Construction of a Dialog System for Talking Using a Topic Maps-Based Online Learning System

Shu Matsuura

Faculty of Education, Tokyo Gakugei University, Japan
E-mail address: shum00@u-gakugei.ac.jp

This study attempts to construct a dialog system to retrieve information from an online learning system. It uses a humanoid robot, NAO, as an interface to convey information.

The learning system was built based on a Topic Maps ontology (ISO/IEC 13250:2002). The dialog model was constructed using the development environment of Choregraphe 2.1.4. The utterance-response interaction between the human and the robot was constructed based on a three-step model comprising (1) human utterance, (2) information retrieval and talking, and (3) updating the list of words recognizable by the robot.

To reduce the misrecognition of the human utterance and allow the theme to be changed flexibly, we introduced a focus interest model that restricted the list of words that could be recognized within a focus of interest. In this model, the theme of the dialog was defined by a topic type, and the members of the theme were determined as its instance topics, which include the focus topic. The list of recognizable words, which is the theme of the dialog, is updated by a transition in the focus topic or the theme.

These two models enabled an efficient dialog for retrieving information from the online learning system. However, the fluency of the dialog depends on how well the user knows the topic map ontology.

Key words: Topic Maps, Online Learning System, Humanoid Robot, Nao

1. Introduction

Smart speakers and the communication robots are becoming our partners in daily life and in education because of their ability to converse. Natural language is expected to become a practical user interface for information systems. Designing dialogs for human-computer interfaces is an important research field [1].

Further, humanoid robots can use image sensing to identify people, and their gestures can produce humorous expressions. Human-like communication is desirable for elderly care, support for autism, and education [2].

In this study, we use a humanoid robot as a lecture partner [3]. It works as an interface to an online learning system, "Everyday Physics on Web" (http://tm.u-gakugei.ac.jp/epw/). We built the learning system based on a Topic Maps [4] ontology [5]. The ontology provides structured knowledge as the backbone of the learning system. In this study, the various types of associations that connect specific topics are used to retrieve information and develop a dialog between the human and the humanoid robot.

2. Method

The sociable humanoid robot used in this study is Nao (SoftBank Robotics). In the autonomous mode, Nao tracks a human's face and tries to keep looking at the person. It signals that it is listening to the human voice by changing its eye color and beeping.
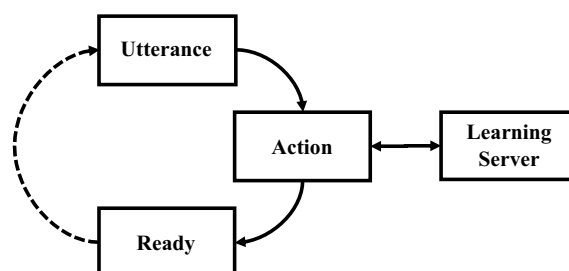

Fig. 1. Schematic illustration of the three-step model.

The programming environment for Nao used in this study was Choregraphe 2.1.4.13. Some memory modules were used to store the text data obtained from the learning system. These texts were spoken by the robot.

Human utterances trigger behaviors in the robot brain. We programmed all possible robot actions in the behavior module of Choregraphe. The dialog needed to trigger each behavior was defined using QiChat Script. The set of words used to signify a common notion is called a concept. Such a set of words will trigger a response or a behavior by the robot. Further, a dynamic concept can update the set of words of a concept at runtime. A dynamic concept can refresh the dialog or change the behavior.

To retrieve information from the online server, a behavior sends a topic map query to the server. The Topic Maps is a standard indexing technology. Construction of the topic map for the learning server began by creating hierarchically structured topic types for subjects, learning resources, web
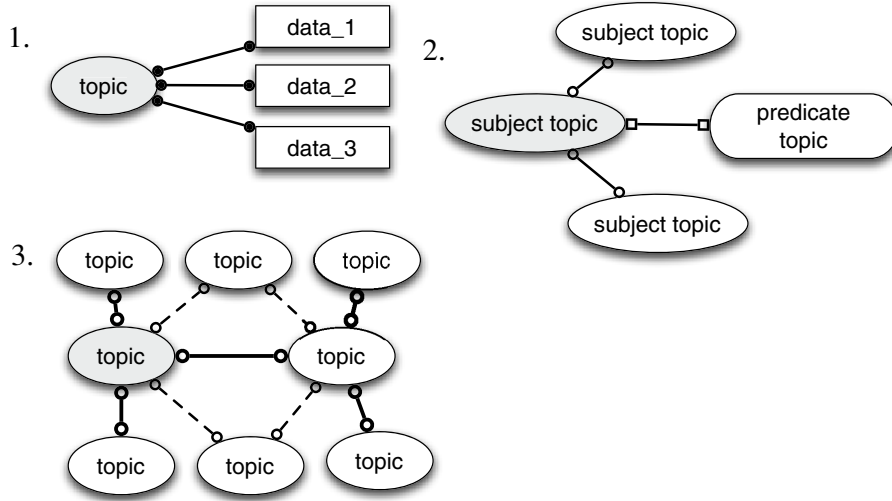
Fig. 2.   Three basic types of dialogs. A gray topic indicates the starting topic, which the robot recognizes. **1**: Occurrence type dialog. **2**: Subject-topic-to-predicate-topic-type dialog. **3**: Topic-to-topic-type dialog.

pages, and interactions with users. Topic instance of each topic types had occurrences to connect them with the actual web resources or internal data. The topic instances were interconnected with each other through typed associations. By defining unique types of topics, associations, and occurrences, one can construct a rich knowledge structure [6]. The topic map development environment herein, i.e., Ontopia, includes a built-in topic map query language, tolog and the topic map remote access protocol (TMRAP), which sends a tolog query through a URL query request.

3.   Results and Discussions
3.1   Basic dialog pattern
The basic dialog pattern is expressed as a mapping of a specific set of the human utterances to a set of the robot responses, as shown by this QiChat script:

u:($\sim$utterances) $\sim$ responses,

where the utterances are defined as a set of utterance words $w_{u1}, w_{u2}, \ldots, w_{uN}$, using the keyword "concept":

concept:(utterances) [$w_{u1}, w_{u2}, \ldots, w_{uN}$],

where [ ] expresses an exclusive OR list of words. Likewise, a set of response words $w_{w1}, w_{w2}, \ldots, w_{wN'}$ expresses the robot's response:

concept:(responses) [$w_{w1}, w_{w2}, \ldots, w_{wN'}$].

Here, one of the response words, $w_{w1}, w_{w2}, \ldots, w_{wN'}$, has to be chosen to utter.

When the robot application starts, the robot loads the default list of words into its memory to recognize the human utterances and initiate a dialog. As the number of words that the robot can recognize increases, the throughput of the robot also increases. At the same time, as the number of homophones or similar words increases, the number of the instances of misrecognition also increases.

In this study, we avoided uploading a considerably long list of candidate words that could be uttered by a human using a dynamic list for a concept, which could be updated at runtime:

$$dynamic : utterances$$
$$u : (\sim utterances) \sim responses. \qquad (1)$$

While retrieving information from the online learning system, the names of topics or associations were used as a set of words for robot recognition. The initial human utterance selected a particular knowledge domain for the dialog, such as "physics," "philosophy," "daily life," and so on.
3.2   Three-step model
Briefly, the dialog system consists of an input manager, dialog manager, and the output generator. The intermediate dialog manager is responsible for updating the internal state of the robot and deciding what action it should take [7].

Here, we model the process from the human utterance to the robot's response in a three-step model, as shown in Fig. 1. The utterance step is when the speech is recognized by the robot, and the ready step sets the dynamic concept of utterance used to recognize the next human utterance.

In the action-step, the robot's memory is refreshed. This step consists of querying the online learning server, updating the robot's memory based on the query results, and speaking. To retrieve topic map elements and their internal data, the robot requests a topic map query through TMRAP. The robot obtains the topic names of the related subjects, as well as an explanation of the subject depending upon the query.

In addition to the topic map, a relational database server that cooperates with the topic map stores the content data of the drills and the users' learning history. To retrieve information from the database, the robot requests the JavaServer Pages that contain specific database queries. First, the robot retrieves the parameters needed to specify the database query through TMRAP. Then, it sends a URL request to the database query page with the parameters.

According to the proposed three-step model, the dialog

Table 1. Examples of retrieving internal data through the occurrences of a topic.

Occurrence type	Meaning	Key phrase in the human utterance	Behavior
Description	Explanation of topic	u("explanation of" [~instance]) u([~instance]とは)	AskDialog
Era	Century	u("era of"[~instance]) u([~instance]の世紀)	AskEra
Date_of_birth	Birth date	u("the birth date of"[~instance]) u([~instance]の生年)	AskBirthDate

Table 2. Examples of topic retrieval through various types of association.

Association type	Meaning	Key phrase in the human utterance	Behavior
broader_narrower	Hierarchy steps	u("lower hierarchy of"[~subject]) u([~subject]の次の階層)	SetSubSubject
intraField_ is_related_with	Intra-field association	u("in relation to"[~instance]) u([~instance]といえば)	AskRelated
transField_ is_related_with	Inter-field association	u("in connection to"[~instance]) u([~instance]につなげて)	AskTransRelated
is_based_on	Base-application	u([~instance]"is based on") u([~instance]の基礎は)	AskBase
transField_ is_based_on	Trans-field base-application	u([~instance]"stems from") u([~instance]につながる基礎は)	AskTransBase
transField_ is_linked with_History	Related history	u("related history of"[~instance]) u([~instance]に関する歴史は)	AskHistory
Is_subject_ of_experiment	Experiment	u("experiment of"[~instance]) u([~instance]の実験は)	AskExperiment

script pattern (1) is modified and implemented as

dynamic : utterances

u :(∼ utterances)ˆrunBehavior (Behavior),   (2)

where the preceding behavior renews the dynamic concept for the utterance. Behaviors are required to access the learning server, generate a text for the response, speaking the text, and raise a memory event to set the dynamic concept in step 3.

The types of association that a behavior uses to send TM-RAP query requests characterize the behaviors in expression (2). Figure 2 shows three fundamental types of topic requests. Figure 2.1 shows the behavior of an occurrence request. The robot requests internal data from the server, such as the description of a starting topic (indicated in gray in the figure), which the human asked about in a dynamic concept. Figure 2.2 shows the predicate request type in which a predicate topic instance describes the starting subject topic. The subject topic is connected with the predicate topic by the is_characterized_by-type association. Figure 2.3 shows a request for the associated subject topics. In this case, one of the associated subject topics should be chosen by the succeeding dialog, with a list of associated topics stored in memory.

The types of behavior run in the action step are selected based on the key phrase added to the name of a topic instance in the human utterance. Thus, the utterance of the expression (2) can be rewritten as

u : ([dynamic concept] + (key phrase))

ˆrunBehavior (Behavior).   (3)

Relation (3) maps a key phrase with the behavior to search the topic and data through the right association. The name of the topic instance that the dynamic concept contains works as a parameter for the selected behavior. Table 1 lists some examples of the occurrence-type dialog with the key phrases. The key phrase selects and executes a specific behavior and queries the occurrence of the specified topic instance. In the table, an "instance" can be a set of topic instances, such as a dynamic concept, but it typically indicates a single focused topic instance (Fig. 3).

Table 2 lists examples of the association types used to search the related instance topics based on the corresponding behaviors; the table also lists the phrases uttered to call the behaviors. In the table, "subject" indicates a subject topic type. It is usually a supertype topic, as described in the next section (Fig. 3).

3.3 Focus interest model

During the three-step dialog, the robot holds a dynamic list of words, i.e., a list of topic names or internal data,
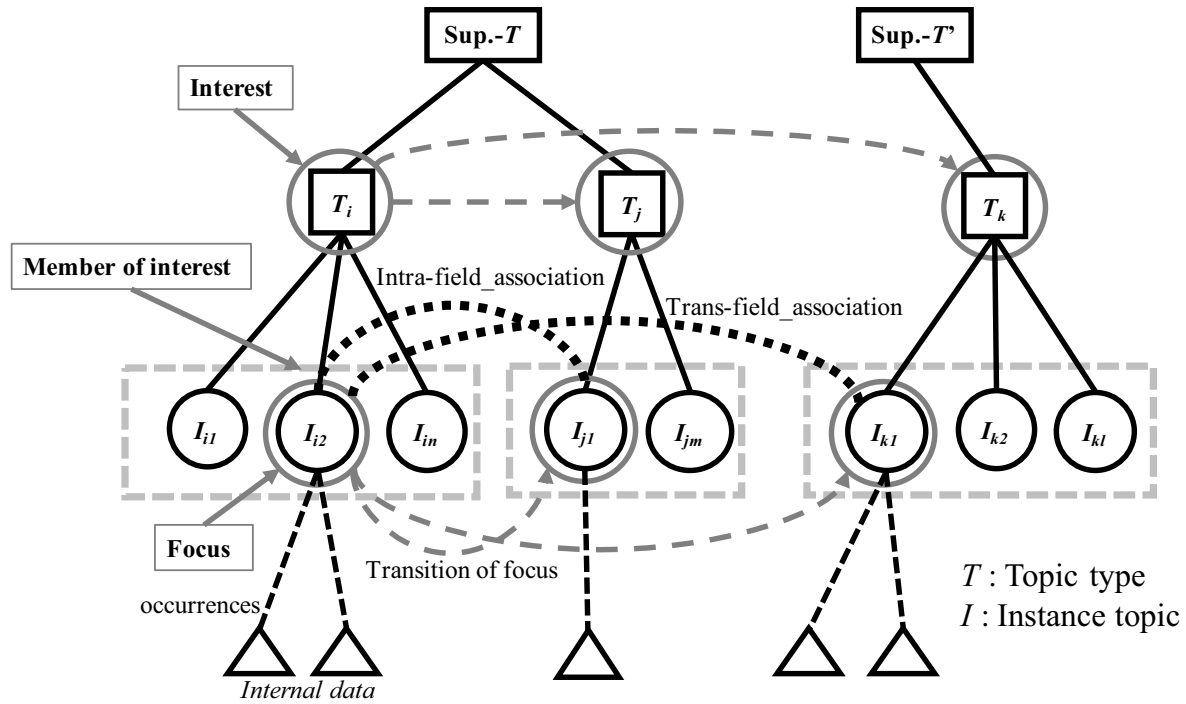
Fig. 3. Focus-interest model applied to a topic map. The letter "*T*" indicates the topic type, and the letter "*I*" indicates the instance topic. A solid straight line indicates a supertype_subtype association, and the dotted black curves indicate associations among instance topics. The black broken lines indicate occurrences that connect the instance topic with internal data.

which it can recognize and respond to. We now define the range of topics and data which are stored temporarily in the robot's memory. Here, we introduce a focus interest model as a representation of temporarily conscious words.

Figure 3 illustrates the focus interest set and its transition along an intra-field association and a trans-field association. In the figure, T_i indicates the ith topic type. If T_i and T_j have the same supertype, i.e., Sup.-T, they belong to a common field of knowledge. The topic instances of type T_i are written as I_{i1}, I_{i2}, ..., I_{in}. If the topic instance I_{i2} is retrieved, then we call I_{i2} the focus topic. The supertype topic T_i is the topic of interest. Further, the instance topics, I_{i1}, I_{i2},..., I_{in}, are the members of interest since they are instances of T_i. Thus, the temporary memory holds internal data for the focused topic I_{i2}, the topic of interest, and the members of interest. The dynamic concept is refreshed in preparation for recognizing words from it in the human utterance.

If the focus moves from I_{i2} to I_{j1} along with an intra-field association, the topic of interest changes to T_j with the members of interest I_{j1}, I_{j2},..., I_{jm}. In addition, if the topic of interest changes from T_i to T_k, the members of interest change to I_{k1}, I_{k2},... I_{kl}. Although the topic of interest can be chosen to be higher or lower in the category tree, the robot will prompt to narrow the scope of the field to the level at which the topic of interest directly has the instance topics as the member of interest.

3.4 Evaluation of dialog generation

Currently, the system has been tested only by a participant who has an expertise in the ontology of this learning system. First, the three-step model for dialog generation updated the robot's memory at every dialog and the list of rec-

ognizable words according to the human utterances. Second, the focus interest model ensured that both the topics within the field of interest and the topics directly associated with the focus topic were always recognizable during the dialog. Third, by changing the topic of interest, the user was able to change the entire topic at any time.

However, we have a major technical problem associated with language processing. Since many *kanji* have multiple pronunciations, the robot cannot identify a topic name unless the human pronunciation matches the robot's reading. The robot's *kanji* reading was in some cases inappropriate.

Furthermore, even though there is flexibility in selecting a topic, the user needs to have some knowledge of the members of interest topics under an interest topic type while talking. If the user is not accustomed to the ontology, the robot should be able to visualize or verbally navigate the theme. This kind of knowledge presentation or navigation is required as a future development.

4. Conclusions

We constructed a dialog generation system by linking a humanoid robot with an online learning system based on the topic map ontology. The dialog system developed in this study was intended for a dialog between a human lecturer and a humanoid, Nao. We proposed a three-step model to generate a dialog, which could update the terms recognizable by the robot during the dialog. Updating this term list in the robot's memory was also a transition in the collection of networked terms. To enable dialog content to be migrated, we proposed a focus interest model, which enabled the migration of a lineage of a set of recognizable topics.

Although this development meant that the dialog and the retrieval of information from the learning server through a humanoid robot were flexible, the users were required to know the ontology to some extent. The development of visual or verbal navigation of the knowledge is expected as future work.

References

[1] R. Luppicini, *Handbook of Conversation Design for Instructional Applications*, Information Science Pub., 2008.

[2] K. Dautenhahn, *et al.*, What is a Robot Companion—Friend, Assistant or Butler? *2005 IEEE/RSJ International Conference on Intelligent Robots and Systems* (2005) 1192–1197.

[3] S. Matsuura and M. Naito, Shaping Dialogues with a Humanoid Robot Based on an E-Learning System, 2016 11th International Conference on Computer Science & Education (ICCSE) (2016) 7–12.

[4] ISO/IEC 13250 Topic Maps, Information Technology, Document Description and Processing Language, Retrieved from http://xml.coverpages.org/TM-iso13250-2nd-ed-v2.pdf

[5] S. Matsuura, Development of a Trans-Field Learning System Based on Multidimensional Topic Maps, Leipziger Beiträge zur Informatik: Band XIX, Linked Topic Maps, University of Leipzig, Vol. 19 (2009) 83–89.

[6] S. Pepper, The TAO of Topic Maps—Finding the Way in the Age of Information, Retrieved from http://www.iro.umontreal.ca/˜lapalme/ift 6281/TopicMaps/TAOTopicMapsPresentation.pdf

[7] M. Nakano, K. Komatani, K. Funakoshi, and Y. Nakano, *Dialogue System, Natural Language Processing Series 7* (Japanese book), Corona Pub. Co. LTD., Tokyo (2015) p. 66.