FAIM/ISCA Workshop on Artificial Intelligence for Multimodal Human Robot Interaction
14-15 July 2018, Stockholm, Sweden

Knowledge-based Dialog Approach for Exploring User's Intention

Katsuyoshi Yamagami [1 2], Hirokazu Kiyomaru [3], Sadao Kurohashi [3],
[1] Advanced Industrial Science and Technology (AIST)
[2] Panasonic Corporation
[3] Kyoto University
yamagami.katsuyoshi@aist.go.jp, kiyomaru@nlp.ist.i.kyoto-u.ac.jp, kuro@i.kyoto-u.ac.jp,

Abstract

In this paper, we address an integrated framework of domain knowledge and dialog system that can understand user's intention as comprehensively as possible through interaction to recommend contents that meet the user's preferences. The essential concept of our framework is exploring user's intention by controlling topic in dialog through domain knowledge. We consider our framework to be a promising solution especially when the user does not have enough knowledge of target domain of contents. We also report our preliminary work toward our research goal.

1 Introduction

We are developing a goal-oriented dialog system for retrieving or recommending contents in a practical use case [Yamagami *et al.*, 2017]. The pilot task is a cooking recipe recommendation.

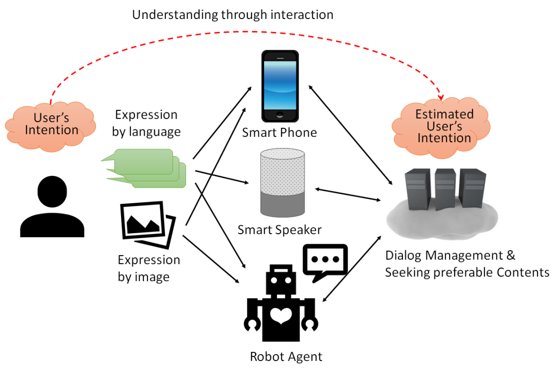

Figure 1: Goal image of our research

Figure 1 is an illustration of our research goal. The dialog system understands the user intention from the expression both by uttered language and by presented image. The reason why we introduce image understanding is that user can utilize image to express his/her intention instead of language when having difficulty to do so with language. Of course, it is possible and reasonable to integrate other interactive modals into our research scope. We have set our goal to be feasible in popularly available devices in recent days or near future.

Goal-oriented dialog system has been an important issue in research of interactive AI and its application. Due to the difficulty of traditional goal-oriented systems, which require a lot domain-specific handcrafting, end-to-end dialog systems have recently been proposed [Bordes *et al.*, 2017]. However, they can deal only with simple tasks based on slot filling due to the difficulty of learning task domain knowledge with wide coverage from dialog corpus. We think that domain knowledge is essential to goal-oriented dialog system, especially in understanding user's intention, and are developing an integrated framework of domain knowledge and dialog management.

When people seek preferable contents by interacting with a dialog system, they have to express their intention or preference in their utterances. However, it is often difficult for ordinary people to state their intentions in a single utterance that can completely express them [1]. Considering the case of shopping, for example, people usually talk with a shop staff in fragmentary expressions which represent or related to their intention, so the shop staff understands their intention gradually through the interaction [Ohtake *et al.*, 2009].

The main problems that make users use fragmentary expression are:

- User does not always have enough knowledge about domain contents.
- User is not always aware comprehensively of his/her intention or preference about domain contents

The word 'knowledge' mentioned above means 'contents' themselves and also means 'concepts' that characterize 'contents.' Therefore, a dialog system for retrieving or recommending contents information should assist the user to express his/her intention. The reason why is that the user may become aware of potential 'concepts' in his/her intention only after the system provides information about these 'concepts.' In addition, the user may often hesitate to express his/her intention because he/she does not know about what 'knowledge' the system has. This is also a problem that the system should solve or mitigate.

[1] A question sentence in TV quiz show is an example of the utterance

10.21437/AI-MHRI.2018-13

Misu [Misu *et al.*, 2011] proposed a dialog system for supporting user's decision on choosing sightseeing spots. His work is one of the promising solutions to the problem mentioned above. He handled a small size of content to focus on the problem of dialog management. However, it is necessary to deal with large-scale contents set for practical applications such as commercial IT services. The cooking recipe provider Cookpad, for example, holds contents more than one million [Harashima *et al.*, 2016]. We think it is a challenge to realize a dialog system dealing with such large-scale content in the field of interactive AI system.

2 Domain Knowledge for Exploring User's Intention

2.1 Requirements for Exploring User's Intention

Table 1: Example of dialog scenario aimed in our research

User utterances and system responses (U: User, S: System)	Domain knowledge applied by system
U: It's cold today.	
S: Warm food is good for today, isn't it?	`cold -> warm food`
U: That's right.	
S: How about Nabe as a warm food?	`warm food -> Nabe`
U: Nabe is preferable.	
S: How about Sake with Nabe?	`Nabe -> Sake`
U: I like beer.	
S: How about appetizer for beer?	`beer -> alcohol -> appetizer`
U: I love it.	

We explain our basic idea of the proposed approach with an example dialog scenario. Table 1 shows an example dialog that we are aiming to realize. Dialog system picks up a word 'cold' in the first utterance as a piece of user intention, then performs inference based on domain knowledge '`cold -> warm food`' in the domain knowledge base. (The notation '`concept_1 -> concept_2`' indicates that there is a semantic relation between '`concept_1`' and '`concept_2`'.) Then, the system responds 'Warm food is good for today, isn't it?' to confirm the user's intention (i.e., preference) with regard to '`warm food`'.

Thus the system explores user's intention by providing a variety of topics (i.e., user preferences that system wants to confirm) in system responses, which are controlled by domain knowledge to keep the dialog context reasonable for the user. We found that dialog context often loses cohesion [2] if dialog topics are controlled only by an efficiency-first mechanism such as decision tree method we introduced in our trial dialog system 'CookChat'[3] for cooking recipe recommendation task [Yamagami *et al.*, 2017].

Analysis of user utterances collected in 'CookChat' indicated that the user might often use expressions representing a feeling or an impression rather than objects (ex. name of material or genre etc.) to state his/her preference of recipe. We call those expression 'topic' instead of 'determinant' mentioned in Misu's work. Misu's dialog model assumes that relation between content and topic is a binary relation (i.e., '0' or '1'), so that the model lacks the flexibility of relation representation.

We consider the following two as fundamental requirements which our domain knowledge should satisfy.

1. Collecting lexical knowledge representing concepts concerning the task domain with high coverage. Especially collecting topic concepts to seek the concepts corresponding to contents.
2. Representing strength of relationships among concepts flexibly.

In our survey, we did not find previous work with a large vocabulary of topic concepts, while there are several works with relatively small one [Matsumoto and Ren, 2011].

2.2 Design of Domain Knowledge Base

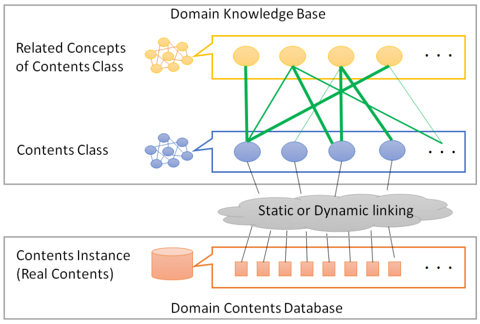

Figure 2: Structure of proposed domain knowledge base

Figure 2 shows the structure of domain knowledge base. It consists of two main categories of knowledge: 'Contents Class' (CC for short) and 'Related Concepts of Contents Class' (RC for short). RC is intended to play a role of topic. 'Contents Instance' (CI for short) are located outside of domain knowledge base. In Misu's work, there is no distinction between CC and CI because it does not matter due to the small size of CI. In case of the large size of CI, relations between RC and large-scale CI are too verbose for dialog manager to handle for exploring the user intention. That is why we introduce CC as a generalization of CI.

As mentioned in subsection 2.1, RC and CC are both used for a system response to explore the user's intention. RC and CC are linked [4] with a variety of relation type. Each link (denoted as a green line) has a link score representing the strength of a relation (denoted with the thickness of the line).

2.3 Dialog Framework based on Domain Knowledge

Figure 3 shows the overview of the dialog framework based on domain knowledge. This framework controls the selection

[2] The dialog looks like diagnosis or interrogation.

[3] https://www.weekcook.jp/trial/cookchat/lp.html (Japanese chat system)

[4] RC and CC both are also linked among themselves.

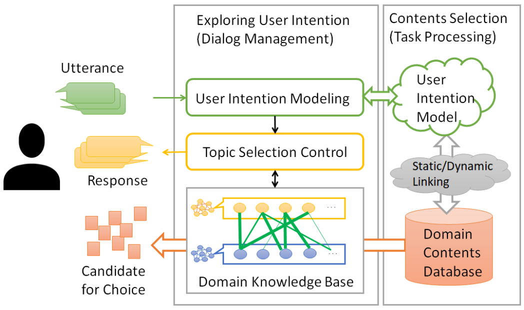

Figure 3: Overview of dialog framework based on domain knowledge

of dialog topic based on the user intention model and domain knowledge. In table 1, the system finds '`cold`' (RC) from the first user utterance, updates the user intention model [5] with it and selects dialog topic '`warm food`' (CC) by tracing the relation '`cold -> warm food`'. In addition, the system can control the selection of dialog topic based on the link score, for example, in the case of multiple relations being possible to select.

3 Preliminary Work

We have built a knowledge base including knowledge about fundamental Japanese cuisines [Kiyomaru *et al.*, 2018] and a trial dialog system using the knowledge base [Yamagami *et al.*, 2018] as preliminary work of our proposal described in section 2.

3.1 Knowledge Base of Fundamental Cuisine

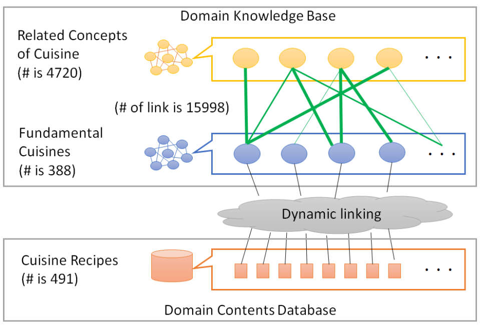

Figure 4: Structure of domain knowledge base of fundamental cuisine

Figure 4 shows the structure of the knowledge base. This consists of two knowledge categories: fundamental cuisines (FC for short) and related concepts of cuisine (RC for short). The both are constructed by a combination of automatic and manual methods. The links between FC and RC have link scores among zero and one. Notably, a large part of related concepts (4396 in 4720) represents impressions that crowd-sourcing workers annotated about cuisine (ex. 'thick,' 'plain' and 'suitable for kids' etc.). The lexical knowledge in this knowledge base covers 78.85% of 766 unknown-words in 6,000 user utterances collected in 'CookChat.'

3.2 Associative Dialog System using Cuisine Knowledge Base

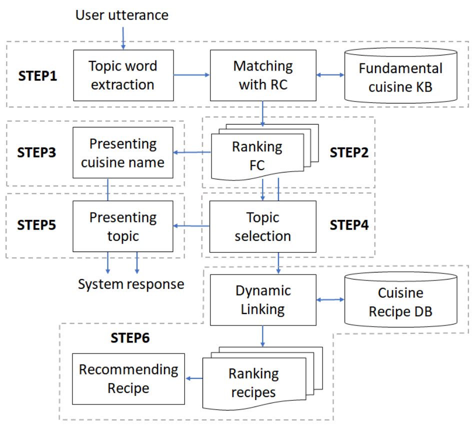

Figure 5: Diagram chart of associative dialog system

Figure 5 shows how the dialog system works. This system works according to following processing steps.

STEP1 Extract topic word from the user utterance, match it with RC in KB, then find FCs related to matched RCs.

STEP2 Update FC ranking using link scores (as user intention model).

STEP3 Present a few top-ranked FC's names to show how the system interprets user utterance.

STEP4 Select topics related to top-ranked FCs presented in STEP3.

STEP5 Present topics selected in STEP4 for exploring user preference.

STEP6 After a certain number of iteration of STEP1 ∼ STEP5, then link top-ranked FCs dynamically with cuisine recipes based on the similarity between FC's name and recipe title, and recommend a few top-ranked recipes.

Table 2 shows an example dialog [6] in the system. The system helps the user to become aware of a variety of cuisine preferences by presenting topics ('plain-tasted food,' 'Korean style food,' 'customizable in taste' and 'taste of Worcester sauce') sequentially. The topic sequence is controlled associatively by tracing their relations in the knowledge base.

We conducted user evaluation [7] of the system by subjective questionnaire about evaluation criteria of 'naturalness of dialog.' The result did not show any statistically significant

[5] In general, user intention model is part of dialog state

[6] Although the example is shown in English, the actual dialog is in Japanese

[7] We do not intend to make our system a text-based one although the subjects used a text-based interface in the evaluation.

Table 2: Example of associative dialog for recipe recommendation

User utterances and system responses	Internal processing
S: This is recipe recommendation chat system. What are you interested in about cuisine?	Initial response
U: I am interested in economy.	
S: Um, I guess 'hamburg steak made of tofu' or 'Korean dish of bean sprout' from 'economy'.	STEP1 $\sim$ STEP3
S: I would like to ask about 'plain-tasted food' or 'Korean style food'	STEP4 $\sim$ STEP5
U: I prefer a recipe using bean sprout.	
S: Um, I guess 'vegetable salad' or 'fried noodle' from 'bean sprout'	STEP1 $\sim$ STEP3
S: I would like to ask about 'customizable in taste' or 'taste of Worcester sauce '	STEP4 $\sim$ STEP5

advantages against the efficiency-first (i.e., non-associative) dialog system we developed before. We could not prove the associative dialog makes contextual cohesion better. However, some evaluators commented: " An associative dialog is also a good approach when my preference is vague", "The system can recommend cuisines beyond my expectation".

4 Summary

In this paper, we addressed an integrated framework of domain knowledge and dialog system for a recommendation of contents that meet the user's preference. Specifically, we proposed an idea of topic control based on that knowledge for exploring the user's intention. The feature of the knowledge base is that it holds rich topic concepts to obtain pieces of intention from the user who does not have enough knowledge of the dialog task domain. In our preliminary work, the fundamental cuisine knowledge base showed reasonably good coverage against utterances in recipe recommendation dialogs, and the associative dialog system with that knowledge base was evaluated as potentially helpful to remind the user of his/her unconscious preference.

We will try the following problems as future work.

Scale Up Domain Knowledge Base
We have to acquire concepts with a cost-effective method of crowd-sourcing. We will introduce gamification approach [Otani *et al.*, 2016] to do that.

Understanding Image
We will deal with language and image from the user with an integrated mechanism, not just with a combination of dialog system and image retrieval. We will introduce a machine learning model generating topics from the images to interpret language and image seamlessly.

Dialog Management
The plan-based approach [Allen *et al.*, 1995] is effective when the user has relatively concrete intention in every dialog turn. However, in our target task, the user does not always have the one. We will formalize a topic selection strategy to explore the user's potential intention.

References

[Allen *et al.*, 1995] James F. Allen, Lenhart K. Schubert, George Ferguson, Peter Heeman, Chung Hee Hwang, Tsuneaki Kato, Marc Light, Nathaniel G. Martin, Bradford W. Miller, Massimo Poesio, and David R. Traum. The TRAINS Project: A case study in building a conversational planning agent. *Jornal of Experimental and Theoretical AI*, 7:7–48, 1995.

[Bordes *et al.*, 2017] Antoine Bordes, Y-lan Boureau, and Jason Weston. Learning End-To-End Goal-Oriented Dialog. In *ICLR 2017*, pages 1–15, 2017.

[Harashima *et al.*, 2016] Jun Harashima, Michiaki Ariga, and Masayuki Ioki. A Large-scale Recipe and Meal Data Collection as Infrastructure for Food Research. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*, pages 2455–2459, 2016.

[Kiyomaru *et al.*, 2018] Hirokazu Kiyomaru, Sadao Kurohashi, Mitsuru Endo, and Katsuyoshi Yamagami. Fundamental Cusine Knowledge Base based on Cooking Recipe and Crawdsourcing. In *Proceedings of the 24th annual meeting of the Association for Natural Language Processing*, pages 662–665, 2018. (Japanese).

[Matsumoto and Ren, 2011] Kazuyuki Matsumoto and Fuji Ren. Construction of Wakamono Kotoba Emotion Dictionary and Its Application. In *The 12th International Conference, CVICLing2011*, pages 405–416, 2011.

[Misu *et al.*, 2011] Teruhisa Misu, Komei Sugiura, Hisashi Kawai, and Satoshi Nakamura. Modeling Spoken Decision Making Dialogue and Optimization of its Dialogue Strategy. *ACM Transactions on Speech and Language Processing (TSLP)*, 7(3):111–129, 2011.

[Ohtake *et al.*, 2009] Kiyonori Ohtake, Teruhisa Misu, Chiori Hori, Hideki Kashioka, and Satoshi Nakamura. Annotating Dialogue Acts to Construct Dialogue Systems for Consulting. In *Proceedings of the 7th Workshop on Asian Language Resources, ACL-IJCNNLP 2009*, pages 32–39, 2009.

[Otani *et al.*, 2016] Naoki Otani, Daisuke Kawahara, Sadao Kurohashi, Nobuhiro Kaji, and Manabu Sassano. Large-Scale Acquisition of Commonsense Knowledge via a Quiz Game on a Dialogue System. In *Proceedings of the Open Knowledge Base and Question Answering (OKBQA) Workshop*, pages 11–20, 2016.

[Yamagami *et al.*, 2017] Katsuyoshi Yamagami, Mitsuru Endo, Takashi Ushio, and Noriaki Horii. Development of Dialog Service Platform. In *The 31st Annual Conference of the Japan Society of Artificial Intelligence*, 2017. (Japanese).

[Yamagami *et al.*, 2018] Katsuyoshi Yamagami, Mitsuru Endoh, Hirokazu Kiyomaru, and Sadao Kurohashi. Associative Dialog System for Recipe Recommendation Using Food Knowledge. In *The 32nd Annual Conference of the Japan Society of Artificial Intelligence*, 2018. (Japanese).