THE IMPACT OF FIFTH-GENERATION COMPUTERS ON LIBRARY AND INFORMATION SERVICES

MICHAEL A. SHEPHERD
SCHOOL OF COMPUTER SCIENCE
TECHNICAL UNIVERSITY OF NOVA SCOTIA
HALIFAX, NOVA SCOTIA, CANADA B3J 2X4

ABSTRACT

Over the past forty years, many predictions have been made of computer systems which would be "intelligent"; able to organize data, infer solutions, and make decisions without human intervention. It appears that the fifth-generation of computers may finally fulfill some of those predictions. In anticipation of such intelligent systems, librarians and information specialists should not ask, "What will fifth-generation systems do for library and information services?", rather, "What do we want library and information services to be by the time fifth-generation systems are readily available?" This paper gives an overview of fifth-generation systems and their potential impact on library and information services.

L'impact des ordinateurs de cinquième génération sur les bibliothèques et les centres de documentation.

Depuis quarante ans, on a fait de nombreuses prédictions au sujet des systèmes d'ordinateur qui seraient «intelligents», capables d'organiser les données, de proposer des solutions et de prendre des décisions sans aucune intervention humaine.

Dans l'attente de ces systèmes intelligents, les responsables des bibliothèques et centres de documentation ne devraient pas se demander: «Qu'est-ce que les ordinateurs de cinquième génération apporteront aux bibliothèques et aux centres de documentation?», mais plutôt: «Que désirons-nous que soient devenus ces bibliothèques et centres de documentation lorsque les systèmes de cinquième génération seront disponibles?»

Ce document donne un aperçu des systèmes de cinquième génération et de leur impact possible sur les bibliothèques et centres de documentation.

THE IMPACT OF FIFTH-GENERATION COMPUTERS ON LIBRARY AND INFORMATION SERVICES

MICHAEL A. SHEPHERD
SCHOOL OF COMPUTER SCIENCE
TECHNICAL UNIVERSITY OF WESTERN ONTARIO
HALIFAX, NOVA SCOTIA, CANADA B3J 2X4

INTRODUCTION

Almost forty years ago, Vannevar Bush recognized that the basic processes of information retrieval could be reduced to machine processes and the effect that this might have on the storage and retrieval of information (Bush, 1945). Twenty years later, Licklider (1965) characterized the computer-based library of the future as being able to:

- handle both documents and facts;
- represent a body of knowledge and improve on that representation through use;
- use heuristics to solve problems;
- automatically improve its performance through learning techniques such as feedback;
- network users for both knowledge input and retrieval;
- negotiate the user query.

Although current systems exhibit some of the above characteristics, there are no systems which have all of the above capabilities. Fifth-generation computers, however, will be "intelligent" and should be able to provide the above capabilities and possibly more.

These intelligent systems are no longer merely the wishful thinking of a few farsighted individuals. The time frame for delivery of fifth-generation systems is the mid 1990's although commercial products incorporating various features of these sytems will be in the market place by 1985 (Schefter, 1983). As the potential impact of such systems on library and information services is profound, planning for their use should begin now, not after the systems are already available.

CHARACTERISTICS OF FIFTH-GENERATION SYSTEMS

The characteristics these machines will exhibit are best exemplified by the three broad targets of the Japanese Fifth-Generation Computer Systems project: the improvement of man/machine communication, the development of systems able to infer solutions to problems, and the storage, organization, and retrieval of nonnumerical information (McCorduck, 1983; Richmond, 1982).

It is expected that man/machine communication in fifth-generation systems will be through continuous speech. These machines should be able to understand continuous human speech with a vocabulary of 50,000 words from a few hundred speakers with an accuracy of 95 per cent.

Typical of the fifth-generation inference systems will be expert systems which should not only be able to translate fluently between languages, but should be able to automatically organize the contents of documents into knowledge bases. Expert systems will make use of these knowledge bases to provide consultative services to both subject specialists and to the general public.

The third broad target of the Japanese project is a system capable of the organization and retrieval of nonnumerical information as required for picture and image processing. Picture and image processing systems require the capability to deal effectively with the characteristics of three dimensional spaces, satellite images, and medical images.

The goal for the fifth-generation computer projects is to produce an inference computer or computer network capable of 1 billion logical inferences per second with a knowledge base of a billion inference rules and a database of hundreds of million of objects. In order to achieve this goal it will require a blending of three areas of science and technology; VLSI, computer architecture, and artificial intelligence.

FIFTH-GENERATION COMPUTERS

FIFTH-GENERATION COMPUTER MEMORY

Much as the personal computer has been given to us by large-scale integration of computer circuits and the supercomputers by very large-scale integration, ultra large-scale integration will play an important part in the delivery of the fifth-generation. These silicon chip-based components provide the internal memory and are the realization of the computer architectures that run the software.

The memory capacity of a silicon chip can be measured in bytes or characters. In order to supply the fifth-generation with the projected memory capacity, a million chips will have to be linked together, each chip with a capacity of 128 kilobytes, for a total of 128 billion bytes of internal memory.

The memory capacity of the silicon chip has doubled every eighteen months since 1974, with the 32 kilobyte chip being developed in 1983 (Pease, 1983). In keeping with this timetable, the 64 kilobyte chip should be developed in mid-1985 and the 128 kilobyte chip the end of 1986. However, in October 1983, IBM announced the production of an experimental 64 kilobyte chip, somewhat accelerating the timetable (IBM Corp., 1983).

In order to put a prototype 128 billion byte computer into perspective, it should be noted that the Encyclopedia Britannica has approximately 200 million characters.

FIFTH-GENERATION COMPUTER ARCHITECTURE

The architecture of current and past generations of computers may be described as the von Neumann architecture, after the mathematician who conceived it. The von Neumann architecture has a global addressable memory to hold program and data objects, with program instructions frequently updating the contents of the memory during execution. It also has an instruction counter which holds the address of the next instruction to be executed. This counter is either implicitly or explicitly updated to provide the machine with a sequence of instructions to execute.

The sequential nature of this architecture has become a limiting factor in the development of faster machines even though the physical components of the machines are faster than ever before. Each instruction to be executed must be fetched from the memory by the CPU along with the operands specified by the instruction. All results from the execution of the instruction must be stored in memory before the next sequential instruction, specified by the instruction counter, is fetched for execution.

In order to break this sequential bottleneck, a variety of parallel architectures have been proposed. In a parallel architecture, instead of one instruction being executed at a time, a number of instructions may be executed at the same time with the results then being used for other parallel operations. Machines based on parallel architectures promise to be orders of magnitude faster than von Neumann architecture machines.

ARTIFICIAL INTELLIGENCE

One of the major areas of research in artificial intelligence (AI) has been the development of "intelligent" machines. An intelligent machine is a computer program that can perform cognitive tasks at high levels of performance, not necessarily mimicing the internal structure of human behaviour (Feigenbaum, 1983).

The field of artificial intelligence will provide the intelligence aspects of fifth-generation computers in the form of expert systems. An expert system is a computer program whose performance is intended to rival that of human experts (Duda and Shortliffe, 1983). Such systems are knowledge-based in that they depend more on the presence of a large body of knowledge than on ingenious computational processes.

There are a number of such expert systems in the market place running on the present generation of computers. These systems are limited by lack of internal memory, lack of processing speed, and lack of a universal knowledge base. The lack of memory and processing speed should be rectified by the fifth-generation computers, after which it will be feasible to build larger, more complex knowledge bases.

Expert systems usually consist of two major components, the knowledge base with its underlying inference structure, and the database. By separating the two it is possible to enlarge the database or change attributes of the database without having to change the knowledge base.

The knowledge and inference structure may be represented in the form of rules, normally of the IF...THEN nature. An example of a rule from the MYCIN system, used in the selection of antibiotics for patients with severe infections, is given in figure 1. Assuming that there is suggestive evidence of pseudomanas aeruginosa as in the example, this knowledge would then be used as the predicate for another IF...THEN rule which may suggest a treatment.

```
IF the infection is meningitis
   AND organisms were not seen
       in the stain of the culture
   AND the type of infection
       may be bacterial
   AND the patient has been seriously burned
THEN there is suggestive evidence that
     pseudomonas aeruginosa is one of the
     organisms that might be causing
     the infection.
```

<u>FIGURE 1: EXAMPLE OF MYCIN RULE</u>

The knowledge in the MYCIN system is matched against the database, in this instance the data on a particular patient. Patient data are stored as <attribute-object-value> triples, as shown in figure 2.

The stain (ATTRIBUTE) of a particular organism (OBJECT) is gram-negative (VALUE).

FIGURE 2: EXAMPLE OF MYCIN DATA

CURRENT LIBRARY AND INFORMATION SYSTEMS

Most current library and information systems can be depicted as having a centralized database accessed from two separate sides; the administrative side and the user side. This is particularly true of library systems which, for the most part, have been designed by librarians for librarians. The user side of such systems have far less scope than the administrative.

On the administrative side, the term, "library automation", has a close analogy in, "factory automation". One of the major functions of a library automation system is inventory control with its cataloguing, circulation, and ordering subsystems.

The user side of such a system allows the user to query the database and retrieve citations. The user cannot update the inventory directly, order new materials, place reserves on items, or view his own circulation history.

FUTURE SYSTEMS

Future systems, based on fifth-generation computers, should allow a merging of the administrative and the user systems. The administrator may access the system for different purposes than the user, but it will be the same system.

The database itself will be stored on computer, not just the references to the database. This database may be distributed across the country but will be accessible anywhere there is an interface and the fact that it is distributed will be transparent to the user.

The interface, whether for the administrator or the user, will be a natural language interface. The dialog may be carried out verbally or through a terminal. The system will remember past sessions and be able use this knowledge to provide a user-friendly interface and refined search strategies. The system will operate at a level indicated by past sessions and by the ongoing query negotiation. In this way, a naive first-time user will have a system dialog tailored to his level.

The knowledge in the database and the underlying inference structure itself will be represented. The user will be able to retrieve documents, facts, and concepts. The system will also be able to analyze, store, and retrieve nonnumerical data such as graphics, still pictures, and moving pictures. The user will be able to ask the system to infer solutions to problems whose answers are not already in the database.

The database will be tied to electronic publishing and updated automatically from the contents of the input data.

The knowledge organization itself will change automatically depending on changes in the database and in patterns of use.

KNOWLEDGE ENGINEERING

If the fifth-generation computer projects deliver systems with capabilities fulfilling the stated goals, there will be a profound change in the manner in which we access and use data.

The very fact that these systems are "intelligent" will require them to be well managed. Procedures regarding database maintenance and access will have to be established. Access procedures determine who has access to what knowledge base as well as what system has access to what data and to what knowledge base. Maintenance procedures determine the vetting of information before it is added to the database and the development and vetting of rules incorporating the knowledge prior to adding the rules to the knowledge base.

Smith (1980) has referenced Tou (1979) who defines knowledge engineering as that discipline that deals with planning, design, construction, and management of knowledge-based systems for transfer, utilization, and extension of knowledge. The goal of a knowledge engineer is to make knowledge more generally accessible; something which librarians have been doing for some time. The main difference is that a new and very powerful tool will soon be at our disposal.

CONCLUSIONS

Fifth-generation systems will provide at least the characteristics of Licklider's library of the future. These systems will be able to:

- handle documents, facts, concepts, and pictures;
- represent the body of knowledge and the underlying inference structures and automatically update both the representation and the data;
- use expert systems to solve problems;
- automatically improve its performance through the use of long-term memory (past sessions) and short-term memory (feedback);
- network users for both knowledge input and retrieval;
- negotiate the user query in natural language, either through voice communication or terminals.

With such a powerful tool soon to be at our disposal, the question is not, "What will fifth-generation systems do for library and information services?", but rather, "What do we want library and information services to be by the time fifth-generation systems are readily available?"

REFERENCES

BUSH, V. "As We May Think." Atlantic Monthly. 176(1). (July 1945): 101-108.

DUDA, Richard O., and SHORTLIFFE, Edward H. "Expert Systems Research." Science 220 (April 1983): 261-268.

FEIGENBAUM, Edward A. "Artificial Intelligence." IEEE Spectrum 20 (November 1983): 77-78.

IBM Corp. "Experimental 512k Chip." Data Product News 5 (October/November 1983): 11.

LICKLIDER, J.C.R. Libraries of the Future. Cambridge, MA: MIT Press; 1965.

McCORDUCK, Pamela. "Introduction to the Fifth Generation." Communications of the ACM 26 (September 1983): 629-630.

PEASE, R.F.W., "Fabrication Issues for Next-Generation Circuits." IEEE Spectrum 20 (November 1983): 102-105.

RICHMOND, Louis S. "Here Comes Computer Inc." Fortune 106 (October 1982): 82-84.

SCHEFTER, Jim. "Fifth-Generation Computers." Popular Science 222 (April 1983): 79-81.

SMITH, Linda C. "Artificial Intelligence Applications in Information Systems". Annual Review of Information Science and Technology. 15 (1980): 67-105.

TOU, Julius T. "Knowledge Engineering". Unpublished Paper available from the author, Center for Information Research, University of Florida, Gainsville, FL 32611. (1979).