Steps towards Ambient Intelligence

Eisaku Maeda† *and Yasuhiro Minami*

Abstract

A research project on "ambient intelligence" recently launched by NTT Communication Science Laboratories aims to envision a new lifestyle made possible by communication science. Research and development of "ambient intelligence" should bridge the boundaries between technological fields and thus cover the entire field of communication science, rather than be limited to specific fields. Besides performing the basic R&D, we are striving to get the concept established and conducting design and publicity activities related to ambient intelligence in a comprehensive and strategic way. This article introduces this new project.

1. Introduction

A research project on the theme of "ambient intelligence" (*kankyo chinou* in Japanese) was launched by NTT Communication Science Laboratories (CS Labs.) in October 2004. It is being primarily undertaken by the Ambient Intelligence Research Group and the Intelligence Integration Open Laboratory of CS Labs. It has two main purposes: the first is to propose a new lifestyle that will emerge from research and development of communication science based on the concept of ambient intelligence and the second is to break down the barriers surrounding existing research disciplines and strategically develop the interdisciplinary area, as symbolized by the term "intelligence integration". In this article we outline the Ambient Intelligence Project and its achievements to date.

2. What is "ambient intelligence" (*kankyo chinou*)?

"Ambient intelligence" implies intelligence that is all around us. The term ambient intelligence is also used for research being pursued in the EU and the USA. The aim of those studies is to produce a design for a future information society by making use of computers and sensors embedded in the environment so as to create more natural and intelligent interfaces. Terms such as "ubiquitous" or "pervasive" come close to describing the concept. However, such R&D studies tend to focus on computers and sensors, and as such are geared toward hardware or devices.

On the other hand, our Ambient Intelligence Project is focused more on such things as hearing, vision, language, and knowledge, which are all related to human intelligence. The aim is to discover what kind of communication we really need and to achieve it. This is where our "ambient intelligence" differs from such concepts as ubiquitous computing or conventional ambient intelligence. A very rough way of distinguishing between them would be to think of ambient intelligence as software and of ubiquitous computing as hardware. Or we could say that ubiquitous computing is an infrastructure and ambient intelligence is the algorithms on that infrastructure. In other words, both are important and complementary fields of research. That said, the ambient intelligence field has a greater number of areas that remain to be explored and is also closer to communication science. Concepts such as ubiquitous computing are premised on the existence of devices and computers and are concerned with hiding and embedding those devices in the environment. On the other hand, "ambient intelligence" is aimed at achieving intelligent communication between human beings and their environ-

† NTT Communication Science Laboratories
Soraku-gun, 619-0237 Japan
E-mail: maeda@cslab.kecl.ntt.co.jp

ment. Whereas ubiquitous computing is a way of monitoring the environment in which we live, ambient intelligence is something that watches over us and talks to us.

Many people are likely to associate the word intelligence with robots. Current studies on robotics are mainly focused on the development of mechanical control systems. While some robots are being designed to play the role of a communication partner, the communication capability achieved so far is still at the level of toys. With this in mind, we plan to apply information processing technology for communication. That is, we will use NTT's areas of expertise—speech, sound, language, dialogue, vision, data retrieval, networking, and so on—to enable us to propose a new style of ambient intelligence that places human intelligence and intellect at the forefront. That will lead to a proposal about a future lifestyle itself. Furthermore, "ambient intelligence" may drive a robot brain or be embedded in an environment of ubiquitous computers and networks.

3. The world of fairies and goblins

"Ambient intelligence" is premised on future communication partners that we want constantly by our side. To develop this concept, we started by considering appropriate fundamental features. For example, such entities will have a personality and be able to evolve, grow, and express their feelings. They might seem to have the same level of intelligence as precocious children. They hide nearby and are unobtrusive, unassertive, and even a little mischievous. They respond when called and are visible to some people but not to others, as shown in **Fig. 1**.

This world of "ambient intelligence" is rather similar to the world of fairies and goblins, which is familiar in both the East and the West. This makes "fairies and goblins" an adequate concept on which to base a common understanding of ambient intelligence. The world of fairies and goblins is close to the true nature of ambient intelligence in two respects. First, it is a world that is familiar to most cultures through their folklore. Indeed, until recently, it was a familiar part of daily life in Japan. When we were children, many of us imagined fairies and goblins living around us and guarding us. With the help of information technology, perhaps we will be able to rediscover various aspects of this forgotten world. Second, from the standpoint of modern science, fairies and goblins are very much a product of the mind. If we search the contents of our own intelligence, we are bound to arrive sooner or later at this world of fairies and goblins. It is in this sense that the human sciences are

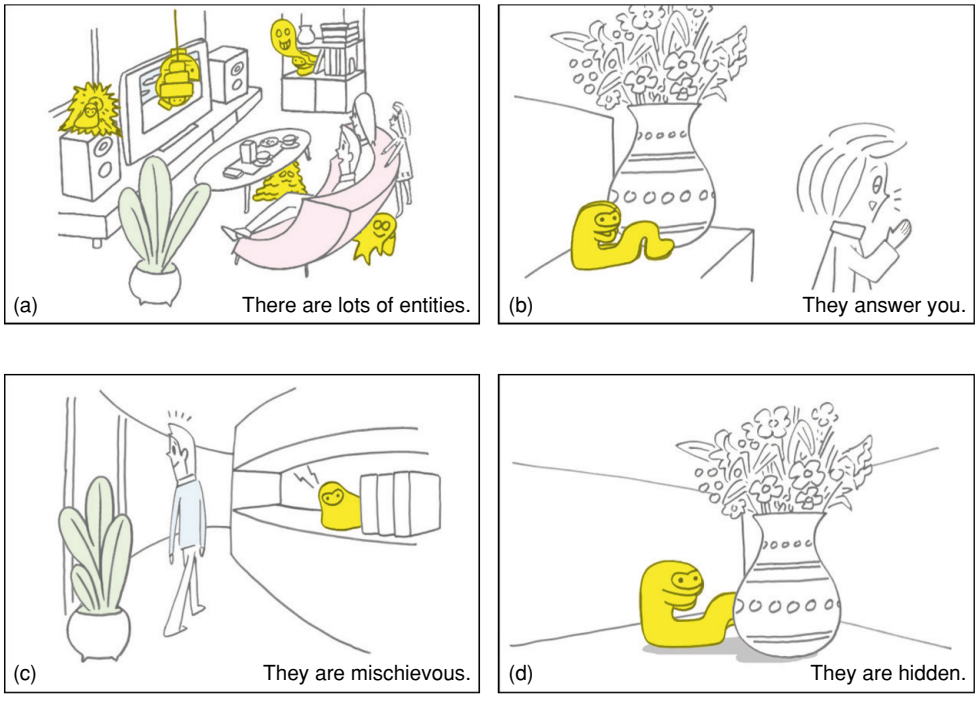


Fig. 1. Ecology of ambient intelligence ("mushrooms").

linked to engineering for the sake of communication. Therefore, the fruits of cognitive psychology research in CS Labs will also contribute to ambient intelligence.

Ambient intelligence also provides a means for communication that speaks to the heart, as fairies and goblins used to do. Today, when information technology has advanced to such a great extent, the time is ripe for thinking about new ways of using this technology in our daily lives as a tool for refreshing the heart and mind. This approach will, we believe, reveal the future of information technology as seen by the person on the street.

4. The world of mushrooms

We gave the name "mushrooms" to the fairies and goblins that are our idealized vision of ambient intelligence. These mushrooms live in a world called "Mush-Room". We gave their shapes and designs coherence and uniformity. Each mushroom has one or two specialized abilities, and their (Japanese) names and shapes correspond to these abilities. Three examples of mushroom types are shown in **Fig. 2**. Keekie has the ability to listen, Chacha can speak and butt in, and Seeshie can provide knowledge acquired from newspaper articles. Their shapes change as they

Fig. 2. Appearance of mushrooms.

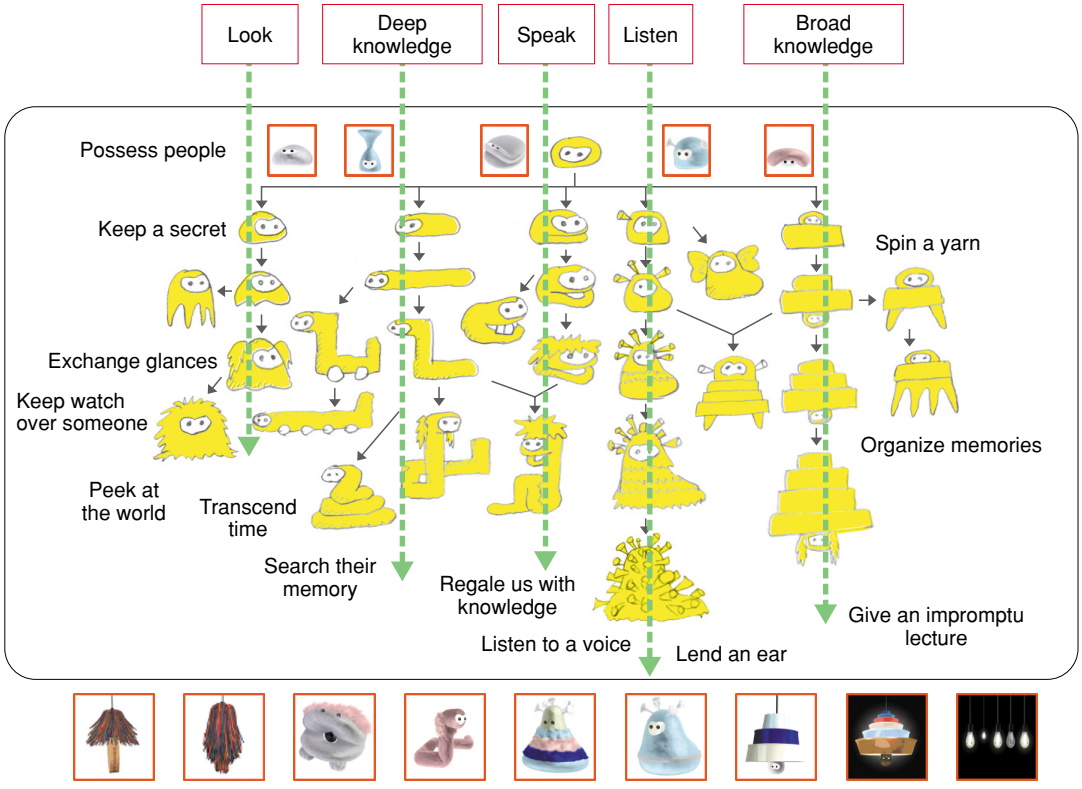


Fig. 3. Evolution of mushrooms.

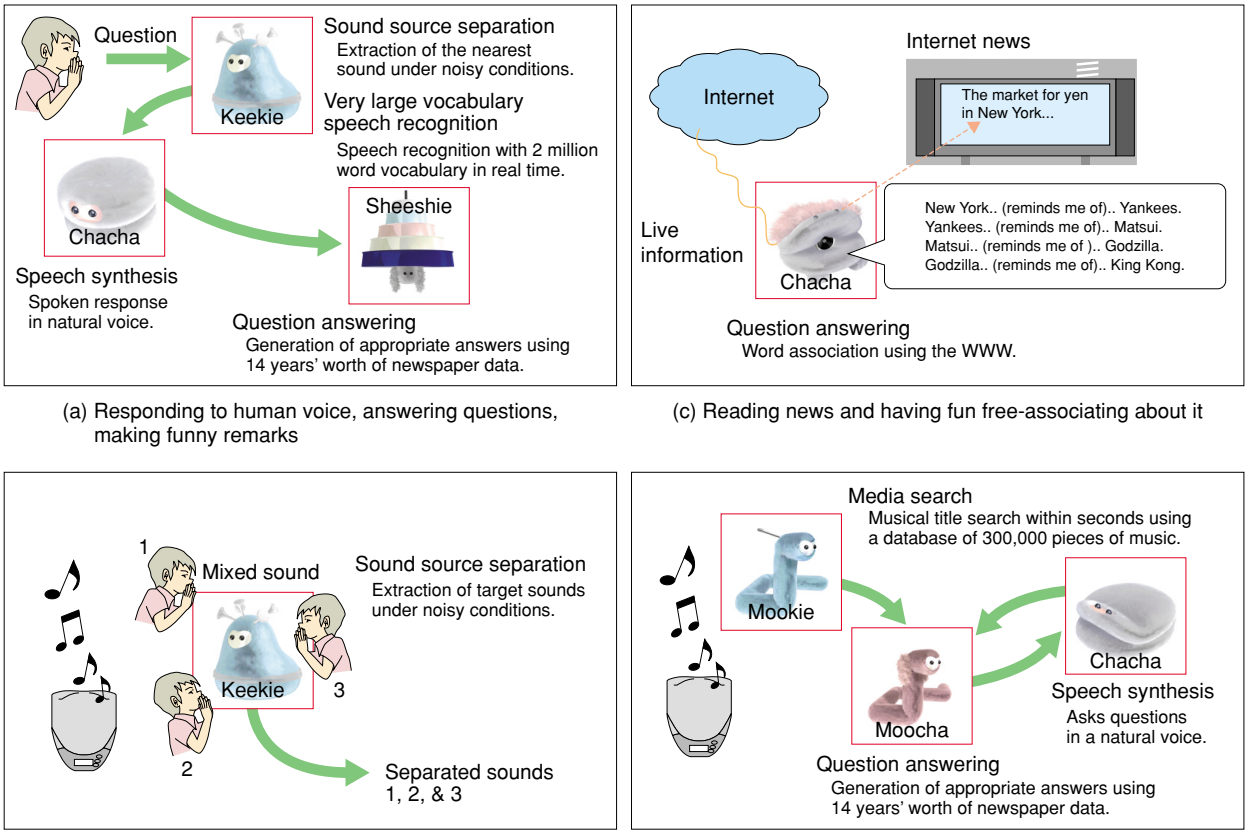
evolve or as their abilities improve. For example, they may become more complicated in shape, lose the dark circles under their eyes, grow longer hair, or acquire more rings around their body as their knowledge increases.

The considerations related to these design issues may be of secondary importance in terms of research and development. Nevertheless, inspired by such designs or shapes, new ideas should emerge. We also believe that such considerations are important from the viewpoint of continuously disseminating information about the concept.

Another important issue in terms of concept design is for the mushrooms to evolve, as shown in **Fig. 3**. The process and the future perspective of mushroom evolution represent the progress of the R&D itself and a road map for the future of "ambient intelligence". As mushrooms evolve, they approach human intelligence or intellect. For example, mushrooms can search their own memories, exchange glances, lend an ear, listen to a voice, convey thoughts, transcend time, organize memories, keep watch over someone, spin a yarn, peek at the world, keep a secret, regale us with knowledge, and give an impromptu lecture. On the other hand, being mischievous little mushrooms, they can also possess people or even poison them. The activities described by these words, whether performed by human beings or by fairies and goblins, conceal the profound intellectual capabilities of human beings. The world of "ambient intelligence" cannot be produced simply by advances in this or that technology. That is why we require a technology for intelligence integration that covers the whole of communication science.

5. Ambient Intelligence 2005

At the CS Labs. Open House 2005, we presented four demonstrations under the general title "Ambient Intelligence 2005: say you want me always beside you, to guard you and to guide you". An outline is shown in **Fig. 4**. Based on the future of ambient intel-



(a) Responding to human voice, answering questions, making funny remarks

(c) Reading news and having fun free-associating about it

(b) Listening in on conversations

(d) Finding the name of a musical work; drawing on a vast store of knowledge to answer questions

Fig. 4. Scenarios from Ambient Intelligence 2005.

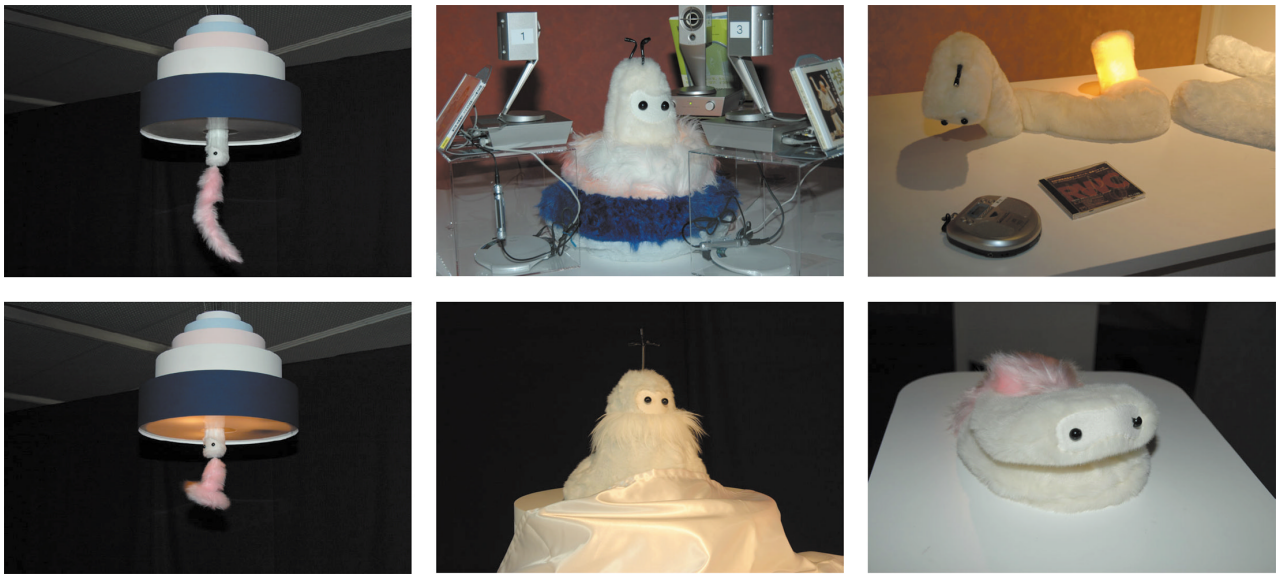


Fig. 5. Pictures of real mushrooms at Ambient Intelligence 2005.

ligence shown in Fig. 1, we created some mushrooms (**Fig. 5**) and presented some of the scenes in a tangible way using our latest research and development. We used five kinds of fundamental technology in the demonstrations.

(1) Sound source separation technology and speech enhancement: Technology for separating a mixed sound of three different sound sources and technology for extracting and enhancing the voice nearest a microphone under noisy conditions [1], [2]. This was implemented in Keekie in Figs. 4(a) and (b).
(2) Very large-vocabulary speech recognition: Technology for recognizing a Japanese speech vocabulary of about two million words. Infrequently used words such as proper nouns or numerical expressions are also recognized [3], [4]. Used in Keekie in Fig. 4(a) and (b).
(3) Japanese question-answering: Technology for estimating answers to a user's questions within a few seconds from 14 years' worth of newspaper articles [5], [6]. Answers can also be found by searching on the World Wide Web in real time. Used in Sheeshie in Fig. 4(a), Chacha in Fig. 4(c), and Moocha in Fig. 4(d).
(4) Ultrahigh-speed media retrieval: Technology for finding the title of music playing on a CD from a database of three hundred thousand pieces of music [7], [8]. Used in Mookie in Fig. 4(d).
(5) Text-to-speech synthesis: Corpus-based text-to-speech synthesis technology called "Cralinet" developed by NTT Cyber Space Laboratories [9]. Used in Sheeshie in Fig. 4(a), Chacha in Figs. 4(a), (c), and (d), and Moocha in Fig. 4(d).

Furthermore, to achieve even wider dissemination of our ambient intelligence concept, we produced a booklet entitled "The World of Mushrooms" and distributed it to visitors to our Open House 2005. The contents can be viewed on our website [10].

6. Concluding remarks

Finally, let us look at the utilization of "ambient intelligence". While some of the fundamental technologies introduced here have already been employed for business, their practical utility must be further enhanced and new application scenarios must be found. New ideas leading to future business need to emerge from the lifestyles created by ambient intelligence ten years from now. We have just started discussions with various people. We would like to continue nurturing this project while interacting with as many people as possible.

References

[1] H. Sawada, S. Araki, R. Mukai, and S. Makino, "Blind Extraction of a Dominant Source Signal from Mixtures of Many Sources," IEEE Proc. ICASSP 2005, Vol. III, pp. 61-64, Mar. 2005.

[2] H. Sawada, R. Mukai, S. Araki, and S. Makino, "Frequency-Domain Blind Source Separation," in Speech Enhancement, J. Benesty, S. Makino, J. Chen, Eds., Springer, pp. 299-327, Mar. 2005.

[3] D. Willett, E. McDermott, Y. Minami, and S. Katagiri, "Time and Memory Efficient Viterbi Decoding for LVCSR Using a Precompiled Search Network," Eurospeech 2001: 847-890, 2001.

[4] T. Hori, C. Hori, and Y. Minami, "Fast On-The-Fly Composition for Weighted Finite-State Transducers in 1.8 Million-Word Vocabulary Continuous Speech Recognition," Proc. ICSLP2004, Vol. 1, pp. 289-292 (2004.10).

[5] Y. Sasaki, H. Isozaki, K. Kokuryo, T. Hirao, and E. Maeda, "NTT's QA systems for NTCIR QAC-1," in Proc. of NII Test Collection for IR Systems Workshop (NTCIR-3), Vol. IV, pp. 63-70, 2002.

[6] H. Isozaki, "An Analysis of a High Performance Japanese Question Answering System," ACM Trans. Asian Language Information Processing (to appear).

[7] K. Kashino, T. Kurozumi, and H. Murase, "A Quick Search Method for Audio and Video Signals Based on Histogram Pruning," IEEE Trans. Multimedia, Vol. 5, No. 3, pp. 348-357, 2003.

[8] K. Kashino, A. Kimura, and T. Kurozumi, "A Quick Video Search Method Based on Local and Global Feature Clustering," Proc. of International Conference on Pattern Recognition (ICPR), Vol. 3, pp. 894-897, 2004.

[9] H. Mizuno, H. Asano, M. Isogai, M. Hasebe, and M. Abe, "Text-to-speech Synthesis Technology Using Corpus-based Approach," NTT Technical Review, Vol. 2, No. 3, pp. 70-75, 2004.

[10] NTT Communication Science Laboratories ed., "The World of Mushrooms," http://www.brl.ntt.co.jp/cs/iol/index.html (Please contact us if you would like to receive a copy of the above booklet (available in Japanese and English versions)).

Eisaku Maeda

Senior Research Scientist, Supervisor, Group Leader, Ambient Intelligent Research Group, NTT Communication Science Laboratories.

He received the B.E. and M.E degrees in zoology and the Ph.D. degree in mathematical engineering from the University of Tokyo, Tokyo in 1984, 1986, and 1993, respectively. In 1986, he joined NTT Integrated Communication Laboratories, NTT. He was a guest researcher at the University of Cambridge, UK in 1996-1997. His research interests are in statistical machine learning, intelligence integration, and bioinformatics. He is a member of IEEE, the Institute of Electronics, Information and Communication Engineers (IEICE) of Japan, the Information Processing Society of Japan, and the Japanese Society of Bioinformatics.

Yasuhiro Minami

Senior Research Scientist, NTT Communication Science Laboratories.

He received the B.E., M.E., and Ph.D. degrees in electrical engineering from Keio University, Kanagawa in 1986, 1988, and 1991, respectively. He joined NTT in 1991 and worked on robust speech recognition. He was a visiting researcher at MIT from 1999 to 2000. Since February 2000, he has been with NTT Communication Science Laboratories. He is interested in modeling for speech recognition. He is a member of IEEE, IEICE, and the Acoustical Society of Japan.