Building a Theory of Communication

Naonori Ueda†

Abstract

At NTT Communication Science Laboratories, we are conducting basic research in communication science with the aim of making telecommunication as accurate and satisfying as face-to-face communication between people. The Feature Articles in this issue describe our research efforts aimed at building a theory of communication as a foundation for communication science and introduce our latest achievements.

1. Introduction

What is information? A well-known answer to this question was provided by the information theory proposed by Claude Shannon in 1949. Shannon laid the foundations of information theory by devising mathematical definitions for quantities of information and entropy. Needless to say, the subsequent expansion of this theory led to the astonishing speed and accuracy of today's information and communications technology. But what is communication? There are limits to how far information theory can currently go towards answering this question. This is because the quality of communication is evaluated only in terms of the probability of the occurrence of an event, which defines the amount of information that it conveys. When humans communicate with each other, the important thing is not how much information is conveyed, but the intrinsic value of this information. In other words, instead of quantitative objective measures such as speed and accuracy, greater importance is attached to qualitative subjective measures like satisfaction. In situations such as job interviews, greater importance has recently been attached to more abstract qualities such as the ability to gain a deep understanding of the other person's intentions and provide accurate answers rather than the ability to speak fluently.

To deliver a rich communication environment, it seems that simply researching information itself is not enough; we must also study the humans who are actually sending and receiving this information. At NTT Communication Science Laboratories (CS Labs), in order to pursue the essential qualities of communication via the dual aspects of computer science and human science, we are conducting basic research on three fronts—creating the communication environment of the future, establishing an intelligent computing platform, and delivering a rich quality of life (QoL) for humans. Specific examples of the work that we have recently been doing in these areas are described below.

2. Creating the communication environment of the future

We are constructing a remote communication system called the t-Room with the aim of implementing more natural video communication in the future, and we are developing middleware to achieve a sense of *co-presence* and analyzing and evaluating remote collaborative work in this environment. Here, co-presence refers to the perception and recognition of people and objects that exist in the room being mutually shared with people in the remote room in order to provide video communication in a more natural way in the future. We are also conducting research aimed at implementing a system that allows people to communicate in a natural way. Recently, we have been developing technology that automatically learns interaction control (back-channeling and questioning) from interactions between humans and from interactions between humans and systems.

We are also building a system called the s-room, which recognizes people's actions if they simply

† NTT Communication Science Laboratories
  Soraku-gun, 619-0237 Japan

wear a camera and acceleration sensor on the wrist. We intend to use this system in applications such as analyzing human actions and watching over the elderly.

3. Establishing fundamental technologies for intelligent computing

Most of the computer science research at the CS Labs is aimed at establishing fundamental technologies for intelligent computing and extends across a wide range of research fields. The main technical fields are introduced below. Techniques for finding specific information from vast quantities of information will become increasingly important in the future.

At the CS Labs, we have spent over ten years working on media search techniques that can rapidly find a desired item of audio or video content from a large collection of material. As far back as 1998, we had developed technology capable of searching six hours of video content in just two seconds (allowing us to find a particular TV commercial from videos containing thousands of commercials). By 2010, we were able to search 60,000 hours of content in one second, and we developed a robust media search technique capable of performing searches very accurately even if the audio and video media are degraded by noise. This technique has been put to practical commercial use in a system for detecting copyright-infringing videos on video-sharing websites.

In our speech recognition research, our technology was capable of recognizing words read out directly into a microphone by the late 1990s, but the technology has recently advanced to the stage where it can recognize conversational speech from more than one person via a remote microphone. In conversations involving more than one person, there are inevitably problems due to effects such as sound blending and reverberation inside the room. To address these problems, we are researching speech and audio processing techniques such as an audio source separation technique that can resolve multiple audio sources in a real environment and a reverberation control technique.

We are also developing technology for recognizing the circumstances of a conversation (identifying who is talking to whom and who is holding the attention of the others and recognizing emotions such as sympathy and antipathy) in order to automatically analyze scene attributes such as the atmosphere of a conversation. Furthermore, we are working hard to develop lossless coding techniques that are free of distortion and to implement international standards for these techniques.

In recent years, as a result of the explosive growth of the Internet, the range of applications for natural speech processing has diversified to include information retrieval, detection of illegal or harmful information (e.g., spam filtering), and analysis of sentiments on websites. As this trend continues, the diversity of languages has become a significant problem. Documents such as blog posts and emails often contain neologisms and non-standard grammar. To tackle this problem, we are developing a natural language processing technique based on a machine learning approach in which a language model is automatically trained from a large volume of language data by using the power of computers. We have already made a number of key achievements using this technique. For example, in the syntax analysis technique that automatically analyzes the dependence relationships of words (a core technique in natural language processing), we utilized a semi-supervised learning technique[*1] to obtain the best analysis precision ever achieved in an international standard benchmark test. Moreover, in Japanese-English machine translation, we have not only developed translation technology but also proposed an automatic evaluation method called RIBES[*2] (rank-based intuitive bilingual evaluation score) that is closer to human evaluation than the conventional evaluation metric of BLEU[*3] (bilingual evaluation understudy). We are also utilizing these cutting-edge language processing technologies to develop a medical information access system that allows non-native speakers of English to obtain the latest medical information from English-language medical literature. As for machine learning technology itself, we have proposed a topic model based on Bayesian statistics[*4], which we have applied to data mining, and we are also engaged in cutting-edge research of topics such as non-parametric Bayesian theory[*5].

---

*1 Semi-supervised learning: A learning technique that uses both supervised (correctly labeled) and unsupervised (unlabeled) data.

*2 RIBES: A method for evaluating a translation system by focusing on words that appear in common between a machine translation and a correct translation and awarding scores on the basis of the correlation coefficient of the sequence in which these words appear.

*3 BLEU: A conventional measure of translation quality for machine translation.

*4 Bayesian statistics: Statistics derived from an inverse probability calculation method based on Bayes' theorem.

*5 Non-parametric Bayesian theory: A generator model where the number of models increases according to the complexity of data and can theoretically increase to infinity.

Research aimed at making a quantum computer, which will deliver ultrafast processing by means of massively parallel computation based on the principles of quantum mechanics, is being conducted by NTT Basic Research Laboratories. At the CS Labs, we are working on a study of the quantum information processing performed in quantum computers. By using a quantum algorithm, it has recently become possible to solve the anonymous leader election problem—a problem that is impossible to solve with classical algorithms. Other topics of theoretical research include the world's fastest generation of physical random numbers using a semiconductor laser and a theory of privacy verification. Details of the latter can be found in the Feature Article "Mathematical Duality between Anonymity and Privacy and Its Application to Law" [1].

4. Delivering a rich QoL for humans

As mentioned in the Introduction, in order to construct a theory of communication, we need to study not only media processing and information processing, but also human information processing mechanisms. At the CS Labs, we are taking a psychophysical and neuroscientific approach to the study of mechanisms that process information in the brains and bodies of humans, such as human sensory and perceptual mechanisms, the actions of emotions, and the mechanisms of movement. We have recently clarified the brain's mechanism for seeing textures, and we have devised a novel method for using the discomfort and illusion of a stopped escalator to generate the sensation of traction forces, which are used by the *Buru-Navi* traction force device [2]. With regard to illusions, we have established the Illusion Forum [3] on the web where people can experience visual and auditory illusions for themselves. We have also made progress in our research related to tactile senses: further details of this can be found in the Feature Article "Communication Research Focused on Tactile Quality and Reality" [4].

References

[1] K. Mano, "Mathematical Duality between Anonymity and Privacy and Its Application to Law," NTT Technical Review, Vol. 9, No. 11, 2011.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201111fa5.html
[2] Buru-Navi haptic interface.
http://www.ntt.co.jp/RD/OFIS/active/2007pdfe/hot/ct/05.html
[3] http://www.brl.ntt.co.jp/IllusionForum/
[4] J. Watanabe, "Communication Research Focused on Tactile Quality and Reality," NTT Technical Review, Vol. 9, No. 11, 2011.
https://www.ntt-review.jp/archive/ntttechnical.php?contents=ntr201111fa6.html

Naonori Ueda
Director, NTT Communication Science Laboratories.
He received the B.S., M.S., and Ph.D. degrees in communication engineering from Osaka University in 1982, 1984, and 1992, respectively. He joined the Yokosuka Electrical Communication Laboratories of Nippon Telegraph and Telephone Public Corporation (now NTT) in 1984. In 1994, he moved to NTT Communication Science Laboratories, Kyoto, where he has been researching statistical machine learning, Bayesian statistics, and their applications to web data mining. From 1993 to 1994, he was a visiting scholar at Purdue University, Indiana, USA. He is a guest professor at the National Institute of Informatics and the Nara Advanced Institute of Science and Technology. He is a Fellow of the Institute of Electronics, Information and Communication Engineers and a member of the Information Processing Society of Japan and IEEE.