

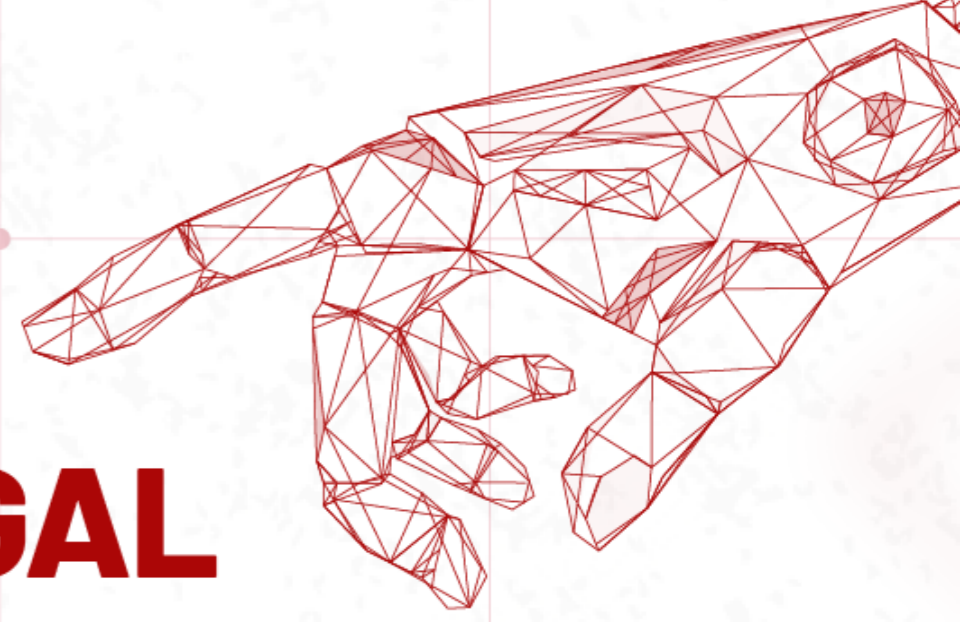
Question Rewriting? Assessing Its Importance for Conversation Question Answering

Gonçalo Raposo

INESC-ID, Instituto Superior Técnico,
Universidade de Lisboa

CMU PORTUGAL
SUMMIT 2022

NEW FRONTIERS IN TECH



Introduction

Conversational question answering systems must correctly interpret interconnected interactions and generate knowledgeable answers, which may require the retrieval of external information. This work presents our conversational question answering system, leveraging neural language models, and reports a detailed analysis on various inputs for the question rewriting module. The results obtained emphasize the importance of the conversation context representation for the overall system performance.

Methods

- Conversational question rewriting using a T5 model fine-tuned on the CANARD dataset.
- Retrieve the top-10 most relevant passages using the BM25 ranking function.
- Generate the answer using a fine-tuned Pegasus model. Answer can be seen as a summary of the retrieved passages, conditioned also on conversation data.
- Trained and tested using QReCC dataset.

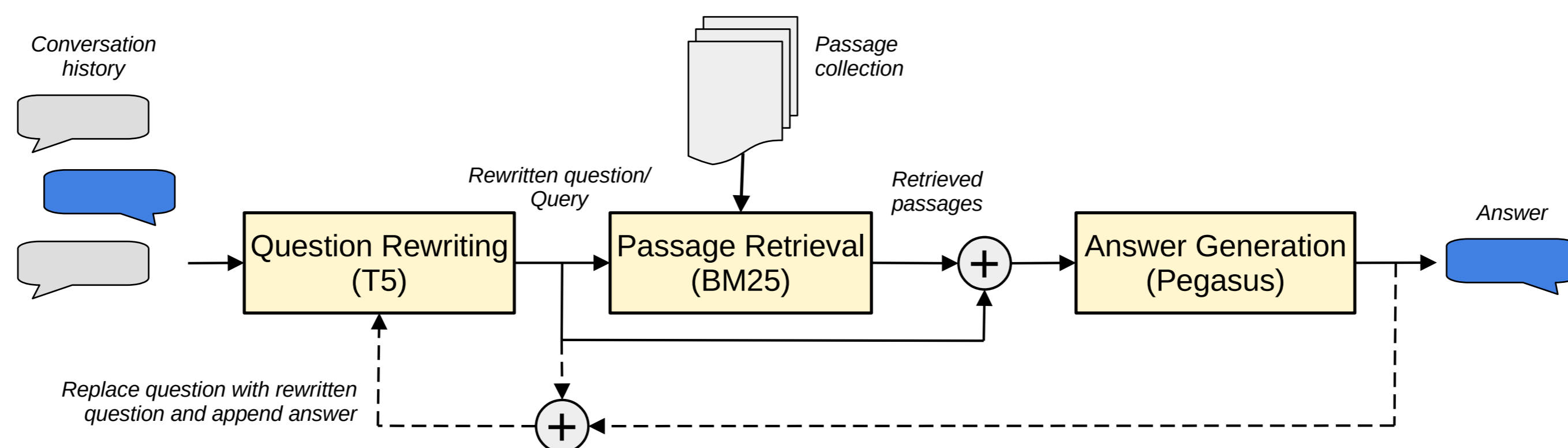


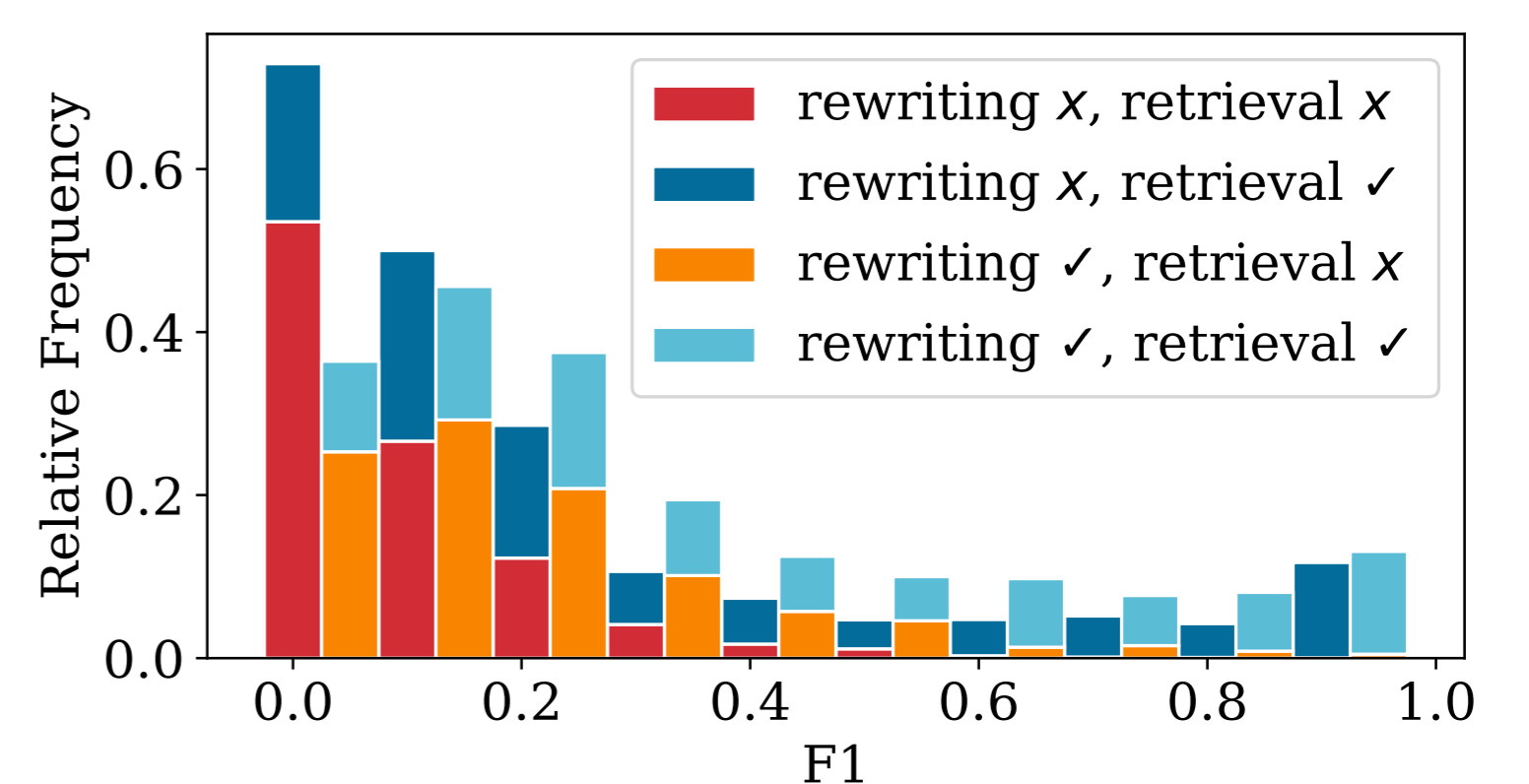
Fig. Proposed conversational question answering system.

Results

Tab. Evaluation of multiple variations of the input used in the question rewriting module.

Description	Rewriting Input	Rewriting			Retrieval		Answer	
		ROUGE1-R	MRR	F1	EM	ROUGEL-F1		
SCAI baseline: question	-	-	-	0.117	0.000	0.116		
SCAI baseline: retrieved	-	(0.571)	0.065	0.067	0.001	0.073		
SCAI baseline: GPT-3	-	-	-	0.149	0.001	0.152		
No rewriting ($h = 1$)	-	(0.571)	0.061	0.136	0.005	0.143		
No rewriting ($h = 7$)	-	(0.571)	0.145	0.155	0.003	0.160		
Questions	(Q) + Q	0.673	0.158	0.179	0.011	0.181		
Questions + answers	(Q + MA) + Q	0.681	0.150	0.179	0.010	0.181		
Rewritten questions	(MR) + Q	0.676	0.157	0.187	0.010	0.188		
Rewritten + answers	(MR + MA) + Q	0.685	0.149	0.189	0.010	0.191		
Ground truth rewritten	-	(1)	0.385	0.302	0.028	0.293		

Fig. Analysis of the influence of question rewriting on answer generation performance.



Conclusions

- 1st place in SCAI shared task 2021.
- Our system surpasses a GPT-3 baseline by 26.8% in answer accuracy.
- Question rewriting improved overall answer accuracy by about 20~30%.
- When the question rewriting succeeded, both the retrieval and answer generation improved – higher scores were about 2x more frequent.

E-mail: goncalo.cascalho.raposo@tecnico.ulisboa.pt

GitHub: <https://github.com/gonced8/rachael-scai>

Acknowledgements

Work supported by national funds through Fundação para a Ciência e a Tecnologia (FCT), under project UIDB/50021/2020; by FEDER, Programa Operacional Regional de Lisboa, Agência Nacional de Inovação (ANI), and CMU Portugal, under project Ref. 045909 (MAIA) and research grant BI|2020/090; and by European Union funds (Multi3Generation COST Action CA18231).