Modeling Mix-initiative Conversational Search

Zhenduo Wang
Kahlert School of Computing, Unverisity of Utah
zhenduow@cs.utah.edu

Overview

Search is ubiquitous and in increasing demand in this age. Conventionally, search is cast as single-turn retrieval: A search system ranks all the pages on the internet according to their relevance to the query and returns the rank list as "ten blue links" [1]. A study [2] revealed that about 60% of the queries are less than two words, and 98.2% of the queries are less than seven words. Short queries are especially challenging for single-turn search systems due to their ambiguity.

Existing commercial search engines attempted to address query ambiguity through search result page (SERP) diversification. However, it is passive and inefficient as the diversified SERP reduces the portion of relevant contents to each user. Further, this approach may not be practical to search scenarios like mobile search or voice search with limited bandwidth for showing results [3].

With the development in machine learning, natural language processing, and computing power, it has become much easier to process and understand complex user statements. This helped develop highly interactive information-seeking systems that support multi-turn user-system interactions beyond user-initiative single-turn systems. They are known as conversational search systems and is becoming an increasingly popular research topic and important frontier of IR [4, 5].

Although existing work has demonstrated benefits of introducing system-initiative interactions in conversational search such as asking clarifying questions, fully mix-initiative conversational search is still a dynamic and complex process that lacks systematic research.

My research focuses on modeling mix-initiative conversational search, in particular implementing a fully mix-initiative conversational search system equipped with a conversational search policy trained and working cooperatively with its core result retrieval and generation models.

To this end, I designed a conversational search policy model that can choose mix-initiative conversational search actions, simulate, and model the risk of each action [6, 7], and can infer the reasoning behind these decisions from historical conversational search logs and improve generalizability [8]. Furthermore, I designed a zero-shot inferencing pipeline to generate high quality clarifying questions based on query facets [9]. As a new search paradigm, there has not been any mature online service for conversational search, which significantly restrain the availability of large conversational search datasets. To facilitate the automatic creation of large conversational search dataset, I developed a user simulation system [10].

How To Train A Conversational Search Action Decision-Making Model (Policy)?

In conversational search, an alternative approach to directly retrieving result while facing an ambiguous query is to proactively ask clarifying questions (CQ) to the user. This approach has drawn attention from both NLP and IR communities in the past. Most of these attentions are focused on generating and selecting CQ, or incorporating CQ with the query. However, their work assumed that users were dedicated to the search session after submitting the search query, and would give an answer to any CQ asked. I argued that asking CQ upon ambiguous query is also a risky decision [7, 6], in that the CQ can be irrelevant, over-specific, etc., making the user uncomfortable (like racial innuendo [11]). Hence asking CQ should not always be taken as an alternative to direct result retrieval. Therefore, a mix-initative conversational search system needs a decision-making model for its actions. There are a few work for addressing this need [12, 13, 14], but these work cast the decision-making problem as single-turn classification. My work [6, 7] was the first to cast the problem as conversational search policy learning and explored the sequential decision-making configuration of the problem. I showed that reinforcement learning (Q-learning) could efficiently learn the conversational search policy.

Figure 1 is my risk-aware conversational search system structure with conversational search policy. Unlike previous systems, it first calls the clarifying question and result retrieval models independently to obtain questions and potential results as action candidates. After that, the conversational search policy (risk-decision module) jointly evaluates the query, context, and the candidates to decide between the clarifying question or the answer. Compared to previous systems, its advantage is that it preemptively evaluates the candidates and reduces the chance of asking bad clarifying questions to the users. Training this policy model was challenging since no annotated conversational search logs were available. To solve this problem, I employed Deep Q-Learning [15] and empirically designed rewards for the Q-learning algorithm.

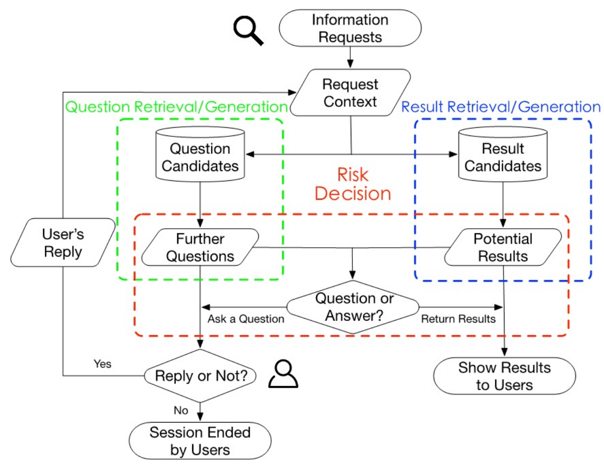

Figure 1: Risk-aware Conversational Search

Through experiments with different types of user simulators, I showed that my system outperformed several deterministic decision baselines on three datasets including MSDialog, Ubuntu Dialog Corpus, and OpendialKG [16, 17, 18]. In conclusion, my proposed risk-aware conversational search system can control the risk in conversational search, boost conversational search result quality, and improve user experience.

How To Generalize Policy Training Without Making User Assumption?

Conversational search systems should be adaptive to a wide range of user types. However, existing conversational search systems are mostly developed with explicit user assumptions. Particularly, their policy models are trained with assumptions about users' patience, tolerance, cooperativeness, etc.[7, 6, 19], which are often instantiated as action rewards in the reinforcement learning process. When tested with different types of users, these systems often fail to generalize. Thus, in real-world applications where user behaviors are unpredictable and varying, this could cause user loss.

My work [8] demonstrated a simple imitation learning (IL) method to train conversational policy without making explicit user assumptions and tuning rewards based on user assumptions. It automatically inferred the reward function from historical conversational search logs containing implicit user behavioral information. The reward-free IL algorithm both reduces the tuning cost and improves generalizability.

The IL algorithm I use is the generative adversarial imitation learning (GAIL) algorithm [20], which alternatively trains a discriminator and a policy model. It shares structural similarities with generative adversarial net (GANs) [21]. The goal of the discriminator is to distinguish model-generated trajectories from the expert conversation trajectories. The goal of the policy model is to generate trajectories more like the best historical conversation and less like what it has already generated so far. With this iterative training process, the policy model is optimized to approximate the expert which generates the ground truth trajectories in the dataset.

By definition, training GAIL needs to identify the best historical conversation trajectories as positive examples. To get the best trajectory, I designed a session-level conversational search evaluation metric named Expected Conversational Reciprocal Rank (ECRR), which balances between the search result quality and search efficiency in its scoring function. I then use this ECRR metric to compute and rank trajectories to find the expert trajectory for each conversation. Through experiments, I showed that GAIL can efficiently learn the expert policy without making user assumptions and generalize significantly better to unseen users compared to previous baselines.

How To Generate High Quality Clarifying Questions When Facing Data Scarcity?

Existing research on generating clarifying questions for conversational search exhibit two limitations about the datasets. The first is the lack of sizable conversational search datasets with ambiguous faceted queries. The second is that requiring a conversation dataset to cover all possible search topics unbiasedly is unrealistic. Therefore, a more practical solution is to generate clarifying questions in a zero-shot setting, which means that the generation system does not need to be trained on any conversational search data. However, directly applying existing zero-shot language generation methods often yield unsatisfactory results because of two reasons: First, they tend to generate superficially relevant contents that do not actually help clarify the search need; such as repeating the original search query. Second, they usually generate narratives instead of asking questions.

In my work [9], I designed a constrained language generation pipeline (Fig. 2) to address these challenges. First, to ensure the generation is useful for clarification, I leverage the idea of constrained language generation, with search facets as the constraints. The constrained language generation algorithm, NeuroLogic Decoding [22], modifies the beam search process and incentivizes the generation model to rank generations beams which contain the facet words higher during beam search. Next, to avoid generating narratives, I give the generation model eight clarifying question prompts to guide the generation process. Then, I use a weighted sequential dependency model [23] to rank the eight prompted generations and return the top question.

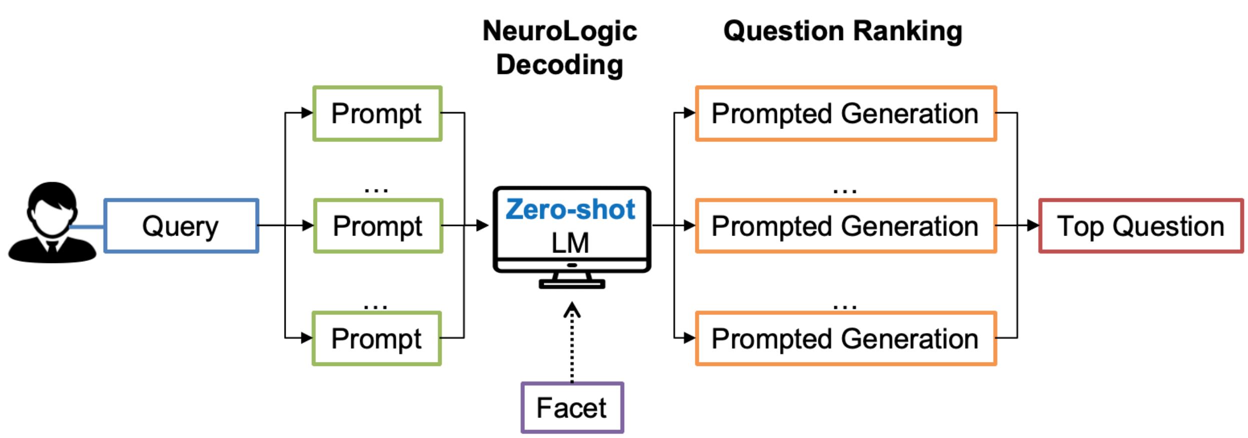

Figure 2: Zero-shot Clarifying Question Generation System

Experiment results evaluated by both automatic language generation metrics and human judge on Clariq-FKw dataset show that my proposed zero-shot system significantly outperforms existing models, even if they are finetuned on the training set and my system is not.

How To Simulate User Responses For Automatic Multi-Turn Dataset Creation?

Most research about conversational search [24, 13, 7, 25, 26, 27, 28] is limited to training their system on datasets with observed or artificial conversation logs. Such a dataset would lack training signal and evaluation references when a conversation veers away from the dataset; especially, when the system generates a question not listed in the dataset and steers the conversation in an unseen direction. Eventually, conversational search systems should be trained, evaluated, and deployed in an open-ended setting. However, training and evaluating them this way is challenging; it requires humans to generate responses to open clarifying questions, which is expensive and unscalable.

Past work [19, 25] has demonstrated that a user response simulator that automatically generates human responses can help evaluate conversational search systems. The goal of a such a system is to generate user-like answers to system-generated clarifying questions based on a query and the user's search intent. I propose that a user response simulation system can also enable studies such as training a multi-turn conversational search system, by generating synthetic conversations and rewards, and perhaps using Reinforcement Learning from Human Feedback (RLHF) [29].

My work [10] showed that current state-of-the-art user simulation system could be significantly improved by replacing it with a smaller but advanced T5 [30] natural language generation model. Further, I present an in-depth investigation of the task of simulating user response for conversational search to supplement existing works with an insightful hand-analysis of what challenges are still unsolved by the advanced model, as well as to propose our solutions for them. The major challenges I identified include (1) dataset noise, (2) existing models struggle to generate the correct answer type, and (3) the standard evaluation will misevaluate generated responses because of user cooperativeness mismatch. Except for the dataset noise issue, I propose two solutions to generate responses with better type accuracy and to avoid the misevaluation. My solution to the former is to improve user simulators with knowledge from question answering tasks, using one of the current state-of-the-art models for QA—UnifiedQA [31]. Further, I propose to train a classifier that predicts answer types to guide the UnifiedQA through constrained generation, using RoBERTa [32] as the classifier, which is a representative of state-of-the-art text classifiers. My solution to the misevaluation is to partition the training and evaluating data, then separately train and evaluate on data with only matched cooperativeness.

I conduct experiments on two popular dataset for user response simulation—Qulac [24] and ClariQ [13]. My proposed solutions lead to further improvements over the T5 baseline.

Future Work

Today, models like GPT-4 [33] and LLaMA [34] pioneer the arrival of the era of large language models (LLMs). In the future, I propose to design a comprehensive mix-initiative conversational search system that will bridge the gap between LLMs and a search system that can perform true conversational search. In particular, I wish to explore the following directions:

1. End-To-End Training For Conversational Search System

My past work has covered various parts of a mix-initiative conversational search system, including conversational search policy, clarifying question generation, and user simulation. However, they are mostly done independently. In reality, these parts are interconnected within the system. For example, the policy model needs to compare the retrieved results and possible clarifying questions to make decisions, while the clarifying question model can benefit from being able to predict users' responses. These potentials make it appealing to train all these models jointly end-to-end.

2. From Large Language Model To Mix-Initiative Conversational Search System

Since OpenAI released ChatGPT [33], LLMs are no longer mere academic research tools but also real online applications. The multi-turn conversation capability of these LLMs is a compelling indicator that a mature online conversational search system is around the corner. However, since these LLMs are not directly designed for doing conversational search, using them as the backbone of a conversational search system raises many potential concerns. For example, **how to leverage their language generation capability in a retrieval task?** and **how to train them and evaluate their generations given that conversational search is a new search paradigm with no industry standard?** To answer these questions, one needs to iteratively test and evaluate various LLM-based conversational search systems configurations and conduct user studies.

By tackling the the above mentioned projects, I hope to contribute to the development of modern conversational search systems to improve user experiences.

References

[1] H. Zamani, J. R. Trippas, J. Dalton, and F. Radlinski, "Conversational information seeking," *arXiv preprint arXiv:2201.08808*, 2022.

[2] A. Spink, D. Wolfram, M. B. Jansen, and T. Saracevic, "Searching the web: The public and their queries," *Journal of the American society for information science and technology*, vol. 52, no. 3, pp. 226–234, 2001.

[3] J. Gao, C. Xiong, P. Bennett, and N. Craswell, *Neural approaches to conversational information retrieval.* Springer Nature, 2023, vol. 44.

[4] A. Anand, L. Cavedon, M. Hagen, H. Joho, M. Sanderson, and B. Stein, "Dagstuhl seminar 19461 on conversational search: Seminar goals and working group outcomes," *SIGIR Forum*, vol. 54, no. 1, feb 2021. [Online]. Available: https://doi.org/10.1145/3451964.3451967

[5] J. S. Culpepper, F. Diaz, and M. D. Smucker, "Research frontiers in information retrieval: Report from the third strategic workshop on information retrieval in lorne (swirl 2018)," in *ACM SIGIR Forum*, vol. 52, no. 1. ACM New York, NY, USA, 2018, pp. 34–90.

[6] Z. Wang and Q. Ai, "Controlling the risk of conversational search via reinforcement learning," in *Proceedings of the Web Conference 2021*, 2021, pp. 1968–1977.

[7] ——, "Simulating and modeling the risk of conversational search," *ACM Trans. Inf. Syst.*, vol. 40, no. 4, mar 2022. [Online]. Available: https://doi.org/10.1145/3507357

[8] Z. Wang, Z. Xu, and Q. Ai, "Reward-free policy imitation learning for conversational search," *arXiv preprint arXiv:2304.07988*, 2023.

[9] Z. Wang, Y. Tu, C. Rosset, N. Craswell, M. Wu, and Q. Ai, "Zero-shot clarifying question generation for conversational search," *arXiv preprint arXiv:2301.12660*, 2023.

[10] Z. Wang, Z. Xu, Q. Ai, and V. Srikumar, "An in-depth investigation of user response simulation for conversational search," *arXiv preprint arXiv:2304.07944*, 2023.

[11] RT, "Trolling tay: Microsofts new ai chatbot censored after racist and sexist tweets," https://www.rt.com/viral/337056-trolling-tay-microsoft-censored/, 2016.

[12] Y. Zhang, X. Chen, Q. Ai, L. Yang, and W. B. Croft, "Towards conversational search and recommendation: System ask, user respond," in *Proceedings of the 27th acm international conference on information and knowledge management*, 2018, pp. 177–186.

[13] M. Aliannejadi, J. Kiseleva, A. Chuklin, J. Dalton, and M. Burtsev, "Convai3: Generating clarifying questions for open-domain dialogue systems (clariq)," *arXiv preprint arXiv:2009.11352*, 2020.

[14] ——, "Building and evaluating open-domain dialogue corpora with clarifying questions," *arXiv preprint arXiv:2109.05794*, 2021.

[15] C. J. Watkins and P. Dayan, "Q-learning," *Machine learning*, vol. 8, no. 3, pp. 279–292, 1992.

[16] C. Qu, L. Yang, W. B. Croft, J. Trippas, Y. Zhang, and M. Qiu, "Analyzing and characterizing user intent in information-seeking conversations." in *SIGIR '18*, 2018.

[17] R. Lowe, N. Pow, I. V. Serban, and J. Pineau, "The ubuntu dialogue corpus: A large dataset for research in unstructured multi-turn dialogue systems," pp. 285–294, 2015.

[18] S. Moon, P. Shah, A. Kumar, and R. Subba, "OpenDialKG: Explainable conversational reasoning with attention-based walks over knowledge graphs," in *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*. Florence, Italy: Association for Computational Linguistics, Jul. 2019, pp. 845–854. [Online]. Available: https://aclanthology.org/P19-1081

[19] A. Salle, S. Malmasi, O. Rokhlenko, and E. Agichtein, "Studying the effectiveness of conversational search refinement through user simulation," in *European Conference on Information Retrieval*. Springer, 2021, pp. 587–602.

[20] J. Ho and S. Ermon, "Generative adversarial imitation learning," *Advances in neural information processing systems*, vol. 29, pp. 4565–4573, 2016.

[21] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial nets," *Advances in neural information processing systems*, vol. 27, 2014.

[22] X. Lu, P. West, R. Zellers, R. L. Bras, C. Bhagavatula, and Y. Choi, "Neurologic decoding:(un) supervised neural text generation with predicate logic constraints," *arXiv preprint arXiv:2010.12884*, 2020.

[23] M. Bendersky, D. Metzler, and W. B. Croft, "Learning concept importance using a weighted dependence model," in *Proceedings of the Third ACM International Conference on Web Search and Data Mining*, ser. WSDM '10. New York, NY, USA: Association for Computing Machinery, 2010, p. 31–40. [Online]. Available: https://doi.org/10.1145/1718487.1718492

[24] M. Aliannejadi, H. Zamani, F. Crestani, and W. B. Croft, "Asking clarifying questions in open-domain information-seeking conversations," in *Proceedings of the 42nd international acm sigir conference on research and development in information retrieval*, 2019, pp. 475–484.

[25] I. Sekulić, M. Aliannejadi, and F. Crestani, "Evaluating mixed-initiative conversational search systems via user simulation," in *Proceedings of the Fifteenth ACM International Conference on Web Search and Data Mining*, ser. WSDM '22. New York, NY, USA: Association for Computing Machinery, 2022, p. 888–896. [Online]. Available: https://doi.org/10.1145/3488560.3498440

[26] K. Bi, Q. Ai, and W. B. Croft, "Asking clarifying questions based on negative feedback in conversational search," in *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*, 2021, pp. 157–166.

[27] I. Sekulić, M. Aliannejadi, and F. Crestani, "Towards facet-driven generation of clarifying questions for conversational search," in *Proceedings of the 2021 ACM SIGIR International Conference on Theory of Information Retrieval*, 2021, pp. 167–175.

[28] H. Zamani, S. Dumais, N. Craswell, P. Bennett, and G. Lueck, "Generating clarifying questions for information retrieval," in *Proceedings of The Web Conference 2020*, 2020, pp. 418–428.

[29] L. Ouyang, J. Wu, X. Jiang, D. Almeida, C. L. Wainwright, P. Mishkin, C. Zhang, S. Agarwal, K. Slama, A. Ray *et al.*, "Training language models to follow instructions with human feedback," *arXiv preprint arXiv:2203.02155*, 2022.

[30] C. Raffel, N. Shazeer, A. Roberts, K. Lee, S. Narang, M. Matena, Y. Zhou, W. Li, and P. J. Liu, "Exploring the limits of transfer learning with a unified text-to-text transformer," *Journal of Machine Learning Research*, vol. 21, no. 140, pp. 1–67, 2020. [Online]. Available: http://jmlr.org/papers/v21/20-074.html

[31] D. Khashabi, S. Min, T. Khot, A. Sabhwaral, O. Tafjord, P. Clark, and H. Hajishirzi, "Unifiedqa: Crossing format boundaries with a single qa system," 2020.

[32] Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, and V. Stoyanov, "Roberta: A robustly optimized bert pretraining approach," *arXiv preprint arXiv:1907.11692*, 2019.

[33] OpenAI, "Gpt-4 technical report," *arXiv*, 2023.

[34] H. Touvron, T. Lavril, G. Izacard, X. Martinet, M.-A. Lachaux, T. Lacroix, B. Rozière, N. Goyal, E. Hambro, F. Azhar *et al.*, "Llama: Open and efficient foundation language models," *arXiv preprint arXiv:2302.13971*, 2023.