

# Reinforcement learning models for search-oriented conversational systems

**Supervisor:** Laure Soulier (laure.soulier@lip6.fr) - MLIA team in LIP6 laboratory (France)  
In collaboration with Ludovic Denoyer (denoyer@fb.com) - Professor LIP6 on secondment at FAIR (France)

**Place of the thesis:** LIP6 lab in the MLIA team.

**Profil:** Ingeneer an/or master student (BAC +5) in Computer Science (Machine learning, Reinforcement Learning, Information retrieval, Natural language processing) or Mathematics.

**To apply:**, send to laure.soulier@lip6.fr and denoyer@fb.com:

- a curriculum vitae, with contact of 2 or more referees
- an cover letter
- a research outcome (e.g. master thesis and/or published papers) of the candidate
- a transcript of grades
- letter of recommendation from previous supervisors or master degree responsables.

**Starting period:** The thesis is planned to start in January 2021. Applications will be processed on the flow and will close at the latest on November 30, 2020.

We are recruiting for the full-time position of a PhD researcher in the fields of Machine Learning (ML)/Information Retrieval (IR) to join the ANR Project on "Search-oriented conversational systems" (SESAMS).

The candidate will be involved in the design of reinforcement learning models for supporting search-oriented conversational systems. The objective of the thesis is to reverse the information retrieval paradigm based on scoring models to design IR-driven reinforcement policy by taking into account users' interactions with search engines and conversational systems.

## Thesis topic

In conversational systems, interactions engaged by the conversational system are inherently correlated with the quality of the outcomes. Indeed, while the goal of chit-chat bots is to provide discussions that are in the continuity of the conversation [RCD11], task-oriented bots generally focus on identifying the right response with respect to a particular question (whether associated to a slot filling step or not) as done in ad-hoc retrieval [BW16]. However, the multi-turn setting imposed by conversational systems offers the possibility to engage the system in a constructive interaction with the user to help the user accomplishing his/her task. In conversational systems, this feature is mainly addressed using decision-making processes based on reinforcement learning models, such as POMDP [YGTW13, SDVM<sup>+</sup>17] or deep reinforcement learning [LCL<sup>+</sup>17, WVM<sup>+</sup>17, ZE16]. These approaches are used to anticipate the next action of the user and accordingly moderate the system response. In the context of conversational search, the framework is different since it assumes that system-to-user interactions are specific to the IR

purposes [RC17]. For instance, the search-oriented conversational system might ask the user to refine the query or to express his/her preferences between queries reformulated by the conversational system or between documents retrieved by the search engine. These bi-directional interactions within search-oriented conversational systems open a new perspective to user-driven system-mediated scenarios, as defined in the context of collaborative search [SST14]. Search-oriented conversational systems have led to exciting research work these last years: the definition of the new paradigm and its underlying challenges [ACH<sup>+</sup>20, RC17], interactive IR or query reformulation models to mimic the behavior of conversational systems [ASD18, MRC<sup>+</sup>20, VLR<sup>+</sup>20], and more recently the introduction of reinforcement learning frameworks for ranking [XWX<sup>+</sup>20] or interactive IR [CTY20, MZA20, ZZZ<sup>+</sup>20]. This thesis is in line with this last research direction and aims at exploiting machine learning techniques to the benefit of IR models.

## Context

The SESAMS project involves several researchers from world-wide laboratory (LIP6 as main investigator, Facebook France, Montréal, Lille). The thesis will be co-supervised by Laure Soulier (MLIA-LIP6) and Ludovic Denoyer (FAIR). The candidate will be host in the LIP6 lab in the MLIA team.

The MLIA team at computer Science lab. at Sorbonne, hosts 10 permanent researchers and 30 PhDs. The lab is equipped with computer facilities including GPU clusters (170 GPU cards) required for machine learning and deep learning projects. The PhD students will have a workplace and access to all the facilities of the lab. Sorbonne University has launched in 2019 a research center on AI (Sorbonne Center on AI). The PhD students will have access to all the facilities provided by the center.

## Required Qualification

Candidates have an outstanding master's degree (or an equivalent university degree) in computer science or another related disciplines (as e.g. mathematics, computational linguistics, information sciences, computer engineering, etc.).

The candidate should master statistical learning methods, information retrieval and/or natural language processing fields.

We expect outstanding analytical competence, strong interest in interdisciplinary research (Machine learning/Information retrieval), experience in software engineering (strong programming skills in Python and ML/RL libraries), as well as superior organization and communication skills.

Fluency in spoken and written English is required.

## References

- [ACH<sup>+</sup>20] Alishke Anand, Lawrence Cavdon, Matthias Hagen, Hideo Joho, Mark Sanderson, and Benno Stein. Conversational search – a report from dagstuhl seminar 19461, 2020.
- [ASD18] Wafa Aissa, Laure Soulier, and Ludovic Denoyer. A reinforcement learning-driven translation model for search-oriented conversational systems. In *2nd International Workshop on Search-Oriented Conversational AI, SCAI@EMNLP*, pages 33–39. Association for Computational Linguistics, 2018.

- [BW16] Antoine Bordes and Jason Weston. Learning end-to-end goal-oriented dialog. *CoRR*, abs/1605.07683, 2016.
- [CTY20] Limin Chen, Zhiwen Tang, and Grace Hui Yang. Balancing reinforcement learning training experiences in interactive information retrieval. In *SIGIR 2020*, pages 1525–1528, 2020.
- [LCL<sup>+</sup>17] Xiujun Li, Yun-Nung Chen, Lihong Li, Jianfeng Gao, and Asli Celikyilmaz. End-to-end task-completion neural dialogue systems. In *8th International Joint Conference on Natural Language Processing*, pages 733–743, 2017.
- [MRC<sup>+</sup>20] Chuan Meng, Pengjie Ren, Zhumin Chen, Weiwei Sun, Zhaochun Ren, Zhaopeng Tu, and Maarten de Rijke. Dukenet: A dual knowledge interaction network for knowledge-grounded conversation. In *SIGIR 2020*, pages 1151–1160, 2020.
- [MZA20] Ali MontazerAlghaem, Hamed Zamani, and James Allan. A reinforcement learning framework for relevance feedback. In Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu, editors, *SIGIR 2020*, pages 59–68, 2020.
- [RC17] Filip Radlinski and Nick Craswell. A theoretical framework for conversational search. In *Conference on Conference Human Information Interaction and Retrieval, CHIIR '17*, pages 117–126, 2017.
- [RCD11] Alan Ritter, Colin Cherry, and William B. Dolan. Data-driven response generation in social media. In *Conference on Empirical Methods in Natural Language Processing, EMNLP '11*, 2011.
- [SDVM<sup>+</sup>17] Florian Strub, Harm De Vries, Jeremie Mary, Bilal Piot, Aaron Courville, and Olivier Pietquin. End-to-end optimization of goal-driven and visually grounded dialogue systems Harm de Vries. In *IJCAI*, 2017.
- [SST14] Laure Soulier, Chirag Shah, and Lynda Tamine. User-driven system-mediated collaborative information retrieval. In Shlomo Geva, Andrew Trotman, Peter Bruza, Charles L. A. Clarke, and Kalervo Järvelin, editors, *SIGIR 2014*, pages 485–494, 2014.
- [VLR<sup>+</sup>20] Nikos Voskarides, Dan Li, Pengjie Ren, Evangelos Kanoulas, and Maarten de Rijke. Query resolution for conversational search with limited supervision. In Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu, editors, *SIGIR 2020*, pages 921–930, 2020.
- [WVM<sup>+</sup>17] Tsung-Hsien Wen, David Vandyke, Nikola Mrkšić, Milica Gasic, Lina M. Rojas Barahona, Pei-Hao Su, Stefan Ultes, and Steve Young. A network-based end-to-end trainable task-oriented dialogue system. In *Association for Computational Linguistics (ACL): Volume 1*, pages 438–449, 2017.
- [XWX<sup>+</sup>20] Jun Xu, Zeng Wei, Long Xia, Yanyan Lan, Dawei Yin, Xueqi Cheng, and Ji-Rong Wen. Reinforcement learning to rank with pairwise policy gradient. In *SIGIR 2020*, pages 509–518, 2020.
- [YGTW13] Steve Young, Milica Gasic, Blaise Thomson, and Jason Williams. Pomdp-based statistical spoken dialogue systems: a review. In *Proceedings of the IEEE*, 2013.
- [ZE16] Tiancheng Zhao and Maxine Eskénazi. Towards end-to-end learning for dialog state tracking and management using deep reinforcement learning. In *Annual Meeting of the Special Interest Group on Discourse and Dialogue SIGDIAL*, pages 1–10, 2016.
- [ZZZ<sup>+</sup>20] Weinan Zhang, Xiangyu Zhao, Li Zhao, Dawei Yin, Grace Hui Yang, and Alex Beutel. Deep reinforcement learning for information retrieval: Fundamentals and advances. In Jimmy Huang, Yi Chang, Xueqi Cheng, Jaap Kamps, Vanessa Murdock, Ji-Rong Wen, and Yiqun Liu, editors, *SIGIR 2020*, pages 2468–2471, 2020.