Setting Copy
KJ

7. THE FIFTH GENERATION II

The Japanese Fifth Generation Project aims to produce computer hardware and software for applications of knowledge engineering. Expert systems, natural language understanding by machine and robotics are examples. In his speech to the Fifth Generation conference, Fuchi coined the phrase "knowledge information processing" as "an extended form of knowledge engineering." He adds, " This, it is thought, will represent the form of information processing in the 90's."

The systems built by the knowledge engineers consist of three main parts. First, there is the subsystem that "manages" the knowledge that is needed for problem solving and understanding. Second is the problem solving and inference subsystem that discovers what knowledge is useful and relevant to the problem at hand, and with it constructs--step by step--a line of reasoning leading to the problem solution, the plausible interpretation, the best hypothesis. Since knowledge -based systems usually provide assistance to human endeavor, they must also include methods of interaction between person and machine, in modes and language that are comfortable and "natural" to the user. Ordinary human natural language is often preferred, but the stylized notations of some fields like Chemistry are more "comfortable" to users. Knowledge base management, problem solving and inference, and human interaction--these are the components of every Expert System.

The Fifth Generation plan organizes the work in just this way. For each component subsystem there is a hardware level and a software level. Between the levels the designers must define a "language" with which the software and the hardware interact.

The knowledge in the Knowledge Base must first be represented in symbolic form, and in memory structures, that can be used by the Problem Solver. The representation can take many forms. One common form is the "Object", a cluster of attributes describing something. An Object is usually associated with other Objects by symbolic references ("links") in the memory. A typical kind of associative network is the taxonomy (known as the is-a hierarchy). "A sparrow is-a kind of bird". Both sparrow and bird are Objects. If the Knowledge Base comes to know that birds can fly, then the Knowledge Base Management System must automatically propagate the little deduction that sparrows can fly. It must also be able to handle the exception cases it is told about, for example as regards penguins.

Another common and useful representation is the rule. A rule consists of a collection of statements called the "IF" part and a conclusion or action-to-be-taken, called the "THEN part. "IF the fog ceiling is below 700 feet and the official weather forecast calls for no clearing within the hour THEN landing is dangerous, will violate air traffic regulations, and diversion to a neighboring alternate field is recommended." To find out if a rule is relevant to reasoning, the problem solver must scan over the "IFs". But that can be an immense task if the knowledge base is large. Here again, the Knowledge Base Management System is put to work to organize the memory so as to reduce the amount of information processing, or at least do things in parallel to speed the search.

In the Fifth Generation plan, knowledge is to be stored electronically in a large electronic file known as a relational data base. The job of automatically updating the knowledge in the file and in organizing appropriate searches for relevant knowledge is performed by the knowledge base management software. The interaction between the hardware file and the software file manager is handled with a type of logical language called a relational algebra.

The prototype of the knowledge base subsystem will handle a modest knowledge base--thousands of rules and thousands of objects--about the size needed for current Expert System applications. An object is to be given a thousand characters of file storage. The ten year goal calls for handling knowledge bases with tens of thousands of rules and one hundred million objects. What is the scope of so much knowledge? An American firm interested in the representation of large bodies of knowledge has estimated that the Encyclopedia Brittanica can be represented in about one hundred million objects.

Knowledge is used as the basis for reasoning by a KIPS, but is not itself sufficient to discover lines of reasoning. Piecing together an appropriate line of reasoning leading to the solution of a problem or the formation of a body of consultative advice is the job of the inference process and the problem solving strategy that employs it. Inference processes can be very much of the common-sense sort, in which relevant knowledge is simply chained. A syllogism (IF x implies y and IF y implies z, THEN x implies z) is an example of such an inference process. Inference processes have been studied by logicians and mathematicians for centuries and many different procedures for inference are known. From this toolkit, artificial intelligence uses routinely just a few. Some methods allow for reasoning "inexactly" from knowledge that is uncertain. One, constructed on a foundation of mathematical logic in the 1960s by the logician Allan Robinson, is subtle, nonintuitive, and especially suited for computer processing (the method with all of its parts is called "resolution").

An inference process is the tool of some problem solving strategy. For example, the strategy for one kind of problem might be to reason backwards from some goal to be achieved to some existing conditions that are known to hold or some present resources that are already available. Imagine how you might reason backwards from the knowledge that you want to be on a particular airplane flight that departs at noon to a good guess about what time to depart from your home or office.

The Fifth Generation project envisions computer hardware engineered to be a logic processor, analogous to well-engineered capabilities of earlier generations of computers to be arithmetic processors. We presently speak of computer capabilities in terms of millions of arithmetic operations per second. The Fifth Generation planners speak of millions of logical inferences per second; by a logical inference they mean the equivalent of a step in a syllogistic line of reasoning.

All scientists and engineers must place their bests for the future on what is known in the present. The Japanese KIPS planners are betting on resolution as the best inference method for which to target their logic processing hardware. This is an approach that has great credibility in Europe, where a computer programming system called PROLOG (for "logic programming") that overlays resolution-based inference was developed (invented in France, polished in Britain). PROLOG has been chosen as the language of interaction between the logic processing hardware and the software that implements the various problem solving strategies (that is, PROLOG is the "machine language" of the logic processor).

For the inference subsystem, the initial milestone is a one-user workstation capable of performing one million logical inferences per second. It is intended to be both a prototype for later development and an intermediate product that may be on the market by the 1985. This prototype would give a factor of ten in performance improvement over software-based PROLOG implementations on today's common mainframe computers like the Digital Equipment 2060. Extraordinarily ambitious is the final target. It aims at an inference supercomputer that can do a hundred million to a billion logical inferences per second. Such extraordinary speed can only be achieved by the insightful use of a great deal of parallel processing in the computer--major advances in the "non-von-Neumann" computer architecture.

Most knowledge based systems are designed to be of assistance to human endeavor; they are almost never designed to be autonomous agents. This makes a human-machine interaction subsystem a design necessity. The imperative in the Fifth Generation plan is to make this interaction as natural as possible for users, in language and mode of interaction. To the Japanese, this means language understanding, the ability to speak to the machine, and the ability to communicate information to the machine by showing it a picture.

To realize these goals across the spectrum of human knowledge and images is one of the most difficult of the long-term goals of artificial intelligence research. But if the subject of the

Research Projects Agency (DARPA) will almost certainly exceed the $850 million amount over a decade, even with no special response to the Japanese challenge. And the ANNUAL R&D budget of IBM for 1982 was about one and a half billion dollars! On the other hand, a myriad of smaller highly innovative companies, whose R&D budgets are relatively small and necessarily focussed on the near-in, would find these budget numbers impressively large and highly motivating. In truth, the large firms too have only small amounts of discretionary funds at the margin of current activities with which to innovate (existing projects tend to be large consumers of funds with big inertia). Seen from these perspectives, the Japanese Fifth Generation project budget is impressive.

Equally impressive is the strategy designed by MITI and Fuchi for managing the Fifth Generation project. ICOT, the "instant institute" drew together, within two weeks of the start of the project, forty researchers from the participating companies, often the best and brightest, handpicked by Fuchi and his managers. The managers were drawn from MITI's Electrotechnical Laboratory, a first-rate scientific lab that was the prenatal home of the project, and from the innovative NEC Research Laboratory. In parallel with the formation of ICOT, closely allied R&D groups in the company laboratories were targeted to track the scientific and technological progress at ICOT and absorb it for proprietary use. MITI funds for support of these company groups will begin to flow in 1983. These funds will flow through ICOT and be disbursed by contract for work performed. This contract mechanism, so familiar to American industry dealing with government funds, is apparently unique among MITI-funded national projects, for which the usual mechanism is a flow directly from MITI to the companies. The contract mechanism will not be applied in a heavy-handed way. Each company has asserted one or more key foci of interest, and ICOT will respect these and work within that framework. The structure seems to have been developed to implement a goal of major importance to MITI: to apply pressure upon Japanese industrial computer scientists to innovate, not merely to evolve existing Western technologies. ICOT, with its farreaching goals and its intellectually aggressive collection of researchers, will nurture young shoots of innovative work and transplant them to the industrial labs. The point of the contract mechanism is to insure that these young shoots receive the necessary and appropriate watering so that they will grow into healthy commercially viable plants.

The transplantation of ideas occurs routinely and frequently. Company researchers assigned to ICOT make visits to their labs of origin to report progress and educate their colleagues when necessary. But the scientific discussions held within the walls of ICOT are not burdened by proprietary boundaries! Such concerns are seen as harmful to the nation at this stage, when a great mixing-up and shaking-out of ideas is most needed. Such cooperative endeavor might agitate a Washington anti-trust regulator, were it to occur in the USA. But ICOT's mission is to foster such cooperation and educate industrial scientists actively by joint project work. Visiting ICOT gives one much the same feeling of openness as one gets on a visit to the major AI laboratories of American universities.

The researchers of ICOT are committed by their companies for a limited period. The early commitment is to the first three year period. Thereafter, most of the first group will transition back to their companies, carrying with them the collective wisdom and the body of techniques collectively brought to fruition; and others will be assigned by the companies to go to ICOT to carry the work forward.

What are the company foci that energize their participation in the Fifth Generation project? NEC has a long-standing interest in the hardware and problem solving software associated with PROLOG machines. In contrast, NTT's Musashino Labs , the "Bell Labs of Japan", has interest in hardware associated with the LISP programming language for symbolic processing, and may build a very high speed LISP machine. Industrial applications of Expert Systems over a wide front motivate Hitachi people at their System Development Laboratory and Energy Laboratory. And Fujitsu's Central Laboratory is motivated in all areas, from hardware (they are building a LISP machine to attach to existing Fujitsu machines), through software and Expert System applications. The other companies will relatively quickly define their foci of interest, with help from ICOT.

interaction is suitably constrained, then the technical problems become tractable, though hard. The Fifth Generation plan sets various constraints of vocabulary and subject matter on the interactions.

Effective processing of the electrical signals that represent speech and pictures requires specialized hardware to determine the most basic features of the words and images. That's only the beginning. Software capable of inducing an understanding of the language spoken or typed or the image shown does most of the work, with the assistance of knowledge in the knowledge base (it is much easier to understand what is being said or seen if you know something about the subject matter!).

That's the essense of the plan: hardware and software for each of three subsystems--Knowledge Base, Problem Solving and Inference, and Human-Machine Interaction. We haven't described here the nature of the ancillary software that is planned, or the intricacies of the technical tactics that constitute the experimentation to be done to approach the technical goals. That would constitute a text for computer scientists.

The realization of visionary engineering goals usually requires much time and much money. The Japanese are accustomed to investing both in their major technology projects. The Fifth Generation project is structured over a ten year period. The first three year period is targeted for climbing the well-known "learning curve"; for building the research teams and laboratories; learning the state of the art; forming the concepts that will be needed in the later work; and in building hardware and software tools for the later stages. The ~~one~~ single-user sequential PROLOG workstation is one of those tools. This workstation itself ~~is~~ will be a prototype of later machines, and so will be its problem solving software. Early Expert system prototype applications will be ~~[illegible]~~ written. Three will be selected from a variety of areas, such as (over)

The second

period, four years, is one of engineering experimentation, prototyping, ~~and the initial~~ continuing experiments at significant applications, and initial experiments at system integration. The first thrust at the major problem of parallel processing will be done in this period. The final period of three years is devoted to advanced engineering, building of major engineering ~~final~~ prototyp~~ing~~es, and system integration (making the subsystems work together smoothly). The earlier work on CAD for VLSI will be used at this stage to assist the design of hardware. Experiments with some difficult applications will be attempted in this period; good engineering requires that one smash the system upon the rocks of hard reality to learn how to fix it so that it is robust and reliable. Finally, in the third period, the results of the R+D will be distilled into a set of specifications for the Fifth Generation commercial products to be produced by the companies, ~~as a result of the project~~.

The budget for the Fifth Generation project is substantial, though not a huge sum by American standards for major projects of this scope. MITI has announced a commitment of $450 Million of the ten year period, starting relatively slowly for the first three year period ($45 million), and accelerating later when the expensive developmental engineering is to be done. The first stage will be fully funded by MITI. In the second and third stages, it is planned that government funds will be matched equally by funds from the participating companies. The total Fifth Generation project budget, therefore, works out to be approximately $850 million. Other MITI-initiated national projects have seen higher ratios of industry to government spending, sometimes two or three to one. It is possible that if the project is going well, and the Japanese economy is strong, the total budget could escalate to well over one billion dollars. Whether such a budget seems large or not depends upon what you scale it against. Budgeted amounts for information processing research and advanced development at the US Defense Advanced

WHAT'S RIGHT?

Often in science and technology, the most important part of the creative act is the asking of the right question, or the placing of the right long-term bet. This act, which may consume only a small fraction of project time and money, is crucial to the ultimate success or failure of the work. The rest is the necessary perspiration that brings the inspiration to fruition.

Creating the Knowledge Industry, with hardware, software, and knowledge system applications, is a great bet! Indeed, it is one of the few great bets sitting out there now in the information processing industry, ready for a major push toward exploitation. Of course the traditional modes of numerical calculation and data processing will continue to develop and prosper. But these will see steady incremental growth, not explosive growth. The exponential growth will be seen in symbolic computation and knowledge-based reasoning by computer.

The key economic insight is correct. For an island trading nation like Japan, wealth is created by the margin of exports over imports. In the Knowledge Industry, the value of the exports is enhanced by indigenous resources--the intelligence, education, and skill of people--and the value of imports is minimal (computers are not material intensive). Further, the KIPS will significantly enahnce the productivity of many other industries, thereby indirectly contributing to the value added.

The time is right for a major initiative in the industrialization of artificial intelligence and the Japanese are siezing the initiative to move out briskly ahead of the pack.

The move had been preceded by a thorough planning effort. The Fifth Generation plan of October 1981 is a strategy and not a set of tactics. Appropriately it sets forth goals that stretch out over a long period of time. It is not, and could not be, a how-to-do-it manual. Its achievement is that it focussed attention on the right set of issues, properly structured. That's important in a complex and difficult project, for it is all too easy to squander resources and time milling about smartly without plan.

The organizational approach that called for the creation of ICOT, the pooling of talent in a cooperative endeavor, the wellcoordinated transfer of technology between ICOT and the parallel labs of the companies, seems inspired.

MITI's concern for nurturing the innovative talents of Japanese computer scientists appears to be well placed. In the first of the "tactical" addenda to the Fifth Generation plan, dated May 1982, ICOT (undoubtedly speaking for MITI) expresses the following worry about the future: In Japan, research and development has hitherto been aimed at catching up with the technologies of the U.S. and the advanced European nations. With Japanese technical achievements, however, the U.S. and the advanced European nations have become wary of providing leading technologies, and we fear that the old style of catching-up research and development will become more and more difficult." This is undoubtedly a correct perception. Trade wars are under way, and blockades are inevitable.

The ten year planning horizon is excellent. Ten years is a long time in the information processing industry, almost inconceivably distant from us. Most of the people in the industry were not even in it ten years ago. Only two prototype Expert Systems had been built. The timesharing of expensive mainframes was a dominant force and maturing technique (what? No personal computers?). Even pocket calculators were relatively expensive. And the Japanese had yet to produce their first commercially viable microelectronic chip. As we live through it, we tend to underestimate the rapidity of technological change.

The Fifth Generation plan is difficult and will require much innovation; but of what sort? In truth, it is more Engineering than Science. Though solutions to the technological problems posed by the plan may be hard to achieve, paths to possible solutions abound. The Japanese are rich with excellent engineering talent, and adequately supplied with cutting-edge computer scientists This mix of talents enables (but of course does not guarantee) a good chance of success.

WHAT'S REAL?

The Fifth Generation project is hard--challenging in every dimension of the information processing science and technology. But as we have said, TEN YEARS IS A LONG TIME! In the magic world of computing, the world of "Always More for Less", where More for Less," where the More is double and the Less is half every two or three years, ten years really is a long time.

The Japanese will have a partial success. Managers of the Fifth Generation project have said that it would not disturb them if only ten percent of the project goals were achieved; others have remarked that the ten year planning horizon should not be taken too seriously; that the project goals were so important that a stretchout over another half or full decade would not be unreasonable.

Partially realized concepts that are superbly engineered can have great utility and be of great economic benefit. At the very least, a partial success can preempt the area, and make it not worthwhile for others to enter playing catch-up. It may be the case that the first twenty percent of the technical achievement will skim off eighty percent of the economic gain that there is to be realized. If this were to be true, then firms in the American industry might never find it in their economic interest to enter the arena. Being late might put them out of the contest. Consider this: though videotaping was invented in the USA, the lengthy and expensive research and development process for the consumer-oriented video cassette recorder led to an all-or-none market share result, with the American industry getting the "none". Even the VCRs with American household words like RCA and Sears attached to their cases are made in Japan.

No matter how partial its success may be, the Fifth Generation project will provide a decade-long learning experience for a new generation of Japanese computer scientists. The will be called upon to confront and perhaps solve the hardest and most challenging problems, looking forward, rather than reengineering traditional systems, looking backward. Since the fundamental ideas, as we have said, are software concepts, they will be living with and learning advanced software concepts in a way that has never been done before in Japan, and has never been widely done in the USA or Europe.

For the Fifth Generation project, technology transfer mechanisms are in place, and the Japanese industry traditionally moves with speed and diligence to bring developments to market. The USA has a substantial lead over the Japanese in virtually every area of Fifth Generation work. But FORTUNE's article on the Fifth Generation project concludes with the observation:

"Even if the U.S. retains its lead in AI research, there is no guarantee that the laboratory work will end up in products. Computer research tends to seep into the American marketplace slowly except when companies perceive a competitive threat. Assuming that ICOT can do even a fraction of what it intends, the results will show up quickly in Japanese computer products. So the U.S. computer industry could be outmaneuvered unless it takes the Fifth Generation seriously."

Let's return to the question of lead. One thing that is real is that the USA, and to a more limited extent the UK, has a big lead at present in this area of information processing technology. If the Japanese did not have a well-planned, well-organized and well-funded effort, that lead would perhaps be ten years. Because the Japanese are moving now, the real lead is perhaps three years. That is still a large lead by the normal standards of Silicon Valley and Route 128, where six month leads confer competitive advantages and twelve months leads are cherished. But our business-as-usual attitude, our near-term R&D planning horizons, our fratricidal competitive zeal and paranoia over proprietary rights, and our planning vacuum at the national level is causing us to spend our valuable lead at the rate of one day per day! That's real to an economic planner and to an executive in the information processing industry. The Japanese have a three year COMMITMENT lead. We must get our act together before we squander the balance of our technical lead.