OVERVIEW TO THE FIFTH GENERATION COMPUTER SYSTEM PROJECT

by Tohru Moto-oka

Department of Electrical Engineering,
The University of Tokyo, Tokyo, Japan

SUMMARY

Computers which have high performances for non-numeric data processing should be developed in order to satisfy and expand new applications which will become predominant fields in information processing of the 1990s.

Knowledge information processing forming the main part of applied artificial intelligence is expected to be one of the important fields in 1990s information processing and the dedicated computers for this have been selected as the main theme of the national project of the Fifth Generation computers.

The key technologies for the Fifth Generation Computer System (FGCS) seem to be VLSI architecture, parallel processing such as data flow control, logic programming, knowledge base based on relational database, and applied artificial intelligence and pattern processing.

Inference machines and relational algebra machines are typical of the core processors which constitute FGCS.

INTRODUCTION

The Fifth Generation computers are defined as the computers which will be used predominantly in 1990s. Supercomputers will be used in scientific and engineering calculations and simulations. Database machines and present mainframe computers will be networked in order to organise worldwide information systems. Many microcomputers will be used as system elements in various social systems. However, many computer industries are already earnestly developing these computers for future use.

Non-numeric data processing, including symbol processing and applied artificial intelligence will play more important roles than at present in the future information processing field. Non-numeric data such as sentences, speeches, graphs, and images will be used in tremendous volume compared to numerical data. Computers are expected to deal with non-numeric data mainly in future applications. However, present computers have much less capability in non-numeric data processing than in numeric data processing.

SOCIAL DEMANDS

The objective of this project is to realise new computer systems to meet the anticipated requirements of the 1990s. Roles that FGCSs are expected to play include the following:

1 To enhance productivity in low-productivity areas among non-standardised operations in the tertiary industries.
2 To overcome constraints on resources and energy by minimising energy consumption and control for optimisation of energy conversion efficiencies.
3 To realise medical, educational, and other support systems for solving ever more complex, multifaceted, social problems including, but not limited to, transition to an elderly society.
4 To contribute to international society and to help internationalisation of Japanese society through international cooperation, machine translation, and in other ways.

Computer application fields will be extremely diversified in the 1990s, and modes of utilisation will be widened to encompass everything from giant information networks with worldwide connections to system components. Everybody will be using computers in daily life without thinking anything of it. For this objective, an environment will have to be created in which a man and a computer find it easy to communicate freely using multiple information media, such as speech, text, and graphs.

TECHNICAL BACKGROUND

The changes from one generation to the next in computer technology have so far been made to accommodate changes in device technology. Such hindsight tells us, then, that there have been no major changes in the basic design philosophy and utilisation objectives of computers.

With Fifth Generation computers, however, the expected generational change is more like a 'generic change' which involves not only a change in device technology, to VLSIs, but also simultaneous changes in design philosophy and in fields of application.

The design philosophy behind conventional von Neumann computers was based on using a minimum of hardware to configure systems of maximum simplicity, capable of efficient processing using adequate software, because in von Neumann's day hardware was expensive, bulky, short-lived, and consumed a lot of power. From this viewpoint, the stored-program, sequentially-controlled systems were superior and high speeds and large capacities were pursued for economic reasons, resulting in the emergence of today's giant computers.

The key factors leading to the necessity for rethinking the conventional computer design philosophy just described include the following:

1 Device speeds are approaching the limit imposed by the speed of light.
2 The emergence of VLSI reduces hardware costs substantially, and an environment permitting the use of as much hardware as is required will shortly be feasible.
3 To take advantage of the effect of VLSI mass production, it will be necessary to pursue parallel processing.
4 Current computers have extremely poor performance in basic functions for processing speeches, texts, graphs, images and other non-numerical data, and for artificial intelligence type processing such as inference, association, and learning.

Computers, as the name implies, were designed as machines to perform numerical computations. Computer applications have, however, expanded without major changes in design philosophy, into such fields as control systems, processing of multiple information media, database management systems, and artificial intelligence systems. However, for computers to be employed at numerous application levels in the 1990s, they must evolve from machines centred around numerical computations to machines that can assess the meaning of information and understand the problems to be solved. For this evolution, in the immediate future, the following will be required:

1 Realisation of basic mechanisms for inference, association, and learning in hardware, making them the core functions of the Fifth Generation computers.
2 Preparation of basic artificial intelligence software in order to fully utilise the above functions.
3 Advantageous use of pattern recognition and artificial intelligence research achievements, in order to realise man/machine interfaces that are natural to man.
4 Realisation of support systems for resolving the 'software crisis' and enhancing software production.

In order to solve the problems faced by current computer technologies, achievements in related technologies such as VLSI technology, software engineering, and artificial intelligence research will have to be integrated, and the interim achievements of this project will be fed back to these technologies so that they can continue to advance.

RESEARCH AND DEVELOPMENT TARGETS

Knowledge information processing systems (KIPS) will be investigated in the Japanese national project.(1)(2) They will be based on innovative theories and technologies, and hence capable of accommodating such functions as intelligent conversation functions and inference functions employing knowledge bases that will be required in the 1990s. The functions of FGCSs may be roughly classified as follows:

1 Problem-solving and inference.
2 Knowledge-base management.
3 Intelligent interface.

These functions will be realised by making individual software and hardware systems correspond. A conceptual image of the system is shown in Figure 1. In this diagram, the modelling (software) system is the project's ultimate target for software development, and the machine (hardware) system the ultimate target for hardware development. The upper half of the modelling system circle corresponds to the problem-solving and inference functions, the lower half to the knowledge-base management function. The portion that overlaps the human system circle on the left corresponds to the intelligent interface function. The diagram also illustrates that the intelligent interface function relies heavily on the two former groups of functions. This diagram shows how the emphasis in computer systems will have shifted decisively towards the human system by significant enhancement of the logic level of the hardware system and by the positioning of the modelling system between the hardware and man.

The interface between the software and hardware systems will be the kernel language. The entire software system will be realised in the kernel language, and the hardware system will directly execute the kernel language.

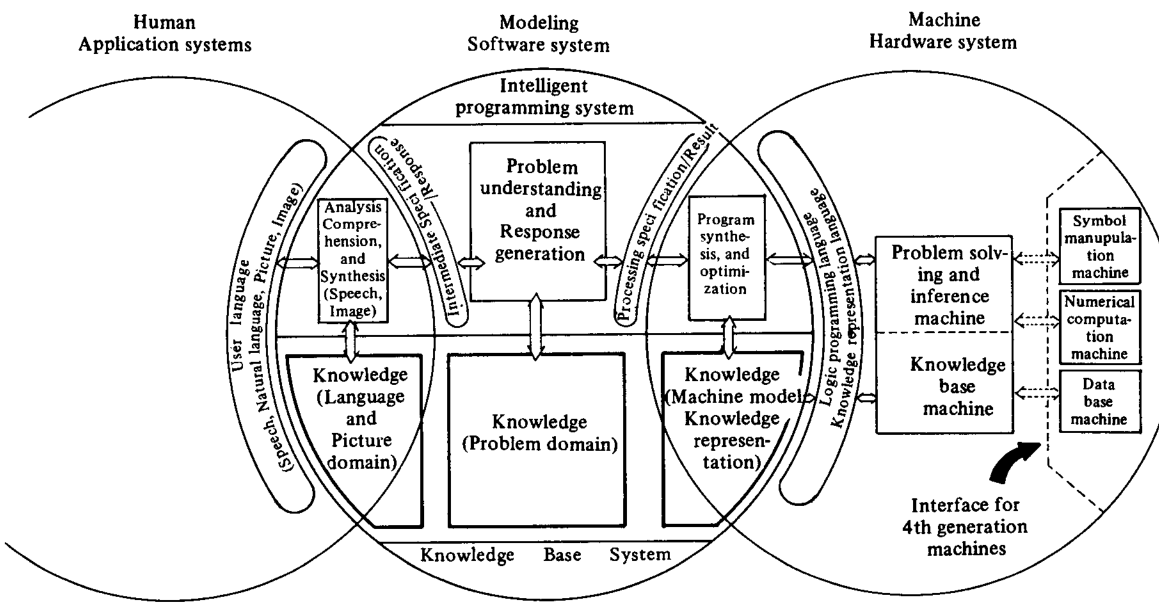


Fig. 1. Conceptual diagram of a Fifth Generation Computer System as viewed from the standpoint of programming

THE PROBLEM-SOLVING AND INFERENCE SYSTEM

Research into the problem-solving and inference mechanism will concentrate on the problem solving and inference functions, the ultimate goal being a cooperative problem-solving system. In such a system, a single problem will be solved by two or more problem-solving systems cooperating with each other.

It will be necessary to develop high performance inference machines capable of serving as core processors that use rules and assertions to process knowledge information. Existing artificial intelligence technology has been developed to be based primarily on LISP. However, it seems more appropriate to employ a Prolog-like logic programming language as the interface between software and hardware due to the following considerations: the introduction of VLSI technology made possible the implementation of high level functions in hardware; in order to perform parallel processing, it will be necessary to adopt new languages suitable for parallel processing; such languages will have to have a strong affinity with relational data models.

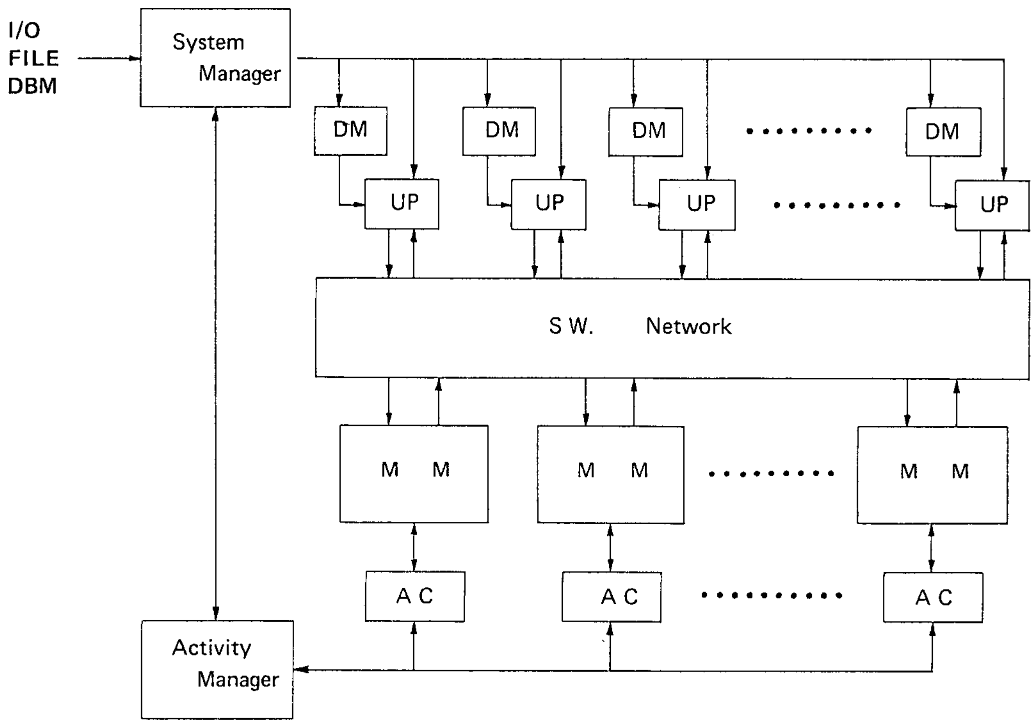
The Fifth Generation Kernel Language (FGKL) has been defined as a language which will determine the interface between the hardware and software of Fifth Generation computers.

The maximum capability of the inference machine is from 100 MLIPS to 1 GLIPS. The inference executing speed, 1 LIPS (logical inference per second), denotes one syllogistic inference operation per second. One inference operation, executed by a current computer is believed to require 100 to 1000 steps; thus, 1 LIPS corresponds to 100 to 1000 instr/sec. The current generation machines are rates at 10^4 to 10^5 LIPS.

To realise such performance capabilities, the essential research and development will concentrate not only on speeding up the basic devices, but also on high-level parallel architectures to support the symbol processing that is the key to inference. A hardware architecture suited to the new parallel inferences based on the data-flow control mechanism will be researched and developed.

As an example of proposed architectures in a parallel inference engine, the parallel processor investigated at the Univ. of Tokyo, which is able to execute Prolog-like language directly, is shown in Figure 2. The main components of this inference engine consist of numerous unify processors with definition memories and numerous memory modules with activity controllers. They are connected through switching networks. In addition, there are a system manager which enables communication with external devices and unify processors by serving as an interface and an activity manager which supervises the activity controllers.

The target scale for the hardware encompasses ultimately about 1000 processing elements, and the requisite VLSI manufacturing technology for such hardware will be researched and developed.



UP: Unify Processor
DM: Definition Memory
MM: Memory Module
AC: Activity Controller

Fig. 2. System Organization of Parallel Inference Engine at the Univ. of Tokyo

THE KNOWLEDGE-BASE SYSTEMS

The intention of software for the knowledge-base management function will be to establish knowledge information processing technology, where the targets will be development of knowledge representation systems, knowledge-base design and maintenance support systems, large-scale knowledge-base systems, knowledge acquisition experimental systems, and distributed knowledge management systems. These systems will then be integrated into a cooperative problem-solving system. One particularly important aim will be semi-automated knowledge acquisition, that is, systems will be equipped with a certain level of learning functions.

For the knowledge-base management function, relational database interfaces and consistency testing functions will have to be realised in the kernel language.

Research into and development of the knowledge-base machine will aim at developing a hardware mechanism that fulfils the demands for knowledge representation systems and large-scale knowledge-base systems, and is capable of efficiently supporting storage, retrieval, and renewal of a large volume of knowledge data. This mechanism will ultimately be integrated in the prototype FGCS.

Regarding the target knowledge-base management function in research into and development of the knowledge-base machine, the aim in performance capabilities based on a database machine with a 100 to 1000 Gbyte capacity, will be to retrieve the knowledge bases required for answering a question within a few seconds.

To realise such performance capabilities, a parallel architecture, capable of speedily supporting the symbol processing function intended to handle a large capacity data management function and knowledge data, will be indispensable. Research and development will be conducted for a parallel processing hardware architecture intended for parallel processing of new knowledge bases, and which is based on a relational database machine that includes a high-performance hierarchical memory system, and a mechanism for parallel relational operations and knowledge operations.

The knowledge base system is expected to be implemented on a relational database machine which has some knowledge base facilities in the Fifth Generation Computer System, because the relational data model has a strong affinity with logic programming. Relational calculus has a close relation with the first order predicate logic. Relational algebra has the same ability as relational calculus in the description of a query. These are reasons for considering a relational algebra machine as the prime candidate for a knowledge base machine.

The relational algebra machine GRACE is being investigated at the Univ. of Tokyo. GRACE adopts a relational algebra processing algorithm based on hash and sort, and can join in O((N+M)/n) time where n is the number of processors and N and M are the cardinalities of two joined relations. GRACE is a typical example of proposed arthitectures of knowledge base machines in the FGCS Project.

The clustering feature of the hash operation can reduce the load of join processing. The simple join algorithm takes time proportional to the product of two relations' cardinalities. However, if two relations are clustered on the join attribute, that is, the tuples are grouped into disjoint buckets based on the hashed value of the join attribute, there is no joining between tuples from buckets of different ids (hashed value of the join attribute). The duplicate elimination task in projection also used to be a big burden in relational data base systems. The above approach can be applied to projection operations quite as well as to join operations.

Another technique utilizing hash is the "Joinability Filter" which can be found in CAFS and LEECH[3].

The two approaches discussed above are mutually independent, so

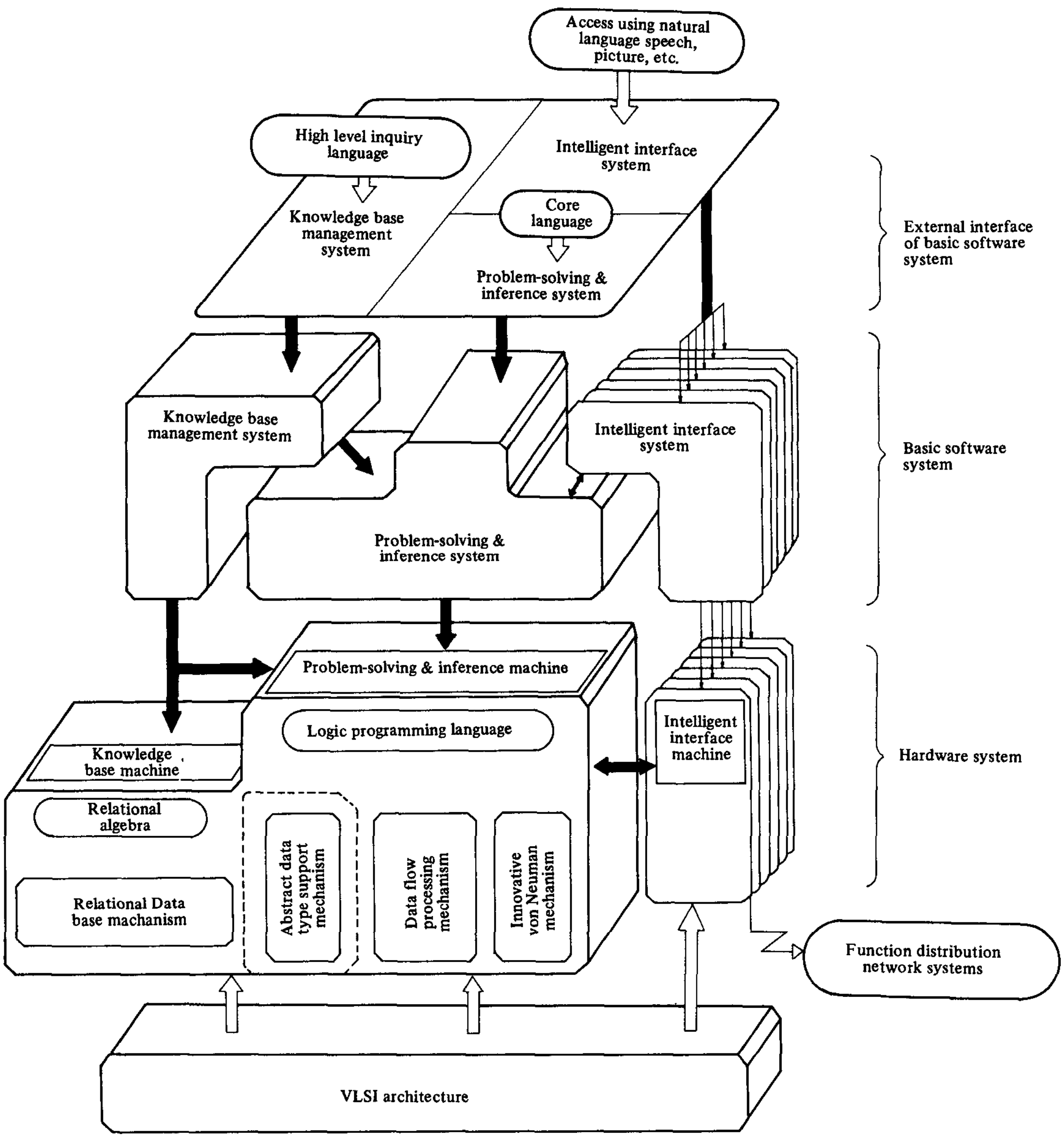


Fig. 3. Basic configuration of the Fifth Generation Computer Systems

that both joinability filter processing which decreases the candidate tuples and clustered processing can be integrated together.

The great reduction of processing load for join and projection operations was realised via the hash based relational algebra processing method.

THE INTELLIGENT INTERFACE SYSTEM

The intelligent interface function will have to be capable of handling man/machine communication in natural languages, speeches, graphs, and images so that information can be exchanged in a way natural to a man. As natural language processing provides the basis for translation, English and other languages are to be included as well as Japanese in the objects for processing. Ultimately, the system will cover a basic vocabulary (excluding technical terms) of up to 10,000 words and up to 2,000 grammatical rules, with a 99% accuracy in syntactic analysis. The fewer the grammatical rules, the higher the system capabilities.

On speech processing, speech input and output systems will be developed. The object of speech inputs will be continuous speech in Japanese standard pronunciation by multiple speakers, and the aims here will be a vocabulary of 50,000 words, a 95% recognition rate for individual works, and recognition and processing three times the real time of speech, though this may vary somewhat depending on hardware capabilities. As for processing of graphs and picture images, the target system to be developed will be capable of structurally storing roughly 10,000 pieces of graph and image information and utilising them for knowledge information processing.

Meanwhile, there will also be research into and development of dedicated hardware processors and high-performance interface equipment for efficiently executing processing of speech, graph, and image data. Furthermore, methods for exchanging information that are natural to man, through parallel utilisation of such multiple media data, will have to be established.

BASIC CONFIGURATION OF THE FGCS

The software and hardware systems that realise the three functions mentioned above will be coupled to form a general-purpose machine. Its conceptual structure is shown in Figure 3.

Because, in actual use, a variety of performance capabilities will be required of each of the three functions, the configuration will have to be flexible enough to provide not only the general-purpose machine, but also various system configurations to accommodate the various performance capabilities required by individual applications, that is, specific function-intensive machines in which some functions are enhanced.

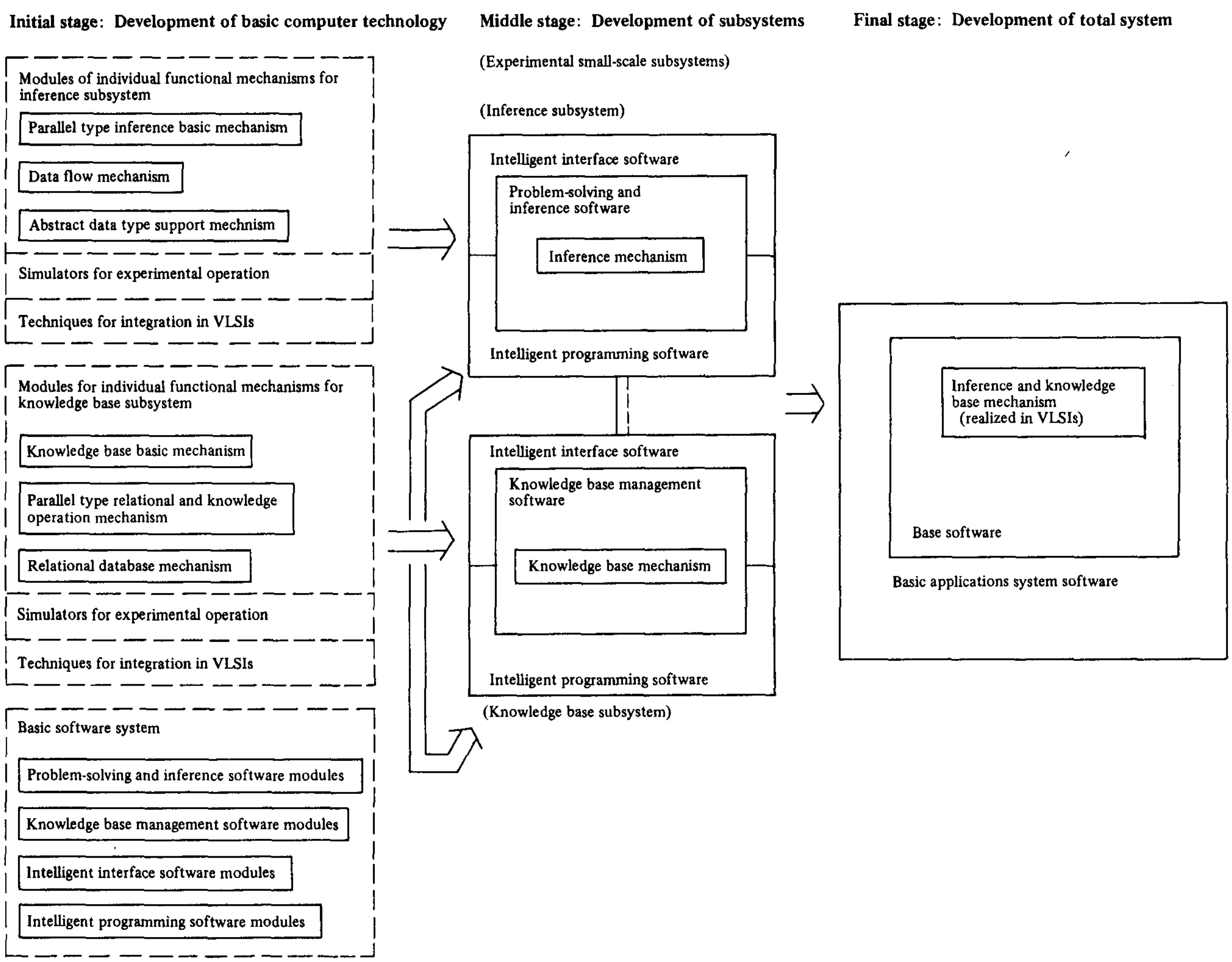


Fig. 4. Stages of Fifth Generation computer research and development

These machines will have the Fifth Generation computer kernel language as their common machine language, and they will be interconnected in networks to form a distributed processing system.

BASIC APPLICATION SYSTEMS IN FGCS

Several basic application systems will be developed with the intention of demonstrating the usefulness of the FGCS and the system evaluation. These are machine translation systems, consultation systems, intelligent programming systems and an intelligent VLSI-CAD system.

CONCLUSION

The research and development targets of the FGCS are such core functions of knowledge information processing as problem-solving and inference systems and knowledge-base systems that cannot be handled within the framework of conventional computer systems.

There is no precedent for this innovative and large-scale research and development anywhere in the world. We will therefore be obliged to move toward the target systems through a lengthy process of trial and error, producing many original ideas along the way.

Since the establishment under this project of the basic computer techniques that will be required in the 1990s is vital, plans will have to encompass as wide an extension of basic techniques as possible. Taking these aspects into consideration, the research and development period set for this project is 10 years, divided, as shown in Figure 4, into an initial stage (three years), a middle stage (four years), and final stage (three years).

As a research and development project involving advanced technologies from many fields, as well as from the viewpoint of its being an international contribution by Japan, the Fifth Generation Computer Project should be promoted through some form of international cooperation. Various formats are conceivable for such international cooperation buy, based on experience gained during the past four years of basic technological development, the research and development work will be served best by extremely close interrelations among the various development themes. Thus a form of cooperation that comes readily to mind is for interested governments or enterprises to promote original research and development at their own expenses, periodically exchanging achievements to the mutual benefit of all involved.

The Japanese national project of FGCS started in April, 1982 and is expected to run for 10 years. The initial step is the preliminary three-year stage in which the project is dealt with by the Institute of New Generation Computer Technology (ICOT).

REFERENCES

(1) T. Moto-oka (Ed.), Fifth Generation Computer Systems, North Holland, 1982.

(2) P.C. Treleaven, I. G. Lime, Japan's Fifth Generation Computer systems, IEEE Computer, Vol. 15, No. 8, Aug. 1982, 79-88.

(3) D. R. McGregar et al., High Performance Hardware for Database System, Systems for Large Data Bases, North-Holland, 1976, 103-116.