Graph Neural Networks for Knowledge Base Question Answering

Daniil Sorokin
@daniilmagpie 👋
https://daniilsorokin.github.io

South England NLP Meetup
January 21, 2019
| Bachelor in Applied Linguistics at St. Petersburg State University, Russia, 2011 |
| Master in Computational Linguistics at University of Tübingen, Germany, 2014 |
| PhD in NLP at the UKP Lab, Darmstadt University of Technology, Germany, since 2015, [http://ukp.tu-darmstadt.de](http://ukp.tu-darmstadt.de) |
What is Princess Leia’s home planet?

Entity Linking

Princess Leia, planet

Semantic interpretation

??

Knowledge Base

Princess Leia

SECRET DAUGHTER

Luke

SIBLING

Rey

HOME WORLD

Alderaan

INSTANCE OF

planet

Answer

Alderaan
Knowledge Base Question Answering

What is Princess Leia’s home planet?

Entity Linking

Semantic interpretation

Knowledge Base

Answer

Alderaan
Knowledge Base Question Answering

*What is Princess Leia’s home planet?*

Entity Linking

Princess Leia  planet

Semantic interpretation

Luke  0.3
Alderaan  0.7

Knowledge Base

SECRET DAUGHTER
Luke  Rey

SIBLING
Princess Leia

HOME WORLD
Alderaan  planet

INSTANCE OF

Answer

Alderaan
What is Princess Leia’s home planet?

Entity Linking

Princess Leia

planet

Semantic interpretation

Structured representation

Princess Leia

planet

Knowledge Base

Luke

Rey

Princess Leia

Alderaan

planet

SECRET DAUGHTER

SIBLING

HOME WORLD

INSTANCE OF

Answer

Alderaan
Knowledge Base Question Answering

What is Princess Leia’s home planet?

Entity Linking

Semantic interpretation
Structured representation

Knowledge Base

Answer
Alderaan
What is Princess Leia’s home planet?

Entity Linking

Semantic interpretation

SPARQL query

```
SELECT DISTINCT ?q WHERE {
  GRAPH <statements> {
    Princess Leia p:homeworld ?q.
    ?q rdf:type planet
  }
}
```
Knowledge Base Question Answering

What is Princess Leia’s home planet?

Entity Linking

Semantic interpretation

Knowledge Base

Answer

Alderaan
Structured Representation Approach

What is Princess Leia’s home planet?
Possible Semantic Graphs

What is Princess Leia’s home planet?

Yih et al. (2015)
Possible Semantic Graphs

$q \xrightarrow{\text{INSTANCE OF}} \text{planet} \xrightarrow{\text{HOME WORLD}} x \xrightarrow{\text{SIBILING}} \text{Princess Leia}

What is the home planet of Princess Leia’s family?
Where was Princess Leia’s first home?
Possible Semantic Graphs

$q \quad \text{ INSTANCE OF } \quad \text{planet}

\text{HOME WORLD} \quad \Rightarrow \quad \text{Princess Leia}
Representation structure is a graph and we model it as such
Previous results on WebQSP

- Yih et al. (2015) - Graph features
- Jain (2016) - Answer scoring
- Reddy et al. (2016) - Start with syntax
- Berant et al. (2014) - Start with syntax

The graph shows the F-score on the y-axis and the number of KB relations per question on the x-axis. The results indicate a decrease in F-score as the number of KB relations increases.
Previous approaches do not integrate the structure of the representation into the model.
Modeling the knowledge base graph and other graphs

1. Graph Convolutional Networks
   - Kipf and Welling (2017) – document classification
   - Marcheggiani and Titov (2017) – encode syntax dependencies for SRL
   - Schlichtkrull et al. (2018) – KB embeddings

2. Gated Graph Neural Networks
   - Scarselli et al. (2009) – Graph Neural Networks
   - Li et al. (2016) – Gated variant
   - Chen et al. (2018) – Structured Dialogue Policy
   - Sun et al. (2018) – QA with Fusion of Knowledge Bases and Text
What is Princess Leia’s home planet?

Knowledge Base Question Answering

Entity Linking

Semantic interpretation

Knowledge Base

Answer

Alderaan
Representation Learning

What is Princess Leia’s home planet?

Entity Linking

Princess Leia  planet

Generate semantic graphs

Encode the question

Select the best matching graph

Encode each graph
Graph Neural Networks – Scarselli et al. (2009)
Graph Neural Networks
Graph Neural Networks
Graph Neural Networks

- Node a
- Node b
- Node c
- Edge R1 from a to b
- Edge R2 from a to c
- Additional nodes and connections in the background
Gated Graph Neural Networks
– Li et al. (2016)

Update and reset gates
Gated Graph Neural Networks
Gated Graph Neural Networks with Labels

Princess Leia

q

HOME WORLD

INSTANCE OF

planet
Propagation Model

Previous state
\[
\begin{bmatrix}
  h_1^{(t-1)T} & \ldots & h_{|\mathcal{Z}|}^{(t-1)T}
\end{bmatrix}
\]

Relation encodings
\[
\begin{align*}
  h' = W_{\rightarrow} h_t \\
  h'' = W_{\leftarrow} h_t
\end{align*}
\]

Update vector
\[
a^{(t)}_v = A_v^T \begin{bmatrix}
  h_1^{(t-1)T} & \ldots & h_{|\mathcal{Z}|}^{(t-1)T}
\end{bmatrix} - A_r^T \begin{bmatrix}
  h_1^T & h_{|\mathcal{R}|}^T, h''_1^T & \ldots & h''_{|\mathcal{R}|}^T
\end{bmatrix}
\]

Update and reset gates
\[
\begin{align*}
  z^t_v &= \sigma(W^z a^{(t)}_v + U^z h_v^{(t-1)} + b^z) \\
  r^t_v &= \sigma(W^r a^{(t)}_v + U^r h_v^{(t-1)} + b^r)
\end{align*}
\]

Update the hidden states
\[
\begin{align*}
  \tilde{h}^{(t)}_v &= \tanh(W a^{(t)}_v + U (r^t_v \odot h_v^{(t-1)}) + b) \\
  h_v^{(t)} &= (1 - z^t_v) \odot h_v^{(t-1)} + z^t_v \odot \tilde{h}^{(t)}_v
\end{align*}
\]
Propagation Model

Previous state
\[
\begin{bmatrix}
  h_1^{(t-1)T} \\
  \vdots \\
  h_{|\mathcal{Y}|}^{(t-1)T}
\end{bmatrix}
\]

Relation encodings
\[
\begin{align*}
  h'_r &= W \rightarrow h'_r \\
  h''_r &= W \leftarrow h'_r
\end{align*}
\]

Update vector
\[
\begin{align*}
  a_v^{(t)} &= A^T \begin{bmatrix}
  h_1^{(t-1)T} \\
  \vdots \\
  h_{|\mathcal{Y}|}^{(t-1)T}
\end{bmatrix} + A^T_r \begin{bmatrix}
  h'_{1T} \\
  \vdots \\
  h'_{|\mathcal{R}|T}
\end{bmatrix}
\end{align*}
\]

Update and reset gates
\[
\begin{align*}
  z^t_v &= \sigma \left( W^z a_v^{(t)} + U^z h_v^{(t-1)} + b^z \right) \\
  r^t_v &= \sigma \left( W^r a_v^{(t)} + U^r h_v^{(t-1)} + b^r \right)
\end{align*}
\]

Update the hidden states
\[
\begin{align*}
  \tilde{h}_v^{(t)} &= \tanh \left( W a_v^{(t)} + U \left( r^t_v \odot h_v^{(t-1)} \right) + b \right) \\
  h_v^{(t)} &= (1 - z^t_v) \odot h_v^{(t-1)} + z^t_v \odot \tilde{h}_v^{(t)}
\end{align*}
\]
Propagation Model

Previous state

\[
\begin{bmatrix}
h_1^{(t-1)T} & \cdots & h_{|\mathcal{V}|}^{(t-1)T}
\end{bmatrix}
\]

Relation encodings

\[
h'_r = W_r h_r, \quad h''_r = W'_r h_r
\]

Update vector

\[
a_v^{(t)} = A_v^T \begin{bmatrix} h_1^{(t-1)T} & \cdots & h_{|\mathcal{V}|}^{(t-1)T} \end{bmatrix} + A_r^T \begin{bmatrix} h_1^{T} & \cdots & h_{|\mathcal{R}|}^{T}, h'_1 & \cdots & h'_{|\mathcal{R}|} \end{bmatrix}
\]

Update and reset gates

\[
z^t_v = \sigma \left( W^z a_v^{(t)} + U^z h_v^{(t-1)} + b^z \right)
\]

\[
r^t_v = \sigma \left( W^r a_v^{(t)} + U^r h_v^{(t-1)} + b^r \right)
\]

Update the hidden states

\[
\tilde{h}_v^{(t)} = \tanh \left( W a_v^{(t)} + U \left( r_v^t \odot h_v^{(t-1)} \right) + b \right)
\]

\[
h_v^{(t)} = (1 - z_v^t) \odot h_v^{(t-1)} + z_v^t \odot \tilde{h}_v^{(t)}
\]
Entity and Relation Embeddings
Entity and Relation Encodings

Word embeddings

Princess Leia

$q$

HOME WORLD

INSTANCE OF
Graph Encoding

Princess Leia

HOME WORLD

 INSTANCE OF

planet
Graph Encoding

$q$ is an instance of a planet, which is Leia's home world.
The Graph Neural Network models all aspects of the semantic graph representation.
What is Princess Leia’s home planet?

Entity Linking

Generate semantic graphs

Encode the question

Select the best matching graph

Encode each graph

Question Answering Pipeline – Representation Learning
Question Answering Pipeline – Representation Learning

What is Princess Leia’s home planet?

Encode the question

Select the best matching graph

Encode each graph

Entity Linking

Generate semantic graphs
Question Answering Pipeline – Representation Learning

What is Princess Leia’s home planet?

Entity Linking

Encode the question

Ranking loss

Generate semantic graphs

Encode each graph

CNN

COS

GNN
Question Answering Pipeline – Representation Learning

What is Princess Leia’s home planet?

Encode the question

Entity Linking

Generate semantic graphs

Encode each graph

CNN

COS

Cosine similarity

Ranking loss
Models

• STAGG (Yih et al. 2015)
• Single Edge
• Pooled Edges
• GNN
• Gated GNN
## Data set

<table>
<thead>
<tr>
<th>WebQSP-WD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Based on WebQSP/WebQuestions</td>
</tr>
<tr>
<td>Mapped to Wikidata</td>
</tr>
<tr>
<td>1033 test questions (293 complex)</td>
</tr>
<tr>
<td>2880 train questions (419 complex)</td>
</tr>
</tbody>
</table>
What is Princess Leia's home planet?

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1.000</td>
<td>1.000</td>
<td>1.000</td>
</tr>
</tbody>
</table>
Evaluation

*What is Princess Leia’s home planet?*

<table>
<thead>
<tr>
<th></th>
<th>Precision</th>
<th>Recall</th>
<th>F-score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.050</td>
<td>1.000</td>
<td>0.095</td>
</tr>
<tr>
<td>Model</td>
<td>Precision</td>
<td>Recall</td>
<td>F-score</td>
</tr>
<tr>
<td>-----------</td>
<td>-----------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>STAGG</td>
<td>0.173</td>
<td>0.214</td>
<td>0.167</td>
</tr>
<tr>
<td>Single Edge</td>
<td>0.250</td>
<td>0.307</td>
<td>0.243</td>
</tr>
<tr>
<td>Pooled Edges</td>
<td>0.245</td>
<td>0.302</td>
<td>0.240</td>
</tr>
<tr>
<td>GNN</td>
<td>0.253</td>
<td>0.293</td>
<td>0.243</td>
</tr>
<tr>
<td>GGNN</td>
<td>0.259</td>
<td>0.305</td>
<td><strong>0.250</strong></td>
</tr>
</tbody>
</table>
WebQSP-WD

![WebQSP-WD Graph](image-url)
WebQSP-WD simple questions
WebQSP-WD complex questions
Example output

Where did Harper Lee attend high school?
Example output Pooled

What language do people speak in Brazil?
Example output GGNN

What language do people speak in Brazil?
Limitations

1. The small training set limits the more complex GraphNNs architecture
2. Small test set
3. WebQSP contains a lot of data set errors and is not entirely covered by Wikidata (around 30% of errors)
4. Not many Wikidata data sets for a meaningful comparison
5. Entity linking is a bottle neck (10% of errors)
Modelling the semantic structure with **GraphNNs** improves the representation learning for knowledge base graphs.

Knowledge base question answering

Your question

Who played Luke Skywalker in Star Wars?
Follow up?

- [https://daniilsorokin.github.io](https://daniilsorokin.github.io) 🔄 slides
- [sorokin@ukp.informatik.tu-darmstadt.de](mailto:sorokin@ukp.informatik.tu-darmstadt.de)
- [@daniilmagpie](https://twitter.com/daniilmagpie)