ACL 2022

The 4th Workshop on NLP for Conversational AI

Proceedings of the Workshop

May 27, 2022

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

Gold Level

Microsoft

Bronze Level



©2022 Association for Computational Linguistics

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 978-1-955917-46-9

Introduction

Welcome to the 4th Workshop on NLP for Conversational AI, at ACL 2022.

Ever since the invention of the intelligent machine, hundreds and thousands of mathematicians, linguists, and computer scientists have dedicated their careers to empowering human-machine communication in natural language. Although the idea is finally around the corner with a proliferation of virtual personal assistants such as Siri, Alexa, Google Assistant, and Cortana, the development of these conversational agents remains difficult and there still remain plenty of unanswered questions and challenges.

Conversational AI is hard because it is an interdisciplinary subject. Initiatives were started in different research communities, from Dialogue State Tracking Challenges to NeurIPS Conversational Intelligence Challenge live competition and the Amazon Alexa Prize. However, various fields within the NLP community, such as semantic parsing, coreference resolution, sentiment analysis, question answering, and machine reading comprehension etc. have been seldom evaluated or applied in the context of conversational AI.

The goal of this workshop is to bring together NLP researchers and practitioners in different fields, alongside experts in speech and machine learning, to discuss the current state-of-the-art and new approaches, to share insights and challenges, to bridge the gap between academic research and real-world product deployment, and to shed light on future directions. "NLP for Conversational AI" will be a one-day workshop including keynotes, spotlight talks, and poster sessions. In keynote talks, senior technical leaders from industry and academia will share insights on the latest developments in the field.

An open call for papers will be announced to encourage researchers and students to share their prospects and latest discoveries. The panel discussion will focus on the challenges, future directions of conversational AI research, bridging the gap in research and industrial practice, as well as audience suggested topics.

With the increasing trend of conversational AI, NLP4ConvAI 2022 is competitive. We received 45 submissions directly to the workshop and 14 submissions through the ACL Rolling Review. After a rigorous review process, we only accepted 18 papers. There are 15 long papers and 3 short papers. The workshop overall acceptance rate is about 30.5%.

We hope you will enjoy NLP4ConvAI 2022 at ACL and contribute to the future success of our community!

NLP4ConvAI 2021 Organizers

Bing Liu, Meta Alexandros Papangelis, Amazon Alexa AI Stefan Ultes, Mercedes-Benz AG Abhinav Rastogi, Google Research Yun-Nung (Vivian) Chen, National Taiwan University Georgios Spithourakis, PolyAI Elnaz Nouri, Microsoft Research Weiyan Shi, Columbia University

Organizing Committee

General Chair

Bing Liu, Meta Alexandros Papangelis, Amazon Alexa AI

Program Chair

Stefan Ultes, Mercedes-Benz AG Abhinav Rastogi, Google Research

Publication Chair

Yun-Nung (Vivian) Chen, National Taiwan University

Diversity Chair

Georgios Spithourakis, PolyAI

Sponsorship Chair

Elnaz Nouri, Microsoft Research

Publicity Chair

Weiyan Shi, Columbia University

Program Committee

Program Chairs

Yun-Nung (Vivian) Chen, National Taiwan University Bing Liu, Meta Elnaz Nouri, Microsoft Research Alexandros Papangelis, Amazon Abhinav Rastogi, Google Weiyan Shi, Columbia University Georgios P. Spithourakis, PolyAI Stefan Ultes, Mercedes Benz Research & Development

Program Committee

Akshat Shrivastava, Meta Alankar Jain, Google Alborz Geramifard, Meta Bin Zhang, Google Bo-Hsiang Tseng, University of Cambridge Chao-Wei Huang, National Taiwan University Chien-Sheng Wu, Salesforce AI Research Chinnadhurai Sankar, Meta Christian Geishauser, Heinrich-Heine Universität Düsseldorf Daniel Cer, Google David Vandyke, University of Cambridge Dilek Hakkani-Tur, Amazon Alexa AI Emine Yilmaz, Department of Computer Science, University College London Evgeniia Razumovskaia, University of Cambridge Gokhan Tur, Amazon Harrison Lee, Google Hongyuan Zhan, Meta Hsien-chin Lin, Heinrich Heine University Düsseldorf Inigo Casanueva, PolyAI Ivan Vulić, University of Cambridge Kai Sun, Meta Lei Shu, Amazon Marek Rei, Imperial College London Michael Heck, Heinrich Heine Universität Düsseldorf Mihir Kale, Google Nikita Moghe, University of Edinburgh Paweł Budzianowski, PolyAI Peng Wang, Google Research Raghav Gupta, Google Seokhwan Kim, Amazon Seungwhan Moon, Meta Shikib Mehri, Carnegie Mellon University Shutong Feng, Heinrich-Heine Universität Düsseldorf Simon Keizer, Toshiba Research Europe Songbo Hu, Language Technology Lab, University of Cambridge

Stefan Larson, SkySync Ta-Chung Chi, Carnegie Mellon University Wei Peng, Huawei Technologies Ltd. Wenqiang Lei, Sichuan University Wolfgang Maier, Mercedes Benz Research and Development Xiujun Li, University of Washington Yang Liu, Amazon Yi-Chia Wang, Meta Yi-Ting Yeh, Carnegie Mellon University Yuan Cao, Google Brain Zhaojiang Lin, Meta Zhiguang Wang, Meta

Invited Speakers

Gokhan Tur, Amazon Alexa AI Zhou Yu, Columbia University William Wang, University of California, Santa Barbara Michael Tjalve, University of Washington + Microsoft Philanthropies Maria-Georgia Zachari, Omilia

Keynote Talk: HybriDialogue: Towards Information-Seeking Dialogue Reasoning Grounded on Tabular and Textual Data

William Wang

University of California, Santa Barbara

Abstract: A pressing challenge in current dialogue systems is to successfully converse with users on topics with information distributed across different modalities. Previous work in multi-turn dialogue systems has primarily focused on either text or table information. In more realistic scenarios, having a joint understanding of both is critical as knowledge is typically distributed over both unstructured and structured forms. In this talk, I will present a new dialogue dataset, HybriDialogue, which consists of crowdsourced natural conversations grounded on both Wikipedia text and tables. The conversations are created through the decomposition of complex multihop questions into simple, realistic multiturn dialogue interactions. We conduct several baseline experiments, including retrieval, system state tracking, and dialogue response generation. Our results show that there is still ample opportunity for improvement, demonstrating the importance of building stronger dialogue systems that can reason over the complex setting of information-seeking dialogue grounded on tables and text. I will also briefly mention a few related studies on dialogue research from the UCSB NLP Group.

Keynote Talk: Dialog Management for Conversational Task-Oriented Industry Solutions

Maria-Georgia Zachari Omilia

Abstract: This talk will focus on how the Omilia Cloud Platform[®] leverages the notion of Dialog Act in order to solve real-life use cases in task-oriented dialog systems for call centers. We will address the challenge of completing tasks efficiently, achieving high KPIs and integrating with a call center, while at the same time building and maintaining a flexible conversational NLU system.

Keynote Talk: Directions of Dialog Research in the Era of Big Pre-training Models

Zhou Yu

Columbia University

Abstract: Big pre-training models (such as BERT and GPT3) have demonstrated excellent performances on various NLP tasks. Instruction tuning and prompting have enabled these models to shine in low-resource settings. The natural question is "Will big models solve dialog tasks?" This talk will first go through big models' impact on several sub-topics within dialog systems (e.g. social chatbots, task-oriented dialog systems, negotiation/persuasion dialog systems, continue learning in dialog systems, multillingual dialog systems, multimodal dialog systems, deployable dialog systems, etc) and then follow up with the speaker's own interpretations of the challenges remaining and possible future directions.

Keynote Talk: Scaling impact: the case for humanitarian NLP

Michael Tjalve

University of Washington + Microsoft Philanthropies

Abstract: Advances in core NLP capabilities have enabled an extensive variety of scenarios where conversational AI provides real value for companies and customers alike. Leveraging lessons learned from these successes to applying the technology in the humanitarian context requires an understanding of both the potential for impact and risk of misuse.

In this talk, we'll discuss how to leverage conversational AI to scale impact for audiences in the humanitarian sector while earning and maintaining trust with the adopters of the technology and with the people they impact.

Keynote Talk: Past, Present, Future of Conversational AI

Gokhan Tur

Amazon Alexa AI

Abstract: Recent advances in deep learning based methods for language processing, especially using self-supervised learning methods resulted in new excitement towards building more sophisticated Conversational AI systems. While this is partially true for social chatbots or retrieval-based applications, it is commonplace to see dialogue processing as yet another task while assessing these new state of the art approaches. In this talk, I will argue that Conversational AI comes with an orthogonal methodology for machine learning to complement such methods interacting with the users using implicit and explicit signals. This is an exceptional opportunity for Conversational AI research moving forward and I will present couple representative efforts from Alexa AI.

Table of Contents

A Randomized Link Transformer for Diverse Open-Domain Dialogue Generation Jing Yang Lee, Kong Aik Lee and Woon Seng Gan 1
Are Pre-trained Transformers Robust in Intent Classification? A Missing Ingredient in Evaluation of Out-of-Scope Intent Detection Jianguo Zhang, Kazuma Hashimoto, Yao Wan, Zhiwei Liu, Ye Liu, Caiming Xiong and Philip S. Yu
Conversational AI for Positive-sum Retailing under Falsehood Control Yin-Hsiang Liao, Ruo-Ping Dong, Huan-Cheng Chang and Wilson Ma21
<i>D-REX: Dialogue Relation Extraction with Explanations</i> Alon Albalak, Varun R. Embar, Yi-Lin Tuan, Lise Getoor and William Yang Wang
Data Augmentation for Intent Classification with Off-the-shelf Large Language ModelsGaurav Sahu, Pau Rodriguez, Issam H. Laradji, Parmida Atighehchian, David Vazquez andDzmitry Bahdanau47
<i>Extracting and Inferring Personal Attributes from Dialogue</i> Zhulin Wang, Xuhui Zhou, Rik Koncel-Kedziorski, Alex Marin and Fei Xia
<i>From Rewriting to Remembering: Common Ground for Conversational QA Models</i> Marco Del Tredici, Xiaoyu Shen, Gianni Barlacchi, Bill Byrne and Adrià de Gispert70
Human Evaluation of Conversations is an Open Problem: comparing the sensitivity of various methods for evaluating dialogue agents Eric Michael Smith, Orion Hsu, Rebecca Qian, Stephen Roller, Y-Lan Boureau and Jason E Weston 77
KG-CRuSE: Recurrent Walks over Knowledge Graph for Explainable Conversation Reasoning using Semantic Embeddings Rajdeep Sarkar, Mihael Arcan and John Philip McCrae
Knowledge Distillation Meets Few-Shot Learning: An Approach for Few-Shot Intent Classification Within and Across Domains Anna Sauer, Shima Asaadi and Fabian Küch
<i>MTL-SLT: Multi-Task Learning for Spoken Language Tasks</i> Zhiqi Huang, Milind Rao, Anirudh Raju, Zhe Zhang, Bach Bui and Chul Lee
Multimodal Conversational AI: A Survey of Datasets and Approaches Anirudh S Sundar and Larry Heck 131
<i>Open-domain Dialogue Generation: What We Can Do, Cannot Do, And Should Do Next</i> Katharina Kann, Abteen Ebrahimi, Joewie J. Koh, Shiran Dudy and Alessandro Roncone 148
Relevance in Dialogue: Is Less More? An Empirical Comparison of Existing Metrics, and a Novel Simple Metric Ian Berlot-Attwell and Frank Rudzicz
<i>RetroNLU: Retrieval Augmented Task-Oriented Semantic Parsing</i> Vivek Gupta, Akshat Shrivastava, Adithya Sagar, Armen Aghajanyan and Denis Savenkov 184

Stylistic Response Generation by Controlling Personality Traits and Intent Sougata Saha, Souvik Das and Rohini Srihari
<i>Toward Knowledge-Enriched Conversational Recommendation Systems</i> Tong Zhang, Yong Liu, Boyang Li, Peixiang Zhong, Chen Zhang, Hao Wang and Chunyan Miao 212
Understanding and Improving the Exemplar-based Generation for Open-domain Conversation Seungju Han, Beomsu Kim, Seokjun Seo, Enkhbayar Erdenee and Buru Chang

Program

Friday, May 27, 2022

- 09:00 09:10 Opening Remarks
- 09:10 09:40 Invited Talk 1 by William Wang
- 09:40 10:10 Invited Talk 2 by Maria-Georgia Zachari
- 10:10 10:40 Oral Paper Session 1

Understanding and Improving the Exemplar-based Generation for Open-domain Conversation

Seungju Han, Beomsu Kim, Seokjun Seo, Enkhbayar Erdenee and Buru Chang

Conversational AI for Positive-sum Retailing under Falsehood Control Yin-Hsiang Liao, Ruo-Ping Dong, Huan-Cheng Chang and Wilson Ma

- 10:40 11:00 *Coffee Break*
- 11:00 12:30 Poster Paper Session

Extracting and Inferring Personal Attributes from Dialogue Zhulin Wang, Xuhui Zhou, Rik Koncel-Kedziorski, Alex Marin and Fei Xia

From Rewriting to Remembering: Common Ground for Conversational QA Models

Marco Del Tredici, Xiaoyu Shen, Gianni Barlacchi, Bill Byrne and Adrià de Gispert

D-REX: Dialogue Relation Extraction with Explanations

Alon Albalak, Varun R. Embar, Yi-Lin Tuan, Lise Getoor and William Yang Wang

A Randomized Link Transformer for Diverse Open-Domain Dialogue Generation Jing Yang Lee, Kong Aik Lee and Woon Seng Gan

Knowledge Distillation Meets Few-Shot Learning: An Approach for Few-Shot Intent Classification Within and Across Domains Anna Sauer, Shima Asaadi and Fabian Küch

Relevance in Dialogue: Is Less More? An Empirical Comparison of Existing Metrics, and a Novel Simple Metric Ian Berlot-Attwell and Frank Rudzicz

Friday, May 27, 2022 (continued)

Stylistic Response Generation by Controlling Personality Traits and Intent Sougata Saha, Souvik Das and Rohini Srihari

Open-domain Dialogue Generation: What We Can Do, Cannot Do, And Should Do Next

Katharina Kann, Abteen Ebrahimi, Joewie J. Koh, Shiran Dudy and Alessandro Roncone

MTL-SLT: Multi-Task Learning for Spoken Language Tasks

Zhiqi Huang, Milind Rao, Anirudh Raju, Zhe Zhang, Bach Bui and Chul Lee

Data Augmentation for Intent Classification with Off-the-shelf Large Language Models

Gaurav Sahu, Pau Rodriguez, Issam H. Laradji, Parmida Atighehchian, David Vazquez and Dzmitry Bahdanau

Toward Knowledge-Enriched Conversational Recommendation Systems Tong Zhang, Yong Liu, Boyang Li, Peixiang Zhong, Chen Zhang, Hao Wang and Chunyan Miao

Are Pre-trained Transformers Robust in Intent Classification? A Missing Ingredient in Evaluation of Out-of-Scope Intent Detection Jianguo Zhang, Kazuma Hashimoto, Yao Wan, Zhiwei Liu, Ye Liu, Caiming Xiong and Philip S. Yu

- 12:30 14:00 Lunch Break
- 14:00 14:30 Invited Talk 3 by Zhou Yu
- 14:30 15:00 Invited Talk 4 by Michael Tjalve
- 15:00 15:20 *Coffee Break*
- 15:20 15:50 Invited Talk 5 by Gokhan Tur
- 15:50 16:50 *Oral Paper Session 2*

Human Evaluation of Conversations is an Open Problem: comparing the sensitivity of various methods for evaluating dialogue agents Eric Michael Smith, Orion Hsu, Rebecca Qian, Stephen Roller, Y-Lan Boureau and Jason E Weston

RetroNLU: Retrieval Augmented Task-Oriented Semantic Parsing Vivek Gupta, Akshat Shrivastava, Adithya Sagar, Armen Aghajanyan and Denis Savenkov

Friday, May 27, 2022 (continued)

KG-CRuSE: Recurrent Walks over Knowledge Graph for Explainable Conversation Reasoning using Semantic Embeddings Rajdeep Sarkar, Mihael Arcan and John Philip McCrae

Multimodal Conversational AI: A Survey of Datasets and Approaches Anirudh S Sundar and Larry Heck

16:50 - 17:00 Closing Remarks