Telme: A Personalized, Context-Aware Communication Support System

Kaoru Sumi, *Communication Research Laboratory*
Toyoaki Nishida, *University of Tokyo*

As communications become more global and instantaneous, we have an increasing opportunity to converse with people who live in different cultures, work in different fields, and speak different languages. New communication devices are now in development, including wearable computers and real-time speech translation systems.[1]

True global communication will require more than just language translation technologies. To fully understand each other, people also need context-specific information. The authors have developed Telme, a support system that gives users real-time information to help bridge the knowledge and experience gap.

Even with such translation systems, however, people will still have difficulty understanding each other if they do not have similar background knowledge or experiences.

To help remedy this, we developed Telme, a communications support system that acts as a mediator between people with varying levels of knowledge and experience. Telme supports real-time communications by presenting information from a knowledge base customized according to the user's profile and operational records.

We model the user's knowledge structure based on two premises. The first is that users' knowledge can be deduced from the questions they ask. In the field of cognitive science, listening to questions is one way of determining how much one knows about a certain subject.[2,3] In our research, this lets us determine what users know without placing a burden on them.

The second premise is that different people can understand the same topic from various viewpoints and use different words and concepts to describe the same thing. We thus use the idea of a "conceptual space" —simple, extensible sets of related keywords—to represent concepts in our system. Although the conceptual space largely depends on the person, some concepts are partially reusable. Cognitive scientists and philosophers have used similar methods to model the human concept-formation process.[4,5]

Telme works by inferring the conversation context based on defined conceptual spaces and the user's own conceptual space, which it generates from the user's profile and his or her system interactions. The system then uses this information to give users "assistant information" about the current topic and its context.

We implemented Telme in two domains—cooking and gardening—and tested the system with 21 subjects, all of whom were computer literate, but had varying degrees of domain knowledge. Based on these tests, we analyzed both the effectiveness and appropriateness of Telme's information for providing real-time communications support.

Framework

Figure 1 shows the Telme framework for wearable computers connected to a central knowledge-base server. The server controls a background knowledge database and downloads data on user request. The computer's display shows four windows:

- the main window, which presents the speaker's dictated words on screen,
- the knowledge conceptual space, which shows the listener's knowledge space;
- the context conceptual space, which shows information about the topic; and
- the assistant window, which shows text and pictures to explain the speaker's words.

Figure 2 shows the wearable computer unit. Telme also works for TV-like broadcasting on a nonportable

1094-7167/01/$10.00 © 2001 IEEE

unit, such as a desk or laptop display.

Presenting assistant information

Assuming that user s is the speaker and user l the listener, we present user l's annotated information in the assistant window $A(l)$ as

$$A(l) = T(d, s) \cdot F(l),$$

where $T(d, s)$ represents the information transformation according to both the topic domain and users' background knowledge and $F(l)$ represents what user l doesn't know. That is, Telme calculates the possibilities of each word's domain based on the current topic domain (d) and user s's background knowledge. From this, the system infers the conversation context and, if the word has several meanings across domains, selects the best meaning. Thus, Telme conveys information about the ongoing topic based on user s's intent and filters it into different information based on what user l doesn't know. When Telme recognizes the speaker's words, it automatically displays information that the listener doesn't know about the current topic domain and the speaker's domain in the assistant window.

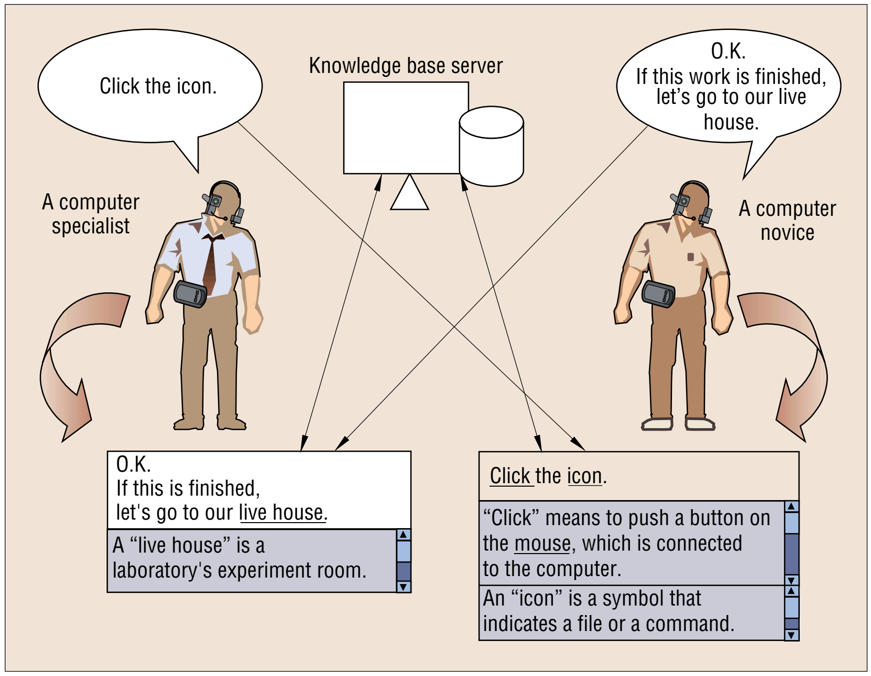

Figure 1. The Telme framework. The wearable computer's display presents speech as dictation in the main window. The assistant window shows text and pictures that describe the speaker's words. Two other windows show additional information: the context conceptual space shows information on the topic and the knowledge conceptual space shows user knowledge of the topic.

Presenting information

Telme's prepared information consists of a knowledge base and defined conceptual spaces. The *knowledge base* consists of terms and question–answer pairs corresponding to each term. The answer part of each question–answer pair is presented as text or pictures, and three kinds of questions are used:

- "*What*," which requests an explanation of the word used;
- "*Example*," which requests an example of the word meaning; and
- "*Why*," which requests the reason for using the word (mainly a verb).

A *conceptual space* consists of keywords defined by their interrelationships. Domain specialists design individual conceptual spaces based on their own viewpoints. For Telme, we designed simple conceptual spaces using documents created by domain specialists. For example, in the gardening domain, we might link two vegetables, such as "tomato" and "eggplant." Such relationships are reversible, and therefore Telme can determine the user's knowledge based on either keyword.

Before using the system, all users must select their domain specialty from among several conceptual spaces. These domains then serve as user profiles.

Inferring user knowledge and conversation context

Telme infers the user's knowledge based on the user's own knowledge conceptual space. The system generates this space based on the conceptual spaces of the specialists and the information space of the current topic; it adapts the space to the user based on his or her operations (such as questions, information checking, and so on).

Nodes in the knowledge conceptual space represent concepts, and links represent relationships between concepts. Both the nodes and links have weight values. The system weights the nodes based on the user's knowledge level and weights links based on the relationships between the linked concepts. Telme assumes that when users question a concept, it is something that they don't know, and weights it and related concepts accordingly. Similarly, when users erase a concept by closing a window, the system assumes user knowledge of it and related concepts.

The system infers users' knowledge level by calculating the weights of concept-related links. Users can also select useful information by clicking a checkbox, which lets them review and print out the information. The

Figure 2. Telme works on wearable computers such as the one shown, as well as on desk or laptop computers.

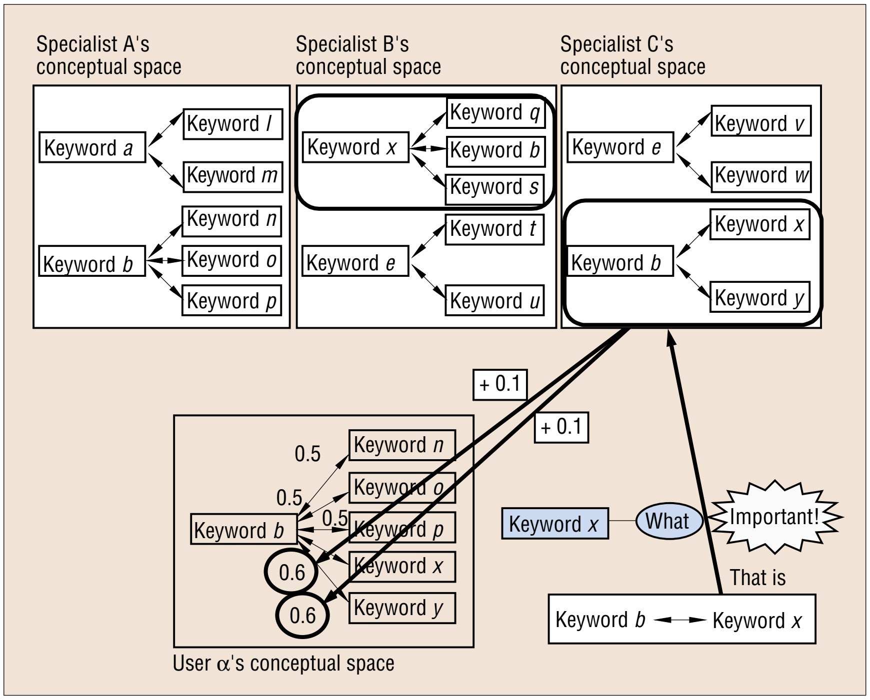

Figure 3. To generate the knowledge conceptual space, the system rewards links related to user operations by increasing or decreasing their weights.

system stores these operational records and uses them to adapt to the user by adjusting the weights of links in the knowledge conceptual space.

How Telme works

First, to extract data from conceptual spaces and apply it to the user's own conceptual space, the system assigns weight values reflecting the listener's background knowledge to all nodes in the knowledge conceptual space. Based on the user profile, the system gives a node maximum value when the listener has knowledge of it and minimum value when the user does not.

For example, when user α queries *Keyword b* of *What*, the system obtains keywords linked to *Keyword b* (such as keywords *n*, *o*, *p*, *x*, and *y*) from the relationship between the number of conceptual spaces in user α's conceptual space. If no links exist in user α's conceptual space, the system gives the keywords initial values based on the values of either the current topic domain *d* or the speaker's background knowledge *s*, whichever is greater. For the links existing in user α's conceptual space, it uses the weights they have at that time.

Next, the system selects information to present. If it finds information that can be presented in the current context, it determines a maximum of five keywords based on the keyword probabilities. When the user drags *Keyword x* and selections *What*, it shows the answer to *Keyword x* of *What*.

Finally, the system updates the user's conceptual space. The user can indicate in the assistant window whether the current information is important or not either by clicking the checkbox or closing the window. For example, if the user clicks the checkbox to save the textual explanation or picture in the assistant window, the system assumes that the explanation is important to the user. If the user closes the assistant window, the system assumes that the explanation is not important.

If user α saves the explanation for *Keyword x* of *What*, the system rewards *Keyword b* ↔ *Keyword x*, which are tracks of inferred keywords. At the same time, the system searches for all conceptual spaces. If a link in the user's conceptual space has the same keyword on either side of *Keyword b* ↔ *Keyword x*, the system rewards the link in the conceptual space. It then updates the user's conceptual space by using conceptual spaces of, for example, specialists *B* and *C* because they contain links to *Keyword b* ↔ *Keyword x*. The link to (*Keyword b* ↔ *Keyword x*) itself is rewarded for updating the conceptual space. Because the links are reversible, it also rewards *Keyword x* ↔ *Keyword b*. Likewise, if a link with the same keyword as that to the left of a rewarded link is available—for example, *Keyword b* ↔ *Keyword y*, *Keyword x* ↔ *Keyword q*, or *Keyword x* ↔ *Keyword s*—it is also rewarded (see Figure 3). If the user indicates that *Keyword x* of *What* is unimportant (by closing the screen), the system lowers the link weights accordingly.

The system assumes that *Keyword b* is in an unknown concept because the user questioned it. The system gives the node a minimum value, then calculates the value of related concepts by multiplying the nodes' weight values by the values of the links to them. If the user erases the concept, the system assumes that the user knows it and gives the nodes maximum values, and (as with the previous example) calculates the values of related concepts by multiplying the nodes' weight values by the values of the links to them. This repeated weight adjustment of keyword relationships corresponds to collecting several people's viewpoints.

To infer the current topic domain, we use a much simpler method based solely on determining word frequency in a specialist's conceptual space.

Presenting the conceptual spaces

The knowledge conceptual space shows the user's own conceptual space and how much knowledge the user has. The context conceptual space shows the current topic and a corresponding specialist's conceptual space.

Telme visually presents conceptual spaces to users in a two-dimensional space by matching the relationships in pairs of keywords to spatial distances using a multidimensional scaling method.[6] By simply glancing at these spaces, the user can easily understand the knowledge structure.

By looking at the knowledge conceptual space, users can see which parts of the knowledge they know or don't know when the assistant window automatically appears. Concepts in the context conceptual space show the user which concepts are related to the current topic. By comparing these two conceptual spaces, users can understand the difference between their concepts and those of a specialist.

Implementation

We implemented the system in two different domains—cooking and gardening—and evaluated it in conversations between people with different knowledge levels. For voice dictation, we used the commercial software application Via Voice by IBM. We installed the system on a Pentium III 600-MHz PC and used the Visual Basic 6.0 programming language.

Figure 4 shows a Telme screen (adapted from the original Japanese version) comprising the four windows. In this case, the system is facilitating a discussion about cooking between a novice and a specialist. The specialist (the speaker) is advising the novice (the listener) about ingredients for minestrone. The dictation area (at the top of the screen) shows the speaker's statement, "I always use garlic, onions, leeks, celery, carrots, kidney beans, potatoes, tomatoes, olive oil, bouillon, and Parmesan."

There are five assistant windows below the dictation area; each defines a word that the system infers the user doesn't know (olive oil, celery, and Parmesan). The lower right shows the context conceptual space, which contains keywords related to the current topic (cooking). The related words ("carrot," "potato," and so on) are shown as dark colored icons so that users can easily distinguish them from other icons. Other terms are arranged according to the relationships between each combination of terms in the field. For example, the vegetables "carrot," "potato," and "onion" are in a group, as are "bouillon" and "Parmesan," which we categorize as Italian seasonings. Herbs are in another group.

In the knowledge conceptual space (upper right corner), the term icons are colored according to weights based on the user's knowledge. The known-knowledge space (selected) shows words that the user knows ranked by priority, with the darkest indicating well-known words and the lighter terms indicating potentially unknown words. Alternately, the user can select the unknown-knowledge conceptual space and the system will prioritize unknown terms.

Based on the user's operational history and user profile, the system makes several inferences. First, using the user's operational history (a conversation about onion gratin soup), the system infers that the user knows "onion" and "garlic," and as a result, does not explain these or related terms ("onion," "garlic," "carrot," "kidney bean," "potato," and "tomato"). The system also infers that the user may not know the meaning of "leeks," "celery," "olive oil," and "Parmesan," based on the user's questioning of the word "leeks."

Figure 5 shows a screen from the gardening domain (also adapted from the Japanese version) where the specialist (the speaker) advises the novice (the listener) on how to care for tomatoes. The dictation area features several underlined words. The system infers that the user doesn't understand the terms "divergent" and "spray leaves" based on the user's questioning of the related term, "determinate" (having growth where a bud or flower terminates at the growing tip). These words are explained in the assistant windows below. When the user clicks on the checkbox in the assistant information window, the specialist's advice can be recalled or printed out later. The system tracks information that the user selects and updates the user's knowledge conceptual space accordingly.

The context conceptual space features several groups, including a group of herbs ("oregano," "sage," and "basil") and a group of synonyms related to growth promotion via pruning ("cropping" and "disbudding"). Finally, based on the conceptual space's contents, the user can assume that "divergent" is a concept related to "bud" and "determinate."

As Figures 4 and 5 show, several terms in the cooking and gardening domains (espe-

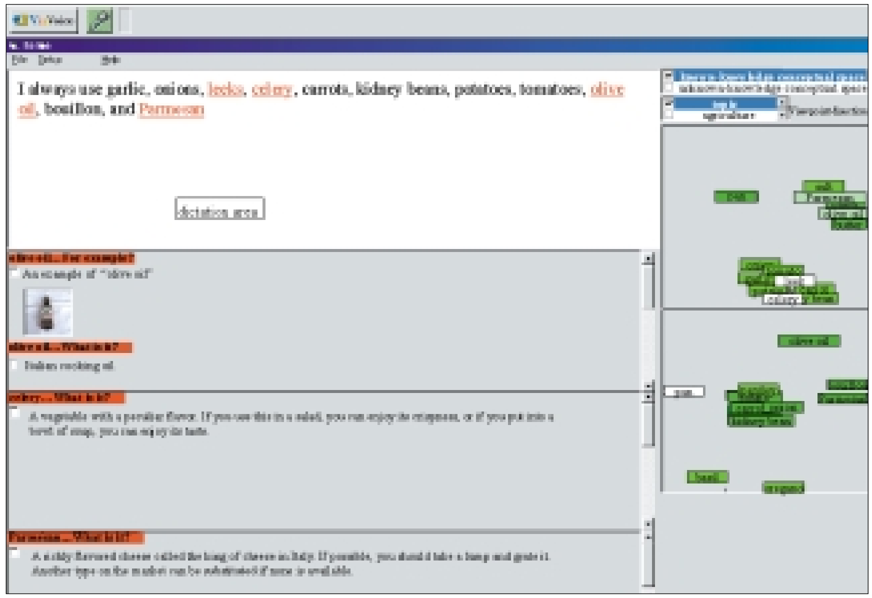

Figure 4. A screen shot of the Telme system. In this example, a novice and an expert are discussing cooking. Dictation appears in the large upper box, with five assistant windows below it (three are shown here). On the right are the knowledge conceptual (upper right) and context conceptual (lower right) spaces.

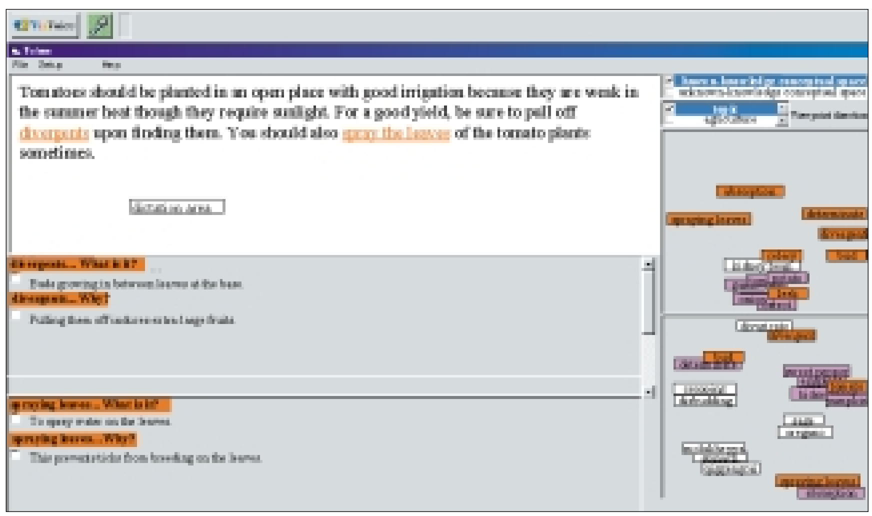

Figure 5. A Telme screen shot from a gardening-related conversation.

cially the names of vegetables) overlap. However, the system treats the terms differently. It presents different explanations for each domain by inferring the current topic and the speaker's specialty. Words related to the terms are also different in each domain, even if the terms are the same (for example, a group of ingredients for Italian cooking would be different than a grouping of leaves and fruits in the gardening domain).

Experiment and evaluation

To evaluate Telme, we performed an experiment using 21 university graduate students and secretaries who had expertise in various subjects (ranging from system science to education and psychology). We selected people who were familiar with computers to circumvent computer-literacy issues.

Because we were mainly interested in evaluating how the system adapts, we prepared small knowledge bases and conceptual spaces. We prepared the scenarios and conversation topics in advance. We asked the subjects to behave as if they were participating in an ordinary conversation and to let the conversation progress naturally. We did not use speech recognition devices. The subjects used the system to read the conversational sentences and assistant information it presents.

The first topic was cooking, followed by gardening. We asked the subjects to continue using Telme until all sessions in the conversation had finished. In our analysis, we focused on the effectiveness and usefulness of the assistant information as well as its appropriateness for real-time communications support.

We analyzed the users' operational histories and the questionnaires we asked them to fill out before and after the test. The questionnaires asked users to rank their impression of the system on a scale from one to five (5 = excellent, 4 = good, 3 = passable, 2 = not so good, and 1 = bad). We considered the values of three or greater to be positive. To gauge the reliability of the answers, we asked the subjects to give reasons for their responses on most of the questions.

Effectiveness of assistant information

When asked to evaluate the effectiveness of the assistant information, 86 percent of the subjects said that Telme presented new information, and 90 percent said that the presented information was useful for understanding the topic. Out of an average of nine instances when assistant information was offered, 2.89 instances were deemed useful. As to relevance, 80 percent of the subjects said that Telme's assistant information was related to the current topic. Other findings include:

- 80 percent said that the knowledge conceptual space presented the relationship between concepts well.
- Most said that Telme's assistant information was compatible with their knowledge level.
- 85 percent said that the context conceptual space presented relationships between concepts well.

In the latter case, one subject said that the icons for the current topic (cooking ingredients) were well-clustered in the conceptual space. Another subject said that it was easy to find the conceptual meanings of unknown words by finding known words that were close to them in the conceptual space.

We divided the subjects into groups according to domain knowledge levels and analyzed the effectiveness of presenting assistant information in each group. We found that as the knowledge level increased, more subjects tended to think that the information presented was relevant to both the topic and their knowledge level. This might be because their advanced understanding of the domain led to greater overall understanding of the content.

Value for real-time communications support

We analyzed the appropriateness of the information in terms of real-time communications support, using both the questionnaires and records of the subjects' system interactions. We first analyzed users' real-time system interaction (such as selecting important information, erasing useless information, and questioning).

All subjects said that they were not reluctant to click the checkbox to save the information, and 85 percent of the subjects said they were willing to use this operation, primarily to reuse and verify useful information. Also, 85 percent said they were willing to click the checkbox to close the window, which erases information. In the comments, most subjects said that erasing was necessary because they became confused when there was too much information on screen. Some subjects gave negative comments, saying that they felt reluctant to discard information that they might need later.

Nearly all subjects (95 percent) said that they would drag unknown terms to the dictation area, then select a question to get more information about the terms. The main reason they gave for their willingness to do this was that it was a quick and easy way to get an answer. One subject said that this operation was especially convenient when she knew something about the question, but was not completely sure of its meaning.

Once the conversation began, on average the user executed the first operation in 0.25 seconds. From this, we conclude that the operation would not necessarily interrupt a conversation. However, 62 percent of the subjects said that they felt the system was slow. Much of this is related to the construction of the conceptual spaces; the assistant information is presented automatically and processed instantaneously. Because Telme uses a multidimensional scaling method to visualize the conceptual spaces, the processing cost is high. However, this is only a problem with the visualization process itself. We can completely solve the problem by processing the visualization individually. For this reason, composing and displaying conceptual spaces should be a background process.

Usefulness for different types of communications support

To explore future possibilities for Telme, we asked for users' impressions about several possible communication devices, including wearable computers. Using the subjects' questionnaires, we also analyzed the general usefulness of presenting assistant information for communications support.

Our analysis captured a peripheral view of the system's support for

- *Computer-chat support.* All subjects agreed that Telme would be a useful computer-chat support system (19 percent checked "5" and 57 percent checked "4"). Among the positive comments were, "We may find it convenient if we challenge something unfamiliar," and "We can talk with someone without the conversation being interrupted." There was also a negative comment: "We may prefer to ask a question directly if it is brief."
- *Generic interview support.* 85 percent of the subjects agreed that Telme would be useful as an interview support system (19 percent checked "5" and 37 percent checked "4"). Among the positive comments were those saying that Telme is very convenient, that it can be used while cook-

Related Work

In the future, communication is likely to become more spatiotemporal, global, and instantaneous. Presenting real-time and personalized information by computer can help people understand each other instantaneously as well. Such technology can also enrich information content.

Current research on using tagging technology to enrich WWW content has received a lot of attention. Semantic Web Development (www.w3.org/2000/01/sw/DevelopmentProposal) is one such project. We believe that helping users better understand a particular topic requires both richer content and the addition of contextual information. Another effort to enrich WWW content is that of Nagao, Shirai, and Squire, [1] but it is not an automatic adaptation system nor does it account for differences in knowledge levels among users.

Bradley Rhodes developed a system that provides information that might be relevant to the context.[2] The system uses a wearable computer and a database of past descriptions (such as memos) as information sources. However, in our research related to human understanding, our desire is to have much richer and more personalized information for different people.

Because Telme adapts information to the user and the context, representing meaning structures is important. Researchers have proposed several question–answer systems for "know-how" knowledge that aim at information distribution among several presenters and users. For example, Answer Garden is based on a relationship network tuned by the user's questions to support the user's knowledge acquisition.[3] If detailed meaning structures can be designed, good results can be generated.

However, because question–answer systems require the information spaces of multiple presenters, such systems require multiple, detailed meaning structures. Building these structures is expensive, and, once built, they cannot be modified easily. Because a meaning structure tends to be biased by the viewpoint of the person who gives it its meaning, the system can only consider specific user concepts; it cannot consider each user's concepts individually. The use of meaning structures can therefore improve the system performance only in a general way, and not for individual users.

We propose a system that uses knowledge-space structures as simple related links. With these structures, the system can generate the user's conceptual space while taking into account his or her individual viewpoints and existing concepts from built-in conceptual spaces.

We previously designed an information navigation system called Takealook.[4] The system infers the user's interests by generating a conceptual space of his or her interests from other users' defined conceptual spaces. With Telme, we used the same method because it can be used to infer knowledge. Although Telme considers only just-in-time adaptation to the current context, the method is flexible because conceptual spaces simply consist of keywords and weights, and the weights can be easily adjusted according to the inferred context.

References

1. K. Nagao, Y. Shirai, and K. Squire, "Semantic Annotation and Transcoding: Making Web Content More Accessible," *IEEE Multimedia*, vol. 8, no. 2, Apr.–June 2001, pp. 69–81.

2. B Rhodes, "The Wearable Remembrance Agent: A System for Augmented Memory," *Personal Technologies*, vol. 1, no. 4, Mar. 1997, pp. 218–224.

3. M.S. Ackerman and T.W. Malone, "Answer Garden: A Tool for Growing Organizational Memory," *Proc. ACM Conf. Office Information Systems*, ACM Press, New York, 1990, pp. 31–39.

4. K. Sumi, "Intelligent Tool for Facilitating Creative Communication," *Dynamic Knowledge Interaction*, Toyoaki Nishida, ed., CRC Press, Boca Raton, Fla., 2000.

ing, and that it provides accurate explanations. A negative comment was that the system might not be as useful as a chat system, wherein users can question each other directly during the conversation.

- *Interview support for people from different fields*. All the subjects agreed that Telme would be useful in this area (29 percent checked "5" and 47 percent checked "4"). Among the positive comments was that the system makes it convenient for the user to learn a word's meaning in the speaker's field, particularly when the meaning might vary across fields. Another positive comment was that a user can learn a word's meaning by questioning the system without interrupting the conversation, noting that this would be especially helpful when the user was the only person in a group who did not know a word. A negative comment was that it is impolite to use a computer when interacting with other people.
- *Real-time conferencing.* 95 percent of the subjects agreed that Telme would be useful as a real-time conference support for meetings between multiple people (24 percent checked "5" and 52 percent checked "4"). Among the positive comments was that the system gives the user an opportunity to quietly analyze unfamiliar words without interrupting the conversation. Another subject said that it might also be useful for checking unfamiliar words when attending a lecture. A negative comment was that people might stop listening to one another and become overly reliant on the machine.
- *Broadcast support:* All the subjects agreed that Telme is useful as a broadcast support system (76 percent checked "5" and 19 percent checked "4"). Among the positive comments were that it would be convenient to immediately understand unfamiliar political terms, and that the system might be useful for constructing knowledge and checking concepts related to the current topic.

Given these results, we conclude that Telme leaves a good impression for any communication style. In the opinions of users—who were merely considering different communication modes—Telme seems more useful as a computer-chat system than as an interview-support system, more useful as a conference system than an interview support system, and generally useful as a broadcast support system. In analyzing user comments, we found that they considered the system most effective when they could not question others directly—for example, when the user might be the only one who didn't understand a word or didn't want to interrupt others to ask a question.

Using Telme, even novice users can easily grasp unfamiliar things or avoid misunderstandings by glancing at the explanations. Telme's assistant information provides users with useful knowledge and a deeper understanding of the topic. As a real-time communication-support system, Telme is relatively fast and easy, and operates without interrupting the flow of conversation.

Because Telme is a customizable medium, it can help overcome the differences in knowledge among people. Personalization according to context will become increasingly important as human networks expand and various kinds of new media appear in the near future. Although we can't foresee what kinds of new communications media will emerge, we believe Telme will be effective for real-time and global communications between people from different fields, cultures, and languages.

In the current version of Telme, we collected most of the information manually, which took much effort. In addition, creating and updating the knowledge base and conceptual spaces were costly endeavors. We need a mechanism to enable inexpensive information collection. Developing such a mechanism will require further improvements in document-tagging technologies that deal with semantic structures.

References

1. T. Takezawa et al., "A Japanese-to-English Speech Translation System: ATR-MATRIX," *Proc. Int'l Conf. Spoken Language Processing* (ICSLP '98), Causal Productions, Adelaide, Australia, 1998, pp. 957–960.

2. A. Flammer, "Towards a Theory of Question Asking," *Psychological Research*, vol. 43, 1981, pp. 407–420.

3. N. Miyake and D.A. Norman, "To Ask a Question, One Must Know Enough to Know What Is Not Known," *J. Verbal Learning and Verbal Behavior*, vol. 18, 1979, pp. 357–364.

4. R.A. Finke, T.B. Ward, and S.M. Smith, *Creative Cognition*, MIT Press, Cambridge, Mass., 1992.

5. M. Polanyi, *The Tacit Dimension*, Routledge & Kegan Paul Ltd., London, 1966.

6. Y. Sumi, K. Hori, and S. Ohsuga, "Computer-Aided Thinking by Mapping Text-Objects into Metric Spaces," *Artificial Intelligence*, vol. 91, no. 1, Mar. 1997, pp. 71–84.

The Authors

Kaoru Sumi is a research associate at Osaka University's Institute of Scientific and Industrial Research. Her research interests are knowledge engineering, communication support systems, user modeling, and human–computer interaction. She received her BS in physics from the Science University of Tokyo, her MS in systems management from the University of Tsukuba, and her PhD in engineering from the University of Tokyo. Contact her at the Institute of Scientific and Industrial Research, Osaka University, 8-1 Mihogaoka, Ibaraki, Osaka 567-0047, Japan; kaoru.sumi@acm.org; www.ei.sanken.osaka-u.ac.jp/members/kaoru/index.html.

Toyoaki Nishida is a professor in the Department of Information and Communication Engineering, School of Information Science and Technology, University of Tokyo, and leader of the Breakthrough 21 Nishida Project at Japan's Communications Research Laboratory. The project aims to understand and assist networked communities. His general research interests are in artificial intelligence; his current research focuses on community computing and support systems, including knowledge sharing, knowledge media, and agent technology. He is an area editor in intelligent systems for *New Generation Computing* and on the editorial boards of *Autonomous Agents and Multiagent Systems*, *Knowledge and Information Systems*, and the *International Journal of Knowledge-Based Intelligent Engineering Systems*. He received his BE, ME, and doctorate of engineering from Kyoto University. Contact him at the Dept. of Information and Communication Eng., Univ. of Tokyo, 7-3-1 Hongo, Bunkyo-ku, Tokyo 113-8656, Japan.