A Multi-Agent Communication Framework for Question-Worthy Phrase Extraction and Question Generation

AAAI 2019

Siyuan Wang¹, Zhongyu Wei¹, Zhihao Fan¹, Yang Liu², Xuanjing Huang³
1 School of Data Science, Fudan University, China
2 LAIX Inc., China
3 School of Computer Science, Fudan University, China
Question Generation (for text)

- Given a sentence or paragraph, construct questions automatically

**Sentence:**
CBS provided digital streams of the game via CBSSports.com, and the CBS Sports apps on tablets, Windows 10, Xbox One and other digital media players (such as Chromecast and Roku).

**Questions:**
- What CBS website provided a stream?
- What version of Windows supported the CBS sports app?
- On what game console was the CBS Sports app available?
Introduction

• Two subtasks in Question Generation
  • What to say: determine the targets that should be asked
  • How to say: produce the surface-form of the question

This paper focus on

Most papers focus on
Motivation

Not all content pieces in the input are significant, the authors propose to use question-worthy phrases to identify which phrases are worthwhile to be asked about.

Moreover, if there are several focuses in an input and we extract question-worthy phrases, we can then generate various questions.
Contributions

• Propose to generate multiple questions given input sentence *without ground-truth answers*.

• Extract *question-worthy phrases* from the input sentence and generate questions based on such information.

• Developing a *multi-agents communication framework* to learn the two tasks simultaneously.
Framework

output: question

Iteratively

Generation Agent

$p_m$

$\mathbf{h}_m^q = [\mathbf{h}; \mathbf{c}]$

Message Passing Module

MP Extraction Agent

$\mathbf{h}_m^q = [\mathbf{h}; \mathbf{c}]$

$h^s$

Generation Agent

$(p_1, ..., p_M)$

Local Extraction Agent

input: sentence
Local extraction agent

- The Boundary Model of Pointer Network
- Pick M pairs of start and end index of phrases
Generation agent

- Seq2Seq Model + Attention Mechanism
- Each time we take both the sentence and an extracted phrase as input, to generate a question
Message Passing (MP) Extraction agent

- Match-LSTM + Pointer Network
- Use representations of input tokens from generation agent as auxiliary information
Experiments

• SQuAD Dataset
  • Answers, extractive from sentences, are treated as target question-worthy phrases
  • More than 30% sentences have multiple questions (One-to-Many here)

<table>
<thead>
<tr>
<th># of questions</th>
<th># of sentences</th>
<th>percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>41,356</td>
<td>67.11%</td>
</tr>
<tr>
<td>2</td>
<td>14,499</td>
<td>23.53%</td>
</tr>
<tr>
<td>3</td>
<td>3,921</td>
<td>6.36%</td>
</tr>
<tr>
<td>4</td>
<td>1,198</td>
<td>1.95%</td>
</tr>
<tr>
<td>≥ 5</td>
<td>649</td>
<td>1.05%</td>
</tr>
<tr>
<td>in total</td>
<td>61,623</td>
<td>100%</td>
</tr>
</tbody>
</table>

Table 1: Distribution of number of questions per sentence in our dataset.
Comparison of Extraction Models

- $\mathcal{E}_{NER}$: take recognized name entities as question-worthy phrases
- $\mathcal{E}_{Local}$: the extraction agent in local layer
- $\mathcal{E}_{MP}$: the extraction agent in message passing layer

<table>
<thead>
<tr>
<th>Model</th>
<th>EM</th>
<th>F1</th>
<th>avg. num</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mathcal{E}_{NER}$</td>
<td>13.12%</td>
<td>17.33</td>
<td>0.86</td>
</tr>
<tr>
<td>$\mathcal{E}_{Local}$</td>
<td>24.27%</td>
<td>38.63</td>
<td>1.43</td>
</tr>
<tr>
<td>$\mathcal{E}_{MP}$</td>
<td>35.77%</td>
<td>46.71</td>
<td>1.38</td>
</tr>
</tbody>
</table>

Table 2: Evaluation results of different phrase extraction models. (underline: diff. with both comparison models ($\mathcal{E}_{NER}$, $\mathcal{E}_{Local}$) $p < 0.01$; **Bold**: the best performance in the column)
Comparison of Generation Models

<table>
<thead>
<tr>
<th>Model</th>
<th>BLEU 1</th>
<th>BLEU 2</th>
<th>BLEU 3</th>
<th>BLEU 4</th>
<th>METEOR</th>
<th>ROUGE _L</th>
</tr>
</thead>
<tbody>
<tr>
<td>\text{NQG}_{\text{Rule}}</td>
<td>38.15</td>
<td>21.03</td>
<td>14.15</td>
<td>9.98</td>
<td>13.38</td>
<td>29.00</td>
</tr>
<tr>
<td>\text{NQG}_{\text{Base}}</td>
<td>43.83</td>
<td>23.80</td>
<td>14.46</td>
<td>9.05</td>
<td>14.63</td>
<td>36.50</td>
</tr>
<tr>
<td>\text{NQG}_{\text{NER}}</td>
<td>44.00</td>
<td>23.79</td>
<td>14.52</td>
<td>9.22</td>
<td>14.89</td>
<td>36.32</td>
</tr>
<tr>
<td>\text{NQG}_{\text{Local}}</td>
<td>44.36</td>
<td>24.58</td>
<td>15.23</td>
<td>9.76</td>
<td>15.15</td>
<td>37.00</td>
</tr>
<tr>
<td>\text{NQG}_{\text{MP}}</td>
<td>45.70*</td>
<td>25.87*</td>
<td>16.33*</td>
<td>10.56*</td>
<td>15.76*</td>
<td>38.09*</td>
</tr>
<tr>
<td>\text{NQG}_{G-t}</td>
<td>47.49</td>
<td>27.81</td>
<td>17.9</td>
<td>11.81</td>
<td>16.84</td>
<td>40.23</td>
</tr>
</tbody>
</table>

Table 3: Evaluation results of different question generation models in terms of BLEU 1-4, METEOR and ROUGE \_L. (underline: diff. with all the comparison models (\text{NQG}_{\text{Rule}}, \text{NQG}_{\text{Base}}, \text{NQG}_{\text{NER}}, \text{NQG}_{\text{Local}}) p < 0.01; *: p < 0.05; \textbf{Bold}: the best performance for each column)
Case Study

Sample 1

**Input:** the panthers finished the regular season with a 15 – 1 record, and quarterback cam newton was named the nlf most valuable player (mvp).

**Phrases**

- **Ground-truth:** 15 – 1, quarterback cam newton.
- **NER:** panthers, (blank).
- **E\_MP:** 15 quarterback cam newton.

**Questions**

- **Ground-truth:** what was the ratio in 2015 for the carolina panthers during their regular season? which carolina panthers player was named most valuable player?
- **NQG\_NER:** who won the regular season? what was the regular session in the afl?
- **NQG\_MP:** how many wins did the panthers win during the regular season? who was named the nlf most valuable player?

Sample 2

**Input:** next to the main building is the basilica of the sacred heart.

**Phrases**

- **Answers:** main building.
- **NER:** sacred heart
- **E\_MP:** next to the main building

**Questions**

- **Ground Truth:** the basilica of the sacred heart at notre dame is beside to which structure?
- **NQG\_NER:** what is next to main building?
- **NQG\_MP:** where is the basilica of prayer?