Template-guided Clarifying Question Generation for Web Search Clarification

Jian Wang and Wenjie Li

Department of Computing,

The Hong Kong Polytechnic University
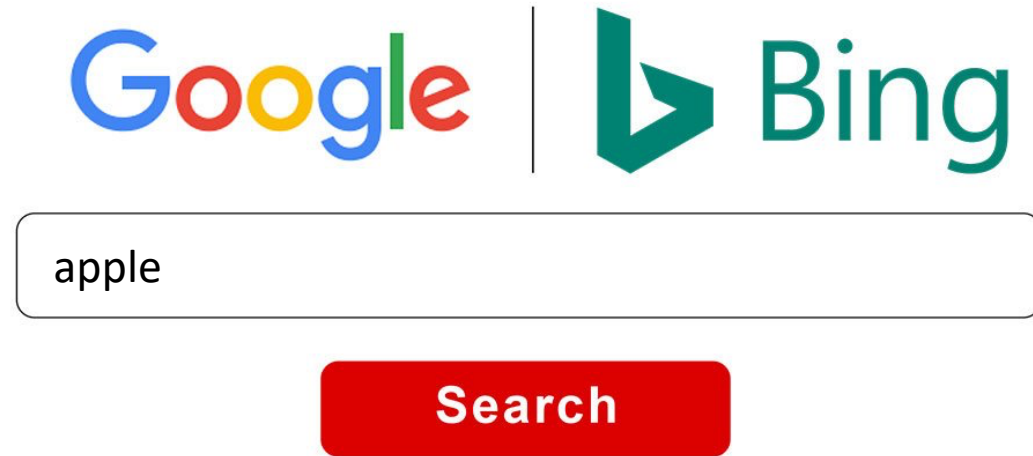
Web Search

apple

Search

(Adapted from https://www.seofuelmarketing.com/basic-differences-google-bing/)

Problems?

Web Search

Problems?

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

Web Search Clarification

User Query

convert string to int

ALL SCHOOL IMAGES VIDEOS

Clarifying Question

What programming language are you looking for?

java · python · c++ · c# · javascript · swift · vba · matlab

Answer Options

Search Engine Result Page (SERP)

Java Convert String to int - javatpoint
https://www.javatpoint.com/java-string-to-int ▾
Java **Convert String to int**. We can **convert String** to an **int** in java using Integer.parseInt() method. To **convert String** into Integer, we can use Integer.valueOf() ...
Java Char to Int · Java String valueOf

How to convert string to int in C#? - TutorialsTeacher
https://www.tutorialsteacher.com/articles/convert-string-to-int ▾
Sep 25, 2019 · The TryParse() methods are available for all the primitive types to **convert string** to the calling data type. It is the recommended way to **convert string** to an ...

How to convert a string to a number - C# Programming Guide ...
https://docs.microsoft.com/en-us/dotnet/csharp/... ▾
Feb 16, 2021 · The following example calls the Convert.ToInt32(String) method to **convert** an **input string** to an **int**. The example catches the two most common exceptions that can be ...
decimal: ToDecimal (String)   **double:** ToDouble (String)
float: ToSingle (String)   **short:** ToInt16 (String)

Java String to Int – How to Convert a String to an Integer
https://www.freecodecamp.org/news/java-string-to-... ▾
Nov 23, 2020 · Use Integer.parseInt() to **Convert** a **String** to an **Integer** This method returns the **string** as a primitive type **int**. If the **string** does not contain a valid ...
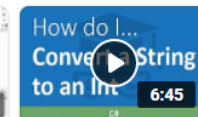**Estimated Reading Time:** 4 mins

Videos of Convert String to Int
bing.com/videos

5:14 String To Int in Java - Learn Ho... 55K views · Aug 8, 2... YouTube › Alex Lee	6:45 How to Convert a String to an Int in... 28K views · Oct 31, ... YouTube › DevIQ
3:31 String to Int in Java - Convert a... 109 views · 10 mont... YouTube › Coding wi...	6:39 4 Ways to Convert String to Int in... 21K views · Jan 4, 2... YouTube › Talented ...

See more videos ›

Related searches

4

Web Search Clarification

Problem Formulation:

Let $Q = \{q_1, q_2, \cdots, q_n\}$ be the set of user queries, for each query $q_i (1 \le i \le n)$, 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 $s_{q_i}^j (1 \le j \le m)$ 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 clarifying question from a large pool of question candidates with high efficiency.

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.

What (.+) are you looking for?

[question template]

What <u>programming language </u>are you looking for?

[slot]

programming language
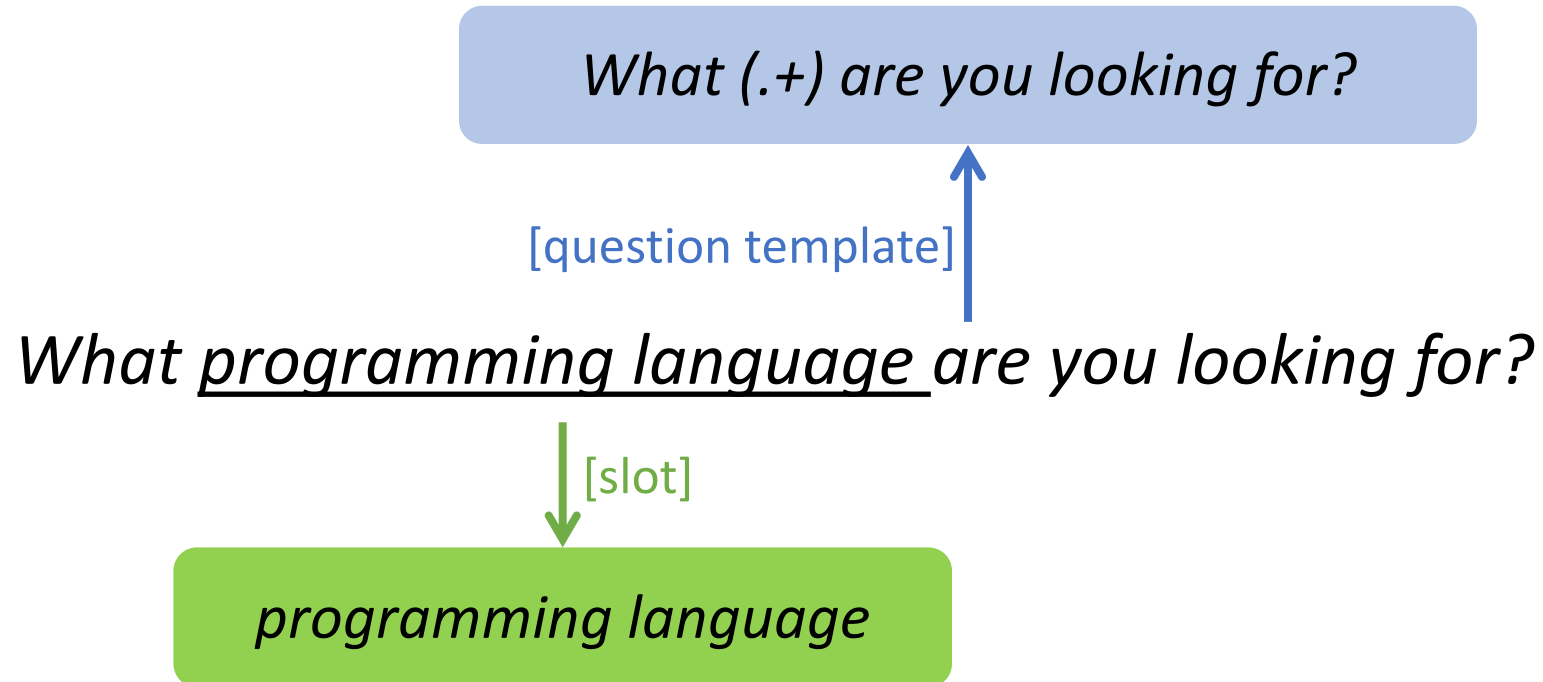
Our Motivation

- 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.

- Preliminary statistical analysis reveals that common question templates can match over 95% of the clarifying questions.

Our Method

- A simple yet effective template-guided clarifying question (TG-ClariQ) generation model, which employs Transformer (Vaswani et al., 2017) 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 — Add & Norm — Point-wise Feed Forward — Cross Attention — Self Attention — Text Encoder — user query + SERP snippets / candidate question template

11
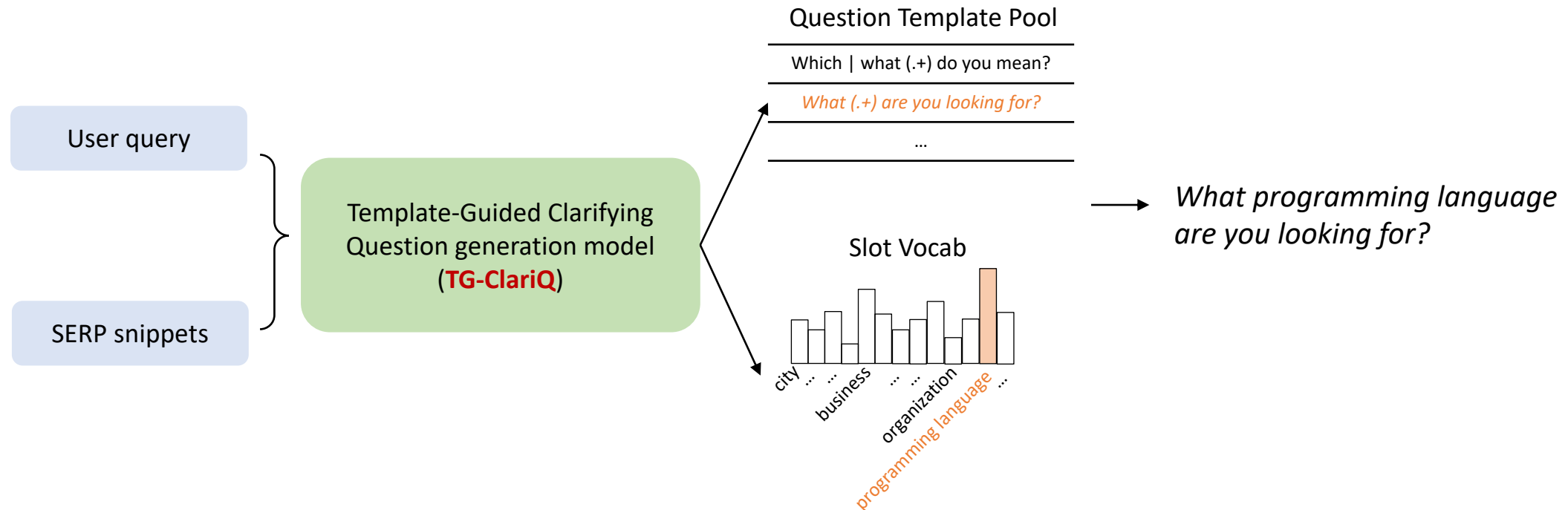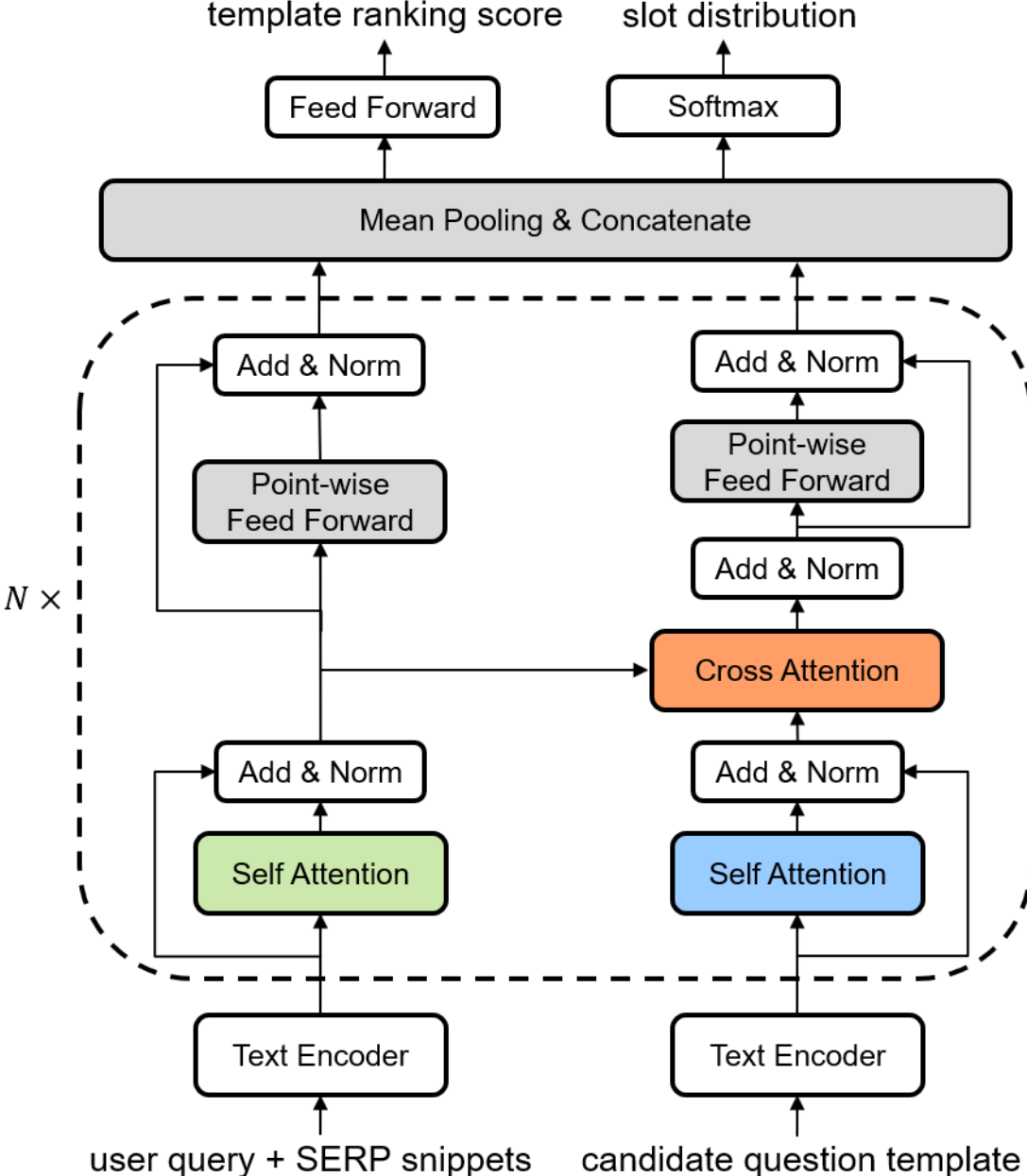
Datasets

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.

Conclusion

- We explore an interesting but under-explored task, which aims to automatically ask clarifying questions with the intention of clarifying the user's ambiguous information needs in web search scenarios.

- We propose a simple yet effective model to solve potential challenges of this task, with the main idea of jointly learning to select the question template and fill in the question slot.

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!

Q & A