ABSTRACT
Modern information retrieval (IR) consists of a series of processes, including query expansion, candidate item recall, item ranking, item re-ranking, etc. The final ranked item list will be exposed to the user, which will accordingly provide feedback through some expected actions such as browsing and click. Such a whole process can be formulated as a decision-making process where the agent is the IR system while the environment is the specific user. This decision-making process can be one-step or sequential, depending on the scenarios or the ways of problem formulation. Since 2013, Deep reinforcement learning (DRL) has been a fast-developing technique for decision-making tasks. The high capacity of deep learning models is incorporated in the reinforcement learning framework so that the agent may successfully handle complex decision-making. In recent years, there have been a bunch of publications attempting to leverage DRL techniques for different IR tasks such as ad hoc retrieval, learning to rank and interactive recommendation. Nonetheless, the fundamental theory, the principle of RL methods or the recognized experimental protocols of decision-making in IR, has not been well developed, making it challenging to evaluate the correctness of a proposed method or judge whether the reported experimental performance is valid. We propose the second DRL4IR workshop\(^1\) at SIGIR 2021, which provides a venue to gather the academia researchers and industry practitioners to present the recent progress of DRL techniques for IR. More importantly, people in this workshop are expected to discuss more about the fundamental principles of formulating a decision-making IR task, the underlying theory as well as the practical effectiveness of the experiment protocol design, which would foster further research on novel methodologies, innovative experimental findings and new applications of DRL for information retrieval.

\(^1\)https://drl4ir.github.io/

CCS CONCEPTS
- Information systems → Information retrieval; Web searching and information discovery; Online advertising; Recommender systems;
- Computing methodologies → Reinforcement learning;

KEYWORDS
Deep Reinforcement Learning, Information Retrieval

ACM Reference Format:

1 BACKGROUND AND MOTIVATIONS
Modern information retrieval (IR) consists of a series of processes, including query expansion, candidate item recall, item ranking, and re-ranking [3]. The final ranked item list will be exposed to the user, which will accordingly provide feedback through some actions such as browsing and click [5]. Moreover, such interactions can be expanded horizontally, e.g., involving multiple search sessions [8, 14, 23] or sequential recommendation spanning over a long time [7, 13]. No matter in one-step or sequential settings, such interaction processes can be formulated as a decision-making process where the agent is the IR system while the environment is the specific user [25]. The decision-making process is naturally counterfactual, i.e., the user would probably provide different feedbacks (behaviors) if another ranked list is provided to him in the historic context [1]. Therefore, directly performing machine learning, particularly supervised learning, over the logged behavior data will make the IR model fail to optimize the decision-making performance.

Deep reinforcement learning (DRL) has been a fast-developing technique for decision-making tasks since 2013 [20], where the high capacity of deep learning models is incorporated in the reinforcement learning framework so that the agent may successfully handle complex decision making to optimize the value in the long run [21, 30]. During the recent years, there have been a bunch of publications attempting to leverage DRL techniques for different IR tasks such as query reformulation [18], relevance feedback [22], learning to rank [15, 29] and interactive recommendation [8, 11, 27, 28].
Nonetheless, to our knowledge, the research along this line has not been clearly established, partially due to the following reasons. (i) The fundamental theory of decision-making in IR tasks may still be unclear. (ii) The problem formulations are different across papers, making it hard to categorize these researches. (iii) A substantial part of works leverage the off-the-shelf DRL methods to deal with the IR tasks, without considering the specific characteristics of IR. (iv) There is almost no well-recognized experimental protocol of decision-making in IR, which makes it infeasible to directly compare the methods proposed in different papers, nor to even judge whether the reported experimental performance is valid.

Therefore, by hosting this workshop, we hope to provide a venue to gather the academia researchers and industry practitioners to present the recent progress of DRL techniques for IR. More importantly, people in this workshop are expected to discuss more about the fundamental principles of formulating a decision-making IR task, the underlying theory as well as the practical effectiveness of the experiment protocol design, which would further research on novel methodologies, innovative experimental findings, and new applications of DRL for information retrieval.

**DRL@IR at SIGIR’20.** Last year we organized the workshop at SIGIR’20 [26], which was one of the most popular workshops and was broadly welcomed by 200 conference attendees. We invited several keynote speakers from academia and industry, e.g., Prof. Yang Yu (NJU), Dr. Zhe Xu (DiDi) and Dr. Alex Beutel (Google Research). We accepted 4 papers from 12 submissions. This year, we will pay attention to a deep investigation of fundamental aspects of the problem as mentioned above, as well as a broader range of topics including but not limited to e-commerce, recommendation, advertising, web search and finance, etc. We plan to have keynotes, paper talks and a panel discussion. We expected 300 participants this year.

## 2 FUNDAMENTAL CHALLENGES

A literature review and some technical challenges are presented in the workshop abstract paper at SIGIR’20 [26]. This year we pay more attention to some fundamental challenges in this field and expect a thorough discussion of them during the workshop.

**Off-policy or offline RL in IR scenario** Due to the system architecture in modern web services, it is almost impossible to perform online reinforcement learning in real-world IR applications [12, 19, 31, 32]. In practice, the ranking functions (or policies) are trained periodically, e.g., daily or weekly, based on the recently collected log data. As such, the ways of training of any DRL-based policies or value functions can only be off-policy or offline. Due to the nature of RL, once the ranking policy is updated, the corresponding data distribution, i.e., occupancy measure, will change, which will make the training biased. In DRL, off-policy methods restrict the new policy around the data-collecting policy while offline RL methods usually constrain the learning policy to support the logged data. So far, there is barely no established research addressing such a problem in IR tasks.

**Formulating ranking as a decision-making task** Ranking is a process that takes a query as input and outputs the ranking of candidate items. Different formulations of ranking lead to different inductive biases, which may fit different IR scenarios. For example, sequentially decide to pick the next-slot item from the remaining candidates fits more to the sequential recommendation tasks but may suffer from efficiency issue [7, 24], while learning a direct ranking function may fail to capture the user’s specific context when browsing the ranked list [6, 10, 17]. To our knowledge, there has not yet been a deep investigation of the effectiveness of different formulation of ranking as a decision-making task.

**Dynamics simulation** As previously mentioned, a user’s behavior when facing a ranked item list can be regarded as a dynamic process, which can be formulated as a prediction or generation process called user click model [4, 9]. Learning such a user click model (i.e., the dynamics for RL-based ranking policies) is still challenging since the logged data is always biased, e.g., position bias and selection bias. Moreover, how well the learned user click model will guide the training of the RL-based ranking policies is non-guaranteed, because the common problem of compounding error of the learned dynamics model for simulation data generation [2]. Meanwhile, there is no commonly recognized simulation benchmark for DRL-based ranking or recommendation tasks, although some attempts have been performed, such as RecSim [16].

Addressing above challenges requires the fundamental investigation of deep reinforcement learning theory and methods from the perspective of IR. Also, the support from industry is a necessity as the data collection from explorative policy would make it more suitable for RL benchmark building, and the principles of building effective simulators come from the observations from the real user behaviors.

## 3 PROGRAM SKETCH

### 3.1 Workshop Format

The workshop is planned to be host a whole day, with 2 keynotes, 4 invited talks and 6 oral research talks. The keynote speakers should be well-recognized professors or scientists working in the area. There are two encouraging types of invited talks and peer-reviewed oral research talks: (i) the academic talk on fundamental research on reinforcement learning with an attempt of application on IR; (ii) the industrial talk on practice of designing or applying deep reinforcement learning techniques for real-world IR tasks.

Each talk is expected to be presented as a lecture with slides. There will be a QA session at the end of each talk.

### 3.2 Tentative Workshop Schedule

The workshop schedule is planned with two half-day sessions.

- **Morning session**
  - 08:30 - 08:40 Welcome & opening
  - 08:40 - 09:10 Keynote 1
  - 09:10 - 09:35 Academic invited talk 1
  - 09:35 - 10:00 Industrial invited talk 1
  - 10:00 - 10:20 Coffee break
  - 10:20 - 12:00 Oral research talks

- **Afternoon session**
  - 13:30 - 14:10 Keynote 2
  - 14:10 - 14:35 Academic invited talk 2
  - 14:35 - 15:00 Industrial invited talk 2
  - 15:00 - 15:20 Coffee break
will be sent to the authors. A website (http://drl4ir.github.io) for this workshop will be made available online right before the lecture is presented. All the relevant materials will be made available on this website, including the talk information, presentation slides, referred papers, speaker information, and related open-source projects, etc.

3.3 Selection Process
Each invited speaker should be well recognized in the community. The invitation should be notified by all organizers with no disagreement.

This workshop opens paper submission with a standard peer-review process. Each paper submission should be reviewed by at least two program committee members and have a recommendation of acceptance or rejection by the senior PCs or the workshop organizers. The detailed review comments and the notification letter will be sent to the authors.

3.4 Online Materials
A website (http://drl4ir.github.io) for this workshop will be made available online right before the lecture is presented. All the relevant materials will be made available on this website, including the talk information, presentation slides, referred papers, speaker information, and related open-source projects, etc.

4 RELATED WORKSHOPS
List of related workshops:
- Deep Reinforcement Learning for Information Retrieval (SIGIR 2020)
- Reinforcement Learning for Knowledge Discovery (KDD 2019)²
- Deep Reinforcement Learning Workshop (NeurIPS 2015-2020)³
- Deep Reinforcement Learning: Frontiers and Challenges (IJCAI 2016)⁴
- Deep Reinforcement Learning Meets Structured Prediction (ICLR 2019)⁵

The Deep Reinforcement Learning Workshop at NeurIPS (2015-2020) and IJCAI (2016) focused on the techniques to combine neural networks with reinforcement learning, and domains like robotics, strategy games, and multi-agent interaction. The Deep Reinforcement Learning Meets Structured Prediction at ICLR (2019) focused on leveraging reinforcement learning paradigm on tasks of structured predictions. The workshops at KDD (2019) focuses on a wide range of real-life reinforcement learning applications. DRL4IR at SIGIR 2020 was the first workshop to focus on deep reinforcement learning for information retrieval, which was broadly welcomed by the conference attendees and achieved great success. This year, the proposed workshop will focus on the recent progress of deep reinforcement learning for information retrieval and pay attention to a broader range of topics. This workshop will bring together experts in information retrieval and reinforcement learning. Our proposed workshop will have invited keynotes and talks, paper presentations, poster session, and panel discussion to help interested researchers gain a high-level view about the current state of the art and potential directions for future contributions. Real datasets and codes will also be released for attendees to practice in the future.

5 ORGANIZERS
Dr. Weinan Zhang, the lead organizer of the workshop, is currently a tenure-track associate professor in Shanghai Jiao Tong University. His research interests include machine learning and big data mining, particularly, deep learning and reinforcement learning techniques for real-world data mining scenarios, such as computational advertising, recommender systems, text mining, web search and knowledge graphs. He has published over 100 papers on first-tier international conferences and journals, including KDD, SIGIR, ICLM, ICLR, JMLR, IJCAI, AAAI, WSDM, CIKM etc. He won the Best Paper Honorable Mention Award in SIGIR 2017, the Best Paper Award in DLP Workshop at KDD 2019, the Best System Paper Award at CoRL 2020, ACM Rising Star Award, Alibaba DAMO Young Scholar Award etc. Weinan has organized workshops and tutorials in SIGIR, KDD, CIKM and ECIR etc.

Dr. Xiangyu Zhao is an assistant professor of the school of data science at City University of Hong Kong (CityU). His current research interests include data mining and machine learning, especially (1) Reinforcement Learning, AutoML, and their applications in Information Retrieval (recommendation, computational advertising and search); (2) Urban Computing and Spatio-Temporal Data Analysis. He has published more than 20 papers in top conferences (e.g., KDD, WWW, SIGIR, AAAI, CIDE, CIKM, ICDM, WSDM, RecSys) and journals (e.g., SIGKDD, SIGWeb, EPL, APS). His research received Criteo Research Award, Bytedance Research Award, Top 100 Chinese New Stars in Artificial Intelligence and MSU Dissertation Fellowship. He serves as the organizers of DRL4KDD@KDD’19, DRL4IR@SIGIR’20, 2nd DRL4KD@WWW’21, 2nd DRL4IR@SIGIR’21, and a lead tutor at WWW’21 and IJCAI’21.

Dr. Li Zhao is currently a Senior Researcher in Machine Learning Group, Microsoft Research Asia (MSRA). Her research interests mainly lie in deep learning and reinforcement learning, and their applications for text mining, recommendation, finance and games. She has co-organized the 3rd Asian Workshop on Reinforcement Learning (AWRL’18), and is one of the invited speakers for AWRL’19. She obtained her Ph.D. degree majoring in Computer Science in July, 2016, from Tsinghua University, supervised by Professor Xiaoyan Zhu. During her Ph.D. studies, she has conducted research on sentiment extraction, text mining and weakly supervised learning. She published several research papers in top conferences, including NeurIPS, KDD, IJCAI, AAAI, EMNLP and CIKM.

Dr. Dawei Yin is Engineering Director at Baidu inc.. He is managing the search science team at Baidu, leading Baidu’s science efforts of web search, question answering, video search, image search, news search, app search, etc. Previously, he was Senior Director, managing the recommendation engineering team at JD.com between 2016 and 2020. Prior to JD.com, he was Senior Research Manager at Yahoo Labs, leading relevance science team and in

2http://www.cse.msu.edu/~zhaoxi35/DRL4KDD/
3https://sites.google.com/view/deep-rl-workshop-neurips2020/home
4https://sites.google.com/site/deeprlijcai16/
5https://sites.google.com/view/iclr2019-drlstructpred/
charge of Core Search Relevance of Yahoo Search. He obtained Ph.D. (2013), M.S. (2010) from Lehigh University and B.S. (2006) from Shandong University. From 2007 to 2008, he was an M.Phil. student in The University of Hong Kong. His research interests include data mining, applied machine learning, information retrieval and recommender system. He published more than 80 research papers in premier conferences and journals, and was the recipients of WSDM2016 Best Paper Award, KDD2016 Best Paper Award, WSDM2018 Best Student Paper Award, and ICHI 2019 Best Paper Honorable Mention.

Dr. Grace Hui Yang is an Associate Professor in the Department of Computer Science at Georgetown University. Dr. Yang is leading the InfoSense (Information Retrieval and Sense-Making) group at Georgetown University, Washington D.C., U.S.A. Dr. Yang obtained her Ph.D. from the Language Technologies Institute, Carnegie Mellon University in 2011. Dr. Yang’s current research interests include deep reinforcement learning, dynamic information retrieval, search engine evaluation, privacy-preserving information retrieval, internet of things, and information organization. Prior to this, she has conducted research on question answering, ontology construction, near-duplicate detection, multimedia information retrieval, and opinion and sentiment detection. Dr. Yang has co-chaired SIGIR 2013 and 2014 Doctoral Consortiums, SIGIR 2017 Workshop, WSDM 2017 Workshop, ICTIR 2017 Workshop, CIKM 2015 Tutorial, ICTIR 2018 Short Paper and SIGIR 2018 Demonstration Paper Program Committees. Dr. Yang served on the editorial board of Information Retrieval Journal from 2014 to 2017.

ACKNOWLEDGEMENT.

This research is supported by “New Generation of AI 2030” Major Project (2018AAA0100900), Chinese National Science Foundation (61632017, 61702327, 61772335), and U.S. National Science Foundation IIS-1453721, IIS1907704, IIS1928278, IIS1714741, IIS1715940, IIS1845081 and CNS1815636.

REFERENCES