



# Design of Coimagination Support Dialogue System with Pluggable Dialogue System - Towards Long-Term Experiment

Seiki Tokunaga(✉) and Mihoko Otake-Matsuura

Center for Advanced Intelligence Project, RIKEN,  
Nihonbashi 1-chome Mitsui Building, 15th floor, 1-4-1 Nihonbashi,  
Chuo-ku, Tokyo 103-0027, Japan  
{seiki.tokunaga,mihoko.otake}@riken.jp

**Abstract.** In this paper, we design a dialogue system which aims to support elderly people's cognitive function with a conversation by face to face. We have a hypothesis that a kind of group conversation method which is called *Coimagination Method (CM)* would have merit for a human at a perspective of cognitive function. However, current CM aims to be designed to provide balanced group conversation among participants. Hence, the method has mainly focuses the group conversation, hence it has some limitations such as user have to join the specific location. But, elderly people sometime face difficulty because of some physically or mental problems to do it. In this paper, we try to design a chat system which provides a one-to-one conversation between human and system which is based on the essence of CM and copes the above limitations. In order to design the system, at first we consider an experimental design which would be conducted remotely during long-term as a system requirement, then we design a system which meets the requirement. Moreover, we also develop a prototype system in order to confirm the feasibility of proposed system design.

**Keywords:** Conversation · Chat bot · Dialogue system · System design · Supporting elderly people

## 1 Introduction

We are facing a super-aged society, and Japan is forecast to become a society with 39.9% of aged people by 2060. Japanese government conduct a questionnaire that how do you feel about aging for elderly people [12]. As a result, over 73% of elderly people fear the problems which relates healthy aspect in their future. Because of the social background, we have to assist them both social system and methods in order to support elderly people with safe.

Although, there are many symptoms which related with aging, the dementia is one of the serious problem since the symptom is strongly effect of elderly

people's lives [4]. Dementia involves impairment of two or more areas of cognition, such as memory, judgement, abstract thinking and so on [1]. On the other hand, many researchers and companies try to develop a method which would be effective to maintain the cognitive function in the world. However, there is no effective method to maintain the cognitive function at an early stage. Hence, society want a method to maintain the cognitive function for elderly people who has not being a dementia.

Our research group tackles to develop a method to maintain a cognitive function for elderly at the perspective of conversation. Concretely, Otake et al. [16] have proposed a conversation method called *Coimagination Method (abbr. CM)*, that is a group conversation method in order to conduct a balanced conversation with sharing pictures and time management. Coimagination method have conducted at various locations (e.g. research institute, nursing home, community center and hospital). Currently, CM is originally designed to coordinate the balance of a group conversation, hence people have to get together a specific location in order to conduct it. However, there are many elderly people who cannot join such a conversation group because of they are hard either physically or mentally problems. In addition to this, when participants conduct a group conversation works either well or not depends on the group compatibility, namely someone who have same things, their opinions match as well. Hence, if we could have a method to provide a way of conversation as if we talk human without gathering the group, this would be helpful to maintain the cognitive function.

Dialogue systems are widely developed since both development of the technologies. For instance, some smart speakers such as Amazon Alexa and Google Home are widely spread in the home. A wide variety of dialogue system is currently developed for individual purpose and tasks [6, 19]. Also, dialogue systems have several techniques in order to meet their purpose such as rule based type [13, 15], neural network based Question-Answering type [17], and so on. Whereas conventional dialogue system aims to improve the performance such as how they can reply for human and how long they can talk with human, hence many conventional researches are aimed to increase the accuracy namely they are technology driven approach. On the other hand, some researchers have tried to develop dialogue systems for elderly people [22]. Although, the dialogue system has potential to provide a conversation for elderly people instead of actual human, no one has shown the effective result of that.

In this paper, we design a dialogue system which aims to maintain a cognitive function for elderly people. Our system's main target users are elderly people who could not communicate with other people, for instance who lives alone and whose family lives far from elder's home. To develop a system in order to meet our needs, first we design an experiment to confirm the effect of proposed system. As the first step, we apply the simple Question-Answering techniques in order to maintain the data with efficiency and easily way.

The outline of paper is as follows. In the next preliminary section, we explain domain knowledge that are about dementia and how to conduct experiment with actual subjects. We also explain the essence of experiment which aims to reveal

the effect of conversation between system and human at the cognitive function perspective. Moreover, we explain the dialogue system and applying the one-to-one conversation with chat bot that is based on the CM. In Sect. 3.1 we analyze a system requirement in order to conduct the experiment which aims to find out the effect of the proposed method, and then we describe the system design and functionalities. In Sect. 4 we describe the system design of detail such as data format between systems and how to store the information data. In Sect. 5, we develop a prototype system and show the feasibility of the proposed system. Then in Sect. 6, we have discussed our system's feasibility and some limitations based on the prototype system's performance which have developed.

## 2 Preliminary

### 2.1 Risk Factor of Dementia and Intervention Method

Dementia is a symptom that declines cognitive function because of the brain damaged. Concretely speaking, memory, communication and language, ability to focus and pay attention, reasoning and judgement and visual perception. To intervention for dementia some risk factors are known as serious for life stage. Concretely speaking, smoking (5%), depression (4%), physical inactivity (3%) and social isolation (2%) contribute for dementia in late life [14]. Especially, people may loss the chance of conversation when they are in social isolation situation. In addition to this, some researchers have researched a relationship between cognitive function and the intervention method. Suzuki et al. have shown that the intervention method by a picture book reading has significantly difference that the rate of memory retention of the intervention group improved after the program completion [18].

### 2.2 Dialogue System

*Dialogue system* provides communication between system and human over the telephone [3, 20, 21, 23]. Because of the development of speech-to-text technologies [10] that enables to input the voice data and outputs as a text which is a transcription of voice data. Currently, also many products have provided as the voice user-interface such as Amazon Echo and Google Home and so on [8]. Hill et al. have researched how communication changes when people communicate with an intelligent agent as opposed to with another human [7].

Generally, dialogue system handles a user's utterance based on similarity, and the system replies the most reliable answer based on the data which we have preliminarily registered. David et al. have developed a digital system that allows people to have an interactive conversation; whose goal is to preserve as much as possible of the experience of the face-to-face interaction [19]. They have also conducted an experiment, as a result most user questions can be addressed and that the audiences are highly engaged with the interaction.

We explain an example of the procedure of user's utterance with existing dialogue system Watson Assistant [8]. Before talking with Watson Assistant,

we have to store a pair of both questions and answers into a workspace which is a logical space that defines the conversation flow preliminar<sup>1</sup> the as training datasets. When a user asks a question to the Watson Assistant, then the system responds an answer which corresponds to the question by the similarity of the text. And the training data is preliminary registered into the system with some pairs of questions and answers. If Watson Assistant cannot respond an answer for a user's question, then the system replies error sentence such as "I am having trouble understanding right now".

### 2.3 Coimagination Method

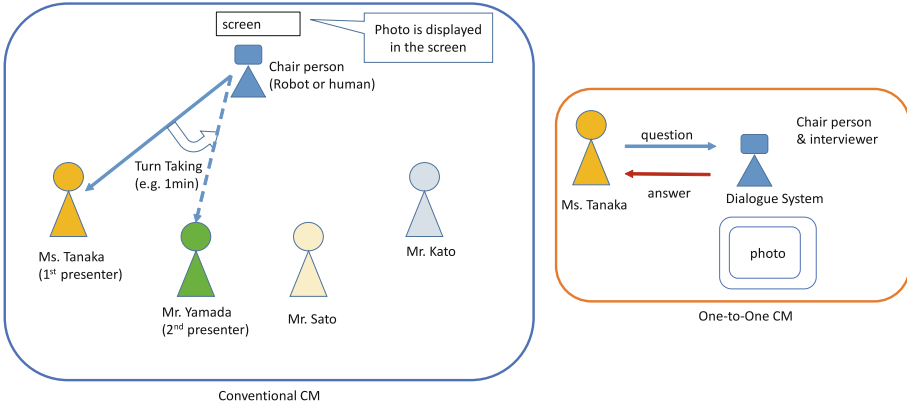
Otake et al. have proposed a group conversation method called *Coimagination Method (CM)* which is designed to prevent the decline of cognitive functions by a well-managed conversation [16]. Concretely speaking, participants take a photo before talking an according to a theme which is selected so as to everybody can join the conversation easily (e.g. favorite food). In summary, CM mainly consists of time-management, the topic which the participants talk and taken photo by individual. In addition to this a conversation unit is defined as *session*, namely the participants take a conversation with looking at the pictures that are taken by themselves. CM consists of two main phase, one is a topic of conversation phase, which the participants take a conversation based on a decided theme. In the topic conversation phase, the participants talk about the photo by sequentially, namely the first speaker finish his/her talk, then the second speaker should talk. The other is a question and answering part, in that part the participants ask to the presenter with sequentially, hence the participants finish asking to the first presenter, then ask to the second presenter. After the session, the participants summarize their conversation with short sentence at most 200 words in Japanese, we name the summarized conversation as *story*.

Figure 1 shows a picture that people join a CM. The blue robot represents a chair person in the top of the Fig. 1. Currently, CM is widely conducted such as research institute, nursing home, hospital and so on in Japan. However, current CM aims to be designed to provide balanced group conversation among participants. Hence, the method has mainly focuses the group conversation, it has some limitations such as user have to join the specific location because of their some physically or mental problems. So if we could develop a system with which the elderly people could conduct the conversation without considering location problems and that might be helpful. Considering the CM, as we already introduced that CM consists of two main parts, one is a topic of conversation part and another is a question and answering part. When we want to implement the two features, we expect that the dialogue system behave as follows.

1. Question and Answer part: System talks about a specific topic, then replies to user's question

---

<sup>1</sup> Strictly speaking, in Watson system has some domain terms *intent*, *entity* and *so on*, but to explain the essential and to avoid complexity, we have skipped the detail of it in this paper.



**Fig. 1.** Image of coimagination method (Color figure online)

2. Understanding part: User talks about a specific topic, then the system asks users to answer the question.

In this paper, we focus on the former question and answer part. Because we have a hypothesis that implementing the feature of questions and answer part is easier than understanding part.

**Table 1.** Design of experiment terms

Terms of experiment	Content
E1. Experimental period	At least 3 months
E2. Number of subject group	2 (Intervention/Control)
E3. Location	Laboratory and individual home

### 3 System Design

#### 3.1 Design of Experiment

In order to confirm an effect to with proposed system, we design how to conduct a feasibility experiment. Table 1 represents the abstract of the feasibility experiment.

At first, we consider that because we have to confirm effect for the cognitive function. In order to confirm the effect of proposed method, we would like to continue at least three months (See Table 1). In order to evaluate the proposed method, we adapt randomized controlled trial (RCT) which separates two groups one is called an intervention group and the other is called a control group, and the RCT method aims to reduce the bias of testing [2]. In addition to this, we

would like to conduct two kinds of locations; one is a laboratory where we can easily support the participants. The other is an individual elderly people's home or nursing home where we would like to conduct the long-term experiment.

### 3.2 System Requirements for Experiment

Our final goal is to find out whether the chat with system may helpful to maintain the cognitive functions. In order to achieve the goal, we consider the requirements in order to develop the system, and we describe the individual requirement in detail.

**R1: Easy Remote Operation by User's Identification.** In order to conduct the experiment remotely, system should be able to have a traceability to track the user's status such as login/logout. Moreover, the user should have user\_id in order to identify who currently use the system and track the operation's log in detail. This system requirement strongly supports the E1 and E3, when the experimental location would be conducted at elderly's home. In addition to this, the system should handle individual user's question. Moreover, we have to organize individual participants during an experiment, because we have designed to separate the participants as two groups.

**R2: Handling Status of the System.** First, the system should handle the status of network, such as currently the client device is online or offline. Basically the system is operated online, however we have to consider if the system becomes offline in order to avoid the process down or something system failure. Moreover, if the system could handle user's status such as login/logout the information would helpful at the system operation perspective. For example, if the user cannot login with some reasons (i.e. forget how to login the system) and the user's login status does not change at the operation side, the system operator would help with checking the user's status.

**R3: Pluggable Dialogue Systems.** The system is able to adapt for several dialogue systems in order to compare which is better choice for the CM's domain. Hence, the system is expected to handle the dialogue systems information such as a workspace. Moreover, the system should store both user's utterance and system's response with a timestamp in order to keep a traceability at the experimental operation perspective.

**R4: User-Centered Design for Elderly People.** The UI of the application should be considered for elderly people carefully. For example, the font size and the size of picture should be big enough to see. In addition to this, the necessary information in the CM domain which includes the limitation of the time and both the transcription of user's utterance and system's utterance should be appeared clearly in the screen. The system should display the essential information which

belongs to CM such as photos, question answering time and the story of photo. Moreover, the both user’s utterance and system’s response should display as the text, because sometime speech-to-text program may mistake, and also some elderly people may have hard of hearing. The usability should be maintained in order not to decrease the motivation of the participants. This is not a functionality requirement, but this is an important to conduct the long-term experiment. Moreover, the system should respond within a certain time, because too late response is difficult to recognize especially for elderly people.

### 3.3 System Design

At first, we design that the application consists of two parts; one is a native application part and the other is a web application which handles user’s utterance and requests to dialogue systems, after receiving the answer from dialogue system, then the system forward the reply to the user. The separations of the application have a technically reasons that as the native application should operate some events such as a starting to speak the story with showing the picture, then starting timer for question-answering time and so on. Hence, these events should be handled in the native application side, because the timer events are strongly related to the UI. In addition to this, the web application which exchange the data with the dialogue systems should be passive, because which have to be adapt many dialogue engines, hence the system’s inner logic should be kept simple in order to maintain with low cost way. We design these system architectures based on the system requirement. Figure 2 shows an overview of proposed architecture. The left of Fig. 2 represents a native application which is called Native Application of Coimagination Human-centered Orchestration

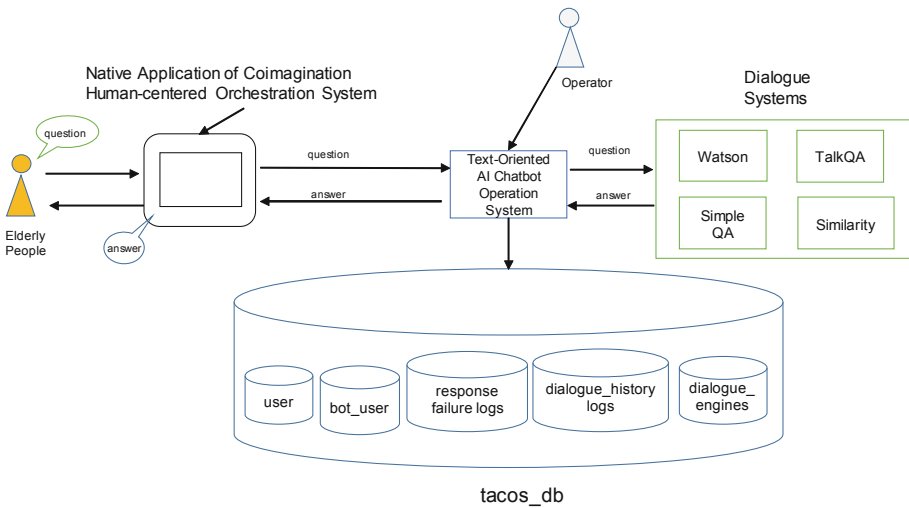


Fig. 2. System architecture overview

System (NACHOS). NACHOS provides UI which is based on CM, and also includes a photo, a story of photo that are stored in the database, a chat like UI to confirm his/her conversation with the dialogue system. The center of Fig. 2 shows a web application, we name it as *Text-Oriented Artificial Chat Operation System (TACOS)* that bridges the native application and the dialogue systems, and finally a square in the right of Fig. 2 represents the dialogue systems which include not only commercial one also original one if we develop a new dialogue system. Also the experimental operator checks and supports if the participants need to support situations such as some system accidents.

In order to transmit and receive between a native application and a web application, we define a common data format (See in Sect. 4.3). Because of the separation and its common data format provides easily to extend the system. The dialogue systems are integrated to reply from a user’s question, we consider to integrate the dialogue systems such as IBM Watson Assistant. In addition to this, TACOS enables us to add flexibility to add dialogue system if we also want to add a new dialogue system. The software separation between a native application and a web application provides easily traceability at the system operation perspective, because when a user login the native application then if the native application transmits the user’s status to a web application. Then an experiment operator could easily to catch up the situation.

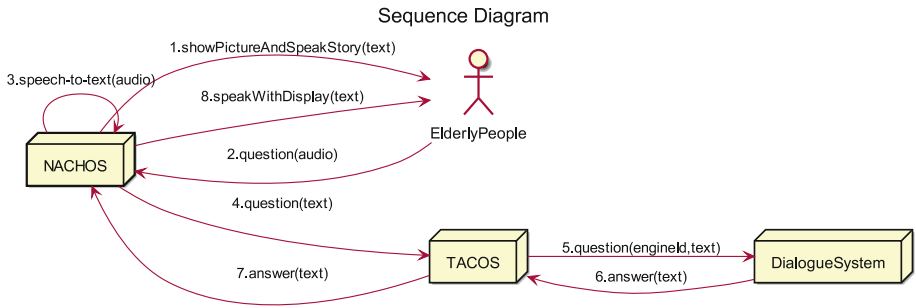


Fig. 3. Sequence diagram with proposed architecture

## 4 System Design Detail

### 4.1 Sequence Procedure of System

Figure 3 shows a sequence diagram which starts from elderly people’s utterance and ends that the system replies to the elderly people through the dialogue system. At first a native application (NACHOS) shows a picture and speaks a story to the user. Once a user speaks to the NACHOS, then the NACHOS converts audio data into text by the speech-to-text technology. Then the data shall be transmitted to the TACOS, then the TACOS forwards message to the



dialogue system. The messages are transformed into specific data format which is described in Sect. 4.3. After TACOS forwards data, namely the TACOS requests to the dialogue system, then the TACOS should receive a response which dialogue system replies. In the following, we have explained the data format, and how to provide identification for each user’s utterance (Sect. 4.3).

### 4.2 Native Application: Design of Native Application of Coimagination-Driven Human-Centered Orchestration System (NACHOS)

We design a native application called Native Application of Coimagination-driven Human-centered Orchestration System (NACHOS). NACHOS would be able to apply domain of CM which photos data which are and time management of questions and answers. Figure 4 is a design of native application user interface. We design that the application’s component is displayed above an elderly people so that the elderly people could easily to look at the UI. In the top left of Fig. 4 is the photo data which is the source of conversation with the dialogue system. Also the topic of photo is appeared bottom of the photo, this information also becomes the source of conversation. The bottom of the left of Fig. 4 shows the progress bar which represents the time for left in 1 session, and usually for the session max time is set as from 1 to 2 min for questioning and answering time in CM.

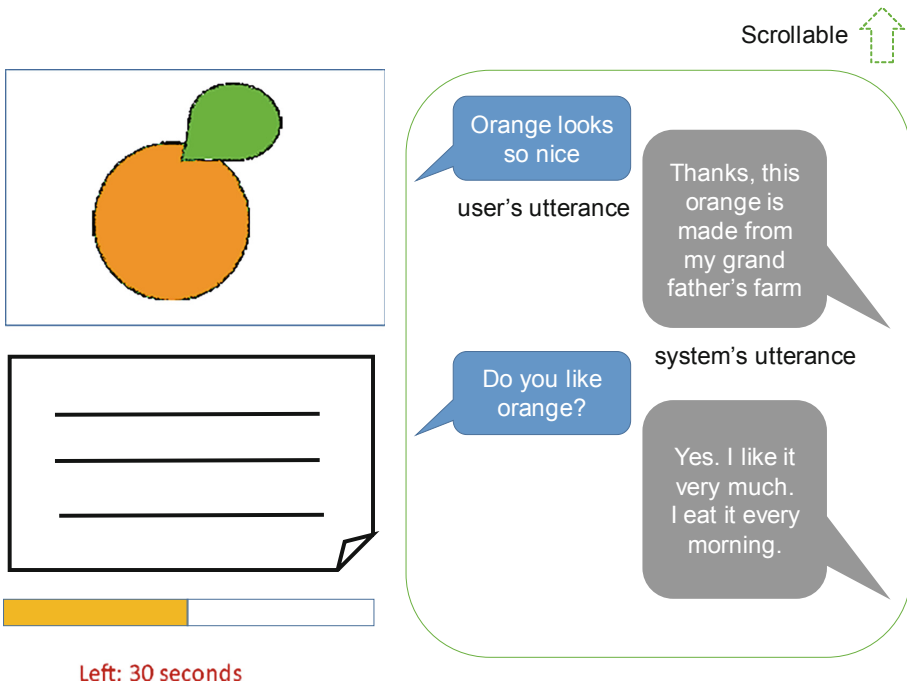


Fig. 4. UI design of NACHOS (Color figure online)

A user can question with looking at the photo in the screen with the native application. In addition to this, user and response from the dialogue engines are appeared as chat like interface. Hence, user can easily to follow what he/she says and what responses from the system.

### 4.3 Web Application: Design of Text-Oriented Artificial Chat Operation System (TACOS)

Text-oriented Artificial Chat Operation System (TACOS) is an intermediate server program which bridges between native application and dialogue system. Specifically, when the mobile app transmits questions data into TACOS, then TACOS forward question to dialogue systems. After that, the TACOS receives a response from dialogue systems then replies to the user as both text data and audio data which is generated by text-to-speech technology. At first, we explain how to TACOS deal with individual user. A user questions by NACHOS, then the data is converted as the specific data format as well as the question belongs to a user, namely the transformed data should have the user\_id which identify the user. We design that each user should have logically separated data channel in order to avoid the race condition while handling the user's data, thus the data transmission between NACHOS and TACOS used a data channel by the individual user. This design guarantees that if the TACOS receives from multiple users at once, then the system exactly replies to the individual user by using own channel. In addition to this, TACOS have a database which aims to have a user management, dialogue systems information and the history of chat in order to analyze the experimental context.

**Exchanged Data Format.** At first, we explain the data format which interchanges three system NACHOS, TACOS and dialogue systems. List 1.1 represents the data format which is transmitted by NACHOS to TACOS. We have applied data format as JSON format [9] and the body of JSON is written by according to Google JSON Style Guide [11] in order to develop and maintain with the readability of the exchanged data between systems.

**Listing 1.1.** Data format (NACHOS to TACOS)

```

1 {
2   "userId": "RX123_tokunaga",
3   "messageId": 1547122048 ,
4   "appName": "1on1FonodialApp",
5   "userUtterance": "Hello",
6   "createdAt": "2019-01-10 21:07:28.654",
7   "scheduleId": "123456"
8 }
```

- `userId` is a string term which aims to identify the owner of the data
- `messageId` should be a number and the number is generated by Unixtime with milliseconds
- `appName` is a fixed string term that represents which application emits this data
- `userUtterance` represents a text data which is generated with speech-to-text technology from user’s voice data
- `createdAt` is a date with timestamp which shows when this message generates by the application
- `scheduleId` aims to identify the conversation schedule between the user and the system.

**Listing 1.2.** Cloud replies to native application

```

1  {
2  "userUtterance": "Hello, where is here?",
3  "output": {
4  "contents": {
5    [
6    {
7      "systemUtterance": "Hi, welcome to west world.",
8      "confidence": 0.65,
9      "context": undefined,
10     "engine": "watson",
11     "contentType": "text"
12    }
13   ]
14  },
15  "replyTo": 1547122048,
16  "createdAt": "2019-02-01 14:40:52.457",
17  "userId": "awesomeBotId"
18  "messageId": 1548999697592,
19  "type": "replyMessage"
20  }
21  }

```

List 1.2 shows a response from TACOS to NACHOS. Specifically, we have designed that the data format has a field whose name is *output* to hold response which request and replies from dialogue system. The output field is mainly consists of two parts; one is a field whose name is *contents* which stores data from dialogue system and the other is a generic fields such as *replyTo*, *createdAt*, *userId*, *messageId*, which are set in order to maintain a traceability of the data at the operation and analysis perspectives. And the *type* represents a response type currently the type is only supported “replyMessage” only. But the type may be added near future, hence we set the data in order to cope the expected changes. The former part is depends on the dialogue system for example, some dialogue system has a confidence which represents the accuracy of the response, but other dialogue systems may not have. Hence, we separate the dependent

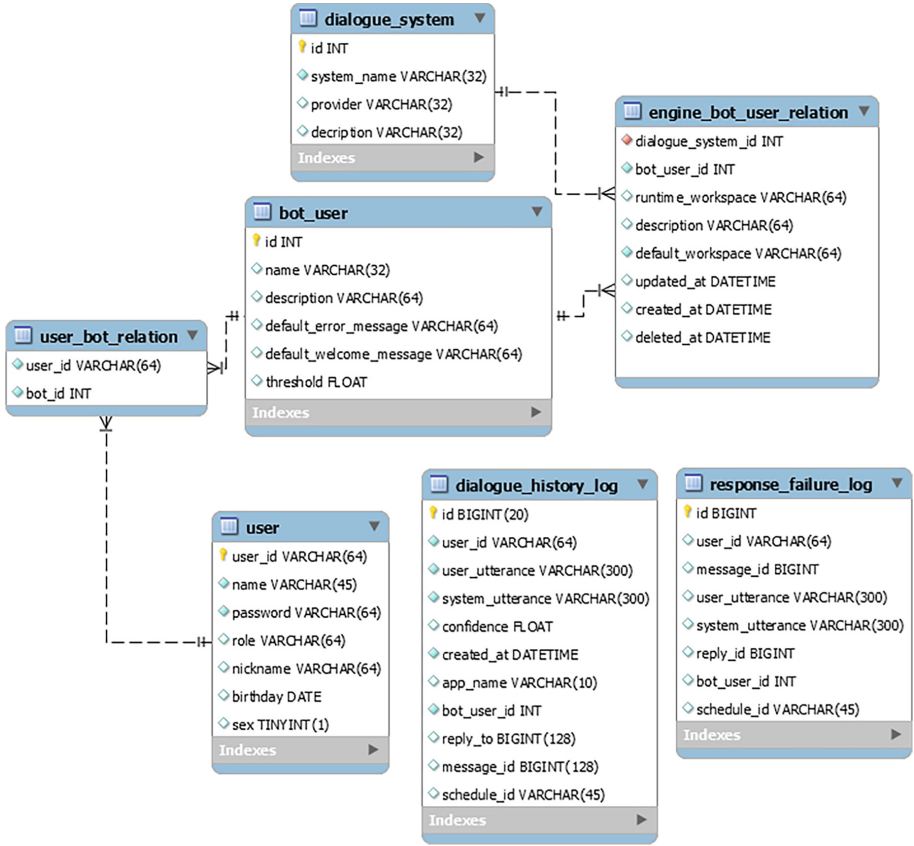


Fig. 5. ER diagram of TACOS (Color figure online)

part as a field to contains. In order to reply multiple responses, we define it as a square bracket [], hence when the dialogue systems reply long message then the message is split several small messages.

#### 4.4 Database Design

Figure 5 represents a ER diagram of TACOS. We describe the ER diagram with crow’s foot notation [5]. A box represents an entity (i.e., table), consisting of an entity name, a primary key and attributes, and the primary key is represented as a yellow key icon. A line represents a relationship between entities, where ++ -- ∈ denotes a one-to-many relationship.

User table stores a basic user information such as, username, nickname, password and so on. The bot\_user table stores the dialogue system’s domain information such as a default welcome message, a error message and a threshold Also, the bot is indirectly connected to dialogue system via engine\_bot\_user\_relation table.

Moreover, the user and bot\_user is one-to-one relationship by user\_bot\_relation. Hence, when we change the relationship with user and bot\_user, then we can easily switch the dialogue system, because a bot\_user is connected to a dialogue system. The *dialogue\_history\_log* table stores a general record such as user's utterance, the response which replies by the dialogue system and the created\_at which refers when the data emits by the native application (NACHOS). Moreover, *response\_failure\_log* stores error response which represents that the dialogue system cannot find an appropriate response. For example, when the dialogue system cannot find an appropriate response for the user's utterance, then most of the dialogue systems response a typical error response such as "I cannot understand you". However, we would like to handle the error response uniformly, because we have to store the data into *response\_failure\_log*, hence the *bot\_user* table has threshold in order to control without depending dialogue engines.

#### 4.5 How to Provide Pluggable Mechanism for Dialogue Systems

**Listing 1.3.** Template of General Dialogue System Response

```

1  {
2  "userUtterance": ${userUtterance},
3  "output": {
4  "contents": {
5    [
6    {
7      "systemUtterance": ${systemUtterance},
8      "confidence": ${confidence},
9      "context": ${context},
10     "engine": ${engine},
11     "contentType": ${type}
12    }
13   ]
14  },
15  "replyTo": ${replyTo},
16  "createdAt": ${createdAt},
17  "userId": ${userId},
18  "messageId": ${messageId},
19  "type": ${type}
20  }
21  }

```

List 1.3 represents a general JSON data which aims to represent the general dialogue systems response. hence which absorbs the different type format of dialogue systems. The enclosed `${}` represents a variable which is replaced at runtime by TACOS and the rule of description is based on JavaScript Template literals. So



Fig. 6. Screenshot of prototype system

once you write some programs which convert the response of the dialogue system into the template JSON, then TACOS could handle a new dialogue system because which has a rule how to replace the template JSON with actual values. Especially the term of *contents* has enough flexibility and it does not depend on a specific dialogue system. Hence, the unified template provides pluggable mechanism for the wide variety of dialogue systems, hence we can easily to add and compare some dialogue systems with ease.

## 5 Prototype System Development

Figure 6 shows a screenshot of prototype system. We have developed TACOS as a prototype system in order to confirm the feasibility of the proposed architecture. Specifically, as the backend system, we use a dialogue system as the Watson Assistant because it provides a lot of Web APIs. The left balloon shows messages that the user's questions. In order to avoid the complexity and easy to debug we have developed a input form in order to ask a dialogue system as text. Also the balloons of left represent responses from the dialogue system. The user's message and the system response is one to one corresponding; hence users can quickly follow what the system responses for the user's message. The number under the text of right balloon shows a confidence value that the dialogue system guarantees and actually the Watson Assistant replies the confidence data with Web API.

## 5.1 Verification with Sample Dataset

We have registered both questions and answers as the dataset to Watson Assistant. The registration number of datasets are 250 pair of questions and answers, and we have collected pairs of questions and answers with crowdsourcing. When we ask the task for the crowd worker, we use the collected stories which were created by elderly people who joined the CM. We create some rules for the crowd worker as follows to conduct crowdsourcing in order to collect appropriate dataset.

- (Description to crowd worker) The purpose of this task is to create question answering datasets which aims to create a conversation data with a chat bot, and when you create the conversation data you should follow following rules.
- Rule1. Please create a pair of question-answer dataset, namely first you ask a question and then the bot answers.
- Rule2. Please create a pair of the question-answer data based on the story which we have provided.
- Rule3. The bot has a persona; hence you should follow the persona when you create an answer dataset as the robot. Concretely speaking, the bot’s persona are as follows; the sex is a man, age (about 70 years old), has a humorous and likes drinking and so on. Moreover, when you create the bot’s answer, please make it as a polite message and also the message should have political correctness and should not include political message, and romantic commitment.

In addition to this, we also have conducted the system verification with given the dataset. Specifically, we have tested 250 questions for Watson Assistant in order to confirm that the system whether replies the expected answer or not. As the result, we have found that 7 answers are wrong which corresponds to 2.8% for total dataset. Moreover, we also have measured the response time from the system, we have tested 250 questions into the system. The response time is evaluated from requesting a question data (e.g. “Do you like drinking?”) to receiving the answer data such as “Yes, I like beer”. As the result, we have confirmed that the system has processed 0.78s per 1 question.

## 6 Discussion

We have confirmed the dialogue system sometimes replies wrong answers. This is because same questions but which have other answer datasets are generated via crowdsourcing. The same questions and different answers pairs are registered, then the system would be expected to response one and the other data would not be used. When we operate dialogue systems for a long time, then the problem is expected to be increased. One thing to prevent the duplicate question answers; we can consider a checking mechanism which notifies to a user that when registration a pair of question answering dataset whether the questions are preliminary registered or not. The other idea is that a mechanism which alerts to a user that the duplicate question answering datasets have registered.

In addition to this, in this paper we have not tested for actual elderly people, hence we would like to show how the proposed architecture is able to cover the user's question based on the registration datasets. Finally, we have confirmed that the response time from the specific dialogue system (Watson Assistant) is less than 1 s the average is 0.78 s. Thus, we think the feasibility at the perspective of real-time is enough based on the response result.

## 7 Conclusion

In this paper, we have designed a new dialogue system architecture which aims to maintain cognitive functions for the elderly people. At first, we have described the requirements of experiment in order to achieve the goal. Then, based on the requirements, we have extracted requirements for the system. Next, we have designed both a native application (NACHOS) and a web application (TACOS) in order to meet the requirements. In addition to this, we have developed the prototype system in order to show the feasibility of proposed architecture. Moreover, we have discussed why dialogue system has wrongly answered based on the preliminary small experiment. Also, we have discussed the feasibility at the perspective of real-time and that of response rate. Our future work is to develop a system which is based on the architecture that we have proposed in this paper. We also have to conduct and confirm the effect of proposed system for the cognitive functions for the elderly people.

## References

1. Bowen, J., Teri, L., Kukull, W., McCormick, W., McCurry, S.M., Larson, E.B.: Progression to dementia in patients with isolated memory loss. *Lancet* **349**(9054), 763–765 (1997). <http://www.sciencedirect.com/science/article/pii/S0140673696082566>
2. Chalmers, T.C.A., et al.: A method for assessing the quality of a randomized control trial. *Control. Clin. Trials* **2**(1), 31–49 (1981)
3. Barker, D.J., Van Schaik, P., Simpson, D.S., Corbett, W.A.: Evaluating a spoken dialogue system for recording clinical observations during an endoscopic examination. *Med. Inf. Internet Med.* **28**(2), 85–97 (2003). <https://doi.org/10.1080/14639230310001600452>
4. Dunkin, J.J., Anderson-Hanley, C.: Dementia caregiver burden. *Neurology* **51**(1 Suppl. 1), S53–S60 (1998)
5. Dybka, P.: Crow's foot notation. <https://www.vertabelo.com/blog/technical-articles/crow-s-foot-notation>
6. Higashinaka, R., et al.: Towards an open-domain conversational system fully based on natural language processing. In: COLING (2014)
7. Hill, J., Ford, W.R., Farreras, I.G.: Real conversations with artificial intelligence: a comparison between human-human conversations and human-chatbot conversations. *Comput. Hum. Behav.* **49**, 245–250 (2015). <http://www.sciencedirect.com/science/article/pii/S0747563215001247>
8. IBM: IBM Watson Assistant. <https://www.ibm.com/cloud/watson-assistant/>
9. I.E.T.F.: The Javascript object notation (JSON) data interchange format. <https://tools.ietf.org/html/rfc8259>



10. Google Inc.: Cloud speech-to-text. <https://cloud.google.com/speech-to-text/>
11. Google Inc.: Google JSON style guide. <https://google.github.io/styleguide/jsoncstyleguide.xml>
12. Japanese Government: White Paper for Eldery People
13. Leuski, A., Patel, R., Traum, D., Kennedy, B.: Building effective question answering characters. In: Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue, SigDIAL 2006, pp. 18–27. Association for Computational Linguistics, Stroudsburg(2006)
14. Livingston, G., et al.: Dementia prevention, intervention, and care. *Neurology* **390**, 2673–2734 (2017)
15. Misu, T., Kawahara, T.: Speech-based interactive information guidance system using question-answering technique. In: 2007 IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP 2007, vol. 4, pp. IV-145–IV-148, April 2007
16. Otake, M., Kato, M., Takagi, T., Asama, H.: Development of coimagination method towards cognitive enhancement via image based interactive communication. In: RO-MAN 2009 - The 18th IEEE International Symposium on Robot and Human Interactive Communication, pp. 835–840, September 2009
17. Serban, I.V., Sordani, A., Bengio, Y., Courville, A.C., Pineau, J.: Hierarchical neural network generative models for movie dialogues. CoRR abs/1507.04808 (2015)
18. Suzuki, H., et al.: Cognitive intervention through a training program for picture book reading in community-dwelling older adults: a randomized controlled trial. *BMC Geriatr.* **14**(1), 122 (2014). <https://doi.org/10.1186/1471-2318-14-122>
19. Traum, D., et al.: New dimensions in testimony: digitally preserving a Holocaust survivor’s interactive storytelling. In: Schoenau-Fog, H., Bruni, L.E., Louchart, S., Baceviciute, S. (eds.) ICIDS 2015. LNCS, vol. 9445, pp. 269–281. Springer, Cham (2015). [https://doi.org/10.1007/978-3-319-27036-4\\_26](https://doi.org/10.1007/978-3-319-27036-4_26)
20. Truong, H.P., Parthasarathi, P., Pineau, J.: MACA: a modular architecture for conversational agents. In: Proceedings of the 18th Annual SIGdial Meeting on Discourse and Dialogue, pp. 93–102. Association for Computational Linguistics (2017). <http://aclweb.org/anthology/W17-5513>
21. Walker, M.A.: An application of reinforcement learning to dialogue strategy selection in a spoken dialogue system for email. CoRR abs/1106.0241 (2011)
22. Wolters, M.K., Kelly, F., Kilgour, J.: Designing a spoken dialogue interface to an intelligent cognitive assistant for people with dementia. *Health Inf. J.* **22**(4), 854–866 (2016)
23. Zhang, S., Dinan, E., Urbanek, J., Szlam, A., Kiela, D., Weston, J.: Personalizing dialogue agents: I have a dog, do you have pets too? CoRR abs/1801.07243 (2018)