The Thirty-Second AAAI Conference
on Artificial Intelligence (AAAI-18)

# A Unified Implicit Dialog Framework for Conversational Commerce

**Song Feng, R. Chulaka Gunasekara, Sunil Shashidhara,**
**Kshitij P. Fadnis, Lazaros C. Polymenakos**
IBM Thomas J. Watson Research Center
Yorktown Heights, NY, USA 10598

## Abstract

We propose a unified *Implicit Dialog* framework for goal-oriented, information seeking tasks of Conversational Commerce applications. It aims to enable the dialog interactions with domain data without replying on the explicitly encoded rules but utilizing the underlying data representation to build the components required for the interactions, which we refer as *Implicit Dialog* in this work. The proposed framework consists of a pipeline of End-to-End trainable modules. It generates a centralized knowledge representation to semantically ground multiple sub-modules. The framework is also integrated with an associated set of tools to gather end users' input for continuous improvement of the system. This framework is designed to facilitate fast development of conversational systems by identifying the components and the data that can be adapted and reused across many end-user applications. We demonstrate our approach by creating conversational agents for several independent domains.

## Introduction

The demands for accessing existing commercial services through chat applications, a.k.a, Conversational Commerce (CC), are increasing rapidly. It is largely pertaining to exchanging information with end users based on the underlying knowledge base of domain services (e.g., car insurance and apartment renting). However, interacting with a knowledge base to fulfill tasks can become very challenging. The developers still have to deal with numerous hand-crafted query rules that are fused with various statistical components repetitively for every single domain. One of the significant yet overlooked drawbacks is the tangling of generic and domain specific knowledge, which makes it nearly impossible in practice to transfer or adapt the existing CC applications to new domains. To address these challenges, we propose a unified `Implicit Dialog` framework for goal-oriented, information seeking conversation systems. It aims to enable the dialog interactions with domain data without replying on the explicitly encoded rules but utilizing the underlying data representation to build the components required for the interactions. The framework is designed for facilitating domain-agnostic, fast prototyping of the interactive search of a domain; more importantly enabling the developers to identify and share common building blocks across various domains.



We assume that a domain knowledge base is available, or can be obtained from the corresponding commercial website. Some of the schema of such websites are illustrated at schema.org. The combination of the domain knowledge base, the permitted queries to the knowledge base, and the application logic can be applied to infer the dialog activities. In particular, we first scan the knowledge base and build a central knowledge representation that can semantically ground multiple dialog subtasks, such as intent labeling, state tracking, and issuing API calls to the domain database. The resulting framework is end-to-end modular and trainable, in line with recently proposed approaches for building conversational agents (rasa.ai 2017), (Truong, Parthasarathi, and Pineau 2017). Those frameworks typically require application developers to code their own dialog logic and provide the training data for dialog subtasks (rasa.ai 2017). However, both the design of dialog management, and annotating the chat data are costly. Our framework is integrated with a dialog simulating module that allows the conversational agent to take advantage of the expertise of a wide variety of stakeholders: end users, domain experts, and human annotators. The module is automatically updated based on the central knowledge representation. It aims to collect targeted feedback that can be directly consumed by the learning modules for continuous improvement of the conversational systems.

We demonstrate the proposed framework on two different domains: apartment renting in NYC and restaurant finding.

## System Overview

As depicted in Figure 1, the proposed framework includes several core modules such as Natural Language Understanding (intent labeler), Inference (state tracker and query generator), a prompt generator and a dialog environment simulator for data collections. The input of these components are all initialized by the relevant part of the Central Knowledge Representation.

### Central Knowledge Representation

We introduce a Central Knowledge Representation (CKR), generated based on the domain knowledge base. It covers the domain entities (e.g., "`apartment`), a set of semantic relations "`has-attribute` between the entities. In particular, CKR provides additional generic characteristics

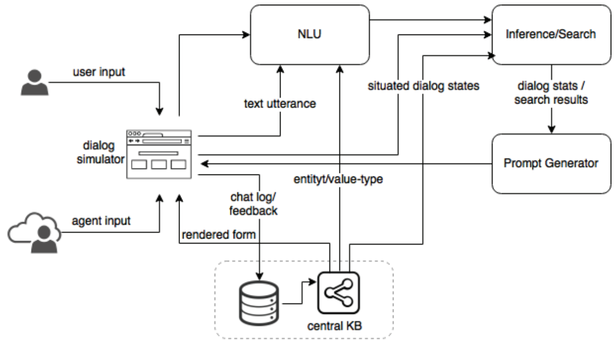


Figure 1: Architecture of the Framework

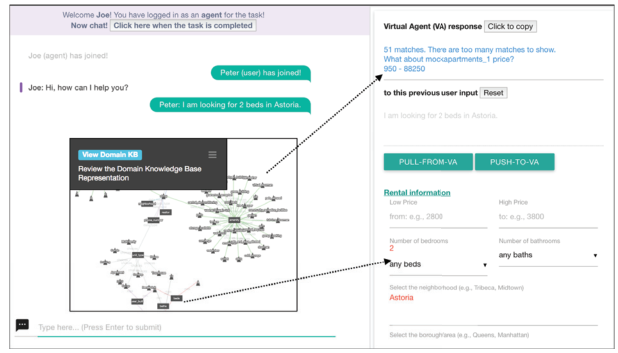


Figure 2: Dialog Simulator Interface

(e.g., expected data types, ranges, operations) that are associated with the entities, which help identifying the content that might be applicable across domains. We then generate the dialog components based on the extracted CKR, thus allowing the reinforcement of the domain changes by updating the CKR.

### Language Understanding and Inference

To facilitate interactive search across domains for CC, we first identify a set of generic operations, such as `ADD`, `DELETE`, `UPDATE` that are useful for updating the dialog states and forming queries. To identify the user intents with any of the operations, we train a natural language classifier (Yang et al. 2016) on chat data[1], where the domain dependent tokens are delexicalised. The best performing classifier shows a promising accuracy of $89\%$ in a test dataset containing over 200 utterances.

Then, we apply semantic matching techniques to identify the slot and values in user utterance to update the dialog state. We map user utterance to the elements in CKR in three consecutive steps, i.e, (1) literal matching, (2) fuzzy matching - which identifies the approximate matches between entities, and (3) vector representation matching - which supports matching word embedding representations of related entities.

Given the last user intent, dialog state and previous search results, the conversational agent either issues an API call using the query graph generated based on CKR or request more information to optimize the search experience. Information gain is used to determine which slot to be requested. Our framework is modular and supports plug-n-play of different methodologies for the language understanding and inference.

### Dialog Simulator

The dialog simulator generates various conversational environments in a Wizard-of-OZ fashion (Schatzmann, Georgila, and Young 2005), which are configured to interface with the crowd-sourcing platforms. Figure 2 presents the UI rendered based on relevant content in CKR. It can simulate the task fulfilling process by human agents: a) accessing the requested information based on identified user intent, and b) converting that information to dialog prompts. The simulator is also able to collect users' feedback on the system's output during the real-time interactions. Compared to many existing chat log annotating tools, our simulator captures certain online information that is more difficult for human annotators to identify offline. Integrating such a simulator with the development framework simplifies the annotation tasks significantly.

[1] https://datasets.maluuba.com/Frames.

## Case Studies

We apply the proposed framework on two independent domains: (1) rental apartments (2) restaurants. For the former, we create a mock database that includes information about `#bedroom`, `transportation`, `location`. As illustrated in Figure 2, the conversational agent could list search results based on users request and provide the guidance on how to optimize the query toward the goal. For the development of the later, we only replace the domain database and re-generate CKR. The conversational agent appears to perform comparably on the new domain.

## Conclusion and Future Work

We propose a unified framework of Implicit Dialog for Conversational Commerce grounded by a centralized knowledge representation. It facilitates fast prototyping and aims at making existing development and chat data reusable and adaptable to new domains. As future work, we plan to conform the central knowledge representation with schema.org to enable the broader practice on domain adaptions.

## References

rasa.ai. 2017. rasa: Open source conversational ai to build next-generation bots and assistants.

Schatzmann, J.; Georgila, K.; and Young, S. 2005. Quantitative evaluation of user simulation techniques for spoken dialogue systems. In *6th SIGdial on Discourse and Dialog*.

Truong, H. P.; Parthasarathi, P.; and Pineau, J. 2017. Maca: A modular architecture for conversational agents. *arXiv preprint arXiv:1705.00673*.

Yang, Z.; Yang, D.; Dyer, C.; He, X.; Smola, A. J.; and Hovy, E. H. 2016. Hierarchical attention networks for document classification. In *HLT-NAACL*, 1480–1489.