Let Me Know What to Ask: Interrogative-Word-Aware Question Generation

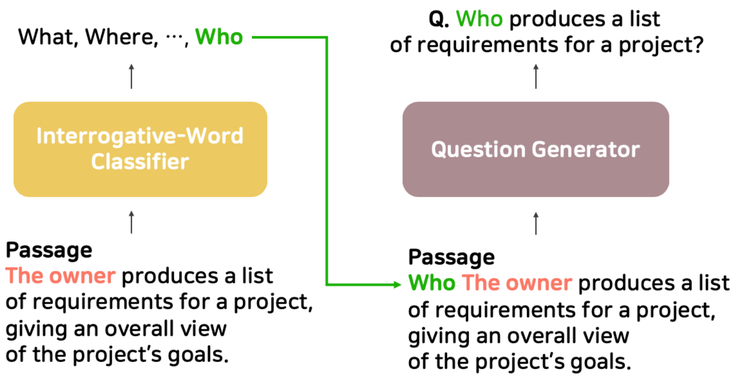
Full paper

Junmo Kang*, Haritz Puerto San Roman* and Sung-Hyon Myaeng

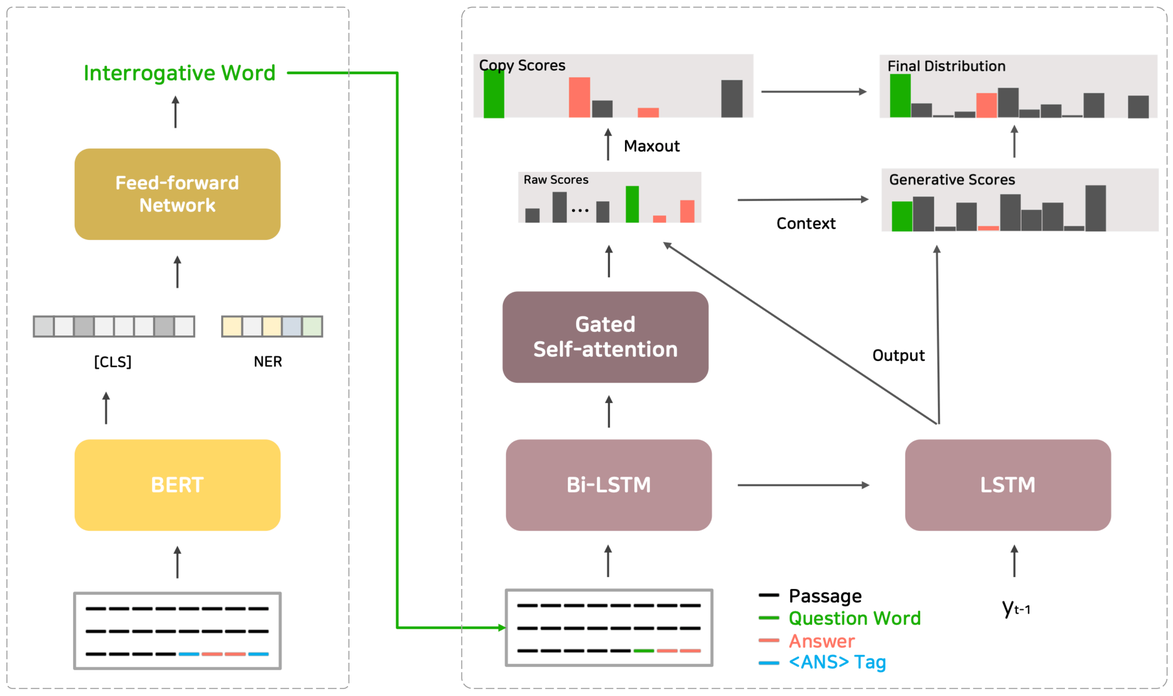
{junmo.kang, haritzpuerto94, myaeng} @ kaist.ac.kr

Introduction

- Two important aspects for **Question Generation**:
  - Interrogative word (i.e., wh-word)
  - Vocabulary & grammar
- The **interrogative word** is a **key component** in a question
- Previous models learn to generate both interrogative word and the rest of the question **simultaneously**
- We propose a method that **separates the two tasks to generate more accurate interrogative words**



Proposed Approach



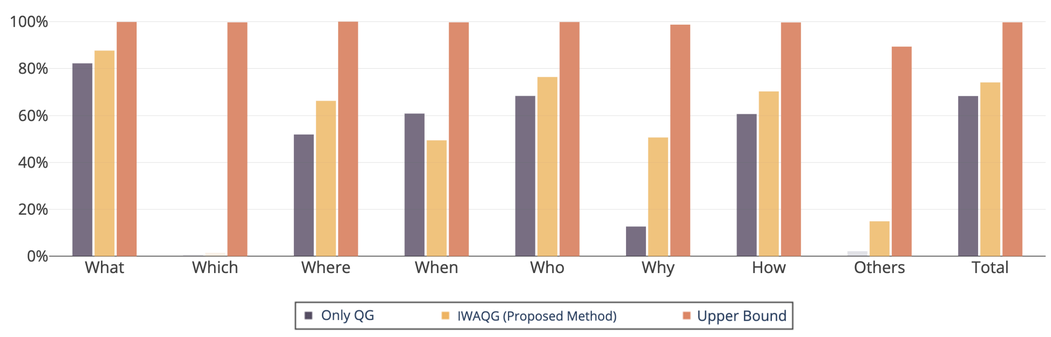
Experimental Results

Experimental Settings

- **Dataset**: SQuAD v1.1
  - In the same way as the baselines, the dev set is split randomly into dev and test set with ratio 50%-50%
- **Metrics**: essentially compute the n-gram similarity between the generated question and the reference question
  - **BLEU**: precision-based evaluation
  - **METEOR**: precision & recall-based evaluation
  - **ROUGE**: recall-based evaluation

Recall of Interrogative Words

- To prove the pipelined approach can predict better interrogative words



Comparison with Baselines

- To demonstrate an independent interrogative-word classifier leads to a better performance

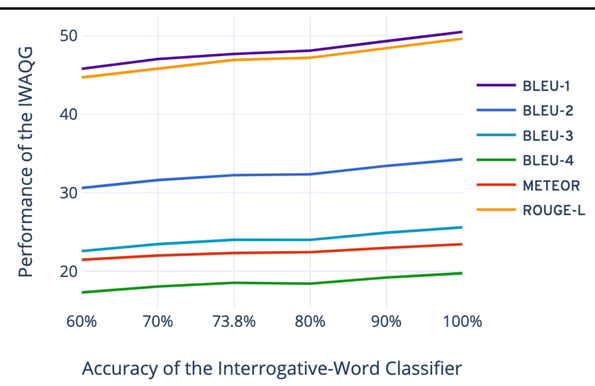
Model	BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR	ROUGE-L
Zhou et al. (2017)	-	-	-	13.29	-	-
Zhao et al. (2018)*	45.69	29.58	22.16	16.85	20.62	44.99
Kim et al. (2019)	-	-	-	16.17	-	-
Liu et al. (2019)	46.58	30.90	22.82	17.55	21.24	44.53
IWAQG	**47.69**	**32.24**	**24.01**	**18.53**	**22.33**	**46.94**

*: our QG module (Only QG)

Upper Bound Analysis

- To show the performance can be improved with better interrogative-word classifiers

Accuracy	BLEU-1	BLEU-2	BLEU-3	BLEU-4	METEOR	ROUGE-L
Only QG*	45.63	30.43	22.51	17.30	21.06	45.42
IWAQG (73.8%)	**47.69**	**32.24**	**24.01**	**18.53**	**22.33**	**46.94**
Upper Bound (100%)	**50.51**	**34.28**	**25.60**	**19.75**	**23.45**	**49.65**



Conclusion and Future Work

Conclusion

- Our approach predicts first an **interrogative word**, and then generates a **question conditioned** on the predicted interrogative word
- An **independent** Interrogative-word classifier **helps identifying the correct interrogative word** for a question
- The proposed pipeline approach **outperforms** the previous models
- **Based on our method**, other modules can be used to improve the overall performance

Future Work

- Testing our approach on other datasets to prove its **generalization capability**
- **Utilizing a QG model** to improve Question Answering systems