AgentIR: 1st Workshop on Agent-based Information Retrieval

Qingpeng Cai
Kuaishou Technology
Beijing, China
cqpcurry@gmail.com

Xiangyu Zhao
City University of Hong Kong
Hong Kong, China
xianzhao@cityu.edu.hk

Ling Pan
Hong Kong University of Science and Technology
Hong Kong, China
lingpan@ust.hk

Xin Xin
Shandong University
Jinan, China
xinxin@sdu.edu.cn

Jin Huang
University of Amsterdam
Amsterdam, Netherlands
j.huang2@uva.nl

Weinan Zhang
Shanghai Jiao Tong University
Shanghai, China
wnzhang@sjtu.edu.cn

Li Zhao
Microsoft Research
Beijing, China
lizo@microsoft.com

Dawei Yin
Baidu
Beijing, China
yindawei@acm.org

Grace Hui Yang
Georgetown University
Washington, D.C., United States
huiyang@cs.georgetown.edu

ABSTRACT

Information retrieval (IR) systems have become an essential component in modern society to help users find useful information, which consists of a series of processes including query expansion, item recall, item ranking and re-ranking, etc. Based on the ranked information list, users can provide their feedbacks. Such an interaction process between users and IR systems can be naturally formulated as a decision-making problem, which can be either one-step or sequential. In the last ten years, deep reinforcement learning (DRL) has become a promising direction for decision-making, since DRL utilizes the high model capacity of deep learning for complex decision-making tasks. On the one hand, there have been emerging research works focusing on leveraging DRL for IR tasks. However, the fundamental information theory under DRL settings, the challenge of RL methods for Industrial IR tasks, or the simulations of DRL-based IR systems, has not been deeply investigated. On the other hand, the emerging LLM provides new opportunities for optimizing and simulating IR systems. To this end, we propose the first Agent-based IR workshop at SIGIR 2024, as a continuation from one of the most successful IR workshops, DRL4IR. It provides a venue for both academia researchers and industry practitioners to present the recent advances of both DRL-based IR systems and LLM-based IR systems from the agent-based IR's perspective, to foster novel research, interesting findings, and new applications.

CCS CONCEPTS

• **Information systems** → **Information retrieval**; • **Computing methodologies** → **Reinforcement learning**.

SIGIR '24, July 14–18, 2024, Washington, DC, USA

ACM ISBN 979-8-4007-0431-4/24/07
https://doi.org/10.1145/3626772.3657989

KEYWORDS

DRL, LLM, Agent-based Information Retrieval

ACM Reference Format:
Qingpeng Cai, Xiangyu Zhao, Ling Pan, Xin Xin, Jin Huang, Weinan Zhang, Li Zhao, Dawei Yin, and Grace Hui Yang. 2024. AgentIR: 1st Workshop on Agent-based Information Retrieval. In *Proceedings of the 47th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '24), July 14–18, 2024, Washington, DC, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3626772.3657989

1 WORKSHOP THEME AND TOPICS

Information retrieval (IR) systems have become an essential component in modern society to help users find useful information, which consists of a series of processes including query expansion, item recall, item ranking and re-ranking [2, 36]. Based on the ranked information list, users can provide their feedbacks, for example, through browsing and clicks [4, 36]. Such user-item interactions can be expanded horizontally, e.g., involving multiple search sessions over a long timescale [7, 8, 10, 11, 27, 36]. In either one-step setting or sequential setting, such interaction processes can both be formulated as a decision-making process where the IR system is the agent while the user acts as the environment [33, 36]. The decision-making process is naturally counterfactual, i.e., the user would probably provide different feedbacks (behaviors) if a different ranked list of items is provided in the corresponding context [1, 36]. As a result, simply using naive supervised learning to regress to the logged user behavior data could not yield optimal performance.

Deep reinforcement learning (DRL) and multi-agent reinforcement learning has become an emerging and promising direction for decision-making tasks in the last ten years [15], where the high model capacity of deep learning is incorporated into the reinforcement learning paradigm to handle complex decision-making scenarios to optimize the expected return (reward) in the long run [16, 19–21, 24, 28, 36, 43]. However, reinforcement learning methods often encounter the issue of limited diversity in the solutions they discover. The emergence of generative flow networks (GFlowNets) has effectively addressed this problem by achieving high rewards

while also maintaining diversity[3, 22, 23, 25, 26, 34, 35]. Recently, there have been an emerging bunch of research attempting to leverage DRL techniques for different IR tasks such as query reformulation [13], relevance feedback [17], learning to rank [12, 42] and recommendation [5, 6, 8, 9, 14, 32, 38, 39, 41]. However, as far as we are concerned, the research regarding DRL for IR tasks has not been well justified, partially due to the following reasons: (i) The fundamental information theory of decision-making may still be unclear; (ii) The evaluation metrics and baselines are not consistent in papers; (iii) A substantial amount of research works just leverage the off-the-shelf DRL methods to handle IR tasks, without considering the specific characteristics and demands of industrial IR, such as the challenges of applying DRL methods to real IR systems.

LLM[18] methods has attract a lot of attention from both academia and industry sides. LLM-based IR[29] systems that focus on applying LLM to either directly recall and rank items, or pre-train the ID embedding, and serves as a controller to call different recommendation modules. Also, the major success of ChatGPT is due to reinforcement learning from human feedback(RLHF). It inspires research works[31] to introduce human preferences into recommender systems. However, the foundations and applications of LLM-based IR are not well defined and verified.

We propose the concept, agent-based information retrieval(AgentIR), where the goal is to construct an intelligent agent that can autonomously handle information retrieval tasks with an emphasis of techniques from LLM, RL, and memory. With LLM technology, the agent can learn language models to improve retrieval efficiency and accuracy. The RL technique enables the agent to continuously optimize its decision-making process to adapt to different retrieval tasks. With adavanced memory design, we can help the agent better manage and utilize existing information, thereby improving retrieval efficiency and quality. We believe that this agent-based information retrieval system will be a highly promising research direction. People in this workshop are expected to discuss more about the fundamental principles of formulating LLM-based IR and DRL-based IR tasks as well as the practical effectiveness of the experimental design, which would foster novel research, interesting findings, and new applications.

DRL4IR at SIGIR'20/21/22 and CIKM'23. Last years we have organized the workshop at SIGIR'20/21/22 [36, 37, 40] and CIKM'23[30], which was one of the most popular workshops and was broadly welcomed by 200 conference attendees each year. This year, we will pay attention to a more deep investigation of fundamental aspects of topics including but not limited to DRL, LLM and agents. We plan to have keynotes, paper talks and a panel discussion. We expected 300 participants this year.

2 PROGRAM SKETCH

2.1 Workshop Format

The workshop is planned to be host a whole day hybrid, 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 DRL 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.

2.2 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
  - 15:20 - 16:20 Oral research talks
  - 16:20 - 17:00 Panel

3 RELATED WORKSHOPS

List of related workshops:

- Deep Reinforcement Learning for Information Retrieval (SIGIR 2020-2022, CIKM 2023)
- Reinforcement Learning for Knowledge Discovery (KDD 2019[1])
- Deep Reinforcement Learning Workshop (NeurIPS 2015-2020[2])
- Deep Reinforcement Learning: Frontiers and Challenges (IJCAI 2016[3])
- Deep Reinforcement Learning Meets Structured Prediction (ICLR 2019[4])

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 2020/2021/2022 and CIKM 2023 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 LLM for information retrieval. This workshop will bring together experts in information retrieval, reinforcement learning and LLM. 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

[1] http://www.cse.msu.edu/~zhaoxi35/DRL4KDD/
[2] https://sites.google.com/view/deep-rl-workshop-neurips2020/home
[3] https://sites.google.com/site/deeprlijcai16/
[4] https://sites.google.com/view/iclr2019-drlstructpred/

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.

4 ORGANIZERS INFORMATION AND QUALIFICATION

Dr. Qingpeng Cai is a Senior Staff Algorithm Engineer in KuaiShou Technology, where he leads the Listwise Recommendation group. He obtained a PhD degree in Computer Science from Tsinghua University. His research interests include reinforcement learning, and recommender system. Currently, he has published more than 20 papers in top international conferences. He was the invited speaker of DRL4IR@SIGIR'22 and VideoRecSys@ACM RecSys'23 with the topic "Reinforcement Learning for Short Video Recommender Systems".

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) Personalization, Recommender System, Online Advertising, Search Engine, and Information Retrieval; and (2) Large Language Models and Reinforcement Learning. He has published more than 100 papers in top conferences and journals. His research received ICDM'22 and 21 Best-ranked Papers, Global Top 100 Chinese New Stars in AI, CCF-Tencent Open Fund, Criteo Research Award, and Bytedance Research Award. 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'21/22 and IJCAI'21.

Dr. Ling Pan is an assistant professor of the Department of Electronic and Computer Engineering at the Hong Kong University of Science and Technology (HKUST). Her research interests mainly include theoretical understanding, algorithmic improvements and practical application of generative flow networks (GFlowNets), reinforcement learning and multi-agent systems. She was also a recepient of Microsoft Research Ph.D. Fellowship (Asia). She has published several high-quality papers on top-tiered machine learning and artificial intelligence conferences and journals including ICML/NeurIPS/ICLR/AAAI/IJCAI/AAMAS, and her works have been selected as spotlight presentations for multiple times. She served as the organizer of the Mila GFlowNet Workshop 2023.

Dr. Xin Xin is an assistant professor of the school of computer science and technology at Shandong University (SDU). Before joining SDU, he got the Ph.D. degree in University of Glasgow. His current research interests include informational retrieval and machine learning, especially (1) Personalized Recommendation, and (2) Reinforcement Learning. He has published more than 20 papers in top conferences (e.g., WWW, SIGIR, ICDM, CIKM, WSDM, IJCAI, ACL) and journals (e.g., TKDE). He was an invited speaker of DRL4IR@SIGIR'21 with the topic of offline reinforcement learning for recommendation.

Dr. Jin Huang is a fourth-year PhD student in University of Amster-dam (UvA). Before joining UvA, she got her BE and ME degree in Renmin University of China. Her current research interests in- clude but are not limited to personalized recommendations, deep reinforcement learning for recommendations, and debiasing recom- mendations. She has published several research papers in top-tier conferences, including SIGIR, WSDM, and RecSys.

Dr. Weinan Zhang 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, ICML, 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. 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 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 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 premium 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 Associate Professor in Computer Science at Georgetown University, Washington D.C. Dr. Yang's current research interests include deep reinforcement learning, interactive agents, search engines, and privacy-preserving information retrieval. Prior to this, she has conducted research on question answering, automatic ontology construction, near-duplicate detection, multimedia information retrieval, and opinion and sentiment detection. Dr. Yang led the effort for the Text Retrieval Conference (TREC) Dynamic Domain Tracks from 2015 to 2017 and SIGIR privacy-preserving information retrieval workshops from 2014 to 2016 and co-organized the SIGIR Deep Reinforcement Learning Workshops since 2020. Dr. Yang is associate editor for ACM Transactions on Information Systems and served on the editorial board of Information Retrieval Journal from 2014 to 2017. She has actively served as an organizing or program committee member in

many top-tier international conferences such as SIGIR, ECIR, ACL, AAAI, ICTIR, CIKM, WSDM, and WWW, and the general co-chair of SIGIR 2024. She is a recipient of the NSF Faculty Early Career Development Program (CAREER) Award and a co-author of the 2016 book "Dynamic Information Retrieval Modeling."

ACKNOWLEDGEMENT.

This research was supported by US NSF grant III-2336768.

REFERENCES

[1] Aman Agarwal, Kenta Takatsu, Ivan Zaitsev, and Thorsten Joachims. 2019. A general framework for counterfactual learning-to-rank. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 5–14.

[2] Ricardo Baeza-Yates, Berthier Ribeiro-Neto, et al. 1999. *Modern information retrieval*. Vol. 463. ACM press New York.

[3] Emmanuel Bengio, Moksh Jain, Maksym Korablyov, Doina Precup, and Yoshua Bengio. 2021. Flow network based generative models for non-iterative diverse candidate generation. *Advances in Neural Information Processing Systems* 34 (2021), 27381–27394.

[4] Alexey Borisov, Martijn Wardenaar, Ilya Markov, and Maarten de Rijke. 2018. A click sequence model for web search. In *SIGIR*. 45–54.

[5] Qingpeng Cai, Shuchang Liu, Xueliang Wang, Tianyou Zuo, Wentao Xie, Bin Yang, Dong Zheng, Peng Jiang, and Kun Gai. 2023. Reinforcing User Retention in a Billion Scale Short Video Recommender System. *arXiv preprint arXiv:2302.01724* (2023).

[6] Qingpeng Cai, Zhenghai Xue, Chi Zhang, Wanqi Xue, Shuchang Liu, Ruohan Zhan, Xueliang Wang, Tianyou Zuo, Wentao Xie, Dong Zheng, et al. 2023. Two-Stage Constrained Actor-Critic for Short Video Recommendation. In *Proceedings of the ACM Web Conference 2023*. 865–875.

[7] Haokun Chen, Xinyi Dai, Han Cai, Weinan Zhang, Xuejian Wang, Ruiming Tang, Yuzhou Zhang, and Yong Yu. 2019. Large-scale interactive recommendation with tree-structured policy gradient. In *AAAI*, Vol. 33. 3312–3320.

[8] Jia Chen, Jiaxin Mao, Yiqun Liu, Min Zhang, and Shaoping Ma. 2020. A context-aware click model for web search. In *WSDM*. 88–96.

[9] Wenqi Fan, Tyler Derr, Xiangyu Zhao, Yao Ma, Hui Liu, Jianping Wang, Jiliang Tang, and Qing Li. 2020. Attacking Black-box Recommendations via Copying Cross-domain User Profiles. *arXiv preprint arXiv:2005.08147* (2020).

[10] Yingqiang Ge, Shuchang Liu, Ruoyuan Gao, Yikun Xian, Yunqi Li, Xiangyu Zhao, Changhua Pei, Fei Sun, Junfeng Ge, Wenwu Ou, et al. 2021. Towards Long-term Fairness in Recommendation. *arXiv preprint arXiv:2101.03584* (2021).

[11] Li He, Liang Wang, Kaipeng Liu, Bo Wu, and Weinan Zhang. 2018. Optimizing Sponsored Search Ranking Strategy by Deep Reinforcement Learning. *arXiv preprint arXiv:1803.07347* (2018).

[12] Yujing Hu, Qing Da, Anxiang Zeng, Yang Yu, and Yinghui Xu. 2018. Reinforcement Learning to Rank in E-Commerce Search Engine: Formalization, Analysis, and Application. *arXiv preprint arXiv:1803.00710* (2018).

[13] Johnny Flame Lee. 2020. *Evaluation of a Deep Reinforcement Learning Method For Query Reformulation*. Ph. D. Dissertation. University of Otago.

[14] Shuchang Liu, Qingpeng Cai, Bowen Sun, Yuhao Wang, Ji Jiang, Dong Zheng, Peng Jiang, Kun Gai, Xiangyu Zhao, and Yongfeng Zhang. 2023. Exploration and Regularization of the Latent Action Space in Recommendation. In *Proceedings of the ACM Web Conference 2023*. 833–844.

[15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. 2013. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:1312.5602* (2013).

[16] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. 2015. Human-level control through deep reinforcement learning. *nature* 518, 7540 (2015), 529–533.

[17] Ali Montazeralghaem, Hamed Zamani, and James Allan. 2020. A Reinforcement Learning Framework for Relevance Feedback. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 59–68.

[18] Long Ouyang, Jeffrey Wu, Xu Jiang, Diogo Almeida, Carroll Wainwright, Pamela Mishkin, Chong Zhang, Sandhini Agarwal, Katarina Slama, Alex Ray, et al. 2022. Training language models to follow instructions with human feedback. *Advances in Neural Information Processing Systems* 35 (2022), 27730–27744.

[19] Ling Pan, Qingpeng Cai, Zhixuan Fang, Pingzhong Tang, and Longbo Huang. 2019. A deep reinforcement learning framework for rebalancing dockless bike sharing systems. In *Proceedings of the AAAI conference on artificial intelligence*, Vol. 33. 1393–1400.

[20] Ling Pan, Qingpeng Cai, and Longbo Huang. 2020. Softmax deep double deterministic policy gradients. *Advances in Neural Information Processing Systems* 33 (2020), 11767–11777.

[21] Ling Pan, Longbo Huang, Tengyu Ma, and Huazhe Xu. 2022. Plan better amid conservatism: Offline multi-agent reinforcement learning with actor rectification. In *International conference on machine learning*. PMLR, 17221–17237.

[22] Ling Pan, Moksh Jain, Kanika Madan, and Yoshua Bengio. 2024. Pre-Training and Fine-Tuning Generative Flow Networks. In *The Twelfth International Conference on Learning Representations*.

[23] Ling Pan, Nikolay Malkin, Dinghuai Zhang, and Yoshua Bengio. 2023. Better training of gflownets with local credit and incomplete trajectories. In *International Conference on Machine Learning*. PMLR, 26878–26890.

[24] Ling Pan, Tabish Rashid, Bei Peng, Longbo Huang, and Shimon Whiteson. 2021. Regularized softmax deep multi-agent q-learning. *Advances in Neural Information Processing Systems* 34 (2021), 1365–1377.

[25] Ling Pan, Dinghuai Zhang, Aaron Courville, Longbo Huang, and Yoshua Bengio. 2022. Generative Augmented Flow Networks. In *The Eleventh International Conference on Learning Representations*.

[26] Ling Pan, Dinghuai Zhang, Moksh Jain, Longbo Huang, and Yoshua Bengio. 2023. Stochastic generative flow networks. In *Uncertainty in Artificial Intelligence*. PMLR, 1628–1638.

[27] Razieh Rahimi and Grace Hui Yang. [n. d.]. Modeling Exploration of Intrinsically Diverse Search Tasks as Markov Decision Processes. ([n. d.]).

[28] Tabish Rashid, Mikayel Samvelyan, Christian Schroeder De Witt, Gregory Farquhar, Jakob Foerster, and Shimon Whiteson. 2020. Monotonic value function factorisation for deep multi-agent reinforcement learning. *Journal of Machine Learning Research* 21, 178 (2020), 1–51.

[29] Likang Wu, Zhi Zheng, Zhaopeng Qiu, Hao Wang, Hongchao Gu, Tingjia Shen, Chuan Qin, Chen Zhu, Hengshu Zhu, Qi Liu, et al. 2023. A Survey on Large Language Models for Recommendation. *arXiv preprint arXiv:2305.19860* (2023).

[30] Xin Xin, Xiangyu Zhao, Jin Huang, Weinan Zhang, Li Zhao, Dawei Yin, and Grace Hui Yang. 2023. DRL4IR: 4th Workshop on Deep Reinforcement Learning for Information Retrieval. In *Proceedings of the 32nd ACM International Conference on Information and Knowledge Management*. 5304–5307.

[31] Wanqi Xue, Qingpeng Cai, Zhenghai Xue, Shuo Sun, Shuchang Liu, Dong Zheng, Peng Jiang, Kun Gai, and Bo An. 2022. PrefRec: Recommender Systems with Human Preferences for Reinforcing Long-term User Engagement. *arXiv preprint arXiv:2212.02779* (2022).

[32] Wanqi Xue, Qingpeng Cai, Ruohan Zhan, Dong Zheng, Peng Jiang, and Bo An. 2023. ResAct: Reinforcing Long-term Engagement in Sequential Recommendation with Residual Actor (ICLR'23). (2023).

[33] Grace Hui Yang, Marc Sloan, and Jun Wang. 2016. Dynamic information retrieval modeling. *Synthesis lectures on information concepts, retrieval, and services* 8, 3 (2016), 1–144.

[34] Dinghuai Zhang, Hanjun Dai, Nikolay Malkin, Aaron C Courville, Yoshua Bengio, and Ling Pan. 2024. Let the flows tell: Solving graph combinatorial problems with GFlowNets. *Advances in Neural Information Processing Systems* 36 (2024).

[35] Dinghuai Zhang, Ling Pan, Ricky TQ Chen, Aaron Courville, and Yoshua Bengio. 2023. Distributional gflownets with quantile flows. *arXiv preprint arXiv:2302.05793* (2023).

[36] Weinan Zhang, Xiangyu Zhao, Li Zhao, Dawei Yin, and Grace Hui Yang. 2021. *DRL4IR: 2nd Workshop on Deep Reinforcement Learning for Information Retrieval*. Association for Computing Machinery, New York, NY, USA, 2681–2684. https://doi.org/10.1145/3404835.3462818

[37] Weinan Zhang, Xiangyu Zhao, Li Zhao, Dawei Yin, Grace Hui Yang, and Alex Beutel. 2020. Deep Reinforcement Learning for Information Retrieval: Fundamentals and Advances. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval*. 2468–2471.

[38] Kesen Zhao, Shuchang Liu, Qingpeng Cai, Xiangyu Zhao, Ziru Liu, Dong Zheng, Peng Jiang, and Kun Gai. 2023. KuaiSim: A comprehensive simulator for recommender systems. *arXiv preprint arXiv:2309.12645* (2023).

[39] Xiangyu Zhao, Long Xia, Liang Zhang, Zhuoye Ding, Dawei Yin, and Jiliang Tang. 2018. Deep Reinforcement Learning for Page-wise Recommendations. In *Proceedings of the 12th ACM Recommender Systems Conference*. ACM, 95–103.

[40] Xiangyu Zhao, Xin Xin, Weinan Zhang, Li Zhao, Dawei Yin, and Grace Hui Yang. 2022. DRL4IR: 3rd Workshop on Deep Reinforcement Learning for Information Retrieval. In *Proceedings of the 45th International ACM SIGIR Conference on Research and Development in Information Retrieval*. 3488–3491.

[41] Xiangyu Zhao, Liang Zhang, Zhuoye Ding, Long Xia, Jiliang Tang, and Dawei Yin. 2018. Recommendations with Negative Feedback via Pairwise Deep Reinforcement Learning. In *KDD*. ACM, 1040–1048.

[42] Xiangyu Zhao, Xudong Zheng, Xiwang Yang, Xiaobing Liu, and Jiliang Tang. 2020. Jointly learning to recommend and advertise. In *KDD*. 3319–3327.

[43] Guanjie Zheng, Fuzheng Zhang, Zihan Zheng, Yang Xiang, Nicholas Jing Yuan, Xing Xie, and Zhenhui Li. 2018. DRN: A Deep Reinforcement Learning Framework for News Recommendation. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*. International World Wide Web Conferences Steering Committee, 167–176.