Abstract

Building dialogue systems that can smoothly communicate with humans is an enduring topic in artificial intelligence. Nowadays, the task-oriented dialogue system is widely used in real-world business to help users complete specific tasks. In the task-oriented dialogue system, the pipeline system is a popular solution so far. Although the end-to-end system was shown as good performance as the pipeline system recently, the pipeline system is advantageable on that each module of the pipeline system can be separately analyzed and studied. This makes the pipeline system more suitable on researching specific problems. Considering this advantage for research, our research focuses on the pipeline system.

With the developments of services and applications, an existing system is required to be extended to new domains that describe new tasks and topics. In adapting to new domains, conventional systems often needed to retrain the model to handle new task-related classes in new domains, which was inefficient in practice. For this problem, one of the ultimate goals of the task-oriented dialogue system was proposed, that is to build a domain-adaptable system that could be adapted to any given domain without training instances, so-called zero-shot domain adaptation.

In the pipeline system, the slot filling module is the immediate module connecting the user and the system. Slot filling aims to extract the user's intended components by predicting slot entities and slot types. The slot entity indicates the tokens that belong to a slot. The slot type indicates the specific slot that the slot entity is belonging to. The intended components carry the information about completing the task. These components are the basis for the subsequential modules' processes. Therefore, the zero-shot capacity of the entire dialogue system relies on the zero-shot capacity of the slot filling module a lot.

Towards the ultimate goal, conventional slot filling methods were inefficient since they needed to retrain to handle unseen slots in new domains. Zero-shot slot filling was proposed to deal with this problem. Zero-shot slot filling aims to train a model on source domains and adapt the model on target domains directly. Previous zero-shot slot filling methods handled target domains mainly relying the domain similarities based on explicit information. However, the improvements in zero-shot slot filling were limited. The main reason is because the explicit information is sensitive to domain shift problems. Specifically, domain shift problems influence the model

performance from three aspects. The first one is the unseen slots that were not appeared in training domains can be encountered in new domains. The second one is seen slots can be differently explained due to the topic changes in new domains. The third one is the context distributions is generally different in new domains. Due to these domain shift problems, it is hard to treat new domains based on the knowledge learned in training.

The overall objective of our research is to mine intrinsic representations that describe intrinsic characteristics of slots and values to alleviate domain shift problems in zero-shot slot filling towards the ultimate goal. Generally, as the intrinsic characteristics of an object could be stable whatever its specific appearance is, the intrinsic characteristics of slots and values could be expected to be more common across domains, thus providing effective transferable information. Specifically, we separately mined intrinsic representations from three aspects to alleviate domain shift problems in zero-shot slot filling. These representations are the inference relation path (IRP), the multi-relation-based representation, and the ontology-based representation.

We proposed the IRP from the knowledge graph to deal with the domain shift problem of unseen slots. We conducted a statistical analysis and showed that IRPs implicitly carry the relationships between slots and the values of specific meanings. Such relationships were not domain dependent. Thus IRPs could be expected to provide transferable information across domains to alleviate domain shift problems. Experimental results of utilizing IRPs in zero-shot slot filling demonstrated that using IRPs improved zero-shot slot filling by alleviating domain shift problems, especially on the unseen slots. However, IRPs were not flexible to be used since the extraction of IRPs needed the slots and values to be identified as entities, while many slots and values could not be identified as entities in practice. Moreover, the absolute improvement by using IRPs on zero-shot slot filling was not high.

To overcome the limitations of IRPs to alleviate domain shift problems, we proposed the multi-relation-based representation to deal with the domain shift problem of different context distributions. The multi-relation-based representation captures general meanwhile specific characteristic of slot entities among a variety of context environments and slots. Thus it could be expected to provide effective transferable information across domains. Experimental results and analysis demonstrated that the proposed representation alleviated the domain shift problem on the slot entity prediction, thus improving zero-shot slot filling. However, improving slot entity predictions alone could not improve zero-shot slot filling for unseen slots much due to the limitations in the slot type prediction.

To fill the gap of lacking knowledge in slot type predictions to handle domain shift problems of unseen slots and differently explained seen slots,

we proposed the ontology-based representation for the slot type prediction. The ontology is a pre-defined knowledge base that describes the intrinsic relationships between slots and their values. In our research, we assume the ontology for each new domain is fully defined, which contains all slots and possible values. When a domain changes from one to another, the definitions in the ontology will not change. Thus the ontology could establish the relationships between slots and values across domains to handle the slots in new domains. Experimental results and analysis demonstrated that using the ontology-based representation significantly improved zero-shot slot filling. We combined the use of the multi-relation-based representation with the ontology-based representation and showed further alleviation of domain shift problems in zero-shot slot filling.

Finally, we investigated whether the improvements of zero-shot slot filling contribute to the performance of the task-oriented dialogue system, which was unclear in previous studies. We compared dialogue systems using different slot filling modules, including the modules based on conventional methods and the modules based on zero-shot methods. Experimental results demonstrated that zero-shot methods generally contribute to improve the zero-shot capacity of the task-oriented dialogue system from different aspects when encountering unseen domains.

In conclusion, by alleviating the domain shift problems in zero-shot slot filling, the proposed intrinsic representations were effective towards the ultimate goal of the task-oriented dialogue system.

Keywords: Task-oriented dialogue system, zero-shot slot filling, neural network, knowledge graph, ontology, machine learning