DADgraph: A Discourse-aware Dialogue Graph Neural Network for Multiparty Dialogue Machine Reading Comprehension

Jiaqi Li, Ming Liu, Zihao Zheng, Heng Zhang, Bing Qin, Min-Yen Kan, Ting Liu

Harbin Institute of Technology
National University of Singapore

July 21
IJCNN 2021
SQuAD, RACE, NarrativeQA, CoQA, QuAC etc.

Oxygen
The Stanford Question Answering Dataset

In the meantime, on August 1, 1774, an experiment conducted by the British clergyman Joseph Priestley focused sunlight on mercuric oxide (HgO) inside a glass tube, which liberated a gas he named "dephlogisticated air". He noted that candles burned brighter in the gas and that a mouse was more active and lived longer while breathing it. After breathing the gas himself, he wrote: "The feeling of it to my lungs was not sensibly different from that of common air, but I fancied that my breast felt peculiarly light and easy for some time afterwards." Priestley published his findings in 1775 in a paper titled "An Account of Further Discoveries in Air" which was included in the second volume of his book titled Experiments and Observations on Different Kinds of Air. Because he published his findings first, Priestley is usually given priority in the discovery.

Why is Priestley usually given credit for being first to discover oxygen?
Ground Truth Answers: published his findings first; he published his findings first; he published his findings first; Because he published his findings first.

Article: Endangered Species Act
Paragraph: “...Other legislation followed, including the Migratory Bird Conservation Act of 1929, a 1937 treaty prohibiting the hunting of right and gray whales, and the Bald Eagle Protection Act of 1940. These later laws had a low cost to society—the species were relatively rare—and little opposition was raised.”

Question 1: “Which laws faced significant opposition?”
Plausible Answer: later laws

Question 2: “What was the name of the 1937 treaty?”
Plausible Answer: Bald Eagle Protection Act
A dialog from *Molweni* dataset.

*jimcooncat*: installing acroread gives me a 404 on maverick -- what to do?  

*jrib*: where are you installing acroread from?  

*elfranne*: people in the same local network?  

*llutz*: not network, on local computer  

*elfranne*: so its only available for ``` localhost ``` and not others on the same local network  

*jimcooncat*: thank you, i had forgot to update  

*llutz*: yes, ``` other users on localhost ```  

**Q1:** Why does *jimcoonact* meet the error?  

**A1:** forgot to update  

**Q2:** Where does *llutz* install acroread?  

**A2:** on local computer  

**Q3:** How did *erUSUL* create a new partition table?  

**A3:** NA.
• Discourse structure of multiparty dialogue.

**jimcooncat**: installing acroread gives me a 404 on maverick -- what to do? $U_1$

**jrib**: where are you installing acroread from? $U_2$

**elfranne**: people in the same local network? $U_3$

**llutz**: not network, **on local computer** $U_4$

**elfranne**: so its only available for ```localhost``` and not others on the same local network $U_5$

**jimcooncat**: thank you, i had **forgot to update** $U_6$

**llutz**: yes, ```other users on localhost```. $U_7$
• Discourse structure has been successfully applied to QA and MRC.

• Graph structure has been proven to effectively represent dialogs.

Discourse structure informs multiparty dialogue MRC performance in modeling long-term dependencies.
Method: DADgraph

Sequential Context Encoding

- $u_1 \rightarrow \text{GRU} \rightarrow g_1$
- $u_2 \rightarrow \text{GRU} \rightarrow g_2$
- $u_3 \rightarrow \text{GRU} \rightarrow g_3$
- $u_4 \rightarrow \text{GRU} \rightarrow g_4$
- $u_5 \rightarrow \text{GRU} \rightarrow g_5$

Discourse Graph Modeling

- Links between utterances from same speaker.
- Links between utterances from different speaker.
- Elaboration relation.
- Question-answer pair relation.
- Acknowledgement relation.
- There are many other discourse relations not in the graph.

MRC

- $h_1 \rightarrow \text{GRU} \rightarrow h_2 \rightarrow \text{GRU} \rightarrow h_3 \rightarrow \text{GRU} \rightarrow h_4 \rightarrow \text{GRU} \rightarrow h_5$
- $w_1, w_i, w_j, w_m$: $i^{th}$ word in dialog;
- $q$: question

Interaction

- $t_1, t_i, t_j, t_m$
• Sequential context modelling
  • The pretrained BERT model is used to represent the utterance to get $u_i$.
  • The sequence structure of dialogue is modeled by GRU.
  • Finally, the utterance representation $g_i$ of fusion context is obtained;
• Discourse graph modelling
  • The utterance representation $g_i$ obtained in the previous step is taken as the input;
  • GCN is used to model discourse structure of dialogue.
  • The updated representation $h_i$ of fusion discourse structure is obtained;
**Method**

- **MRC module**
  - First, let the $h_i$ attend with each word in the input dialogue.
  - Weighted sum of utterances representations and obtains $t_i$.
  - Predict the probability of the *span* $(i, j)$ via $t_i$ and $t_j$. 

WS: weighted sum
• Dataset: *Molweni* corpus.

<table>
<thead>
<tr>
<th></th>
<th>Train</th>
<th>Dev</th>
<th>Test</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dialogs</td>
<td>8,771</td>
<td>883</td>
<td>100</td>
<td>9,754</td>
</tr>
<tr>
<td>Utterances</td>
<td>77,374</td>
<td>7,823</td>
<td>845</td>
<td>86,042</td>
</tr>
<tr>
<td>Questions</td>
<td>24,682</td>
<td>2,513</td>
<td>2,871</td>
<td>30,066</td>
</tr>
</tbody>
</table>

### Main results

<table>
<thead>
<tr>
<th>Model</th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>BiDAF</td>
<td>22.9</td>
<td>39.8</td>
</tr>
<tr>
<td>DocQA</td>
<td>42.5</td>
<td>56.0</td>
</tr>
<tr>
<td>BERT</td>
<td>45.3</td>
<td>58.0</td>
</tr>
<tr>
<td>DialogueRNN</td>
<td>45.4</td>
<td>60.9</td>
</tr>
<tr>
<td>DialogueGCN</td>
<td>45.7</td>
<td>61.0</td>
</tr>
<tr>
<td>DADgraph (Our)</td>
<td>46.5</td>
<td>61.5</td>
</tr>
<tr>
<td>Human performance</td>
<td>64.3</td>
<td>80.2</td>
</tr>
</tbody>
</table>
## Experiments

- **Ablation results**

<table>
<thead>
<tr>
<th></th>
<th>EM</th>
<th>F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>DADgraph</td>
<td>46.5</td>
<td>61.5</td>
</tr>
<tr>
<td>- w/o discourse relations</td>
<td>44.9</td>
<td>60.6</td>
</tr>
<tr>
<td>- w/o discourse structure</td>
<td>44.7</td>
<td>60.5</td>
</tr>
</tbody>
</table>
Experiments: case study

sipher: bacon5o there 's no `` fixmbr '' with ubuntu.

morfic: xaa is old acceleration architecture, exa is the new one, font rendering is so much filepath

bacon5o: i dont want ubuntu, it does n't support my internet, thus i can not use it

morfic: your internet is different from mine? damn bush and his internets!

bacon5o: my internet is different! why you ask?

morfic: your possessive ` ` my '' on the internet

bacon5o: i use a wireless accesspoint that plugs into my usb which then goes into my motherboard

Q1: What does bacon5o not want to use?
Gold: ubuntu DialogueRNN: a wireless accesspoint DialogueGCN: it does n't support my internet Our model: ubuntu

Q2: Which one is new acceleration architecture?

Q3: What is missing from ubuntu?
• Propose **DADgraph** model for multiparty dialogue MRC task.
• Prove the discourse structure can help understand the dialogue.

Thank you!