

Domain-Adaptive Conversational Agent with Two-Stage Dialogue Management

Jin-Hyuk Hong and Sung-Bae Cho

Dept. of Computer Science, Yonsei University,
134 Shinchon-dong, Sudaemoon-ku, Seoul 120-749, Korea
hjinh@sclab.yonsei.ac.kr, sbcho@cs.yonsei.ac.kr

Abstract. The conversational agent understands and provides users with proper information based on natural language. Conventional agents based on pattern matching have much restriction to manage various types of real dialogues and to improve the answering performance. For the effective construction of conversational agents, we propose a domain-adaptive conversational agent that infers the user's intention with two-stage inference and incrementally improves the answering performance through a learning dialogue. We can confirm the usefulness of the proposed method with examples and usability tests.

1 Introduction

Conversational agents have been studied as an effective interface to understand and respond to users. Conversations are not only convenient but also abundant and flexible for users to express what they want [1]. Pattern matching, one of popular methods for constructing the conversational agent, works favorably at a sentence, but it is not feasible to understand a dialogue in which context should be considered [2]. And the performance of conversational agents is mainly subject to the quality of dialogue management and the quantity of knowledge-base. Even with an excellent dialogue manager, the answer performance might be limited by the size of knowledge-base. Constructing knowledge-base has many difficulties encountered in knowledge engineering for expert systems. Scripts are usually designed by hand, and it costs much time and effort [3]. Furthermore, designers should be intelligent in analyzing the domain and designing scripts, since the quality of scripts affects on the answering performance.

For the construction of effective conversational agents, at first in this paper, we develop the architecture of two-stage dialogue management. Then, we propose an interactive knowledge acquisition method to construct scripts for the pattern matching of the proposed conversational agent.

2 Knowledge Acquisition from Natural Language Sources

As a basic means of the representation of information, natural language sources are obtainable in large quantities from the web and other materials. It might be successful

when we utilize them to construct the knowledge base of intelligent agents instead of manual processing. If the data are static, the acquisition of information from them might be easy. In many cases, however, natural language sources are variable so that it is very difficult to extract useful information from them [4]. The knowledge acquisition from natural language sources has long been studied, and recently there are many works on extracting semantic or structural information from web pages [5].

Model-based approach is a conventional knowledge acquisition method, which employs in constructing the knowledge base of expert systems [4]. It defines the static knowledge representation of a target domain as a model, and collects necessary information from documents to complete the model. It uses templates or frames on the target domain to effectively process information, and it provides a brief process to understand [6]. Recently, the knowledge acquisition using the ontology awakes interest. Comparing with the model-based approach, it extracts rather deep semantic and structural information from documents. Especially it is now applied to acquire information by analyzing semantic structures not only within a web page but also between web pages [7].

There are some other works on the knowledge acquisition from dialogue not from simple natural language documents. Lauria *et al.*[8] tried to construct the course of the movement of a robot by natural language instructions. However, it just manages the dialogues of lower level such as simple commands, contrary to real conversation.

3 Proposed Conversational Agent

As shown in Fig. 1, the proposed conversational agent is composed of two parts: dialogue manager and knowledge manager. Dialogue manager deals with and responds to the user's query using the inference engine. If dialogue manager dose not understand a

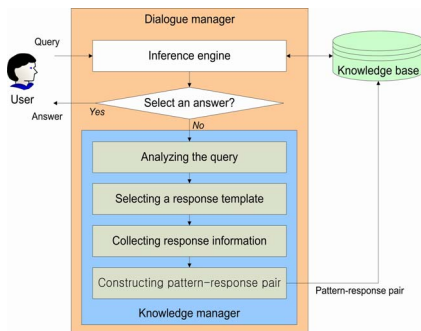


Fig. 1. The proposed conversational agent

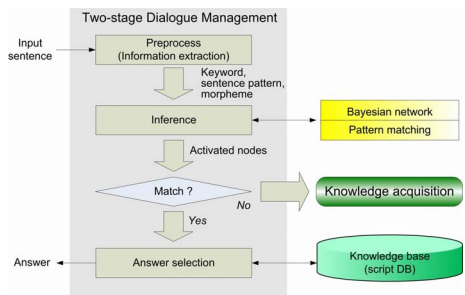


Fig. 2. Overview of the two-stage dialogue management

query, the query is regarded as a learning dialogue and knowledge manager is executed. With a series of knowledge acquisition procedures, a script composed of patterns and responses is generated to be appended to knowledge-base. Since knowledge acquisition is conducted by conversation, it is easy to construct a domain-adaptive knowledge-base.

3.1 Two-Stage Inference for Dialogue Management

For efficient inference, useful words are extracted from the input query by preprocessing. These words are used to infer the user's intention as shown in Fig. 2. The inference process is divided into two steps: Bayesian network and pattern matching. Analyzing queries in stages makes it feasible to infer the detailed intention of the user and to model the context of conversation. Dividing knowledge-base, furthermore, improves the scalability and portability of conversational agents.

3.1.1 Topic Inference Using Bayesian Network

As a counter proposal of simple pattern matching, Bayesian network has been applied to dialogue management for the effective and accurate response. We design Bayesian network to infer the topic of conversation, which leads to define the scope of dialogue. Bayesian network is hierarchically constructed with three levels based on their functions: keyword, concept, and topic. Keyword level consists of words related to topics in the domain. Concept level is composed of the entities or attributes of the domain, while topic level represents the cases of entities whose attributes are defined. It is sometimes failed to infer the intention at one try, rather accumulating information through conversation is more efficient to infer the intention.

Words extracted from preprocessing are used as evidences in this paper. Through inference, the highest probabilistic topic node is selected if the probability is over threshold. Then pattern matching associated with a topic selected is executed to select a proper answer. If a topic is not selected, the proposed method tries mixed-initiative interaction to understand the user's intention. Since concept level contains a part of information, the agent asks the user to collect more information. After the user provides more information about his intention, the agent attempts inference again.

3.1.2 Answer Selection Using Pattern Matching

Once a topic is selected as the topic of an input query, pattern matching using knowledge-base associated with the topic is performed to select an answer. When there are

<pre> <SCRIPT> <TOPIC> topic </TOPIC> <PATTERN> words </PATTERN> <RESPONSE> answer </RESPONSE> </SCRIPT> </pre>
<pre> <SCRIPT> <TOPIC> The location of school </TOPIC> <PATTERN> where is school </PATTERN> <PATTERN> know location school </PATTERN> <RESPONSE> The school is located in Shinchon Dong. </RESPONSE> <RESPONSE> It is in Shinchon Dong. </RESPONSE> </SCRIPT> </pre>

Fig. 3. Definition of a script

lots of scripts, the performance declines because of the redundancy of information. In this paper, we divide scripts into several groups based on their topics so as to reduce the number of scripts compared. A script is stored as XML form, and Fig. 3 shows the definition and example of a script.

3.2 Pattern-Response Pairs Learning Through Dialogue

The pattern of a script is composed of the topic and a set of words. The topic of a query is obtained by Bayesian network, while words are extracted by preprocessing. Therefore, a pattern is simply defined by combining them and generating the pattern part of a script.

```

(def-index template-name
  Definition :
    :class      (dialogue-act)
    :question   (question-script)
    :requirement (concept-of-construction)
    :answer     (script input-position)
  )

(def-index no. 10
  Example :
    :class      (?DescriptionQuestion)
    :question   ("What kind of location information should be provided?")
    :requirement (descript_value)
    :answer     ("The location is %s.", descript_value)
    :answer     ("It locates in %s.", descript_value)
  )
    
```

Fig. 4. Structure of the answer template

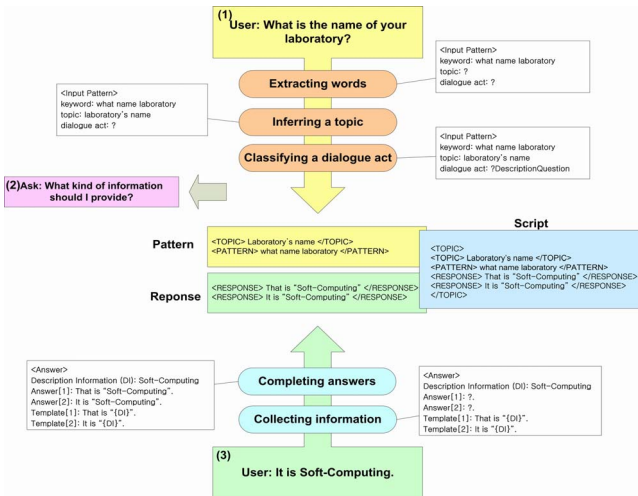


Fig. 5. Overall procedure of the knowledge manager

We consider the dialogue act of an input query to select a proper answer template for the generation of scripts. An automaton extracts a dialogue act, and 30 automata are designed for 30 dialogues. The subsumption architecture is adopted to select one dialogue act for a query. Each dialogue act has the corresponding answer template for

constructing a response. An answer template is defined as shown in Fig. 4, of which “class” means the dialogue act, and “question” is a sentence to collect information. An answer script is constructed by filling “requirement (information collected from the user)” into “answer (a template).” In this paper, we define 64 templates: three templates for each question dialogue act, and six positive/negative templates for each statement dialogue act. After all, an answer is constructed by completing the template, and saved as the response part of a script.

The pattern-response pair is generated with a pattern and the responses. Fig. 5 shows the overall procedure of the generation of a pattern-response pair.

4 Experiments

4.1 Examples

The change of topics is common in real conversation, and the agent has to maintain the context of dialogue. Bayesian network of the proposed method models the context, and Dialogue 1 shows the example. Dialogue 2 shows the example of knowledge acquisition. Contrary to manual operations, a user only needs to provide what the agent requires through dialogue. It automatically performs all processes such as stemming words, defining dialogue act, and constructing answers.

User: What is your professor's telephone number?
 Word (what, number, telephone, professor)
 BN (topic: Professor's phone number)

Agent: Our professor's phone number is 2123-XXXX.

User: What about you?
 Word (what, you)
 High probability nodes (phone-number, professor)
 BN (topic: The phone number of Agent)

Agent: I'm sorry, I don't have any phone.

User: Hmm, so what about email address?
 Word (what, email, address)
 High probability nodes (phone-number, Agent)
 BN (topic: The email address of Agent)

Agent: My email address is chatbot@sclab.yonsei.ac.kr.

Dialogue 1. Context maintenance using Bayesian network

User: Where is the laboratory?
Agent: Hmm, I don't know. Change mode to Knowledge acquisition.
 => analyzing query
Agent: Where is it?
User: It locates at 529, the 3rd engineering building, 134, Yonsei university.
 => analyzing answer (requirement: 529, the 3rd engineering building, 134, Yonsei university)
 => generating the pattern-response pair

```

<SCRIPT>
<TOPIC> Laboratory's location </TOPIC>
<PATTERN> where laboratory </PATTERN>
<RESPONSE> It locates at 529, the 3rd engineering building, 134, Yonsei university </RESPONSE>
<RESPONSE> 529, the 3rd engineering building, 134, Yonsei university </RESPONSE>
<RESPONSE> It is at 529, the 3rd engineering building, 134, Yonsei university </RESPONSE>
</SCRIPT>
    
```

User: Where is the laboratory?
Agent: It locates at 529, the 3rd engineering building, 134, Yonsei university

Dialogue 2. Knowledge acquisition by the proposed method

4.2 Usability Test

At first, we have collected a number of dialogues from 8 subjects performing 3 tasks that request to search information. 25 simple queries, 11 context-maintaining dialogues, and 11 mixed-initiative dialogues are collected. These numbers show that missing or spurious words are included in real conversation. We have compared with simple pattern matching (SPM) using these dialogues, and Table 1 shows the result. The result shows that the proposed method is superior to SPM, since it manages various types of dialogues while SPM fails to respond.

Table 1. Comparative result with simple pattern matching

Retrieval rate (RR)		Task 1		Task 2		Task 3	
Average interactions (AI)		RR	AI	RR	AI	RR	AI
SPM	Training	92.5%	1.5	95%	2.7	93.2%	4.4
	Test	87.3%	1.6	86.3%	2.7	87.7%	4.6
Proposed method	Training	100%	1.2	100%	2.1	100%	4.0
	Test	92.3%	1.4	96%	2.3	91.7%	4.5

For knowledge acquisition, we have experimented the usability test comparing with the manual construction of the knowledge base. 50 dialogues are used as training data to construct pattern-response pairs, while other 50 dialogues are used as test data. Experts and novices perform the experiment. Table 2 shows the result of the usability test for knowledge acquisition.

Table 2. Result of the usability test for knowledge acquisition

		Manual construction		Proposed method	
		Experts	Novices	Experts	Novices
Pattern-response pairs generated		50	50	44	44
File size		28 KB	30 KB	24 KB	24 KB
Construction time		20 min.	1 hour	5 min.	8 min.
Accuracy	Training	92%	84%	100%	96%
	Test	84%	82%	88%	86%
User satisfiability (0~5)		2	1	4	4

For queries having the same pattern, designers do not recognize them manually while they notice that case with the proposed method. Therefore, it reduces the size of knowledge base. In terms of construction time and the users' satisfiability, the proposed method is outstandingly superior to the manual construction.

5 Concluding Remarks

In this paper, we have constructed the conversational agent using two-stage dialogue management and its knowledge acquisition. Bayesian network used in dialogue management provides more flexible and detailed inference to manage various types of dialogues. In order to construct domain adaptive conversational agents, we have also proposed interactive knowledge acquisition. It is familiar and effective to collect useful knowledge to construct pattern-response pairs from learning dialogues. The example and the usability test have demonstrated the usefulness and power of the proposed method comparing with conventional approaches, especially for novices.

Acknowledgements

This paper was supported by Brain Science and Engineering Research Program sponsored by Korean Ministry of Science and Technology.

References

1. V. Zue and J. Class, "Conversational interfaces: Advances and challenges," Proc. of the IEEE, 88(8), pp. 1166-1180, 2000.
2. J.-H. Hong and S.-B. Cho, "A two-stage Bayesian network for effective development of conversational agent," LNCS 2690, pp. 1-8, 2003.
3. E. Horvitz, et al., "The lumiere project: Bayesian user modeling for inferring the goals and needs of software users," Proc. of the 14th Conf. on Uncertainty in Artificial Intelligence, pp. 256-265, 1998.
4. C. Schmidt and T. Wetter, "Using natural language sources in model-based knowledge acquisition," Data & Knowledge Engineering, 26(3), pp. 327-356, 1998.
5. H. Alani, et al., "Automatic ontology-based knowledge extraction from web documents," IEEE Intelligent Systems, 18(1), pp. 14-21, 2003.
6. A. Arruarte, et al., "A template-based concept mapping tool for computer-aided learning," Proc. IEEE Int. Conf. Advanced Learning Technologies, pp. 309-312, 2001.
7. R. Navigli, et al., "Ontology learning and its application to automated terminology translation," IEEE Intelligent Systems, 18(1), pp. 22-31, 2003.
8. S. Lauria, et al., "Personal robots using natural language instruction," IEEE Intelligent Systems, 16(3), pp. 38-45, 2001.