# Chatbot Who Wants to Learn the Knowledge: KB-Agent

Younggyun Hahm, Jiho Kim, Sangmin An, Minho Lee, and Key-Sun Choi

Semantic Web Research Center, School of Computing, KAIST
291 Daehak-ro, Yuseong-gu, Deajeon, South Korea
{hahmyg, hogajiho, asm427, pathmaker, kschoi}@kaist.ac.kr

**Abstract.** The usual chatbot is to response for the user's query based KB or already made Question-Answering pairs with its ranked result. KB-agent is a chatbot that is intended to ask for completing the knowledge in the given KB and find the new entity and its attributes trough the conversations with users. In this paper, we introduce KB-agent chatbot, its modular architecture, and its preliminary implementation.

**Keywords:** Chatbot, Knowledge Acquisition, Knowledge Base

## 1 Introduction

The usual chatbot is to response for the user's query based on Knowledge Base (KB) [1] or already made Question-Answering pairs [4] with its ranked result. In terms of KB-based question answering, to answer user's query it is required to construct the *complete* KB which can provide facts for all of user's query sufficiently. KB completion is a task to address this problem. In this context, the dialog interface of chatbot systems has several advantages: 1) It can choose what to ask based on a given dialog context, 2) it can acquire knowledge from user by asking for missing knowledge in KB directly to complete KB, 3) and it can encourage users to participate through dialogs.

In this paper, we introduce KB-agent that asks for completing knowledge to users. We will 1) describe a modular architecture of KB-agent to learn the knowledge through dialogs, 2) show a preliminary prototype, and 3) describe remaining issues.

## 2 Related Work

Lately, researches have shown increasing interest in dialog systems, and there are many chatbot applications, such as Alexa, Siri, Google Assistant, and so on. These applications are focusing on short, and task-oriented dialogs (e.g. "*play music"*) to provide relevant services for a given user's request. Alexa Prize [5] focus on engaging dialog systems for longer conversation. For example, [3] consist of several modules to respond to various user's needs such as jokes and opinions. Dialog System Technology Challenge (DSTC) and bAbI project focus on task-oriented dialogs [2][6], such as restaurant reservation and recommendation system. These dialog systems respond to user's utterance with conversation for providing information or for asking missing information in

user's request. By responding user's request, these systems achieve their purpose of a task the user intended to perform. These dialog systems are designed to achieve the user's goals by responding user-initiative dialogs. In this paper, we implement KB-agent which is designed to achieve the purpose of the system to acquire knowledge for KB by asking system-initiative questions.

## 3 System Architecture

In our scenario, KB-agent monitors a given KB and listing what is missing in KB. KB-agent is designed to ask one of these missing knowledge to users when predefined keywords appeared in a given conversation context. This means that KB-agent generate questions about topic entities in conversation that users are interested in. In this scenario, KB-agent conduct the following dialog process: 1) First, KB-agent generate a list of questions for each entity in a given KB, and then 2) KB-agent ask question(s) to user when the relevant keywords for the list of question. And 3) KB-agent identify entities from a response of the user, and make it KB data. The scope of KB-agent implemented in this paper is limited to the *singer* domain, more specifically, to entities which is belongs to `MusicalArtist` class in Korean DBpedia.

### 3.1 NLU

The Natural Language Understanding Module (NLU) is a module to understand user's indent from a given user's utterance in context.

*Domain classification*

First, it identifies the domain of user's dialogs. For example, if KB-agent deals with the *singer* domain only, it will identify whether the current user's utterance is about a singer or other out-of-scope topic. The current version of KB-agent is based on a keyword-based approach.

*Intent classification*

This module classifies user's intent based on the dialog act taxonomy[1]. The current KB-agent is based on a rule-based approach to determine whether a given user's utterance is a positive or negative answer (`Agree`, `Disagree`), and whether is an informative or answering utterance (`Inform`, `Answer`).

*Topic identification*

For example, for a given question "What is your favorite song of *SINGER_A*?", a user can answer "I like *SONG_A*. It is exciting". In other words, an answer of users would be not only a noun phrase, but also sentences. The NLU module identifies entities and consider that is an answer for system's question.

[1] https://dit.uvt.nl/

### 3.2 DM

The Disambiguation Module (DM) determines what to do in the next turn in conversation with user, and generates contents of system's utterance. The following submodules are included:

*Dialog Handler*

This module determines the next action of KB-agent in conversation with user. This module takes an input from the NLU module. When a given user's utterance deals with its domain, and includes entities which belongs to that domain, this module choose the next action *turn-taking* and then generates questions by calling the KB-question generator module. In other case, it calls chitchat modules to generate responds.

*KB-question generator*

The KB-question generator is a module that listing the incomplete knowledge for a given KB. In more detail, KB consist of triplets <e1, relation, e2> and this module generate a list of high frequent relations for entities belonging to same class in the KB. And if some entities have no triplet of the high frequent relation, this module considers that is the incomplete knowledge of the entity so that an incomplete triplet <e1, relation, _ > is generated as an output of this module.

*Error Handler*

The purpose of this module is to handle unexpected answers from users. For example, when KB-agent ask "Who is your favorite singer?", KB-agent expects users answer is about the singer, however, in many case, it is hard to be sure that the user's answer is a singer. If there are ambiguity, this module asks again to check whether the user's answer is a singer. This is a necessary process for adding new entities that are not in the existing KB.

### 3.3 NLG

The Natural Language Generation Module (NLG) generates natural language sentences for the contents generated by the DM module. It uses the AIML framework based on a rule-based approach.

### 3.4 UI Controller

We used the KakaoTalk messenger interface[2] to provide a front-end user interface. Most of the messenger interface supports the REST API service, which returns a response for an input utterance. KB-agent generates utterances based on the context of the conversation. This module controls messenger service and KB-agent interactively by storing dialog history of each user to send it to KB-agent.

[2] https://github.com/plusfriend/auto_reply

## 4 Conclusion and Discussion

In this paper, we introduced the chatbot, KB-agent, which is intended to ask for completing the knowledge in the given KB and find the new entity and its attributes trough the conversations with users. KB-agent has the following features: 1) It monitors the given KB and lists incomplete knowledge, 2) and it ask questions to user based on a given dialog context. Then 3) it acquires knowledge from the user. The current version of KB-agent which is implemented in this paper is based on a rule-based approach, and there are remaining tasks to improve the system: 1) How to generate appropriate questions for a given dialog history, and 2) how to encourage users for long conversations to acquire knowledge.

## Acknowledgements

This work was supported by Institute for Information & communications Technology Promotion (IITP) grant funded by the Korea government(MSIT) (2016-0-00562, Emotional Intelligence Technology to Infer Human Emotion and Carry on Dialogue Accordingly)

## References

1. Athreya, Ram G., Ngomo, Axel-Cyrille Ngonga., Usbeck, Ricardo.: Enhancing Community Interactions with Data-Driven Chatbots–The DBpedia Chatbot. In Companion of the The Web Conference 2018 on The Web Conference 2018. International World Wide Web Conferences Steering Committee, p. 143-146 (2018)
2. Bordes, Antoine., Boureau, Y.-Lan., Weston, Jason.: Learning end-to-end goal-oriented dialog. arXiv preprint arXiv:1605.07683, (2016)
3. Fang, Hao., Cheng, Hao., Clark, Elizabeth., Holtzman, Ariel., Sap, Maarten., Ostendorf, Mari., Choi, Yejin., and Smith, Noah A.: Sounding Board–University of Washington's Alexa Prize Submission. Alexa Prize Proceedings (2017)
4. Kato, Makoto P., Yamamoto, Takehiro., Manabe, Tomohiro., Nishida, Akiomi., Fujita, Sumio.: Overview of the NTCIR-13 OpenLiveQ Task. In Proceedings of the 12th NTCIR Conference. (2016)
5. Ram, Ashwin., Prasad, Rohit., Khatri, Chandra., Venkatesh, Anu., Gabriel, Raefer., Liu, Qing., Nunn, Jeff., Hedayatnia, Behnam., Cheng, Ming., Nagar, Ashish., King, Eric., Bland, Kate., Wartick, Amanda., Pan, Yi., Song, Han., Jayadevan, Sk., Hwang, Gene., Pettigrue, Art.: Conversational AI: The Science Behind the Alexa Prize. arXiv preprint arXiv:1801.03604, (2018)
6. Wen, Tsung-Hsien., Vandyke, David., Mrksic, Nikola., Gasic, Milica., Rojas-Barahona, Lina M., Su, Pei-Hao., Ultes, Stefan., Young, Steve.: A network-based end-to-end trainable task-oriented dialogue system. arXiv preprint arXiv:1604.04562, (2016)