Controlled Conversational Models through Conversation-Dedicated Ontology

Ph.D proposal - LORIA, Nancy, France
Supervisors: Mathieu d'Aquin, Gaël Guibon

1 Context

Through the development of large language models (LLMs), conversational agents have made great progress recently, producing outputs that are increasingly realistic and acceptable to end users. When applied to specific tasks however, it remains important to be able to control the flow of conversation so that, as the conversation evolves, it is driven towards a positive outcome. Achieving this requires knowledge of different aspects of the conversation, including its objective, but also of various facets of the current conversational flow. The objective of this Ph.D. is to represent such knowledge in an ontology so that it can be used to drive a language model towards better control.

Ontologies are conceptual and logical models of knowledge in a field. They allow the representation of taxonomies of concepts that are relevant to the domain, of relations between those concepts, and of formal definitions, making it possible to carry out inferences of specific situations from such knowledge. As such, in the context of conversational agents, they have typically been used to represent domain knowledge that the agent could use to generate content [1, 7]. However, few works have tried to use ontologies to represent information pertaining to the conversations themselves as a way to drive, in a naturally interpretable way, the way the conversation evolves (flow) in time. One recent exception is Convology [3, 10], an ontology built to represent aspects of conversations that will be used by conversational agents. This ontology and its application are, however, focused on a specific application domain (health) and designed for the specific goal of supporting management and understanding, rather than to improve the efficiency of conversations by controlling their flow.

2 Objectives

As introduced briefly above, the main hypothesis at the foundation of this Ph.D. is that structured, explicit knowledge about the typical characteristics of conversations and of their evolution can help conversational models in producing outputs that are more valuable to the user. A first objective of this Ph.D. will therefore be to use knowledge extraction approaches to build an ontology to represent such knowledge as obtained from analyzing large corpora of conversations annotated with various characteristics, including, in particular, dialog acts, dialog states, expressed emotions and sentiments, among others. Various methods will be used to automatically learn both a typology of conversational flows from those characteristics, the dimensions through which they can be recognized, and how they relate to success or failure of conversations in relation to their objectives.

Based on such an ontology, the ultimate objective of the Ph.D. will be to establish methods to drive pre-trained conversational models towards outputs that are more likely to support successful conversations. While several approaches considered dedicated modules to represent dialog and knowledge memory [6, 11], they focus on task-oriented conversations and totally ignore interpretability/controllability of the model. Moreover, as of now, related research focus mainly either on open-domain conversations [12] or on language-agnostic approaches for task-oriented conversations [5, 9, 12]. In this Ph.D, the objective is to design task-agnostic models for this, which could be achieved, for example, by using the ontology as a dedicated discriminator in a Generative Adversarial Network (GAN) [4] approach. Such methods from the Ph.D could aim at enhancing

conversation representation in the model in an target task-agnostic way, which is yet to be tackled by the research community.

In addition to supporting more efficient conversational models, we expect that the use of the ontology will also increase the interpretability of conversational agents, as decisions to use certain outputs over others will be associated with the ontological dimensions of the current conversation that have driven those choices, thus, moving from common fully black-box NLP models [11]. This could also help these models be less harmful to both users and providers, along with controllable ethics in the way the conversation is analyzed.

3 Methodology

We expect the research in this Ph.D. to be iterative. The two objectives described above could naturally be seen as two phases: we first build the ontology (1) and then use it to control conversational models (2). However, multiple iterations of those two phases will be carried out with successive versions of the ontology, focusing on certain specific characteristics of conversational flows, being used to demonstrate increasing improvements in their ability to control conversational models. We choose this approach to minimize the risks associated with such ambitious research and to ensure early validation of the hypotheses on which it is based.

Build the Ontology. The ontology will be built following a dataset-based approach. This means considering multiple data sources [8, 2] and types would highly benefit the robustness of the ontology in how it will represent conversation flow and structure-related information. This approach yields the benefit of not having to require a fully exhaustive representation of the conversation, thus simplifying the iterative process previously mentioned in the meantime.

Control Conversational Models. Following the hypothesis stating that ontology-based control will yield better conversational models not limited to task-oriented dialogs, we plan to integrate the built ontology in several ways. The ontology integration does not have to be full at first. The most important part is to ensure a steady and reliable improvement from the ontology on several tasks. To do so, we plan to evaluate on multiple target domains and multiple conversation-related tasks (slot filling, dialog acts, etc.). Human interaction and explicability are the ultimate outputs expected in this Ph.D.

4 Supervision Team

This Ph.D. will take place in the LORIA laboratory in Nancy, France, and will be supervised by two researchers from two different teams in LORIA with respective expertise in conversational/language models and in ontology engineering.

Gaël Guibon (SyNaLP Team) is an Associate Professor of Computer Science at the University of Lorraine. His research focuses on Natural Language Processing and Machine Learning towards conversational data, dialog analysis, and model adaptation, especially on user-generated content from different sources.

Mathieu d'Aquin (K Team) is Professor of Computer Science at the University of Lorraine and head of the K Team at LORIA. His research focuses on the way knowledge engineering approaches can contribute to the performance and interpretability of a wide variety of artificial intelligence processes.

References

[1] Hadeel Al-Zubaide and Ayman A. Issa. Ontbot: Ontology based chatbot. In *International Symposium on Innovations in Information and Communications Technology*, pages 7–12, Nov 2011.

[2] Emile Chapuis, Pierre Colombo, Matteo Manica, Matthieu Labeau, and Chloé Clavel. Hierarchical pre-training for sequence labelling in spoken dialog. In *Findings of the Association for Computational Linguistics: EMNLP 2020*, pages 2636–2648, Online, November 2020. Association for Computational Linguistics.

[3] Mauro Dragoni, Giuseppe Rizzo, and Matteo A Senese. Convology: an ontology for conversational agents in digital health. In *Web Semantics*, pages 7–21. Elsevier, 2021.

[4] Ian Goodfellow, Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. Generative adversarial nets. In *Advances in neural information processing systems*, pages 2672–2680, 2014.

[5] Fei Mi, Minlie Huang, Jiyong Zhang, and Boi Faltings. Meta-learning for low-resource natural language generation in task-oriented dialogue systems. In *Proceedings of the 28th International Joint Conference on Artificial Intelligence*, pages 3151–3157, 2019.

[6] Bowen Qin, Min Yang, Lidong Bing, Qingshan Jiang, Chengming Li, and Ruifeng Xu. Exploring auxiliary reasoning tasks for task-oriented dialog systems with meta cooperative learning. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 35, pages 13701–13708, 2021.

[7] Abdul Quamar, Chuan Lei, Dorian Miller, Fatma Ozcan, Jeffrey Kreulen, Robert J Moore, and Vasilis Efthymiou. An ontology-based conversation system for knowledge bases. In *Proceedings of the 2020 ACM SIGMOD International Conference on Management of Data*, pages 361–376, 2020.

[8] Hannah Rashkin, Eric Michael Smith, Margaret Li, and Y-Lan Boureau. Towards empathetic open-domain conversation models: A new benchmark and dataset. In *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pages 5370–5381, Florence, Italy, July 2019. Association for Computational Linguistics.

[9] Evgeniia Razumovskaia, Goran Glavas, Olga Majewska, Edoardo M Ponti, Anna Korhonen, and Ivan Vulic. Crossing the conversational chasm: A primer on natural language processing for multilingual task-oriented dialogue systems. *Journal of Artificial Intelligence Research*, 74:1351–1402, 2022.

[10] Milene Santos Teixeira, Vinícius Maran, and Mauro Dragoni. The interplay of a conversational ontology and ai planning for health dialogue management. In *Proceedings of the 36th annual ACM symposium on applied computing*, pages 611–619, 2021.

[11] Jintao Wen, Dazhi Jiang, Geng Tu, Cheng Liu, and Erik Cambria. Dynamic interactive multiview memory network for emotion recognition in conversation. *Information Fusion*, 91:123–133, 2023.

[12] Yicheng Zou, Bolin Zhu, Xingwu Hu, Tao Gui, and Qi Zhang. Low-resource dialogue summarization with domain-agnostic multi-source pretraining. In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 80–91, 2021.