Reinforcement Learning Meets Information Seeking: Dynamic Search Challenge

Zhiwen Tang
Infosense, Department of Computer Science
Georgetown University
zt79@georgetown.edu

Grace Hui Yang
Infosense, Department of Computer Science
Georgetown University
huiyang@cs.georgetown.edu

ABSTRACT

We introduce the Dynamic Domain Track, which was a challenge held in Text REtrieval Conference (TREC) from 2015-2017. The challenge aims to improve search engine for interactive scenarios, where we believe reinforcement learning would play an important role. We share our design ideas of the challenge and discuss what we have learned from it. This paper has been published in the Machine Learning Competitions for All (CiML) Workshop at NeurIPS 2019.

ACM Reference Format:
Zhiwen Tang and Grace Hui Yang. 2020. Reinforcement Learning Meets Information Seeking: Dynamic Search Challenge. In *The 1st Workshop on Deep Reinforcement Learning for Information Retrieval (DRL4IR '20), July 30, 2020, Virtual Event, China.* ACM, New York, NY, USA, 2 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

1 INTRODUCTION

Search engines are the most convenient and widely used tool for online information seeking. A seeking process involves several rounds of retrieval where the user forages newly-obtained knowledge and achieves better understanding of a search topic; search engine learns user's interests and adjusts its algorithm. This information seeking process is termed as Dynamic Search by Yang et al. [6]. The search system learns user's interests by trial-and-error, aiming to satisfy user's information need in the end. The interactive and goal-oriented nature of dynamic search makes it a natural application of reinforcement learning.

Multiple evaluation protocols have been proposed in the past in the Text REtrieval Conference (TREC) to evaluate interactive search systems. An early attempt is the Interactive Track [5], where real human user and search system are evaluated as a whole. The required participation of real human users makes it labor-intensive in its experimental remits and non-reproducible. Another attempt is the Session Track [1]. It is based on search logs, where the system is expected to optimize the last round of retrieval after it reads the logs of previous rounds. It does not support the live interaction between the search engine and the user.

None of the above tracks provide live and reproducible testbeds for evaluating dynamic search systems. The Dynamic Domain Track [7] tackles those challenges by introducing a simulated user, enabling cost-effective and real-time interaction between the system and the user. Third-party assessors were first asked to find documents pertaining to the search task as complete as possible. The user simulator have full access to the annotated data and thus is capable of providing relevance judgement for any documents the system retrieves. After simple setup, researchers can implement their own algorithms and interact with the simulated user for unlimited times. The responses from the simulated user remain constant over time, which makes the experiments reproducible.

DRL4IR '20, July 30, 2020, Virtual Event, China

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM…$15.00
https://doi.org/10.1145/nnnnnnn.nnnnnnn

	Search DD17-10
User:	*Leaning Towers of Pisa Repairs*
System:	Return document 0290537
User:	Non-relevant document.
System:	Return document 0298897
User:	Relevant on subtopic 320 with a rating of 2, *"No one doubts that it will collapse one day unless preventive measures are taken."*

Table 1: Example Interaction History.

2 THE PROTOCOL

Create the ground truth: TREC DD created knowledge-intensive search topics. In 2017, search topics are created from New York Times Annotated Corpus [4], which contains 1.8 million archived newswires in 30 years. Each search topic consists of several subtopics. Third-party assessors were asked to find documents relevant to each subtopic as complete as possible. Assessors were required to highlight relevant passages and rate each of them in a scale of 1-4, where 4 means key results and 1 means marginally relevant. The relevance rating may serve as the reward.

Interact with the user simulator: The simulated user is preloaded with ground truth data. It starts the search task by issuing a query indicating the information need. After that, at each turn, the system retrieves five documents and send them to the simulated user. The simulated user then checks the ground truth and provides feedback regarding those documents. The system then retrieves another five documents and this loop continues. Feedback from the simulated user includes the relevance rating and highlighted passages from the ground truth. Figure 1 illustrates the interaction loop between the system and the simulated user. A toy interaction history is shown in Table 1 where the search topic is *Leaning Towers of Pisa Repairs* (id: DD17-10), which consists of 4 subtopics (id: 318-321).

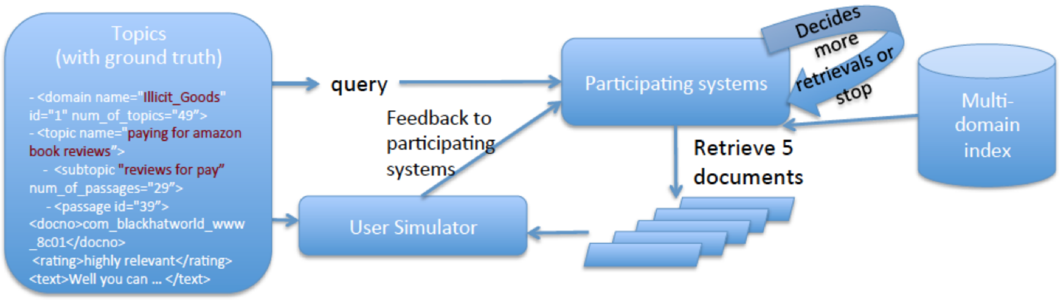

Figure 1: Evaluation Procedure of TREC Dynamic Domain Track, by [7]

Metrics: TREC DD adopts several information retrieval (IR) metrics to evaluate search systems. The IR evaluation is based on the interaction transcript logged by the simulated user. Metrics used includes Cube Test [3], sDCG [2] and Expected Utility [8]. They measure the relevant information the system found and the effort of the user in different ways.

3 LESSONS LEARNED

The evaluation procedure relies largely on the ground truth data we created. However, it is time-consuming to find a relatively complete set of documents pertaining to a search topic. The annotation took six assessors more than one month and only 60 search topics they created meet our expectation, where more than 5000 passages and 3800 documents are considered as relevant. The amount of the annotated data is smaller than we expect and we look forward to more efficient annotation procedure.

An important feature the simulated user fails to model is the learning path for each search topic. Human users cannot learn the new knowledge until prerequisites have been satisfied. But the simulated user does not consider the order in which different pieces of information are collected. We hope to fix this issue in a reasonable way in the future.

Even though RL naturally fits the setting of TREC DD, most participating systems are still using classical IR approaches, which are based on non-differentiable retrieval functions and hand-crafted optimization rules. We are anticipating more RL-based approaches.

4 CONCLUSIONS

TREC DD introduced a simulated user, which provides standard interface for the search system to interact with. With simplification, the user simulator makes live and reproducible experiment possible. We wish to expand the size of annotated data and improve the simulator, to attract researchers from a broader background to participate in this task.

ACKNOWLEDGMENTS

This research was supported by U.S. National Science Foundation IIS-145374. Any opinions, findings, conclusions, or recommendations expressed in this paper are of the authors, and do not necessarily reflect those of the sponsor.

REFERENCES

[1] Ben Carterette, Paul D. Clough, Mark M. Hall, Evangelos Kanoulas, and Mark Sanderson. 2016. Evaluating Retrieval over Sessions: The TREC Session Track 2011-2014. In *SIGIR '16*.
[2] Kalervo Järvelin, Susan L. Price, Lois M. L. Delcambre, and Marianne Lykke Nielsen. 2008. Discounted Cumulated Gain Based Evaluation of Multiple-Query IR Sessions. In *ECIR '08*.
[3] Jiyun Luo, Christopher Wing, Hui Yang, and Marti A. Hearst. 2013. The water filling model and the cube test: multi-dimensional evaluation for professional search. In *CIKM '13*.
[4] Evan Sandhaus. 2008. The new york times annotated corpus. *Linguistic Data Consortium, Philadelphia* 6, 12 (2008), e26752.
[5] Ellen M. Voorhees. 2002. Overview of TREC 2002. In *TREC '02*.
[6] Grace Hui Yang, Marc Sloan, and Jun Wang. 2016. *Dynamic Information Retrieval Modeling*. Morgan & Claypool Publishers.
[7] Grace Hui Yang, Zhiwen Tang, and Ian Soboroff. 2017. TREC '17 Dynamic Domain Track Overview. In *TREC '17*.
[8] Yiming Yang and Abhimanyu Lad. 2009. Modeling Expected Utility of Multi-session Information Distillation. In *ICTIR '09*.