Challenges for the Conversational Entity Dialog Model

Wolfgang Maier and **Stefan Ultes**
Mercedes-Benz Research & Development
Sindelfingen, Germany
{wolfgang.mw.maier,stefan.ultes}@daimler.com

Abstract

The conversational entity dialogue model (CEDM) (Ultes et al., 2018) offers an intuitive way of modeling task-oriented dialogues in a statistical spoken dialogue system around objects and relations instead of task domains. We address several open challenges around the CEDM and possible extensions of the model.

1 Introduction and Motivation

Research in statistical spoken dialog systems (SDS) (Young et al., 2013) has produced successful systems for task-oriented human-machine dialogues (Lison, 2011; Wang et al., 2014; Budzianowski et al., 2017, among others). In such systems, dialogues are generally modeled using a *multi-domain dialogue model* (MDDM), i.e., they are modeled around single or multiple task domains.

The MDDM focuses on task domains. Therefore, it is hard to describe particular objects of the same type, or to address relations between them. As an example, consider the following example, where two objects (restaurant and hotel) are requested which share an attribute (area).

user i am looking for a restaurant and a hotel in the same area

The first to provide a principled approach for intuitively modeling objects and their relations in the context of SDS have been Ultes et al. (2018). In their *conversational entity dialogue model* (CEDM), dialogues are modeled around objects and the relations between instead of domains.

The CEDM leaves open a number of challenges, which we want to address in this paper. In the following section, we briefly describe the CEDM, and then discuss challenges and possible extensions in Sec. 3 and 4. We close by presenting perspectives for future work.

2 Conversational entity dialog model

The CEDM (Ultes et al., 2018) defines a *conversational entity* as a virtual conversational entity that exists in the context of the current conversation and that is either a conversational object or a conversational relation. A conversational *object* is a conversational entity with a certain type together with a set of attributes which may or may not map to a real-world entity. A conversational *relation* is a conversational entity which connects objects or attributes of objects. Object instances reside in a conversational world that can be derived from the user input, or be predefined.

Dialogues using the MDDM can be modeled using the CEDM, by treating a domain as a conversational object of a specific type, and the slots as the attributes of the type (Ultes et al., 2018, Sec. 4.4). As the CEDM additionally allows for the modeling of relations, it is more expressive than the MDDM.

For more details regarding the handling of belief tracking, etc., please refer to Ultes et al. (2018).

3 Hierarchical Extension

The CEDM can model objects with attributes and relate the attributes. However, types in the CEDM are flat. For instance, the type *hotel* is no more related to the type *guesthouse* than it is to the type *lamp*. This makes it challenging to model (or rather talk about) semantic relations between objects such as hypernymy or hyponymy. Similar to relations in the CEDM that relate attributes of object (e.g., *price2price*), one could allow conversational objects that relate types using a WordNet-style (Fellbaum, 1998) relation (e.g., τ_1 *is a* τ_2 for two types τ_1, τ_2). Consider the following example:

user i am looking for a running outfit
sys here is a suggestion for a jacket, shirt, pants, underwear, and shoes
user i want only black *outerwear*

Proceedings of the 25th Workshop on the Semantics and Pragmatics of Dialogue, September 20–22, 2021, Potsdam / The Internet.

Here, *outerwear* would be a new conversational object with attribute *color=black*, such that for all previously objects except the underwear stand in an *is a* relation with it.

Note that this method explicitly models a relation between the actual objects in the conversation, i.e., it is more expressive than that the ontological knowledge from a backend knowledge base alone.

4 Further Challenges

Count The count of objects present in the conversation can be essential.

user i want to book the tour for tomorrow at 8am
sys how many people will participate
user four
sys please tell me the name of the first person
⋮
user did I say four, make that three persons.

In this case, the count of the *person* objects could be a conversational object itself which needs to be linked to the count of the person objects in the conversational world (Ultes et al., 2018).

Additional knowledge In case several objects are present in the conversation, a user can refer to a subset of the objects in various ways. Above, we have described the case of using a hyperonym (*outerwear*). It is easy to find a similar example where no direct hyperonym is involved.

user i am looking for a running outfit
sys here is a suggestion for a jacket, shirt, pants, underwear, and shoes
user please *no pockets*

In this case *no pockets* addresses a subset of the objects in the conversation, probably *jacket* and *pants*. It could be resolved through the backend knowledge base of the SDS. The subset itself could be handled in the same way as proposed in the previous section.

One can think of more sophisticated ways of addressing subsets of objects in the conversation, such as the following example.

user i am looking for a running outfit
sys here is a suggestion for a jacket, shirt, pants, underwear, and shoes
user too boring above the waistline, can you suggest something else

In such cases, it would be harder to determine what the actual subset is, as it depends, e.g., on a user model (cf. *boring*) or on more general world knowledge. However, the approach of addressing the actual subset once it is found can be the same as before.

Relating multiple attributes The CEDM introduces binary relations between the attributes of objects in the conversation. The following example however implies an n-ary *equals* relation for the *color* attributes.

user i am looking for a running outfit
sys here is a suggestion for a jacket, shirt, pants, underwear, and shoes
user can you suggest something with identical colors?

A similar challenge is the selection of the correct subset of objects based on an attribute for which a particular relation holds, such as *color=red* in the following example.

user i am looking for a running outfit
sys here is a suggestion for a jacket, shirt, pants, underwear, and shoes
user can you switch the red items for yellow ones?

More complex relations Ultes et al. (2018) present the *equals* relation, as for instance in the following example, where it holds between the *area* attributes of the restaurant and the hotel objects.

user i am looking for a hotel and a restaurant *in the same area*

Also, other relations are mentioned such as *less than*. A trivial extension would be to allow for relations on strings such as *startswith*:

user show my contacts with last name starting with 'Z'

One can also think of other knowledge-based relations, such as *matching* in the following example.

user i am looking for a running outfit
sys here is a suggestion for a jacket, shirt, pants, underwear, and shoes
user i want shirt and pants to have *matching* colors

5 Future Work

In this paper, we have sketched open challenges of the conversational entity dialogue model (Ultes et al., 2018). Currently, we are working on an integration of some of the presented aspects into the PyDial dialogue system (Ultes et al., 2017) with the goal of a proper evaluation.

References

Paweł Budzianowski, Stefan Ultes, Pei-Hao Su, Nikola Mrkšić, Tsung-Hsien Wen, Iñigo Casanueva, Lina M. Rojas-Barahona, and Milica Gašić. 2017. Sub-domain modelling for dialogue management with hierarchical reinforcement learning. In *Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue*, pages 86–92, Saarbrücken, Germany. Association for Computational Linguistics.

Christiane Fellbaum, editor. 1998. *WordNet: An electronic lexical database*. MIT Press.

Pierre Lison. 2011. Multi-policy dialogue management. In *Proceedings of the SIGDIAL 2011 Conference*, pages 294–300, Portland, Oregon. Association for Computational Linguistics.

Stefan Ultes, Paweł Budzianowski, Iñigo Casanueva, Lina M. Rojas-Barahona, Bo-Hsiang Tseng, Yen-Chen Wu, Steve Young, and Milica Gašić. 2018. Addressing objects and their relations: The conversational entity dialogue model. In *Proceedings of the 19th Annual SIGdial Meeting on Discourse and Dialogue*, pages 273–283, Melbourne, Australia. Association for Computational Linguistics.

Stefan Ultes, Lina M. Rojas-Barahona, Pei-Hao Su, David Vandyke, Dongho Kim, Iñigo Casanueva, Paweł Budzianowski, Nikola Mrkšić, Tsung-Hsien Wen, Milica Gašić, and Steve Young. 2017. PyDial: A multi-domain statistical dialogue system toolkit. In *Proceedings of ACL 2017, System Demonstrations*, pages 73–78, Vancouver, Canada. Association for Computational Linguistics.

Zhuoran Wang, Hongliang Chen, Guanchun Wang, Hao Tian, Hua Wu, and Haifeng Wang. 2014. Policy learning for domain selection in an extensible multi-domain spoken dialogue system. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 57–67, Doha, Qatar. Association for Computational Linguistics.

Steve Young, Milica Gašić, Blaise Thomson, and Jason D. Williams. 2013. POMDP-based statistical spoken dialog systems: A review. *Proceedings of the IEEE*, 101(5):1160–1179.