IBM Watson

Berfu Büyükozu - Mert Çıkla
CmpE 561

Boğaziçi University

May 9, 2017
1. Jeopardy!

2. IBM and QA

3. DeepQA
   - Question Analysis
   - Resource Acquisition
   - Candidate Answer Search & Generation
   - Candidate Ranking with Evidence
   - Structured Data

4. Watson as a Jeopardy player
   - Making Watson Fast
   - In-Game Strategy Optimization

5. Conclusion

6. References
Considered a hard game for humans
Even harder for any computer system at the time
Turkish adaptation Riziko!

Category: Types of Financial Aid
Involves money that does not have to be repaid, and generally based on merit.
*Scholarship*

Category: Number Word Play
This number, one of the first 20, uses only one vowel (4 times!)
*Seventeen*
<table>
<thead>
<tr>
<th></th>
<th>% attempted</th>
<th>% correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>PIQUANT</td>
<td>70</td>
<td>16</td>
</tr>
<tr>
<td>Winning Players</td>
<td>40-50</td>
<td>85-95</td>
</tr>
<tr>
<td>Watson’s Goal</td>
<td>70</td>
<td>85</td>
</tr>
</tbody>
</table>
DeepQA Pipeline
Question Analysis

- Watson never assumes any component perfectly understands the question
- Keeps collecting evidence from hundreds of algorithms and ranks them
- AdaptWatson is used to develop, test and optimize the algorithms
Success of any QA system with limited resources is very dependent on the quality of its data resource. Wide domain range of Jeopardy makes this even more critical. Watson tackles this in three separate ways:

1. Fill the gaps in knowledge space
2. Transform existing knowledge into new information bits
3. Increase existing knowledge by adding new lexical/syntactic variations
Prismatic Knowledge Base

- Employs *is-a* relations derived from large corpora
- Amounts to 2.4% overall increase in precision
Finding an exact answer is rare (2 % at best)
The rest is extracted from many algorithms with 3 main strategies

1. Title-in-clue Based Passage Search
2. Passage search in unstructured resources
3. Direct answer lookup & Prismatic
Evidence Scoring

Type Coercion- TyCor

- Performed after candidate generation and before ranking
- Assign scores to candidate answers based on whether it has a specified lexical type.
Who wrote The Hobbit?

“Dan Brown wrote several books and has read The Hobbit.”

“Tolkien, an English author born in the late nineteenth century, wrote The Hobbit.”
<table>
<thead>
<tr>
<th>Jeopardy! question</th>
<th>Relation detected (relations are from the DBpedia knowledge base)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MOTHERS &amp; SONS: Though only separated by one year in real life, she played mother to son Colin Farrell in &quot;Alexander.&quot;</td>
<td>Starring (she, &quot;Alexander&quot;)</td>
</tr>
<tr>
<td>THE DEVIL: &quot;The Screwtape Letters&quot; from a senior devil to an under devil are by this man better known for children's books.</td>
<td>Author (man, &quot;The Screwtape Letters&quot;)</td>
</tr>
<tr>
<td>THE LONE REPRESENTATIVE: Michael Castle from this state with 3 counties: New Castle, Kent and Sussex.</td>
<td>Residence (&quot;Michael Castle&quot;, state)</td>
</tr>
</tbody>
</table>
1. Large online databases like Wikipedia
2. Large collections of automatically extracted data from unstructured sources
3. Small amount of manually added sources to account for differences between task domain and the source
4. Small amount of manually added formal knowledge targeting most common questions/answers.
Making Watson Fast

From 2 hours to 3 seconds with 2880 processors, preloading to RAM to eliminate disk I/O and preprocessing.
In-Game Strategy Optimization

- Daily Doubles
- Final Jeopardy
- Square selection
- Attempt or pass decision
Conclusion

Watson beating humans at Jeopardy marks a milestone for QA systems.

Diverse problems encountered during its development helped further research in NLP immensely by introducing new state of the art techniques for these problems.
Thank you for listening
References