

Estimating the Usefulness of Clarifying Questions and Answers for Conversational Search

Ivan Sekulić *
 Weronika Łajewska ^
 Krisztian Balog ^
 Fabio Crestani *



* Università della Svizzera italiana, Lugano, Switzerland, {name.surname}@usi.ch

^ University of Stavanger, Stavanger, Norway, {name.surname}@uis.ch

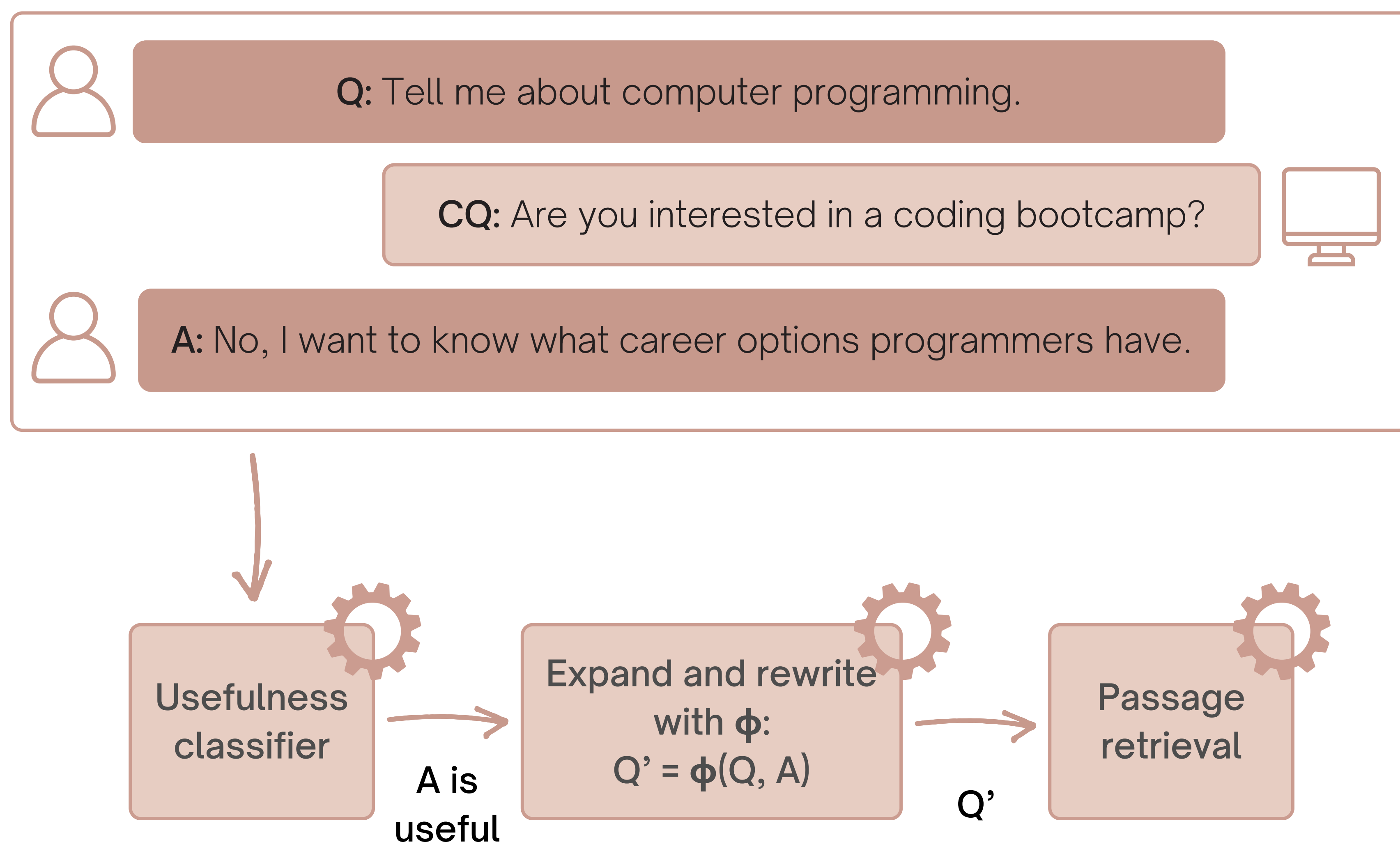
BACKGROUND

- Large body of work on constructing clarification questions (CQs), but **scarce work on processing and comprehending user's answers** (As) to such questions
- Not all CQs and given answers are useful for downstream retrieval
- **Useful CQs and As contain information addressing user's information need**

CONTRIBUTIONS

- Classifier-based answer comprehension method
- Filtering-based clarifying question selection

METHODOLOGY



- Pairs of CQs and As from ClariQ [1] annotated for their usefulness
- Classifier (MI-Clf) trained to predict the usefulness
- Given an initial query Q, CQ, and A, predict usefulness of CQ and A:
 - Affirmative A → append CQ to Q
 - Expression of information need in A → append A to Q
 - Affirmative A and additional info on information need → append CQ and A to Q
 - None of the above → do not expand Q
- Expanded Q rewritten with T5 and further used by the retrieval pipeline

Initial query Q	Clarifying question CQ	Answer A	Useful	Prevalence
I'm looking for information on hobby stores.	Do you want to know hours of operation?	No.	None	32%
Tell me information about computer programming.	Are you interested in a coding camp?	No, I want to know what career options programmers have.	Answer	53%
Find me map of USA.	Do you want to see a map of US territories?	Yes.	Question	11%
All men are created equal.	Would you like to know more about the declaration of independence?	Yes, I'd like to know who wrote it.	Question and Answer	6%

RESULTS

Approach/RunID	R@1000	MAP	MRR	NDCG	NDCG@3	NDCG@5
BM25_T5_automatic	0.3244	0.1498	0.5272	0.2987	0.3619	0.3443
BM25_T5_manual	0.4651	0.2309	0.7155	0.4228	0.5031	0.4831
our_baseline (DuoT5)	0.3846	0.1680	0.4990	0.3392	0.3593	0.3502
+ MI-All	0.4441	0.1741	0.5297	0.3594	0.3722	0.3508
MI-Clf	0.4302	0.1776	0.5144	0.3613	0.3697	0.3581

- Significant improvements over non-MI baselines
- Similar performance to the naïve method that always expands Q with both CQ and A → however, MI-Clf uses significantly less information

• **Predicting the usefulness of clarifying questions and corresponding answers can help improve retrieval effectiveness.**

[1] Aliannejadi, M., Kiseleva, J., Chuklin, A., Dalton, J., Burtsev, M.: Building and evaluating open-domain dialogue corpora with clarifying questions. In: Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing. pp. 4473–4484. EMNLP'20 (2021)