Using a KG-Copy Network for Non-Goal Oriented Dialogues

[Reproduced Research Paper]

Debanjan Chaudhuri  University of Bonn/Fraunhofer IAIS
Md Rashad Al Hasan Rony  University of Bonn/Fraunhofer IAIS
Dr. Simon Jordan  Volkswagen Group Research
Prof. Dr. Jens Lehmann  University of Bonn/Fraunhofer IAIS
A Primer on Dialogue Systems

- Automatic (human-like) conversation with users on a particular domain/topic
- Generally a major challenge in AI
- Some use cases:
  - Emphatic Assistance
  - Therapy Chatbot
  - Customer Facing Agents (Call Centres)
  - Website Navigation (for complex websites)
  - Training and Education
Types of Dialogue Systems & Focus of the Talk

- Based on Application:
  - Goal-oriented/Task-oriented Systems
  - Non-goal-oriented Systems
- Based on Implementation:
  - Generative Systems
  - Retrieval Based
Generative Dialogue Systems

- Systems able to generate dialogues given an input utterance
- Often implemented using **Seq-to-seq** based neural network models
- At each time-step, the neural network picks a word from the vocabulary
Generative systems can produce **natural sounding dialogues** compared to other approaches.

However, generative architectures **cannot produce coherent, knowledge-grounded responses**.

---

**Can the Semantic Web community help to solve a major AI challenge?**

---

* Which Stadium does Arsenal play in?

* I think Lionel Messi
Research hypothesis:

By including information from a knowledge graph into the output of a generative system, the correctness of the dialogue system can improve while the answers remain natural.
Contributions in this Paper

1. Football **knowledge graph** with information regarding famous club teams and countries

2. New **training dataset** for non-goal oriented conversations over football consisting of both knowledge-grounded dialogues and chit-chats

3. Novel, end-to-end **neural network architecture** (KG-Copy) for integrating knowledge graphs into the dialogue generation process
(1) Football Knowledge Graph

Created a Knowledge Graph about football teams (club and country) from crawling Wikipedia articles, following Bergmann et. al. [2].

<table>
<thead>
<tr>
<th>Total # of Triples in KG</th>
<th>4433</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total # of Entity</td>
<td>934</td>
</tr>
<tr>
<td>Total # of Relations</td>
<td>30</td>
</tr>
</tbody>
</table>
(2) Why a New Dataset for Non-Goal Oriented Dialogues

- Eric et. al. [1] introduced a dataset (in-car) for goal-oriented dialogues over knowledge graphs.
- Dodge et. al. [4] introduced dataset for having factoid conversations in the domain of movies, but the factoid replies are not well-articulated.
- **Gap:** Non-goal-oriented dataset with well-articulated responses using both chit-chat and factual information.

Which Stadium does Arsenal play in?
- Emirates Stadium

Which Stadium does Arsenal play in?
- Arsenal’s home ground is **Emirates Stadium**.
(2) Knowledge-grounded Non-Goal Oriented Dialogue Dataset


- Asked turkers to converse about these teams, asking factual questions as well as having chit-chats/small-talks
- Furthermore, we did another task to keep only coherent conversations from the previous dataset
(2) Statistics of Football Conversation Dataset

<table>
<thead>
<tr>
<th>Split</th>
<th># of Dialogues</th>
<th># of Utterances</th>
</tr>
</thead>
<tbody>
<tr>
<td>Train</td>
<td>2,493</td>
<td>12,243</td>
</tr>
<tr>
<td>Validation</td>
<td>149</td>
<td>737</td>
</tr>
<tr>
<td>Test</td>
<td>348</td>
<td>1,727</td>
</tr>
</tbody>
</table>

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Size of Vocabulary</td>
<td>4,782</td>
</tr>
<tr>
<td>Total # of Teams</td>
<td>30</td>
</tr>
<tr>
<td>Avg. # of Conversations per team</td>
<td>83</td>
</tr>
</tbody>
</table>
(3) Challenges integrating KGs for Non-goal Oriented Dialogue Systems

- System should be able to produce well-articulated as well as knowledge-grounded responses
- Natural Language Generation from triples is challenging [8]
- The systems should be able to detect when to use the knowledge graph and when to use elements from the vocabulary
3) KG Copy Network Architecture
(3) Sentient Gating Mechanism

Who is the captain of Argentina?

Knowledge Graph (KG)

<table>
<thead>
<tr>
<th>Subject</th>
<th>Relation</th>
<th>Object</th>
</tr>
</thead>
<tbody>
<tr>
<td>Argentina</td>
<td>coach</td>
<td>Lionel Scaloni</td>
</tr>
<tr>
<td>Argentina</td>
<td>captain</td>
<td>Lionel Messi</td>
</tr>
<tr>
<td>..........</td>
<td>..........</td>
<td>............</td>
</tr>
</tbody>
</table>

Encoder

Encoded Vector Representation

Vocab Distribution

[-0.68, 1.075, ...]

Object Distribution

[0.264, 0.289, ...]

Sentient Gate

Decoder

lionel messi

t = 0
### (3) Evaluation Results

<table>
<thead>
<tr>
<th>Model</th>
<th>BLEU</th>
<th>Entity F1</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Valid</td>
<td>Test</td>
</tr>
<tr>
<td>Vanilla Seq2Seq</td>
<td>1.04</td>
<td>0.82</td>
</tr>
<tr>
<td>Mem2Seq [7]</td>
<td>1.30</td>
<td>0.52</td>
</tr>
<tr>
<td>KG Copy</td>
<td>2.56</td>
<td>2.05</td>
</tr>
</tbody>
</table>

*Results on Football Conversations Dataset*

<table>
<thead>
<tr>
<th>Model</th>
<th>BLEU</th>
<th>Entity F1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Vanilla Seq2Seq</td>
<td>8.4</td>
<td>10.3</td>
</tr>
<tr>
<td>Mem2Seq [7]</td>
<td>12.6</td>
<td>33.4</td>
</tr>
<tr>
<td>KG Copy</td>
<td>9.6</td>
<td>52.8</td>
</tr>
</tbody>
</table>

*Results on In-car [1] Test Set*

<table>
<thead>
<tr>
<th>Model</th>
<th>Correctness (1-5)</th>
<th>Human-like (1-5)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mem2Seq [7]</td>
<td>1.30</td>
<td>2.44</td>
</tr>
<tr>
<td>KG Copy</td>
<td>2.26</td>
<td>3.88</td>
</tr>
</tbody>
</table>

*Qualitative Evaluation*
(3) A Conversation with the KG Copy Model

- The first response is from the vocabulary and the others copy facts from the KG
- Performs implicit coreference resolution

```
i like the team pretty much
i don't think they're a lot of winning
who is the captain of argentina?
lionel messi is the captain
do you know the name of their coach?
lionel scaloni is the coach
```
(3) Shortcomings

- The model is **only** able to handle **simple, factoid based questions**
- In the football conversation dataset, many questions are complex, such as "**who is the youngest player in nigeria**" → KG Copy cannot handle those
- **Memory** is based on similarity between the question and the entity / relation in the memory → more sophisticated techniques could perform better
Summary: How did we improve SOTA for knowledge-grounded non-goal oriented dialogues?

1. A football knowledge graph with recent facts about teams.
2. A real-world dataset of conversations over football across club teams and countries.
3. A novel, end-to-end neural network architecture: **KG Copy Network** - confirmed our research hypothesis that the inclusion of KG can improve correctness while maintaining natural dialogues
References


