SereTOD 2022

Towards Semi-Supervised and Reinforced Task-Oriented Dialog Systems

Proceedings of the Workshop

December 7, 2022

The SereTOD organizers gratefully acknowledge the support from the following sponsors.

Gold

Order copies of this and other ACL proceedings from:

Association for Computational Linguistics (ACL)
209 N. Eighth Street
Stroudsburg, PA 18360
USA
Tel: +1-570-476-8006
Fax: +1-570-476-0860
acl@aclweb.org

ISBN 978-1-959429-21-0

Introduction

Welcome to the Workshop - Towards Semi-Supervised and Reinforced Task-Oriented Dialog Systems (SereTOD), co-located with EMNLP 2022!

Task-oriented dialog (TOD) systems are designed to assist users to accomplish their goals. Recently, neural generative approaches have received increasing attention. Unfortunately, building TOD systems remains as a label-intensive and time-consuming task. The process still heavily relies on manually labeled dialog data and annotated task-related knowledge base. However, unlabeled data are often easily available in many forms such as human-to-human dialogs, open-domain text corpus, and unstructured knowledge documents. The purpose of this Workshop is to invite researchers from both academia and industry to share their perspectives on building semi-supervised and reinforced TOD systems, discuss challenges and advance the field in joint effort.

In parallel, we open up a challenge, in which we collect and share a newly released, large-scale, human-human dialog dataset, called the MobileCS (Mobile Customer Service) dataset to foster this line of research. The Challenge consists of two tracks: Information extraction from dialog transcripts (Track 1), and Task-oriented dialog systems (Track 2). Congratulations to the 15 teams, who submitted effective results, out of the total of 62 teams registered for the SereTOD Challenge!

We received submissions from all levels of methodologies, algorithms, models, system developments, applications and datasets towards semi-supervised and reinforced TOD systems. Given the high-quality submissions received and the capacity of the Workshop, the selection process was very competitive. We accepted 11 papers accounting for 47% of the submissions. Further, authors of a total of 13 Findings papers on Dialog have confirmed to present at the Workshop (as nonarchival presentations). In total, we have 24 papers included in the program, splitting into 4 oral sessions and 1 poster session.

SereTOD Workshop is co-located with EMNLP on December 7, 2022 (virtually with EMNLP main venue and on-site in Beijing). In additional to the paper presentations, the program also features 3 invited talks, a panel, as well as awards for the SereTOD Challenge.

We would like to take this opportunity to thank the Program Committee for their support and thorough reviews. We are deeply honored to have excellent talks from our invited speakers - Pascale Fung, Dilek Hakkani-Tur, and Jason Williams. We are especially thankful for the support from Joint Institute of Tsinghua University - China Mobile Communications Group Co. Ltd. Finally, we are grateful for the extensive help from EMNLP 2022 workshop co-chairs, Daniel Hershcovich and Asli Celikyilmaz.

We sincerely hope you will enjoy a memorable SereTOD Workshop!

The SereTOD Workshop General Chairs,
Zhijian Ou, Tsinghua University
Junlan Feng, China Mobile
Juanzi Li, Tsinghua University

Organizing Committee

Organizers

Zhijian Ou, Tsinghua University
Junlan Feng, China Mobile
Juanzi Li, Tsinghua University

Program Committee

Reviewers

Lu Chen, Shanghai Jiaotong University
He He, New York University
Yi Huang, China Mobile Research
Bing Liu, Facebook
Yang Liu, Amazon
Verena Rieser, Heriot-Watt University
Rui Yan, Renmin University of China
Koichiro Yoshino, RIKEN
Kai Yu, Shanghai Jiaotong University
Jiangjiang Zhao, China Mobile Online Services

Keynote Talk: Ingesting Knowledge from Diverse Sources to Open Domain Social Conversations

Dilek Hakkani-Tur
Amazon

Abstract: Following the recent advancements in language modeling and availability of large natural language datasets, the last decade has been flourishing for conversational AI research. The progress also helped emphasize the importance of reasoning over a diverse set of external knowledge and task completion resources for forming relevant, informative, and accurate responses, discussing with the users when the available solutions/information are not sufficient, and making proactive suggestions. For ingesting knowledge in conversations, recent work has mainly grounded conversational responses on knowledge snippets from wikipedia and web documents, with the goals of preventing hallucination and providing users diverse and accurate responses. However, much of the world's knowledge is dynamic and it is spread across diverse resources. Some of these are already structured, such as knowledge graphs. But a majority of them are not structured, for example, news articles and books. And some of them also include subjective information, such as customer reviews. In this talk, I will discuss our recent work on integrating knowledge to conversation responses from such a diverse set of resources, challenges associated with these, and progress we made so far.

Bio: Dilek Hakkani-Tür is a senior principal scientist at Amazon Alexa AI focusing on enabling natural dialogues with machines. Prior to joining Amazon, she was a researcher at Google Research, Microsoft Research, International Computer Science Institute at University of California, Berkeley, and AT&T Labs - Research. She received her BSc degree from Middle East Technical Univ., and MSc and PhD degrees from Bilkent Univ., Department of Computer Engineering. Her research interests include conversational AI, natural language and speech processing, spoken dialogue systems, and machine learning for language processing. She has over 80 patents that were granted and co-authored more than 300 papers in natural language and speech processing. She received several best paper awards for publications she co-authored on conversational systems, from IEEE Signal Processing Society, ISCA, EURASIP and others. She served as an associate editor for IEEE Transactions on Audio, Speech and Language Processing (2005-2008), a member of the IEEE Speech and Language Technical Committee (2009-2014), an area editor for speech and language processing for Elsevier's Digital Signal Processing Journal and IEEE Signal Processing Letters (2011-2013), the Editor-in-Chief of the IEEE/ACM Transactions on Audio, Speech and Language Processing (2018-2021), and an IEEE Distinguished Industry Speaker (2021). She also served on the ISCA Advisory Council (2015-2019) and the IEEE Signal Processing Society Fellows Committee (2019-2022). She was elected as a fellow of the IEEE (2014) and ISCA (2014).

Keynote Talk: Insights on the relationship between usage frequency, user proficiency, and interaction quality for a virtual assistant

Jason D. Williams
Apple

Abstract: For a virtual assistant, it seems clear that users who have a higher-quality experience would tend to use the assistant more. But causality is less obvious — for example, does higher usage frequency result from higher-quality interactions, or is higher usage frequency a reflection of higher user proficiency? How does user proficiency change over time? In this talk I'll cover a quantitative investigation into the relationships between usage frequency, user proficiency, and interaction quality for a real-world virtual assistant. The insights from this study may help inform reward or loss functions for virtual assistants optimized with reinforcement or semi-supervised learning. This is joint work with colleagues Zidi Xiu, Kai-Chen Cheng, David Q. Sun, Jiannan Lu, Hadas Kotek, Paul McCarthy, Yuhan Zhang, Christopher Klein, and Stephen Pulman.

Bio: Jason D. Williams leads a team that builds language understanding for Siri at Apple, where he has been since 2018. Prior to Apple, he was a Research Manager at Microsoft Research, leading research groups on conversational systems and reinforcement learning. Jason has published over 60 peer-reviewed papers on dialog systems and related areas, with over 8,000 citations and five best paper/presentation awards. Jason initiated the Dialog State Tracking Challenge series in 2012; shipped components of the first release of Microsoft Cortana in 2014; and launched Microsoft's Language Understanding Service (www.luis.ai (http://www.luis.ai/)) in 2015. Jason has previously served as an elected member of the IEEE Speech and Language Technical Committee (SLTC) in the area of spoken dialogue systems for 3 terms, President of SIGDIAL, senior area chair at ACL and EMNLP, and general chair and technical chair of IEEE ASRU.

Keynote Talk: Responsible & Empathetic Human Robot Interactions

Pascale Fung
Hong Kong University of Science & Technology

Abstract: Conversational AI (ConvAI) systems have applications ranging from personal assistance, health assistance to customer services. They have been in place since the first call centre agent went live in the late 1990s. More recently, smart speakers and smartphones are powered with conversational AI with similar architecture as those from the 90s. On the other hand, research on ConvAI systems has made leaps and bounds in recent years with sequence-to-sequence, generation-based models. Thanks to the advent of large scale pre-trained language models, state-of-the-art ConvAI systems can generate surprisingly human-like responses to user queries in open domain conversations, known as chit-chat. However, these generation based ConvAI systems are difficult to control and can lead to inappropriate, biased and sometimes even toxic responses. In addition, unlike previous modular conversational AI systems, it is also challenging to incorporate external knowledge into these models for task-oriented dialog scenarios such as personal assistance and customer services, and to maintain consistency. In this talk, I will introduce state-of-the-art generation based conversational AI approaches, and will point out remaining challenges of conversational AI and possible directions for future research, including how to mitigate inappropriate responses. I will also present some ethical guidelines that conversational AI systems can follow.

Bio: Pascale Fung is a Professor at the Department of Electronic & Computer Engineering and Department of Computer Science & Engineering at The Hong Kong University of Science & Technology (HKUST). Prof. Fung received her PhD in Computer Science from Columbia University in 1997. She worked and studied at AT&T Bell Labs (1993 1997), BBN Systems & Technologies (1992), LIMSI, CNRS, France (1991), Department of Information Science, Kyoto University, Japan (1989 1991), and at Ecole Centrale Paris, France(1988). She is an elected Fellow of the Association for Computational Linguistics (ACL) for her significant contributions towards statistical NLP, comparable corpora, and building intelligent systems that can understand and empathize with humans. She is an Fellow of the Institute of Electrical and Electronic Engineers (IEEE) for her contributions to human-machine interactions and an elected Fellow of the International Speech Communication Association for fundamental contributions to the interdisciplinary area of spoken language human-machine interactions. She served as Editor and Associate Editor for Computer Speech and Language, IEEE ACM Transactions on Audio, Speech and Language Processing, Transactions for ACL, IEEE Signal Processing Letters. She served as a Committee Member of the IEEE Signal Processing Society Speech and Language Technology Committee (SLTC) for six years. She is a past president t and a Board Member of the ACL Special Interest Group on Linguistics Data and Corpus Based Approaches in NLP (SIGDAT).

Table of Contents

Program

Wednesday, December 7, 2022

04:50 - 05:00 *Opening Remarks*

05:00 - 05:40 *Invited Talk 1 - Dilek Hakkani-Tur*

05:40 - 06:00 *Break*

06:00 - 07:00 *Oral Session 1 (Semi-Supervised Dialogue Systems)*

Semi-Supervised Knowledge-Grounded Pre-training for Task-Oriented Dialog Systems
Weihao Zeng, Keqing He, Zechen Wang, Dayuan Fu, Guanting Dong, Ruotong Geng, Pei Wang, Jingang Wang, Chaobo Sun, Wei Wu and Weiran Xu

Prompt Learning for Domain Adaptation in Task-Oriented Dialogue
Makesh Narsimhan Sreedhar and Christopher Parisien

DIALOGIC: Controllable Dialogue Simulation with In-Context Learning
Zekun Li, Wenhu Chen, Shiyang Li, Hong Wang, Jing Qian and Xifeng Yan

Robust Task-Oriented Dialogue Generation with Contrastive Pre-training and Adversarial Filtering
Shiquan Yang, Xinting Huang, Jey Han Lau and Sarah Erfani

07:00 - 08:00 *Oral Session 2 (Information Extraction and Knowledge-Grounded Dialogue Systems)*

Information Extraction and Human-Robot Dialogue towards Real-life Tasks A Baseline Study with the MobileCS Dataset
Hong Liu, Hao Peng, Zhijian Ou, Juanzi Li, Yi Huang and Junlan Feng

Explainable Slot Type Attentions to Improve Joint Intent Detection and Slot Filling
Kalpa Gunaratna, Vijay Srinivasan, Akhila Yerukola and Hongxia Jin

Topic-Aware Response Generation in Task-Oriented Dialogue with Unstructured Knowledge Access
Yue Feng, Gerasimos Lampouras and Ignacio Iacobacci

Doc2Bot: Accessing Heterogeneous Documents via Conversational Bots
Haomin Fu, Yeqin Zhang, Haiyang Yu, Jian Sun, Fei Huang, Luo Si, Yongbin Li and Cam-Tu Nguyen

Wednesday, December 7, 2022 (continued)

08:00 - 09:00 *Lunch*

09:00 - 09:40 *Invited Talk 2 - Jason Williams*

09:40 - 10:40 *Oral Session 3 (Reinforced Dialogue Systems)*

A Generative User Simulator with GPT-based Architecture and Goal State Tracking for Reinforced Multi-Domain Dialog Systems
Hong Liu, Yucheng Cai, Zhijian Ou, Yi Huang and Junlan Feng

Offline-to-Online Co-Evolutional User Simulator and Dialogue System
Dafeng Chi, Yuzheng Zhuang, Yao Mu, Bin Wang, Jianzhu Bao, Yasheng Wang, Yuhan Dong, Xin Jiang, Qun Liu and Jianye Hao

State-Aware Adversarial Training for Utterance-Level Dialogue Generation
Yi Huang, Xiaoting Wu, Wei Hu, Junlan Feng and Chao Deng

Is MultiWOZ a Solved Task? An Interactive TOD Evaluation Framework with User Simulator
Qinyuan Cheng, Linyang Li, Guofeng Quan, Feng Gao, Xiaofeng Mou and Xipeng Qiu

10:40 - 11:00 *Break*

11:00 - 11:40 *Invited Talk 3 - Pascale Fung*

11:40 - 12:40 *Oral Session 4 (Dialogue Datasets)*

CMCC: A Comprehensive and Large-Scale Human-Human Dataset for Dialogue Systems
Yi Huang, Xiaoting Wu, Si Chen, Wei Hu, Qing Zhu, Junlan Feng, Chao Deng, Zhijian Ou and Jiangjiang Zhao

DOROTHIE: Spoken Dialogue for Handling Unexpected Situations in Interactive Autonomous Driving Agents
Ziqiao Ma, Ben VanDerPloeg, Cristian-Paul Bara, Yidong Huang, Eui-In Kim, Felix Gervits, Matthew Marge and Joyce Chai

AssistSR: Task-oriented Video Segment Retrieval for Personal AI Assistant
Stan Weixian Lei, Difei Gao,, Yuxuan Wang, Dongxing Mao, Zihan Liang, Lingmin Ran and Mike Zheng Shou

DialogUSR: Complex Dialogue Utterance Splitting and Reformulation for Multiple Intent Detection
Haoran Meng, Xin Zheng, Tianyu Liu, Zizhen Wang, He Feng, Binghuai Lin, Xuemin Zhao, Yunbo Cao and Zhifang Sui

12:40 - 13:40 *Poster Session*

A GlobalPointer based Robust Approach for Information Extraction from Dialog Transcripts
Yanbo J. Wang, Sheng Chen, Hengxing Cai, Wei Wei, Kuo Yan, Zhe Sun, Hui Qin, Yuming Li and Xiaochen Cai

A Token-pair Framework for Information Extraction from Dialog Transcripts in SereTOD Challenge
Chenyue Wang, Xiangxing Kong, Mengzuo Huang, Feng Li, Jian Xing, Weidong Zhang and Wuhe Zou

Disentangling Confidence Score Distribution for Out-of-Domain Intent Detection with Energy-Based Learning
Yanan Wu, Zhiyuan Zeng, Keqing He, Yutao Mou, Pei Wang, Yuanmeng Yan and Weiran Xu

Oh My Mistake!: Toward Realistic Dialogue State Tracking including Turnback Utterances
Takyoung Kim, Yukyung Lee, Hoonsang Yoon, Pilsung Kang, Junseong Bang and Misuk Kim

History-Aware Hierarchical Transformer for Multi-session Open-domain Dialogue System
Tong Zhang, Yong Liu, Boyang Li, Zhiwei Zeng, Pengwei Wang, Yuan You, Chunyan Miao and Lizhen Cui

Modeling Complex Dialogue Mappings via Sentence Semantic Segmentation Guided Conditional Variational Auto-Encoder
Bin Sun, Shaoxiong Feng, Yiwei Li, Weichao Wang, Fei Mi, Yitong Li and Kan Li

Diving Deep into Modes of Fact Hallucinations in Dialogue Systems
Souvik Das, Sougata Saha, Yiwei Li and Rohini Srihari

Keep Me Updated! Memory Management in Long-term Conversations
Sanghwan Bae, Donghyun Kwak, Soyoung Kang, Min Young Lee, Sungdong Kim, Yuin Jeong, Hyeri Kim, Sang-Woo Lee, Woomyoung Park and Nako Sung

13:40 - 13:50 *SereTOD Challenge Awards*

13:50 - 14:30 *Panel, Closing*