Chatbot and Dialogue Demonstration with a Humanoid Robot in the Lecture Class

Shu Matsuura(✉) and Riki Ishimura

Faculty of Education, Tokyo Gakugei University,
4-1-1 Nukuikita, Koganei, Tokyo 184-8501, Japan
shumats0@gmail.com

Abstract. This paper describes ideas to extend the way of verbal lecture visually and verbally. Firstly, we conducted questionnaires to the first-year university students of the science education. As for the attention to the verbal and visual information, we noticed participants' worry about keeping following with the verbal lecturing. We also noticed the participants' preference for the visual presentation. These imply a potential necessity of visual presentation that supplements the lecturer's speech in a real time. On the other hand, the results showed a preference for the learning under familiar atmosphere. For the visualization of lecture talk, we created a chatbot on the platform of Api.ai to exhibit a brief explanation or a question of the words as the lecturer utters them. This text visualization provides the basis of the lecture in sync of lecturer's talk. We also used a humanoid robot for verbal presentation as a partner of the lecturer. The Topic Maps was introduced to build the dialogues for analogy and association. The analogy topic map was constructed based on a Japanese entertainment riddle. Robot dialogues were conducted by connecting an online learning system based on Topic Maps.

Keywords: Chatbot · Topic maps · Humanoid robot NAO · Online learning system

1 Introduction

For lecturers of higher education, the lecturing is a performance and interaction with the participants. Such a performance is an interface of knowledge and information for the participants. Also, it is an opportunity to think by themselves or with others on the related issues. The role of the lecturer is not simply transmitting the knowledge anymore as the information is becoming more and more ubiquitous for the people of any generation. Rather, the lecture has to inspire the participants to become positive in thinking, feeling, acting, and establishing the knowledge.

Skills of lecturing as a performing art have been studied elaborately [1]. Recently, more and more new technology has been applied to make the lecture interactive and experience rich based on the learning theories. Applications of the artificial intelligence and communication robot are two hot topics to explore the future learning. However, we should be aware of the demand of actual participants to the lecture room, and at the same time, how the technology can be unified with the lecturer's specific art of performing.

M. Antona and C. Stephanidis (Eds.): UAHCI 2017, Part III, LNCS 10279, pp. 233–246, 2017.
DOI: 10.1007/978-3-319-58700-4_20

In this study, we conduct a questionnaire survey to clarify the classroom participants' tendencies. Then, we attempt to introduce a text visualization that supplements the lecture's talk on sync. Further, we use a humanoid robot to introduce a dialogue that stimulates participants' intuition within the lecture's logical talk. In this study, we describe these attempts and the results of participants' attitude survey.

2 Method

To survey the tendency of students' preference for the visual and verbal presentation, questionnaires with five ranks Likert scale were made and conducted to the 1st-grade university students in a science education classroom. The total number of students participated was 53. The questions were on the logical tendency and the preference for expression and learning. The survey was carried out after a brief introductory practice of the following terminology visualization and robot talk.

The terminology visualization we developed is a chatbot using the api.ai conversational user experience platform [2]. The terminologies used in the lecture are registered for detection through voice recognition of a PC. For the response words of the chatbot, we registered the short texts of the meaning, definitions, comments, or simple questions on the terms. The chatbot service was developed and implemented on the api.ai website. At the classroom, opening the chatbot page, the lecturer turns on the voice recognition when to pronounce the registered words. Then the chatbot replies the texts to explain the term. The texts shown in the chatbot page are displayed in the LCD hanged from the roof.

For verbal communications, we used a humanoid robot NAO [3]. NAO has vision and voice recognition and can talk with the human. NAO's emotional expression has also been applied to the interactive teaching [4]. NAO can transmit URL requests through Wi-Fi and receives XML or JSON data.

One of the authors has developed an online learning system, "Everyday Physics on Web (EPW) [5]," based on Topic Maps, a standard of indexing technology (ISO/IEC 13250:2003). This website was meant for the support of introductory science and science education lectures.

Topic Maps constitute of "*topic*," "*association*," and "*occurrence*," where a *topic* represents an entity, an *association* links two topics, and *occurrence* links a *topic* to the actual web materials on the *topic*. Various types of *associations* are defined so that the *topics* and the *topic* instances are structured using various types of *associations*.

EPW has been developed on the basis of the Ontopia development and runtime environment for Topic Maps [6, 7]. Ontopia Topic Maps remote access protocol (TMRAP) enabled URL requests by the Topic Maps query language. EPW ontology and the way to access EPW from NAO have already been described briefly in [8].

3 Result and Discussion

3.1 Questionnaire

Table 1 shows the list of the questions together with their shortened index phrases and the average values of the answers obtained. The index phrases are used to retrieve the questions from the figures in this section. Error values indicate the corrected sample standard deviations. These questions were delivered after a brief trial of the text visualization and the robot dialogue with an instructor. The list was arranged in the descending order of the mean value.

Table 1. Questions on learning and average values.

No.	Question	Shortened index of question	(average) ± (sample standard deviation)
Q1	Do you understand things better, when you feel familiar with them?	Familiarity Brings Understanding	4.3 ± 0.7
Q2	Do you obtain information mainly by vision?	Obtain Information By Vision	4.0 ± 0.9
Q3	Do you think the structured knowledge is most important to comprehend things?	Structured Knowledge Important	3.9(8) ± 0.9
Q4	Do you memorize others' words well?	Memorize Others' Words	3.8 ± 1.0
Q5	Do you understand things better, when you are smiling?	Understanding While Smiling	3.8 ± 0.9
Q6	Do you understand things better, when your feel relaxed?	Understanding While Relaxing	3.7 ± 1.1
Q7	Do you think you mainly use words to think?	Think With Words	3.6(9) ± 1.2
Q8	Do you like to be provided of inspiring expressions on the things you need to comprehend or attain?	Inspiring Information	3.6(6) ± 1.0
Q9	Do you think a sense of unity between you and the things is most important before comprehending it with words?	Sense Of Unity	3.6(0) ± 1.1
Q10	Do you prefer rational understanding?	Rational Understanding	3.5(6) ± 1.2
Q11	Are you interested in the dialogue between two or more person rather than the monologue in speech?	Prefer Dialogue	3.5(2) ± 1.0
Q12	Do you tend to understand things intuitively?	Intuitive Understanding	3.4(5) ± 1.1

(*continued*)

Table 1. (*continued*)

No.	Question	Shortened index of question	(average) ± (sample standard deviation)
Q13	Do you feel the inspiration is most important to comprehend things?	Inspiration To Comprehend	3.3 ± 1.3
Q14	Do you remember a monologue better than dialogue?	Remember Monologue Better Than Dialogue	3.2(2) ± 1.2
Q15	Does the speech of one person rather than the dialogue between two or more persons attract you?	Speech Attracts Better Than Dialogue	3.2(0) ± 1.1
Q16	Can you keep concentration on lecturer's long talks?	Keep Concentration On Talk	3.0 ± 1.3
Q17	Do you understand things better, when you feel tense?	Understanding While Feeling Tense	2.9 ± 1.2
Q18	Do visual images added to the text distract you?	Image Distracts	2.5 ± 1.2

We found five underlying tendencies from Table 1 as described below.

Tendency 1: Q1, 5, 6, and 17 show that the participants feel they can understand things better under pleasant and relaxed state than when they feel tense.

Tendency 2: Q2, 4, and 18 show that the participants memorize things well from the verbal information, while many of them obtain information from visual symbols. However, the result of Q16 shows that many of them feel difficulty to keep their concentration on the instructor's long talks.

Tendency 3: Q3, 7, and 10 show that many of the participants infer rationally based on the established knowledge. This rationality might reflect that the participants belong to the science education division of the faculty of education, and most of them have strength in the subjects of sciences. They can make use of the established scientific concepts to consider the real problems.

Tendency 4: Q8, 9, 12, and 13 show that some of the participants have a tendency of comprehending intuitively. They might feel that inspiration is important.

Tendency 5: Q11, 14, and 15 indicate that a part of participants prefers a straightforward explanation in a monologue style. Also, a part of participants feels interested in the dialogue better than monologue.

Figure 1 shows a histogram of participants' answers concerning the tendency 1. 32% of them agreed or strongly agreed that they understand well while feeling tense. As shown in Table 2, the correlation coefficient ρ between "Understanding While Feeling Tense" and "Understanding While Relaxing" indicates that they are negatively correlated. Thus, as a learning condition, some students may prefer relaxing atmosphere while others prefer tense. Furthermore, 85% of participants agreed or strongly agreed that they have a better understanding when they feel familiarity on the subject.

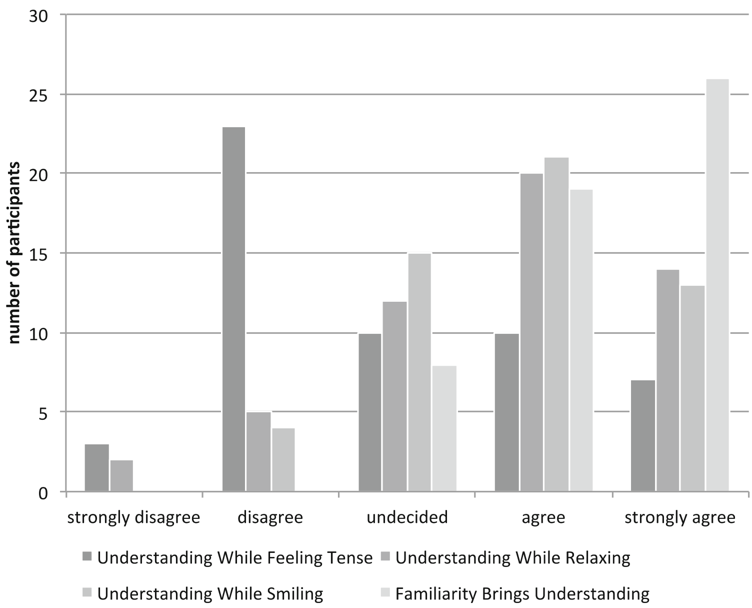

Fig. 1. Histogram of answers for tendency 1

Table 2. Correlation coefficients ρ for tendency 1

Indices of questions examined	ρ	p-value
"Understanding While Feeling Tense" vs. "Understanding While Relaxing"	-0.49	2.0×10^{-4}
"Familiarity Brings Understanding" vs. "Understanding While Relaxing"	0.51	1.0×10^{-4}

"Familiarity Brings Understanding" is also moderately correlated with "Understanding While Smiling" with the correlation coefficient of 0.51. This correlation implies that at least for a part of students, a moderate humor possibly makes their mind open for thinking and learning the subject.

Figure 2 shows the histogram of the answers on visual and verbal information gain indicated as tendency 2. 75% and 70% of participants agreed or strongly agreed that they obtain and memorize information visually and verbally, respectively. The answers for "Obtain Information By Vision" and "Memorize Others' Word" showed no correlation, with the coefficient of 0.09. However, to understand a lecturer's talk, one has to keep attention to the talk, thinking on the meaning of what they hear synchronously. 38% of participants disagreed or strongly disagreed with "Keep Concentration On Talk." This anxiety in keeping attention suggests participants' needs on the supplemental visual presentation. In many cases, the lecturer uses the blackboard or the presentation tools for summarizing or enrich the verbal explanation.

Figure 3 shows the histogram on the tendency 3 concerning the participants' understanding. 74% of participants agreed or strongly agreed on the importance of the

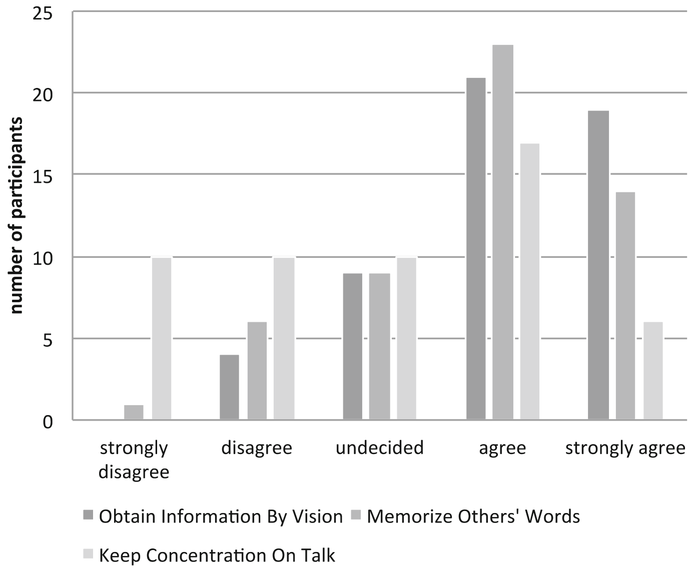

Fig. 2. Histogram of answers for tendency 2

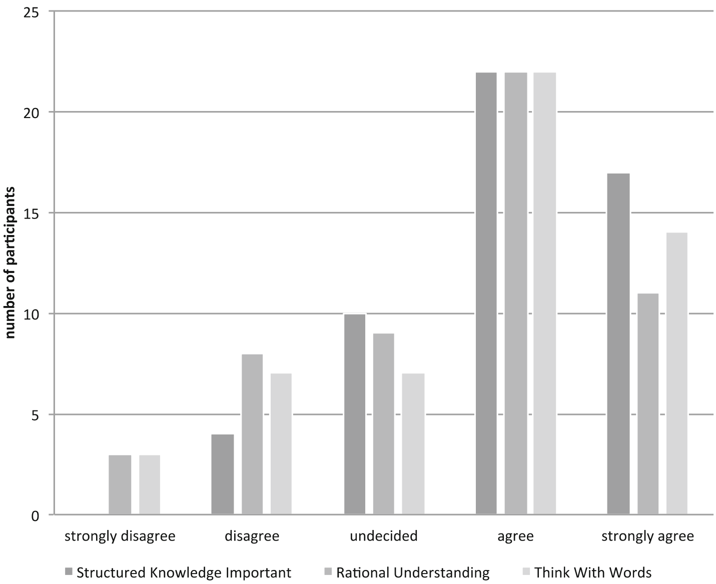

Fig. 3. Histogram of answers for tendency 3

Table 3. Correlation coefficients ρ for tendency 3

Indices of questions examined	ρ	p-value
"Rational Understanding" vs. "Think With Words"	0.53	4.7×10^{-5}
"Rational Understanding" vs. "Structured Knowledge Important"	0.45	7.0×10^{-4}

structured knowledge as the base of learning. The science-oriented students are particularly trained to consider things based on the established knowledge on nature or mathematics. The answers on "Rational Understanding" was found moderately correlated with "Think With Words" with the correlation coefficient of 0.53, as well as with "Structured Knowledge Important" with that of 0.45 as shown in Table 3. This implies the particular importance of the participants' use of words on scientific concepts for both thinking and learning. Therefore, it may be meaningful to reinforce the students' knowledge structure by providing such as the definition of terminology or the knowledge that supports the concepts that the lecturer uses.

On the other hand, inspiration and intuition are also essential factors for the scientific inference. Particularly, the heuristic consideration or abductive reasoning in the scientific exploration cannot be without inspiration and intuition. Figure 4 for the tendency 4 shows the histogram on these functions of the mind. 47% of participants agreed or strongly agreed with the both of the questions "Inspiration To Comprehend" and "Intuitive Understanding." As shown in Table 4, the moderate correlations are

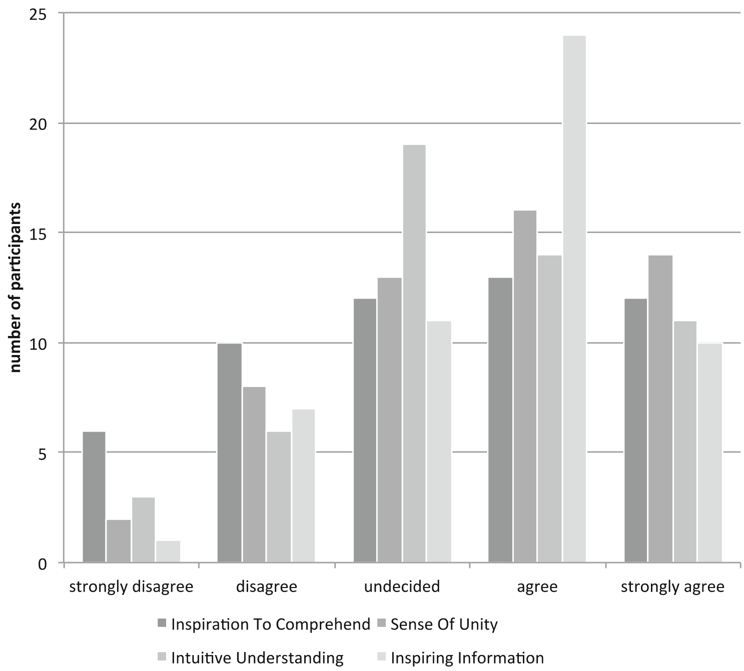

Fig. 4. Histogram of answers for tendency 4

Table 4. Correlation coefficients ρ for tendency 4

Indices of questions examined	ρ	p-value
"Inspiration To Comprehend" vs. "Sense Of Unity"	0.52	7.0×10^{-5}
"Sense Of Unity" vs. "Intuitive Understanding"	0.56	1.3×10^{-5}
"Intuitive Understanding" vs. "Inspiration To Comprehend"	0.63	5.1×10^{-7}

found among the results of the above items and "Sense Of Unity." On the other hand, 64% of participants agreed or strongly agreed with the expectation of "Inspiring Information." The last question showed no correlation with the former three questions. Thus, the students seem to require an inspiring component in the instruction, regardless of whether they have intuitive or rational tendency of reasoning.

Figure 5 shows the histogram on the tendency 5, which is about the participants' preference for the monologue and dialogue form of verbal presentations. In the traditional large classroom lecture on natural sciences, the lecturer makes a step-by-step explanation in the monologue form, using a blackboard. This style is efficient as an explanation of the theory construct, as far as the students can catch up with the talk. This will hold for 36% of the participants who agreed or strongly agreed with "Speech Attracts Better Than Dialogue," and will also be true for 45% of participants who agreed or strongly agreed with "Remember Monologue Better Than Dialogue."

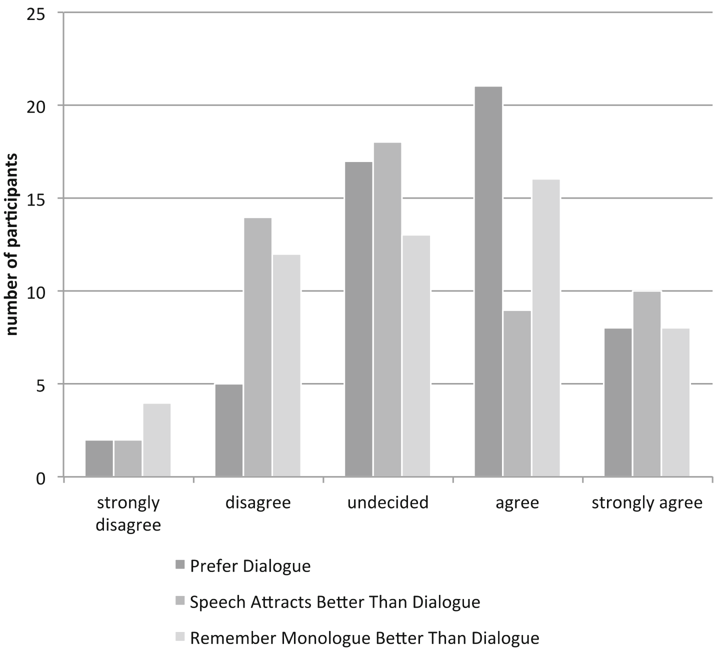

Fig. 5. Histogram of answers for tendency 5

Table 5. Correlation coefficients ρ for tendency 5

Indices of questions examined	ρ	p-value
"Remember Monologue Better Than Dialogue" vs. "Speech Attracts Better Than Dialogue"	0.53	4.7×10^{-15}
"Remember Monologue Better Than Dialogue" vs. "Prefer Dialogue"	−0.51	9.0×10^{-5}

However, on the other hand, 55% of participants agreed or strongly agreed with "Prefer Dialogue." This result indicates that the dialogue between two or more peoples stimulates the participants more than monologue by a single lecturer. Also, as seen in Table 5, "Prefer Dialogue" showed a negative correlation with "Remember Monologue Better Than Dialogue." Thus, in the classroom, a part of participants prefer dialogue form of presentation.

3.2 Visualization of Speech Contents

Terminology visualization chatbot was applied in three sessions of successive lectures. This lecture aimed at providing opportunities to the students to reconsider the roles and the issues of science education based on the human history and the tasks for the construction of future society. Therefore, sometimes the students need to review the meanings or definitions of the sociological terminology.

We provided brief explanations of a number of terms that the lecturer utters in the classroom, using the terminology visualization chatbot. The examples of the terms visualized are, "layer," "identity," "nation state," "1st, 2nd, 3rd, 4th industrial revolution," "capitalism," "frontier," "democracy," "big government," "small government," "neoliberalism," "state socialism, " "the world is flat," "fluidity of labor," etc. In case speaker's verbal input is not matched by any of the words defined, one of the phrases predetermined for simple response are replied, such as, "indeed," "well," "then?" etc. To display the terms, the lecturer started verbal recognition on PC just before speaking the term. The display of each terminology was not repeated many times. Rather the lecturer showed the visualization mostly once for each term in the introductory phase of the lecture.

Table 6 shows the summary of the participants' comments on the terminology visualization. These comments were collected at the end of second practice using the e-learning system. At this time, the participants were already so familiar with the display that some of them said they forgot to notice the display most of the lecture time. 48% of the participants commented positive effect of the display. The lecturer can display the terms at an appropriate timing according to his intent and can continue to talk recognizing that the participants notice the display.

However, 26% of participants feared about losing attention to the lecturer by noticing the display. Particularly, loss of concentration seemed to occur when one notice the miss-recognition of speech. Application of the machine learning of verbal recognition and adaptation to the lecture might improve the system.

Table 6. Summary of the comments on the terminology visualization

Content of comment	Percentage of participants [%]
By displaying explanations of the terms, it became possible to understand the content of the lecture smoothly	48
It will be of the assistance to students with hearing disabilities or inconvenience	26
The degree of concentration on the lecture will be reduced if we see the miss-display associated with miss-recognition of speech	26
Improvement of display or user interface is necessary	17
It may be particularly useful in certain situations, such as when conducting science experiments, practical training or large lecture meetings	13

13% of participants noted that this tool might help the students who have any inconvenience in the hearing. Normally, the lecturer prepares a visual presentation or the blackboard usage. Terminology visualization will supplement them effectively.

Besides, some participants pointed out the following problem. Although checking the uncertain terminology on time is quite smooth for following the lecture, it could be important that the learner review by him/herself the uncertain terms after the lecture. Some rearrangements might be necessary for the lecture design, to avoid making the learners passive on gaining knowledge.

3.3 Dialogue with Humanoid Robot

We made use of the humanoid robot NAO to provide verbal communication in the form of dialogue with the lecturer. In comparison with the terminology visualization, which supports the establishment of structured knowledge in the lecture, the dialogue with the humanoid robot rather contributes to building a humorous atmosphere. Also, we attempt to build a possible way to include intuitiveness or inspiration in the dialogue form. Particularly, we focus the analogy and the association in the following sections.

Analogy and *Nazokake*. The purpose of this dialogue is to use an analogy to make the scientific concept easier to understand. We connect scientific notions with daily life things that are imaginable for the learners of a wide range of the age.

Nazokake is a form of the riddle that is often done by the professional entertainer [9]. *Nazokake* connects two completely different entities with a common hidden feature. The homophony in the Japanese language is often used for this common feature. Also, conceptual similarity or equality is used as well. *Nazokake* is successful if the difference between the two entities is surprising and the finding of the hidden connection is amazing. Since the rule of *Nazokake* is clear, there are some attempts to generate *Nazokake* automatically using semantic relationships [10].

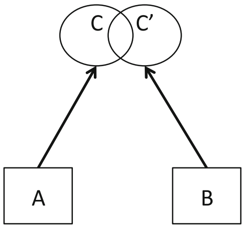

Fig. 6. Relationships of *Nazokake* elements

Figure 6 shows the relationships of *Nazokake* elements. A and B are two different entities. In the *Nazokake* play, the starting entity A is given as a question. After the player answers B to the question A, he is asked what is hidden behind them. Then he reveals C(C'), as the main part of the humor. In the figure, C is associated with A, and C' is associated with B. C and C' are homophonic words or they have the similar meanings. To answer *Nazokake*, one firstly finds several C's, and then finds the candidates of C' and B. Finally, the answerer selects the best combination of A and B.

Here, we consider the case that B is an analogy of A, and A is a scientific concept. Although A can be understood logically, learners might map the new notions to the already attained knowledge established in their minds to recognize the notion. That is, it is helpful to understand A by an analogy of B that is already familiar to the learner. To find the familiar topic B via our online learning system, we next apply the method of solving *Nazokake*.

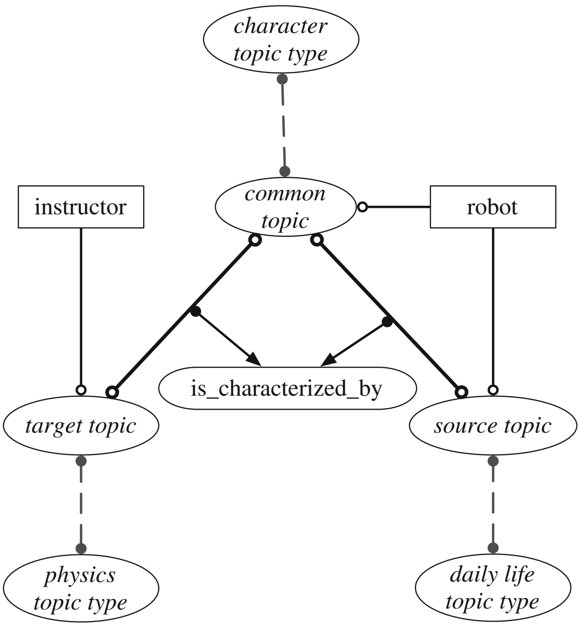

Fig. 7. Topic map ontology of analogy

Figure 7 shows a topic map ontology to find a *source topic* B from a *target topic* A. Let the *target topic* in question A is a *topic instance* of physics subject. Let A be “evaporation” for example. Assume the instructor intend to explain a temperature decrease caused by the evaporation. That is, evaporating molecules take away the kinetic energy from the liquid. The liquid molecules collide with each other transferring kinetic energy. As some molecules obtain enough kinetic energy to get out of the intermolecular force of liquid molecules, they evaporate, and their kinetic energy is taken away from the liquid. This situation may be named as “deprived of heat.” An analogous situation in the *daily life topic* is for example “a debut of a fellow.” A fellow in an amateur band group gets public fame. Then, he/she gets away to be professional, and, in turn, the band gets less active for a while. This amateur band *topic* B can be related with A in the context of “deprived of heat.” We register a *topic* of “deprive of heat” as C. Now the *topic* C (in general partially) characterizes both A and B. Both A - C and B - C are associated by is_characterized_by *association* as shown in the figure. Finding an appropriate C in agreement with the intent of the instructor, the system can retrieve B as an analogy of A.

Association. Association also refers to the connection of two different entities that share some common natures or similarities. Association is thought to contribute to the construction of inspiration-oriented dialogue. In this study, we consider “speaking of” pattern introduction of a *topic*.

Figure 8 shows the topic map for generating the following simple dialogue.
Instructor: This is the year of Rooster. Speaking of the Rooster Zodiac?
NAO: Speaking of the Rooster Zodiac, I recall the chicken.
Instructor: Speaking of the chicken?
NAO: Speaking of the chicken, I recall the egg.
Instructor: Can you tell me an experiment on the egg?
NAO: Spinning egg.

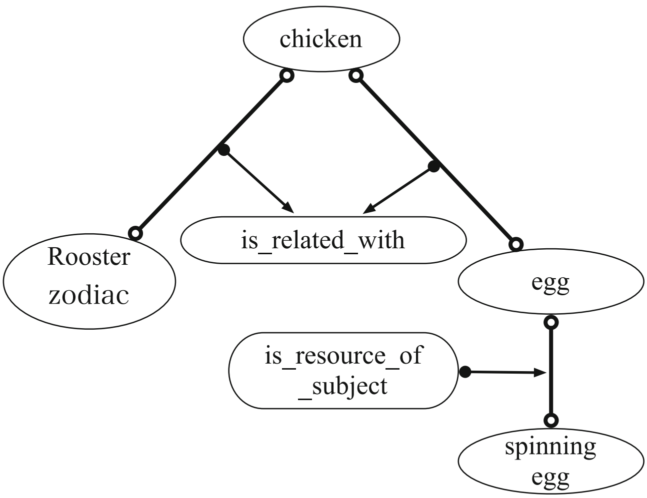

Fig. 8. Topic map for simple dialogue to find associated *topic*

The instructor can talk more in the intervals of these dialogue patterns. In the topic map, a neutral *association*, is_related_with, links "Rooster Zodiac," "chicken," and "egg" with each other. A specific *association*, is_resource_of_subject, links "egg" with a simple experiment demonstration "spinning egg."

The *topic* "chicken" may have relations with the subjects other than "egg" and "Rooster Zodiac." The robot reaches "egg" in the repetition of a fixed phrase "Speaking of?" Implementation of more targeted recalling is a future problem.

Participants' responses. Table 7 shows the questions on the humanoid robot demonstration and the average values of the results. Table 8 shows the values of the correlation coefficient between any two of the questions. Participants felt the humanoid robot and its dialogue enjoyable and showed a positive attitude towards the man-machine interactions. The answers were moderately correlated with each other, implying that participants' impressions on the communicative machine are consistent. In the participants' free descriptions, most of the impressions were about the friendly atmosphere brought by the robot and fun of its dialogue. In this sense, the dialogue with the robot is suitable for producing intuitiveness and entertainment.

Table 7. Questions on the humanoid robot demonstration and average values of answers.

Question	Shortened index of question	(average) ± (sample standard deviation)
Do you feel enjoyable if you have a communication robot close to you?	Enjoyable With Robot	3.6 ± 1.1
Can the communication robot be a partner of you?	Robot Partner	3.1 ± 1.3
Do you think more interaction will be required between man and machine?	More Man-Machine Interaction	3.4 ± 1.2

Table 8. Correlation coefficients ρ for answers on the robot.

Indices of questions examined	ρ	p-value
"Enjoyable With Robot" vs. "Robot Partner"	0.52	7.0×10^{-5}
"More Man Machine Interaction" vs. "Enjoyable With Robot"	0.56	1.3×10^{-5}
"More Man Machine Interaction" vs. "Robot Partner"	0.63	5.1×10^{-7}

4 Conclusion

A Likert scale survey question on the visual and verbal presentation and comprehension were conducted to students of university classroom to explore the tendency of students' preference. The results indicated that the students learn better under a relaxed situation rather than feeling tense. They obtain knowledge both from visual and verbal information. Although they can follow instructor's monologue, they feel difficulty to keep concentration on the long monologue, and a part of them feel interested in the more interactive dialogue style.

Based on the results of the survey, we developed a terminology visualization that provided the brief explanation of the terms using voice recognition and Api.ai chatbot. This chatbot supplements lecturer's verbal explanation with visual information in sync with the explanation. It was found from a survey that this terminology visualization was helpful to understand the lecture smoothly. However, loss of concentration occurs when miss-recognition and inappropriate display happen.

For dialogue presentation, we utilized the humanoid NAO that communicated with an online learning system constructed based on a topic map. Using dialogue form, analogy and association were implemented to create intuitive and relaxing part of the talk. From the result of the survey, it was found that the participants felt the dialogue between the humanoid robot and lecturer was enjoyable and brings about a friendly atmosphere in the lecture.

The visual and verbal presentations described in this study add new elements that are adaptable to the lecture. At the same time, it was pointed out that they could be distracting elements. The future task will be to refine them to become unified expressions with the lecturer.

Acknowledgement. This study was funded in part by a Grant-in-Aid for Scientific Research (C) 15K00912 from the Ministry of Education, Culture, Sports, Science and Technology, Japan.

References

1. Davis, J.: Dialogue, monologue and soliloquy in the large lecture class. Int. J. Teach. Learn. Higher Educ. **19**(2), 178–182 (2007)
2. api.ai. https://api.ai
3. SoftBank Robotics. https://www.ald.softbankrobotics.com/en
4. Anis, M., Elnaggar, A., Reichardt, D.: Exploring interactive teaching of a multi-modal emotional expression of a humanoid robot. In: 2016 Future Technologies Conference, pp. 908–915 (2016)
5. "Everyday Physics on Web," http://tm.u-gakugei.ac.jp/epw/
6. Matsuura, S.,: Development of a trans-field learning system based on multidimensional topic maps. In: Leipziger Beiträge zur Informatik: Band XIX, Linked Topic Maps, vol. 19, pp. 83–89. University of Leipzig (2009)
7. Ontopia. https://github.com/ontopia/ontopia
8. Matsuura, S., Naito, M.: Shaping dialogues with a humanoid robot based on an e-learning system. In: 2016 11th International Conference on Computer Science & Education (ICCSE), pp. 7–12 (2016)
9. Szirmai, M.: Raising multicultural awareness by teaching humour in foreign language classes. In: Alao, G., Plard, M.D., Suzuki, E., Roger, S.Y. (eds.) Didactique plurilingue et pluriculturelle: L'acteur en context modialisé, (Archives Contemporaires), pp. 193–202 (2012)
10. Uchimura, K., Nadamoto, A.: Automatic generation of NAZOKAKE by using web content. IPSJ SIG Technical Report, Vol. 2009-DBS-148, No. 25, pp. 1–6 (2009)