ABSTRACT

Conversational agents, or commonly known as dialogue systems, have gained escalating popularity in recent years. Their widespread applications support conversational interactions with users and accomplishing various tasks as personal assistants. However, one key weakness in existing conversational agents is that they only learn to passively answer user queries via training on pre-collected and manually-labeled data. Such passiveness makes the interaction modeling and system-building process relatively easier, but it largely hinders the possibility of being human-like hence lowering the user engagement level. In this tutorial, we introduce and discuss methods to equip conversational agents with the ability to interact with end users in a more proactive way. This three-hour tutorial is divided into three parts and includes two interactive exercises. It reviews and presents recent advancements on the topic, focusing on automatically expanding ontology space, actively driving conversation by asking questions or strategically shifting topics, and retrospectively conducting response quality control.

KEYWORDS

Proactive conversational agents, task-oriented dialogue systems, conversational search, conversational AI

1 MOTIVATION & BRIEF OUTLINE

Building conversational agents capable of understanding human language and making proper responses to humans is a long-cherished goal of Artificial Intelligence (AI). Such agents can be broadly categorized into two main types: (1) Chit-chat systems that are designed to engage users and provide emotional support. It conducts chit-chat type of conversations on open topics without having any specific goal to complete, and (2) Task-oriented dialogue systems that are designed to assist users in completing certain tasks. Currently, most of the popular personal assistants, such as Amazon Alexa, Apple Siri, Google Home, and Microsoft Cortana, are task-oriented. They are primarily designed to take natural language verbal instructions or commands, understand them and translate them into actions to be executed by the underlying application systems.

Despite the increasing attention and widespread adaptations, most existing conversational agents can only perform pre-defined actions or offer factual information in response to a user’s command or query. They heavily rely on existing training conversations and exploit knowledge associated with these for exchanging information, recommending something, and completing tasks [21, 25, 63]. Such systems usually passively generate responses to questions instead of leading the conversation or asking questions [5, 50, 51]. This reactive, passive style of response constrains the types of conversations one could have. For example, situations involving exploratory search and complex decision-making where a conversation with both sides actively engaging is desired.

In the past few years, researchers in multiple fields, including natural language processing (NLP), dialogue systems, conversational search, conversational recommender systems, and machine learning [8, 12, 14, 35, 62, 64] have started to address the issue to enable conversational agents the ability for two-way, proactive conversations [30, 32, 52, 60, 65]. Various aspects or angles have been touched, such as proposing new intents and conversational slots via ontology expansion and actively analyzing failures in unseen situations [17, 18, 23, 27–29, 34, 40, 41, 48, 49, 59, 61], learning to raise questions to move the conversation forward [22, 36, 66, 67], leveraging task information and domain knowledge to guide purposeful topic shifting and exploration [26, 38, 39, 42, 56], and controlling quality of the response generation [2, 11, 33, 43, 46] as well as improving evaluation [16, 44] etc. In this tutorial, we review the literature on this emerging research topic, and introduce and discuss methods in a structured way to equip conversational agents with the ability to proactively interact with end users in performing various tasks.

This three-hour tutorial will be divided into three parts and two interactive exercises. The first part of the tutorial focuses on introducing the paradigm of proactive conversation agents and discusses system architecture design, the differences and challenges in realizing the proactive conversational AI paradigm. It includes an interactive exercise to engage the audience while emphasizing the important issues to address. In the second part, we present recent advancements on the topic, focusing on automatically expanding ontology space and how to drive conversation by asking questions actively. In the third part, we review and discuss methods that support strategic planning of topic shifting and retrospectively conducting response quality control. At the end of the lectures, we arrange a session of interactive exercises for the attendees to chat with a few conversational agents supported by major variations of the methods we show in the tutorial and test their own fixes to solve some of the issues that will be present. Finally, we conclude...
We observe a steadily increasing interest in conversational agents and present some open challenges for future research.

2 INTENDED AUDIENCE & PREREQUISITE
The intended audience includes researchers, graduate or senior undergraduate students, and industrial practitioners interested in conversational search, conversational recommendation, and task-oriented dialogue systems. This timely tutorial will particularly benefit those who are building or intend to build conversational agents for being proactive is becoming necessary for next-generation conversational agents. The tutorial will assume the audience’s basic knowledge and experience in machine learning, deep learning, information retrieval, and NLP.

3 RELEVANCE TO THE COMMUNITY
We observe a steadily increasing interest in conversational agents in recent IR, AI, and NLP conferences. For instance, there is a Conversational Assistant Track in TREC 2019 to 2022[8], and the primer IR conference, SIGIR, has recognized ‘Dialogue and Interactive systems’ to be one of its major research topics. Our tutorial focuses on building next-generation conversational agents that can interact with end users proactively and contribute more to the shared ongoing conversation. With the wide use of customer service conversational agents, task completion conversational agents, and virtual personal assistants etc. in our daily life, there has been urgent demand to make these conversational agents more proactive. As more and more usage of conversational interfaces in downstream tasks becomes prevalent in both commercial and daily use cases, we believe the content of this tutorial would be interesting, beneficial and important to both IR researchers and practitioners.

Similar Tutorials in Related Conferences: Several tutorials on general dialogue systems and conversational recommendation have been given in recent related conferences, such as Conversational Information Seeking [7] in SIGIR’22 and Conversational Recommendation Systems [13] in RecSys’20 and in WSDM’21. However, none of them had focused on the proactive aspect of conversational agents. With a completely different and new perspective, our tutorial expands the horizon of conversational agents in a timely fashion and presents the latest advances on topics that are complementary yet relevant to existing mainstream approaches (e.g., topic shifting, ontology expansion, out-of-distribution machine learning, inappropriate dialogue management, and conservative reinforcement learning). We are unaware of any other proactive conversational agents tutorials in WSDM or related conferences.

4 FORMAT AND DETAILED SCHEDULE
Our tutorial will mainly be lectures, with interactive exercises created for both in-person and online attendees. We will also maintain an online forum to host question-answering sessions and encourage discussions during and after the tutorial.

- Schedule of the Tutorial [3hrs]

Part One:
(1) Introduction ——————————————— [15 mins]
   (a) Dialogue & Interactive Systems: Background

(b) State-of-the-art conversational agents: Weaknesses
(c) Motivation & Aim of this tutorial

(2) Interactive Exercise-1 ——————————————— [15 mins]

(3) Proactive Conv. Agents: Overview ———————————— [25 mins]
   (a) Existing conversational agents architectures
   (b) Proactive extensions
   (c) Applications: Search, recommend, and more [15, 54]

(d) Evaluation: Metrics, procedures, and challenges [16, 44]

Part Two:
(4) Proactive Ontology Expansion ———————————— [25 mins]
   (a) New intent discovery [28, 61]
   (b) Novel slot induction [24, 49, 53, 55, 59]
   (c) Reflecting experiences for unseen situations [58]

(5) Learning to Ask ———————————— [25 mins]
   (a) Mixed initiatives [19, 47]
   (b) Learning to ask [22, 36, 66, 67]
   (c) Question selection and generation [3, 9, 57]

Part Three:
(6) Purposeful Topic Shifting ———————————— [25 mins]
   (a) Target-guided topic shifting [26, 42, 56]
   (b) Counterfactual data augmentation [4, 19, 43, 45]
   (c) Out-of-distribution label generation [1, 31]

(7) Response Quality Control ———————————— [25 mins]
   (a) Types of inappropriate responses [10]
   (b) Language detoxification [2, 11, 33]
   (c) Conservative reinforcement learning [4, 6, 20, 37, 43, 46]

(8) Interactive Exercise-2 ———————————— [15 mins]

Ending:
(9) Concluding Remarks ———————————— [10 mins]
   (a) Summary
   (b) Challenges & Future directions

As listed in the scheduled above, we will have two interactive exercises. In the first one that takes place during Part One, we will instruct the participants to pair up. One of them in each pair will act as a user and the other one will act as an agent. We will then assign the users a task that requires gathering information and making a decision. An example task would be planning a trip to Singapore. The person acting as the agent can respond to a question coming from the user by looking up relevant information (e.g., from a search engine or a website) and constructing a short answer, similar to a typical conversational assistant will do. We will spend about 5 minutes setting up the exercise, 5 minutes having the participants do the role-play, and 5 minutes to come back and share what we discovered. This exercise will do two things: (1) it will allow us to see some of the limitations of the current systems that will set the
stage for the next part of the tutorial; and (2) it will set up a baseline for the next interactive exercise.

The second interactive exercise will take place in Part Three, almost toward the end. By this time, the participants will have learned various methods that can allow a conversational agent to be proactive. We will now repeat the exercise from before, but this time asking the persons playing the agent to use appropriate technique (e.g., ontology expansion, topic shifting) while constructing their responses. Once again, we will come back and share our experiences. This will allow everyone to not only practice what they learned in the tutorial, albeit in a pseudo-system fashion, but also start thinking beyond the specific techniques to also account for human experiences in a proactive conversation.

- Reading List
Please see the References section for the Reading list.

5 SUPPORT MATERIALS
We will release our presentation slides, exercise questions and answers, and reference lists to the attendees. We also agree to allow video recording of the tutorial and make the videos available to the tutorial attendees. During and after the tutorial, we will collect questions from the attendees, post our answers to their questions, and maintain a discussion forum during the entire conference.

6 PARTICIPATION AND ENGAGEMENT
Weeks before the conference, we will promote the tutorial on social media and on professional email lists (e.g., the SIGIR mailing list). We will set up and maintain a tutorial website and promote the content there, too. We plan to contact colleagues in multiple research areas, including NLP, AI, Speech, Dialogue Systems, ML, Recommender systems, and IR, to forward the news about the tutorial. During the tutorial, we will have multiple sessions of interactive exercises with the attendees. We will also actively solicit questions from the attendees and post our answers to the questions in a forum (e.g., piazza.com) that will be created specifically for this tutorial. We will maintain the forum during the entire conference. The goal is to create the discussion and keep the communications going. For vivid presentations and hands-on experiences, conversational devices, such as Amazon Alexa and Apple Siri, will be installed in the tutorial room. They will be used in some examples in our slides and in interactive exercises with the attendees.

7 CONTACT INFORMATION OF PRESENTERS
(1) Lizi Liao (Main Contact), Singapore Management University lzliao@smu.edu.sg
(2) Grace Hui Yang, Georgetown University grace.yang@georgetown.edu
(3) Chirag Shah, University of Washington chirags@uw.edu

8 BIOS OF PRESENTERS
Dr. Lizi Liao is Assistant Professor of Computer Science at Singapore Management University, leading the CoAgent group. She obtained her Ph.D. from National University of Singapore in 2019. Dr. Liao’s research interests center on task-oriented dialogues, proactive conversational agents, and multimodal conversational search and recommendation as the application target. She publishes regularly in top AI and Data Science conferences and journals such as SIGIR, WWW, ACM MM, TKDE, TACL etc. One of her work was nominated in the ACM MM Best Paper Final List in 2018. She serves as (senior) PC member of these prestigious conferences and reviewer of the top journals. She also serves as organizing committee members of ACM MM 2019, NLPC 2022 and WSDM 2023, while chairs sessions at KDD 2021 and SIGIR 2022.
Dr. Grace Hui Yang is Associate Professor of Computer Science at Georgetown University, Washington D.C., U.S., leading the InfoSense (Information Retrieval and Sense-Making) group. Dr. Yang obtained her Ph.D. from Carnegie Mellon University in 2011. Her current research interests include conversational agents, deep reinforcement learning, and information retrieval. Dr. Yang is a recipient of the U.S. National Science Foundation Faculty Early Career Development Program (CAREER) Award. She has served on the organizing or program committees in recent conferences, such as SIGIR, ECIR, ACL, AAAI, ICTIR, CIKM, WSDM, and WWW, and on the editorial boards of ACM TOIS (2022-) and Information Retrieval Journal (2014-2017). Dr. Yang has experience conducting tutorials in the WSDM conferences – she delivered tutorials on Dynamic Information Retrieval at SIGIR’14 and WSDM’15, and on Privacy-Preserving IR at ICTIR’17 and WSDM’18. Dr. Yang also taught in the first ACM SIGIR AFIRM Summer School in 2019.
Dr. Chirag Shah is Professor at University of Washington in Seattle. He has taught undergraduate and graduate courses in IR, HCI, Data Science, and Machine Learning. He has also taught several courses and tutorials on topics related to web search and recommender systems at different international places, including at conferences such as RecSys, SIGIR, and WSDM, Russian Summer School on Information Retrieval (RuSsIR), and Asian Summer School in Information Access (ASSIA). He has developed Coursera course on Social Media Data Analytics, and a SAGE course on Machine Learning for Data Science. He has authored multiple books, including a textbook on Data Science (2020) and soon to be published textbook on Machine Learning.

REFERENCES


