

A Chat Information Service System using A Humanoid Robot

Atsushi Hoshino and Kazuhiko Kato
Institute of Information Science and Electronics
University of Tsukuba
 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8573, Japan

Johane Takeuchi and Hiroshi Tsujino
Honda Research Institute Japan, Co., Ltd.
 8-1 Honcho, Wako, Saitama 351-0188, Japan

Abstract—In this paper, we describe development of an information service system using a humanoid robot. This system provides information such as topical news articles through conversations. Our goal is to evolve a conversational robot that can carry on conversations and enlighten the user. We think that achieving fun conversations is the basis for creating real symbiotic robots. This robot provides news information whether users explicitly ask for it or not during the conversation in order to attract users to conversations with the robot. The developed conversation mechanism is a natural extension of chatbot. Priorities for possible answers in response to user inputs are determined according to the dialogue strategy. This scheme is used for searching items from a news article database. With this method, one does not have to define an explicit domain model. This system can use pre-fixed databases but also renewable ones. We demonstrate the proposed idea on the humanoid robot, ASIMO.

Index Terms—multimodal conversation system, humanoid robot, symbiosis, information service robot

I. INTRODUCTION

Symbiosis of robot and human is a current topic in the study of human-robot communication [1]. For a communication robot, we may feel the need to allow communication with it as if it were a partner, and have it provide us with a variety of information. We think that conversations in daily life need fun, surprise and unpredictable elements. Only obeying the user's intention is not enough for a real symbiotic robot to act as if it were a partner, even though it is a difficult problem to figure out a user's intention. We do not focus on

the problem of recognizing a user's intention, in this paper. We present a conversational robotic system, what we call a chat information service robot that is useful for users, while simultaneously engaging in fun conversations (figure 1).

The proposed system provides information to a user during conversation with the robot. This information provision can be viewed as a kind of Internet search engines using natural language. The main difference from the typical Internet search engines is that the purpose of information delivery is to promote the evolution of the conversation. A person does not always want to search items from the Internet in daily life. However, if a person wants to know the more detail about a specific topic during a conversation with a robot, the robot provides the information related to the topic. The proposed conversation scheme brings fun to users through chatting with the robot, with the robot providing new and interesting information that was not predictable before beginning the conversations. As a first step, we limited Internet information to news obtained from a specific site. This conversation system is achieved by extending conversation methods used in chatbot.

We implemented and demonstrated the system using the Honda humanoid robot, ASIMO [2]. This paper describes the detailed mechanism of the chat-information-service system.

II. PREVIOUS WORK

While there have been several studies of computer interfaces that can engage in conversation with humans, most of this work uses "conversation" as the interface to perform requested tasks. For example, Sato et al. proposed the "word-spotting association strategy" as a dialogue strategy for interactive robots [3]. Their proposal on dialogue strategy seems to be related to the mechanism of chatbot. They built a "ubiquitous home" to demonstrate ubiquitous computing services in a house. In their project called the "Yukari Project," small robots worked as user-interfaces against their "ubiquitous home." Mori et al. created the combination mechanism of chatbot and their scripted presentation system called MPML [4], [5]. They achieved interactive presentations by means of character agents. Although both the Yukari project and MPML systems perform interactive interfaces, they both

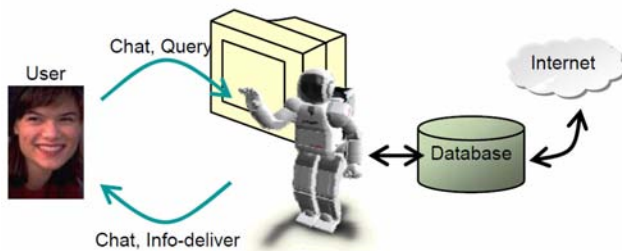


Fig. 1. Schematic image of the chat information service. The information delivery will be done by means of not only voice but also displays in the room. For the first step, a particular news site is used to retrieve news articles.

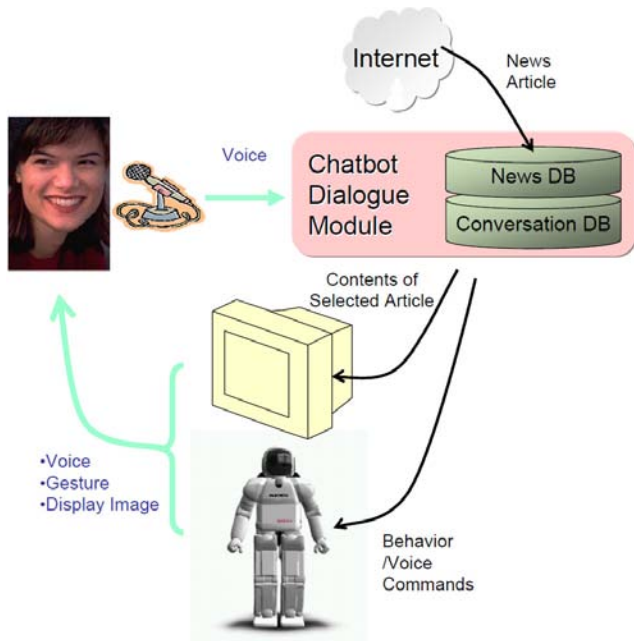


Fig. 2. Schematic diagram of the semantic system design.

have the constraint that the database or the rule has to be prescribed. Another limitation of both systems is they need to implement mechanism to select proper domains to distinguish the user's utterance (or input sentence).

There are also studies that are oriented to symbiotic robots [6], [7]. We share the concept of the necessity of an entertainment factor for humanoids. The referenced authors seem to emphasize behaviors of robots rather than conversations and their conversations are limited to what was described to them beforehand.

The other major studies on conversational robots concern to what we call task-oriented dialogue [8]. These approaches are intended to accurately recognize human intentions and providing information to the human. We think that this task oriented dialog is important to create a useful conversational robot. It is, however, unclear whether naive users want to spend long periods of time to get information through real robots in daily life.

III. CHAT INFORMATION SERVICE SYSTEM

In order to achieve information service through spoken dialogue, it is generally an important issue to select proper conversational domains in accordance with the user's intentions. Many systems providing conversational information service in multiple domains implement mechanisms to select the proper domain automatically. If the system only treats the direct queries from users, the problem becomes relatively tractable. As long as the systems are applied to voice web services, or telephone services, the approach limiting the user's queries works well. The symbiotic robot living with

humans should be able to treat not only explicit queries but also implicit user's intention because direct queries using spoken dialogue are sometimes cumbersome for people. It is not easy for the system to solve the problem of user's intentions. Conventional approaches use describing domain ontology. For topical news that may include topics from politics to show businesses, defining domain models working in conversational services is too hard. Moreover, since the contents of the domain are varying day by day, it is almost impossible to determine the model. We think that the difficulty comes from making conversational systems only to obey the user's intention.

In order to realize our goal of a chat information service robot, we utilized the so-called chatbot [9]. The dialogue strategy of chatbot is intrinsically scalable. We extended the dialogue strategy of chatbot to be able to treat not only chatting, described in advance, but also news topics. The chatbot system processes only superficial word orders of the user's utterance so that we do not have to describe the domain models explicitly.

The conversational system using chatbot strategy does not have an explicit mechanism to distinguish the user's intentions. In human conversation, giving a person new information is equivalent to responding to their intention. The conversation system of information service robots should have the both capabilities. We think that the task-oriented dialogue system [10] is also important. We are planning to combine and integrate both approaches in future work.

At home, the ability to entertain users is important for conversational humanoid robots. Even in the case of service robots, chatting with persons and informing them on related topics during conversation are important capabilities to carrying out task-oriented dialogue for humanoids. We think that the humanoid robots should serve in a human-like way. By behaving as a human, the humanoid robot can compete with other media such as CG character on PC.

Figure 2 shows the diagram of the system. When a user says something, an input device such as a microphone receives the information. Chatbot-based dialogue module chooses a proper conversation from its databases in accordance with its policy. ASIMO behaves according to the commands sent from this module. When a news item is selected, the display is used to show the article. Details are described in the following subsections.

A. Chatbot

The main component of chatbot is a database that contains questions asked by the users and the corresponding answers. The current conversation system of the chat information service is based on that of chatbot. In our system, the database is written in XML style. Figure 3 shows an example of chatbot database script. The basic design of chatbot is specified as voice communications using the current speech

```

<?xml version="1.0"?>
<category>
  <talk>
    <name text="[Name]">
      <planetext>
        What is your name
      </planetext>
      <answer priority="50"
        feeling="neutral"
        motion="congratuate"
        event="none" ans="none">
        I am ASIMO.
      </answer>
      <keyword text="[You]" />
      <keyword text="name" />
    </name>
  </talk>
</category>

```

Fig. 3. Example of the current version of chatbot database script. The actual one is written in Japanese.

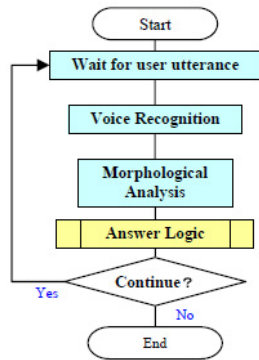


Fig. 4. Main loop of conversation system.

recognizers that sometimes fail to catch utterances precisely. <keyword> tags define the template of input text that describes the word order pattern of mainly nouns and verbs in input text. The text element of <answer> tag defines the output text corresponding to the matched word order pattern. <answer> tag has several attributes: feeling defines emotions, motion defines motion of ASIMO, and priority defines the priority of the output. The details of this priority value are described in the design part. Additionally, we also made a special editor of this script to ease the construction of the database.

B. Design of the Conversation System

In this section we explain the design of the conversation system. Figures 4, 5 and 6 are a flow chart of the answer component.

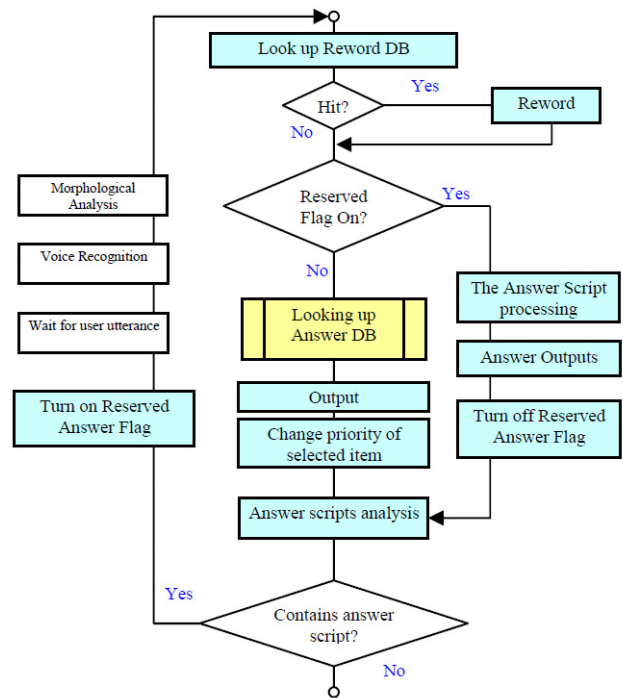


Fig. 5. The part of Answer Logic

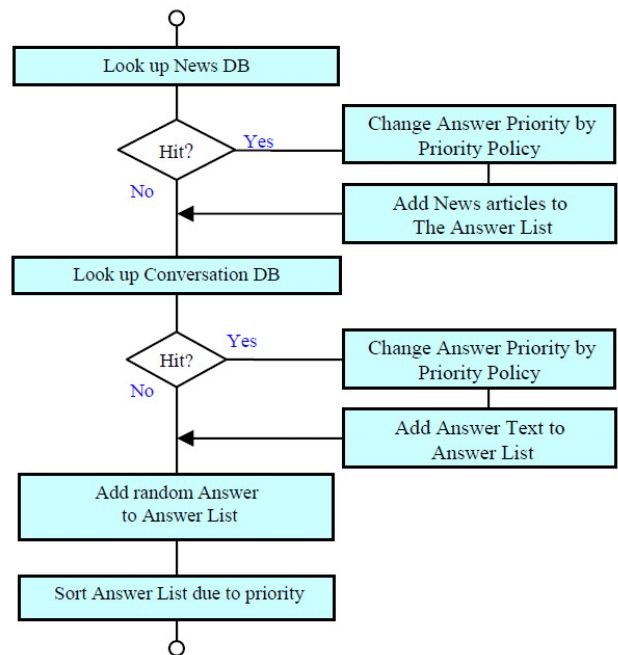


Fig. 6. The part of Looking up Answer DB.

Example: "You dislike what, ASIMO?" Keyword template: [you] dislike
Candidates List
News Article : priority=54 article="The method to cope with a child who dislikes vegetables."
Conversation : priority=80 answer="I terribly hate thunder."
Random : priority=40 answer="What did you say?"
Selection
Conversation : answer="I terribly hate thunder."

Fig. 7. Example of selection of output.

First, this system performs voice recognition of the user's utterance. "Julius", a Japanese dictation kit made by Kyoto University [11], is used for voice recognition. Second, a morphological analysis performed on the inputted string. "ChaSen" [12], which is a Japanese morphological analyzer, is used for this analysis. The next operation is a synonym change. In this section, the reword database is accessed, and the input string is changed into the standardized word.

Next, this program retrieves the answer to the extracted part of speech. This retrieval is done for two or more databases such as the conversation database and the knowledge database. In the current implementation, the conversation database and the news article database are used. As for the retrieved answers, chatbot has as policy the value of easiness as priority in selection, and the policy of chatbot controls the priority. Chatbot permutes the answers that are hits to the retrieval by using the priority, and selects one answer from the high-ranking ones by this priority (figure 7). Finally, chatbot outputs the conversation response and scripts of the humanoid behavior.

The policy of chatbot updates the priority correction value. For instance, the same answer is prevented from being repeated. Moreover, by analyzing the conversation reservation script included in the answer data, the flag of conversation reservation to the following input of the user is decided. By performing the conversation reservation, chatbot is able to speak in a conversation that relates to the previous content, and can produce a more natural conversation. When the conversation reservation has been generated, the usual retrieval is not done, alternatively, the conversation response according to the script is done.

C. Multimodal Expression

We make the multi-modal conversation system not only a voice conversation but also one that contains gestures and expressions. Figure 8 is an example of the operations and feelings expression of the ASIMO.

We use a speech synthesizer that is a commercially available one, FineVoice by NTT-IT. This speech synthesizer allows changing parameters such as pitch, speed and pitch range through its APIs. Varying these parameters, we try

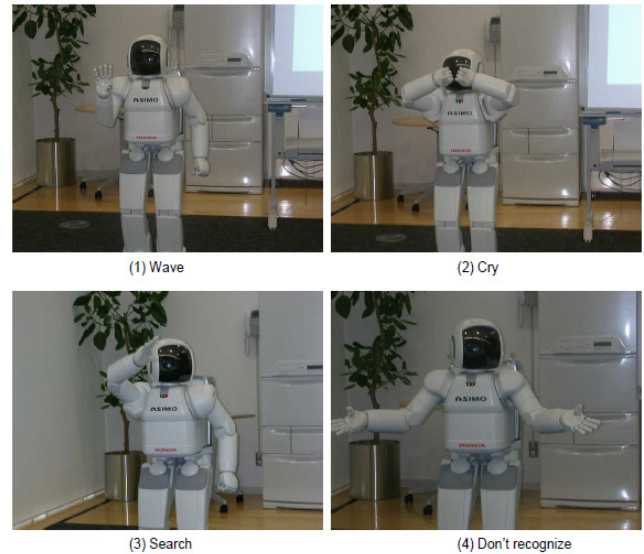


Fig. 8. Gesture examples of ASIMO.

to express emotions of the robot, although this cannot be accomplished perfectly.

If the robot reads out a long news article, the user will get bored. For this, the system also utilizes a display in the room to show news articles to users.

To realize these multimodal expressions, we ported MPML system (mentioned in section II) to ASIMO, which will be described in another article [13]. Humanoid robots are suitable devices for doing the multimodal-communication in the real world. Some authors also argue in [14] that the multimodal presentations that use ASIMO are more attractive than character agents on PC. We think that presence and behavior in the real world are important factors in continuous interaction between machines and humans.

D. News Retrieval

We installed a news retrieval function in the system. News articles are retrieved based on the words that the user speaks in daily conversation, and chatbot gives verbal answers to the user. The priorities of news articles are determined by the degree of text matching and the date of the articles. The nouns in the user's speech are collected. A tentative score for an article is determined by the degree of matching with the collected noun words. There is another kind of tentative scores defined by the most recent dates of the articles. The multiplication of the both scores yields the priority of an article.

We made programs for retrieving original news articles by using Microsoft.Net Framework and commercial XML native database software. The XML database server used here has enough specs even in enterprise usage. The gathered news articles are translated to XML data in compliance with RSS and stored in the XML database server.

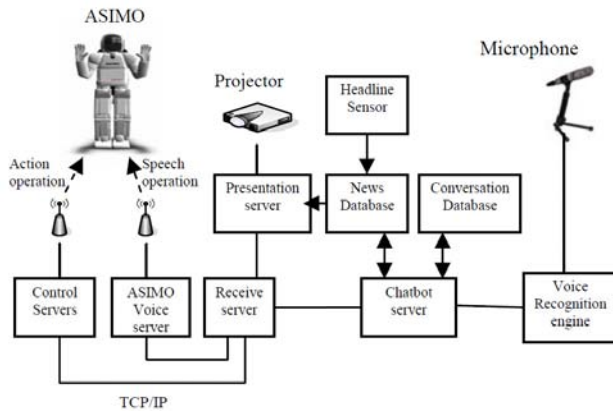


Fig. 9. System configuration.

E. Summary of Design

In our concept, topical news is used for prompting conversations. Whether users intentionally desire news topics or not, this system provides the related news article when the priority value of the news article is the highest. In addition, if we want to treat a user's direct query for news article, we just add the corresponding script to the chatbot conversation database. This system has the scalability through the separation of implementation and data.

IV. IMPLEMENTATIONS

A. Humanoid Robot

Several humanoid robots have been developed. In this research, we used the biped humanoid robot "ASIMO." It stands 1200 mm tall and has mass of 52 kg. ASIMO has height similar to a human child. Among biped pedaled humanoids, this size is relatively big so that the impression on users makes conversation easier.

B. System Configuration

Figure 9 shows the configuration of our prototype system. It consists mainly of the following components:

- ASIMO,
- Chatbot Server (Windows PC),
- Control Servers (Linux PC),
- Receive Server (Windows PC),
- ASIMO Voice Server (Windows PC),
- Conversation Databases (Windows PC),
- News Database (Windows PC),
- Headline Sensor (Windows PC),
- Voice Recognition Engine (Windows PC),
- Presentation server (Windows PC), and
- LCD projector, speaker and the large vertical screen.

The Voice Recognition Engine converts a human's voice into words and sends the words to Chatbot Server. Chatbot Server sends commands to the Receive Server. The Conversation

Databases and the News Database are described in XML style.

Headline Sensor collects news from the Internet and accumulates it in the News Database. The Control Servers control the humanoid robot ASIMO via wireless connection. The Receive Server receives commands from the Chatbot Server. This Server sends the ASIMO commands to the Control Servers, and sends the speak commands to the ASIMO Voice Server. The Presentation Server produces the presentation screen, which is an image with a news article. Then, this image is projected on the big wall screen with a projector.

The Chatbot Server, the Receive Server, the Control Servers, the ASIMO Voice Server, the Voice recognition engine, and the Presentation Server are mutually connected via TCP/IP network.

The Receiver Server controls the synchronization of ASIMO movements, speech, and presentations with the projector so that ASIMO can speak and use gestures. In addition, a simple head nodding mechanism during speaking is also implemented so that ASIMO appears more life-like.

V. RESULTS AND DISCUSSION

Figure 10 shows an example of chat news service dialogue. This system can handle a simple conversation, and always retrieves the news by word matching using words from the conversation of the user. When the news article is retrieved, this system offers the user the topic from the news. If the user wishes, the system offers a detailed article about the offered topic using a wall display.

We described a conversation system for humanoids to achieve fun conversations. Operations including searching for items through spoken dialogue with machines are sometimes cumbersome when compared with other conventional media such as keyboard, mouse or touch panels. We believe that this chatting ability is the basis of a conversational service robot. The chatbot system also can guide users to other services of the robots. Starting with chatting with the robot, users will use other services provided by the robot. Providing topical information is necessary, because only chatting with no information is also boring. For symbiotic humanoids, spoken dialogue will be the main interface to operate the robot. To be a symbiotic partner of the user the main interface should have an entertainment factor.

For future work, we intend to increase the conversational ability and improve the chatbot's policy so as to produce a more natural conversation. For future enhancing, we want to develop a universal interface so that all computers and men can communicate by using the humanoid chat-information-service system for ubiquitous computing services.

VI. CONCLUSION

In this work, we describe a chat-information-service system using a humanoid robot. We installed news retrieval on the system as one example of the application of a chat

Atsushi:はじめまして
(Hi!)

ASIMO: **[with bowing]**
 どうもはじめまして。僕はASIMO。よろしくね。
 (I am ASIMO. thank you.)

Atsushi:私はアツシ、よろしく。
(I am Atsushi.)

ASIMO: **[with a nod]**
 うん。
 (Yes.)

Atsushi:ロボットと会話できるなんて夢のある話だね。
(It is a dream to be able to talk with a robot.)

ASIMO: **[with opening hands]**
 僕は人とロボットがいっしょに生活することが夢なんだ。
 (For me it is a dream that I live with humans.)

ASIMO:ところで君の夢は？
(By the way, what is your dream?)

Atsushi:イチローみたいな、野球選手になりたいよ。
(I want to become a baseball player like Ichiro.)

ASIMO:イチローといえばこんなニュースがあるよ。
(There is news is for Ichiro.)
 「イチロー17試合連続安打」
 ("Ichiro hit in 17 consecutive games")

ASIMO:もっと詳しく知りたい？
(Do you want to know in greater detail?)

Atsushi:うん。もっと知りたい。
(Yes. I want to know more.)

ASIMO: **[With displaying the news article on wall screen]**
 米大リーグは27日、各地でオープン戦が行われ、マリナーズのイチロー外野手はホワイトソックス戦で5打数2安打3打点。初戦のパドレス戦からの連続試合安打を17に伸ばした。(As for Major League Baseball, an exhibition game was held on various places on the 27th, and outfielder Ichiro of the Mariners had 3 RBIs of 5 times at bat. Two hits at the White Sox game. His continuous hit game was lengthened to 17 in the Padres game.)

Atsushi:ASIMOってなんでも知ってるんだね。
(You know anything.)

ASIMO: **[Only gesture displaying doubtful]...**

Atsushi:ところでASIMOって本当にわかっているの？
(By the way, Are you really understood?)

ASIMO:そうかもね。
(Is it so?)

Fig. 10. Example of dialogue in Japanese.

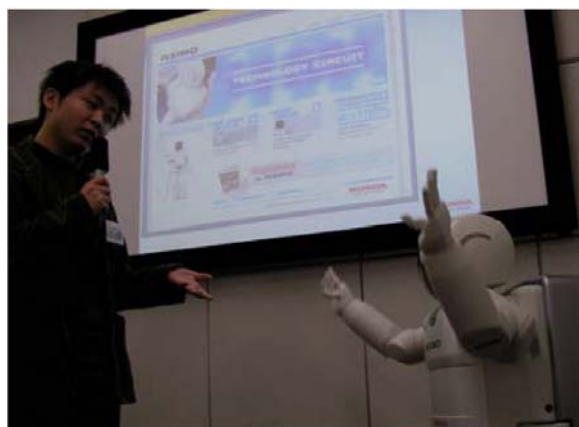


Fig. 11. Dialogue with the demonstration system.

information service. To demonstrate the system we made a preliminary implementation of the system on the humanoid robot, ASIMO.

ACKNOWLEDGMENT

We thank Dr. Mikio Nakano of Honda Research Institute Japan and Prof. Hiroshi Okuno of Kyoto University for helpful discussions.

REFERENCES

- [1] N. Hagita, "Symbiosis with Communication Robots," In 6th EAJ International Symposium, pp. 44-51, October, 2004.
- [2] Y. Sakagami, R. Watanabe, C. Aoyama, S. Matsunaga, N. Higaki, K. Fujimura, "The intelligent ASIMO: system overview and integration," In Proc. IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS-2002), pp. 2478-2483, 2002.
- [3] J. Sato, H. Ueda, M. Kidode, "A Service Description Method and a Dialogue Strategy for Interactive Robot" (in Japanese), Human Interface 6(3), pp. 69-72, 2004.
- [4] Multimodal Presentation Markup Language (MPML) Resources. Available at <http://www.miv.t.u-tokyo.ac.jp/MPML/en/>
- [5] K. Mori, A. Jatowt, M. Ishizuka, "Enhancing Conversational Flexibility in Multimodal Interaction with Embodied Life-Like Agents," In Proc. International Conference on Intelligent User Interface (IUI-2003), pp. 270-272, 2003.
- [6] T. Kanda, H. Ishiguro, M. Imai, T. Ono, K. Mase, "A constructive approach for developing interactive humanoid robots," In Proc. IROS-2002, pp.1265-1270, 2002.
- [7] Y. Hoshino, T. Takagi, U. Diprofo, M. Fujita, "Behavior description and control using behavior module for personal robot," In Proc. ICRA-2003, pp.4165-4171, 2003.
- [8] H. Asoh, Y. Motomura, F. Asano, I. Hara, S. Hayamizu, K. Itoh, T. Kurita, T. Matsui, N. Vlassis, R. Bunschoten, B. Kroese, "Jijo-2: An office robot that communicates and learns," IEEE Intelligent Systems 16(5), pp.46-55, 2001.
- [9] J. Weizenbaum, "ELIZA—A Computer Program For the Study of Natural Language Communication Between Man and Machine", Commun. ACM 10(1), pp. 36-45(1966).
- [10] M. Nakano, Y. Hasegawa, K. Nakadai, T. Nakamura, J. Takeuchi, T. Torii, H. Tsujino, N. Kanda, H.G. Okuno, "A two-layer model for behavior and dialogue planning in conversational service robots," to be appear in Proc. IROS-2005.
- [11] A. Lee, T. Kawahara, and K. Shikano, gJulius — an open source realtime large vocabulary recognition engine, in Proc. 7th Eurospeech, pp.1691-1694, 2001.
- [12] Y. Matsumoto "Morphological Analysis System ChaSen" (in Japanese), Info. Proc. Soc. of Jpn. Mag. 41(11), pp.1208-1214, 2000.
- [13] K. Kushida, Y. Nishimura, H. Dohi, M. Ishizuka, J. Takeuchi, H. Tsujino "Humanoid Robot Presentation through MultimodalPresentation Markup Language MPML-HR," accepted by 4th International Joint Conference on Autonomous Agents and Multi Agent System (AAMAS-2005).
- [14] Y. Nishimura, K. Kushida, Y. Nishimura, H. Dohi, M. Ishizuka, J. Takeuchi, H. Tsujino "Development and psychological evaluation of Multimodal Presentation Markup Language for Humanoid Robots," in submission to IEEE-RAS International Conference on Humanoid Robots (Humanoids2005).