NEW GENERATIONS OF HUMAN-COMPUTER INTERACTION

Brian R Gaines
Knowledge Science Institute
University of Calgary
Calgary
Alberta, Canada T2N 1N4

Abstract. The problems of developing adequate human-computer interaction in fifth and sixth generation computing systems are analyzed in terms of two overall frameworks. The first framework is a model of the tiered structure of learning curves that make up the infrastructure of information technology, its development and applications. This shows human-computer interaction technology to be in a phase of theory and moving into one of automation. The second framework gives foundations for theory and automation through the layered protocol underlying the operation of integrated embedded information systems, including their human components. It is based on knowledge structures with systemic foundations that clarify the similarities and differences between the human and technological components of the system.

Keywords. Artificial intelligence, computer interfaces, hierarchical systems, human factors, system theory

INTRODUCTION

The human-computer interface still presents major problems in the specification, development and evaluation of computer-based systems. There are four major problems:

- The rapid pace of change in the technology, forms of interaction and system designs.
- Inadequate psychological models of the human operator. There is still a major gap between psychological and cognitive science theories of people, and their behavior in practice.
- Inadequate systemic models of computer-based systems. The flexibility of computing technology makes it very difficult to develop general theoretical foundations that have sufficient content to capture critical issues in practical situations.
- Inadequate interaction models of the human operator and computer-based system coupled together. Even when psychological and systemic models are restricted to the domain of human-computer interaction, the residual flexibility and complexity makes it difficult to generate adequate models. Empirical studies encompass too narrow a range of systems. Guidelines from experience are superficial and may fail to generalize to new situations. Foundations are currently too weak to be of practical significance.

The opportunities offered by technological developments and applications are counter-balanced by major problems stemming from the pace of change and the lack of foundations:

- The obvious changes are in the *computing technology* where the availability of low-cost graphic displays, graphic input devices, and windowing software has enabled the keyboard and textual interfaces of the 1970s to be replaced, or supplemented, by the window and mouse interfaces of the 1980s. Developments in speech, gesture, and direct sensory interfaces will continue the pace of change in the 1990s.
- Other obvious changes are in the *user psychology* where the spectrum of users has extended to include those with no computer experience at one extreme, to those with wide-ranging experience of a variety of past systems at the other.
- There are also major classes of users who are *anomalous* in the experience dimension—the widespread availability of computers in schools has generated a class of users who have experienced many forms of computer interaction but have little or no experience of real-world applications.

There are also major, but less obvious changes occurring in the *types of interaction* occurring between computers and people:

- The step from batch-processing to *interactive use* was a widely recognized transition generating a need for in-depth studies of human-computer interaction and usability.
- The step from episodic, occasional use of computers to *continuous use* as an integral part of the job environment was equally significant, but the change in work patterns and job concepts involved has been less widely studied.
- The step from highly specific computer applications to *integrated systems* having a role in every aspect of one's working life is also leading to major changes in the forms of human-computer interaction and the conceptual structures underlying individual and social activities.
- The increasing complexity of *embedded* interfaces combining special-purpose and general-purpose interactive devices is generating a need to integrate classical ergonomics and human-computer interaction approaches to the interface.

As the technologies, user psychologies and modes of interaction change, so do the theories underlying human-computer interaction and many of the guidelines for effective system design. This is an intrinsic feature of the rapidly changing technology and patterns of use, which is exacerbated by the extreme flexibility of both computers and people. Computer technology has only a limited grounding in the physical world. Its applications are diverse and may have little inter-relation. Its forms are as much determined by our creative imaginations as they are by its applications. Its applications in themselves may create new infrastructures for social and economic interaction that did not previously exist.

Despite these problems, the development and application of interactive computer-based systems is proceeding and highly significant both commercially and socially. Product innovation continues, and products evolve through a commercial selection process in which those with poor human-computer interfaces either improve them or fail in the market place. The cost is high, however, and the proportion of product development time and resources devoted to the interface has steadily risen so that it now exceeds fifty per cent and may reach ninety in major product lines.

What can be done to improve the interface, reduce, or at least contain, development costs, and improve understanding of human-computer interaction? There is no simplistic answer—we are dealing with situations that are already complex and are growing in complexity. This paper describes two frameworks which are helpful in encapsulating the complexity and making its sources and its control more comprehensible:

- Modeling the role of human-computer interaction technologies in the development of information technology.
- Developing a fundamental model of embedded information systems in terms of an extension of layered protocols.

FIFTH AND SIXTH GENERATION OBJECTIVES

The concept of a "generation" of computers was mainly regarded as one of marketing convenience until the the Japanese initiative in 1981 of scheduling a development program for a "fifth generation" of computers (Moto-oka 1982, Gaines 1984) led to a reappraisal of the state of information technology and the structures underlying its progress. According to the Japanese proposals, fifth generation computing systems would integrate advances in *very large scale integration, database systems, artificial intelligence,* and *human-computer interaction* into a new range of computers that were closer to people in their communication and knowledge processing capabilities.

Human factors considerations were originally stated to be fundamental to fifth generation objectives. Mota-oka's "conceptual diagram" (Figure 1) shows the user interface to be through speech, natural language, pictures and images to a knowledge-based system. He notes:

"In these systems, intelligence will be greatly improved to match that of a human being, and, when compared with conventional systems, man-machine interface will become closer to the human system." (Moto-oka 1982)

and Karatsu reiterates the human-computer interaction theme:

"Until today, man had to walk toward the machine with effort to fit precisely. On the contrary, tomorrow, the machine itself will come to follow the order issued by man." (Karatsu 1982)

However, these human-computer interaction objectives have not been reflected in the research program at ICOT, and their neglect is apparent in the conceptual diagram of Figure 1. The hardware and software rings are full of structure which is expanded in detail in the research program, but the human ring is empty. Fuchi, the Director of ICOT, recognizes this problem in his reply to the interview question:

"Are you saying that the design of the fifth-generation may be modeled by learning more about the human thinking process?"

answering:

"Yes, we should have more research on human thinking processes, but we already have some basic structures. For more than 2000 years man has tried to find the basic operation of thinking and has established logic. The result is not necessarily sufficient; it's just the one that mankind found. At present we have only one solution—a system like predicate calculus. It is rather similar to the way man thinks. But we need more research. What, really, is going on in our brain? It's a rather difficult problem." (Fuchi, Sato & Miller 1984)

Fuchi's concerns are echoed in the recent Japanese 'sixth-generation computing proposal', *Promotion of R&D on Electronics and Information Systems That May Complement or Substitute for Human Intelligence.*(STA 1985, Gaines 1986) proposing a research program focusing on the development of *knowledge science.*

The report is divided into three major sections: *basic considerations in the promotion of knowledge science; present state and future development of science and technology related to computer science;* and *methodology for promoting knowledge science.* The first section argues that computers were invented as devices to enhance human intelligence, but conventional computers may be inadequate as social and economic structures change, and as non-specialists become extensive users. For an information society computers must be developed with intelligent abilities: intuitive recognition of situations; inductive inference; accumulation of knowledge; and learning. The man-machine interface must also be improved. To develop integrated artificial intelligence systems with human capabilities it is necessary to: clarify the brain's thought processes; construct a basic model for the technical development of intelligent functions; and establish theories of the model.

Knowledge Science is defined as *science and technology for clarifying advanced intelligent functions of human beings, i.e. recognition, learning, inference, etc., while developing these functions technically for the construction of various systems.* Knowledge science will revolutionize many science and technology fields and contribute to progressive global change. Four objectives are given for promoting knowledge science: *innovations in frontier high technologies; contributions to societal, economic and cultural advancements; contributions to the expansion of human potential;* and *establishing a foundation for creative sciences.* The need for interdisciplinary research is emphasized and *physiology, psychology, linguistics* and *logic* are identified as relevant to the objectives of the program.

The second section analyses the state of the art in: these four relevant sciences; eight technologies based on them; and four major application areas. Figure 2 summarizes the coverage of this section. It is particularly interesting to note the operationalism of the section on brain physiology which calls for *"a model of the information processing function of the brain...a reasonable facsimile of the human cognitive process...modified throughout the investigation by using computer simulation and a mathematical method."* The Japanese proposal requires mathematically well-founded structures as the basis for developments in artificial intelligence and human-computer interaction, a contrast to the heuristics common in the West.

The psychology of human-computer interaction is analyzed in the science part of the second section in these terms:

"The man-machine system has been considered thus far in terms of human engineering in which the configurational characteristics of man are discussed. However, it is also important to reconsider it in terms of cognitive science, in which knowledge gained by analyzing human cognitive processes is used for technical design, so as to develop a system for human beings not only physically but also cognitively."

The man-machine interface is also analyzed in the technology part in terms of the requirements of non-specialist users, the assignment of intelligent functions and the understanding of human factors.

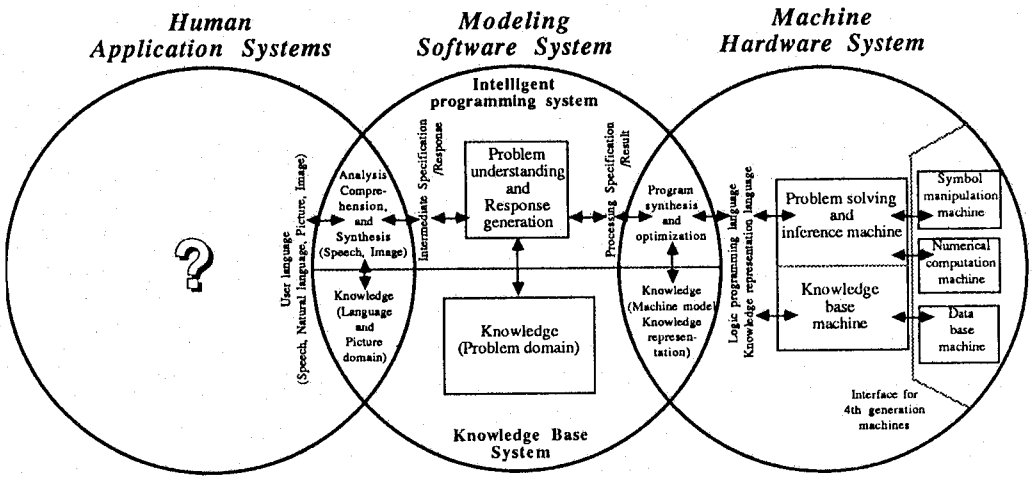


Fig.1 Conceptual diagram of fifth generation computing (Moto-oka 1982)

The third section outlines how the program might be based on the development of manpower, international cooperation, and a multi-disciplinary approach combining research on: the clarification of human intelligence; systems for the acquisition, utilization and control of knowledge ; and intelligent man-machine communication systems. It is noted that:

"the life cycle of R&D in knowledge science is much longer than that of other scientific topics...we must exert effort to train researchers, expand research institutes which have already been established, build more institutes where research can be integrated, gather researchers from related fields, and arrange R&D systems continuously from a long-term point of view with industry-university government cooperation."

This report was part of the preparation for the Human Frontier research proposal which is now under consideration by the "Group of 7" and involves international collaboration in research on problems of the human species (CHFP 1986). It is not clear how much of the work will be funded, but the proposal provides very clear objectives for sixth generation computing.

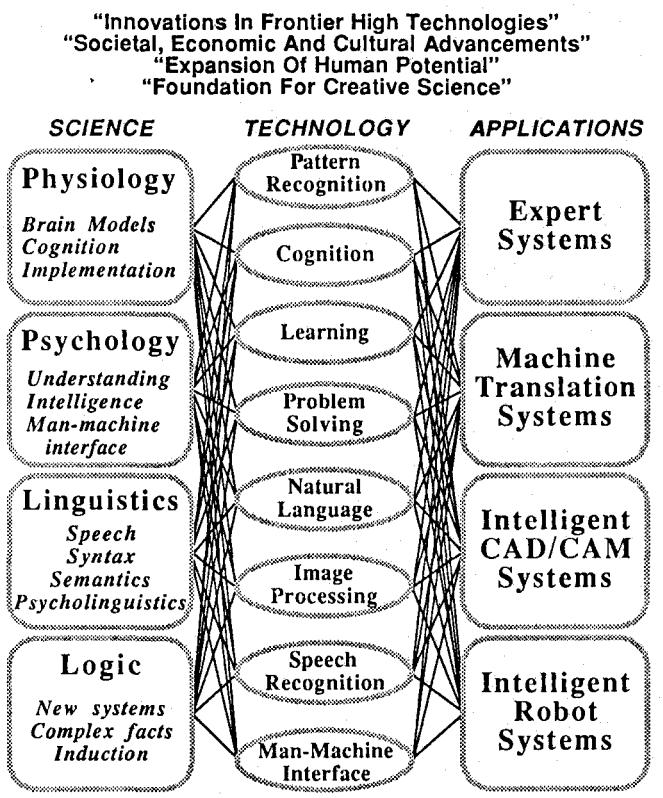
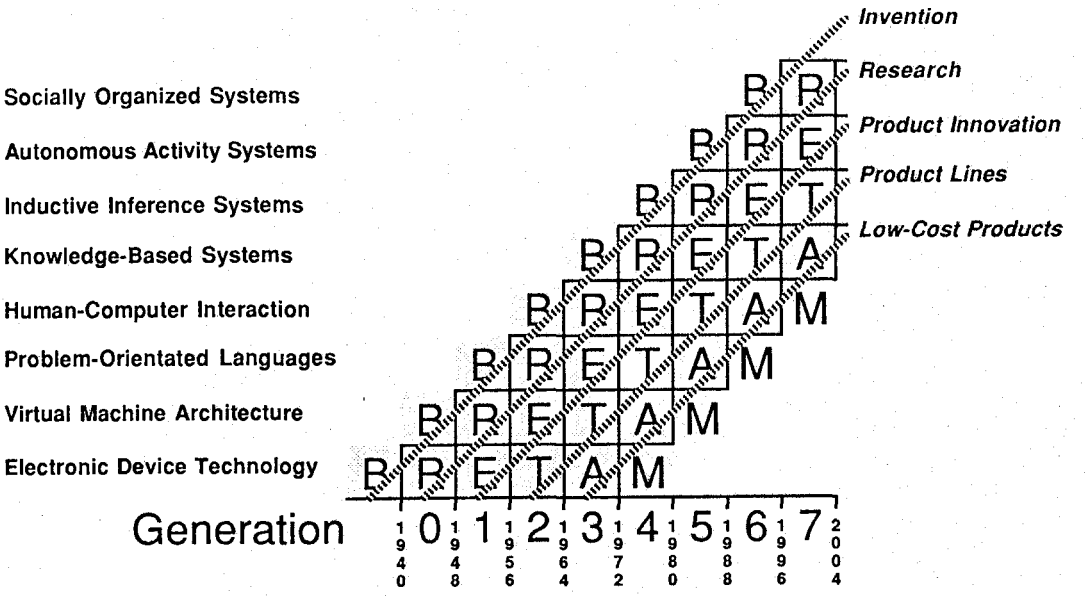


Fig.2 The Japanese Sixth Generation Proposal: Knowledge Science, Technology and Applications

GENERATIONS—THE INFRASTRUCTURE OF INFORMATION TECHNOLOGY

The sixth generation proposals emphasize the significance of multi-disciplinary research in physiology, psychology, linguistics and logic. Human-computer interaction research in has a history of multidisciplinary interaction with brain science, cognitive science, psycho-linguistics, philosophy, and systems theory. However, this diversity of relationships has made it difficult to comprehend the place of such research within the infrastructure of computing technology with its emphasis on hardware and software engineering.

The role of human-computer interaction in fifth and sixth generation computing can be analyzed through study of the infrastructure of computing. Like other basic technologies (Marchetti 1981), the growth of computing is comprised of the envelope of learning curves in successive underlying technologies (Ayres 1968). Figure 3 shows this structure through the zeroth through fifth generations projected to the sixth and seventh. Each technology has a learning curve in which a **breakthrough** leads to a phase of: **replication** in which the results are duplicated; **empiricism** in which rules for design are derived from experience; **theory** in which generative principles are derived; **automation** in which the theory is made operational; and **maturity** when the knowledge becomes proven, widely available and widely used. In the development of computing each phase corresponds to a generation of about eight years linked in timing to the medium term business cycle (Gaines & Shaw 1986).



- Breakthrough: creative advance made
- Replication period: experience gained by mimicing breakthrough
- Empirical period: design rules formulated from experience
- Theoretical period: underlying theories formulated and tested
- Automation period: theories predict experience & generate rules
- Maturity: theories become assimilated and used routinely

Fig.3 The infrastructure of information technology through eight generations

At the base is the learning curve for different electronic device technologies: 1940 zeroth generation using relays; 1948 first using vacuum tubes; 1956 second using transistors; 1964 third generation using using integrated circuits; 1972 fourth using large-scale integration; 1980 fifth using VLSI; 1988 sixth projected to use ultra large-scale integration with some 10 million transistors on a chip; 1996 seventh projected to use grand-scale integration with 1,000 million on a chip. The definition of generations in terms of electronic device technology captures some important aspects of computing such as cost decreases, size decreases, power increases, and so on. However, it fails to account for the qualitative changes that have given computing its distinct character in each generation. These appear through the tiered succession of learning curves of higher level technologies based on the lower level developments.

The first generation breakthrough was the introduction of stored program and subroutine concepts around 1947 which detached computing as a separate discipline from electronics by substituting software for hardware in a virtual machine architecture. The second generation breakthrough was to bridge the gap between machine and task through the development of problem-orientated languages such as FORTRAN in 1956. The third generation breakthrough was to bridge the gap between the computer and the human with the development of interactive time-shared computers in 1964 allowing close human-computer interaction. The fourth generation breakthrough was in the early 1970s with developments in expert systems based on knowledge-processing. The fifth generation breakthrough was in 1980 with developments in machine learning and inductive inference systems. One may speculate that the growth of robotics will provide the next breakthroughs in which goal-directed, mobile computational systems will act autonomously to achieve their objectives, and that interaction between these systems will become increasingly important in enabling them to cooperate to achieve goals.

This model of the development of computing is important in analyzing the role of human-computer interaction studies and their changing priorities. In particular the emphasis of workers in a particular activity changes as the learning curve progresses so that: **invention** is focused at the **BR** interface where new breakthrough attempts are being made based on experience with the replicated breakthroughs of the technology below; **research** is focused at the **RE** interface where new recognized breakthroughs are being investigated using the empirical design rules of the technology below; **product innovation** is focused at the **ET** interface where new products are being developed based on the empirical design rules of one technology and the theoretical foundations of the technology below; **product lines** are focused at the **TA** interface where product lines can rest on the solid theoretical foundations of one technology and the automation of the technology below; **low-cost product lines** are focused at the **AM** interface where automation has allowed cost-reduction in high-volume mature products.

For example, in the fourth generation (1972-79): **BR**: recognition of the knowledge acquisition possibilities of KBS led to the breakthrough to inductive-inference systems; **RE**: research focused on the natural representation of knowledge through the development of human-computer interaction, e.g. the Xerox Star direct manipulation of objects; **ET**: experience with the human-computer interaction using the POL BASIC led to the innovative product of the Apple II personal computer; **TA**: the simplicity of the POL RPG II led to the design of the IBM System/3 product line of small business computers; **AM**: the design of special-purpose chips allowed the mass-production of low-cost, high-quality calculators.

In the current fifth generation (1980-87): **BR**: recognition of the goal-seeking possibilities of IIS is leading to the breakthrough to autonomous-activity systems in robotics; **RE**: research is focused on learning in KBS; **ET**: the advantages of the non-procedural representation of knowledge for human-computer interaction led to the innovative designs of the Visicalc spread-sheet business product and the LISP machine scientific product; **TA**: the ease of human-computer interaction through a direct manipulation POL led to the Apple Macintosh product line of personal computers; **AM**: the design of highly-integrated language systems has allowed the mass-production of low-cost, high-quality software such as Turbo Pascal.

The horizontal **BRETAM** sequence for human-computer interaction has the **breakthrough** in 1963-64 with the development of systems such as MIT MAC (Fano 1965). In the **replication** period such systems came into widespread use well before the human factors principles underlying their design were understood. Hansen's tabulation in 1971 of some user engineering principles for the design of interactive systems marks the transition to the **empirical** period (Hansen 1971). The transition to **theory** at the beginning of the 1980s was marked by studies of human-computer interaction developing theoretical foundations based on cognitive science (Norman 1984). A reasonable expectation in the theory phase of fifth generation human-computer interaction is a set of principles that systematically generates rules for dialog engineering grounded in system theory, computer science and cognitive psychology. The principles should be applicable to a range of possible dialog styles and technologies,, and operational so that they can be embedded in standard dialog shells.

The vertical **BRETAM** sequence for the current fifth generation era shows the interplay between human-computer interaction and the other strata of computing. Fifth generation product lines are built upon the solid foundations of electronics, machine architecture and problem-orientated languages. The most recent developments to impact them are in human-computer interaction and these are the most critical to product differentiation. Fifth generation product innovation is based on the same technologies with the critical one being that of knowledge-based systems supporting intelligent user interfaces. Fifth generation research concentrates on the inductive aspects of knowledge acquisition, systems learning from people and from experience.

In 1988 we moved into the sixth generation and the final phase of the learning curve for human-computer interaction. For any discipline in this phase research becomes harder, standards of refereeing become harsh and many researchers drop out (Crane 1972). Obtaining results in the final 10% of the curve requires a thorough understanding of the known 90%. It requires careful experimental design, precise theoretical formulations, meticulous engineering practice in system development. In short, it is a phase of professionalism.

In each generation technologies interact to mutually support one another's development. Figure 4 shows these interactions with human-computer interaction technology in the fifth and sixth generations. The transition from theory to automation in interaction technology is supported from below by the maturity of technologies in hardware (large memories, fast processors), architectures (virtuality, graphic processing), and software (abstract data types, object-oriented languages). It is supported from above by evolving theory in knowledge systems (knowledge representation, expert systems), empirical developments in inductive inference (user modeling), and will feel the impact of breakthroughs in goal-directed and socially organized systems (planning, cooperation).

In the next section an analysis of human-computer interaction in integrated embedded information systems is developed in terms of the information and knowledge flows involved.

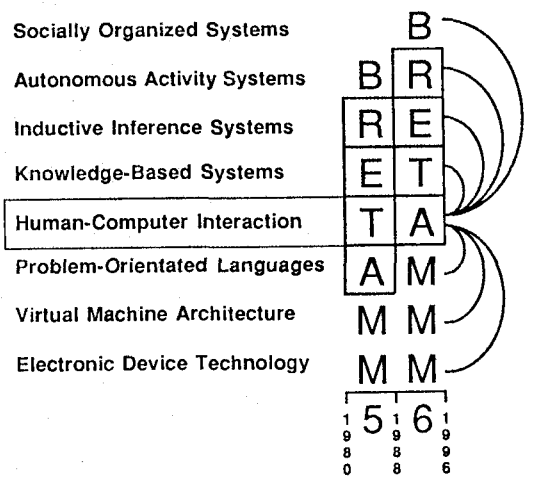


Fig.4 Interactions between human-computer interaction and other technologies in the fifth and sixth generations

EMBEDDED HUMAN-COMPUTER INTERACTION IN INTEGRATED SYSTEMS

Figure 5 shows the information systems in a typical organization using integrated embedded information systems. The outer ring comprises eight basic information technologies that together support the organization and its interface to the physical and organizational worlds. The center shows the organization itself with its corporate structure, and the rings surrounding it show the human interfaces and communication protocols necessary for the systems' management and integrated operation. The figure has been simplified to the minimum structure necessary to support an effective integrated system, and the complexity shows the need to consider the total context of the information technology. Such a system will not necessarily succeed if any one component, or even several, are advanced state-of-the-art technologies. However, it will probably fail if one component is inadequate or the total system is not well-integrated.

The Organization

The central core is the non-technology part of the system, the people that manage and operate the plant, their organization and value systems. The overt operation of the integrated system is part of the *technical culture* of the organization in Hall's (1959) terms. It can be communicated, documented and partially programmed. It can also be changed in response to new requirements and technologies. The management structure of the organization is part of its *formal culture*. Only part of it is overt and documented, and it is resistant to change. What has come to be called the "corporate culture" (Deal & Kennedy 1982) is in Hall's terms the *informal culture*. It pervades the organization, is communicated by example, rarely documented, and changes only slowly or under pressures of survival.

The managerial structures and social processes of the organization are not normally treated as part of its information technology. However, effective technical systems must be consistent with organizational structures and processes. This applies particularly to inter-personal communications where computer networks will not be used if they violate the norms of the organization, and to command and control paths which have to conform to existing authority and procedures that are not necessarily those of the organization chart.

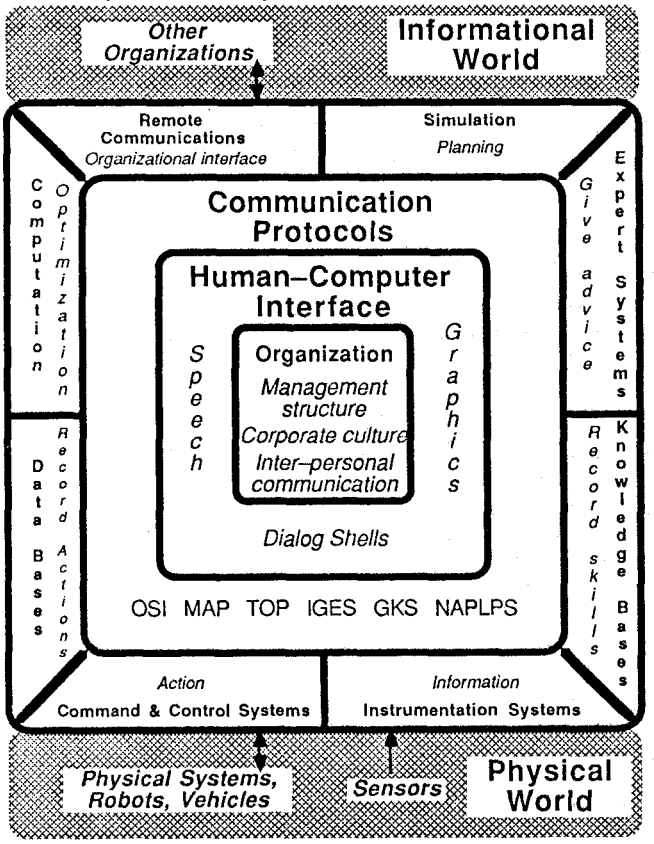


Fig.5 Integrated embedded information systems

Human-Computer Interface

Recognition that the managerial structure of the organization dominates its style of operation draws attention to human-computer interaction as a critical technology. A technically powerful system will be wasted unless it is comprehensible and accessible to its users, and gives them confidence in the basis of its operations.

Communication Protocols

The days of single-vendor, special-purpose systems designed against a static, well-defined specification are long gone. Equipment from many vendors, purchased at different times, for different purposes, by different parts of the organization, has to operate together. The technical problems are beginning to be resolved at the lower levels by the adoption of industry standards in communication protocols such as the open systems interconnection standard (Day & Zimmerman 1983).

The Applications Ring

The outer ring in Figure 5 shows the main information technologies in integrated systems and their typical applications. Moving counter-clockwise around the ring:

Remote Communications are necessary to interface the organization with the organizational world in which it operations A particular organization is part of an infrastructure each part of which is trying to optimize its own operations.

Computation Systems perform the classical number-crunching tasks under the control of algorithms that with modern software engineering have a high degree of accuracy, conformity to specification and reliability.

Database Systems have become the classic core of an organization's information system, recording all its activities in a co-ordinated fashion and making the information available on a controlled basis to those systems and people with a legitimate need to know.

Control Systems are now increasingly integrated with information systems so that decisions can be implemented directly in the physical world.

Instrumentation Systems are also increasingly integrated to provide timely information on the state of the physical world and the effectiveness of operations.

Knowledge Bases are distinguished from databases by containing inference rules rather than facts. If databases may be said to record an organization's activities then knowledge bases may be said to record its skills.

Expert Systems are the means to apply a knowledge base to the solution of a particular problem. They stand in relation to the knowledge base as does conventional computation to the data base.

Simulation has been separated from other forms of computation because it is computation based on a model of some other system and it is the relation of the simulated model results to that of the system modeled that is of paramount importance.

HIGHER LEVEL PROTOCOLS IN INTEGRATED SYSTEMS

The diagram of Figure 5 captures the key problem areas for integrated systems in a neat picture, but it gives no basis for treating them in an integrated way. How do the corporate culture, human-computer interface and network protocols relate, and how does this affect operations? The only strongly formalized and standardized structures within Figure 5 are the network protocols, and it is useful to examine these in more detail because they involve concepts that are useful in formalizing more of topics encompassed in Figure 5.

Figure 6 shows the Open Systems Interconnection Reference Model. It models communications as a protocol with seven layers each having a distinct function that can be defined fairly independently of the other layers. Within each layer the system can be viewed as providing services to the layer immediately above using the capabilities of the layer immediately below. The first level defines standards for the physical transfer of messages—electrical levels, timings, connectors, and so on.

Remote Communications

Organizational interface

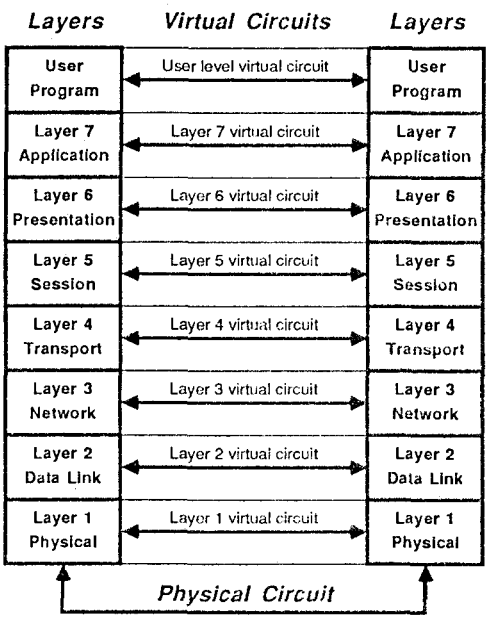


Fig.6 Open systems interconnection layered protocol

The seventh level defines standards for the content of messages in relation to applications programs. Intermediate levels are concerned with intra-message validity, network addresses, end-to-end message transfer checking, correct sequencing, and standard data formats for very general data structures such as graphics.

The most important concept underlying the OSI model is that of *virtual circuits*. Designers of a particular level are concerned primarily with its relation with the equivalent level in the system with which it is communicating. They need have little concern with the levels above or below it. Hence they think in terms of messages passing directly between equivalent levels, through a "virtual circuit," rather than through the multiple levels of the actual system.

The top level above the OSI defined layers in Figure 6 is the "user level virtual circuit," representing the interactions between user programs through the communications network, and hence the interactions between users themselves, and their activities, if the programs are interactive. This concept of virtual circuits between users, and user activities, is very powerful in enabling Figure 5 to be restructured to provide a more integrated framework for the other sections not directly included in the communications protocol. It can be redrawn in such a way that the relation of managerial and operations considerations to higher-level protocols in integrated communications networks is apparent.

Figure 7 shows the same conceptual blocks as those in Figure 5, but now in hierarchical form with the human-computer interface, information systems, organizational structure and external worlds shown as higher level "layers" above the network layer. In OSI terms these upper layers may be seen as subsuming and extending the presentation and application layers of the protocol, and give substance to the user level virtual circuit.

The virtual circuits between the human-computer interface layer and the information systems layer in Figure 7 are easy to understand. The user of a computer terminal does not think of himself as communicating through a network but rather as directly interacting with some data-processing sub-system. This layer may go beyond the OSI specification but it seems a natural extension of it to the technical aspects of the interaction between people and computing systems. However, the "virtual circuits" at the level above, between the organizational layer, and the external world layer, may seem rather less credible. It is a reasonable metaphor that the sales manager does not see himself as interacting with a complex administrative and productive system, but rather as communicating directly with a customer organization by supplying it with goods.

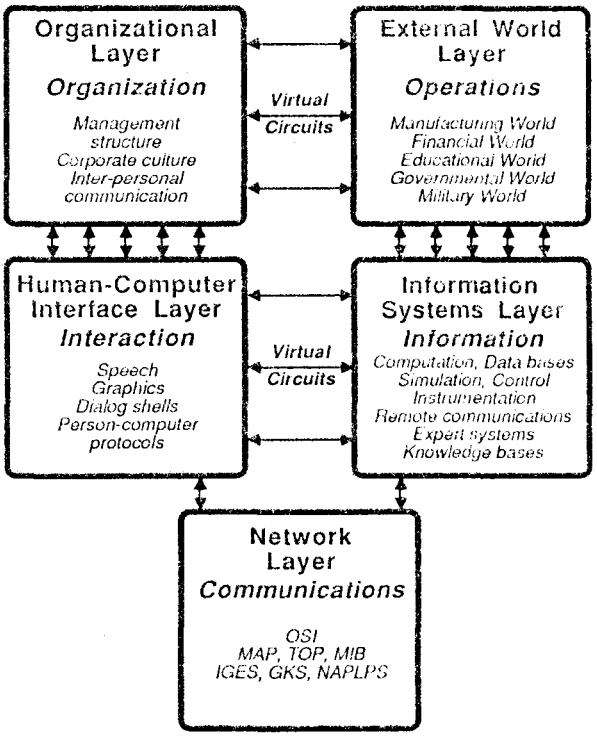


Fig.7 Integrated embedded information system layered protocol

It is a reasonable metaphor that the chief executive does not see himself as concerned with that level of communication but rather as communicating directly with the board and shareholders through a flow of profits and dividends. However, can these metaphors be given technical substance through extension of the layered hierarchical model of virtual communication paths?

The main problem in extending the model to the management level is that people are directly and necessarily involved, and they have different dynamics from technical systems. However, this is also true for the level below where human-computer interaction is involved, and models have been developed at this level that integrate the human and technical components with a hierarchical layered communication model (Taylor 1987). The following section gives systemic foundations for the analysis of integrated systems.

SYSTEMIC FOUNDATIONS FOR INTEGRATED SYSTEMS

In analyzing human-computer interaction it is important to establish what are the critical characteristics of a person and what are those of a computer. How are they differentiated from one another and from other forms of system? Habermas (1981) has emphasized the essential differences between the dynamics of society and that of physical systems. The "laws" that govern human behavior are largely conventions embedded in our society and propagated through our culture (Schutz & Luckman 1973). They do not have the necessity of the laws underlying physical systems and their analysis differs in many ways from those of physical systems (Ulrich 1983).

Human beings create their future through acts of choice that are constrained by their forward-looking intentions. Physical systems unroll their future through acts of necessity that are determined by their previous states. From this point of view, in human-computer interaction we are dealing with phenomena of the life-world which are essentially different from those of the physical world and cannot be encompassed by causal models.

This perspective is radically different from a behaviorist psychological model of people which sees them as conditioned through reinforcement into patterns of behavior that mimic the causal dependencies of physical systems (Skinner 1974, Mackenzie 1977). Habermas' point of view is more akin to a constructivist model of people which sees them as anticipating the future through the development of personal construct systems that are, however, always reconstructable (Kelly 1955).

It is reasonable to suppose that people exhibit both these phenomena, and that any model of human-computer interaction should be able to encompass those aspects of human behavior that are best modeled as causally based and those which are best modeled as anticipatorily based.

The computer component of the system seems straightforward because it has a physical, causal basis. However, computers are programmed by people and hence their behavior also involves human choice. Once they are operating the causal model is appropriate, but in considering system design and human-computer relationships it is important to take the human component into account. The users' mental models may be influenced by their impressions of the designers' intentions. The system designer may be seen as behaving through the computer to anticipate the future behavior of users.

To encompass the possibilities discussed, a conceptual framework is required that can account for these essential differences and similarities, differentiating people from technology and computing from non-programmed physical systems. Popper's (1968) *3 worlds theory* can be used to provide foundations for the distinctions underlying physical processes, human activities, and information technologies. He makes the three-fold distinction: "If we call the world of things—of physical objects—the first world and the world of subjective experience the second world we may call the world of statements in themselves the third world (world 3)." These concepts capture the differences between physical systems and people, and the role of abstraction in moving the distinctions into a third realm regardless of source.

Figure 8 shows the existential hypotheses underlying Popper's 3 worlds as a conceptual framework based on very general distinctions made about distinctions. It introduces a 'World 0' of distinctions in their own right before Popper's criteria are applied. This world is one where distinctions are made without reference to any criteria except that they are useful. It captures the notions of value and of the underlying anarchy that Feyerabend (1975) sees in the scientific process. Any distinction generates a system and hence that World 0 as defined here is that of general systems (Gaines 1980).

- World 0 arises from the assertion that there exist distinctions

—this defines a world of *systems*

—a key concept is that of *anarchy*, that distinctions may be made freely

—the conceptual framework involved is that of *axiology*, that distinctions have value

—truth in this world is *utilitarian*, that distinctions made are useful

—inference is *pragmatic*, based on chains of reasoning that appear to work

- World 1 arises from the assertion that there exist *physically necessary* distinctions

—this defines a world of *physical systems*

—a key concept is that of *actuality*, that the distinctions made are actual

—the conceptual framework involved is that of *epistemology*, that the distinctions have to be drawn from the actual world

—truth in this world is by *correspondence*, that the distinctions made correspond to the properties of physical systems

—inference is *causal*, based on chains of physical cause and effect

- World 2 arises from the assertion that there exist *humanly chosen* distinctions

—this defines a world of *mental systems*

—a key concept is that of *agency*, that distinctions have a maker

—the conceptual frameworks involved are those of *psychology and sociology*, that distinctions exist in the mental worlds of individuals and groups

—truth in this world is *performative*, that the distinctions chosen will be continue to be chosen

—inference is *conventionalist*, based on accepted chains of reasoning that are chosen within a culture

- World 3 arises from the assertion that there exist *formal systems* of distinctions

—this defines a world of *coherent systems*

—a key concept is that of *abstraction*, that systems of distinctions have their own existence

—the conceptual framework involved is that of *ontology*, that some distinctions are made independent of any criteria other than intrinsic structure

—truth in this world is by *coherence*, that the systems of distinctions have an acceptable intrinsic structure

—inference is *structuralist*, based on chains of reasoning that conform to the structure

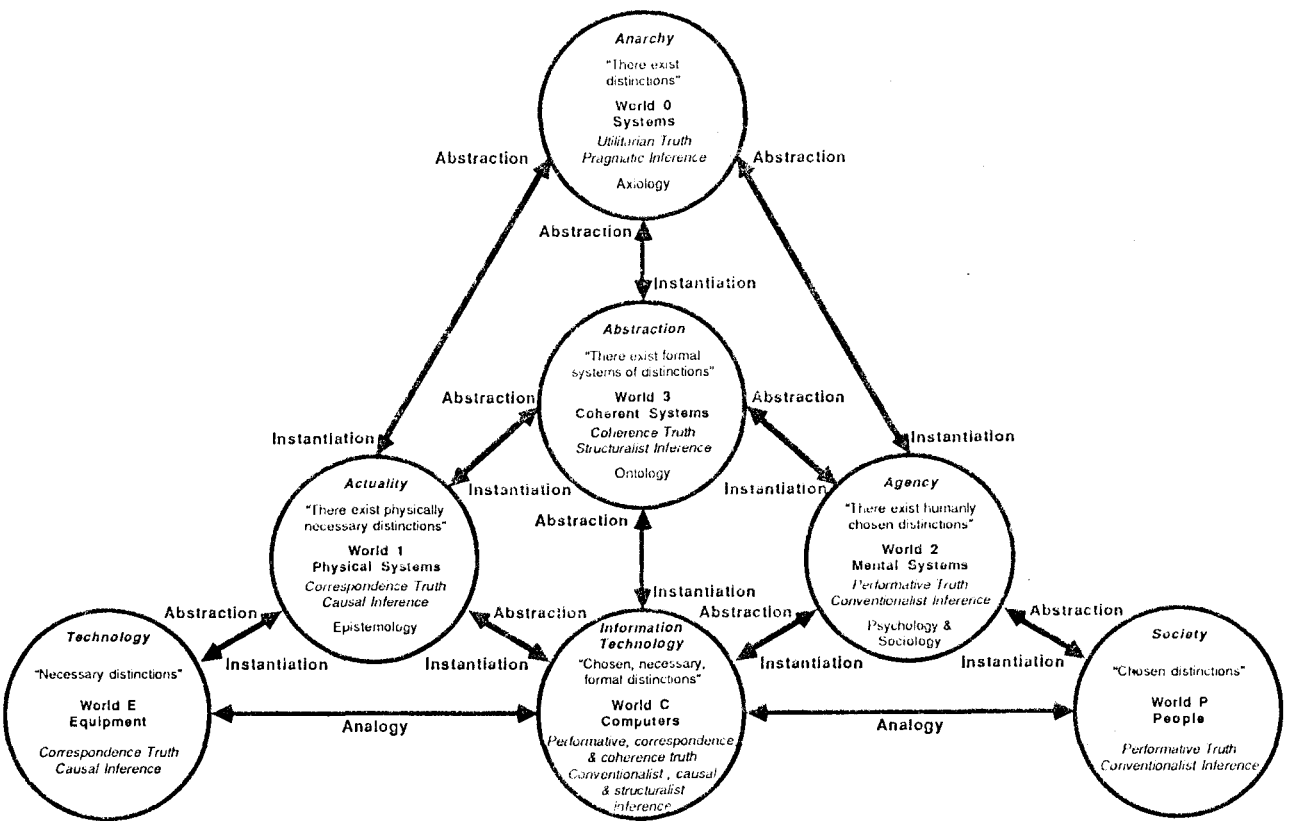


Fig.8 Knowledge structure for abstractions and analogies relating to equipment, computers and people

Worlds 1, 2 and 3 may be seen as components of World 0 and also as particular instances of it. Similarly Worlds 2 and 3 may be seen as instances of World 1. Conversely World 0 may be seen as an abstraction from Worlds 1, 2 and 3, and World 3 may be seen as an abstraction from Worlds 2 and 3. The conceptual framework of Figure 1 shows how the worlds relate through this *abstraction* and *instantiation*. It also shows the *analogy* between worlds 1 and 2 when we attribute agency to the causal dynamics of physical objects or necessity to the social conventions of human activity.

Figure 8 extends the conceptual framework to show equipment as an instance of a physical system, people as an instance of mental systems, and computers as an instance of both physical and mental systems, inheriting properties through both paths.

The peculiar property of electronic digital computers is that they are the archetype of deterministic causal systems, modeled in their behavior as finite-state automata whose next state and output are precisely determined by their current states and inputs. However, they are also programmed devices where the transition tables of the automaton are completely under the control of people—what computers do is what we chose them to do. Hence, they are also the archetype of performative, conventional systems, modeled in their behavior as the intentional artefacts of people, and whose next state and output are in major part determined by open choices made in programming.

This dual nature of computing systems underlies many of the problems and controversies in human-computer interaction. The computer is a technological system and it may be advisable to present it to its users as such, avoiding animistic considerations. However, the behavior of the computer is prescribed by people and this underlying prescription may also be evident to the user. The value systems and inter-personal attitudes of the system designer and programmer may have become embedded in the system behavior, making an animistic perspective unavoidable.

The performative truth and conventionalist inference aspect of computers is implicit in rules for programming effective human-computer dialog. For example (Gaines & Shaw 1984), "The system should be consistent in operation—the commands should always do the same thing throughout—the information presentation should always mean the same thing throughout" and "The system should be uniform in operation—the facilities which users have learned to use in one part of the package should be available to them in other parts if they might reasonably expect this", are both example of performative truths allowing inference from reasonable conventions.

The analogies between computers, people and equipment shown in Figure 8 are particularly interesting. If the conventions of the life-world become widely accepted in a culture then they assume the same status as the necessary distinctions of the physical world and the behavior of people within a culture may appear as completely causal. In informal terms, one may be just as hurt attempting to break a moral convention as attempting to walk through a brick wall. In human-computer interaction skilled operators may have highly practiced patterns of behavior that may be modeled through causal dynamics. They are no longer making choices at the lower levels of functioning. This is the basis for some negative "transfer of training" phenomena when functioning developed for one system is carried over to another where it may result in inappropriate choices. All technology has some aspects of human choice embodied in it so that the distinction between technology in general and information technology shown in Figure 8 is not a hard boundary.

The computer is also shown in Figure 8 as an instance of a World 3 system, that is an abstract entity defined in terms of its coherent structure. Category-theoretic models of software systems exemplify the possibility for a high degree of abstraction while still capturing essential features of computing systems (Goguen and Meseguer 1983) and analogies between them, and between them and people (Gaines 1975). The abstract World 3 entity will generally lose some features of both the World 1 physical dynamics that underlies its implementation and the World 2 psychological dynamics that underlies the choice of what to implement.

SYSTEM ORIGINS AND ACTIVITY

The principles underlying Figure 8 provide systemic foundations for the forms of reasoning by analogy and abstraction that are commonly used in the analysis and design of human-computer systems. To develop these concepts more formally and provide a framework for the overall dynamics of human-computer interaction, we need to introduce a further distinction, that between the activity and the origins of a system, between its behavior and its structure. In abstract terms the behavior of a system provides a description of what the system does, and the structure of a system provides a description of what the system is. One of the most important problems of system theory is the analysis of the relations between behavior and structure—in one direction, given the structure of a system, to derive its behavior—in the other direction, given the behavior of a system, to derive its structure. In the study of physical systems, mathematical techniques have been developed for moving in both directions with causal models (Gaines 1977, Klir 1985), that is interrelating the necessities of World 1 behavior and structure. However, as already noted, integrated computing systems involve the choice behavior of people and hence show phenomena of the life-world which are essentially different from those of the physical world and cannot be encompassed by causal models.

The dynamics of human behavior are best modelled as those of an *anticipatory system* (Rosen 1985), enhancing its survival by modeling the world, both passively and actively, in order to better anticipate the future. This corresponds to the choice component of World 2 phenomena, that agents are not bound by rigid necessity but can plan and chose certain aspects of their behavior. Figure 9 shows an analysis of the relations between a system, its origins and its activities, when the *actuality—agency* distinction of Figure 8 is also taken into account. The origins of a system have two components: its *causal structure* relating to how it was created; and its *anticipatory structure* relating to why it was created. The activities of a system also have two components: its *causal behavior* relating to how it carries out its activity; and its *anticipatory behavior* relating to why it carries out its activity.

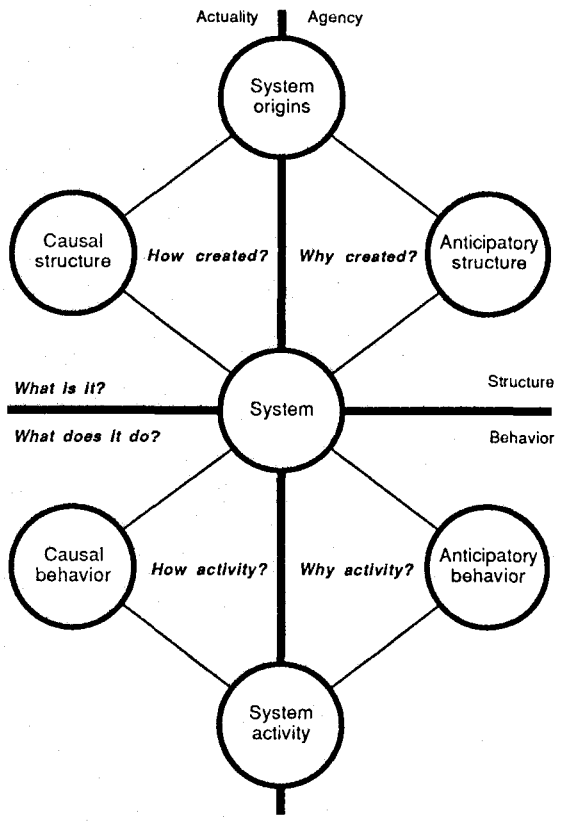


Fig.9 Structures defining system origins and activity

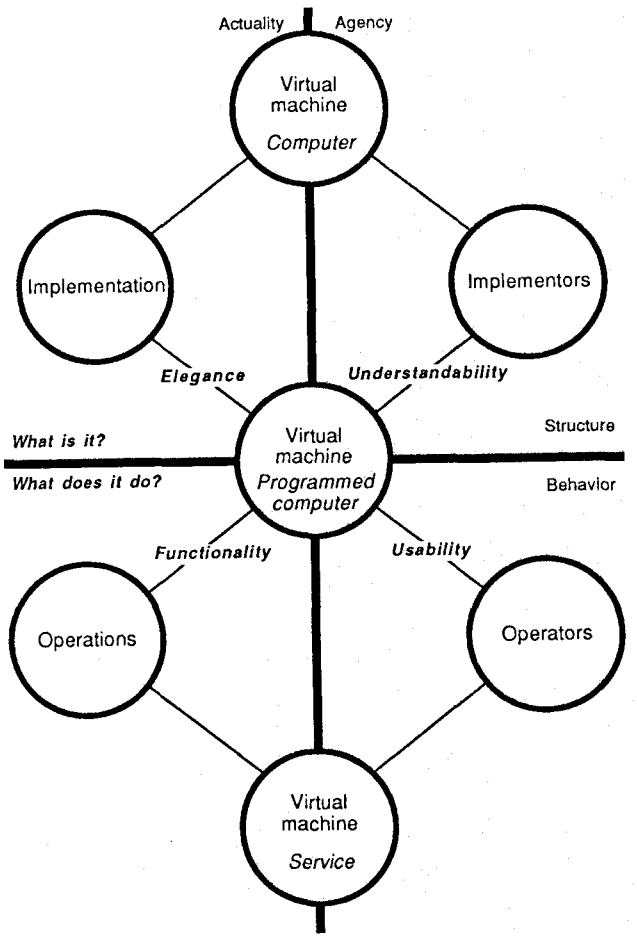


Fig.10 Virtual machines in terms of origins and activity

The basic distinctions of *actuality—agency* and *structure—behavior* are sufficient to define the virtual machine hierarchy fundamental to computing, and to account for its technical and human components and their relations. Figure 10 fills in Figure 9 to show how the levels of virtual machine arise from the basic distinctions. The origins of any computer-based service are that a *virtual machine*, a computing system, has been programmed by agents, the *implementors*, to create an actuality, the *implementation*. This results in a second virtual machine, the programmed computing system. The activities of the computer-based service are that a *virtual machine*, a programmed computing system, is being operated by agents, the *operators*, to create an actuality, the *operations*. This results in a third virtual machine, the service itself. The relations between the central virtual machine and its structural and behavioral components represent fundamental concepts in system implementation and operation (Edwards and Mason 1987):

- *elegance* considerations, capturing the notion of efficiency in design, concern the relation between the system virtual machine and the underlying resource with which it is implemented—this is another virtual machine capturing the characteristics of the high-level language, operating system, and so on, used in implementation;
- *understandability* considerations, capturing the notion of comprehensibility in design, concern the relation between the system virtual machine and the implementors responsible for creating it;
- *functionality* considerations, capturing the notion of capabilities provided, concern the relation between the system virtual machine and the tasks for which it is being implemented.
- *usability* considerations, capturing the notion of capabilities suitably interfaced, concern the relation between the system virtual machine and the operators who use it.

The logic underlying Figures 9 and 10 may be iterated to interpolate as much detail as required between the upper and lower levels—what is "activity" at one level becomes "structure" for the lower level. Figure 11 shows an analysis of the key features of a computer system in terms of six layers:

- At the top level the overall computer system originates in terms of purpose and structure as part of the culture within which it embedded. This *cultural layer* captures the milieu within which the system has been generated and can itself be subject to detailed analysis.
- At the next level the activities generated through this structure and purpose lead to a system of intentionality in which anticipatory activity leads to the acquisition of knowledge, or models of the world, and the generation of goals.
- At the next level the activities generated through anticipation lead to actions and communications—a linguistic distinction between activities directed to World 1 effects and World 2 effects, respectively. This *knowledge* layer captures the internal processes supporting the modeling and control activities of an anticipatory system.
- At the next level the activities generated through action and communication have to be transmitted to some external medium. This *action* layer captures the internal processes supporting the interfacing of an anticipatory system to the world.
- At the next level the activities generated through transmission have to be expressed in such a way as to have the desired effect in the world. This *expression* layer captures the internal processes supporting the encoding of actions and communications.

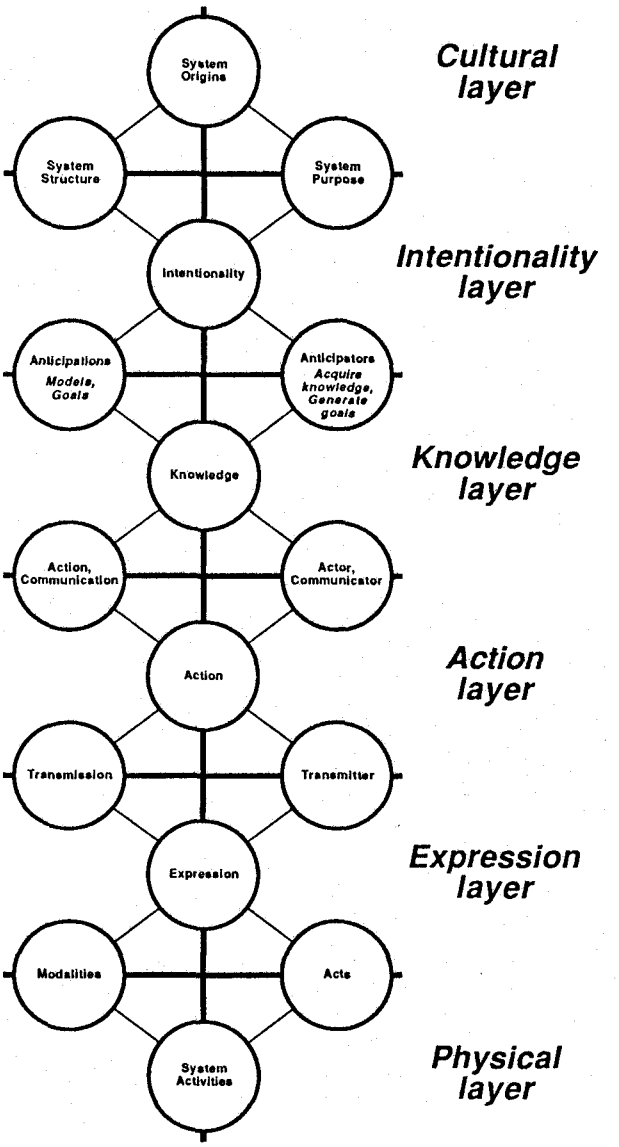


Fig.11 Hierarchy of layers in an intelligent system

- At the lowest level the activities generated through expression exist physically in the external world. This *physical* layer captures the external effects of the activities of an anticipatory system.

It is interesting to relate the iterated levels of Figure 11 back to the terminology of Figure 10. First, consider Figure 11 as a conceptual framework for a person. The culture is the 'implementor' of the person, taking the genetic 'virtual machine' and socializing it to have intentions and knowledge that conform with the purpose of a society. The intentions of the person are carried out by the formation of anticipatory systems that acquire knowledge and generate goals. These in turn implement goal-directed actors or communicators, which themselves implement processes to transmit the activities of the actors to the world, ultimately as sequences of acts that are directly expressed in the physical world. Second, consider Figure 11 as a conceptual framework for a computer system, an intelligent integrated system. The diagram from bottom up represents the evolution of computer technology, from the physical device technology to carry out acts, through the programmed control of action, through the knowledge-based derivation of action, through the goal-creating activities of intentionally-specified systems, to the origins of information technology in our culture at the top level.

CONCLUSIONS

The human-computer interface presents major problems in the specification, development and evaluation of computer-based systems. This paper has presented two frameworks within which to understand and overcome these problems:

- Modeling the role of human-computer interaction technologies in the development of information technology.
- Developing a fundamental model of embedded information systems in terms of an extension of layered protocols.

ACKNOWLEDGEMENTS

This paper has been influenced by discussions over many years and I am grateful to Walter Doherty, Jack Edwards, Peter Facey, Saul Greenberg, Martin McGuire, Mildred Shaw, Ben Shneiderman, Martin Taylor, Harold Thimbleby, Ian Witten and various anonymous referees. Financial assistance for this work has been made available by the Natural Sciences and Engineering Research Council of Canada.

REFERENCES

Ayres, R.U. (1968). Envelope Curve Forecasting. **Technological Forecasting for Industry and Government**. pp. 77-94. New Jersey: Prentice Hall.

CHFP (1986). Concept of human frontier program. Japanese Embassy (March).

Crane, D. (1972). **Invisible Colleges: Diffusion of Knowledge in Scientific Communities**. Chicago, Illinois: University of Chicago Press.

Day, J.D. & Zimmerman, H. (1983). The OSI reference model. **Proceedings IEEE, 71**(12), 1334-1340 (December).

Deal, T.E. & Kennedy, A.A. (1982). **Corporate Cultures: The Rites and Rituals of Corporate Life**. Reading, Massachusetts: Addison-Wesley.

Edwards, J.L. & Mason, J.A. (1986). Evaluating the intelligence in dialogue systems. **International Journal of Man-Machine Studies, 27**, to appear.

Fano, R.M. (1965). The MAC system: a progress report. Sass, M.A. & Wilkinson, W.D., Eds. **Computer Augmentation of Human Reasoning**. pp. 131-150. Washington D.C., USA: Spartan Books.

Feyerabend, P. (1975). **Against Method**. London: NLB.

Fuchi, K., Sato, S. & Miller, E. (1984). Japanese approaches to high-technology R&D. **Computer**, 17(3), 14-18 (March).

Gaines, B.R. (1975). Analogy categories, virtual machines and structured programming. Goos, G. & Hartmanis, J., Eds. **GI—5.Jahrestagung. Lecture Notes in Computer Science, 34**, 691-699. Berlin: Springer Verlag.

Gaines, B.R. (1977). System identification, approximation and complexity. **International Journal of General Systems, 3**, 145-174.

Gaines, B.R. (1980). General systems research: quo vadis? Gaines, B.R., Ed. **General Systems 1979**. Vol. 24, pp. 1-9. Kentucky: Society for General Systems Research.

Gaines, B.R. (1984). Perspectives on fifth generation computing. **Oxford Surveys in Information Technology**, 1, 1-53.

Gaines, B.R. (1986). Sixth generation computing: a conspectus of the Japanese proposals. **ACM SIGART Newsletter**, (95), 39-44 (January).

Gaines, B.R. & Shaw, M.L.G. (1984). **The Art of Computer Conversation: A New Medium for Communication**. New Jersey: Prentice Hall.

Gaines, B.R. & Shaw, M.L.G. (1986). A learning model for forecasting the future of information technology. **Future Computing Systems, 1**(1), 31-69.

Goguen, J.A. & Meseguer, J. (1983). Programming with parametrized abstract objects in OBJ. **Theory and Practice of Programming Technology**. Amsterdam: North-Holland.

Habermas, J. (1981). **The Theory of Communicative Action: Reason and the Rationalization of Society**. Boston: Beacon Press.

Hall, E.T. (1959). **The Silent Language**. New York: Doubleday.

Hansen, W.J. (1971). User engineering principles for interactive systems. **Proceedings of the Fall Joint Computer Conference, 39**, 523-532. New Jersey: AFIPS Press.

Karatsu, H. (1982). What is required of the fifth generation computer - social needs and impact. Moto-oka, T., Ed. **Fifth Generation Computer Systems**. pp. 93-106. Amsterdam: North-Holland.

Kelly, G.A. (1955). **The Psychology of Personal Constructs**. New York: Norton.

Kennedy, T.C.S. (1974). The design of interactive procedures for man-machine communication. **International Journal of Man-Machine Studies, 6**, 309-334.

Klir, G.J. (1985). **Architecture of Systems Problem Solving**, New York: Plenum Press.

Mackenzie, B.D. (1977). **Beahviourism and the Limits of Scientific Method**. London: Routledge and Kegan Paul.

Marchetti, C. (1981). Society as a learning system: discovery, invention and innovation cycles revisited. Report RR-81-29 (November). Laxenburg, Austria: International Institute for Applied Systems Analysis.

Martin, J. (1973). **Design of Man-Computer Dialogues**. New Jersey, USA: Prentice-Hall.

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

Norman, D.A. (1984). Stages and levels in human-machine interaction. **International Journal of Man-Machine Studies, 21**(3) (September).

Popper, K.R. (1968). Epistemology without a knowing subject. Van Rootselaar, B., Ed. **Logic, Methodology and Philosophy of Science III**. pp. 333-373. Amsterdam, Holland: North-Holland Publishing Co..

Rosen, R. (1985). **Anticipatory Systems**. Oxford: Pergamon.

Schutz, A. & Luckmann, T. (1973). **The Structures of the Life-World**. London: Heinemann.

Skinner, B.F. (1974). **About Behaviourism**. London: Jonathan Cape.

STA (1985). Promotion of R&D on Electronics and Information Systems That May Complement or Substitute for Human Intelligence. Tokyo: Science & Technology Agency.

Taylor, M.M. (1987). Layered protocols for computer-human dialogue: I Principles. **International Journal of Man-Machine Studies, 27**, to appear.

Ulrich, W. (1983). **Critical Heuristics of Social Planning**. Bern: Haupt.

Wasserman, T. (1973). The design of idiot-proof interactive systems. **Proceedings of the National Computer Conference, 42**, M34-M38. New Jersey: AFIPS Press.