting data / versioning
- Missing and incomplete data
- Spatial data queries
- Temporal questions
Non-factoid Questions

- What is the difference between impressionism and expressionism?
- How do histone methyltransferases cause histone modification?
- Is honesty the best policy?
- Are donuts tasty?
What is the best fruit to eat?
→ personal preference

Who will be the next president of USA?
→ unknown

If I ate myself, would I be twice the size as now or would I disappear completely?

Will QA systems be expected to detect unanswerable queries (and provide entertaining answers) in the future?
4 Question Answering Approaches
Semantic Parsing Approaches

- Explicitly represents the grammatical and semantical relationships between the words of the sentences
- Step 1: Forms a semiformal intermediate structure devoid of lexical and semantic variations
- Step 2: Informal structure is then converted into a logical form (see e.g. AskNow, XSER systems)

http://dbpedia.org/resource/Mount_Everest

http://dbpedia.org/ontology/elevation
Addressing the Lexical Gap

- Add multiple labels to the knowledge graph
- String normalization (Stemming/Lemmatizing)
- String similarity functions
- Automatic query expansion e.g. via WordNet → improves recall, reduces precision
- More difficult: properties
  - No continuous substrings: In which place was Keira Knightley born? (relation birthPlace)
  - Expressed as nouns or verbs: Who wrote Braveheart? , Who was the writer of Braveheart? (relation writer)
  - ”Solution“: pattern libraries (BOA, Patty, Paralex)
- Word embeddings
Addressing Ambiguity

Who founded Apple?

Can I pickle apples?
Addressing Ambiguity

- Same phrase has different meanings (classical example: Jaguar)
- High lexical gaps reduce recall, ambiguity reduces precision
- Context can be used for disambiguation (neighbor words, POS tags, parse tree structure)
- Exploiting the structure of the knowledge graph i.e. maximizing semantic relatedness measures of entity candidates in the question
- User interaction e.g. inductive logic programming on given answers
Addressing Complex Queries

- **IBM Watson:**
  - Decomposes complex question into a set of simple question
  - Hypothesis generation and evidence scoring using multiple analysis algorithm.
  - Final confidence merging and ranking to generate answer

- **AskNow:**
  - Represent complex natural question as an intermediate canonical structure
  - Predicate and entity linking
  - Conversion of canonical queries into SPARQL using template fitting algorithm
## Publications for Challenges (up to early 2015)

<table>
<thead>
<tr>
<th>Year</th>
<th>Total</th>
<th>Lexical Gap</th>
<th>Ambiguity</th>
<th>Complex Operators</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>absolute</td>
<td>percentage</td>
<td>percentage</td>
</tr>
<tr>
<td>2010</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2011</td>
<td>16</td>
<td>11</td>
<td>12</td>
<td>3</td>
</tr>
<tr>
<td>2012</td>
<td>14</td>
<td>6</td>
<td>7</td>
<td>2</td>
</tr>
<tr>
<td>2013</td>
<td>20</td>
<td>18</td>
<td>12</td>
<td>5</td>
</tr>
<tr>
<td>2014</td>
<td>13</td>
<td>7</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>2015</td>
<td>6</td>
<td>5</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>all</td>
<td>70</td>
<td>46</td>
<td>42</td>
<td>12</td>
</tr>
<tr>
<td>2011</td>
<td></td>
<td>68.8</td>
<td>75.0</td>
<td>18.8</td>
</tr>
<tr>
<td>2012</td>
<td></td>
<td>42.9</td>
<td>50.0</td>
<td>14.3</td>
</tr>
<tr>
<td>2013</td>
<td></td>
<td>85.0</td>
<td>60.0</td>
<td>25.0</td>
</tr>
<tr>
<td>2014</td>
<td></td>
<td>53.8</td>
<td>61.5</td>
<td>15.4</td>
</tr>
<tr>
<td>all</td>
<td></td>
<td>65.7</td>
<td>60.0</td>
<td>17.1</td>
</tr>
</tbody>
</table>
Motivation of End-to-End Systems

Problem of Semantic Parsing Approaches:

- Complex pipelines
- A lot of manual configuration
- Error propagation along the pipeline $\rightarrow$ accuracy (a value between 0 and 1) after each step (roughly) multiplies
- $\rightarrow$ Interest in end-to-end approaches
End-to-End Question Answering

- No traditional Natural Language Processing (NLP) – the QA System essentially knows nothing about language
- No intermediate logical structures – answers are generated directly from the questions („end-to-end“)
- Current approaches answer simple questions, i.e. involving one relation and one entity
  - Where is the southern cross travel insurance Australia located?
  - Which characters played in Star Trek?

```
Spock
  - played
    - characterIn
      - genre
        - starredIn
          - Leonard Nimoy

Science Fiction
  - genre
    - starredIn
      - Star Trek

Obi-Wan Kenobi
  - characterIn
    - played
      - Star Wars

Leonard Nimoy
  - starredIn
    - Star Trek

Alec Guinness
  - starredIn
    - Star Wars
```
Lukovnikov et. al (WWW’2017):

- Questions are mapped to vector representation (encoding)
- Entities (nodes) and predicates (edges) in the knowledge graph are encoded as well
- Representation of question and answer is compared using similarity measures
End-to-End Question Answering

- Each word represented via word embeddings (capturing the semantics of the word)
- Each character of the word is also passed through encoder (capturing the rare/unknown words)
## Evaluation Facebook AI bAbI QA Dataset

<table>
<thead>
<tr>
<th>Approach</th>
<th>Year</th>
<th>Accuracy %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fraunhofer IAIS Dev</td>
<td>2017</td>
<td>74.8</td>
</tr>
<tr>
<td>Fraunhofer IAIS WWW</td>
<td>2017</td>
<td>71.2</td>
</tr>
<tr>
<td>Golub and He [A]</td>
<td>2016</td>
<td>70.9</td>
</tr>
<tr>
<td>Yin et al. [B]</td>
<td>2016</td>
<td>68.3</td>
</tr>
<tr>
<td>Bordes et al. [C]</td>
<td>2015</td>
<td>62.7</td>
</tr>
<tr>
<td>Dai et al. [D]</td>
<td>2016</td>
<td>62.6</td>
</tr>
</tbody>
</table>

- Comparisons with other end-to-end systems (pre-processing in other system achieves up to approx. 75% accuracy)


## Semantic Parsing or End-to-End?

<table>
<thead>
<tr>
<th>Type</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Semantic Parsing</td>
<td>Less/no training data needed&lt;br&gt;Can handle complex questions&lt;br&gt;Can handle complex questions&lt;br&gt;Fast once the system is trained</td>
<td>More manual adaptation work</td>
</tr>
<tr>
<td>End-to-End</td>
<td>Less manual adaption work&lt;br&gt;More training data&lt;br&gt;No robust theory for complex questions</td>
<td>More training data</td>
</tr>
</tbody>
</table>
5 Future Directions
Future Challenges: Context

- Context Information
- Personal Assistant
- Social Media Feeds

Diagram illustrating connections between a user and various technologies and social media platforms.
Future Challenges: Hybrid QA

Name some institutions located in Bonn?

Structured:
- U.N.
- Bonn
- DFG

Unstructured:
The headquarters of Deutsche Post DHL and Deutsche Telekom, both DAX-listed corporations, are in Bonn. The city is home to the University of Bonn and a total of 19 United Nations institutions.
Future Challenges: QA from Several Sources

- **Location Data**
  - How far is the school from now?

- **Unstructured Data**
  - What's the latest news from DW?
  - Who is the president of United States?

- **Structured Data**
  - Will it rain today in Bonn?

- **Weather Data**
  - How much is five times ninety five?
Future Challenges: APIs and QA

<table>
<thead>
<tr>
<th>System</th>
<th>User</th>
<th>System</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hey Ian, you're about to run out of gas!</td>
<td>Oh! Okay. Can I make it to the school and back?</td>
<td>I estimate that we can go about 10 km with the fuel. The school is 8.4 km from here, and there are no fuel stations on the way. So you would not be able to make it to school and back.</td>
</tr>
<tr>
<td>Sure, calling Miranda ...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- Vehicle Position
- Map of City
- Destination Position
- Vehicle Stats
- Vehicle Performance Estimate
- Context Aware Planning
- Spatial Reasoning
- Personalized Information
- Integration to User’s Phone
Future Challenges: Out-of-scope Questions

- QA system should know when question cannot be answered (very important in IBM Watson)
- Examples of OOS:
  - Knowledge base cannot answer it: Is my mother at home now?
  - Question refers to the future with no forecasts available: When will the world record for the fastest supercomputer be broken?
  - Unknown question type (e.g. asking non-factual question to factoid system): How honest is the government with me?
- Systems need to be benchmarked on OOS questions, otherwise they are encouraged to guess
Future Challenges: QA Architectures

Qanary - an Open Knowledge-based Architecture for Question Answering
Further Challenges

- Complex queries in end-to-end systems
- Better relation extraction
- Multilinguality
- Automatic template learning for semantic parsing systems
- Spatial and temporal questions
- User interaction requiring minimal user effort
- Integration of QA and dialogue systems
Want to build and test your own system?

HOBBIT Project

Data Collection
Industry data

Benchmark Creation
Benchmark 1
Benchmark 2
Benchmark n

Challenges
Solution 1
Solution 2
Solution k

Reports
Want to build and test your own system?

QALD (“Question Answering over Linked Data“)
- 7 editions so far
- Co-located with major conferences
- Datasets:
  - Multilingual question answering over DBpedia
  - Hybrid question answering (text and structure)
  - Large-scale QA → runtime performance
  - Wikidata question answering
- Coming up: 5000 question dataset including a number of interesting SPARQL query patterns!
Take Home Messages

- Knowledge graphs gaining popularity
- Wide range of applications for question answering over knowledge graphs
- Two main approaches: semantic parsing and end-to-end systems
- Active areas of research with many challenges – still many unsolved problems
- Benchmarks are available – you can explore and test your own ideas
Thanks for your attention!


I'd like to buy a computer - to use the internet. May I ... ?

Sure! You know how to use it?

I do!


Is this going to take long?

Well, I guess about 10 to 20 years... "Commander"