A Double Phases Generation Network for Yes or No Question Generation (Student Abstract)

Jiayuan Xie,[1,2] **Feng Chen,**[1,2] **Yi Cai,**[1,2*] **Zehang Lin**[3]

[1] School of Software Engineering, South China University of Technology, Guangzhou, China
[2] Key Laboratory of Big Data and Intelligent Robot (South China University of Technology), Ministry of Education
[3] Department of Computing, Hong Kong Polytechnic University, Hong Kong, China
{sexiejiayua, sefengchen}@mail.scut.edu.cn, ycai@scut.edu.cn, cszhlin@outlook.com

Abstract

This paper aims to solve the task of generating yes or no questions, which generates yes/no questions based on given passages. These questions can be used for evaluation automatically. We propose a double phases generation network that can identify specific phrases related to facts from the input passage and use them as auxiliary information for generation. Specifically, the 1st-phase prediction uses the extracted phrases as assistance to generate an initial question. Then, the 2nd-phase prediction utilizes an attention network to focus on the relevant phrases related to the initial question in the passage to generate questions that are more relevant to the specific facts contained in the initial question. Extensive experiments we performed on BoolQ dataset demonstrate the effectiveness of our framework.

Introduction

Question generation (QG) has obtained great attention in the field of education as it meets the requirements of quick and automatic evaluation and consistent scoring. However, most existing researches on QG problems mainly focus on generating some questions that require students to ask the question with some specific text fragments, which need to evaluate the answer manually.

In this paper, we study a new novel task of yes or no question generation (YNQG), which generates a yes/no question from a given passage. Specifically, the generated question only requires students to answer "yes" or "no", which is easy to mark and can even be automatically scored by a computer. Most existing studies on answer-aware QG utilize the given answer to help the model focus on certain specific fragments of the passage for better performance. However, the given answer (i.e., "yes" or "no") in YNQG does not contain useful information. Therefore, we propose to identify focuses from the passage as assistance to generate yes/no questions. That is because a complete extraction of phrases related to a specific fact in the passage can help the model to better generate yes/no questions.

To this end, we propose a novel framework named double phases generation network (DPGNet), which can extract as many phrases as possible related to a certain fact as auxiliary information for question generation. Specifically, in the 1st-phase, we use an extraction module to extract fact-related phrases from the passage to generate an initial question. In the 2nd-phase, we utilize an attention network in re-extraction to extract more phrases related to the content of the initial question from the passage. Besides, in order to reduce the negative impact of extracted meaningless phrases, e.g., the stopwords, a gate mechanism is proposed to control whether the extracted phrases are used for the re-generation module. With the help of the attention network and gate mechanism, DPGNet can generate questions that are more related to a specific fact than the initial questions. We conducted experiments on BoolQ dataset (Clark et al. 2019) and evaluated them showing the effectiveness of our framework.

Our Proposed Framework

Given an input passage x=(x_1,...,x_n), where x_i is a token, YNQG aims to generate a yes/no question y. Specifically, we propose to extract the fact-related phrases p=(w_1,...,w_n), $w_i \subset x$ as assistance to generate questions. The overall architecture of the DPGNet which contains two phases prediction is illustrated in Fig. 1.

1st-Phase Prediction

The 1st-phase prediction contains an extraction module and a generation module. In the extraction module, we use bi-LSTM followed by a fully-connected layer and a softmax activation to predict whether each token is a fact-related phrase. Specifically, the target phrase of the training set is the fragment where the ground truth question coincides with the passage. In the generation module, we utilize an encoder-decoder model with attention mechanism for generation. In order to take advantage of the additional information, i.e., the extracted phrases, we concatenate the word embedding and BIO tagging embedding (Wang et al. 2019) as input to the encoder-decoder to generate an initial question $\tilde{y}$.

2nd-Phase Prediction

The 2nd-phase prediction contains a re-extraction module and a re-generation module. To extract more complete phrases related to the fact that contained in the $\tilde{y}$, the re-extraction module first utilizes a passage-to-question attention network (Seo et al. 2017) to extract the phrases from

*Corresponding author: Yi Cai, ycai@scut.edu.cn

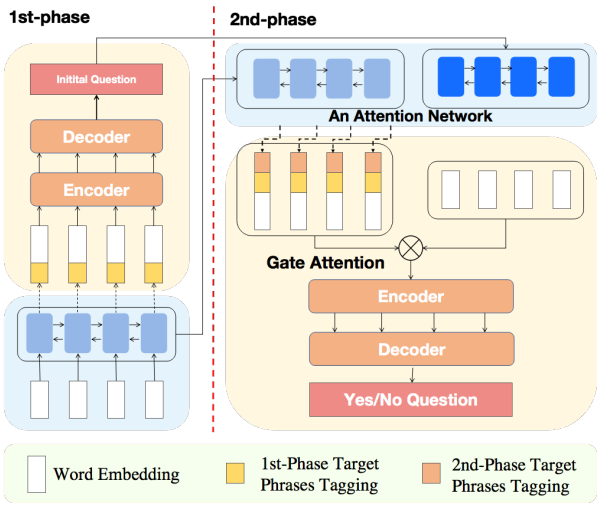

Figure 1: Overview of Double Phases Generate Network

the original passage according to the $\tilde{y}$. In addition, the re-generation module utilizes a gate mechanism to control whether the additional information (i.e., the BIO tagging embedding for 1st-phase and 2nd-phase) is used. Then, the final input can be obtained by the gate mechanism,

$$f_i = relu(W_f E(x_i)) \quad (1)$$

$$g_i = sigmoid(W_g E(x_i)) \quad (2)$$

$$\tilde{E}(x_i) = g_i \odot f_i + (1 - g_i) \odot e_i \quad (3)$$

where $E(x_i)$ represents the feature of the i-th word, i.e., concatenating the word embedding and BIO tagging embedding. W_f and W_g are trainable parameters.

Finally, the re-generation module takes the final initialization feature as input, and adopts the same encoder-decoder structure as the 1st-phase to generate yes/no questions.

Experiment

Dataset

We conduct experiments on BoolQ dataset, which is divided into three subsets: 9,427 for training, 1,635 for development and 1,635 for testing. Each sample includes a paragraph, an answer ("yes" or "no") and a question.

Experimental Details

In our implementation, we use the 300-dim GloVe embeddings as a fixed word embedding. We randomly initialize a trainable 50-dim phrase-embedding for BIO tags. We use Adam optimizer with a learning rate of 0.001 and train our models with a mini-batches for 15 epochs. We use a single layer in all LSTM and the number of hidden units is 256.

Baseline Methods

In order to prove the effectiveness of our framework, we make a comparison for the following generation models. NQG$_{base}$ uses an encoder-decoder model with an attention mechanism that takes passage as input. NQG$_{true}$ and NQG$_{1p}$ take ground truth phrases, and phrases extracted by 1st-phase as additional information for NQG$_{base}$ structure.

	BLEU 1	BLEU 4	ROUGH$_L$
NQG$_{base}$	0.391	0.140	0.349
NQG$_{true}$	0.446	0.177	0.397
NQG$_{1p}$	0.403	0.144	0.360
DPGNet	**0.420**	**0.151**	**0.371**

Table 1: Evaluation results of different generation models. We highlight in **bold** the best performance except NQG$_{true}$ in each column.

Experimental Results

We conduct some metrics for evaluation: BLEU n (Papineni et al. 2002) and ROUGE$_L$. Table 1 shows the results obtained by different models on BoolQ. Specifically, since the phrases most relevant to ground truth are used as additional information, NQG$_{true}$ is the upper bound of this task. We can find that the performance of our model is very close to the performance of the upper bound, which proves the effectiveness of our proposed two phases extraction method.

Conclusion

In this paper, we introduce the task of yes/no question generation and propose a double phases generation network (DPGNet) which achieves promising performance on BoolQ dataset. DPGNet divides the generation task into two phases prediction, which utilizes an attention network and a gate mechanism to generate a question that is more related to a specific fact. The experimental results show the effectiveness of our proposed framework.

Acknowledgements

This work was supported by the Fundamental Research Funds for the Central Universities, SCUT (No.D2182480), the National Key Research and Development Program of China, the Science and Technology Programs of Guangzhou (No.201802010027, 201902010046), National Natural Science Foundation of China (62076100).

References

Clark, C.; Lee, K.; Chang, M.; Kwiatkowski, T.; Collins, M.; and Toutanova, K. 2019. BoolQ: Exploring the Surprising Difficulty of Natural Yes/No Questions. In *NAACL-HLT*, 2924–2936.

Papineni, K.; Roukos, S.; Ward, T.; and Zhu, W. 2002. Bleu: a Method for Automatic Evaluation of Machine Translation. In *ACL*, 311–318.

Seo, M. J.; Kembhavi, A.; Farhadi, A.; and Hajishirzi, H. 2017. Bidirectional Attention Flow for Machine Comprehension. In *ICLR*.

Wang, S.; Wei, Z.; Fan, Z.; Liu, Y.; and Huang, X. 2019. A Multi-Agent Communication Framework for Question-Worthy Phrase Extraction and Question Generation. In *AAAI*, 7168–7175.