

Conversational Computing

Yuji Nomura

Abstract

The advancement of voice recognition technology with enhanced context/emotion interpretation will make natural and seamless people-to-technology interactions possible. Such intelligently interactive systems will change human behavior, societal interactions, and decision making.

Keywords: voice recognition technology, chatbot, emotional recognition

1. Expansion of interactive interfaces

It is now customary to see people talking to their smartphones. In fact, a personal voice assistant installed in smartphones and smart devices is being used in a variety of ways from checking the day's schedule and weather, to searching for nearby restaurants. It has even become commonplace to use text messages to inquire about, order, pay, or send money via wire transfer for products. A recent survey indicates that approximately 89% of users wish to use a messaging tool for business communication [1].

In addition to smartphones, voice assistant terminals installed in residences and offices are also becoming popular in both Europe and the USA. These terminals are always ready, so when a user gives them instructions simply by talking to them, they will start music, adjust air-conditioner temperatures and room lights, or order products.

Behind the popularization of interactive interfaces is the fact that as smartphones have become popular, voice conversations with computers and the use of text chatting have taken root in our culture. The number of active global users who use messaging applications (apps) at least once a month is now over 3 billion. One of the factors for this may be that users have simply gotten tired of using ordinary apps. Users feel it is bothersome to install and learn to use each individual app. In fact, one in four users will quit using a new app after using it only once. In contrast, messaging tools use natural languages, which are familiar to humans, and users can immediately start using these tools irrespective of their information technology literacy.

The advanced processing capability made possible through interaction technology, which was enabled by the evolution of artificial intelligence (AI) in recent years, is also boosting the popularity of messaging tools. Chatbots, programs that use AI to automate communication, are becoming particularly popular. Chatbots can provide a wide range of services from flight reservations to real estate suggestions. In addition, plans are in place to use chatbots in banks to provide account balances and wire transfers and even financial plans based on usage patterns. Interactive services in coordination with AI will likely continue to increase in the future.

2. Development of conversation support technology

The advances in deep learning technology in recent years have significantly increased the accuracy of recognizing images and other patterns. In October 2016, the error rate of voice recognition technology reached 5.9%, equivalent to that of speech-to-text experts. In addition to voice recognition, lip reading with AI has achieved an accuracy of 46.8%, approximately four times as accurate as professional lip readers. This may enable users to interact with computers in situations where they are not able to speak or are in a distant location. Computers cannot only read but are also capable of varying tone and pausing speech depending on situations, producing voices close to human voices.

The use of technology to recognize emotions from voice, expressions, and text has also been spreading. For example, analysis of emotional changes in

viewers enables the identification of effective scenes in commercials and differences in responses depending on countries and cultures. This technology is used widely in marketing. Call centers also use similar technology to support their customers with consideration for their emotions. Future uses of computers will expand from human assistance purposes to reading human emotions for direct interaction. For example, a computer that perceives anger in a user driving a car might talk to the user to calm him or her down.

Unlike humans, computers still cannot understand contextual meanings that hide behind words. For this reason, humans may feel stressed and disappointed about the interaction, and there may come a period of disillusionment with these technologies. In the near future, however, a context-understanding technology may be launched just as technologies such as voice and emotional recognition, and voice synthesis have been developed. These technologies, which support human interaction with computers, will enable more natural interaction, as they more accurately understand the nuances of a conversation, including a user's intent and emotion. It is anticipated that this will further expand the use of interactive computing.

3. Interactive computing that innovates the world

Interactive computing in coordination with AI will result in ultimate personalization. Traditional personalization methods use past behaviors such as viewing and purchase histories. This method is often used to present to the user products already purchased and related products in which the user is not even inter-

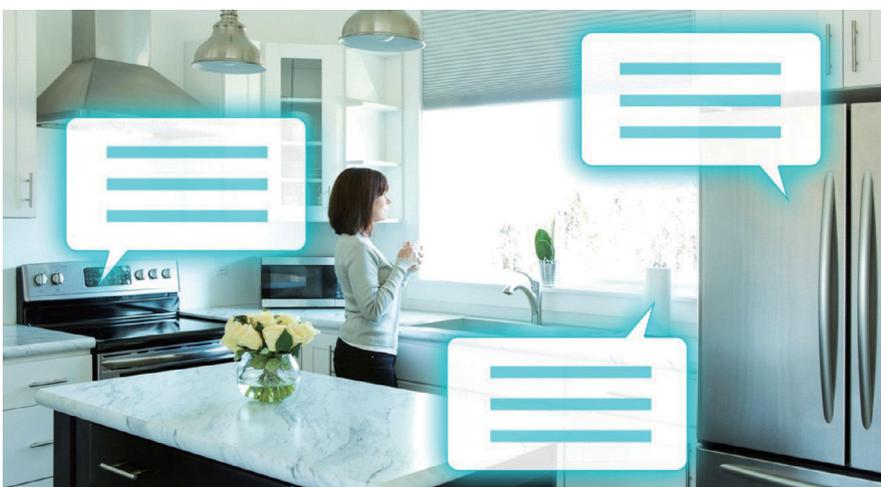
ested. However, an interactive system can use conversations to interpret a user's intention and respond accurately even to complex requests. This system, when combined with information from sensors, can respond to a user's individual situation immediately. Personalized information will be available that defies comparison with what is used today, changing customer service, marketing, and advertising.

Conversation is the most natural communication tool used by humans. With a conversation-based interface, many future actions will be completed using interactive apps. Instead of opening a different app for each purpose, a system will likely be developed in which a personal assistant app listens to a user's requests and then distributes the tasks to other apps and chatbots based on the content. As automatic responses from AI become possible, immediate and appropriate responses will be available 24/7, which will have an enormous impact on the relationship of individuals to society. The traditional individual-to-society connection mainly consisted of one-way information by companies, such as emails and online advertisements, and a temporary connection by telephone or direct visits to brick-and-mortar stores. With interactive computing, however, the company-to-individual, and individual-to-individual connections will become bidirectional and continual.

Interactive computing even has the potential to change the means of decision making. Currently, users need to select relevant data from an overwhelming amount of information and make decisions such as product purchases and travel arrangements. In interactive computing with AI, the information required is narrowed down step by step through interaction, leading to more natural decision making. Thus, not only will the process of decision making change, but satisfying and prompt decision making will also become possible.

4. The future of interactive computing

Interactive computing goes beyond a mere convenient interface based on speech. Accumulation of interaction histories and sensor information will



increase the accuracy of behavior predictions based on the situation and preferences of the user. By forecasting users' intentions ahead of time and providing the necessary information from the system, interactive computing may well become a virtual enabler of human behaviors prior to individuals becoming aware of the desire for the behavior. For example, the system may be able to provide users with the information and data they need before they realize it.

The ultimate communication may not be via language, but the immediate and accurate transfer of intentions. Studies on a brain-computer interface, which reads human intentions using brain waves to control devices, are being actively pursued. As a result, it may become possible to control computers and vehicles by merely thinking about something. The ability to read complex thoughts may eliminate the problem of not being able to express one's thoughts in words, giving rise to interactions with no

gap in communication.

Interactive computing is not merely an app on a smartphone or a personal computer. It will be installed as a standard feature in all kinds of devices, while making its way to becoming a new computing infrastructure. It will not be long before we can simply talk to a device to get any information or product, at any time and from anywhere such as at home, at a store, or in a car.

Reference

- [1] Twilio, "Global Mobile Messaging Consumer Report 2016."
<https://www.twilio.com/white-papers/global-mobile-messaging-consumer-report-2016>

Trademark notes

All brand names, product names, and company names that appear in this article are trademarks or registered trademarks of their respective owners.



Yuji Nomura

Deputy Manager, Strategy Development Section, Research and Development Headquarters, NTT DATA Corporation.

He received an M.S. in science and technology from Keio University, Kanagawa, in 2005. Since joining NTT DATA in 2005, he has researched and developed a text processing technology system centered on information extraction technology. He is a member of the Information Processing Society of Japan.
