GENERALISATION OF END-TO-END TASK-ORIENTED DIALOG SYSTEMS TO UNSEEN KNOWLEDGE

DINESH RAGHU

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 2024

GENERALISATION OF END-TO-END TASK-ORIENTED DIALOG SYSTEMS TO UNSEEN KNOWLEDGE

by

DINESH RAGHU

Department of Computer Science and Engineering

Submitted

in fulfillment of the requirements of the degree of Doctor of Philosophy

to the

INDIAN INSTITUTE OF TECHNOLOGY DELHI

APRIL 2024

Certificate

This is to certify that the thesis titled **Generalisation of End-to-End Task-Oriented Dialog Systems to Unseen Knowledge** being submitted by **Mr Dinesh Raghu** for the award of **Doctor of Philosophy** in the Department of Computer Science and Engineering is a record of bonafide work carried out by him under my guidance and supervision at the **Department of Computer Science and Engineering, Indian Institute of Technology Delhi**. The work presented in this thesis has not been submitted elsewhere, either in part or full, for the award of any other degree or diploma unless otherwise stated explicitly. In particular, some parts of work in Chapter 3 was done jointly with undergraduate and master's students, some parts of Chapter 4 was done jointly with a master's student, and some parts of work in Chapter 5 was done jointly with an undergraduate student and a research developer. In each case, the part done by the collaborators appeared in their respective theses.

Mausam
Professor
Department of Computer Science and Engg.
Indian Institute of Technology Delhi
New Delhi- 110016

Acknowledgements

I would like to extend my heartfelt gratitude to my advisor, Prof. Mausam, whose unwavering patience, guidance, and mentorship have been instrumental in this journey. I am truly fortunate to have had him as my mentor and am grateful for the invaluable wisdom and support provided along the way. I also extend my thanks to my Student Research Committee, Prof. Parag Singla, Prof. Sumeet Agarwal and Prof. Srikanta Bedathur, for their insightful feedback and constructive criticism. I had the privilege of collaborating with Surag Nair, Nikhil Gupta, Atishya Jain, Shantanu Agarwal, and Suraj Joshi. Their presence not only made the journey more enjoyable but also significantly enriched my learning experience.

I wish to express my sincere thanks to Sachindra Joshi, my manager at IBM Research, for his unwavering support during my part-time PhD. His encouragement not only played a pivotal role in initiating my doctoral journey but also remained a constant source of guidance throughout this endeavour. I wish to express my deepest gratitude to my friends and colleagues who not only provided unwavering encouragement and companionship during this academic endeavor but also generously offered constructive comments on my various paper manuscripts. I am sincerely thankful to the following individuals for their invaluable contributions: Sandeep Chandran, Rajshekar K, Gayathri Ananthanarayanan, Prathmesh Kallurkar, Sakshi Tiwari, Neetu Jindal, Hadi Brais, Solomon Abera, Ankit Anand, Yatin Nandwani, Danish Contractor, Dinesh Khandelwal, Gaurav Pandey, Ajay Gupta, and Dhiraj Madaan. Their collective support has significantly enriched the quality of my work, making this journey a collaborative and rewarding experience.

Finally, I am grateful to my family for the unending support and understanding throughout the challenges of this academic journey. A special shout out to my wife, who not only stood by me during both the triumphs and tribulations but also took on more than her fair share of responsibilities, especially in caring for our child. I extend heartfelt thanks to my parents. Their sacrifices and belief in my abilities have shaped not only my education but also the person I am today. This achievement is as much theirs as it is mine, and I am blessed to have them as my pillars of strength.

Abstract

Task-oriented dialog (TOD) systems converse with users to help them with specific tasks such as calendar enquiry, restaurant reservation, and tourist package recommendation. To achieve task success, TOD systems must interact with a task-specific knowledge source. For example, a restaurant reservation system has to be grounded on a knowledge base that contains the names of restaurants, and their details. End-to-end TOD systems are a type of TOD system that are trained using past human-to-human dialogs and do not require any handcrafted state representations and their corresponding annotations on each dialog.

There are several challenges that prevent end-to-end TOD systems from being used for real-world applications. In this thesis, we address three such challenges. The common theme across the contributions is that the proposed TOD systems that are aimed at tackling these challenges generalize well to knowledge unseen during train. (1) To improve the performance of end-to-end TOD systems grounded on structured knowledge bases, we propose three novel neural models – BoSsNet, MLM-Net and CDNet, (2) to reduce the training cost, we define and study the novel problems of training end-to-end TOD systems without the need for explicit KB query annotations, and (3) to extend the types of knowledge sources supported by end-to-end TOD systems, we define and study the novel task of end-to-end learning of a TOD system that troubleshoots user's problems by using a flowchart and a corpus of FAQs.

सार

टास्क–ओरिएंटेड डायलॉग (टी.ओ.डी.) सिस्टम्स उपयोगकर्ताओं की कैलेंडर पूछताछ, रेस्तरां आरक्षण, और पर्यटक पैकेज सुझाव जैसे विशेष कार्यों में मदद करते हैं। कार्य सफलता प्राप्त करने के लिए, टी.ओ.डी. सिस्टम्स को कार्य–विशिष्ट ज्ञान स्रोत के साथ बातचीत करनी चाहिए। उदाहरण के लिए, एक रेस्तरां आरक्षण सिस्टम को एक ज्ञान आधार पर आधारित होना चाहिए जिसमें रेस्तरां के नाम और उनके विवरण शामिल हों। एंड–टू–एंड टी.ओ.डी. सिस्टम्स एक प्रकार के टी.ओ.डी. सिस्टम होते हैं जो पिछले मानव–से–मानव डायलॉग्स का उपयोग करके प्रशिक्षित किए जाते हैं । इन्हें प्रत्येक डायलॉग पर हैंडक्राफ्टेड स्टेट रिप्रेजेंटेशंस और उनसे संबंधित एनोटेशन्स की आवश्यकता नहीं होती है।

एंड–टू–एंड टी.ओ.डी. सिस्टम्स को वास्तविक दुनिया के अनुप्रयोगों में उपयोग करने में कई चुनौतियाँ हैं। इस शोध अध्ययन में, हमने ऐसी तीन चुनौतियों का समाधान किया है। सभी योगदानों में एक सामान्य विषय यह है कि प्रस्तावित टी.ओ.डी. सिस्टम्स, जो इन चुनौतियों को संबोधित करने के लिए बनाए गए हैं, प्रशिक्षण के दौरान अनदेखे ज्ञान के लिए भी अच्छी तरह से सामान्यीकृत होते हैं। (1) संरचित ज्ञान (स्ट्रक्चर्ड नॉलेज) आधारों पर आधारित एंड–टू–एंड टी.ओ.डी. सिस्टम्स के प्रदर्शन में सुधार के लिए, हम तीन नवीन न्यूरल मॉडल्स प्रस्तावित करते हैं –– बॉसनेट, एमएलएम–नेट और सीडीनेट । (2) प्रशिक्षण लागत को कम करने के लिए, हमने एक्सप्लिसिट के.बी. एनोटेशन्स का उपयोग किए बिना एंड–टू–एंड टी.ओ.डी. सिस्टम्स को प्रशिक्षित करने की नई समस्याओं को परिभाषित किया और उनका अध्ययन किया है। (3) एंड–टू–एंड टी.ओ.डी. सिस्टम्स द्वारा समर्थित ज्ञान स्रोतों के प्रकारों का विस्तार करने के लिए, हमने एक नया कार्य परिभाषित किया और उसका अध्ययन किया है। इसमें हम फ़्लोचार्ट और एफ.ए.क्यू. की एक संग्रह का उपयोग करके टी.ओ.डी. सिस्टम का एंड–टू–एंड प्रशिक्षण करते हैं।

Contents

List of Figures

List of Tables