IJCoL

Italian Journal of Computational Linguistics

7-1, 2 | 2021

Special Issue: Computational Dialogue Modelling: The Role of Pragmatics and Common Ground in Interaction

Response to the discussion points by Lina Varonina and Stefan Kopp

Lina Varonina and Stefan Kopp

Electronic version

URL: <https://journals.openedition.org/ijcol/815>

DOI: 10.4000/ijcol.815

ISSN: 2499-4553

Publisher

Accademia University Press

Printed version

Number of pages: 30-31

Electronic reference

Lina Varonina and Stefan Kopp, "Response to the discussion points by Lina Varonina and Stefan Kopp", *IJCoL* [Online], 7-1, 2 | 2021, Online since 01 December 2021, connection on 26 December 2021.
URL: <http://journals.openedition.org/ijcol/815> ; DOI: <https://doi.org/10.4000/ijcol.815>

IJCoL is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License

Knowledge Modelling for Establishment of Common Ground in Dialogue Systems

Response to the discussion points
by Lina Varonina
and Stefan Kopp

Here, we would like to respond to the questions and discussion points raised by Francesco Cutugno and Maria Di Maro in the wake of our paper on knowledge modelling for common ground establishment in dialogue systems. These points are concerned with the connection between the modules for common ground (CG) modelling and dialogue state tracking (DST), and how recent developments with the introduction of deep learning methods to DST can influence CG and knowledge modelling.

Recent research on DST has started to recognise the importance of connecting dialogue context with background information about a domain (Zhou and Small 2019; Ouyang et al. 2020; Chen et al. 2020; Liao et al. 2021). While this is by no means a novel insight, it has not been incorporated much into technical modelling approaches to DST. Further, even in smaller task-oriented domains it is often necessary to look beyond one turn to understand the user and successfully accomplish the task, as one turn will most likely not carry all the information the system needs. In contrast, contemporary voice assistants focus on one-shot request-response interactions with limited needs for context understanding. This is one of the problems users encounter with commercially available voice assistants, which usually expect users "by design" to provide all the necessary information and to ensure that it is understood by the system. When integrating background knowledge about the domain with the information provided by the user during dialogue, however, a system can resolve under-specified requests by making assumptions about user goals (Ouyang et al. 2020): If the user booked a hotel and a restaurant for two people and then wants to also book a taxi, it is highly likely that the taxi will be required for the same two people to transfer from the hotel to the restaurant.

The bigger point behind this argument is that communicative goals of the conversation partners are part of their mental state and the ability to infer mental state facilitates the construction of CG as per the definition of (Clark and Brennan 1991; Lee 2001) that is used in our paper, i.e. mutual knowledge, beliefs and assumptions that the parties have in common either due to similar background or because they were negotiated during the interaction. We argue that to solve the DST challenge one needs to re-recognise the importance of such mental state modelling for human-agent conversational interaction in the future. The current focus of research seems to lie on inference and prediction, while one of the main aspects of CG seen in the above-presented definition is the negotiation of knowledge as the interlocutors cannot be sure that their interpretation of the other's mental state is correct (Kopp and Krämer 2021). Thus, future research should aim to extend the capabilities of dialogue systems to include representations of knowledge that go beyond taking the information recovered from dialogue history as "objective truth". Instead, these representations should incorporate aspects of the interlocutor's mental state, such as epistemic stances or degrees of belief, in order to

account for different degrees of user's understanding or agreement with regards to a particular piece of information.

As we describe in our paper, graphs are an important representation of knowledge in the context of dialogue systems and research on DST often uses this form of representation in combination with deep neural networks (Zhou and Small 2019; Chen et al. 2020; Liao et al. 2021). However, we do believe that there is a need for bringing these methods closer together and that graph-based representations can be embedded into neural dialogue state trackers to enhance the quality of the dialogues through the introduction of cross-turn grounded context and general domain knowledge as argued above. Cutugno and Di Maro correctly note that it is crucial to account for the dynamic nature of knowledge in the context of building CG. Different types of knowledge can exhibit different temporal dynamics of change. Those distinctions can be traced back to Description Logics (Baader, Horrocks, and Sattler 2008) with its concepts of T-Box (terminological box) and A-Box (assertional box). The T-Box contains descriptions of properties and roles of general domain concepts and the relationships between these concepts (comparable with a database schema). The A-Box, on the other hand, contains properties of and relationships between individual instances of these concepts (comparable with data within a database). Looking at modern knowledge representation approaches, one can see parallels of the T-Box and A-Box with the concepts of ontology and knowledge graphs, respectively. In fact, ontology languages such as OWL (Bock et al. 2012), a widely-used web ontology language developed by W3C OWL Working Group, are often based on Description Logics.

Many examples of knowledge modelling discussed in our paper consider knowledge that is static within the use case, e.g., factual domain knowledge. However, we argue that in order to build truly conversational dialogue systems capable of co-constructing CG with their human user, the dynamic aspect of knowledge cannot be discarded. That is, characterising a dialogue state only based on static domain knowledge is insufficient because, even if the topic of a conversation is not changing, user's stance with regards to it may. These notions are currently being introduced into DST research. The work in (Zhou and Small 2019) features a dynamically changing knowledge graph for DST to represent relationships between slots and their values. In their work, they also consider the labels "not mentioned" and "user doesn't care" with regards to possible slot values. This can be considered a basic expression of dynamically changing stance about knowledge within conversation.

The other way around, even the system can have its own stance that evolves throughout the discourse. Depending on the type of the dialogue and the communicative goals of the human and the agent, it may then be necessary to align their beliefs about the domain through interaction, for instance by explanation or argumentation. Hereby it is important to separate representations of the agent's beliefs about the domain and its beliefs about the user's beliefs about the domain. Especially challenging here is that changes in user's beliefs about the domain are never directly observable and can only be inferred under uncertainty from communicative responses or feedback signals. An interesting research question is thus whether a conversational system can reduce this uncertainty with specific dialogue strategies and feedback elicitation in order to more efficiently infer the mental state of the user.

References

- Baader, Franz, Ian Horrocks, and Ulrike Sattler. 2008. Description Logics. In Frank van Harmelen, Vladimir Lifschitz, and Bruce Porter, editors, *Handbook of Knowledge Representation*, volume 3 of *Foundations of Artificial Intelligence*. Elsevier, pages 135–179.

- Bock, Conrad, Achille Fokoue, Peter Haase, Rinke Hoekstra, Ian Horrocks, Alan Ruttenberg, Uli Sattler, and Michael Smith. 2012. OWL 2 Web Ontology Language Structural Specification and Functional-Style Syntax (Second Edition). Online. Accessed on 26.10.2021.
- Chen, Lu, Boer Lv, Chi Wang, Su Zhu, Bowen Tan, and Kai Yu. 2020. Schema-guided multi-domain dialogue state tracking with graph attention neural networks. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (AAAI-20)*, New York, NY, February 7-12.
- Clark, Herbert H. and Susan E. Brennan. 1991. Grounding in Communication. In L. B. Resnick, J. Levine, and S. D. Behrend, editors, *Perspectives on Socially Shared Cognition*. American Psychology Association, Washington, D.C., pages 222–233.
- Kopp, Stefan and Nicole Krämer. 2021. Revisiting human-agent communication: The importance of joint co-construction and understanding mental states. *Frontiers in Psychology*, 12:597.
- Lee, Benny P.H. 2001. Mutual knowledge, background knowledge and shared beliefs: Their roles in establishing common ground. *Journal of Pragmatics*, 33:21–44.
- Liao, Lizi, Le Hong Long, Yunshan Ma, Wenqiang Lei, and Tat-Seng Chua. 2021. Dialogue state tracking with incremental reasoning. *Transactions of the Association for Computational Linguistics*, 9:557–569.
- Ouyang, Yawen, Moxin Chen, Xinyu Dai, Yinggong Zhao, Shujiang Huang, and Jiajun Chen. 2020. Dialogue state tracking with explicit slot connection modeling. In *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, Online, July.
- Zhou, Li and Kevin Small. 2019. Multi-domain dialogue state tracking as dynamic knowledge graph enhanced question answering. *arXiv preprint*, (arXiv:1911.06192).