Key-Value Memory Networks for Directly Reading Documents (EMNLP16)

Alexander H. Miller¹, Adam Fisch¹, Jesse Dodge¹,², Amir-Hossein Karimi¹, Antoine Bordes¹, Jason Weston¹

¹Facebook AI Research
²Language Technologies Institute, Carnegie Mellon University
Question Answering From Text

Goal:
- Read documents and answer questions about them

Challenges for Raw Text:
- Information is unstructured
- Answers are indirectly or ambiguously expressed
- Scattered across multiple documents

Work-around: Knowledge Bases?
- Large data sources: e.g. FreeBase, Dbpedia, etc
- Have been effective tools for QA

Poor performance.

But … intrinsically limited?
Knowledge Bases vs Text

Limitations of Knowledge Base Learning:
- Availability
- Knowledge bases are often incomplete
- Expressiveness

Advantages of Raw Text Learning:
- Plenty of data
- Lossless
  More difficult to interpret
Our Work

Hypothesis:
Intuitively, systems answering directly from text should be more powerful than those using KBs. They should be able to address more complex tasks.

Problem…
In the current state of the art, they are not as accurate, even on basic tasks.

This work tries to close that gap.
Contributions

WikiMovies
- A new dataset for QA in the movies domain
- Raw text + Knowledge Base formats side by side
- A useful analysis tool for comparing methods

Key-Value Memory Networks
- A new, more effective method for memory-based QA
- Makes directly reading documents more viable
Jurassic Park directed by Steven Spielberg

Jurassic Park written by David Koepp, Michael Crichton

Jurassic Park release year 1993

Jurassic Park has_tags action, adventure, science fiction, …
Who directed Jurassic Park?  
Steven Spielberg

Can you describe Jurassic Park in a few words?  
A science fiction, action, and adventure movie

In Jurassic Park, what is the name of the island where the park is built?  
???

KBs can suffer from missing information and fixed schemas
Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. The film centers on the fictional Isla Nublar, an islet located off Costa Rica's Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.
Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. The film centers on the fictional Isla Nublar, an islet located off Costa Rica's Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.

Who directed Jurassic Park?
Can you describe Jurassic Park in a few words?
In Jurassic Park, what is the name of the island where the park is built?
### Question Answering + Open IE

#### Wikipedia Page: Jurassic Park (film)

Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. **The film centers on the fictional Isla Nublar, an islet located off Costa Rica's Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.**

Who directed Jurassic Park?

Can you describe Jurassic Park in a few words?

In Jurassic Park, what is the name of the island where the park is built?

#### Open IE Knowledge Base

- Jurassic Park *directed* by Steven Spielberg
- Jurassic Park *written* by David Koepp and Michael Crichton
- Jurassic Park *is* a 1993
- Jurassic Park *is* a American science fiction
- **Jurassic Park centers on Isla Nublar**

Steven Spielberg

A science fiction, action, and adventure movie

Isla Nublar

---

Open IE is not an easy problem
### Wikipedia Page: Jurassic Park (film)

Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. The film centers on the fictional Isla Nublar, an islet located off Costa Rica's Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.
Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. The film centers on the fictional Isla Nublar, an islet located off Costa Rica's Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.

Who directed Jurassic Park?
Steven Spielberg

Can you describe Jurassic Park in a few words?
A science fiction, action, and adventure movie

In Jurassic Park, what is the name of the island where the park is built?
Isla Nublar

Much more information available than in KB QA is harder but no need for open IE
WikiMovies

Objective:
We want to compare the performance of various learning methods between all three settings (raw, open IE, KB).

A new analysis tool for QA
- A set of 100k question-answer pairs (based on SimpleQuestions)
- 3 Knowledge Sources
  - A KB based on OMBD/MovieLens
  - Raw text extracted from Wikipedia
  - An imperfect KB built with open IE methods
- Answers to all questions appear in both the KB and Wikipedia text
Key-Value Memory Networks
Memory Networks
(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)

- Incorporates reasoning with attention over memory (RAM)
- Uses collection of dynamic, addressable slots for storing information
- Learns to read and weight memory slots
- Combines input and memories to choose a response
- Has been successfully applied to QA tasks (among others)
  - e.g. bAbi tasks, Children’s Book Tasks, etc.
Memory Networks for QA
(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)
Memory Networks for QA

(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)
Memory Networks for QA

(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)
Memory Networks for QA
(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)

Embed question

$q$

Embed memories

Inner product to compute attention scores

Memories

Embed question

Inner product to compute attention scores
Memory Networks for QA
(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)

Embed question

$q$

Embed questions

Inner product to compute attention scores

$p_{su}$

Weight memories

Weighted average

O

Memories
Memory Networks for QA
(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)

Question

Embed memories

Weight memories

Repeat

Hops $j = 1, \ldots, H$

Combine input + output

Inner product to compute attention scores

Memories

Embed question

$q$

$q_{j+1}$

$R_j$

$p_{wu}$

$o$

$q_{H+1}$

Weighted average
Memory Networks for QA

(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)

Embed question

Combine input + output

Inner product to compute attention scores

Embed memories

Weight memories

Embed candidates

Repeat

Hops \( j = 1, \ldots, H \)

Inner product

Softmax

Weighted average

Memories

Question

\( q \)

\( q_{j+1} \)
Memory Networks for QA
(Weston et al., ICLR15; Sukhbaatar et al., NIPS15)

Embed question

Combine input + output
Inner product to compute attention scores

Embed memories

Weight memories

Score
Inner product

weighted average

argmax

Answer
Memory Networks for QA from KB
(Bordes et al., arxiv15)

Knowledge Base
Jurassic Park directed_by Steven Spielberg
Jurassic Park written_by David Koepp, Michael Crichton
Jurassic Park release_year 1993
Jurassic Park has_tags action, adventure, science fiction, ...

Question
Who directed Jurassic Park?

Candidates
Steven Spielberg
1993
Blade Runner
Jeff Goldblum
...

Answer
Steven Spielberg
Memory Networks for QA from Text
(Hill et al., ICLR16)

Wikipedia Article
Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. The film centers on the fictional Isla Nublar, an islet located off Costa Rica's Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.

Question
Who directed Jurassic Park?

Answer
Steven Spielberg

Candidates
- Steven Spielberg
- 1993
- Blade Runner
- Jeff Goldblum
- ...

Memories

$m_1$: is a 1993 American science fiction action adventure film.
$m_2$: directed by Steven Spielberg and is
$m_3$: on the 1990 novel of
$m_4$: name by Michael Crichton, with
Comparison on WikiMovies

![Comparison Graph]

- **Standard QA System on KB (Bordes et al. 2014)**
  - KB: 78.5%
  - IE: 63.4%
  - Wikipedia: 69.9%

- **No Knowledge (Embeddings)**
  - KB: 54.4%

**Response Accuracy (%)**

- Memory Networks:
  - KB: 93.5%
  - IE: 74%
  - Wikipedia: 69.9%

**Hits @ 1**
Structuring Memories

Using symbolic structure

Use part of the memory to detect relevance (memory addressing), and the other part for returning a value (memory reading)

Leverage prior knowledge on the task

- One part designed to match the question ← The key
- One part designed to match the answer ← The value
Structuring Memories

KB Triple:

Jurassic Park directed by Steven Spielberg
Structuring Memories

KB Triple:

Key
Jurassic Park directed by

Value
Steven Spielberg
Structuring Memories

KB Triple:

Key
Jurassic Park \textit{directed by}

Who directed Jurassic Park?

Value
Steven Spielberg
Structuring Memories

Text Window:

directed by Steven Spielberg and is
<table>
<thead>
<tr>
<th>Key</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>directed by Steven Spielberg and is</td>
<td>Steven Spielberg</td>
</tr>
</tbody>
</table>
Who directed Jurassic Park?

Key:
- directed by Steven Spielberg and is

Value:
- Steven Spielberg
Structuring Memories

Text Window: Can we add prior knowledge to help?

Key

directed by Steven Spielberg and is

Who directed Jurassic Park?

Value

Steven Spielberg

Steven Spielberg
Can we add prior knowledge to help?
Key-Value Memory Networks

- Generalizes the way context is stored in memory
- Uses a pair of feature extractors — one each for “keys” and “values”
- Memories stored in the form of \((key, value)\) pairs
- Provides a simple way to tailor memories
  - Better recall for lookup
  - Better precision for value reading
- Note: Standard Memory Networks are a special case when key = value
Key-Value Memory Networks from KB

Knowledge Base
Jurassic Park directed_by Steven Spielberg
Jurassic Park written_by David Koepp, Michael Crichton
Jurassic Park release_year 1993
Jurassic Park has_tags action, adventure, science fiction, …

Question
Who directed Jurassic Park?

Candidates
Steven Spielberg
1993
Blade Runner
Jeff Goldblum
…

Answer
Steven Spielberg
Jurassic Park is a 1993 American science fiction action adventure film. The film was directed by Steven Spielberg and is the first installment of the "Jurassic Park" franchise. It is based on the 1990 novel of the same name by Michael Crichton, with a screenplay written by Crichton and David Koepp. The film centers on the fictional Isla Nublar, an islet located off Costa Rica’s Pacific Coast, where a billionaire philanthropist and a small team of genetic scientists have created a wildlife park of cloned dinosaurs.
Comparison on WikiMovies

<table>
<thead>
<tr>
<th>Dataset</th>
<th>Memory Networks</th>
<th>Key-Value Memory Networks</th>
</tr>
</thead>
<tbody>
<tr>
<td>KB</td>
<td>78.5</td>
<td>93.9</td>
</tr>
<tr>
<td>IE</td>
<td>63.4</td>
<td>68.3</td>
</tr>
<tr>
<td>Wikipedia</td>
<td>69.9</td>
<td>76.2</td>
</tr>
</tbody>
</table>

Response Accuracy (%)

- Standard QA System on KB (Bordes et al 2014)
- No Knowledge (Embeddings)

Closing the gap

Hospitalization Accuracy (%)

- 75.0
- 72.0
- 69.0

KB: Knowledge Base; IE: Information Extraction; Wikipedia
WikiQA

QA Benchmark in the answer selection setting

- Given question + set of sentences (paragraph), select which sentence contains the answer
- Key-value Memories: (window, sentence)

A glacier cave is a cave formed within the ice of a glacier
WikiQA

QA Benchmark in the answer selection setting

• Given question + set of sentences (paragraph), select which sentence contains the answer

• Key-value Memories: (window, sentence)

  A glacier cave is a / A glacier cave is a cave formed within the ice of a glacier

  a cave formed within the / A glacier cave is a cave formed within the ice of a glacier

  the ice of a glacier / A glacier cave is a cave formed within the ice of a glacier

• Training size is very small (~1k examples)
  
  - Word embeddings pre-trained on Wikipedia then fixed
  
  - Input, memory, and label dropout regularization
## WikiQA

<table>
<thead>
<tr>
<th>Method</th>
<th>MAP</th>
<th>MRR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Word Cnt</td>
<td>0.4891</td>
<td>0.4924</td>
</tr>
<tr>
<td>Wgt Word Cnt</td>
<td>0.5099</td>
<td>0.5132</td>
</tr>
<tr>
<td>2-gramm CNN (Yang et al., 2015)</td>
<td>0.6520</td>
<td>0.6652</td>
</tr>
<tr>
<td>AP-CNN (Santos et al., 2016)</td>
<td>0.6886</td>
<td>0.7069</td>
</tr>
<tr>
<td>Attentive LSTM (Miao et al., 2015)</td>
<td>0.6886</td>
<td>0.7069</td>
</tr>
<tr>
<td>Attentive CNN (Yin and Schütze, 2015)</td>
<td>0.6921</td>
<td>0.7108</td>
</tr>
<tr>
<td>L.D.C. (Wang et al., 2016)</td>
<td>0.7058</td>
<td>0.7226</td>
</tr>
<tr>
<td>Memory Network</td>
<td>0.5170</td>
<td>0.5236</td>
</tr>
<tr>
<td>Key-Value Memory Network</td>
<td>0.7069</td>
<td>0.7265</td>
</tr>
</tbody>
</table>
Conclusion

Key-Value Memory Networks
Promising model for jointly using symbolic and continuous memory
- Can be trained end-to-end using backpropagation + SGD
- Introduces great flexibility in how memories are designed
- Future work: selecting key and value content, applying to other domains

WikiMovies (https://research.facebook.com/research/babi/)
- New tool for evaluating and comparing learning algorithms
- Training / Evaluation sets of reasonable size
- Future work: close gap between KB and raw text