Template-guided Clarifying Question Generation for Web Search Clarification

Jian Wang and Wenjie Li

Department of Computing,

The Hong Kong Polytechnic University

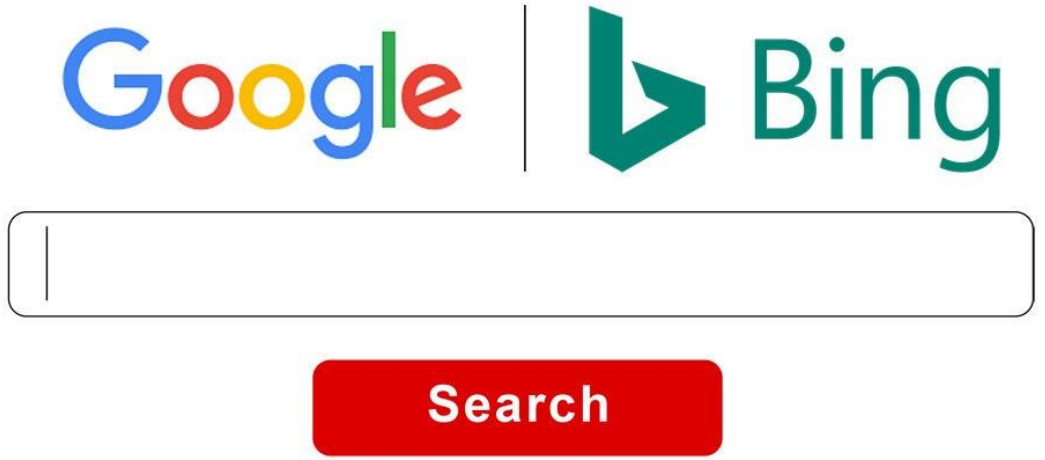
Web Search

Google | Bing

Search

Problems?

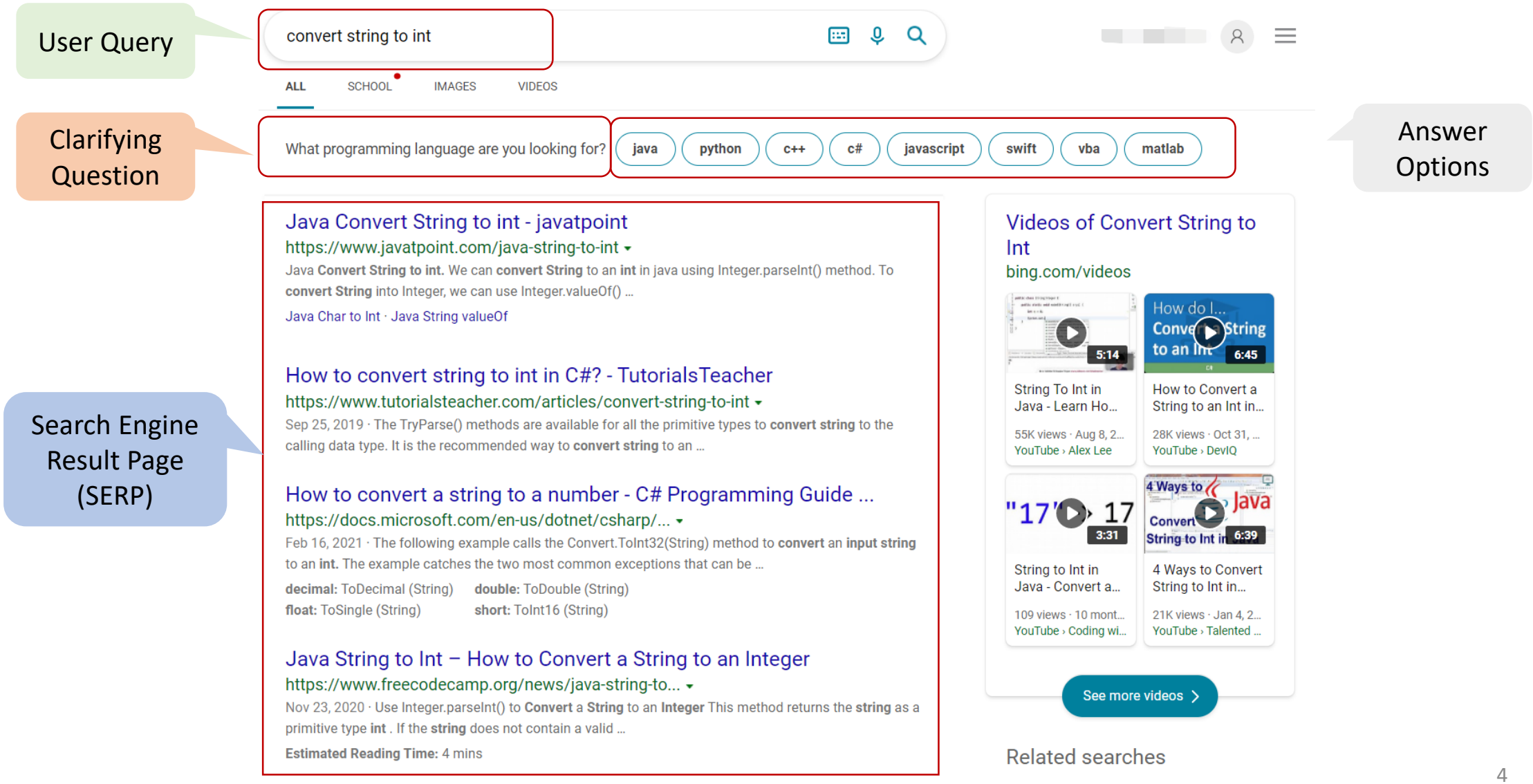
Web Search



Problems?

- Search queries are often short, the underlying user intents are often ambiguous.
- It's challenging for search engines to return the appropriate results that pertain to the users' actual information needs.

Web Search Clarification



Web Search Clarification

Problem Formulation:

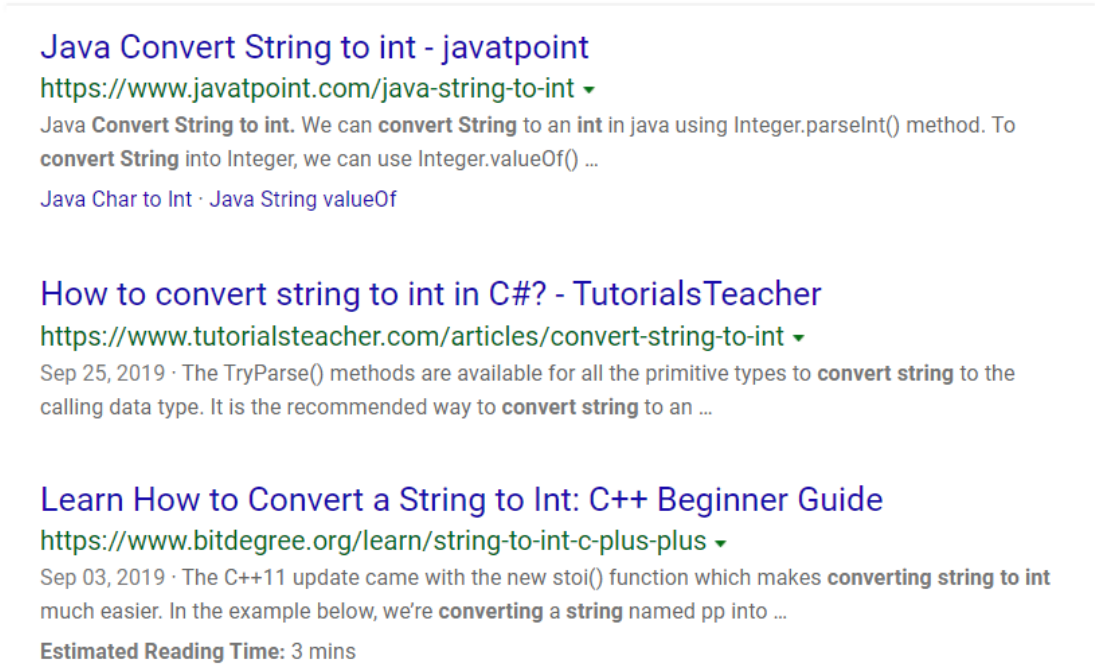
Let $Q = \{q_1, q_2, \cdots, q_n\}$ be the set of user queries, for each query $q_i (1 \leq i \leq n)$, let $S_{q_i} = \{s_{q_i}^1, s_{q_i}^2, \cdots, s_{q_i}^m\}$ denote top-m search engine result pages (SERP) in response to q_i, where the content of each $s_{q_i}^j (1 \leq j \leq m)$ is the snippet of the Web page returned by the search engine.

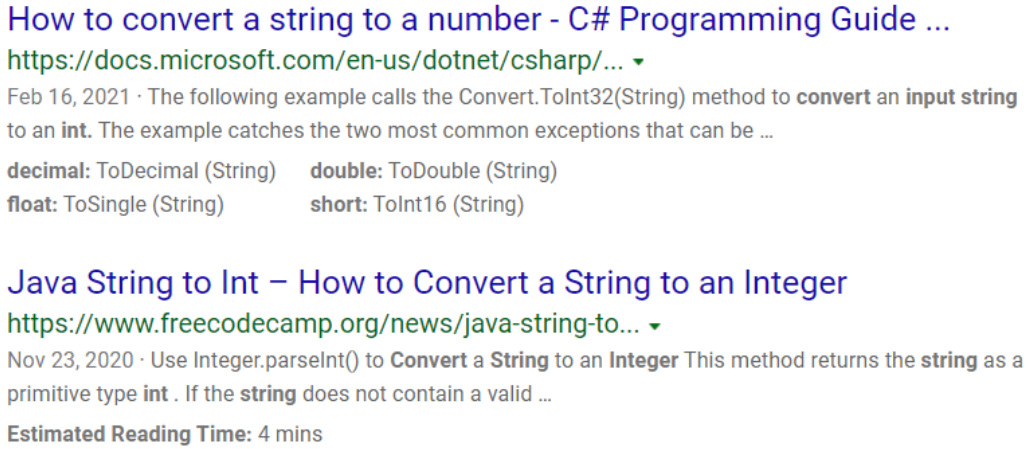
Given a user query q_i and SERP snippets S_{q_i}, the task of Web search clarification is to automatically ask a clarifying question c_i with the intention of clarifying the user's ambiguous information need.

Web Search Clarification

Challenges?

- **Generative method.** For sequence-to-sequence methods (Sutskever et al., 2014; Bahdanau et al., 2015) to generate clarifying questions directly, they can hardly well capture the intra-semantics of each SERP and the inter-patterns between different SERPs, which are crucial for what is to be clarified.





Web Search Clarification

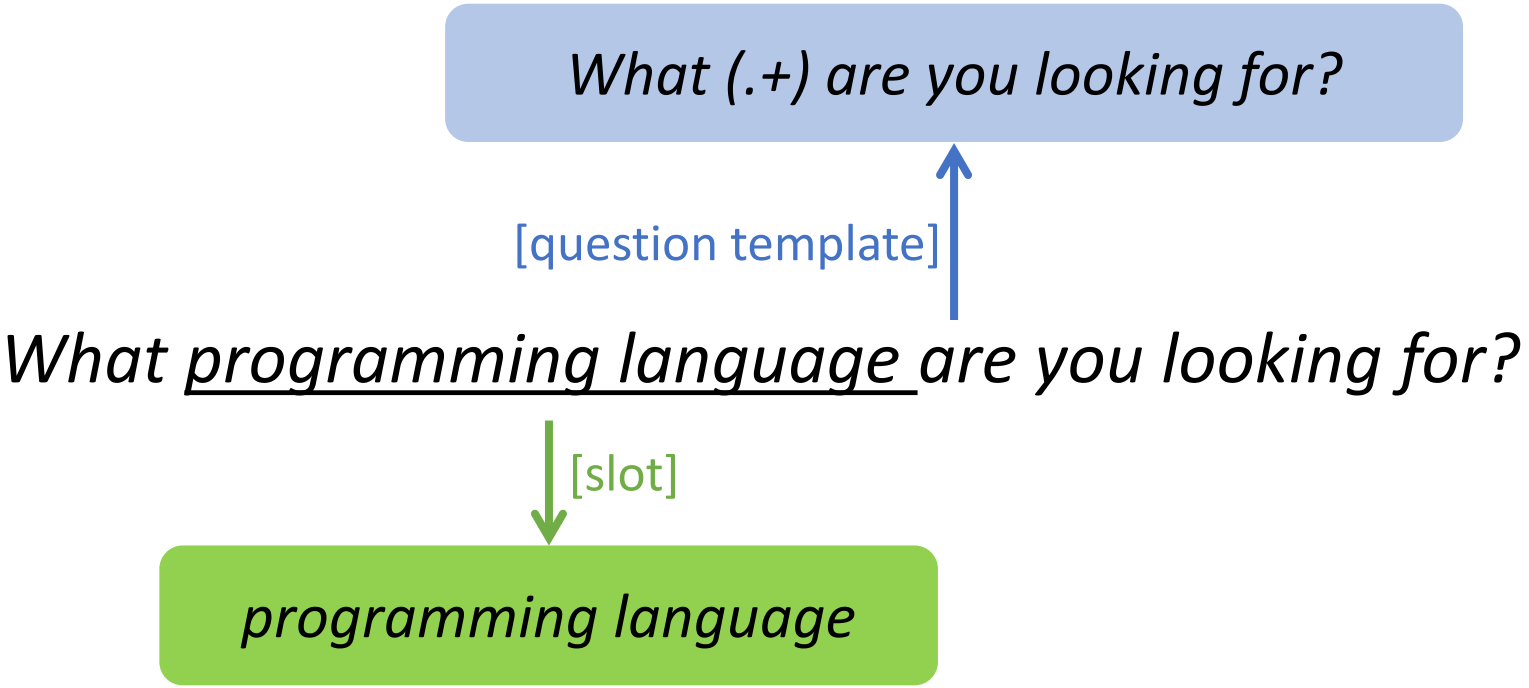
Challenges?

- **Generative method.** For sequence-to-sequence methods (Sutskever et al., 2014; Bahdanau et al., 2015) to generate clarifying questions directly, they can hardly well capture the intra-semantics of each SERP and the inter-patterns between different SERPs, which are crucial for what is to be clarified.

- **Retrieval method.** The bottleneck is to select the most appropriate one from a large pool of question candidates.

Our Motivation

- Clarifying questions often follow a few types of templates according to their purposes like disambiguation, comparison, asking for preference, or asking for sub-topic information.



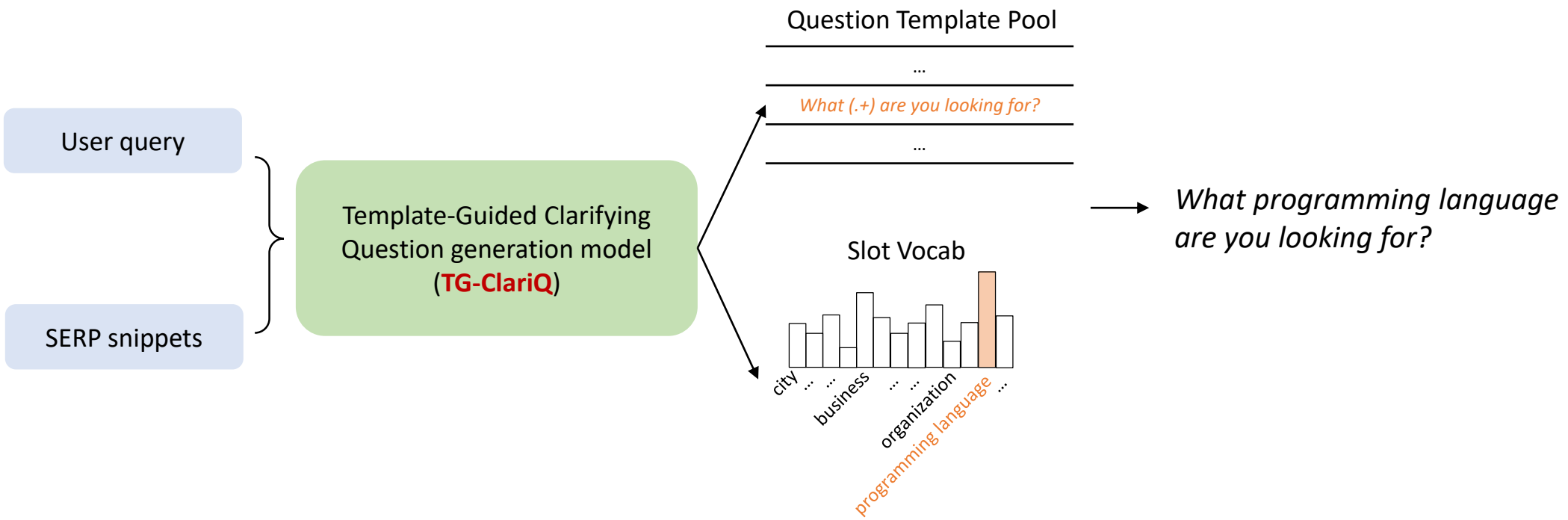
Our Motivation

- Clarifying questions often follow a few types of templates according to their purposes like disambiguation, comparison, asking for preference, or asking for sub-topic information.
- Statistical analysis on the MIMICS (Zamani et al., 2020) dataset reveals that the common question templates can actually match with over 95% of the clarifying questions.

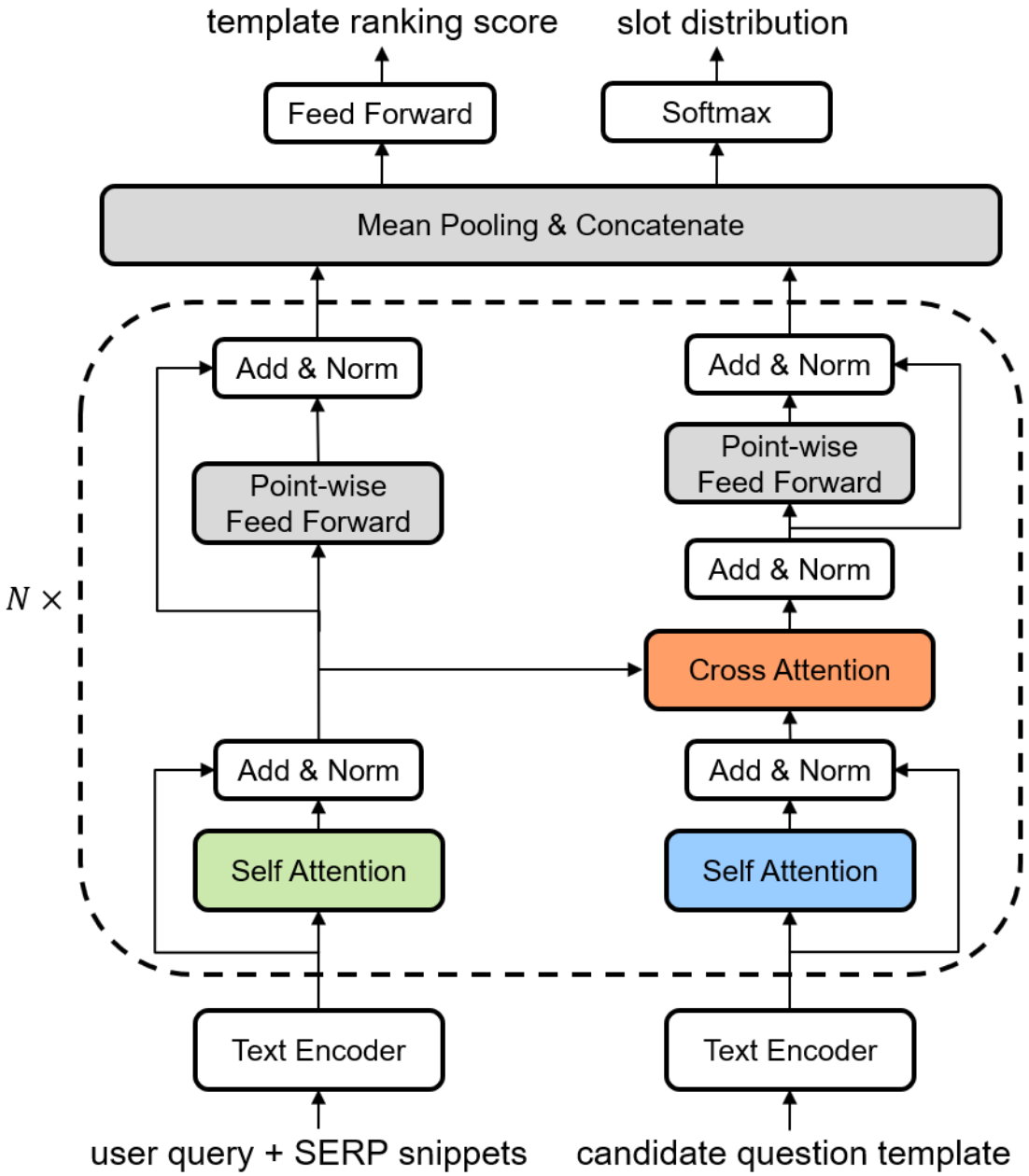
ID	Clarifying question template	#train	#dev	#test
T1	select one to refine your search	12,000	325	308
T2	what (do you want | would you like) to know about (.+)?	10,662	230	233
T3	(which | what) (.+) do you mean?	8,607	147	151
T4	(what | which) (.+) are you looking for?	4,645	130	127
T5	what (do you want | would you like) to do with (.+)?	1,988	89	105
T6	who are you shopping for?	300	40	37
T7	what are you trying to do?	227	30	29
T8	do you have any (specific | particular) (.+) in mind?	79	9	10

Our Method

- A simple yet effective template-guided clarifying question generation model, which employs Transformer (Vaswani et al., 2017) networks to enable deep interactions between user queries and SERP contents.
- Jointly learning to select the question template from a list of template candidates and fill in the question slot from a slot vocabulary.



Our Method


template ranking score
slot distribution
Feed Forward
Softmax
Mean Pooling & Concatenate
$N\times$
Add & Norm
Point-wise Feed Forward
Add & Norm
Self Attention
Add & Norm
Cross Attention
Add & Norm
Point-wise Feed Forward
Add & Norm
Self Attention
Text Encoder
Text Encoder
user query + SERP snippets
candidate question template

Datasets

We use the MIMICS (Zamani et al., 2020) data collection.

- We extract <query, clarifying question> pairs, each pair is associated with at most top-10 SERP snippets returned by the Bing's search API.
- Training/validation/testing: 38,508/1000/1000 samples
- We obtain 8 question templates in total, which cover all clarifying questions in the samples.

ID	Clarifying question template	#train	#dev	#test
T1	select one to refine your search	12,000	325	308
T2	what (do you want | would you like) to know about (.+)?	10,662	230	233
T3	(which | what) (.+) do you mean?	8,607	147	151
T4	(what | which) (.+) are you looking for?	4,645	130	127
T5	what (do you want | would you like) to do with (.+)?	1,988	89	105
T6	who are you shopping for?	300	40	37
T7	what are you trying to do?	227	30	29
T8	do you have any (specific | particular) (.+) in mind?	79	9	10

Baseline Methods

- Clarifying question selection (CQS):

  Extract clarifying questions from a large pool of candidate questions

- Clarifying template selection (CTS):

  Directly select clarifying question templates

- Clarifying question generation (CQG):

  Generate clarifying questions in an end-to-end manner

Experimental Results

	Methods	Accuracy	MRR@3	BLEU	Entity F1
CQS	BM25	0.355	0.399	n.a.	0.414
	RankNet	0.308	0.384	n.a.	0.203
	LambdaMART	0.490	0.564	n.a.	0.214
	BERT	0.394	0.440	n.a.	0.356
CTS	BM25	0.095	0.191	n.a.	n.a.
	RankNet	0.323	0.455	n.a.	n.a.
	LambdaMART	0.564	0.621	n.a.	n.a.
	BERT	0.676	0.794	n.a.	n.a.
CQG	Seq2Seq-LSTM	n.a.	n.a.	45.30	0.166
	Seq2Seq-LSTM+Copy	n.a.	n.a.	52.64	0.495
	Seq2Seq-Transformer+Copy	n.a.	n.a.	55.37	0.546
	TG-ClariQ-LSTM	0.659	0.791	55.05	0.682
	TG-ClariQ-BERT	**0.722***	**0.827***	**60.49***	**0.788***

Discussions

- If there are two or more slots that need to be filled in a question template, our model can be extended by adding additional slot generation layers and designing extra strategies to determine the order of slot filling. (Due to the single slot nature of the dataset, we leave this as a direction for future investigation.)

Discussions

- If there are two or more slots that need to be filled in a question template, our model can be extended by adding additional slot generation layers and designing extra strategies to determine the order of slot filling. (Due to the single slot nature of the dataset, we leave this as a direction for future investigation.)
- Asking clarifying questions is an essential step for Web search clarification. We intend to further explore how to generate the answer options that are paired with the clarifying questions.

Discussions

- If there are two or more slots that need to be filled in a question template, our model can be extended by adding additional slot generation layers and designing extra strategies to determine the order of slot filling. (Due to the single slot nature of the dataset, we leave this as a direction for future investigation.)
- Asking clarifying questions is an essential step for Web search clarification. We intend to further explore how to generate the answer options that are paired with the clarifying questions.

Our code is available at: https://github.com/iwangjian/TG-ClariQ

References

- Ilya Sutskever, Oriol Vinyals, and Quoc V Le. 2014. Sequence to sequence learning with neural networks. In *Advances in Neural Information Processing Systems*. 3104–3112.
- Dzmitry Bahdanau, Kyunghyun Cho, and Yoshua Bengio. 2015. Neural Machine Translation by Jointly Learning to Align and Translate. In *3rd International Conference on Learning Representations, ICLR* 2015.
- Hamed Zamani, Gord Lueck, Everest Chen, Rodolfo Quispe, Flint Luu, and Nick Craswell. 2020. Mimics: A Large-scale Data Collection for Search Clarification. In *Proceedings of the 29th ACM International Conference on Information & Knowledge Management (CIKM).* 3189–3196.
- Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. 2017. Attention is All you Need. In *Advances in Neural Information Processing Systems 30: Annual Conference on Neural Information Processing Systems (NeurIPS)* 2017. 5998–6008.

Thank you!