

ACL 2020

First Workshop on Natural Language Interfaces

Proceedings of the Workshop

July 10, 2020

Online

©2020 The 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 978-1-945626-00-5 (Volume 1)
ISBN 978-1-952148-16-3 (Volume 2)

Introduction

Ahmed Hassan Awadallah, Yu Su, Huan Sun, Wen-tau Yih

Natural language interfaces (NLIs) have been the "holy grail" of human-computer interaction and information search for decades. However, early attempts in building NLIs to databases did not achieve the expected success due to limitations in language understanding capability, extensibility and explainability, among others. The last 5 years have seen a major resurgence of NLIs in the form of virtual assistants, dialogue systems, and semantic parsing and question answering systems. The horizon of NLIs has also been significantly expanding beyond databases to, e.g., knowledge bases, robots, Internet of Things, Web service APIs, and more.

This has been driven by a number of profound revolutions: (1) In the big data era, and as digitalization continues to grow, there is a rapidly growing demand for interfaces that connect users to the ever-expanding data sources, services and devices in the computing world. NLIs represent a very promising technology to accomplish that as they provide users with a unified way to interact with the entire computing world using language, their natural way of communication, and (2) the renaissance and development of deep learning have brought us from rule and feature engineering to a world of neural architecture and data engineering, promising better language understanding, adaptability and scalability. As a result, many commercial systems like Amazon Alexa, Apple Siri, and Microsoft Cortana, as well as academic studies on NLIs to a wide range of backends have emerged in recent years.

Many research communities have been advancing NLI technologies in recent years: NLP and machine learning, data management and databases, programming language, human-machine interaction, among others. This workshop aims to bring together researchers and practitioners from related communities to review the recent advances and revisit the challenges that led to the failure of earlier NLI systems, and discuss what the remaining challenges are and what to expect in the short- and long-term future.

This workshop is featured by a strong, diverse lineup of invited speakers: Monica Lam's research on virtual assistants stems from a programming language perspective and emphasizes openness and privacy, Percy Liang is a leading researcher on natural language interaction and machine learning, and is also leading Microsoft's development of next-generation conversation systems, H V Jagadish is a pioneer on database usability and NLIs to databases, Joyce Chai has made fundamental contributions to human-robot interaction, Luke Zettlemoyer is a world-renowned expert in semantic parsing, question answering, and natural language processing in general, and Imed Zitouni is an industrial leader on conversational AI with decades of experience at IBM, Microsoft, and Google. The program is also featured by many innovative and forward-looking papers that explore different aspects of NLIs, ranging from compositionality to personalization to data to generation to deployment and interactive learning. We warmly welcome everyone to the workshop, and hope you will enjoy the rich program that covers the past, the present, and the future of NLIs from different research communities.

Organizers:

Ahmed Hassan Awadallah, Microsoft Research
Yu Su, The Ohio State University
Huan Sun, The Ohio State University
Wen-tau Yih, Facebook AI Research (FAIR)

Program Committee:

Danqi Chen, Princeton
Li Dong, Microsoft Research Asia
Xinya Du, Cornell
Kevin Duh, Johns Hopkins University
Ahmed Elgohary, University of Maryland
Hao Fang, Microsoft Semantic Machines
He He, New York University
Lifu Huang, UIUC
Srinivasan Iyer, University of Washington
Robin Jia, Stanford
Bernhard Kratzwald, ETH
Jayant Krishnamurthy, Microsoft Semantic Machines
Victoria Lin, Salesforce
Honglei Liu, Facebook Research
Jian-Guang Lou, Microsoft Research Asia
Yi Luan, Google Research
Siva Reddy, Stanford
Matthew Richardson, Microsoft Research
Jinfeng Rao, Facebook Conversational AI
Sameer Singh, UC Irvine
Shuyi Song, Zhejiang University
Yangqiu Song, HKUST
Xiang Ren, USC
Lingfei Wu, IBM Research
Ziyu Yao, The Ohio State University
Mo Yu, IBM Research
Zhou Yu, UC Davis
Hamed Zamani, University of Massachusetts Amherst

Invited Speakers:

Joyce Chai, University of Michigan
H V Jagadish, University of Michigan
Monica S. Lam, Stanford University
Percy Liang, Stanford University
Luke Zettlemoyer, University of Washington & Facebook AI Research (FAIR)
Imed Zitouni, Google

Table of Contents

<i>Answering Complex Questions by Combining Information from Curated and Extracted Knowledge Bases</i> Nikita Bhutani, Xinyi Zheng, Kun Qian, Yunyao Li and H. Jagadish	1
<i>Towards Reversal-Based Textual Data Augmentation for NLI Problems with Opposable Classes</i> Alexey Tarasov	11
<i>Examination and Extension of Strategies for Improving Personalized Language Modeling via Interpolation</i> Liqun Shao, Sahitya Mantravadi, Tom Manzini, Alejandro Buendia, Manon Knoertzer, Soundar Srinivasan and Chris Quirk	20
<i>Efficient Deployment of Conversational Natural Language Interfaces over Databases</i> Anthony Colas, Trung Bui, Franck Deroncourt, Moumita Sinha and Doo Soon Kim	27
<i>Neural Multi-task Text Normalization and Sanitization with Pointer-Generator</i> Hoang Nguyen and Sandro Cavallari	37

Conference Program

Friday, July 10, 2020

08:15–08:30 *Opening remarks*

08:30–09:30 *Invited Talk: Joyce Chai*

09:30–10:30 *Invited Talk: Percy Liang*

10:30–10:45 *Break*

10:45–11:00 *Answering Complex Questions by Combining Information from Curated and Extracted Knowledge Bases*

Nikita Bhutani, Xinyi Zheng, Kun Qian, Yunyao Li and H. Jagadish

11:00–11:15 *Unnatural Language Processing: Bridging the Gap Between Synthetic and Natural Language Data (Cross-submission)*

Alana Marzoev, Samuel Madden, M. Frans Kaashoek, Michael Cafarella and Jacob Andreas

11:15–11:30 *Towards Reversal-Based Textual Data Augmentation for NLI Problems with Opposable Classes*

Alexey Tarasov

11:30–12:30 *Invited Talk: H V Jagadish*

12:30–13:30 *Break*

13:30–14:30 *Invited Talk: Imed Zitouni*

14:30–14:45 *Examination and Extension of Strategies for Improving Personalized Language Modeling via Interpolation*

Liqun Shao, Sahitya Mantravadi, Tom Manzini, Alejandro Buendia, Manon Knoertzer, Soundar Srinivasan and Chris Quirk

14:45–15:00 *A Multi-Modal Agent that Learns Concepts and Conditionals from Natural Language and Demonstrations (Cross-submission)*

Toby Jia-Jun Li, Marissa Radensky, Justin Jia, Kirielle Singarajah, Tom Mitchell and Brad Myers

15:00–15:15 *Efficient Deployment of Conversational Natural Language Interfaces over Databases*

Anthony Colas, Trung Bui, Franck Deroncourt, Moumita Sinha and Doo Soon Kim

Friday, July 10, 2020 (continued)

15:15–15:30 *Neural Multi-task Text Normalization and Sanitization with Pointer-Generator*
Hoang Nguyen and Sandro Cavallari

15:30–15:45 *Break*

15:45–16:45 *Invited Talk: Monica Lam*

16:45–17:45 *Invited Talk: Luke Zettlemoyer*

17:45–17:50 *Closing remarks*

