Pursuing a Dialogue System for Making Society More Harmonious

Ryuichiro Higashinaka
Senior Distinguished Researcher, NTT Media Intelligence Laboratories

Overview

The spread of smartphones and artificial intelligence (AI) systems that speak means that the opportunities for general users to come into contact with AI are increasing. Users are now expecting improved ease of use and more new technologies from research and development. Under these circumstances, NTT has continuously gained attention—both domestically and globally—in announcing pioneering technologies in the fields of question answering functions and language processing. We visited Ryuichiro Higashinaka, Senior Distinguished Researcher at NTT Media Intelligence Laboratories, and asked him about the progress and outlook of his research, his attitude as a researcher, and how the development of AI and a future in which robots and humans can talk smoothly will bring about changes in our society.

Keywords: artificial intelligence, natural language processing, question answering technology

Research on dialogue systems takes steady work

—Tell us about the research that you are currently working on and the initiatives undertaken so far.

I joined NTT in 2001, and since then, I have been researching language processing, artificial intelligence (AI), and dialogue systems. A dialogue system is a technology for interacting with a computer. It is probably easier to understand if imagined as an animated robot character like Doraemon or Astro Boy who can talk smoothly with humans. The idea of "smoothness," that is, "naturalness," is the point of my research, and for that reason it is necessary to unravel what kind of elements a person's conversation is made up of.

As shown by the fact that 60% of human conversation constitutes chat (casual conversation), chat is very important. A relationship between people cannot be established simply by having a conversation about work. Chat serves as cushioning material that makes it possible to know the other person's personality, thereby encouraging cooperative work. Let's consider that one person in a conversation is replaced with a dialogue system (i.e., a computer): if the computer does not understand what kind of person the person is, and if the person does not understand what kind of thing the computer is like, the conversation will not go smoothly. In my past research, though, I focused on tasks rather than chat, so I had the idea that a dialogue between a person and a computer should be brief. However, I came to realize the importance of communication and building relationships in achieving natural conversation, so I am currently focusing on chat. In regard to the study of an actual dialogue system, I'm repeating the straightforward work of
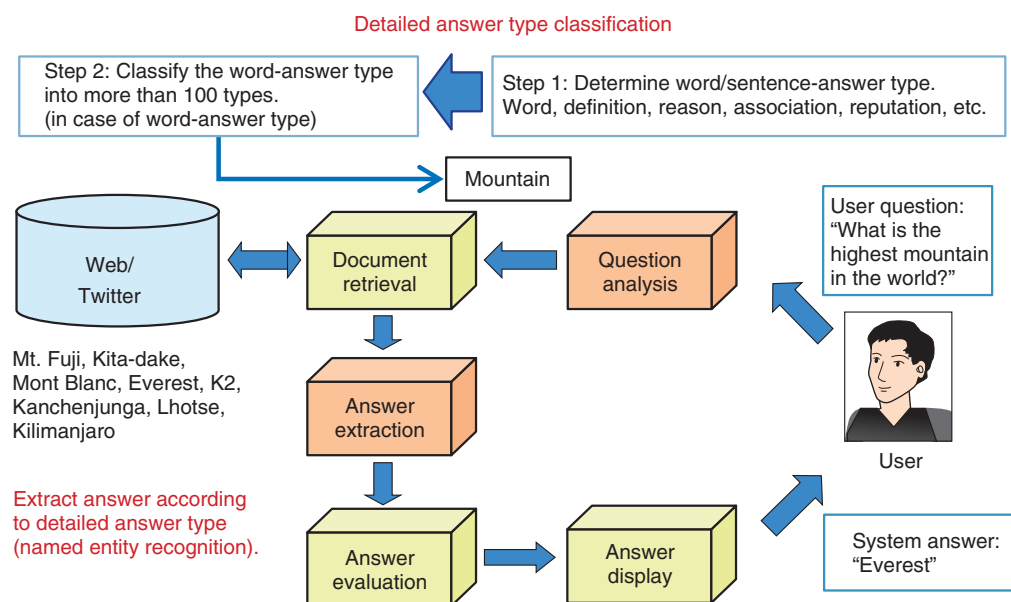
Fig. 1. Question answering logic used by "Shabette Concier."

observing conversations and building the system, getting people to talk with the system and checking unnatural parts of speech, and feeding the results back to the system. In that sense, research on dialogue systems is step-by-step.

—*Can you introduce the research that you have worked on specifically?*

There is NTT DOCOMO's voice-agent service called "Shabette Concier" (talking concierge), which was launched in March 2012 in Japan. I was in charge of the logic used by the question answering system supporting this service (**Fig. 1**). The question answering system was adopted as a centerpiece function when Shabette Concier was upgraded in June 2012. It recognizes and analyzes a user's words and searches for relevant information on the Internet and finds the best answers within a few seconds. It was designed to work smoothly even if many users are connected at the same time. It took about half a year for us to make the system practical, and as the practical application of question answering technology—which had not been developed on a large scale then—it received much attention from the research community. Incidentally, even though I was directly in charge of the final stage of development of the function, the basic research that formed its foundation spanned over 10 years before that stage. I don't think that we would

have been successful without that basic research.

As for AI, I was involved in the project called "Can a robot get into the University of Tokyo?" led by the National Institute of Informatics (**Figs. 2** and **3**). This is a project to get AI to solve problems and answer questions in a university entrance exam. I was in charge of the subject of English in collaboration with joint-research institutes. At that time, we applied language processing technology with the goal of strengthening the English skills of *Torobo-kun* (the name of the AI. *Torobo* stands for the University of Tokyo Robot, and *kun* is an honorific title in Japanese normally used for boys). We had Torobo-kun study English in order to take the National Center Test for University Admissions, but I actually think that was more difficult than studying by oneself. In particular, clarifying what was difficult was problematic. That admission test is made up of a variety of questions such as pronunciation problems, filling-the-gap problems, long-sentence reading comprehension, and listening comprehension. By utilizing dictionaries and big data, we were able to reach a deviation score of 50.5; however, we have not yet reached the acceptance standard of the University of Tokyo. All the researchers involved are pursuing the next step. By solving this admission test, I hope to improve language processing technology and develop more advanced natural dialogue systems.

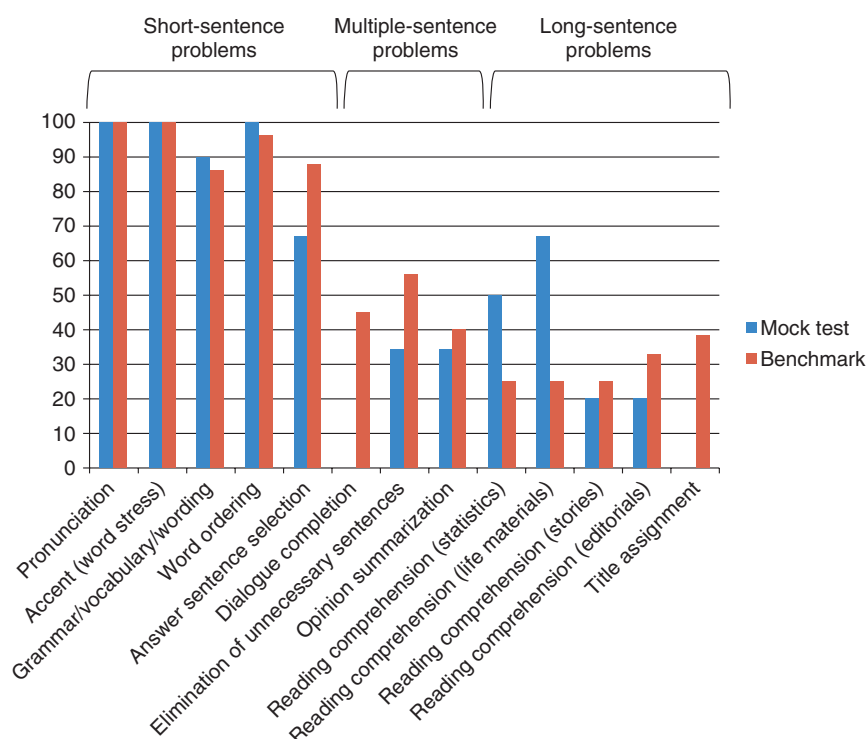Also, in collaboration with Professor Hiroshi

Fig. 2. Scoring rate achieved by Torobu-kun in comparison with benchmark (test data created at laboratory) for various problems in University of Tokyo entrance English test.

Ishiguro of Osaka University, I'm researching an android that can hold a conversation. Professor Ishiguro is at the forefront of research on humanoid robots. He found the point at which human-like robots make us feel uncomfortable and is now pursuing that *something* that makes us humans. Meanwhile, I want to approach human nature by exploring what the essence of dialogue is. In common with these efforts, we have been conducting joint research with Professor Ishiguro such as creating "Matsukoroid," which looks like the famous Japanese television personality Matsuko Deluxe, in 2015.

Moreover, in a collaborative experiment with Dwango Co., Ltd., I have begun work on the development of AI of "Ayase" (a character in a light-novel series called "*My Little Sister Can't Be This Cute.*") to thereby create "Ayase AI" (**Fig. 4**). This project is aimed at developing AI with personality with the user's involvement. A certain amount of data is required to make a computer personality like a person; however, currently there is not enough data in the whole world to express personality, and from the viewpoint of privacy, it is difficult to obtain personal data. Ayase AI started from the idea that if the data are

not available, people should make the data. By asking users to become Ayase and talk on the web, we can collect fundamental data and foster the personality of Ayase based on such data. As the amount of data grows, the dialogue system becomes unique. Users can feel as if they're really talking to someone, so I think we are getting closer to finding the essence of dialogue.

—What is the significance of your research?

I think that the significance of my research is to pursue the essence of human beings. We as humans have elements of ourselves that we do not understand. Humans are social creatures that cannot live by themselves, so we have developed communication skills for the purpose of living with others. I believe that if we can clarify that communication on a scientific basis, humans will get closer to understanding each other. I believe that if human mutual understanding progresses, cooperative work will become smoother, we will feel happier, and so on, leading to improved quality of life. I hope to pursue my goal of developing a dialogue system as a shared property that will make

問5 Most of the students voted ☐12☐ Tom's proposal, and it will be put into practice soon.

   ① at   ② for   ③ into   ④ to

(a) Grammar, vocabulary, wording (completion of declarative statement) (short-sentence problems)

問1 ☐29☐

A one-way trip takes you through the heart of Australia, traveling 2,979 km on one of the world's greatest train journeys. ①The Ghan train was named after the Afghan camel drivers who reached Australia's unexplored center. ②Camels are known to be the best animals for desert transport. ③Starting in Adelaide in the south, it takes 20 hours to reach Alice Springs in the middle of Australia. Furthermore, it takes another 24 hours to reach the final stop, Darwin, in the north. ④Passengers can enjoy the stunning, untouched scenery of the real Australia in comfort. There is a choice of luxury private cabins or the more sociable row seating to suit all budgets.

(b) Elimination of unnecessary sentences (multiple-sentence problems)

Rick's Burger Shack

Since 1926, we've served Laketown's finest hamburgers. Try this month's four special burgers !

#1 Cowgirl $9
The Cowgirl is just like our Classic burger but with 150 grams of Kobe beef. We've added spicy barbecue sauce, bacon, and cheddar cheese.
450 calories   Fat 35g   ⑤ ⑱ T

#2 Firefighter $11
WARNING: Your mouth will catch fire ! The Firefighter is made with 150 grams of Angus beef, jalapeño peppers, and our chef's special hot sauce.
390 calories   Fat 25g   ⑤ ⑱ C

#3 Sailor $12
The Sailor is perfect for seafood lovers. It's made with 150 grams of fresh Ahi tuna cooked to perfection.
310 calories   Fat 5g   ⑤ ⑱ H

#4 Farmer $8
Our customers have asked for a veggie burger for years, and we've delivered ! The Farmer is made with tofu, shiitake mushrooms, and black beans.
250 calories   Fat 0g   V H

⑤ = Small size (100g) for $1 less   ⑱ = Extra large size (225g) for $3 more

V = Vegetarian   H = Heart-healthy

C = Chicken available at no extra charge   T = Turkey available at no extra charge

Extras
Add cheese (brie, cheddar, or pepper jack), avocado, or microgreens for $1 per addition. Add ketchup, mustard, or wasabi mayonnaise for free.

問3 Two friends are visiting Rick's Burger Shack to try its Ahi tuna burgers, which have received great reviews from restaurant critics. One friend wants a regular-sized burger with avocado and cheddar cheese. The other wants a smaller portion with wasabi mayonnaise. How much will they pay for their meal ? ☐41☐

   ① $24   ② $25   ③ $26   ④ $27

(c) Reading comprehension (life materials) (long-sentence problems)

Fig. 3. Examples of problems.

My Little Sister Can't Be This Cute.

Ayase's room

Archive

Ranking

Anonymous user

Please take our survey

Q. Anonymous user: What part of your body do you wash first, Ayase?

Answer

A. Ayase: My left arm first….what kind of question is that!? I'll get you for that!

Current number of question answer pairs

Ask Ayase a question (50 characters maximum)

Ask a question

Fig. 4. Screenshot of question answering bulletin board.

society more harmonious and allow us to lead better lives.

Shifting from liberal arts to sciences, and taking time and effort to overcome adversity

—How did you end up on the path to become a researcher?

Actually, my origins lie in the liberal arts. I thought that I would like to study law at university, but I chose a university and faculty leaning towards public policy after many twists and turns. At university, computer education was vigorous, and we used email—which was rare at that time—to exchange coursework and learn programming. I fit in really well in that environment, so I ended up pursuing programming, and when I was in graduate school, I spent a year and a half as a student researcher at IBM Tokyo Research Laboratory (TRL). While at IBM TRL, I encountered natural language processing, so I decided to take the researcher's path and entered the NTT laboratories.

Because I'm also very interested in foreign languages and have studied in the UK, when I joined NTT, I offered to specialize in research on translation. However, at that time, translation was not a statistical process, like it is now, but a rule-based process conforming to rules such as grammar. Since the translation research I wanted to do was shifting to a commercial basis and not being tackled at basic research laboratories, I was assigned to a department responsible for dialogue systems. From that point on, I decided to engage in research on dialogue for the first time. However, that research was really difficult, and even seemingly simple tasks like booking a meeting room by talking to the dialogue system were difficult.

With that difficulty in mind, I became more interested in why people can converse, and gave myself up to research on dialogue. Around 2001, AI was still in a period of winter-like hardship, and dialogue systems were receiving little attention in a field considered to be for diehard researchers only. However, since I began working on dialogue systems back then, I have had many opportunities to present my research—which has been ongoing for 17 years—and this research has started drawing attention.

—What is the driving force behind your research activities?

The driving force behind my research is "curiosity." I consider that everything is interesting, so I take a stance in which I never refuse what comes my way. And since nothing can be done in one leap, I think it is better to do a lot of experiments and find things out one by one. Anyway, I value taking time and effort in my pursuits. In that sense, I think of myself as a craftsman. As well as taking time, research does not necessarily lead to successful outcomes, so I think that it is enough to obtain one or two successes out of 100 attempts.

When I joined NTT, I was coming from a liberal arts background but was with many employees who came from a science background, so I was immersed in things I did not understand. To overcome this hurdle, I worked on problems more carefully than others, and I applied trial and error as much as possible. My experiences of trial and error led to the present and have built confidence in me.

—Do you have anything particularly memorable concerning your research activities?

We participated in a large-scale event held in the USA, called "SXSW" (South by Southwest), which combines a music festival, film festival, and interactive festival, in two consecutive years (2016 and 2017) (**Fig. 5**). In 2016, we showcased the dialogue system of an android that talked in English in Professor Ishiguro's demonstration, and it was well received. In 2017, I was invited as a featured speaker and took the podium with Professor Ishiguro. It was a very honorable occasion, since among Japanese people, only very few have received such an invitation. That year, we wanted to demonstrate more advanced technology than what we did the previous year, so we decided to set up a discussion between humans and robots. We set up a situation in which robots outnumbered humans, based on the assumption that we might create a world in which robots are a majority. We wanted to present a world in which a lot of robots that are smarter than humans exist and to get people to think about what human beings really are.

By the way, the demonstrations in 2016 and 2017 were packed with drama. In 2016, the demo system was only completed the day before the demonstration. The system needed a fairly complicated program, which I continued to write after arrival, not to mention on the plane to the event. We finally got the program to run when the system and the dialogue partner were "face to face," so to speak. Although the situation at that time is now the stuff of legend, it was truly a miracle that the program ran.

Fig. 5. Demonstration of argumentative dialogue system at SXSW 2017 (in collaboration with Ishiguro Laboratory, Osaka University).

On taking the podium in 2017, the system did not work as we expected. I managed to keep talking for about 30 minutes, and in the meantime, we made various adjustments to the system, and it finally started working. And when it worked, everyone broke into applause. Because I was in front of more than 1000 spectators, I felt frustrated when it didn't work right away. Even so, I was optimistic that it would work at some point. Since I knew that dialogue systems are complicated and don't work that easily, I was probably just keeping my nerve.

Let's find places of interest, things, and other matters, and move forward continuously

—*How will research on dialogue systems proceed in the future?*

The level of Japanese research in this field is relatively high and has developed rapidly recently. Our research is of course often influenced by the research and development done by major global companies. Nevertheless, many concepts that I thought about have been achieved since I joined the company, and I have a "sense of the future." As for my goal, I'd like the act of speaking with a dialogue system to become an everyday event. Since AI and computers are not good at everything, I think that it is better to divide the tasks that humans will bear responsibility for and the tasks left to computers, and that division will allow us to enter an era in which we can live a richer life. To that end, I want to build a good relationship between humans and computers (robots).

Robots can work longer than people can, and only robots can do advanced simulation calculations. And some tasks are better done if they are not done by humans. From the viewpoint of privacy protection, there are certain occasions, such as counseling, when it is better to leave it to the robot. Recently, dialogue systems have started to be applied in the counseling field. Efforts to introduce dialogue systems in facilities for the elderly and to apply them with the aim of dementia prevention are underway—which is said to be a promising area.

However, I think that whether AI is able to respond to unknown events or whether it could acquire an ego are major challenges. Current AIs do not possess an ego, and discussions on ethics and legal matters in regard to whether an ego should be given to AI are ongoing. Although guidelines concerning those matters are already being created, the industry has not yet reached a consensus. On the other hand, I think that we cannot carry out dialogue, in the true sense of the word, with things that have no ego. In the future, we will also conduct research on giving AI an ego, and I'd like to pursue the concept of "What is a human being?"

—Please say a few words to all young researchers.

I have been very blessed to have good co-workers and leaders. Now that we are in a technological boom, the number of people working on dialogue systems has increased accordingly, but when I started, there were few such researchers. Fortunately, NTT had a group dedicated to research on dialogue, and many employees were engaged in that research. For the time being, I have been concentrating on academic activities in order to publicize the value of the dialogue system and increase the number of like-minded researchers. We organize a symposium called Dialogue System Symposium, and it has expanded to a large-scale event attended by 200 people. Some of our projects have been exhibited at events in the USA and have become internationalized. Through such projects, we have come to share a common sense of value with researchers in various fields and with people in different positions, which is very meaningful.

It is currently difficult for fourth-year undergraduate students aiming to become researchers, postgraduate students in master's programs, and new employees to start research on dialogue systems if there are no colleagues in the same field. I have tried various things to foster the next generation in this field, including publicizing data to reach such people, planning events, and so on. Through these kinds of activities, I'd like to continue to increase the number of fellow researchers.

I think that it is better to share new concepts and ideas to get closer to the truth. It is a good time to be an AI researcher now. The status is also getting better, and compared to when there was no Internet, the amount of data is now abundant, and we have all the tools needed. We are in an era where you can push forward with your own ideas and anyone who has the ability can do anything. Since the world is becoming borderless, aim to be top class and keep the world in mind. Our ability is our output. There are various ways of producing output such as getting papers published, writing programs, and implementing systems, and if you leave your mark through output and results, you will be acknowledged as a fellow researcher. For that reason, trial and error is important. Although there are many different research styles, let's find interesting things that will connect to good outputs.

Acknowledgments

■ **Interviewee profile**
Ryuichiro Higashinaka

Senior Research Scientist (Senior Distinguished Researcher), Knowledge Media Project, NTT Media Intelligence Laboratories.

He received a B.A. in environmental information, a Master of Media and Governance, and a Ph.D. from Keio University, Kanagawa, in 1999, 2001, and 2008. He joined NTT in 2001. His research interests include building question answering systems and spoken dialogue systems. From November 2004 to March 2006, he was a visiting researcher at the University of Sheffield in the UK. He received the Maejima Hisoka Award from the Tsushinbunka Association in 2014 and the Prize for Science and Technology of the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology in 2016. He is a member of the Institute of Electronics, Information and Communication Engineers, the Japanese Society for Artificial Intelligence, the Information Processing Society of Japan, and the Association for Natural Language Processing.