論文 / 著書情報
Article / Book Information

Title	Introduction of the 21st century COE program Framework for systematization and application of large-scale knowledge resources
Author	Sadaoki Furui
Journal/Book name	SCIS & ISIS 2004, Vol. , No. , pp. THP-5-2
発行日 / Issue date	2004, 9

Introduction of the 21st Century COE Program "Framework for Systematization and Application of Large-scale Knowledge Resources"

Sadaoki Furui

Department of Computer Science, Tokyo Institute of Technology
2-12-1 Ookayama, Meguro-ku, Tokyo, 152-8552 Japan
E-mail: furui@cs.titech.ac.jp

Abstract

This paper describes the new five-year COE program "Framework for Systematization and Application of Large-scale Knowledge Resources," that has recently been launched at Tokyo Institute of Technology. The project is conducting a wide range of interdisciplinary research combining humanities and technology to create a framework for the systematization and the application of large-scale knowledge resources in electronic forms. Spontaneous speech, written language, materials for e-learning and multimedia teaching, classical literature and historical documents, as well as information on cultural properties are just some examples of the real knowledge resources being targeted by the project. These resources will be systematized according to their respective semantic structures. Pioneering new academic disciplines and educating knowledge resource researchers are also objectives of the project. Large-scale systems for computation and information storage, as well as retrieval, have already been installed to support this research and education.

1. Introduction

To function at optimal levels of efficiency, the 21st century, which is being called the century of knowledge resources, will require the construction and utilization of large-scale knowledge resources in every domain of research, education and daily life. The term "knowledge" is related to "information" and "data". A framework clarifying these concepts and their differences is summarized in Table 1. Knowledge is the representation of regularity and general rules in observing content that is used in interpreting information, solving problems, and creating new information through reasoning. Knowledge is the comprehensive integration of information that has been verified as valid for a specific topic, and is, therefore, a more significant entity than information. In contrast, data is a less significant entity than information. Unless a person possesses the appropriate knowledge, they will not be able to process available data to obtain relevant information. The relationships between the concepts of data, information and knowledge are shown in Fig. 1.

Table 1 – Differences between data, information and knowledge

Knowledge	**Basis for the creation and interpretation of information**
Information	**Entities describing things and events in the world and claims about their properties**
Data	**Physical information that forms the basis in obtaining abstract information**

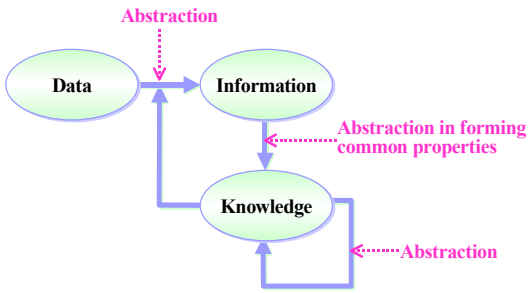

Fig. 1 – Relationships between data, information and knowledge.

As a knowledge resource is the large-scale accumulation of usable knowledge combined with meta-knowledge, it represents a much wider concept than just mere content. While various individual knowledge bases already exist, being developed with inconsistent concepts and dispersed among various individuals, research organizations and research domains, these knowledge bases are usually not easy to manage, extend or utilize. For example, although various dictionaries are used within research organizations and projects, as basic tools for natural language processing and spoken language processing, it can be difficult to establish equivalent relationships between them. A serious obstacle for web-search systems is that a common concept can be represented in different ways, even with spelling variations of the same word. This is also a serious problem for question-answering (QA) systems.

In order to coordinate various kinds of knowledge resources, such as teaching materials, documents and videos, it is necessary to develop a sophisticated structure within which to integrate the meanings of constituent elements. This, in turn, requires research that encompasses various

related sciences and technologies. The concept for the COE program evolves out of this kind of awareness for what is needed for future progress.

2. Plans of the COE program

The program is conducting a wide range of interdisciplinary research, combining humanities and technology to develop frameworks for the systematization and the application of large-scale knowledge resources in electronic forms. To this aim, it is necessary not only to accumulate various knowledge resources, but also to create a wide variety of fundamental technologies. The program is been implemented with a team of 20 professors, as listed in Table 2.

Table 2 – Members of the COE program

Leader	S. Furui (Dept. Computer Science)	Speech
Sub-leader	N. Makoshi (GSIC)	Language resources
Dept. Human System Science	M. Nakagawa, H. Akama	Classical literature
Dept. Value & Decision Science	A. Tokosumi, K. Yamamuro	Semantics, Historical documents
GSIC	H. Yokota, S. Matsuoka, M. Mochizuki	Database, Grid comp., E-learning
Int. Student Center	K. Nishina	Foreign student education
Dept. Computer Science	H. Tanaka, M. Nakajima, M. Saeki, T. Tokuda, T. Tokunaga, T. Sato, K. Shinoda, N. Yonezaki, H. Kamei	Natural language, Image, Software, Web, Logic, Speech, Security, Data mining
Precision and Intelligence Lab.	M. Okumura	Summarization

GSIC: Global Scientific Information and Computing Center

Multimedia documents, spontaneous speech, written language, materials for e-learning and multimedia teaching, classical literature and historical documents are just some examples of the real knowledge resources being targeted by the project. The pioneering of new academic disciplines and the education of knowledge resource researchers are also objectives of the program. Large-scale systems for computation and information storage, as well as retrieval, have already been installed to support this research and education [1].

Figure 2 illustrates the targets of the COE program, namely, the construction of frameworks for the standardized systematization and the application of large-scale knowledge resources to the various domains mentioned above. Research on semantic structures with which to integrate and systematize various forms of knowledge resource and on the construction of systematized large-scale knowledge resources is being conducted using a diverse range of technologies, such as speech and language analysis, databases, semantics, and dictionaries. Technologies to develop applications of the large-scale knowledge resources include data mining, information retrieval, visualization, WWW, and security, as well as mobile communication and networking. Through these processes, the program aims to establish a basic structure that will make it easy for the general population to construct and systematically use knowledge resources.

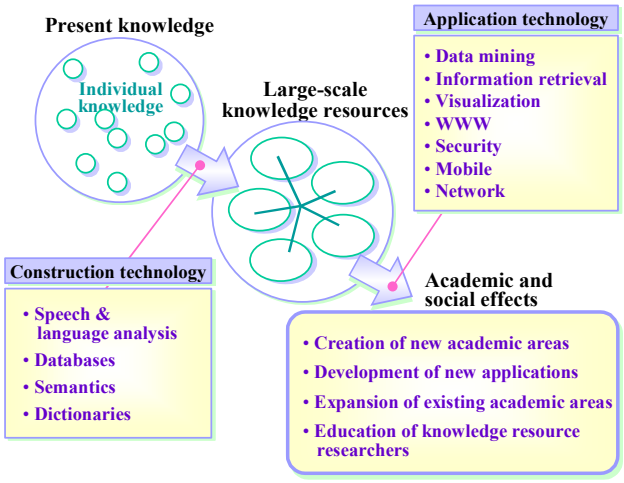

Fig. 2 – Targets of the COE program.

The program will also establish an open research center for the storage and utilization of knowledge resources and for the creation of new knowledge resources. Through this activity, the program is seeking to extend existing academic areas, develop new applications, and to create new sciences in interdisciplinary areas. The new knowledge gained from such activities is expected to accelerate new scientific developments within pedagogy, history, philology, and archaeology, as well as speech science and linguistics. Knowledge resources being targeted by the program also include classical literature, historical materials and classical art, which are directly related to Japanese traditions, culture and language. Large-scale knowledge resources will be created by the activities of the program, and through their storage and systemization, new insights, unobservable at the level of more restricted individual knowledge, will be gained.

The program will also contribute to the construction of materials for e-learning and the autonomous learning of second languages. Outcomes from the program will be applicable to various information processing systems, including human-computer interaction using automatic speech recognition (ASR) and the automatic transcription of lectures, presentations and news broadcasts.

Various IT technologies, such as large-scale computation, multimedia processing, database systems, knowledge-base systems, XML, data mining, and information retrieval, as well as visualization, security, mobile communication and networking, are indispensable for the construction and full utilization of large-scale knowledge resources. The activities of the program will, in turn, make major contributions to the advancement of these IT technologies.

3. Multimedia/multimodal archiving and retrieval

One of the core research projects within the COE program is multimedia/multimodal archiving and retrieval using various combinations of media technologies, such as speech, image,

video, and text processing technology (see Fig. 3). A key issue for this research into the development of useful systems is how to integrate various knowledge sources according to their semantic representations. Another interesting research topic is the problem of building models of actions and gestures. As already reported [2], multi-modal speech recognition, combining lip movements with audio signals, is an effective method of improving recognition accuracy for noisy speech recognition.

Research targets also include topic detection and extraction. Segmenting transcriptions into thematically distinct stories and categorizing the stories by topic can, for example, increase readability and comprehensibility. In many studies of this kind, a text or speech would be classified under a topic or category based on the relevance of keywords to a set of categories or topics. Thus, a document would be classified under the category with the highest relevance score, after summing all relevance scores to the various categories assigned to all keywords. In our method [3], however, categories are replaced with topic words and multiple topic words of high relevance to a document are extracted. So, this method involves two lists of words—words in the documents and topic words—where the mapping of document words to topic words can be estimated statistically from training data. As the two lists are different, the topic extraction model can extract appropriate topic words, even when they do not appear in the document. The method has been applied to topic extraction for speech from Japanese news broadcasts. In order to compensate for poorer performance levels due to speech recognition errors, we have incorporated a likelihood weighting mechanism among a set of N-best candidates, and have achieved practical levels of accuracy.

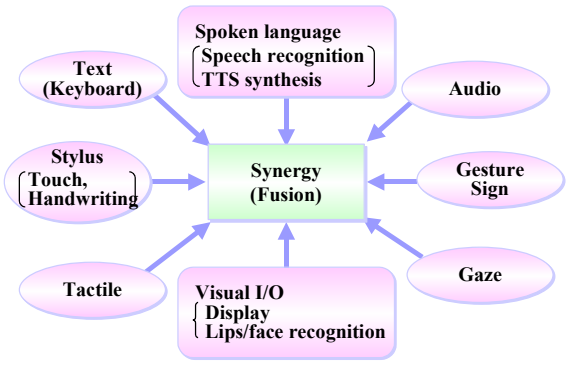

Fig. 3 - Multimedia content technology.

4. Three research groups

The research activities of the program are organized under three major groups. To foster close cooperation between the groups, each member of the program belongs to at least one group.

4.1 The linguistic and document knowledge resource group

This group consists of three subgroups, as shown in Fig. 4 overleaf. The interdisciplinary collaborations among the subgroups are creating a new research field within the humanities; research on large-scale linguistic and document resources based on innovative information technology concepts and methodology.

(1) Subgroup for the intelligent acquisition of knowledge resources

Several projects in this subgroup are working towards the automatic and intelligent acquisition of large-scale linguistic resources from the Internet. For instance, the project for the automatic collection of example sentences is utilizing a user-agent program (also known as a crawler or spider) that can search hyper-linked sites all over the Internet and collect natural language sentences, in Japanese and other languages, in the order of several hundred thousand sentences [4]. When augmented with annotation concerning syntactic structures and semantic content, the collected sentences will provide a firm basis for machine translation systems and computer-assisted language learning systems. While another project targeting the construction of a document archive has proposed a distributed input and evaluation mechanism that enables people on the Internet to collaborate in building a large-scale distributed archive of literary and academic texts [5]. This subgroup is also building a large-scale Japanese CFG for syntactic parsing [6].

(2) Subgroup for the development of analysis system engines

This subgroup is working to introduce a probabilistic model of language within the humanities and, by combining natural language processing with philological science, is advancing a new research field that can be called computational humanities. A set of web-based software tools are currently being developed that will facilitate latent context detection and precise matching of variant texts, by extending text mining algorithms to lexical co-occurrence. Specifically, two tools are now being developed: a) Tele-COEX which implements a number of different functions to set up a kind of window centered on keywords, and b) Tele-Synopsis, which has been designed to gather information relating to word usage under certain conditions and to further enhance a statistical approach to tracing the origin and genealogy of parallel and variant texts [7]. Research on the Gospels, newspaper headlines, and personal pronouns in French has already been carried out using these analytical engines.

(3) Subgroup for the development of a search engine capable of handling metaphoric key words

The aim of this subgroup is to develop, based on the statistical analysis of language corpus data, a search method that can accept metaphoric expressions in key words. For example, similes, such as ‘a dog like a cloud’, are being used

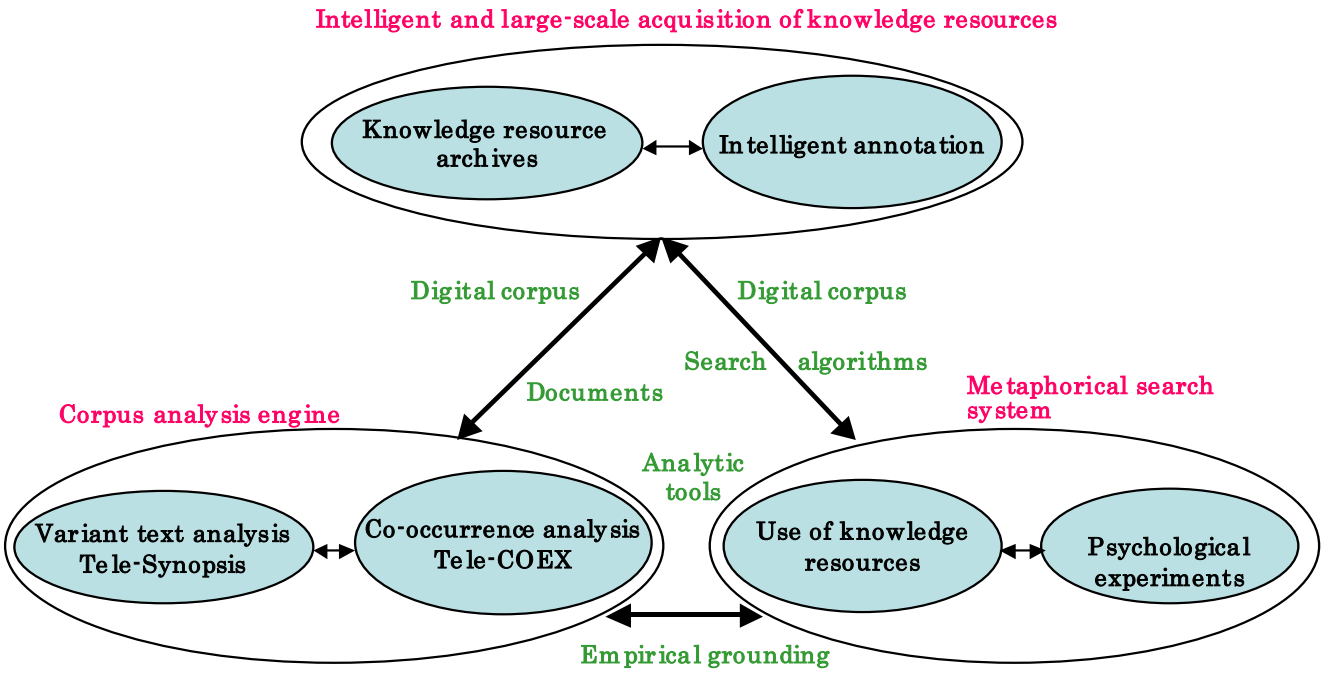

Fig. 4. - Collaborative structure of the three subgroups within the linguistic and document knowledge resource group.

in the present version of the system as key words in searching for information, such as images of 'a white and fluffy dog'. This research is working to develop a system that includes human-like metaphorical search methods. Initial efforts have focused on constructing a simulation model of human comprehension and generation processes for metaphoric expressions, especially, similes of the form 'A like B'. Psychological experimentation will also be necessary in developing these models. Experiments drawing on the large-scale data will be conducted to investigate the processes involved in the human comprehension and generation of metaphoric expressions, for comparison of experimental results with those of statistical analysis.

4.2 The education and learning group

This group consists of interdisciplinary experts from speech processing, multimedia data processing and integration, natural language processing, computing technology for distant education, and Japanese language teaching (second language acquisition). The group is conducting research to create Web-based education systems and e-learning content. The research includes teaching materials for Japanese language and culture, teaching materials for other foreign languages, and e-learning content, including recorded Tokyo Tech lectures, as well as a system for retrieving a large number of technical papers from both domestic and international resources. Three practical systems, ASUNARO, UPRISE and PRESRI, have already been built.

(1) ASUNARO

ASUNARO is a support system for Japanese language learners, designed to morphologically parse sentences and retrieve target words in multiple languages, including Chinese, Thai, Indonesian and Malay. The structure of each sentence is displayed to enhance visual learning. The system, which is accessible at http://hinoki.ryu.titech.ac.jp/, is being used by both local and overseas learners. The system will be extended for the learning of other languages, such as English and Chinese.

(2) PRESRI (Paper Retrieval System using Reference Information) system

PRESRI is a system that effectively extracts relationships between technical papers and displays research histories, using paper referencing information between related papers [8]. When a reference item is input at the bibliographic information retrieval screen, this system graphically displays retrieved information, such as reference papers and referred papers. Already available on the Web, http://presri.pi.titech.ac.jp:8000/, this is a useful tool for writing review papers.

(3) UPRISE (Unified Presentation Content Retrieved by Intelligent Search Engine)

UPRISE is a system that effectively retrieves presentation materials, by drawing on an integrated database of videos for presentations and information about the circumstances in which the materials were used [9]. Figure 5 presents an example screen displaying retrieval results, combining video and presentation slides. The system will be an especially useful tool for students taking lectures at remote locations, including foreign language classes, allowing them to review materials later.

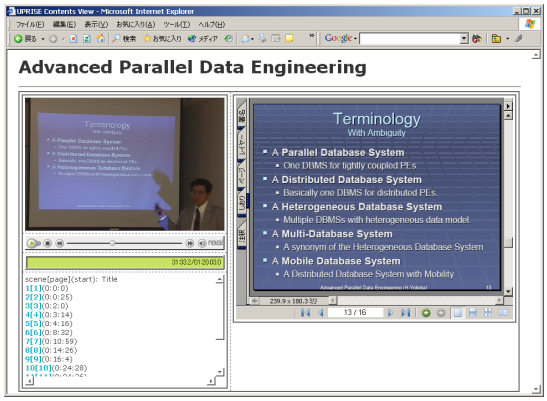

Fig. 5 – Example screen for the UPRISE system

The education and learning group aims to integrate and extend the systems outlined above, and to provide other multi-media systems for science and technology students in Japan and other countries. These systems will be further improved through technical support from specialists of spoken language processing and remote lecture technology. In developing systems to provide Tokyo Tech educational content to both local and overseas learners, the group will also contribute to the education of large numbers of students for international leadership.

4.3 The audio, visual, and sensory knowledge resource group

This group is studying the systemization and the application of audio, visual, and sensory knowledge resources. Because of problems in directly utilizing the raw signals of audio and visual resources, it is difficult to extract useful information from them. Various pattern recognition techniques, including speech and image recognition, have been developed to tackle this problem. Using these techniques, the group is conducting research on search and summarization for multimedia content.

Audio and visual information can sometimes be more important than text information, as they contain information that is extremely difficult, if not impossible, to symbolize. One example is the prosodic information, representing speaker intentions or emotions, contained in spontaneous speech. This research group is developing a large speech corpus to study how to extract and systematize this kind of information.

A particularly interesting topic, that has not previously been investigated, is whether we can extract audio and visual information from textual information, to create new forms of knowledge. Exploring this approach, the group is applying acoustic phonetics to the analysis of classical literature.

Three major research topics of the group are outlined briefly below.

(1) Search and summarization of multimedia content

Nowadays, the amount of multimedia content, such as TV programs, is rapidly increasing. Accordingly, it is becoming increasing important to develop search and summarization methods to handle their content. Techniques currently under investigation include those of speech recognition for closed-captioning of news broadcasts and recognition methods for human motion and gestures in daily life. Based on these techniques, novel search and summarization techniques, capable of simultaneously utilizing both audio and visual information, are now being investigated, including joint research with NHK Science & Technical Research Laboratories. International collaborative research on speech summarization has also started.

(2) Corpus of Spontaneous Japanese

Prior to the launch of the current COE program, the Corpus of Spontaneous Japanese (CSJ), containing about 7.5 million words and rich annotations, was built (see Sec. 5 for more details of the CSJ). The CSJ is now being widely used for the purposes of spontaneous speech recognition and summarization, and the results are attracting much interest from domestic and international researchers.

Although CSJ is one of the best databases of spontaneous speech in the world in terms of both its quality and quantity, there are clear limitations: most notable is the fact that monologues currently represent more than 90% of the CSJ speech. Accordingly, it is necessary to enlarge the range of speech types, so that the corpus will cover the whole spectrum of spoken language. The aims of this research group include establishing an effective way of compiling a “balanced” corpus of spoken language. Such a corpus will contribute greatly to the development of speech summarization systems and the linguistic study of spoken language.

(3) Analysis of Japanese classical literature from a spoken-language perspective

Passed down as a treasured Japanese legacy, a large body of classical literature is now publicly available as digitized cultural resources from the National Institute of Japanese Literature and other organizations. While many classical works have been studied as forms of written literature, some were originally created as oral performance or narration, and only subsequently transcribed in writing. However, over the centuries, these performance works may have influenced the written style of Japanese literal culture.

The research group is carrying out statistical studies of the underlying properties of classical written texts from a spoken-language perspective, and is examining the characteristics and records of spoken and written Japanese. As an initial contrastive analysis, the phonemic and prosodic characteristics of the Tale of Genji and the Tale of Heike are now being analyzed. This study is a typical example of how the COE program is bring humanities and technology together.

5. Speech recognition for the construction of knowledge resources

As speech is the most natural and effective method of communication between human beings, various important speech documents are produced everyday, including lectures, presentations, meeting records and news broadcasts. However, if such speech documents are simply recorded as an audio signal, it is not easy to quickly review, retrieve, selectively disseminate, and reuse them. The automatic transcription of speech using speech recognition technology is, therefore, a crucial aspect in the creation of knowledge resources from speech. Although high levels of recognition accuracy can be obtained easily for speech that is read from a text, such as the speech of anchor persons during news broadcasts, the technology for recognizing spontaneous speech is still rather limited [10].

Spontaneous speech and speech based on a written script are very different, both acoustically and linguistically [11]. In order to improve recognition performance for spontaneous speech, it is essential to build acoustic and language models for spontaneous speech. The desire of current statistical techniques for model training material is well captured in the aphorism "there's no data like more data." Large structured collections of speech and text are essential in order to make advances in speech recognition. However, establishing a good speech database for speech recognition is not a simple task. There are basically two broad issues that require careful consideration; the first relates to content and its labeling and annotation, while the second is the collection mechanism. Labeling and annotation for spontaneous speech can easily become unmanageable. For example, how should we annotate speech repairs and partial words, how can phonetic transcribers reach a consensus concerning acoustic-phonetic labels when there is ambiguity, and how do we represent semantic notions? Errors in labeling and annotation reduce system performance. Thus, a major concern is just how to ensure the quality of annotated results. Research on automating or creating tools to assist in the verification procedure is, in itself, a challenging topic.

It was in this context that a 5-year Science and Technology Agency Priority Program entitled "Spontaneous Speech: Corpus and Processing Technology" was conducted in Japan during the period 1999 to 2003, which focused on three major themes [12][13].

1) Building a large-scale spontaneous speech corpus. The Corpus of Spontaneous Japanese (CSJ) consists of roughly 7M words with a total speech time of 650 hours, based primarily on recorded monologues, such as lectures, presentations and news commentaries. The recordings have been manually given orthographic and phonetic transcriptions. One-tenth of the utterances, referred to as the *Core*, have been manually tagged and used in training a morphological analysis and part-of-speech (POS) tagging program with which to automatically analyze the entire 700-hours of utterances. The *Core* has also been tagged with para-linguistic information including intonation (see Figs. 6 and 7).

2) Acoustic and linguistic modeling for spontaneous speech recognition and comprehension, employing both the linguistic and para-linguistic information in speech.

3) Investigating spontaneous speech recognition and summarization technology [14][15].

By constructing acoustic and linguistic models with the CSJ, recognition errors for spontaneous presentations have been reduced to roughly half compared to the levels obtained with models constructed based on read speech and written texts [13][16]. Technology created under the Spontaneous Speech project will have a wide range of applications, such as the indexing of speech data (news broadcasts, etc.) for information extraction and retrieval, transcription of lectures, preparing the minutes of meetings, closed captioning, and aids for the handicapped. Major challenges for spontaneous speech recognition still to be overcome, however, include how to handle filled pauses, repairs, hesitations, repetitions, partial words, and disfluencies.

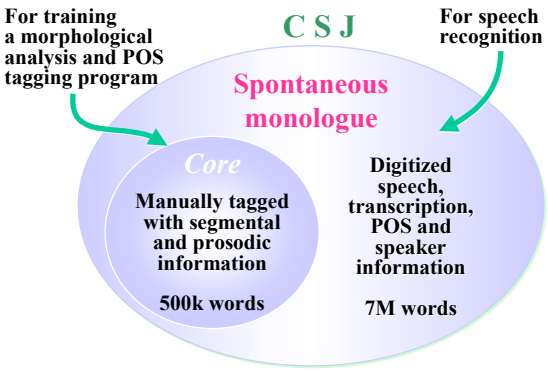

Fig. 6 – Overall design of the Corpus of Spontaneous Japanese (CSJ).

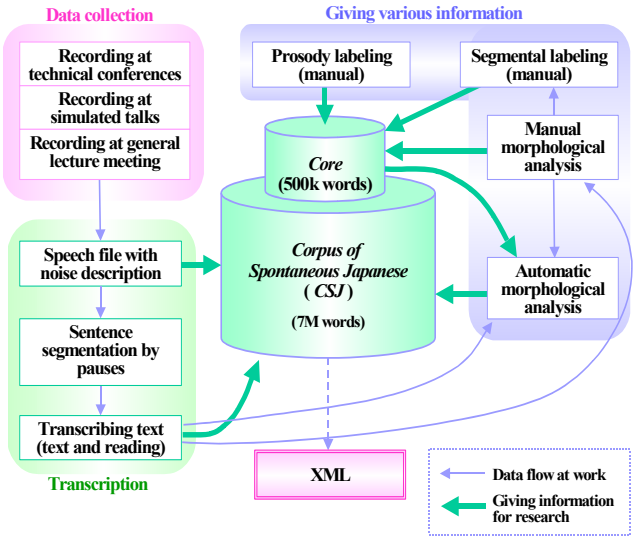

Fig. 7 – CSJ corpus construction.

The CSJ will be stored in XML format within a large-scale database system developed by the COE program, so that the general population can access and use it easily for research purposes. As recognition accuracy for spontaneous speech is still rather low, the data collection work will continue under the COE program, in order to increase the coverage of different varieties of spontaneous speech within the corpus.

6. Knowledge resources for speech recognition

In automatic speech recognition, it is extremely important to create modeling methodologies and collection schemes for associated data that have the capacity to be applied generally to many different tasks. One criterion for maximizing performance is to ensure that as far as possible the data accurately reflects the actual environment. This calls for a database collection plan that is consistent with the task. Data collection will quickly become totally unmanageable, if the system designer has to undertake separate collections of data for each and every application that is being developed. It is, therefore, desirable to design a task-independent data set and a modeling method, which can both deliver reasonable performance on first use and cope easily with in-field trials for subsequent revisions as more task-dependent data becomes available. Research in this area can yield benefits in terms of reduced costs for application development.

Hybrid language models (LMs) are commonly used, in which several LMs specific to a particular topic or style of language are generated and stored. In general, an application domain can be characterized by a set of subtasks, where each task is, in turn, characterized by a topic or a set of topics. A particular collection of documents, such as the articles of a newspaper, can be categorized into specific text clusters according to a given set of topics. Newspaper articles are usually manually classified under different genres, such as news, sports, and movies. Based on this information, it is possible to derive text clusters that are topic specific, and then, for each cluster, generate a topic specific LM. Automatic text clustering for topic assignment can also be used. Probability estimates from component LMs can be interpolated to produce an overall probability, where interpolation weights can be chosen to reflect the topic or style of language currently being recognized. Difficult problems arise, however, from that the facts that new topics are always created and that different representations are frequently used for the same topic. Accordingly, one crucial issue is how to dynamically model the set of topics.

Voice QA systems that respond to queries given by voice to retrieve exact answers or articles from a wide range of domains are important and useful applications of speech recognition technology. One of the most pressing research issues for such systems, in terms of usability, is just how to reduce speech recognition errors for the query utterances and their effects on retrieved answers. We have been investigating a method to generate effective domain-dependent language models for voice query recognition and a new dialogue strategy [17]. In the proposed interactive dialogue strategy using a multimodal user interface, users are requested to indicate correct keywords, with inappropriate keywords being automatically replaced by the most probable keywords from an N-best list based on domain-dependent word co-occurrence probabilities. Word co-occurrence probabilities are used as a long-distance language model to augment the trigrams used in voice query recognition. A preliminary QA system using voice input and a graphic user interface has been implemented using NTT's SAIQA open-domain QA system.

Human speech recognition is a matching process, where incoming speech is matched to various forms of existing knowledge, as shown in Fig. 8 [18]. In order to make significant advances on current levels of speech recognition performance, there is a pressing need to create new paradigms and to make more extensive use of the various sources of knowledge involved in the recognition process. Future research on speech summarization will include investigations of other kinds of information/features for important unit extraction. Accordingly, key issues for this research will be how to systematize and utilize various kinds of knowledge, such as domain and topic-related knowledge [3], context and speaker identity.

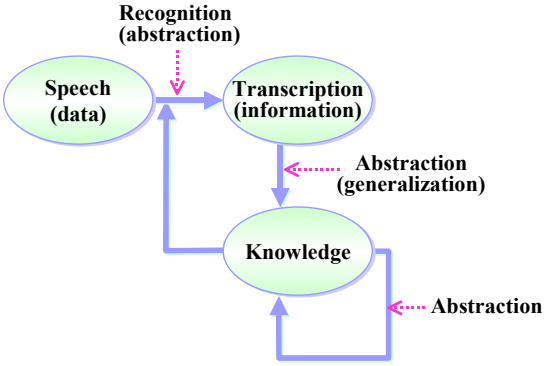

Fig. 8 - Knowledge resources for speech recognition.

7. Summary

This paper has presented an overview of the new five-year COE program "Framework for Systematization and Application of Large-scale Knowledge Resources" that has recently been initiated at the Tokyo Institute of Technology. The expected outcomes of the COE program are summarized below.

(a) Systematization of various knowledge resources,
(b) Construction of resources and application technology for e-learning and second language learning,
(c) Analysis of mutual relationships and systematization of textbooks, historical documents, and classical literature,
(d) Fostering intelligent creativity in knowledge research,
(e) Diverse and efficient education, and the creation of new inter/trans-disciplinary academic studies,
(f) Development of various kinds of information processing technology,
(g) Foundation of a large-scale knowledge resource research center, and
(h) Education of knowledge resource researchers with extensive interdisciplinary knowledge, high academic abilities, liberal arts awareness, logical thinking abilities, and international understanding.

We will strive not only to accumulate and systematize various knowledge resources, but will also work to establish the fundamentals of interdisciplinary science and technology.

References

[1] H. Yokota, "An information storage system for large-scale knowledge resources," Proc. Int. Symp. Large-scale Knowledge Resources (LKR 2004), pp. 87-90 (2004)

[2] S.Furui, "Robust methods in automatic speech recognition and understanding," Proc. EUROSPEECH, Geneva, pp.1993-1998 (2003)

[3] K. Ohtsuki, T. Matsuoka, S. Matsunaga and S. Furui, "Topic extraction based on continuous speech recognition in broadcast news speech," IEICE Trans. Inf. & Syst., pp. 1138-1144 (2002)

[4] M. Taguchi, K. Jamroendararasame, K. Asami and T. Tokuda, "Comparison of two approaches for automatic construction of Web applications: Annotation approach and diagram approach," Proc. Int. Conf. on Web Engineering, Munich (2004)

[5] N. Matsumoto and A. Tokosumi, "A distributed acquisition and evaluation system for literary text archive maintenance," Proc. Distributed and Collaborative Knowledge Capture Workshop, Sanibel, pp.14-18 (2003)

[6] T. Noro, T. Hashimoto, T. Tokunaga and H. Tanaka, "Building a large-scale Japanese CFG for syntactic parsing," Proc. 4th Workshop on Asian Language Resources, pp.71-78 (2004)

[7] M. Miyake, H. Akama, M. Sato, M. Nakagawa and N. Makoshi, "Tele-synopsis for biblical research: Development of NLP based synoptic software for text analysis as a mediator of educational technology and knowledge discovery," Proc. Conf. on Educational Technology in Cultural Context (ETCC), Finland (2004)

[8] H. Nanba, T. Abekawa, M. Okumura and S. Saito, "Bilingual PRESRI – Integration of multiple research paper databases," Proc. RIAO, pp. 195-211 (2004)

[9] H. Yokota, T. Kobayashi, T. Muraki, and S. Naoi, "UPRISE: Unified Presentation Slide Retrieval by Impression Search Engine," IEICE Trans. Inf. & Syst., pp. 397-406 (2004)

[10] B.-H. Juang and S. Furui, "Automatic recognition and understanding of spoken language – A first step towards natural human-machine communication", Proc. IEEE, 88, 8, pp. 1142-1165 (2000)

[11] S. Furui, "Toward spontaneous speech recognition and understanding," in Pattern Recognition in Speech and Language Processing, W. Chou and B.-H. Juang (Ed.), CRC Press, pp. 191-227 (2003)

[12] K. Maekawa, H. Koiso, S. Furui, H. Isahara, "Spontaneous speech corpus of Japanese," Proc. LREC2000, Athens, pp. 947-952 (2000)

[13] T. Shinozaki, C. Hori and S. Furui, "Towards automatic transcription of spontaneous presentations," Proc. EUROSPEECH, Aalborg, pp.491-494 (2001)

[14] C. Hori, S. Furui, R. Malkin, H. Yu and A. Waibel, "A statistical approach to automatic speech summarization," EURASIP Journal on Applied Signal Processing, pp. 128-139 (2003)

[15] T. Kikuchi, S. Furui and C. Hori, "Two-stage automatic speech summarization by sentence extraction and compaction," Proc. ISCA-IEEE Workshop on Spontaneous Speech Processing and Recognition, Tokyo, TAP10 (2003)

[16] S. Furui, "Recent advances in spontaneous speech recognition and understanding," Proc. ISCA-IEEE Workshop on Spontaneous Speech Processing and Recognition, Tokyo, MMO1 (2003)

[17] D. Kim, S. Furui and H. Isozaki, "Language models and dialogue strategy for a voice QA system," Proc. ICA2004, Fr3.H.4 (2004)

[18] S. Furui, "Overview of the 21st Century COE Program "Framework for systematization and application of large-scale knowledge resources," Proc. Int. Symp. Large-scale Knowledge Resources (LKR 2004), pp. 1-8 (2004)