HOP, STEP, and JUMP

Kazuhiro Fuchi

Director
ICOT Research Center
Tokyo, Japan

One of the objectives of FGCS'88 is for us at ICOT to present the results of the middle stage of research being conducted in the fifth generation computer project.

Another objective of this conference is for researchers from all over the world to exchange the results of their research, aiming towards a new generation of computers.

Our fifth-generation project started in 1982. This year is its seventh year, and the end of the middle stage. The previous FGCS conference, FGCS'84, was held at the end of the project's third year, four years ago. The three stages of the project can be represented as three stages of movement : hop, step, and jump. We have already hopped and stepped, and now we are about to jump into the final stage of the project.

I would like to outline the FGCS project briefly. It is proceeding according to plan, along the lines that I envisaged at the beginning of the project, and research is yielding results at the pace expected at the beginning of the project.

It goes without saying that these results could never have been obtained without the concerted efforts of the researchers involved in this project, and without the support and cooperation of the many people around ICOT.

The efforts and cooperation of all these people are bound up with the results, but the excellent project set-up and basic policies have made it possible for everyone to do their best.

I am sure that I have said this many times before, but the key words that express the technological features of this project are parallel inference. I will go into these key words in more detail later, but, briefly, parallel inference was the objective of the project from the start. This objective was not changed in the initial or middle stages. This research has formed a base for the final three years of this project, and the last stage will see the establishment of a basic parallel inference technology, as envisaged whenthe project was set up.

The initial stage, or "hop" stage, was from 1982 to 1984. It can be thought of as the personal sequential inference machine, or PSI, stage. The research results can be seen in the technical reports and papers of this period, but it is easier to understand by looking at examples.

The PSI is a sequential computer that uses a logic language as the machine language. It was designed as a tool and workstation for software research in the middle and final stages.

In addition to the hardware, the sequential inference machine programming and operating system, or SIMPOS, was being developed as the operating system on the PSI. It was the first operating system and the first large-scale software built using a logic language.

The PSI and SIMPOS are representative results of the initial stage.

In the initial stage, research tools were developed, fundamental software problems were researched based on the concept of logic programming, and progress was made in research on natural languages. Other basic research was also conducted. The initial stage can be roughly thought of as such age.

The middle stage, from 1985 to this year, was the "step" stage. One of the features of research during this period was the start of research oh parallelization. On the language side, Guarded Horn Clauses, or GHC, was proposed. It acted as a springboard for research on parallelization. GHC is a logic language with parallel operation added.

The problem of how to add parallel control to a logic language has been a major item in logic programming research worldwide over the last few years. ICOT has contributed greatly to this field by proposing GHC, which formed the base for fundamental research on parallel programming.

With GHC as the base, KL1 was decided on as the kernel language in the first plans. The kernel language is the software base, and is also the starting point for machine architecture. Research on parallel architecture started, based on KL1.

One of the results of the middle stage was the Multi-PSI. The Multi-PSI uses an upgraded version of the PSI developed in the initial stage. Sixty-four PSIs form one Multi-PSI system. The purpose of developing it was to provide an environment for research on parallel software as quickly as possible. To achieve this purpose, a prototype of the PIMOS, the operating system for parallel inference machines, are being developed on the Multi-PSI.

In parallel with the Multi-PSI, a 100 processor element parallel inference machine, or PIM, system is being developed. This PIM is the middle stage version of the target parallel inference machine, and is expected to be completed during the first half of next year. It will be linked with research in the final stage.

Another feature of the middle stage is research on software based on the PSI and SIMPOS developed in the initial stage. The PSI and SIMPOS have been upgraded. There are now more than 300 PSIs, which are being used as workstations for software research in this project.

One of the main themes in this project is researching logic programming in depth. Parallel programming and constraint programming are major research themes all over the world. ICOT is also researching constraint programming, with the aim of integrating it with parallel programming.

In the framework of logic programming, the development of meta-programming, program transformation, and partial evaluation have been in progress since the initial stage. In the middle stage, parallelization has also been emphasized, based on GHC.

Research on natural language understanding is part of core research on artificial intelligence. It will be essential in future man-machine communication. It has been a major part of research at ICOT since the initial stage. Research on models based on situation semantics theory is also under way.

One result of natural language research in the middle stage is DUALS3, Discourse Understanding Aimed at Logic-based System, a language understanding system that has been upgraded from the system developed in the initial stage. The logic programming library developed through research during this period is the language tool box, or LTB, for language processing.

Another part of core research on artificial intelligence is knowledge processing. Testing expert systems in various areas is another feature of the middle stage. ICOT is researching knowledge processing not for immediate practical application, but in order to pioneer high-level techniques and to prepare for parallelization. This research will help us to work on knowledge acquisition, induction, learning, hypothetical reasoning, and non-monotonic reasoning.

Compared to the initial stage, research in the middle stage has been more active, and results have diversified. In a broad sense, this is a preparation for integrating and developing the basic concepts held from the beginning in the final stage.

Next year will see the beginning of the final stage, the "jump" stage. This will be the period for in-depth research on parallel inference software. Until now, parallel software has been researched by simulation, but now that a new and powerful research environment has appeared in the form of the Multi-PSI, we hope that research will move faster.

On the hardware side, the target is to build a 1000 processor element parallel inference machine, integrating a knowledge base machine architecture.

By using this machine to the full, the abstract and theoretical research that has been done until now, such as research on resource management, will give birth to specific problems.

The parallelization of knowledge processing, natural language processing, and application programs of a reasonable scale will be a major area of research in the final stage. This is an area on which almost no work has been done as yet.

To realize this research, systemization is necessary, along with research that returns to the basics. At the same time, it is essential to set up a parallel programming environment.

One of the software targets of the final stage is to group parallel programming techniques in the form of knowledge programming, using the environmental conditions provided up to the middle stage.

As I have shown, the project has been developing with parallel inference as the warp of its material, because we expect parallel inference to be the core of future information processing, and because we believe that it will lead to a new type of computer.

One of the assumptions in planning this project was the prediction that the future of information processing and the basic direction of its progress lie in knowledge information processing, usually thought of as artificial intelligence. The situation in 1982, when the project started, and the situation now are completely different, because artificial intelligence is becoming popular all over the world, and is becoming an integral part of society.

However, we have emphasized from the beginning that this project is not an AI project as such. To put it more accurately, the aim of the project is new computers for future AI. The aim of this project is not AI systems that can be used in the near future, nor is it computers that are an extension of present computers.

To diverge, I believe that AI should be viewed as at least three dimensions. Like height, width, and depth in space, AI should also be the volume obtained by multiplying these three dimensions.

The present AI boom is mainly width. The width has been increased as applications for solving problems have been found. Not long ago, there was almost no width in AI, but as AI became more popular, it broadened.

Depth is the depth of research. Research on artificial intelligence has been going on for 30 years, but how far have we managed to clarify what intelligence is? How good are our models of intelligence? This is the scientific side of artificial intelligence.

An example of the height of artificial intelligence is the realization of AI functions and the system with which to realize them. This is the technical dimension.

If obtaining a good volume is a problem, it is necessary to increase each dimension. The width of applications has increased, which is valuable, but is that all that is necessary? It is said increasingly often that the importance of basic research on AI lies in the overall, healthy growth of AI in the dimension of width.

We also believe that the third dimension, height, is as important as the other dimensions. Our project emphasizes the third dimension.

Artificial intelligence systems are multi-layer structures, where the computer, or hardware, is the base, and there are many layers of software. Is the present system form suitable? It is said that knowledge is important, and so is software. This is true, but this does not mean that hardware is not important.

There is still room for growth in present computer technology, and it is certainly nearing maturity. The world of information processing, however, is expected to keep on growing. This poses the question of whether an extension of present computer technology can support growth.

The progress of electronics technology, supporting computers, is truly remarkable. It shows that electronics technology is starting to have enough power even if computers cast off their current form.

If we review computers from the viewpoint that AI and knowledge information processing are the future direction of information processing, it is evident that the structure of today's computers is not ideal. We must therefore review the principles of making computers from a new viewpoint. The starting point of our project was to show that the time for this is coming.

Taking the view that a healthy body houses a healthy mind, we need new types of computers to form a stronger base for future, more advanced information processing. The technical potential is now apparent.

From the point of view of hardware, the guiding principle of parallelization to give more powerful, higher performance has appeared in ultra VLSI.

On the software side, from the viewpoint of high-performance software such as AI, logic programming comes to mind. The basic operation is the inference in logic.

There is an internal relationship between hardware and software by taking parallelization as the form in which inference will be realized. The phrase "parallel inference" is the guiding principle and the key words for new computers.

However, I am not claiming that parallel inference is the only research required for future information processing.

Research on artificial intelligence is spreading throughout society. It is impossible to handle all artificial intelligence in a single project, and, indeed, the time has come when it is better to spread it over many projects. Each project should be formed by placing its strong points on some place in the three dimensional space : height, width, and depth. Our national project has tried to do this, and one of the features of our project is the establishment of basic technology for new computers, not the extension of existing technology.

This is one of the reasons why I have emphasized a little while ago that this project is not an AI project as such.

I said that the increase of artificial intelligence will be a major pointer of the future direction of information processing. This is not the whole story. It is an important trend, but it is more reasonable to consider that the world of information processing is wider than that. There are, however, some areas that are not yet considered as artificial intelligence, particularly in the narrow sense. Systems with more advanced functions and better performance are also essential in these areas.

One major problem that comes to mind here is how to create large-scale, complex, advanced software. It can be said that the desire to use scientific methodology in creating software is the mainstream of computer science.

The concept of logic programming can be looked at from this point of view. Automatic programming is a dream. To realize this dream, we need to identify the logical properties of programs. To develop program transformation and other techniques, however, we should build programming methodology and programming languages on clearer logic. This means rebuilding computer science. Logic programming can be viewed as a possible approach.

Based on this, inference can be taken as the basic concept in building software. If we view parallelization as a method of high-speed realization, parallel inference is necessary to advance future information processing and to give it meaning.

I keep on talking about parallel inference because it's the main theme of our project.

As I said at FGCS'84, it is vital to have simple, clear basic concepts in major long-term research projects, and to keep them in mind right through the project. In the period of transition between the initial stage and the middle stage, we planned development this way.

Even basic concepts must be reviewed periodically. We take great care in doing this, and, so far, have not had to change our basic concepts. In the seven years that research has been developing in this project, I feel that we have verified the initial basic concepts again and again. Of course, we now have many more ideas than we did at the start, and these have been reinforced. As time has passed, we have gained more confidence in our approach.

This afternoon, there will be a more detailed report on plans for the final stage. Our basic approach to these plans has been to re-verify and clarify the basic concepts since the initiation of the project. We have been working on a central theme and improving our plans. We are trying to simplify our plans for the final stage.

To digress again, I believe that there are several things we must take note of if we want a national research project like this to succeed.

The first is the temptation to compromise.

All over the world there is a tendency to lean towards practical topics in order to establish support for a project or to show its success. If the project is set up for practical topics or the research goes much better than expected, this is all right, but in many cases, this tendency means that the initial ideals and targets must be lowered. This in turn lowers the researchers' morale, and means that results that could be attained with a little more hard work are not achieved.

Topics that show a likelihood of practical application are usually best left to the spontaneity and creativity of a private company. Of course, this depends on the nature and scale of the research theme.

In a national project, especially in an industrialized country, not the extension of existing technology, but the creation of new technology, that is, work with risks, is proper.

Another temptation is to set targets too high. There is a natural limit on targets, imposed by time and human resources. There have been cases where the target was too high, or was misinterpreted as being too high, causing an adverse reaction, and damping research.

Targets that are too high are often caused by the naive optimism of researchers. There are cases where dreams or things that seem good are popular, and, whether consciously or unconsciously, the researchers go along with those ideas. These cases are good examples of disguised compromise.

Another problem is the danger of fearing danger. Fear of danger is the biggest incitement to compromise. Fearing danger, compromising time and again, and trying to create new things are certainly contradictory, and may be unattainable targets. However, it is no exaggeration to say that the greatest danger in creative research is compromise.

Considering that many people follow the principle of safety today, surely it is better to follow the principle of adventure.

As our project nears its final stage, we are often asked about its conclusion. Of course, we would like a secure conclusion, but it is not easy to wind up, using the remaining time well.

Rather, we must work on the problem of parallel inference. I like to think of the last seven years as a preparation period for the next three years, in which we start a new project.

This may provoke criticism or questions as to whether we are trying to deal with an impossible problem or trying to deal with the problem in too short a time.

This problem is a huge challenge, even if we limit it to parallel inference. The words are simple, but they embody high-level, diverse research themes. It may be impossible to solve all the problems in that time, but I believe that we are very likely to break through at a basic point.

My basic, optimistic view is that if the problem is correctly set and a good environment is given, researchers can be stimulated far more than expected. The past seven years are the preparation and efforts towards doing that.

The nature of our project is not such that the second half will end in total success, nor that we are embracing impossible challenges.

I have talked mainly about our project today, but research towards a new age of computers and a new age of information processing is going on all over the world. Our project is one of many, and can only be one of many, in this worldwide movement. This is not an expression of Japanese modesty.

We want to contribute to making new history, and are working very hard to do so. Our efforts are part of the vast flow of history, in fact, part of historical inevitability.

Research must take many approaches. In this field, these approaches will bear fruit not by simple diffusion, but when unity can be seen in their diversity. The simultaneous pursuit of diversity and unity is vital for the development of research, and the desirability of open exchange of research by various groups is bound up with it.

We have stressed international exchange and co-operation from the beginning of our project, because we believe that there are people all over the world working towards making history in research in this field, that their numbers are increasing, and that history will be made by the joint efforts of all these researchers. We believe that this is necessary for the future of all mankind.

To realize this, not rivalry between researchers, but co-operation founded on understanding of different approaches is needed. At no time have we needed this as much as now.

This conference is a link in the chain of activities aiming towards the future. We will hear both the interim report of FGCS activities and reports of the research activities of our colleagues all over the world. I would like to thank you all for attending the FGCS'88 conference, and to express my hope that your research will bear fruit through this exchange, and that this conference will be a stepping-stone whereby history can advance towards a new age.