Advanced Methods of Information Retrieval
- Question Answering Applications -

Dr. Elena Demidova
SS 2018
Open questions from the last lecture
Questions from last lecture: MAP vs. NDCG

NDCG gives higher scores to result lists that have more relevant results at the top (through gain & discount computation).

\[
NDCG(Q, k) = \frac{1}{|Q|} \sum_{j=1}^{|Q|} \sum_{m=1}^{k} \frac{2^{R(j,m)} - 1}{\log_2(1 + m)}
\]

Mean Average Precision (MAP) averages precision at all recall levels \(\{d_1, \ldots, d_{mj}\}\). \(R_{jk}\) is a set of ranked retrieved results from the top up to \(d_k\).

\[
MAP(Q) = \frac{1}{|Q|} \sum_{j=1}^{|Q|} \frac{1}{m_j} \sum_{k=1}^{m_j} \text{precision} (R_{jk})
\]
Recap:

Knowledge Graphs
SPARQL
Information Extraction
A Knowledge Graph

A Knowledge Graph is an entity-centric knowledge repository that:

• describes real-world entities and their interrelations
• defines possible classes and relations of entities
• allows for interrelating arbitrary entities
• covers various topical domains

(Paulheim, 2016)
Which maize dishes are popular in the United States?

SELECT ?dish ?name WHERE {
}
Selected Information Extraction methods

• Tokenisation
• Disambiguation
• Part-of-speech tagging
• Semantic role labels
• Co-reference resolution
• Relation extraction
• Named entity recognition
• (and many other)
NE Extraction Pipeline

- Sentence splitting: *Dubrovnik is located in the region of Dalmatia.*
- Tokenization: 
  - **Dubrovnik** is located in the region of Dalmatia.
- Lemmatisation or stemming: 
  - E.g. “located” -> “locate”
- POS tagging: Nouns, adjectives and verbs
- Chunking, parsing: 
  - “Dubrovnik is located”, “region of Dalmatia”
- Co-reference resolution: 
  - “Dubrovnik” and “This city”.

(Examples: Stanford NER)
Entity Linking

Entity Linking (EL): detecting entities and linking them to the entries of a Knowledge Graph

**Dubrovnik** is located in the region of Dalmatia.

-> [http://dbpedia.org/resource/Dubrovnik](http://dbpedia.org/resource/Dubrovnik)

-> [https://www.wikidata.org/wiki/Q1722](https://www.wikidata.org/wiki/Q1722)
Relation Extraction

- **Relations** between two or more entities, which relate to one another in real life.
  - A relation is defined in the form of a tuple $t = (e_1, e_2, ..., e_n)$, where $e_i$ are entities in a predefined relation $R$ within document $D$.

- **Relation extraction**: 
  - is a task of detecting relations between entities and assigning relation types to them.

- **Binary relations**: a relation between two entities
  - located-in(Dubrovnik, Croatia)
  - married-to(Angelina Jolie, Brad Pitt)
Aims of the session “Question Answering Applications”

- Lecture:
  - Provide an overview of:
    - Question Answering approaches
    - Semantic Question Answering
  - Analyse:
    - Challenges in Semantic Question Answering
    - Some ways to address them
  - Discuss:
    - Differences between Question Answering and keyword search

- Hands-on:
  - Get practical experience with:
    - Question Answering interfaces
    - Semantic Question Answering pipelines
Definition of Question Answering

A Question Answering (QA) system is a system where users can:

- Ask a question in a natural language (NL)
- Use their own terminology
- Receive a concise answer

(L. Hirschman and R. Gaizauskas. 2001)
Definition of Question Answering

A Question Answering (QA) system is a system where users can:
• Ask a question in a natural language (NL)
• Use their own terminology
• Receive a concise answer

(L. Hirschman and R. Gaizauskas. 2001)

What is the difference between QA and keyword search?
Question Answering example

- Ask a question in a natural language (NL)
  - “Which maize dishes are popular in USA?”
- Use their own terminology
  - “What are the famous maize dishes in America?”
- Receive a concise answer
Applications of Question Answering

- Personal assistance
  - Amazon Alexa, Google Home, Siri, Cortana …
- Search and customer support in enterprises
- Integral part of Web search engines

What is the weather in Hannover?

How much is 1917/5?
QA evolution: BASEBALL (1961)

BASEBALL: First QA system (1961)
Questions about US baseball league

“Baseball is a computer program that answers questions phrased in ordinary English about stored data. The program reads the question from punched cards. […] The program's present context is baseball games; it answers such questions as:

"Where did each team play on July 7?"

(Green et al., 1961)
QA evolution: WolframAlpha (2009)

What is the total length of all roads in Spain?

Input interpretation:

| Spain  | total road length |

Result:

666,300 km (kilometers) (2003 estimate)
QA evolution: IBM Watson (2011)

• 2011: IBM Watson won Jeopardy!
• Jeopardy! is an American television game show (a quiz)
• Participants are presented with general knowledge clues in the form of answers and phrase their responses as questions.
QA evolution: Google Knowledge Graph (2012)

Who is the first cosmonaut?

Yuri Gagarin

April 12 was already a huge day in space history twenty years before the launch of the first shuttle mission. On that day in 1961, Russian cosmonaut Yuri Gagarin (left, on the way to the launch pad) became the first human in space, making a 108-minute orbital flight in his Vostok 1 spacecraft. Apr 13, 2011

NASA - Yuri Gagarin: First Man in Space
https://www.nasa.gov/mission_pages/shuttle/sts1/gagarin_anniversary.html
QA evolution: Google Assistant, Amazon Alexa (2015/16)

News:
  “Ok Google, what was the result of the last FIFA World Cup game?”

Routing with background knowledge:
  “Ok Google, I want to go home.”

Reminder:
  “Ok Google, remind me on the exam tomorrow.”

Do homework:
  “Ok Google, what is 485 times 24?”
FAQ evolution: „AI develops its own language“ (2017)

Facebook engineers panic, pull plug on AI after bots develop their own language


Facebook-AI entwickelt eigene geheime Sprache

http://www.pcwelt.de/a/facebook-ai-entwickelt-eigene-geheime-sprache,3447605

Facebook stellt sich verselbständigendes KI-System ab

http://www.silicon.de/41654841/facebook-stellt-sich-verselbständigendes-ki-system-ab/?inf_by=59710ccd681db834188b4608

Künstliche Intelligenz - Facebook stoppt KI, nachdem sie neue Sprache erfindet


No, Facebook Did Not Panic and Shut Down an AI Program That Was Getting Dangerously Smart


Does it really?
Question Answering approaches

- **Semantic question answering [R1]**
  - convert Natural Language Query (NLQ) to a logical form
    - (e.g. a SPARQL query or to a $\lambda$-calculus expression)

- Information retrieval based techniques
  - convert NLQs to an answer, usually without any explicit intermediary form

- **End-to-end systems:**
  - use deep learning methods
Question Answering approaches

• **Semantic question answering [R1]**
  - convert Natural Language Query (NLQ) to a logical form
  - (e.g. a SPARQL query or to a λ-calculus expression)

• Information retrieval based techniques
  - convert NLQs to an answer, usually without any explicit intermediary form

• End-to-end systems:
  - use deep learning methods

• **What is the difference between semantic Question Answering and keyword search in knowledge graphs?**
Example semantic Question Answering pipeline

- Question
- Entity Linking
- Relation Linking
- Query Building
Example semantic Question Answering pipeline

Question

"Was Winston Churchill the prime minister of Selwyn Lloyd?“
Example semantic Question Answering pipeline

Question
"Was Winston Churchill the prime minister of Selwyn Lloyd?"

Entity linking
http://dbpedia.org/resource/Selwyn_Lloyd
http://dbpedia.org/resource/Winston_Churchill
Example semantic Question Answering pipeline

**Question**

"Was Winston Churchill the prime minister of Selwyn Lloyd?"

**Entity linking**

http://dbpedia.org/resource/Selwyn_Lloyd

http://dbpedia.org/resource/Winston_Churchill

**Relation linking**

http://dbpedia.org/ontology/primeMinister
Example semantic Question Answering pipeline

Question
"Was Winston Churchill the prime minister of Selwyn Lloyd?“

Entity linking
http://dbpedia.org/resource/Selwyn_Lloyd
http://dbpedia.org/resource/Winston_Churchill

Relation linking
http://dbpedia.org/ontology/primeMinister

Query Building
Challenges in Semantic Question Answering

- Lexical gap
- Ambiguity
- Variations
- Complexity
- Scalability
- Non-factual questions
- Non-answerable questions
Challenges: lexical gap

Which corn dishes are popular in America?

```
SELECT ?dish ?name WHERE {
}
```
Approaches: lexical gap

- Entity Linking
  - Add multiple labels to the entities in the knowledge graph
  - Normalize strings
  - Use string similarity
  - Perform query expansion using lexicons
- Relation Linking
  - Use pattern libraries
  - Use context of relations in the question and in the knowledge graph

- In general more difficult than linking entities
  - *Where was Jack London born?*
  - *dbo:birthPlace*
Challenges: lexical knowledge

Give me all cosmonauts.

SELECT ?uri WHERE {
}

cosmonaut

/ˈkɒzmənət/ 🎧

noun

a Russian astronaut.

synonyms: astronaut, spaceman, spacewoman, space traveller, space pilot, space flyer, space cadet; informal jock
Challenges: variations and errors in schema usage

Give me all cosmonauts.

SELECT ?uri WHERE {
}

Note: this query returns useful results, but fails to retrieve Juri Gagarin (and many others).

http://dbpedia.org/page/Yuri_Gagarin

<table>
<thead>
<tr>
<th>dbo:birthPlace</th>
<th>dbo:nationality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>dbr:Russian_SFSR</td>
</tr>
<tr>
<td></td>
<td>dbr:Soviet_Union</td>
</tr>
<tr>
<td></td>
<td>dbr:Klushino</td>
</tr>
<tr>
<td></td>
<td>dbr:Soviet_space_program</td>
</tr>
</tbody>
</table>
Challenges: ambiguity

Give me a London book

place of publication, narrative location, headquarters location, location of final assembly, published in / London (capital of England and the United Kin...) / instance of, depicts, ...

- Nineteen Eighty-Four
- Animal Farm
- Guinness World Records
- Around the World in Eighty Days
- Oliver Twist
- Victoria History of the County of Warwick: Volume 7, the City of Birmingham
Approaches: ambiguity

• Context in the question
  • terms
  • tags
  • parse tree structure
• Neighbourhood in the knowledge graph
  • semantic relatedness
  • user interaction
Challenges: complexity

Q: List the movies whose music composers have won the BAFTA Award for Best Film Music?

SELECT DISTINCT ?uri WHERE {
    ?x dbp:award dbr:BAFTA_Award_for_Best_Film_Music .
    ?uri dbp:musicComposer ?x
}
Approaches: complexity

- Decomposition in simpler questions
- Use of intermediate canonical structures
Entity and relation linking approaches: sequential

- Entities linked first, followed by relation linking

- "Was Winston Churchill the prime minister of Selwyn Lloyd?"
  - [http://dbpedia.org/resource/Selwyn_Lloyd](http://dbpedia.org/resource/Selwyn_Lloyd)

- Uses schema knowledge
- Reduces search space for relation linking
- Correct entity matching is crucial / error propagation
Entity and relation linking approaches: independent

- Entities and relations are linked independently

- "Was Winston Churchill the prime minister of Selwyn Lloyd?"
  
  http://dbpedia.org/resource/Selwyn_Lloyd
  http://dbpedia.org/resource/Winston_Churchill
  http://dbpedia.org/ontology/primeMinister

- Schema knowledge is not used
- Relations and entities are not used for mutual verification
- Correction of errors in entity linking is possible at a later step
Entity and relation linking approaches: joint

• Joint optimisation of entity and relation linking

• "Was Winston Churchill the prime minister of Selwyn Lloyd?"
  • Selwyn Lloyd -> http://dbpedia.org/resource/Selwyn_Lloyd

- dbo:primeMinister
  - dbr:Alec_Douglas-Home
  - dbr:Harold_Macmillan
  - dbr:Anthony_Eden
  - dbr:Sir_Winston_Churchill

• Potentially better disambiguation [R2]
• Higher time complexity
Factual / non-factual / unanswerable questions

• Factual questions
  • "What is the population of Paris?"
    • 2.244 million (2010) according to Google

• Non-factual
  • “Which city is the best place to live?”
    • (personal preferences)

• Unanswerable
  • “Who will win the FIFA World Cup 2018?”
    • (not known yet)

• Detecting unanswerable questions is a challenge
Evaluation campaigns

- Aim: Quantitatively compare QA systems using benchmarks
- Focus on performance of the system as a whole
- Question Answering over Linked Data (QALD)
  - yearly evaluation campaign since 2011
  - open domain Semantic QA on lexicographic facts of DBpedia
  - multilinguality, hybrid (textual and Linked Data)
  - Semantic QA on statistical data
- BioASQ
  - domain-specific challenge of biomedical data (till 2015)
- TREC LiveQA
  - unanswered Yahoo Answers questions (from 2015)
- Scalable Question Answering Open Challenge (SQA) 2017-2018
  - [https://project-hobbit.eu/open-challenges/sqa-open-challenge/](https://project-hobbit.eu/open-challenges/sqa-open-challenge/)
Benchmarks

Benchmark forms:
- Natural language question – answer pairs
- Natural language question – logical form (e.g. SPARQL query) pairs

Benchmark characteristics:
- Size: number of questions
- Variety: lexical and syntactic variations
- Complexity: number of triples in the SPARQL query, number of entities, …
- Target KG
Question Answering Evaluation: Benchmarks

[B1] LC-QuAD: Large-Scale Complex Question Answering Dataset: 5000 questions and intended SPARQL queries. Target KG: DBpedia. Syntactic and structural variations of questions. Complex questions, i.e. queries include >1 triple pattern.

Collect questions: Questions begin with a wh-word and contain exactly one entity collected using Google suggest API.
Annotations: manually selected answers, as given by Freebase (using AMT). Target KG: Freebase.

[B3] Simple Questions

Simple and complex questions
Example question from the LC-Quad benchmark

https://figshare.com/articles/LC-QuAD_QALDformat/5818452

"Which comic characters are painted by Bill Finger?"

"sparql": "SELECT DISTINCT ?uri WHERE {?uri
   <http://dbpedia.org/ontology/creator>
   <http://dbpedia.org/resource/Bill_Finger> . ?uri
   <http://www.w3.org/1999/02/22-rdf-syntax-ns#type>
   <http://dbpedia.org/ontology/ComicsCharacter>},

http://dbpedia.org/resource/Batman

..
References and further reading


Benchmark references


Summary “Question Answering Applications”

- Lecture:
  - Provided an overview of:
    - Question Answering approaches
    - Semantic Question Answering
  - Analysed:
    - Challenges in Semantic Question Answering
    - Selected ways to address them
  - Discussed
    - Differences between Question Answering and keyword search
- Hands-on:
  - Get practical experience with:
    - Question Answering interfaces
    - Semantic Question Answering pipelines
Thank you!

Questions, Comments?

Dr. Elena Demidova

L3S Research Center
Leibniz University of Hannover

email: demidova@L3S.de
www: https://demidova.wordpress.com