ISSN 0910-6707

1992

Japan Computer Quarterly

見本

Japan Information Processing

Development Center

Real World Computing
&
Related Technologies

No. 89

Japan Computer Quarterly

1992

Japan Computer Quarterly (JCQ) is published quarterly by the Japan Information Processing Development Center (JIPDEC), Kikai Shinko Kaikan Bldg., 3-5-8 Shibakoen, Minato-ku, Tokyo 105 Japan.

Publisher: Eiji Kageyama, President
Editor: Yuji Yamadori, Director Research & International Affairs

JIPDEC is a non-profit organization founded in 1967 with the support of the Ministry of International Trade and Industry, the Ministry of Posts and Telecommunications and related industry circles for the purpose of promoting information processing and the information processing industry in Japan.
JCQ, formerly called the *JIPDEC* Report, was first published in September, 1970 and is prepared with the assistance of the Japan Keirin Association through its Machine Industry Promotion Funds.

NOTE: The opinions expressed by the various contributors to the Japan Computer Quarterly do not necessarily reflect those views held by JIPDEC.

Translated and Printed by The Translation Institute of Technology, Science & Culture
Printed in Japan, March, 1992

CONTENTS

No. 89

From the Editor

With the progress of technology, the role of computers has become increasingly important nowadays. It is said that present society could not exist without computers. However, the principle of the computer essentially has not changed compared to 40 years ago when mechanical computing technology was first accomplished. In addition, the principle of the computer basically included the difficult problem that it is not suited to large volume information processing as is required today. The principle is derived from the basic principle of computers today, which includes the fact that processing procedures must be completely stored as computer operations in advance and that operations are performed on a sequential basis.

Conventional computers are called von Neumann computers. The new types of computers, based on new non-von Neumann computer architecture, have now been introduced. They include fifth-generation computers such as parallel processing/inference machines, but in recent years neuro-computers have attracted a substantial amount of public attention.

JIPDEC once did basic research on new generation computers for about a three year period. When we submitted the results of this research to the government, MITI (Ministry of International Trade and Industry) reported that the theme of new generation computers was extremely important for the advancement of Japan's future information technologies. MITI formed a special research and development organization called ICOT (Institute for New Generation Computer Technology) and began R&D at this institute as a large-scale national project. The project began in 1982 as a ten-year plan, and has now entered its final phase. We have many expectations for the future advancement of research based on the results ICOT has achieved to date.

The human brain has about 14 billion neurons, and these form a network (the neural network). Intellectual activities of humans are said to be performed through the interchange of signals between these neurons. Therefore, it naturally follows that research and development on neuron computers is being carried out by a variety of research organizations and manufacturers. The research is based on the concept that a neuron computer can be realized if each neuron is equated with a processor and the network is realized technologically. Neuron computers have enormous potential because of their learning capabilities, which permit large-scale improvement in program preparation work, and the acquisition of knowledge and rules that human beings may have difficulty in acquiring. In this respect, the early development of neuron computers is clearly desirable.

It is said that human information processing consists of two functions: the logical/mathematical function and the intuitive/pattern recognition function. Conventional computer technology is designed solely for the realization of the former function. The computers which we can expect to see in the future are those which will possess the latter function; and such computers will have a greater affinity with human beings.

To bring about the realization of neuron computers, there are many tasks to be accomplished. The principal task is that of establishing "flexible information processing" with "super-parallel" and "super-distributed" techniques using optical technology.

The New Information Processing Technology (NIPT) project has also been in progress for the past few years, with full support from MITI. A large number of scholars, researchers and specialists have participated in this project, which is aimed at the 21st century. As with the fifth-generation computer project, JIPDEC has been the administrative office for the NIPT project, while at the same time concentrating our efforts on basic research over the past three years. The basic research is already in the final phase. Based on the results of the research, the Japanese government will create a new organization. This organization will administer a new national project focused on innovative information processing technology, which seeks an optimum solution for the handling of incomplete information in a complex, real society. The watchword for the project will be RWC (Real World Computing). The government is about to start the research and development with cooperation from foreign countries.

This time, we will give an introduction to the R&D on various new information processing technologies that have been implemented in Japan. We will also provide an introduction to the technologies which are necessary for the realization of RWC. We hope this will be helpful to all of our readers.

Yuji Yamadori
Director
Research & International Affairs

Outline of Real World Computing Program

Kenji Torizuka
Technical Official
Machinery and Information Industries Bureau
Ministry of International Trade and Industry (MITI)

1. Outline

(1) Computers of conventional design have already been shown to be inconvenient tools for processing the incomplete or erroneous data that is so often encountered in today's society. However, ease in handling complex data will be a vital need that society must satisfy as it grows ever more reliant on sophisticated information in the next century. Thus techniques must be developed to process incomplete data within realistic time limits, and it will be essential to develop new and revolutionary processing techniques (Real World Computing) based on learning and recognition functions that require no expertise to perform data processing.

The present research and development drive will focus on fundamental, revolutionary technologies. Considering its global significance, cooperation from all countries will be called for as it progresses.

(2) Details of the R&D drive

① Planning and executing research activities
A new body, tentatively named the Real World Computing Partnership, will play the central role in implementing and executing the R&D projects. It will lay down R&D plans by drawing on fundamental theory, new system architectures, opto-electronics (optical computing), and other developing fields. The R&D activities run by the partnership will be based on these plans.

② Joint international studies
The research partnership will act as an intermediary in building a system to promote joint research or research cooperation with overseas institutions.

③ Creating the development environment
Linkages will be established between Japanese and overseas research organizations engaged in studies on Real World Computing Program, and computer networks will be established to help improve the efficiency of research and development activities.

(3) Expected results

Real World Computing, when realized, will make processing of large volumes of incomplete data possible within realistic time limits, and the processing will be done easily without requiring the intervention of experts.

This will widen the range of application of computers in such areas as early diagnosis of cancer, completion of telephone systems incorporating automatic translation functions, and measurement of the carbon dioxide balance in the global environment. All these are real problems that cannot be attended to easily as of today.

2. Background

(1) State-of-the-art technology

Thanks to advances in hardware and software, computer technology has reached an astounding level of development today. In handling regular calculations, documents (texts), and databases, and more recently in logical data processing involving inferences, computers already seem to have an edge on man. On the other hand, they are far less flexible in information processing involving intuition, so it will be essential to expand the range of available functions related to this.

Flexible data processing is now being demanded in all of today's hightech fields, as it has been found to be indispensable for sophisticated functions such as pattern or knowledge data processing, including for use with intelligent robots.

(2) Target technologies

The term "Real World Computing Paradigm" refers to the basic concepts underlying the research program. In short, it represents the judgement and problem solving capabilities needed for the kind of data processing that will become necessary in the next century to handle the jumble of vague, incomplete, and variable data that pervades real life, in its actually existing form. Pursuing the concepts above will make it a technical target to grope for a further expansion of information processing function with computers.

The following is needed to realize the Real World Computing possible:

① Theoretical base: "Flexible information processing" with flexibility and abilities of learning and adaptation to handle real-world information with incompleteness, ambiguity and changeability.

② Computational base: "Massively parallel and distributed information processing" with robustness capable of flexible information processing with limited amount of resources and within limited time.

3. Guidelines for promoting technical development

(1) Basic concepts

① Fundamental research oriented:
To establish the fundamental and generic technologies for the highly advanced information society in the 21st century.

② International contribution oriented:
To promote international cooperation and to report openly the results of research activities to the world.

(2) Basic promotional policy

① A flexible formation of research organization
The tasks to be attended to will be divided appropriately to propel the studies ahead successfully along their courses. Thus, a centralized laboratory will be taken for studies involving a high level of integration, symbolism, and sharing of resources. On the other hand, distributed

laboratories will be adopted for studies that are highly individualistic and elemental in nature. The two must be linked intimately and flexibly.

② Introduction of competitive principle
In the initial stage of research and development, mutually competing theories must be tried out (that is, various rival approaches must be given a chance to compete with each other for more effective study results). The results thereof will be evaluated at an intermediate stage, and on this basis, the actual problems to be researched further will be narrowed down.

③ Inter-disciplinary and international cooperation
Inter-disciplinary and international linkages will be necessary to meet the basic challenges of this R&D project. Thus, all possible efforts should be made to perform joint studies with the Electrotechnical Laboratory and university research institutions. At the same time, studies must be re-consigned to research institutions affiliated to universities in Japan or abroad.

④ Making research results public
Progress in the studies of the various research institutions will be reported in conferences held both in Japan and overseas. Also, symposiums and workshops will be held regularly. In this way the program will actively make its result public.

⑤ Establishing a research infrastructure
A high-speed network environment will be established and made to serve as the basis for global-scale distributed research efforts. This will make the plan for the execution system flexible and will also support the public release of study results.

4. R&D plans evaluation system

(1) Outline

Next, we will take up the system followed and the items covered in the fiscal 1991 feasibility studies. As Figure 1 shows, a workshop steering committee and a working group for system evaluation were formed under the feasibility study committee for new information processing technologies. These bodies have been evaluating the basic R&D plans and the basic elements that make up the research and development system. The workshop steering committee are sponsoring workshops on a global scale and are planning activities on the basis of technical discussions among Japanese and overseas researchers working in related fields.

(2) The feasibility study committee

The object of this committee is to determine guidelines for the following items. Each item is of importance for promoting the Real World Computing Program.

① Formulation of master plan for promotion of the Real World Computing Program

<Items to be covered in the master plan>
- Outline of the program
- Targets for technical development
- Methods, budget, and schedule for development activities
- Evaluation methods
- Evaluation yardsticks etc.

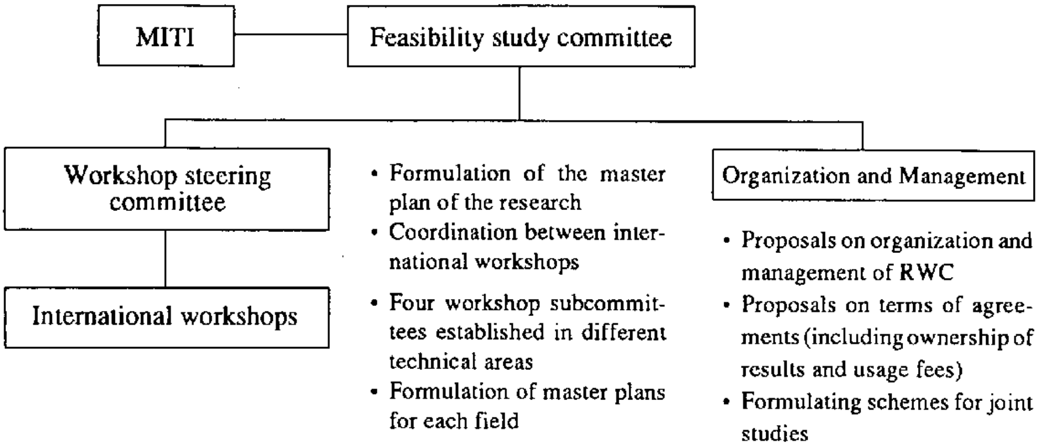

Figure 1 Organization of the feasibility study committee

② Specifying the basic items for the technical development system

<Main items>

Organizing and managing the main bodies in charge of technical developments

Items covered in agreements reached for consignment (including how the industrial property rights are to be handled)

Rules to apply to joint studies

(3) International workshops

① Object

Workshops relating to the basic research plans of Japanese and overseas researchers will be conducted according to the results of a two-year long survey starting in 1989. The respective basic plans will delineate the objects of the research undertaken, the research schedules, budget outlines (funding plans), methods of evaluation, the research system, and other relevant details.

The workshops will be split into four different areas as shown in Figure 2.

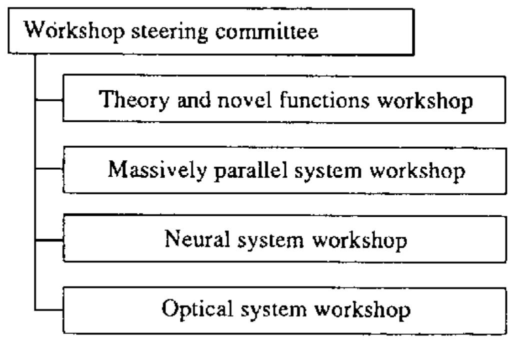

Figure 2 Workshop composition

② Scope of activities of the different workshops

<Theory and novel functions workshop>

This workshop will deal with the theoretical basis for real world computing and will investigate the new functions to form the basis for applications. Real world computing will make it possible to integrate a wide range of information — including incompleteness and errors — and handle it flexibly in the realtime mode. This investigation will help develop a concrete image for application systems. The main items to be investigated by this workshop are:

- Theoretical base:
  Constraint satisfaction
  Statistical theories, theory of multivariate analysis
  Theory of learning, applications, and self-organization

- Novel functions:
  Sophisticated, flexible recognition and understanding functions
  Flexible information bases and inference based on their use

- Problem solving:
  Flexible control
  Human interface and simulation

<Massively parallel system workshop>

This workshop will investigate R&D themes relating to architectures, operating systems, languages, and other areas which are necessary to achieve a general purpose massively parallel system. The main topics to be attended to are:

- Architecture related studies:
  Processor architecture
  Mutual linkage networks
  Maintenance architecture
  Robustness
  System architecture
  Calculation models

- Operating system related studies:
  Process control
  Network control
  Resource control
  Load distribution
  Improved reliability

- Language related studies:
  Object oriented languages
  Parallel procedure type languages
  Language models
  Programming environment

<Neural system workshop>

This workshop will attend to problems related to actualizing neural models, developing hardware for neural systems, and integration with massively parallel systems. Such problems will be of vital importance in developing large-scale sophisticated neural systems. The main problems this workshop will be concerned with are:

- Studies relating to neural models:
  Modularization
  Hierarchical management
  Self-organization
  Learning models
  Associative memory
  Integrating with symbol processing

- Hardware development:
  Neurochips
  Neurocomputers
  Scalable architectures
- Integration:
  Integrating neural and parallel systems

<Optical system workshop>

This workshop will study R&D problems related to the optical technology necessary for giving shape to real world computing. Real world computing will make use of the large volume capabilities and flexible connectivity of light as an information medium. The main topics to be attended to by this workshop are:

- Optical interconnection technology to

solve the signal transmission problems encountered in electronic systems

- Optical neural systems that make realtime learning and associative recall possible
- General-purpose massively parallel optical digital systems that help to attain a high level of operational accuracy and massively parallel processing

Toward Flexible Information Processing — Theory and Novel Functions —

Nobuyuki Otsu
Director
Machine Understanding Division
Electrotechnical Laboratory (ETL)

1. Introduction

The real-world environment is full of uncertainty and changeability. However, we humans can recognize what is in the environment, make decisions and act in quite a flexible way by processing and integrating a variety of information which is usually incomplete and ambiguous.

Such information processing as seen in humans could be characterized by the term "flexible information processing" or "real-world computing" in contrast to conventional, rigid information processing by computers, which assumes that all information is given in a pre-assumed world or problem domain.

The objective of research and development in the New Information Processing Technology (recently named the Real-World Computing) program is to lay the theoretical foundation and to pursue the technological realization of human-like flexible information processing as a new paradigm for information processing in anticipation of the highly information-based society of the 21st century, in which great increases in both the quantity and in the variety of information to be handled can be expected .

In what follows, an outline of the research and development of the theories and novel functions that will form the bases for this new realm of information processing will be briefly described.

2. Theoretical Foundation

2.1 Research framework

So far, much theoretical research has been done in respective research fields related to flexible and parallel distributed information processing. Such fields are, for example, pattern recognition, multivariate data analysis, probabilistic and statistical inference, fuzzy logic, and neuro-computing, machine learning, regularization and various optimization methods, and so on.

In order to provide a theoretical foundation for flexible information processing, it is important not only to continue in-depth study in these research areas but also to clarify the theoretical framework for the "soft logic" commonly underlying these fields and to aim at constructing a new unified theoretical base for dynamic real-world computing. Here, probabilistic and statistical formulation of problems and nonlinear dynamics in conjunction with learning and self-organization will be a key approach.

For this purpose, it will be necessary to expand and generalize the conventional framework of information processing in all aspects of information representation, processing and evaluation, and to systematize basic theories and elementary novel functions for application on the flexible framework.

The following are conceived of as the two most fundamental issues of common concern:

- Integration of multi-modal information (and of heterarchical processing modules),
- Learning and self-organization (optimization and adaptation).

Implementing these in a generalized and flexible framework of massively parallel and distributed information processing will be a major key point in developing novel functions for applications such as flexible recognition, inference and control.

It will also be important to learn from and receive inspiration from nature, namely, to take into account new findings in scientific research into the brain, evolutionary processes of creatures, and ecological systems.

In what follows we will investigate the topics of theoretical research that will serve as the theoretical foundation for novel functions in various applications and possibly for new computing architectures.

2.2 Research topics

2.2.1 Flexible representation of information

In order to treat various kinds of information in the real world (such as images, speech sounds and languages) and to construct heterarchical flexible systems, it is first of all necessary to establish a flexible framework for information representation. The framework should be flexible enough not only to represent various kinds of information in a unified manner but also to represent the certainty of information. It should also be suitable and efficient for implementing associative memory and learning (or self-organization) procedures.

Important research topics include distributed representation (patterns) vs concentrated representation (symbols), spatial representation vs temporal representation, probabilistic representation, topological representation, hierarchical representation, and the representation of information and knowledge as constraints, etc.

2.2.2 Evaluation of information and processing models

For actively interacting with the real world and learning or self-organizing from experience, real-world computing (denoted hereafter as RWC) systems should have a flexible and systematic framework (criteria) for evaluating the importance of the various kinds of input information, the output information that is a result of processing, and the processing models themselves.

One important issue will be to extend the evaluation framework from the conventional hard type that mainly deals with "true/false" to more flexible and quantitative types such as information criteria (e.g., AIC and MDL) and energy, which can serve as objective functions for optimization and regularization including for the information integration processes and processing models themselves.

2.2.3 Flexible storage and recall of information

The highly sophisticated functioning of the human brain's memory is a key to the flexible information processing of humans. RWC systems should have such flexible memory functions, to store and associatively recall various kinds of information. Thus research should aim at establishing methods of theoretical analysis for associative memory and at developing new, efficient mechanisms for flexible associative memory.

Important research topics are association using probabilistic reasoning, association using structural similarity, associative memory for storing time series, and associative memory using non-linear dynamical systems, etc.

2.2.4 Integration of information and processing modules

Information processing such as inference, prediction and planning can be considered to be a process for integrating various kinds of information and knowledge. It is important to carry out theoretical research to analyze these information integration processes and to develop new, flexible ways of controlling the processes.

Multivariate data analysis methods will provide the methods for information integration, although most of these are limited to linear transformations. Nonlinear extension of those methods will be important. Neural network models are thought to be able to provide a kind of non-linear extension. Regularization theory will provide a method for incorporating various kinds of constraints.

Cooperative processing by a huge number of processing modules is also an important issue in information integration. Research on the integration of processing modules will be important. Investigating the information integration process of human cognitive systems will also be important as a source of inspiration.

2.2.5 Learning and self-organization

In addition to information integration, learning (or self-organization) is one of the most important topics in flexible information processing.

It will play an important role in constructing adaptive autonomous systems, complex heterarchical systems or flexible databases. Thus, it is necessary to construct a computational theory of learning/self-organization for exploring the learnability of various concepts or structures and to develop novel, efficient algorithms.

Important research topics are learning from uncertain information, learning of probabilistic knowledge, learning of heterarchically structured knowledge, learning algorithms using active information acquisition, methods for incorporating existing knowledge with learning process, and selection of learning models, etc.

2.2.6 Optimization methods

Information integration processes can be formalized as optimization processes. Learning/self-organization can also be considered to be an optimization procedure. Solving these optimization problems usually demands a huge amount of computation. In order to surmount this difficulty, developing approximately correct optimization methods that can be executed effectively on massively parallel systems will be an important area of research .

Important research topics include probabilistic optimization methods such as simulated annealing, genetic algorithms, ecological algorithms, evolutionary algorithms, optimization using nonlinear dynamical systems such as neural network models, and other nonlinear optimization techniques, etc.

3. Novel Functions for Application

3.1 Research framework

RWC systems support various human activities by acquiring and processing various kinds of information in the real world such as images, speech sounds, tactile sensations so on. This information is massive and modal and moreover is subjected to incompleteness and uncertainty by nature. Thus RWC will require novel functions with flexibilities of various kinds. Attributes such as robustness, openness and real-time, reflect different attributes of flexibility. Therefore the key issue for novel functions is how to realize such flexibility.

The objective of research and development is to investigate elemental novel functions in the various application fields, closely cooperating with the theoretical foundations, and to embody those in respective application fields or integrate them to demonstrate new flexible information systems.

The novelty of the functions should emerge from new theoretical concepts or algorithms suitable to RWC. The bottle-necks of conventional information processing should be broken through by new kinds of flexible functions such as integration of symbols and patterns and learning/self-organization. Merely combining conventional technologies or making ad-hoc systems for specified tasks is not what is desired.

Since flexible information processing intends to expand the abilities of information processing beyond the conventional limitations, the range of expected fields of application is quite wide. These fields can be divided into four categories as follows:

- Flexible recognition and understanding of various kinds of information such as pattern information like images and speech sounds and symbolic information like natural languages,
- Flexible inference and problem solving based on flexible information bases which admit direct treatment of information and have capabilities for learning and self-organization,
- Flexible human interface and simulation for realizing mutual interaction between humans and the real world,
- Flexible control and autonomous systems interacting with the real-world environment.

The following two categories are conceived of as important directions to pursue for realization of integrated systems of novel functions:

- Real-world adaptable autonomous systems,
- Information-integrating interactive systems.

The former means cooperation with the real world. These will be flexible systems that can autonomously understand and control the environment through active and adaptive interaction with the real world and work for the purpose of partial replacement of human activities in the real world. Here it will be necessary to cope with the uncertain, incomplete and changeable characteristics of the real world. Necessary novel functions are understanding of the environment, modeling of the real world, planning for action sequences and optimal control for adaptation, etc.

The latter means cooperation with humans. These will be flexible systems which support and enhance human intelligence capabilities such as problem solving and information creation through enlarged communication channels between human and systems. Here it will be required to both understand and integrate various kinds of information flexibly to assist human beings in solving problems and creating new information. Necessary novel functions are question and answer by spoken natural language, understanding of intentions from various types of information produced by humans, realizing intelligent and interactive assistance for retrieving and presenting valuable information from a large amount of data in databases, intelligent simulation to create new information findings and forecast transient states in the real world, integration methods for combining human factors and a computational model of the real world, etc.

The fundamental novel functions should be evaluated both from the viewpoint of how they contribute to realizing these two systems and from the viewpoint of how they make breakthroughs in their own technological fields. The breakthroughs will be related to the attributes of real-world computing, namely, robustness, openness and real-time.

In the following, we will list the research themes which seem to be necessary for the realization of both systems. This does not necessarily mean that this program will pursue all of these research themes.

The computational bases for realizing these novel functions will be massively parallel and distributed systems including neural systems or optical computing systems.

3.2 Research topics

3.2.1 Flexible recognition and understanding

We can easily and flexibly recognize and understand various kinds of information in the real world such as images and speech sounds. This ability is a typical example of intuitive information processing and will indispensable for RWC systems to be able to possess autonomy and smooth communication with humans. It will also form the base for inference and problem solving at higher levels.

Many efforts have been made in research to mechanize these functions in even restricted domains, but the state of the art is still far behind the level of human ability. The difficulty of realizing recognition/understanding in the real-world environment lies in the incompleteness, uncertainty and ambiguity inherent to natural patterns and further in the ill-defined and ill-posed nature of the problems themselves.

Therefore, a new framework for solution is needed that admits these difficulties and can integratively include various constraints (knowledge) and also subjective value judgments, for example flexible frameworks such as optimization, constraint satisfaction, Bayesian inference, hidden Markov models, learning/self-organization from examples, or a new paradigm of parallel computation unifying these.

Important research topics are as follows:
For image understanding, the formulation of processing modules for early visual information (such as color, shape, movement) and integration mechanisms for these, interactive segmentation and understanding of scene and motion images on the basis of flexible modeling of objects and the world environment, understanding of facial expressions and gestures to infer human intentions, etc.

For speech understanding, noise suppression and robust recognition under noisy environments, speaker-independent understanding of conversational speech, formulation of heterarchical processing modules from the signal level to the linguistic level and integration mechanisms for these, flexible mechanisms for syntactic and semantic analyses, etc.

For natural language understanding, robust parsers applicable to large volumes of real-world coded sentences containing incompleteness, flexible methods to manage such sentence data as data-bases or electronic dictionaries and to utilize those for understanding, algorithms for extracting conceptual information, computational models for understanding situations and for integrating knowledge units on the basis of conceptual/structural distance to treat dialogue, explanation-oriented understanding of natural language, etc.

3.2.2 Flexible inference and problem solving

Conventional AI realizes intelligence by reducing human intelligence such as inference and problem solving to problems of logical operations and retrieval in knowledge representation by symbols. Such methods are efficient for restricted or rigorously abstracted problem domains, but for the variety of problem solving required in the real world, these face many difficulties such as the limitations of symbolic representation and manipulation, combinatorial explosion and knowledge acquisition.

In the real world, much information and knowledge contains inconsistency, incompleteness and uncertainty. There are many cases where the problems to be solved themselves cannot be described completely, and problem formulation is required based on incomplete statements by the user. In addition, real-time problem solving is also an important factor. Therefore, a new scheme of flexible inference and problem solving is necessary to cope with those problems. This will be attained by expanding the conventional framework to a more flexible one.

Important research topics will include flexible information representation integrating symbols and patterns, flexible knowledge acquisition from real-world information (learning of probabilistic structures or causal relationships), stochastic inference and analogy based on soft logic which integrates probability theory and fuzzy logic, constraint satisfaction or network dynamics to solve ill-posed problems or problems with many tacit assumptions as optimization problems, and massively parallel processing for the fast solving, modeling of the processes of inference and problem solving in humans.

3.2.3 Flexible information bases

Flexible information bases to handle various kinds of information and knowledge in the real world are important in the sense that they will support the intellectual activities of humans in the forthcoming era of high utilization of information networks and also in the sense that they will form a basis for novel functions such as flexible understanding/recognition, inference, problem solving, and action control.

So far, several types of databases have been proposed and developed, such as relational, object-oriented, deductive, and knowledge bases. However there are still many remaining problems with conventional databases and knowledge bases, which mainly rely upon symbolic and logical representation and retrieval.

Two new technologies should be focused on here. One is the flexible representation and memorization of the variety of real-world information within a unified framework for an inner information model ensuring tractability. The other is the flexible and semantically correct retrieval of the required items corresponding to the intentions of users.

Related important research topics include flexible information representation and data structures able to reflect the topological characteristics of objects and suitable for treating hierarchies of information, self-organization functions to cope with large-scale real-world information, learning of correspondences and cooperative relationships between various kinds of information, evaluation of information value and feature extraction for information abstraction or automatic indexing, detection of lacking information and active information acquisition, inference of user's intentions from incomplete or ambiguous requirements and completion through dialogue, high speed retrieval by massively parallel processing, etc.

3.2.4 Flexible human interface and simulation

Novel ways of using RWC systems must be developed in parallel with the development of the individual novel functions mentioned above. This new information processing will provide a new environment which enables broad, cooperative relationships between humans and computers and enlarges human intelligence activities. Users will be able to interact with computing systems by such natural means as spoken language, gestures and facial expressions, and will be able to receive information by means of real-time 3-D images, etc. This means that users can concentrate their intelligence efforts on more creative activities, freed from learning the exhausting skills required for communication with conventional computers through narrow channels.

Related research topics to realize such an information processing environment will include broad-band multi-modal interface (linguistic and visual communication including recognition of human body actions, facial expressions, etc.), information display systems with virtual reality, cognitive and behavior models including sensor fusion to understand human intentions.

RWC will also realize novel simulation technologies to solve very complicated and difficult problems. This will assist human thinking, creative activity and decision making by means of visualizing phenomena that have no inherent visual appearance. Instead of performing expensive and time consuming physical experiments which may not always be precise, computing systems will provide users with powerful tools to simulate very large scale complex systems and to predict their behavior in real-time. Prediction of untapped phenomena and future events, such as the global environment and weather forecasting, will be an important application of real-world computing.

Related research topics are learning/adaptation type simulation, prediction and control of complex/chaotic time series, and large scale simulation and decision making support for solving environmental, economical and traffic problems.

3.2.5 Flexible autonomous control

Research is needed to focus on the development of the technologies needed to realize flexible autonomous coordinated systems operating in the real world in real time. Robots are typical examples of this, and applications also include aids for the elderly or physically handicapped.

Such systems will be composed of various functional modules which interact with each other in perception, decision making and action control. On the other hand, the real world where the system operates is subjected to incompleteness and ambiguity in the available information, dynamic changes in the physical environment and limitations on available time and space. Thus, the important problems to be solved are how to integrate these various functions and how to control the interactions between the functional modules in order to achieve desired goal states under such real world constraints.

Related research topics are flexible modeling of the environment, task and control, active and distributed sensing and sensory integration, on-site planning and distributed cooperative searching, structuralization and coordination of multi-agents (for sensing, planning and action) for real-time skillful manipulation of objects, and maintenance of consistency between the internal world model and the dynamic real world, etc.

4. Concluding Remarks

An outline of the research on theory and novel functions in the RWC program has been shown in the above. This is, indeed, a basic and challenging research attempt to pioneer new horizons of information processing technology beyond the conventional frameworks.

Therefore, it will be necessary and important to promote interdisciplinary and international cooperation and to ensure flexible management of the program, not only in the research fields related to theory and novel functions, but also in other fields including massively parallel systems and optical computing systems, and other related research fields such as neural and cognitive sciences.

It seems that now is the time to bravely proceed one more step toward challenging the construction of a theoretical and technological foundation for the advanced and prosperous information society of the 21st century.

References

[1] N. Otsu: Toward soft logic for the foundation of flexible information processing, Bull. ETL, Vol.53, No.10, pp.75—95 (1989 Oct.).

[2] Report of the research committee on new information processing technology, Industrial Electronics Division, Machinery and Information Industries Bureau, MITI (1991 March).

[3] The Master Plan of RWC Program, Industrial Electronics Division, Machinery and Information Industries Bureau, MITI (1992 April).

Massively Parallel Systems

Toshio Shimada, Manager
Yoshinori Yamaguchi, Chief Scientist
Computer Architecture Section
Computer Science Division
Electrotechnical Laboratory (ETL)

1. Research Framework

Real-world computing requires a computation framework which canprocess various kinds of information flexibly and with integrity. Some novel functions for real-world computing applications areflexible recognition and understanding, flexible inference andproblem solving, flexible human interface and simulation, andflexible autonomous control. A system which implements one of these applications is likely to consist of many modules, which can exploit parallel and distributed processing at several levels both within and between modules.

Several parallel paradigms applicable to the above applications have been proposed, including , concurrent object oriented, dataflow, and data parallel, neural network, probability based information processing.

Real-world computing will probably be realized by some combination of these paradigms. In such a flexible information processing system, there are many levels of parallelism to be exploited. These new paradigms are naturally adopted to a massively parallel system and they require a huge amount of computation to solve practical problems within a reasonable amount of time. These observations show that a massively parallel system is necessary to support efficient computational power of real-world computing and it must also be a system which can efficiently execute the multi-paradigm.

The massively parallel system which is the computational base of real-world computing should be flexible in itself. The system flexibility comes in part from hardware, in part from architecture and in part from software. Hardware robustness and reliableness are essential. It should be general-purpose from an architectural point, because a massively parallel system should support multi-paradigms efficiently. Software will be needed to realize the functions of adaptation, self-organization and optimization. The system is one that can adapt itself to application environments for optimal performance. This flexible system will deliver maximum performance while minimizing the burden on the users. The goals of general-purpose massively parallel systems research in this program concern the development of architecture, operating systems, languages, and system environments of massively parallel systems.

2. Research Topics

2.1 Massively Parallel Architectures

The following are fundamental technologies which should be pursued in the development of general-purpose massively parallel systems.

(1) Model

Flexible execution models which can be bases of general-purpose architectures should have the ability to fill the gaps between the language models and hardware. Flexibility which allows a mapping of a virtual computer onto actual processing elements should also be pursued.

(2) Architecture

The architecture must support various paradigms efficiently. The massively parallel system should be based on a general-purpose architecture which does not focus on any special applications. It is also important to study hardware architecture in consideration of future device technology and packaging technology for an efficient implementation of massively parallel system.

(3) Interconnection network

The interconnection network should provide high speed communication which is comparable to computation speed.

It should also provide support for dynamic load distribution, global synchronization, global priority control. In implementing the high speed interconnection network system not only silicon technologies but also optical technologies should be considered.

(4) Robustness/Reliability

Hardware oriented robustness which can tolerate expected component failures in massively parallel systems should be examined.

System components should have self checking and self repairing facilities. The total system should have a maintenance architecture or facilities to maintain system reliability.

The system to be developed in the Real-World Computing Program should be a general-purpose massively parallel system.

The general-purpose nature may be achieved by mapping the multi-paradigm model to the hardware execution model in software. If the hardware can be modified and adapted to a variety of applications, it may be considered as general-purpose. Since a huge number of modules or processes must be processed concurrently in real-world computing, architectural mechanisms which can support rapid process switching and cheap synchronization should be implemented as an integral component of the system.

In the first half of this program, a prototype system which consists of 10^4 processing elements should be developed. Several hardware prototype systems are developed in this stage. One of these prototypes should be designed and developed as a platform. A platform system is used for tools of software development and a research platform on novel functions. Fundamental researches on massively parallel model and architecture are also concurrently studied.

In the second half, a massively parallel system to be developed is expected to have the order of 10^6 processing elements. It will have the ability to execute various kinds of real-world computing applications at real time speed. The architecture will be based on the new massively parallel computing model to be studied in this program.

2.2 Operating Systems for Massively Parallel Systems

The primary purpose of the operating systems are to assign computer resources to processes for efficiency and to build a user-friendly software environment behind which hardware details are hidden. Like operating systems for a serial computer, operating systems for massively parallel systems must be general-purpose, interactive, multi-user parallel systems.

The operating system for a massively parallel system should be designed to support the execution of various processes concurrently with high throughput. To administer massively parallel processors, the operating system may have a hierarchical structure much like human society. Support for the simultaneous execution of a number of parallel programs requires the operating systems to coordinate the partitioning of resources among the programs, as well as any dynamic repartioning at run-time. The locality, the concept of grouping activities and the principle of balancing are guidelines for partitioning.

One goal of an operating system is to provide the programmers a higher level of abstraction of a general-purpose parallel systems in order to make parallel programming easier. Through the abstraction, the operating system manages the resources and coordinates several user programs according to the above guidelines. The research issues of an operating system for the massively parallel system are as follows:

(1) Hierarchical structure

To realize a functionally distributed management system for flexible processor management, the operating system may require a hierarchical structure. Hierarchical structure makes the system scalable for a massively parallel system.

It is required that some efficient mechanisms to control activities and its hierarchical structure in terms of hardware and software. The reduction of overhead for controlling parallelism or executing critical sections should be considered.

(2) Network management

Advanced intelligent routing, addressing, synchronization, deadlock prevention, flow control, and failure avoidance should be incorporated into a flexible network management system.

(3) Resource management and load distribution

In the massively parallel system, the elimination of synchronization overhead, access contention and communication overhead will become more serious issues. To overcome these problems, the operating system should be able to collect management information autonomously and undertake statistical management or adaptive management. Memory management and virtual systems for several resources should also pursued for efficient scheduling and load distribution. For example, paging and virtual memory are the most successful ideas in memory management for a conventional computer system. It seems reasonable to expect these benefits in parallel systems. In order to balance load among processors, it is sometime desirable to migrate a group of processes to other set of processors. Efficient hardware support is needed to support process migration efficiently as well as the operating system facility.

(4) Fault tolerance

In the massively parallel system, resource management should be carried out in a manner which allows for expected component failure rates. Therefore, it is necessary to handle the failure avoidance system as a normal process. Multiple-route processing will also be required in order to deal with failures. In order to manage the massively parallel system efficiently, advanced intelligent routing and failure avoidance should be incorporated into a flexible network management system.

2.3 Languages for Massively Parallel Systems

The language for the massively parallel system must be able to describe the coordinated operations of a number of processes. Several languages based on several paradigms should be considered. The problem is how to extract the available parallelism in the problem domain, and be able to execute it with as much parallelism as the underlying system can provide. The user programs for such a system will be very complex. Development, verification, testing, debugging, and maintenance of large-scale parallel programs will be far more complex compared to the sequential counterparts. Various compilation techniques and run-time implementation techniques scalable to nearly one million processors should be studied. Design and prototyping of the language model and creation of the high level language models should be done.

Some high level languages which can be used for practical applications should be implemented.

The following items should be considered.

(1) Language model

Language model is a description model for flexible programming languages of massively parallel systems. The language model can provide a basic programming abstraction for the underlying architecture. The model must be simple and be sufficiently close to the underlying architectures so as not to restrict their computing power, and at the same time provide powerful means of abstraction to promote software programmability, portability, and reusability. Here, one viable candidate is the object-oriented concurrency model.

In the research of language model, the fundamental research on supporting flexible language, model of describing coordinating and cooperating actions, inheritance and reflections should be pursued. Inheritance serves to adapt the objects to a certain computing environment. Recent studies have provided deep theoretical foundations to object-oriented concurrency with regards to the concurrent behavior as well as their inheritance behavior. The idea of reflection in a language is to make the representation of the underlying program execution scheme manipulable within the language. Reflection allows the parallel system to encompass scheduling, communication, and load-balancing within the programming language framework.

(2) High level languages for massively parallel systems

The primary goals of high level languages for massively parallel systems should be ease of programming and also the ability to describe computation on the scale of a million processors. Several languages which will based on multiple parallel paradigms will be considered. One viable candidates of such languages will be an appropriate amalgamation of concurrent object-oriented, functional, and declarative constraint-based approaches. This amalgamation will create layers of programming languages.

The common base language make it possible to provide the basic programming abstraction of the several high level languages. Efficient inter-layer compilation techniques must be developed so as not to preclude the use of high level languages. Support for speculative computation within the parallel constraint satisfaction framework will be also exploited.

As for the concurrent object-oriented languages, currently available object oriented models are not intended to process more than million processes, so the following extensions will be needed.

- Introduction of a description system permitting hierarchical decomposition of complexity,
- Diversification of message propagation systems,
- Introduction of self-reflection functions for adapting and evolving objects,
- Declarative description of object relationships.

2.4 Environments for System Development and Programming

(1) Programming environment

Multi-paradigm can give users the freedom to choose among various programming languages. The need to develop a programming environment that can support multi-paradigm programming is expected. Tools for debugging, graphically monitoring and analyzing load balance, communication characteristics, etc. will also be required. Since these functions may need hardware support, the architectural design should take these requirements into consideration.

(2) System development support environment

The requirements of an environment supporting the development of the massively parallel system include two features different from conventional ones. One is the support for the interconnection network development.

In the development of the interconnection network, the requirements and limitations of several goals should be examined, including robustness, dynamic load distribution, and global synchronization mechanisms. The overall functions and performance of the interconnection network should be evaluable in advance by system-level simulation. The other is the support for the architecture development of processing element. In the development of an architecture for the processing elements of the massively parallel system, a set of basic functions for processing elements should be determined through a functional assessment of the various subsystems, including the interconnection network.

2.5 System Evaluation

The application fields should be picked to verify the effectiveness of massively parallel computational bases for real-world computing. The simulations that are dedicated to the prediction of future events from established models are examples for this purpose. These simulations can predict macro-behavior by describing interactions between micro-elements. Massively parallel systems are expected to handle situations such as non-linear or many-body problems by direct mapping paradigm. They will also be expected to handle simulations where the governing equations are difficult to formulate because of the complexity of the phenomena or of the incompleteness of our knowledge to describe the phenomena stochastically.

3. Research Schedule

The final massively parallel system to be developed is expected to have the order of a million processing elements.

However, it would be risky to develop such a large system in one step; prototyping is therefore essential. The prototype systems will be constructed in the middle of the program. The development technologies used will be intermediaries between the currently available ones and those required for the final systems. The progress of semiconductor technologies must be watched closely to set up the development plan. A platform system which is used for tools of software development and a research platform of novel functions should be constructed. One of the prototypes will be designed and developed as a platform. The software development on this platform system is very important for using the platform system as computational bases in the program. The final systems is the integration of the massively parallel systems and the neural systems.

The research and development of massively parallel systems is divided into two stages as follows.

3.1 The first stage

Several hardware prototype systems are developed in this stage. One of these prototypes is called platform system which is used as a platform for software implementation and evaluation. Prototype systems should be designed and developed as intermediate goals of the final massively parallel systems. At the end of the first half of the program, these prototype systems including a platform system are evaluated from the standpoint of designing final systems.

Fundamental research on massively parallel model and architecture are also concurrently studied. The establishment of the concept for languages and operating system, and prototyping on the platform system is pursued in this stage. Design and implementation of the common base language is done first. The research of the language model is also pursued to provide means of abstraction of high level specialized languages. Some high level languages and operating systems which can be used for practical applications in the second stage should be implemented on the platform system. Various compilation techniques and run-time implementation techniques will also be studied on the prototype systems. Also, a new user programming environment must be created to facilitate millions of concurrent objects executing simultaneously. We must provide special purpose high level languages which will accommodate the users' needs.

3.2 The second stage

The final system (or systems) are designed in the second half of the program based on the evaluation of the prototype systems which is developed in the first stage. The bases of the final system are selected from the prototype systems. The important technologies which are developed in the first stage are integrated into the design of the final system. The second half of the program is dedicated to developing the massively parallel system based on the massively parallel model which will be studied in the first stage. It will also be based on the results of research on theoretical foundations and novel functions in the first half of the program. The final system and the neural system will be integrated into a integrated information processing systems.

As for the software research, the second half of the program will focus on the implementation of the massively parallel operating systems and language systems. The extension of the high level language models is also pursued. Optimized compilation techniques is implemented on the platform and its prototyping is studied on the final system. Furthermore, in order to harness the complexity, hierarchical decomposition methodology of large-scale parallel programs and analysis framework should be studied.

Many researchers in this program should be able to use the platform system from their respective research institutes. This will be accomplished by connecting the research institutes and the platform system through high speed global network. The research and development of the total structure on such network environments should be pursued. In addition, the system software should support a standard user interface and a currently used system environment. The installation of small platform systems with a smaller number of processing elements at each research institutes will be considered. The ability for many researchers to work in common circumstances, giving them common concepts and technologies for real-world computing, is very important in the promotion of this program.

Neural Systems

Tatsumi Furuya
Manager
Computation Models Section
Computer Science Division
Electrotechnical Laboratory (ETL)

1. Introduction

One of the main objectives of real world computing is to break through the limits of conventional artificial intelligence (AI) techniques and implement information processing systems that perform in a manner much closer to the way information is processed by humans. The neural network, on which a great deal of research has been done in recent years, is modeled on the human brain. It is hoped, therefore, that the neural network will have much the same flexibility in information processing capability as humans have. Thus, with the neural network as a central focus, the research work under this program will pursue the theory, models, and finally, the hardware of neural networks in an integrated manner. The main research subjects that are at present considered to be necessary for the materialization of real world computing as well as the research schedules to be followed are mentioned below.

1.1 Background

The information processing speed of neurons is five or six orders of magnitude slower than silicon logic gates. In some cases, however, the human brain processes information far faster than even the fastest digital computers. A neural network is based on a model constructed so as to use the same principle of organization as is employed by living things. It is expected to exhibit intelligent behavior, processing specific information at high speed. Information processing by neural network is drawing attention because processing is based on the principle of "coordination and competition" between a large number of simple processing units that closely exchange information among themselves. Furthermore, since the programs that run on conventional computers are expressed in terms of intensity of connection, existing numerical optimization techniques can be used when systems learn or adapt themselves. As judged by these features, the neural network is considered to be capable of providing a platform for flexible information processing.

A variety of neural networks have been proposed in recent years. It has been shown that they are suitable for pattern processing, the solution of optimization problems through approximation, constraint satisfaction processing and so forth (Figure. 1). But those networks are limited to small scale applications. The neural models used in those applications are very simple, and back propagation is adopted as the learning method in most cases. Flexible information processing will be more difficult to implement than were past applications. We will have to search for new neural network possibilities with respect to learning methods, self-organization and so on.

		Number of Neurons (10^2, 10^3, 10^4, 10^5 . . . .)
Forecasting	Diagnosis	Diagnosis of small-scale plants; Diagnosis of large-scale plants
	Econom-ics	Short-term power demand; Short-term stock prices; Short-term economics; Long-term power demand
	Natural phenom-ena	Local weather; Earthquakes; General weather; Creation of environmental models
Pattern recognition	Voice	Vowels; Unspecific speaker: 26 words; Specific speaker's vocabulary voice recognition
	Charac-ters	Handwritten numerals & alphabet; Kana-kanji conversion 300 handwritten kanji; 4000 printed kanji
	Images	Simple noise elimination; Extraction of characteristics; 2-dimensional optical flow; Object shape recognition; Automatic target recognition; Semiconductor testing equipment; Cubic view; Medical images; Concept formation; Image restoration; Image under-standing
Control		Inverse kinematics; Inverse dynamics; Manipulator orbit plan; Sensor fusion; Autonomous automobile
Problem solution		Judgment of homonyms; Personal dictionary; Othello; Common dictionary; Translation of technical documents; Large-scale rule base; Game of "go"

Figure 1. Neural Network Applications

How to realize high speed processing on a neural network is another important subject. Today, neural networks are constructed on conventional type computers, and executed through simulation. Simulation speed is very low, especially on large networks. This problem will probably be substantially solved by general purpose super-parallel computers. But when cost and the scale of integration are considered, it will be desirable to have special hardware.

In this program, we will explore the possibilities for large neural networks to create flexible information processing systems that can operate in the real world. We will also construct a neural system made up of hardware and software based on a new neural model. The neural model we will work on under this program must be capable of learning new knowledge through interaction with the outside world and changing its own structure adaptively. Furthermore, it is expected that the hardware system will support one million neural network units and that it will realize a processing speed of 10 TCUPS (Tera Connections Updates Per Second). In the final stage, the neural system will be integrated with a super-parallel system to make flexible information processing a reality.

2. Research Subjects

2.1 Neural Models

In this project, we are trying to develop a new neural model for implementing a flexible information system. A flexible information system can be implemented using a large neural network that changes its own structure adaptively through interaction with the real world. The network grows dynamically in the real world through receiving help from teachers or through learning and self-organization. But as it grows, it must not forget the knowledge it already has. To realize these capabilities, the following fields of research are considered to be essential:

1. Neuron unit model

   Simple neuron unit models have so far been used with success in limited areas. But more advanced processing demands that a more sophisticated neuron model be developed. To begin with, the possibilities of already-proposed models, such as the chaos neuron model, the complex number neuron model and the neuron logic model, must be evaluated. At the same time, research must be done on new neuron unit models.

2. Modularization and hierarchization

   A large neural network used in the real world must have functions for interacting with the external environment and for changing the network structure through adaptation and training. In these situations, it needs the ability to acquire new knowledge without destroying the knowledge it already has and the ability to fetch information efficiently. To meet these requirements, the modularization and hierarchization of knowledge are necessary. We must develop a distributed learning method and a learning control method for the purpose of realizing modularization, structuralization and functional specialization. To this end, the following subjects will have to be considered:

- Functional specialization through adaptation to the external environment and the construction of hierarchical structures
- The autonomous growth of neural networks through interaction with the real world
- Modularization with the use of centralized or distributed control
- Interaction among modules
- Techniques for evaluating modularization and hierarchical structuralization

3. Learning and self-organization

Real time learning is very important in real world computing. It is vital, therefore, to work out a learning method in which new knowledge can be added without destroying existing knowledge and existing knowledge can be deleted while maintaining consistency. Many large neural networks have recursive loops. Hierarchical neural networks are strong in spatial pattern recognition while recurrent neural networks are expected to be effective for time pattern recognition or time pattern creation. Recurrent neural networks can also be applied to optimization problems because they have functions for satisfying constraints. It is thought that they will play an important role for the next ten years or more. Thus it is vital to construct a learning and self-organization mechanism for recurrent neural networks.

The topology and the size of a network are among the most critical parameters in the determination of its ability to be applied to general purposes. Learning procedures for constructing an optional self-organizing network must be developed. Such a network must be large enough for problem-based learning and small enough to be applied to general purposes.

4. Associative memory

Association is one of the basic functions created by neural networks. Various information such as spatial patterns and time patterns is memorized distributively and recalled on the principle of best match. It is necessary for us to clarify theoretically the principles of this associative function and to work out an engineering mechanism to implement this function.

5. New analog computing principles

Information processing by a neural system is based on the analog nonlinear dynamics of the system. We must clarify the principles of neural system analog computing, including chaos dynamics, from this point of view.

6. Integration of different paradigms

Research must be done on models for integration of different paradigms. For example, the integration of a neural network and logical processing, and the combination of pattern processing and symbol processing will be required for the implementation of a neural system. Integration models must be executed in a parallel processing environment. The input/output expressions will function as interfaces when different paradigms are integrated. Thus the selection of the input/output expressions for the neural network will substantially affect the processing performance of the network. In the development of a neural system, therefore, it will be important to do research on input/output expressions both from a theoretical standpoint and from an experimental standpoint.

2.2 Neural System Hardware

The hardware of the neural system must support a large neural network that is modularized so that it changes its structure adaptively through interaction with the external environment. The neural system is expected to process a network on the scale of one million units at high speed, a network made up of a large number of sub-neural networks in which thousands of neurons are completely connected. The target processing speed is 10TCUPS (Tera Connections Updates Per Second). In the design of special hardware, general purpose, extensible mechanisms must be incorporated partly because we do not have a sufficiently clear neural network model at present and partly because various other new models will be unveiled in the future. In the early stages of this research program, research must be done on a variety of architectures.

Hardware for neural systems can be classified into the following three types:

- Neuro-accelerators
- VLSI neuro-chips
- Engineering implementation of neural networks

A neuro-accelerator consists of special parallel processors developed for neural network processing. A large number of architectures for this have been proposed. In a typical architecture, a neuro-accelerator is made up of hundreds of element processors and realizes 1GCPUS. A VLSI neuro-chip is the hardware for neuron units that simulate a neuron unit model. The domain of neuro-chip architectures is so wide that it ranges from digital processing chips to analog processing chips. The engineering implemententaion of a neural network is the third approach in which the functions of the neural network are implemented through hardware logic engineering techniques, without the use of a neuron unit model.

These respective approaches cannot be compared since each one is unique in terms of learning ability, scalability and so on. For our neural system, we will focus our attention on the second and the third approach. The following choices are available:

1. Digital circuit neuro-chips

   Digital circuits have a number of advantages, such as high noise resistance and high processing accuracy. They are suitable for stable processing in a large system. We can directly apply computer manufacturing technology to the fabrication of digital neuro-circuits. In addition to the conventional approach, a large number of variations using the pulse density model or the like have been proposed. It is also anticipated that these approaches may produce new neuro-chip possibilities.

2. Analog circuit neuro-chips

   Because it has fewer operation circuits, an analog system makes it possible to reduce the hardware volume. This is a very attractive factor in the development of a large network. Moreover, analog circuits have potential for the implementation of dynamic networks and complex neural networks such as the chaos neuron network.

3. Hybrid neuro-chips

   It is also possible to consider the use of hybrid neuro-chips, which combine the strengths of both digital and analog neuro-chips.

4. Engineering implementation

   The neuro-chips mentioned above are designed to implement a neuron unit model directly. On the other hand, there is another, completely different approach to the implementation of the functions of a neural network. This approach uses logic hardware that does not simulate a neuron unit model. Such a system is constructed through the repetition of simple logic circuits.

In a neural system, all neurons fundamentally exchange active values among themselves. The interconnection network architecture is an important point in design. A time multiplex system or a frequency multiplex system are possible solutions to this problem. The design of connection networks is closely related to packaging technology. The band width of connection networks is determined by restrictions on the number of pins on chips and boards. The technologies to solve this problem will include integration on the wafer scale, three-dimensional architecture and optical technology. CAD, silicon compilers and so on are also considered to be important design tools.

2.3 Neural System Software

A variety of neural software systems are required for neural system research and development.

1. Simulation system

   The simulation system is an important tool in the research and development of neural networks. Most neural network researchers have their own unique simulators. A flexible, general purpose neural simulator for large neural networks would be a powerful tool. Among the requirements for such a simulator are machine independency, extensibility, convenient user interface, high speed and a large number of utility routines. It is also desirable that such a simulator have mathematical analysis tools designed to control the convergence and other characteristics of individual networks.

2. Neural network language

   Neural network processing should be described using a high level language. The design of such a language demands that research on the following research and development themes be pursued:

   - Expression of ambiguous information
   - Description of best match operations
   - Integration with logic programs
   - Integration with simulators

3. Operating system

   The number of hardware neurons is generally smaller than the number of neurons in the neural network. Therefore, a virtual mechanism to fill in this gap is an important part of a neural system.

- Mechanism for mapping the neural network on the hardware
- Scheduling of resources

2.4 Integration with a Super-parallel System

It is highly probable that a neural system will be one of the processors for such specific purposes as pattern recognition, associative memory and the pursuit of optimal combinations. This automatically means that the neural system will be combined or integrated with other systems. The forms of integration with other systems range from close connection to loose connection. The following forms of integration have already been proposed. An example of close connection: the neural system is connected as an associative memory to an element processor of a super-parallel system. An example of loose connection: super-parallel processors throw problems such as optimization problems to the neural system.

3. Research Schedules

The development of the neural system is divided into two stages, the early stage and the late stage. In the early stage, we will research a wide range of approaches with a view to expanding the theoretical, model and hardware architecture possibilities. For hardware, we will construct a prototype system that can function as a test bed for experiments on new functions in the late stage. In the late stage, we will make our theory and model more sophisticated on the basis of the results in the early stage. At the same time, we will construct a final system on the scale of one million units by utilizing our experience with the prototype. The final system will be integrated with a super-parallel system. A number of application systems will be mounted on it for evaluation of the total system.

4. Conclusion

In the neural network field, future visibility is quite low. Some researchers say the field will be defunct ten years from now while others think it is the only field in which a new information processing breakthrough is conceivable. For this reason, research on neural systems is both challenging and exciting. Progress in science and technology is so rapid that it is difficult to predict what the situation will be ten years from now. In a good sense, this means that we can expect that this research plan will be rewritten in the future so that we can progress in a more fruitful direction.

Optical Technology

Hiroyoshi Yajima
Manager
Optical Information Section
Optoelectronics Division
Electrotechnical Laboratory (ETL)

1. Introduction

An optical computer can be defined as a machine that uses light for the transmission and processing of information. This definition is symmetrical to the definition of an electronic computer, which uses electrons for the transmission and processing of information. Then why are optical computers drawing a great deal of attention today? It is because by using light as a medium of information, an optical computer may be able to actualize forms of information processing that are hard to accomplish with an electronic computer.

2. Light as a Medium of Transmitting Information

2.1 Information Transmission Capacity

As an example, let's compare the information transmission capacity of 5 GHz electric waves and 500 THz visible light. Since the volume of information carried by an electromagnetic wave is generally proportional to its frequency, the information transmission capacity of light is in principle five orders of magnitude greater than that of electric waves.

Let's also compare the two in terms of information transmission capacity per unit of cross-sectional area. Since wave guide cross-sectional area is proportional to the square of the wavelength, the information transmission capacity of light per unit of cross-sectional area is ten orders of magnitude larger than that of electric waves. As a result, the difference between the information transmission capacity of light and that of electric waves is expressed by the volumetric ratio shown in Figure 1.

2.2 Flexible Connectivity

A major feature of light is its flexible connectivity, which results from its short wavelengths. The short wavelengths of light makes it possible to generate beams that have excellent directivity in propagation through space. Light also enables broadcast type transmissions from one point to many points by use of lenses and one-to-one connections between mutually remote ultra-small areas. In contrast, due to its long wavelength, an electric wave can be used only as a means of information transmission of the broadcast type.

Let's then make a comparison with respect to transmission paths. Electric wiring consists of conductors, and inter-wire cross talk proportional to the signal frequency occurs through the medium of distributed capacitance. Cross talk also occurs between optical lines but it is not related to the signal frequency. Thus optical lines enable ultra-high speed signal transmission as well as more flexible

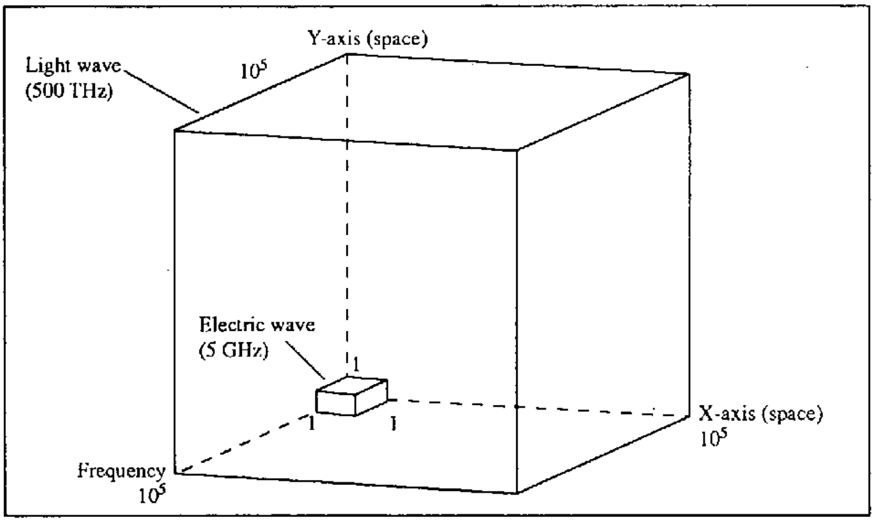

Figure 1 . Comparison of Information Transmission Capacity between Light and Electric Waves

connections through combination with optical switching.

3. Light as a Medium of Processing Information

3.1 Operating Speeds of Optical Elements

The maximum operating speed of electronic elements has by now reached the order of several picoseconds. But it is still thought that it will be difficult to implement sub-picosecond operating speeds. The extent to which the rate of optical pulses can be increased is determined by the uncertainty principle of time and energy. The maximum is considered to be 1 femtosecond in the case of visible light. But we do not have the technology to control such high speed optical pulses. In view of integration, the main types of optical devices that will be used in the future will be semiconductor devices. The operating speed of semiconductor devices is limited by the movement of the electrons in the semiconductor. Under these circumstances, the operating speed of optical devices will not be greatly higher than that of electronic devices.

3.2 Parallel Processing

One of the main features of optical processing is its parallel processing potential. Figure 2 shows an image of an optical computing system. The basic computing process, with a feedback loop, is the same as that of an electronic computer. In this process, however, information with spatial expanse is entered directly and all the individual units of information propagate through the system simultaneously and in parallel. Now let's make a comparison with respect to the degree of integration of the element technologies that make up the system. The minimum size of optical elements is limited by the wavelength of light, whereas the minimum size of electronic elements is determined by the wave-

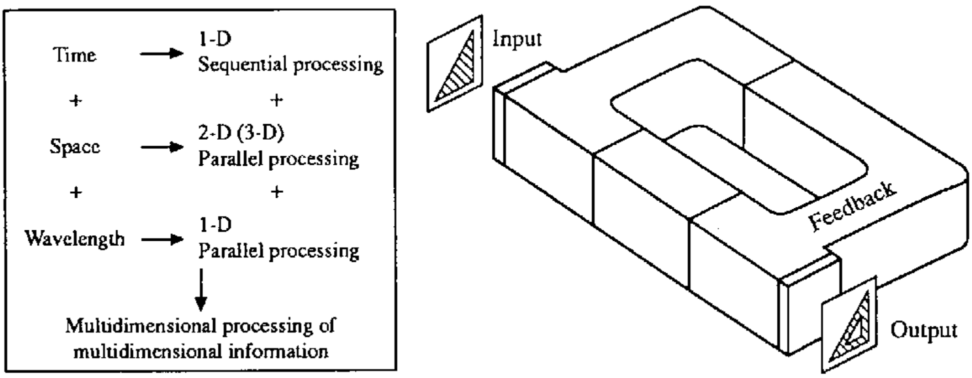

Figure 2 . Image of Optical Computing System

length of electron waves. Thus the limit to the degree of integration of elements per unit area is higher in the case of electronic elements. But a three-dimensional system such as the one shown in Figure 2 can be more easily structured with an optical system, which can use space propagation, than with electronic circuits, which require actual wiring. Also, while the information handled by an electronic system must in principle be time signals, an optical system can directly handle not only time signals but also information that has spatial expanse or expanse along the wavelength axis. With an optical system, therefore, we can expect the multidimensional processing of multidimensional information.

4. Optical Device Technology

Optical technology is supported by optical device technology. Just as the basic functional device that supports electronics is the transistor, the basic functional device that supports optical technology is the optical transistor. The important factor in optical technology is the relation between light and electrons. Electrons will play some kind of role in the operation of optical devices. Therefore, an optical transistor should more precisely be called an optoelectronic (OE) transistor. Figure 3 shows the classification of two- and three-terminal devices using electricity (E) and light (O) as input/output terminals. The combination of light and electrons provides a wide variety of device structures and functions. It is seen in this classification that purely electronic devices and purely optical devices make up only a part of the total range of Optoelectronic devices. As for the

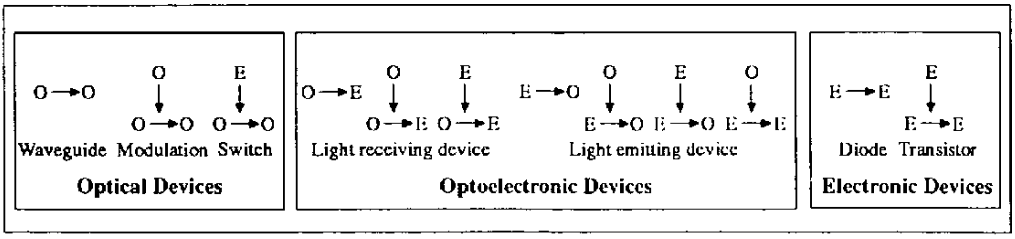

Figure 3. Optoelectronic Two and Three Terminal Devices

information handled, electric terminals can handle only time signals, due to their principle of operation, while optical terminals can handle not only time signals but also information that has spatial expanse or expanse along the wavelength axis. OE transistors will be integrated into optoelectronic IC, (OEIC). As they are improved, OEICs will make it possible to construct optical systems [1, 2].

5. RWC and Optical Technology

The target of real world computing (RWC) programs is to realize "flexible information processing." Based on the premise that logical information processing and intuitive information processing can be integrated, the goal is to replace conventional sequential information processing with flexible information processing, and to develop new forms of information processing marked by "distribution and cooperation" or by "parallelism and flexibility." The implementation of flexible information processing will require a means of realizing the flow and processing of huge volumes of information.

The role of optical technology in RWC is to provide the tools necessary for the implementation of the information paradigms aimed for in RWC by taking advantage of the features of light as a medium of information, that is, huge information transmission capacity, massively parallel operation and flexible connectivity.

In the RWC program, the tasks involved in the development of optical technology can be divided into three major categories; optical interconnection, optical neural systems and optical digital systems. Optical interconnection is a technical issue related to smoothing out the flow of information in information systems. At the same time, it is a part of the basic technology required for the construction of both optical neurosystems and optical digital systems. The optical neurosystem can be positioned as a system technology that uses light to accomplish the super-distributed processing of information aimed for in RWC, while the optical digital system is viewed as a system technology to use light to accomplish massively parallel processing. Figure 4 illustrates the relationships between new information processing (RWC) and optical technology.

5.1 Optical Interconnection

Today's electronic computers are supported by silicon device technology, and the processing speeds of the devices themselves have been rising year by year. But along with this progress in the combination and parallel operation of systems, system performance has come to be affected to a large extent by the time required for the conveyance of information within systems. As a result, communication bottlenecks are now a major obstacle to improving the performance of information systems.

Optical interconnection is a way to solve these shortcomings of electric wiring by means of the huge information transmission capacity and the flexible connectivity of light.

5.2 Optical Neurosystems

The aim of developing optical neurosystems is to materialize the same information processing processes as performed in image recognition by the human visual system and in

	Neural operation	General-purpose massively parallel operation
Type of operation	Correlation, Learning	Logic, Symbol Processing
Problems with conventional computer	Communication bottleneck	Limitation of sequential processing

⇧

		Interconnection	Massively parallel operation	
			Analog	Digital
Optical approach	Time area	Time-multiplexing, ultra-high speed	Linear	Non-linear
	Space area	Space-multiplexing (parallel), space propagation, wave-guide control	Wavefront control	Space coding
	Spectral area	Wave-length multiplexing (parallel)	Wavelength control	Wavelength coding

Figure 4. Relationship between New Information Processing and Optical Technology

the association, learning, intuitive judgment and other functions of the human brain by utilizing light-using neural networks and the unique physical properties of light. In principle, light permits massively parallel operations and high density wiring, which are considered to be promising technological advantages for the large neural systems of the future.

5.3 Optical Digital Systems

Based on the combination of the massive parallel operability of light and the principles of logic operations, the purpose of developing optical digital systems is to develop the technology for the high speed, precise processing of information with a certain expanse, such as image information, as well as wavelength information.

Optical digital technology will not compete with electronic computers for higher performance in the same arena. Instead, the aim is to develop new aspects of optical information processing technology, which has conventionally centered on analog handling of information, through the direct digital processing of optical information. This is considered to be a promising general purpose technology that can improve the level of optical technology as a whole.

6. Conclusion

One problem with optical technology is that in spite of the potentialities of light as a medium of information, the device technologies required for realizing those potentialities are in general immature. Under the present circumstances, therefore, people researching optical computing architectures are having a great deal of difficulty in securing the optical device technologies they need while people involved in optical device development can hardly spare any time and energy on research into optical computing architecture due to the difficulties of hardware development.

The RWC programs are intended to form a common platform for expediting exchanges between these two groups of people and to thereby make flexible information processing that utilizes the potentialities of light a reality.

Literature

[1] H.Yajima: Present and Future of Optical Devices, JITA Now & Future, 7, 25, p2, 1991

[2] H.Yajima: OEICs in Japan, JITA Now & Future, 7, 25, p.5, 1991

RWC Related Technologies (1) Virtual Reality

Michitaka Hirose
Associate Professor
Dept. of Mechano-Informatics
The University of Tokyo

1. Introduction

Virtual reality is the technology for generating a virtual space around the human being by presenting information synthesized by computers directly to the human sense organs.

In other words, it provides an advanced human interface that can present such a great deal of information to the users that it causes them to feel that those surroundings are reality.

In this sense, this technology is very suggestive of what the relationship of the new generation of computers to humans should be and what form inputs and outputs should take.

Up until the first half of the 1980s, research and development efforts on computers had primarily been centered on improvement in performance of the computer itself. In recent years, however, much importance has also come to be attached to the overall state of human and computer technology for interfacing computers to the external world, that is, to human interface technology. At the background of this tendency is the idea that no matter how advanced the information processing performed in computers is, it is completely meaningless if it is not conveyed to people.

The first point of novelty that virtual reality has can be expressed by the term "sense of reality" or "existence". Presenting three-dimensional images that are quite like actual things, enabling the observers themselves to enter the world of the image, and using not only images and sounds but also the other senses such as feel are part of the novelty unique to virtual reality. A second point of novelty is that a displayed object can be viewed from various directions and can be picked up. That is, a second characteristic is having an interactive element in which the observer not only can receive information from the system, but can also positively influence it.

These two elements add up to the composition of a sense of shallow reality. But when a deeper sense of reality is to be created, a simulative element is required as a third element. In other words, it must be possible to depict events in the virtual world that have consistency and coherence, such as events in which a virtual object thrown up in the air bounces back upon bumping against another virtual object.

The technology of virtual reality is based on the solid fusion of these three elements. Then, what kinds of systems are necessary to make these things possible? First, a variety of displays are necessary to present information to the various sense organs. Second, in order to implement interactiveness, a system is required that can directly enter conscious and unconscious human actions into the computer. Third, a system that defines certain kinds of causal relations is necessary for connecting the displays and the input system. These three systems are explained below. Cases from the author's research are offered as examples when needed.

2. Implementation of Virtual Reality System

(1) Input system

An input system is required to implement interactiveness, or to enable a person to work positively on the system. The most typical input system is probably the system called a "Data Glove" (Figure 1). The Data Glove uses the principle that the light transmittance of optical fibers attached to the fingers will change as the fingers are bent. This device makes it possible to handle a virtual object present in a computer intuitively.

Besides this, in order to support movements in space caused by larger actions such as walking, a system can be used that enables movement in a large virtual space by using human motion on a tread mill as input. This system was originally developed by the University of North Carolina. It is useful in applications that simulate walking around inside a large building, for example.

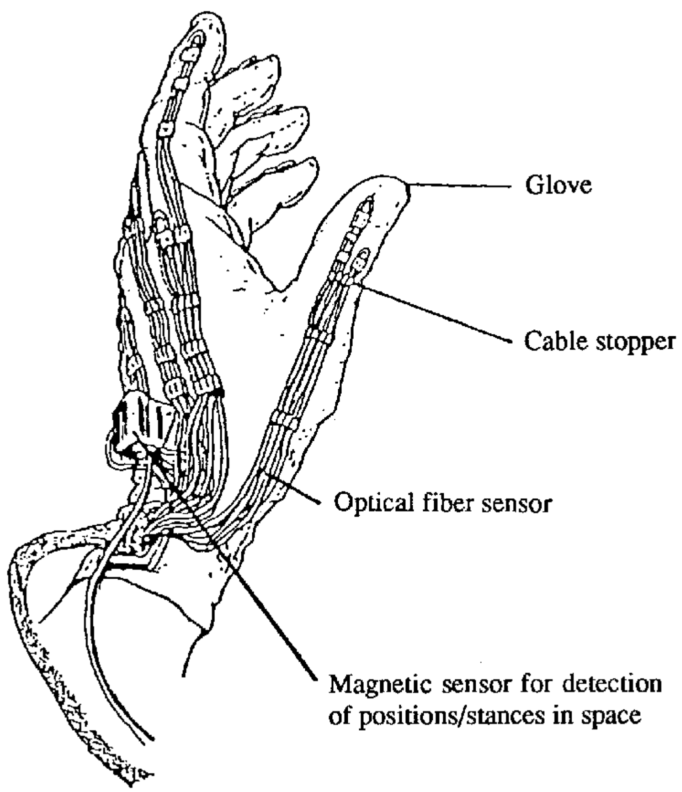

Figure 1. Data Glove
(A product of VPL Corp.)

In addition to the above-mentioned systems that use attached devices, various systems of completely non-attaching types are also being considered. One of these is a system in which the human body is monitored with a video camera and the motions of the individual parts of the body are measured through image processing. In view of the recent progress of real time image processing technology, this field may provide opportunities for substantial growth.

(2) Display system

Next the displays, or the output system, will be described. What is required of the display system is the ability to produce a visual sensation. The quality of the indicated contents is, of course, an important factor in realizing this ability. It is also necessary that most of the range of human senses be replaced by displays.

A head mounted display (HMD) can support both a wide viewing angle and a wide look-around range. Thus it is effective for obtaining a high degree of visual reality. The user of the HMD can totally enter into the world of the image. Shown in Figure 2 is a see-through type HMD. This device makes it possible to implement the fusion of the virtual world and the visible world.

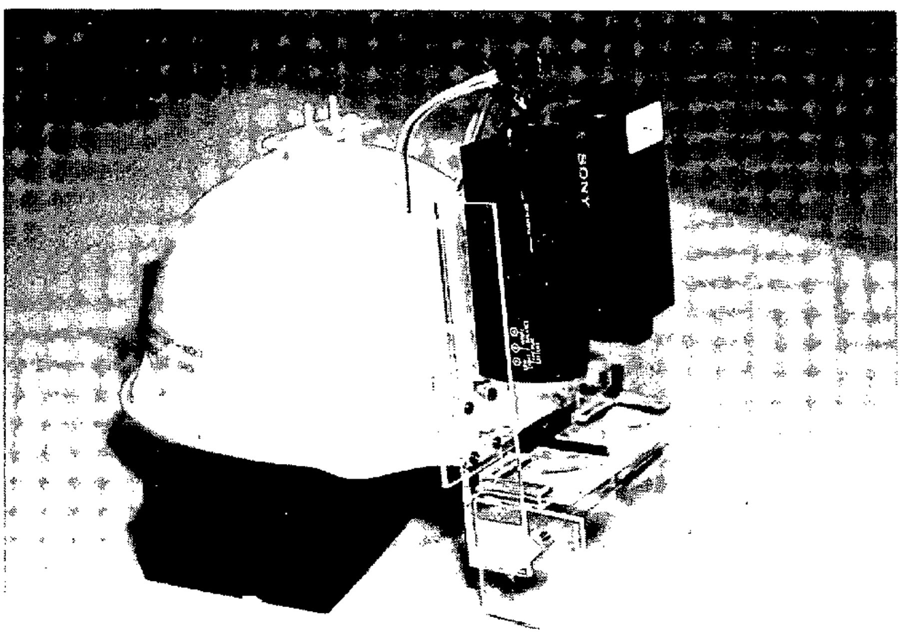

Figure 2. See-through HMD (The University of Tokyo)

Substantial downsizing is necessary for making the practical use of the HMD possible. For this purpose, it will be vital to develop very small, highly accurate display devices. The liquid crystal display (LCD) is considered to be most promising candidate for this application. As shown in Figure 3, for example, it is now possible to fabricate a fairly small HMD (160gw). By world standards, the LCD technology of Japan is at a high level, and Japan is expected to play a leading role in this field.

The sound display can be the first type of sensory display to considzer in addition to the visual display. Using headphones, there are systems that can position a sound image at any point in a three-dimensional space.

The system is designed to reproduce the transfer function at the time when sound waves from a sound source reach the eardrum in actual space by computer on a real time basis. The execution of such complex calculations in real time owes a great deal to the diffusion of digital signal processors (DSP).

There is no definitive technology as yet with respect to the sense of touch. As an example, Figure 4 shows a mock surface technique in which a surface that does not actually exist is expressed by moving a small part for surface presentation together with the fingertip. Using a rectangular coordinate manipulator that operates at a sufficiently high speed, this technique makes it possible to present a small

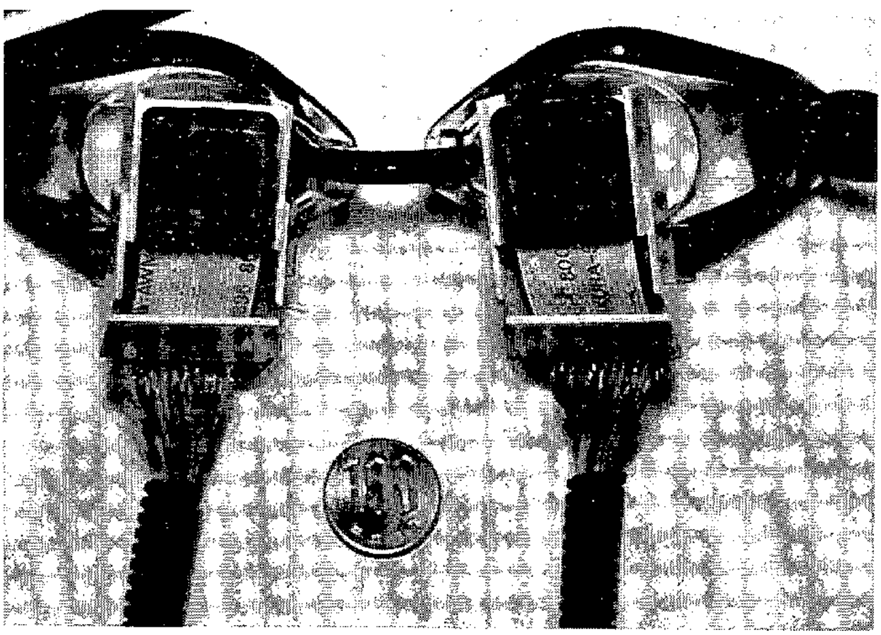

Figure 3. Small HMD
(The University of Tokyo)

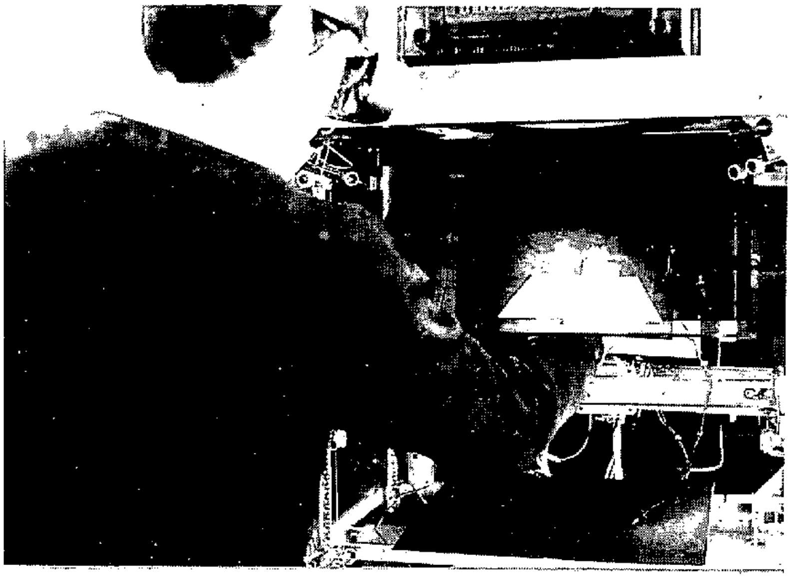

Figure 4. Operation of Putting a Virtual Object into a Hole
(The University of Tokyo)

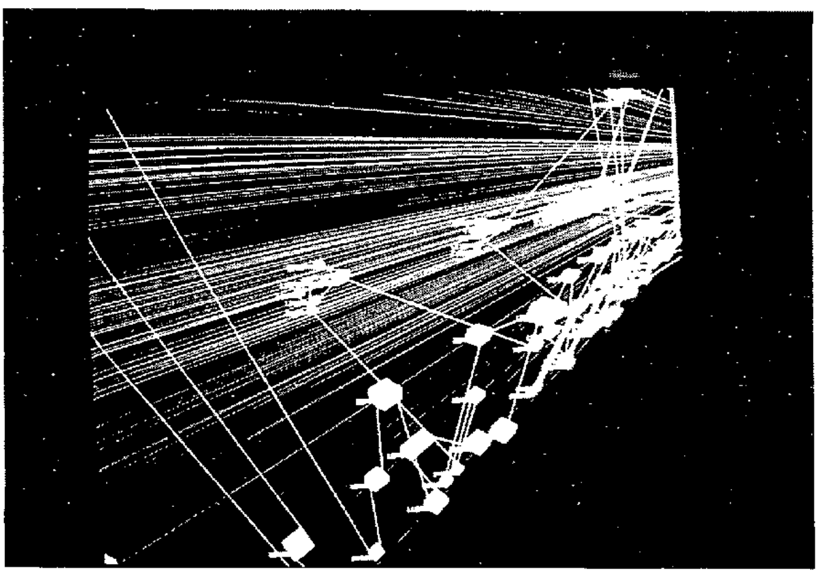

Figure 5. Visualization of Software
(The University of Tokyo & Tokyo Electric Power Company)

tangent plane near an object when its surface is approached by the finger.

The system shown here is called virtual holography. The visual display that is used is a stereo display with a liquid crystal shutter for head-linked display. The three-dimensional image it generates is displayed before the eyes with a half mirror. By using this system, we can grasp hold of a CAD model inside a computer and check, for example, the possibility of assembling two or more parts.

(3) Simulation system

After the interface devices described above are readied, software will be required to drive them. This software will include a database to define the shape of the objects that exist inside the virtual world. This shape database will be much the same as the shape database used for ordinary CAD systems.

In addition, we need a group of procedures to describe and simulate functional aspects, such as constraints among objects (e.g. the motion of an object on a rail is restricted to one-dimensional motion, and objects cannot enter into other objects) and the rules that govern the motion of objects (e.g. an object falls under gravity, and an elastic body rebounds upon bumping against a wall).

The technical difficulty lies in running all this complex software on a real time basis. When a somewhat complex virtual world is to be generated, such for checking of complex laws of the world or generating "real" images, a volume of calculations that far exceeds the capabilities of existing computers must be performed. Therefore, super-parallel and super-distributed computer hardware is an indispensable technical element for the implementation of virtual reality.

3. Virtual Reality as a Visualization Technology

With these tools, virtual reality systems are expected to be used in a wide variety of applications. The application considered most promising is the visualization of information. For example, it will become possible to propose a system that is one step ahead of conventional CAD systems. In construction CAD, for example, the conventional technology cannot possibly express vividly how frightening it is to look down from the top of a slope with a gradient of 30 degrees.

Further, aided by computer, it will be possible to synthesize things that cannot be seen with the naked eye. Software itself is a typical example of such things. A team set up jointly by the Tokyo Electric Power Co., Ltd. and the University of Tokyo is trying to develop a system to visualize large scale software. It is extremely difficult to grasp the total scope of large scale software. A macroscopic methodology has been presented for this system. This methodology is quite different from the microscopic methodology studied in software engineering today. That is, as shown in Figure 5, efforts are made to visualize software, which cannot actually be seen, with the use of virtual space.

4. Conclusion

One major feature of the technology of virtual reality is that it presents a new methodology for positively using the information processing capabilities of the persons who are the users, as opposed to conventional computer science, which has concentrated primarily on improving the performance of the computers themselves. In this sense, this technology is one of the most promising technologies of the 1990s and should be pursued as a new field of information processing technology.

RWC Related Technologies
(2) Optical Neurocomputing

Masatoshi Ishikawa
Associate Professor
Dept. of Mathematical Engineering &
Information Physics
The University of Tokyo

1. Introduction

Optical computing is drawing attention as a promising computing mechanism for the future. As shown in Figure 1, optical computing features high speed, parallelism, high density interconnections, and so on. Optical computing is an attempt to break through the limits of the present computing, which is based primarily on the conventional electronic technologies. It has long been pointed out that optical computing is possible in principle, but few optical computing systems have implemented to date, partly because there are few parallel algorithms, and partly because it is difficult to provide versatility in operations.

Neurocomputing can provide optical computing with a highly practicable architecture that has operational flexibility. Conversely, the high parallelism of optical computing is a very attractive feature to base neurocomputing on.

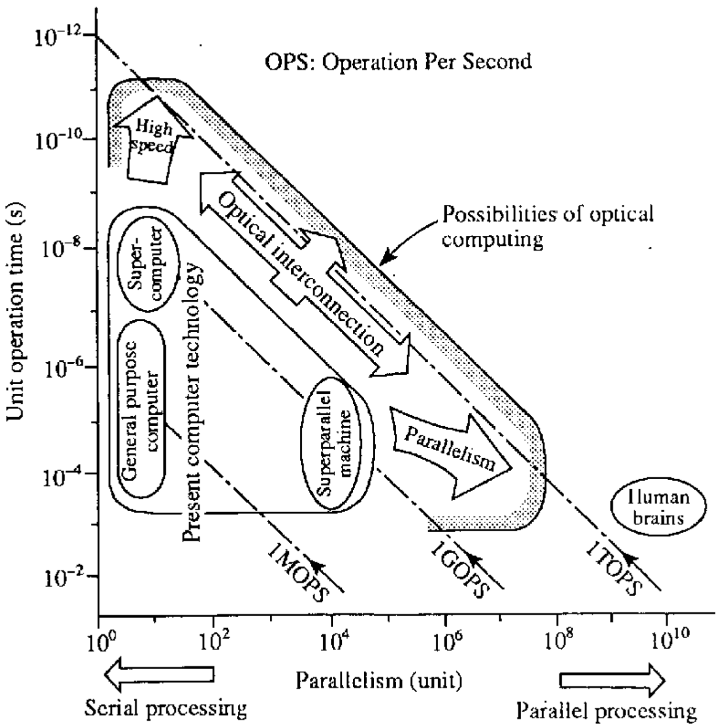

Figure 1 Possibilities of Optical Computing

This paper focuses on the fusion of optical computing, parallel learning, and information processing and describes its structure and features through concrete examples.

2. Features of Optical Neurocomputing

Conventional computers express a logical procedure (a program) in a series of operational structures. In contrast, neurocomputing employs the technique of extending an operational structure spatially. It features large scale parallelism in which a neuron model is the basic operational element and learning mechanism. The neuron model replaces the program model to generate operational flexibility.

2.1 Large Scale Parallelism

It is a major task to use existing computer technologies for large scale parallel processing. Attention is focused on the development of practical operational techniques that can be implemented parallely.

Large scale parallel neurocomputing requires a network that uses a large number of low competance single-function operational elements (equivalent to neurons) with parallel high density connections.

2.2 Learning

Conventional optical computing can utilize intelligent processing in such specific applications as Fourier transformation analysis, but it still falls short in providing such versatile operation in any given required operational structure. It has been said, in particular, that conceptualizing parallel algorithms is difficult.

Learning makes it possible for optical computing, which is otherwise lacking in flexibility, to implement an operational structure that fits any specific object of operation or environment.

2.3 Optical Interconnection

As a medium of information, light features inter-signal incoherence and a wide band in addition to spatial parallelism. Moreover, it is free from inductive noise and capacitive delay, and it requires no grounding. Furthermore, it permits high density connections in the direction vertical to the operation plane. This presents a striking contrast to the wiring on the substrate or inside of a LSI circuit, which is confined to a plane of operation.

The technique of using light, which provides these advantages, for communication among operational elements is called optical interconnection. Aside from using light for the operations themselves, it can also be used to interconnect electronic circuits.

2.4 Direct Processing of Images

One of the merits in using light as a medium of information is that it can be used in the direct processing of external visual information. It is expected, therefore, that an intelligent sensor that combines a detection mechanism with some sort of recognition mechanism can be created.

3. Proposed Systems

Based on such concepts, a large number of optical neurocomputing systems, especially optical associative memory systems, have been proposed to date.

3.1 Retinal Chips

With respect to the direct processing of visual information, C. Mead et al. have proposed a "silicon retina" in which light sensors and computing elements are combined on a silicon chip.[1] It approximates a retina using a silicon chip that has sensors and processing circuits. It consists of light sensors and simple analog electronic circuits. Utsugi and Ishikawa are developing a method to represent external coordinate systems internally with a resistance network.[2]

3.2 Implementation of the Hopfield Model

The Hopfield model is a calculation model in which a dynamic system is made by feeding back the outputs of a single-layer network to its inputs, and the solution of a problem is obtained by minimizing the energy of the system.

Farhat, Psaltis, et al. were the first to propose an optical parallel operation circuit that is designed to implement the model.[3] In their method, a common optical matrix vector operation circuit is used and the result of an electronic circuit's threshold operation output is fed back to the input side. Because the intensity of light is used as the medium of information, negative values can not be expressed. Therefore, positive and negative memory matrices are written to separately and require separate positive and negative elements. Then, deductions are made through an electronic circuit. In the meantime, Ohta et al. have implemented the Hopfield model by developing an optical neurochip that integrates an LED array and a PD array.[4] Figure 2 shows this chip.

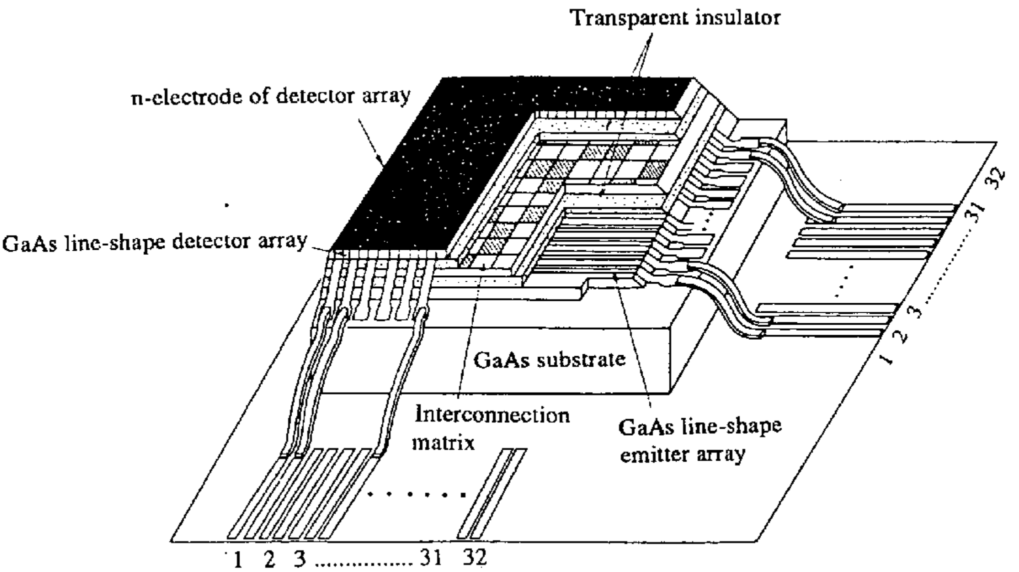

Figure 2 Illustrative scheme of GaAs/AlGaAs optical synaptic chip

4. Implementation of Learning

Various other systems using holograms, optical disks, and electronic neurochips have been proposed. However, most of them have no self-learning mechanisms. Others, though having self-learning mechanisms, are merely proposed. Few systems have precisely demonstrated the essence of parallel learning and information processing.

In recent years, however, Ishikawa et al. have made a prototype of an optical associative memory system with learning capabilities[5] and conducted some learning experiments with it.[6), 7] With this system as an example, the creation of a learning mechanism will be discussed below.

4.1 Optical Associatron

Ishikawa et al. have created an optical autoassociative memory system with a learning capability by using a spatial light modulator. They have named it an "optical associatron."[5] A spatial light modulator is an optical device that changes the contents of memory matrices as required during learning. In a two-dimensional optical pattern, it permits such operations as write, read, add, subtract, and erase.

The optical associative memory system must make positive use of optical matrix operations and the spatial light modulator's analog operation and memory functions. Moreover, it must be structured so as not to require the use of negative values.

In consideration of these points, the optical associatron has adapted an orthogonal learning method. This is a learning method of the closed loop type. This method is based on successive least square approximations of input/output relations.

Figure 3 shows the structure of the optical associatron. All optical systems are first conceived. Then, only the main parts (i.e. recalling and learning of memory matrices) of the system are implemented with optical operations, and the other parts are replaced by computers and electronic circuits. That is, it is a hybrid system. The spatial light modulator 1 is used to memorize the memory matrices (16×16, analog). Learning is done through a revising the pattern presented by the light emitting diode 1. The function of the spatial light modulator 2 is to convert the pattern into coherent light as an input pattern multiplexed into the form that fits a memory matrix. The light created by the intensity modulation of the memory matrices through the spatial light modulator 1 is then entered into the spatial light modulator 2 to be subjected to intensity modulation and input pattern formation. Thus parallel multiply operations are performed. The results are detected by the PTR for output operation.

In a neural network model, two-dimensional patterns are usually converted into vector expressions because input/output patterns are expressed in vectors. In the system in question, however, two-dimensional patterns can be operated on as they are by a unique spatial coding method. The number of neurons is 16 while inputs and outputs are 16 (4×4)-dimensional.

Figure 4 shows an example of a memory matrix obtained through this system's learning process. In this example, the three patterns constitutes one instance of learning, and the recalling output shown is the result of applying a threshold operation to an analog output. Figure 4(a) shows an instance in which learning is still insufficient. It is seen that with the progress of learning, the memory matrix changes and the output comes closer to the correct pattern as shown in Figure 4(b).

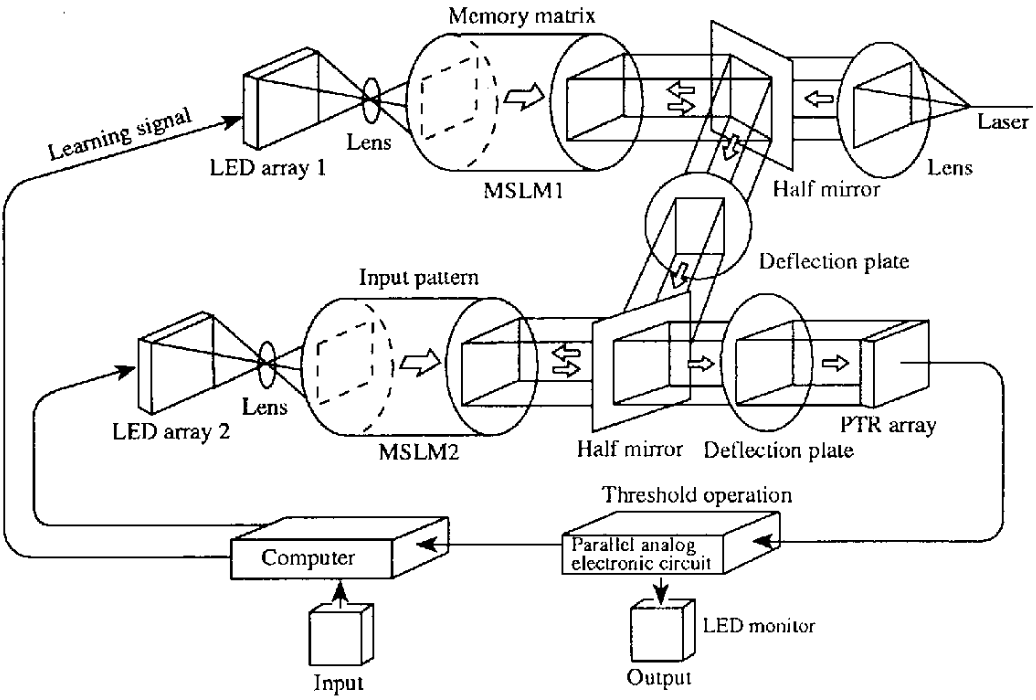

Figure 3 Optical Associatron

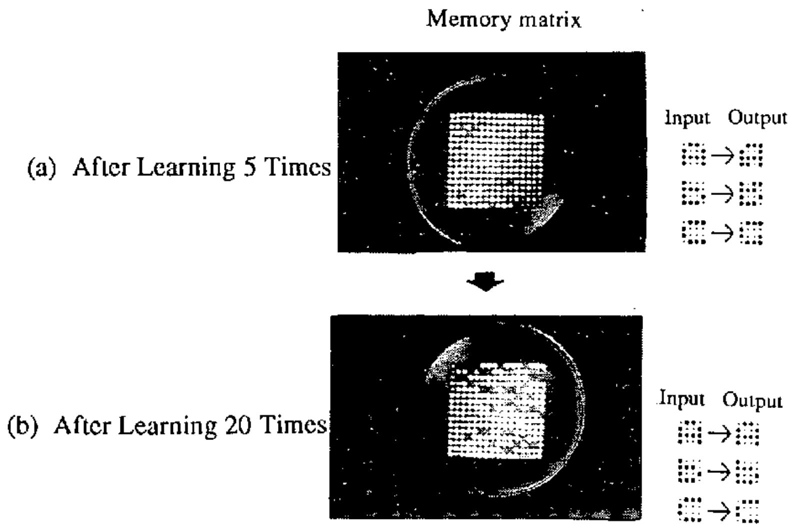

Figure 4 Convergence of a Memory Matrix

4.2 Combination with Primary Optical Processing

This optical associatron is a single-layer network. But when it is connected to the stage following conventional optical operations, it can give an increased flexibility to operation. As an example of such a system, Ishikawa et al. use an experimental system in which optical Fourier transformation system and the optical associatron are connected.[7] This combination makes it possible to set output patterns freely, and the Fourier transformation can make it shift invariant.

4.3 Significance of Learning

The significance of introducing such learning lies first in the fact that flexible processing can be attained as the processing circuit itself acquires a processing function instead of having its processes dictated by an outside program.

A second point of significance is that the nonuniformity of optical devices can be compensated for. Optical devices are accompanied by various forms of nonuniformity. The effects of such nonuniformity are reflected on output values. Therefore, the use of closed loop type learning, such as orthogonal learning, can compensate for such nonuniformity by the formation of memory matrices.

5. Conclusion

Optical neurocomputing has a large number of problems yet to be solved. Practical limits must be made clear with respect to spatial parallelism, higher speeds of operation, and downsizing. It is also an important task in the future to develop a new method of learning. To improve performance, it is vital to develop an effective method of learning that takes the features of optical computing into consideration.

References

1) C. Mead: Analog VLSI and Neural Systems, Addison-Wesley (1989)
2) A. Utsugi and M. Ishikawa: Construction of Inner Space Representation of Latticed Network Circuits by Learning, Neural Networks, 4-1, pp.81–87 (1991)
3) N.H. Farhat, et al.: Optical Implementation of Hopfield Model, Appl. Opt., 24-10, pp.1469–1475 (1985)
4) J. Ohta, et al.: GaAs/AlGaAs Optical Synaptic Interconnection Device for Neural Networks, Opt. Lett., 14-16, pp.844–846 (1989)
5) G.J. Dunning, et al.: All-Optical Associative Memory with Shift Invariance and Multiple-Image Recall, Opt. Lett., 12-5, pp.346–348 (1987)
6) M. Ishikawa, et al.: Optical Associatron: A Simple Model for Optical Associative Memory, Appl. Opt., 28-2, pp.291–301 (1989)
7) M. Ishikawa, et al.: Experimental Studies on Learning Capabilities of Optical Associative Memory, Appl. Opt., 29-2, pp.289–295 (1990)
8) M. Ishikawa, et al.: Optical Associative Memory Combining with Optical Preprocessing, Proc. Optical Computing '90, pp.160–161 (1990)

Current News

* NTT Enters Semiconductor Market

NTT started marketing LSIs for use with information systems for security protection through its subsidiary. UP until now, NTT entrusted contracts to outside makers to produce LSIs based on NTT's R&D results and procured the products for NTT's business. However, the corporation has decided to sell the LSIs for the purpose of business stability throughout NTT Group Companies.

The subsidiary is NTT Electronics Technology. For the present, the company will sell LSIs for security protection that have been developed by NTT's Network Information Systems Laboratories. Annual sales are expected to be around several hundred million yen. A Japanese maker will be entrusted with producing the LSIs on a commission basis, and a new sales department will be created in the subsidiary for sales to non-NTT customers.

Since its inception in 1982, NTT Electronics Technology has been responsible for development and installation of facilities for NTT's research departments, and also for the design and sales of LSIs for communications systems used in industry. Since 1990 the company has been selling semiconductors on a trial basis to a number of measurement instrument manufacturers. Current sales revenue is about 6 billion yen per year. The company plans to sell more than half of its semiconductors to non-NTT customers. NTT has high-level semiconductor technologies which are focused on the communications field. Therefore, it is thought that NTT's new entry into the outside sales market will have a major impact on the semiconductor industry both in Japan and the rest of the world.

* DEC and Seiko-Epson Reach Alliance Agreement in Small Computer Field

DEC and Seiko-Epson have reached an agreement in the field of small computers centering on PCs. For the time being, the two companies will go forward and jointly develop, market and sell products on the Japanese market, and plan to eventually establish a joint-venture company for new business purposes. After signing the formal agreement, Nihon DEC, a wholly-owned subsidiary of DEC, and Epson will organize a joint project team, and embark on joint product development in the area of PC-based small-scale network systems. This alliance is intended to mutually supplement the strengths in products and technologies between DEC, who is strong in system construction based on the minicomputer, and Epson who has expertise in low-cost PCs. They plan to begin marketing products using sales routes of the two companies halfway through 1992. There is also the possibility that these firms will form cooperative relationships overseas if this new alliance goes well.

* Japan, UK, Germany and France Cooperate in International Communications Service

British Telecom (BT), NTT, DBP Telekom of Germany and France Telecom (FT) have reached basic agreement for cooperation in the field of international communication services for business use.

The four common carriers will form a consortium by making a joint-equity investment in Syncordia, which BT created in 1991 as a subsidiary for international communications service in the US. Syncordia is a company specializing in communications services, principally for multi-national corporation customers. The specialized services include design, construction and operation of intra-company communications networks which connect foreign locations all over the world. Hereafter, the four common carriers will discuss the equity percentage, the scope of management participation, etc. However, it is forecasted that the final equity ratio will be 40% for BT and 20% each for NTT, DBP Telekom and FT. As overseas advances by companies have become more active, intra-company communications networks have gone global, and therefore this has become a promising market for carriers. Common carriers in countries have cooperated with each other in the area of international communications in past years. In expectation of intensifying competition in the future, large common carriers in Japan, the USA and Europe have begun groping for plans to cooperate in this area.

* NTT and KDD to Tie up for Overseas Business

NTT and KDD will cooperate in the field of overseas communications business. As communications liberalization progresses on a global scope, both NTT and KDD have been invited to participate in the projects of countries, who aim at the enhancement of their communications business through the introduction of foreign capital. These countries include those in Southeast Asia, Eastern Europe, and the Community of Independent States (CIS). Under these projects, corporations are often requested to manage communications operations for both the domestic and international services. As it has become necessary to have an operating organization managed by one company, NTT and KDD agree to create a new joint-venture company. They then form a consortium with trading companies, communications equipment manufacturers, etc., and participate in international tenders. The consortium will be responsible for focusing on consulting services, fund raising, equipment installation and communication business management of the new joint-venture company. NTT, who is eager to enter on the international communications field, is aiming at a full-scale advance into the international communications market, and is taking the opportunity to do so through the establishment of this new company.

* IBM Japan Closes Order with Sanwa Bank for Large-Scale Computer System

IBM Japan has closed an order with the Sanwa Bank for "Global System", a next generation overseas operations system that is scheduled to be operational in March 1993. Twelve IBM mainframe computers will be installed at six locations worldwide, including Tokyo, New York and London. In addition, 550 units of RS/6000 workstations with UNIX, about 600 units of PS/55 personal computers, etc. will be installed at about 50 locations in 27 countries worldwide. Workstations will be used for dealing and trading of foreign exchange, etc. The personal computers will be used for account settlements at branches.

Under the current system at Sanwa Bank, medium-sized computers are installed at major overseas locations, where account-related business and foreign exchange business are processed separately by location. However, as the processing volume has increased, the bank has decided to substantially increase volume capacity and to construct an integrated system that links overseas locations by network.

This is the first time that IBM Japan has successfully closed a UNIX-based system with a bank. The order amount is approx. 10 billion yen. The order for workstations is the largest-scale for the entire IBM group worldwide. If the system is completed, it will be the largest 24-hour integrated job processing system for a bank's overseas system.

Back Issues of Japan Computer Quarterly are as follows:

Published in 1992
- No. 88: Information-related Examinations in Japan

Published in 1991
- No. 87: Workstations in Japan
- No. 86: VAN Services in Japan
- No. 85: CIM in Japan
- No. 84: Laptop Computer in Japan — Market & User Strategies —

Published in 1990
- No. 83: Distribution Information Systems in Japan
- 82: Computer Security in Japan
- 81: Financial Information Systems in Japan
- 80: EDI in Japan

Published in 1989
- No. 79: Neurocomputers and Fuzzy Theory – R & D Trends in Japan –
- 78: Japan's Approach to Privacy Protection
- 77: State of CAL (CAI) in Japan
- 76: Software Industry in Japan – Striving for Increased Productivity –

Published in 1988
- No. 75: Personal Computers in Japan – An Unabridged Account –
- 74: Globalization of Telecommunication Services
- 73: The Microcomputer Industry – Training Engineers, Creating Applications –
- 72: Informatization – Handling Tomorrow's Problems Today –

Published in 1987
- No. 71: Systems Security – The Fight Against Computer Crime –
- 70: The Informatization of Small and Medium Businesses
- 69: Expert Systems in Japan
- 68: Large-scale Projects in Japan

Published in 1986
- No. 67: Information Services in Japan
- 66: IC Cards – Cards with Brains –
- 65: Database Services in Japan
- 64: Machine Translation – Threat or Tool –

Published in 1985
- No. 63: EDP Certification ExamLand, Japan –
- 62: Liberalizing Telecommunications
- 61: VIDEOTEX: A Glimpse of The 21 Century
- 60: The Day of the Robot

Published in 1984
- No. 59: Financial Revolution – Electronic or Plastic –
- 57: The PC Phenomenon
- 56: Information Services Japan '83

Exclusive Representative of Japan Computer Quarterly

ODS

ODS Corporation

TEL : TOKYO (03) 3400-7090
FAX : TOKYO (03) 3407-8035
TELEX : J26487 ODSTHINK
CABLE : ODSTHINKTANK TOKYO

Please send the ORDER FORM to:

GMS Div.
ODS Corporation
Dai-Ni Kuyo Bldg., 5-10-2, Minami-Aoyama
Minato-ku, Tokyo 107, JAPAN

ORDER FORM

Please send me the items checked below:

☐ Japan Computer Quarterly
Annual Subscription $ 105 (including air mail charge)

I would like to receive the following back copies:
No. ________ $ 27 per copy (including air mail charge)
____
____
____

☐ Bill me

Total: $ ________

Signature: ________ Date: ________

Name: ________

Position: ________

Company: ________

Address: ________

Telephone: ________ Fax: ________

JIPDEC

Japan Information Processing Development Center