Int. J. Man–Machine Studies (1987) **26,** 453–472

An overview of knowledge-acquisition and transfer

BRIAN R. GAINES

Department of Computer Science, University of Calgary, Calgary, Alberta, Canada T2N 1N4

A distributed anticipatory system formulation of knowledge acquisition and transfer processes is presented which provides scientific foundations for knowledge engineering. The formulation gives an operational model of the notion of expertise and the role it plays in our society. It suggests that the basic cognitive system that should be considered is a social organization, rather than an individual. Computational models of inductive inference already developed can be applied directly to the social model. One practical consequence of the model is a hierarchy of knowledge transfer methodologies which defines the areas of application of the knowledge-engineering techniques already in use. This analysis clarifies some of the problems of expertise transfer noted in the literature, in particular, what forms of knowledge are accessible through what methodologies. The model is being used as a framework within which to extend and develop a family of knowledge-support systems to expedite the development of expert-system applications.

Introduction

Growing recognition of the significance of knowledge-based computing systems has focused attention on processes of knowledge acquisition and transfer. Commercial application of expert systems is being impeded by the knowledge-engineering bottleneck and has led to the development of rapid prototyping tools. These have proved practically useful but the range of applications of existing tools is limited, and it is not clear what limitations are inherent in the prototyping tools, what arise from the shells and their knowledge representation and inferencing procedures, and what arise from our lack of understanding of the underlying processes of knowledge acquisition and transfer.

Knowledge acquisition and transfer processes in human society are complex and poorly understood. Many disciplines are involved in their study but none provides a comprehensive framework let alone overall theory. Neurology, psychology, linguistics, education, sociology, anthropology, philosophy and systems theory all have significant contributions to make. However, integrating them, disambiguating jargon, and combining different objectives and perspectives are major intellectual problems. The task is attracting increasing effort but this is unlikely to have significant short-term impact on knowledge-based system development.

Knowledge acquisition and transfer are now important aspects of advanced computer technology, and neither the pragmatic development of rapid prototyping techniques nor the dependence on partial theories from other disciplines can be regarded as satisfactory foundations. We need to be eclectic in acquiring techniques from all possible sources, but we also need to complement our pragmatism with effort to develop scientific foundations for our activities. We cannot be satisfied until

0020-7373/87/040453 + 20$03.00/0 © 1987 Academic Press Inc. (London) Limited

we have a precise understanding of the processes underlying expertise, its operation, acquisition and transfer. This is not a simple requirement since it entails understanding the nature of knowledge, its dynamics and application. The foundations of computer technology in the physical sciences are no longer adequate and need extension into the humanities. The philosophy, psychology and sociology of knowledge processes are highly significant to future computing, and we have to make theories operational and obtain precise answers to questions that have long been regarded as certainly controversial and possibly intractable.

This paper sketches a framework for encompassing the variety of processes and phenomena of knowledge acquisition and transfer that provides a convenient classification, an indication of overall theoretical foundations, and a guide to research methodologies.

Phenomena of expertise-acquisition and -transfer

The problem of expertise elicitation from a skilled person is well known in the literature of psychology (Nisbett & Wilson, 1977; Broadbent, Fitzgerald & Broadbent, 1986). Hawkins (1983) has analysed the nature of expertise and emphasizes its severe limitations and dependence on critical assumptions which are often implicit. Dixon (1981) has surveyed studies showing that much human activity is not accessible to awareness. Collins (1985) has studied knowledge-transfer processes among scientists and suggests that some knowledge may not be accessible through the expert, not only because he cannot express it, but also because he may not be aware of its significance to his activity. Bainbridge (1979, 1986) has reviewed the difficulties of verbal debriefing and notes that there is no necessary correlation between verbal reports and mental behavior, and that many psychologists feel strongly that verbal data are useless. Clinical psychologists see the problem as one of cognitive defences that impede internal communication, and have developed techniques of verbal interaction to identify underlying cognitive processes (Freud, 1914; Kelly, 1955; Rogers, 1967). These can be used to bypass cognitive defences, including those resulting from automization of skilled behavior. Hayes-Roth, Waterman & Lenat (1983) have given guidelines about knowledge transfer from experts to expert systems, and Welbank (1983) has surveyed the psychological problems of doing this.

The main problems identified in accessing an expert's knowledge are:

"Expertise" may be fortuitous. Results obtained may be dependent on features of the situation which the expert is not controlling;

Expertise may not be available to awareness. An expert may not be able to transmit the expertise by critiquing the performance of others because he is not able to evaluate it;

Expertise may not be expressible in language. An expert may not be able to transmit the expertise explicitly because he is unable to express it;

Expertise may not be understandable when expressed in language. An apprentice may not be able to understand the language in which the expertise is expressed;

Expertise may not be applicable even when expressed in language. An apprentice

may not be able to convert verbal comprehension of the basis of a skill into skilled performance;

"Expertise" expressed may be irrelevant. Much of what is learnt, particularly under random reinforcement schedules, is superstitious behavior that neither contributes nor detracts from performance;

Expertise expressed may incomplete. There will usually be implicit situational dependencies that make explicit expertise inadequate for performance;

"Expertise" expressed may incorrect. Experts may make explicit statements which do not correspond to their actual behavior and lead to incorrect performance.

In the development of knowledge engineering methodologies and rapid prototyping techniques for expert systems the emphasis has been on interviewing experts, and hence on linguistic transmission of expertise. A rich variety of alternative methods for expertise transfer exists in human society:

Expertise may be transmitted by managing the learning environment. A trainer may be able to establish effective conditions for an apprentice to acquire expertise without necessarily understanding the skill or himself being expert in it;

Expertise may be transmitted by evaluation. A trainer may be able to induce expertise by indicating correct and incorrect behavior without necessarily understanding the skill in detail or himself being expert in its performance;

Expertise may be transmitted by example. An expert may be able to transmit a skill by showing his own performance without necessarily understanding the basis of his expertise.

There are many sources of expertise that do not involve others:

Expertise may be acquired by trial and error learning. This is the basic inductive knowledge acquisition process that is always in operation although heavily overlaid by the social transfer processes already discussed;

Expertise may be acquired by analogical reasoning. The transfer of models and skills from one situation to another is an important source of knowledge;

Expertise may be acquired by the application of general laws and principles to new situations. The use of physical laws and systemic principles to generate specific expertise is the basis of scientific and engineering expertise.

These many aspects of the problems of knowledge acquisition and transfer are further confounded by the combinatorial possibilities resulting from the recursive nature of knowledge processes. We can acquire knowledge about knowledge acquisition. We can mimic the behavior of an expert coach managing a learning environment. We can give evaluative feedback on analogical reasoning or the application of principles and laws. We can express the principles behind the effective management of learning in specific domains. The knowledge processes of human society are rich and complex, and, in practice, many of these possibilities will be instantiated as parallel, interrelated activities in any knowledge acquisition and transfer situation.

Socio-systemic foundations of expertise

Figure 1 shows some of the relations between a knowledge engineer, or knowledge transfer system, an expert in a problem domain, and the physical, social, knowledge and problem environments.

An "expert role" has been shown rather than an expert person to emphasize that expertise is focused on specific task domains and is not an indication of global capability. Within that role, various significant psychological processes have been noted: perception, action and language as interfaces with the world; memory, awareness, intentions, emotions and attention as associated internal phenomena. Some of these processes go beyond the basic information-processing models of current cognitive science (Norman, 1980), but they are essential to knowledge acquisition and transfer.

The specific problem environment for the expert is shown as intersecting three general environments that are common to most problems: physical, social and knowledge. The knowledge engineer is part of the social environment but outside the problem environment. He interacts with the expert and also with his environments. For example, if the expert is performing a task in the physical environment then the knowledge engineer will attempt to understand the basis of that task, observe the expert performing it, and may manage the task environment to create scenarios in which to observe specific features of the skill.

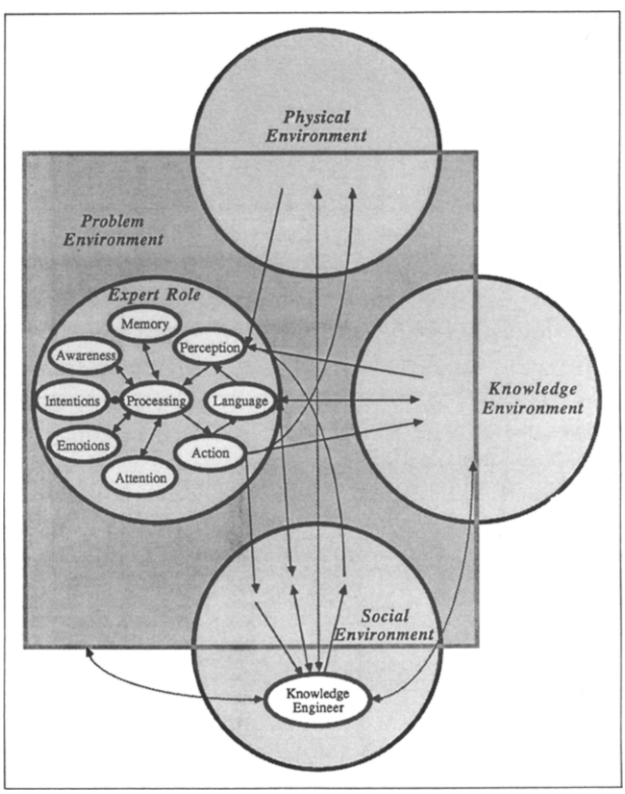

FIG. 1. Relations between knowledge engineer, expert and environments.

The three environments correspond to Popper's (1968) basic delineation of three *worlds*:

World 1, the physical environment, is the world of "things", of physical objects and their dynamics;
World 2, the social environment, is the world of "subjective experience", of people and the dynamics of their mental experience;
World 3, the knowledge environment, is the world of "statements in themselves", of theories and the dynamics of their logical development.

All the phenomena of expertise acquisition and transfer discussed in the previous section can be represented and analysed in terms of this diagram. The expert role develops through interaction with the three environments shown: direct experience of the dynamics of the physical, social and knowledge environments; learning by example, evaluation and linguistic communication from other experts in the social environment; managed experience and transmission; and so on. Figure 1 provides a general framework for knowledge acquisition and transfer, including knowledge engineering for expert systems.

However, the diagram only gives expression to the phenomena, making some significant distinctions, but giving no underlying model for the processes involved. The remainder of the paper develops such an overall model of knowledge acquisition and transfer.

Humanity as a distributed anticipatory system

It is difficult to know where to begin in analysing human knowledge processes. The theoretical foundations to be developed are based on the concept that humanity can be characterized as a *distributed anticipatory system*—a coordinated collection of autonomous systems whose primary dynamic is the anticipation of the future. The motivation for this concept links back to the processes underlying the survival of the species. To survive, a living system must have access to the necessary resources and be capable of coping with threats to its survival. Systems evolve to maximize their access to resources and minimize their vulnerability to threats.

If the universe were static a simple model of resource availability and prey/predator relations would determine the dynamics of living systems. Until the advent of humanity, this planet was a static universe over long periods of evolutionary time. The beginnings of the humanity were set in this static universe but its activities soon began to change that universe so that uncertainty about the future began to dominate our survival processes. Changes in the earth that would have taken millions of years began to occur over millenia, and now they occur within our lifetimes. Humanity developed resources far beyond their natural availability, extinguished most predators other than itself, and changed the ecology of the earth. Much of this planet is now a human construct and the distinction between natural and artificial has become meaningless.

Figure 2 shows the system dynamics underlying humanity. The shaded areas show the response to uncertainty. The keys to survival in an uncertain world are three-fold: first to maximize territorial dispersion so that some part of the system is outside the range of a threat; second to maximize cultural diversity so that some part

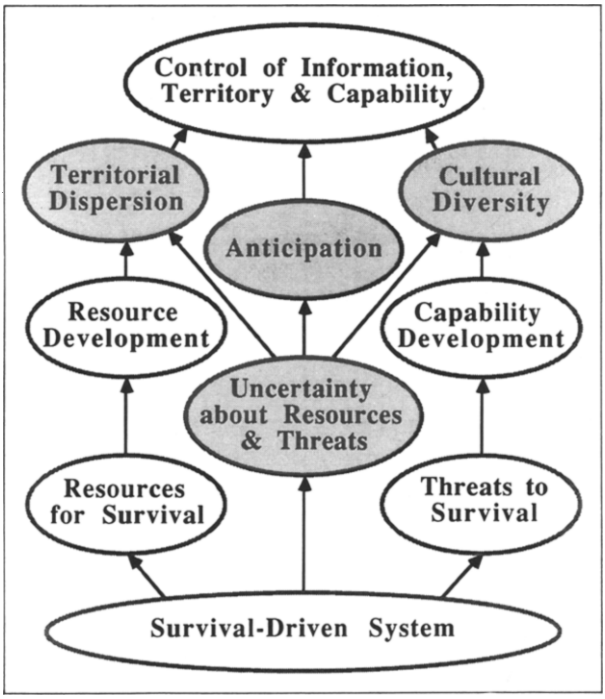

FIG. 2. Dynamics of a survival-driven system.

of the system has the capability to survive a threat; third to improve anticipation of the future, passively to predict threats in advance, and actively to rebuild the universe so that they do not occur.

Figure 2 encompasses the key technological phenomena of our time. Information technology increases control of information, giving computing when implemented in silicon and genetic engineering when implemented in DNA. The space program aims to achieve territorial dispersion and access to new resources. Genetic engineering aims to achieve cultural diversity and improvement of existing resources. Information technology is being used to model our anticipatory processes, leading to their greater understanding through cognitive science and their enhancement through artificial intelligence.

Characterizing humanity as a distributed anticipatory system follows from the logic of Fig. 2. The survival-driven system is the species rather than the individual, and the main knowledge-acquisition system is the species. Humanity has achieved territorial dispersion and cultural diversity through its structure as distributed collection of autonomous entities. Language has evolved as a means of coordinating the activities of these autonomous entities. It is essential to many of the processes of knowledge acquisition and transfer. Interaction with the social and knowledge environments is at least as important as interaction with the physical world. The evolutionary pressures have been very strong in selecting genes giving the capability for the species to act as a single distributed individual, combining autonomy and cohesion through enhanced linguistic communication.

Knowledge acquisition as a modeling process

The characterization of humanity as a distributed anticipatory system, modeling the world so as to increase its probability of survival, places the systematic foundations of knowledge processes in theories of modeling. Klir (1976, 1985) has provided the most general account of modeling, and proposed an epistemological hierarchy accounting for the main components of any modeling systems and their interrelations. Gaines (1977) gave a mathematical formulation of the general problem of modeling as a relation between order relations at different levels of this hierarchy. The hierarchy has proved a valuable conceptual tool in analysing a wide variety of modeling systems both in terms of their ontological presuppositions and their epistemological processes.

Underlying the modeling hierarchy is the notion of a distinction as the primitive concept underlying the representation of knowledge. It is a sufficient primitive to give foundations for systems theory including the notion of a system itself (Gaines, 1980). In its psychological form, as a personal construct (Kelly, 1955), the notion has been used to derive very effective techniques for knowledge transfer from experts to expert systems (Boose, 1985). Its foundational role in knowledge acquisition is evident in the hierarchical representation of distinctions in a modeling system shown in Fig. 3. The levels of the hierarchy itself are the results of distinctions that we make so that no additional primitives are introduced—in Klir's (1976) terminology:

> The source system is distinguished as those distinctions that the particular modeling system makes—it is a distinction about distinctions defining the construct system of an individual;

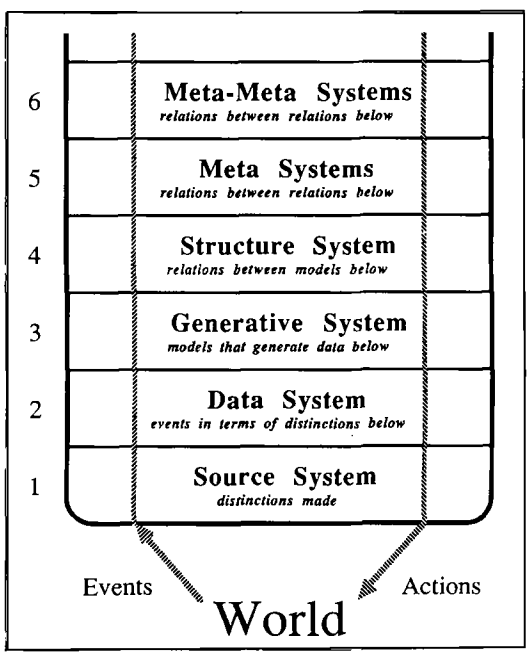

FIG. 3. Epistemological hierarchy of a system modeling a world.

The data system is distinguished as those distinctions that have been made about a particular event—a distinction about distinctions defining an event;

The generative system is distinguished as a set of distinctions that also defines an event—these are model-generated rather than event-generated—it is the match between the model-generated and event-generated distinctions that determines the degree of approximation of the model to the world—this is a distinction about distinctions among distinctions that defines goodness of fit;

The structure system is distinguished as a set of distinctions that compare models—it is the order relation of simplicity/complexity on models that determines the preference for the simplest model that is an adequate approximation to the world—this is a distinction about distinctions that defines our preference for simple models;

The meta system is distinguished as a set of distinctions that specify the basis of these comparisons;

The meta–meta system, and higher levels, are distinguished as sets of distinctions that specify further relations among the distinctions on the level below.

Note that the upper levels of modeling are totally dependent on the system of distinctions used to express experience through the source system.

Distinctions are not just static partitions of experience. They may be operations: actions in psychological terms; processes in computational terms. Whether a system finds a distinction in the world, imposes it passively as a view of the world, or imposes it actively as a change in the world, is irrelevant to the basic modeling theory. It makes no difference to the theory whether distinctions are instantiated through sensation or action. In knowledge engineering we have to incorporate both the expert's prediction and control processes.

Knowledge acquisition in the hierarchy

The hierarchy of Fig. 3 accounts for knowledge acquisition as the modeling of events enabling adequate prediction and action. A modeling schema results from distinctions about distinctions at each level in the hierarchy. In prediction the key distinction is to what degree a level accounts for the information flowing through it and hence this distinction may be termed one of *surprise* (Gaines, 1977), in the sense used by the economist Shackle (1955). Surprise goes in opposition to the degree of membership of a predicted event to an actual event and the expected surprise is a form of entropy. Surprise at the lowest level of the hierarchy corresponds to distinctions being inadequate to capture events; surprise at the next level to inadequate variety to experience events; at the next level to inadequate approximation to predict events; at the next level to inadequate simplicity to explain events; at the next level to inadequate comprehensiveness to account for events.

The formal theory of modeling is one in which models are selected at each level down the hierarchy to minimize the rate at which surprise is passing up the hierarchy. The criteria for model selection independent of the data are generally thought of as being ones of simplicity/complexity: of two models which fit the data equally well choose the simplest. However, notions of simplicity/complexity are not well-defined nor intrinsic to the class of models. The simplicity/complexity ordering is arbitrary and in its most general form is just one of *preference*. Hence the general modeling schema is one in which surprise flows up the hierarchy and preference

flows down. In situations that are mathematically well defined, such as determining the structure of a stochastic automaton from its behavior, such a model schema gives the correct results (Gaines, 1977). Conversely, the success of the schema in stabilizing with regard to a given world defines the characteristics of that world.

Thus the basic modeling schema for learning from experience is one in which surprise flows up the hierarchy and preferences flow down. In primitive organisms only the lower levels of the hierarchy are developed, surprise is generated from experience and preference is genetically encoded. In higher organisms the modeling process generalizes both surprise and preference to cope with novel environments. Humanity has developed the upper levels of the hierarchy and detached surprise from experience and preference from its genetic roots. Surprise can flow up from a level without flowing into it from below because the processes at that level have generated novelty. Preference can be generated at a high level detached from both experience and genetic origins and flow down to affect the relations of the organism to the world.

There is neurological and behavioral evidence of the existence within the brain of the two channels of communication shown in Fig. 3 (Tucker & Williamson, 1984). The arousal system passes surprise up to the cortex from the limbic region as unexpected events occur. The activation system passes preferences down from the cortex to the motor regions.

Cognitive-science interpretation of the hierarchy

The loop in Fig. 1 from events through distinctions up through the modeling hierarchy and then down again to predictions and actions corresponds to the epistemological model of man as an anticipatory system developed by Kelly (1955). Thus, the systemic hierarchy of Fig. 3 has an analog in psychological terminology as shown in Fig. 4:

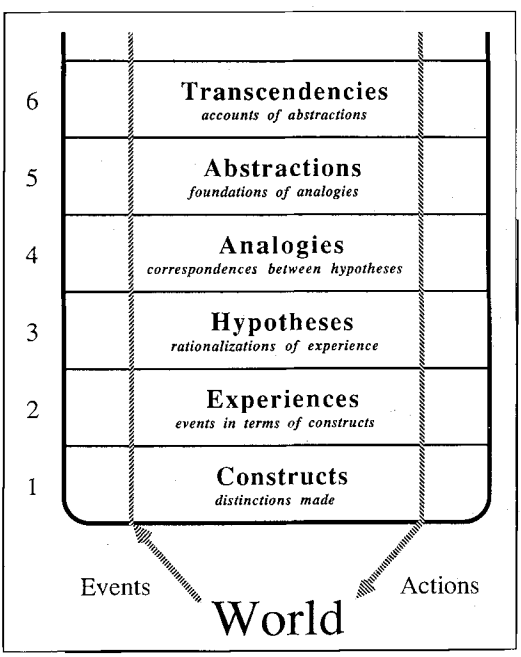

FIG. 4. Construction hierarchy of a person modeling a world.

The source level is one of constructs, distinctions made in interacting with the world;
The data level is one of experiences, events which happen to us, and we make happen, in terms of the distinctions already made;
The generative level is one of hypotheses which are rationalizations of experience;
The structure level is one of analogies which are correspondences between these rationalizations;
The meta level is one of abstractions which are foundations of analogies;
The meta–meta level is one of transcendencies which are accounts of abstractions.

Expertise, roles, groups and society

The anticipatory processes of the modeling hierarchy may be extended to the operation of society by viewing groups as cross-sections comprising multiple individuals. This concept may be given deeper significance by considering the inductive inference process underlying knowledge acquisition. Whereas deductive logical inference is well understood and well founded, the inductive inference that underlies knowledge acquisition is not. Deduction guarantees to take us from valid data to valid interfaces, but the inferences are thereby part of the data—no new knowledge is generated. Induction takes us from valid data to models of that data go beyond it—by predicting data we have not yet observed, and by giving explanations of the data in terms of concepts that are unobservable. Induction generates new knowledge but, as Hume (1739) pointed out over 200 years ago, the process is not deductively valid and it is a circular argument to claim that it is inductively valid.

Philosophers have continued to debate Hume's arguments and search for justification of the inductive process. Goodman (1973) proposed that we accept the circularity but note that it involves a dynamic equilibrium between data and inference rules as shown in Fig. 5: "A rule is amended if it yields an inference we are unwilling to accept; an inference is rejected if it violates a rule we are unwilling to amend". Rawls (1971) in his theory of justice terms this a reflective equilibrium. Stich & Nisbett (1984) noted flaws in Goodman's argument and repaired them by proposing that the equilibrium is social not individual: "a rule of inference is justified if it captures the reflective practice not of the person using it but of the appropriate experts in our society". This gives an operational definition of the role of experts within an anticipatory system as referential anticipatory sub-systems coordinating the overall knowledge-acquisition process. It also leads to an concept of the role of expert systems in society, as making the referential process more overt and widely available.

The extension of the modeling hierarchy to social processes is straightforward since Fig. 3 presents a general modeling schema and applies as much to groups of people, companies and societies as it does to the roles of a person. The epistemological hierarchy of a person is a cross-section of the epistemological hierarchy of the society generating their life-world. Pask's (1975) concept of P-Individuals as the basic units of psycho-socio-processes allows roles, people, groups, organizations and societies to be treated in a uniform framework. An individual is defined in cognitive terms as a psychological process (Pask, 1980) and more complex

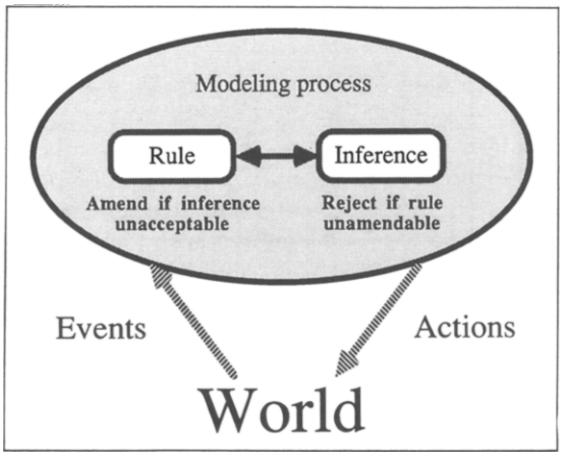

FIG. 5. Reflective equilibrium in inductive inference.

psychological and social structures may be defined similarly by taking into account the possibilities of timesharing, process switching and distributed processing with psychological processors. For example, one person may assume many psychological roles (process switching), whereas a group of people working together may act as a single goal-seeking entity and hence behave as one process (distributed processing).

Representation of expert skills within the hierarchy

In the analysis of expertise, the skills to achieve goals in the world are the crucial capabilities of the modeling system, and the usual hierarchical models of skills (Powers, 1973; Rasmussen, 1983) are naturally subsumed within the modeling formulation. Figure 6 shows the basis for action at different levels in the modeling hierarchy:

- At level one, the activation of a construct may be linked directly to a primitive act, another construct. This corresponds to reflex actions and stimulus-reponse connections. In system-theoretic terms this level might be implemented by conditional probability calculations giving confirmation-theoretic inductive inference;
- At level two, constellations of experience may be linked to complete action sequences through rules derived from similar experience. In system-theoretic terms this level might be implemented by fuzzy production rules giving generalization-based indictive inference. These constellations may be regarded as prototypical schema, or frame-like structures in computational terms;
- At level three, a generative model of experience, may be used to compute an optimal action sequence. In system-theoretic terms this level might be implemented by a state-based modeling scheme giving model-based inductive inference.

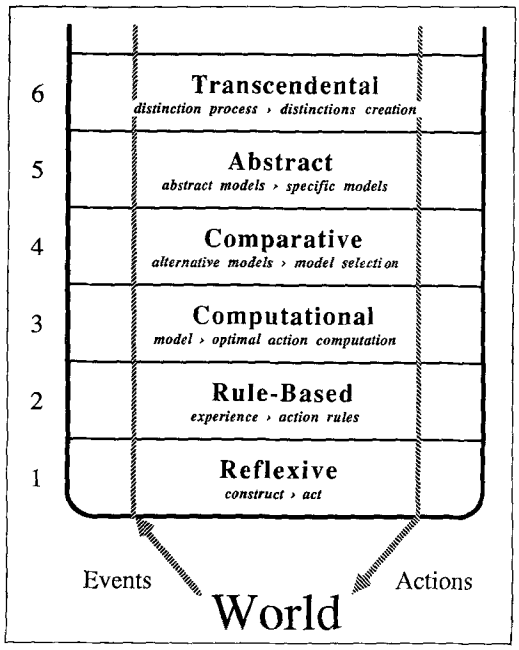

FIG. 6. Action hierarchy of system modeling a world.

At level four, a variety of alternative models may be compared as a basis for selecting one appropriate to the required goals. In system-theoretic terms this level might be implemented by a category-theoretic functional analysis scheme giving analogical inductive inference.

At level five, generalized abstract models may be used as templets from which to instantiate one appropriate to the required goals. In system-theoretic terms this level might be implemented by a category-theoretic factoring scheme abstracting the mathematical form of an analogy and giving abstractional inductive inference.

At level six, the entire process described may be transected through a recognition that it is based on distinctions being made at various level, and an attempt to rationalize these distinctions and create new ones. In system-theoretic terms this level might be implemented by a distinction-based analysis scheme giving what might be termed transcendental inductive inference.

It is an interesting comment on the state of the art in computer science that it has proceeds "middle–outward" in its representation of the knowledge involved at different levels of the hierarchy. Information technology has been primarily concerned with level three activities, and is only now beginning through expert systems to emulate level two activities.

Language and culture in knowledge acquisition

The creation of new knowledge takes place through the surprise/preference flows within the hierarchy and it is these processes that determine the rate of scientific

discovery, technological invention and product innovation. The capability for an entire society to act as a distributed knowledge acquisition system depends on communication processes to coordinate activity at a given level of the hierarchy across different people. This process whereby each person does not have to undertake all aspects of the inductive process but can share the results of such processing by others supports what is generally termed the culture of a society (Vanderburg, 1985). People use language for much of this communication but they also have in common with other animals the capability to communicate cultural information without the use of language. Mimicry is an important mechanism for knowledge acquisition as is reinforcement through reward and punishment.

The human development of language enables coordination to take place in a rich and subtle fashion that greatly enhances, but does not replace, the more basic mechanisms in common with other species. It is particularly important at the upper levels of the hierarchy where direct expression is difficult. From an individual point of view, language is a way of bypassing the normal modeling procedures and interacting directly with the system at any level. In particular it can directly affect the preference system. Much skilled activity is not directly accessible through language, but even when language cannot mediate the direct transmission of knowledge it may be used to achieve the same effect by the indirect support of other mechanisms. For example, one can describe a good learning environment, or a behavior in sufficient detail for mimicry.

Figure 7 shows the cultural support for knowledge acquisition at different levels in the modeling hierarchy:

The reflexive knowledge at level one has no verbal component and comes directly from experience, often that of mimicking the behavior of others. This level has been termed informal to correspond to Hall's (1959) definition of cultural transmission of behavior of this type.

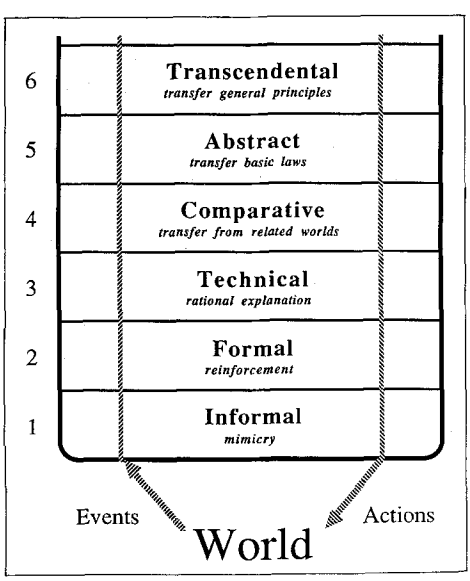

FIG. 7. Cultural transmission hierarchy of people modeling a world.

The rule-based knowledge at level two is usually transmitted by reinforcement of behavior, verbal rules, or is induced from the patterns of knowledge at level 1. This level has been termed formal to correspond to Hall's definition of cultural transmission of behavior of this type.

The computational knowledge at level three is usually transmitted by technical explanation, or is induced from the patterns of knowledge at level two. This level has been termed technical to correspond to Hall's definition of cultural transmission of behavior of this type.

The comparative knowledge at level four is usually transmitted by simile and metaphorical analysis, or is induced from the patterns of knowledge at level three. Hall does give a name to this level but the term comparative captures his own activity of highlighting the features of one culture by contrasting it with others.

The abstract knowledge at level five is usually transmitted through mathematical representation, or is induced from the patterns of knowledge at level four.

The transcendental knowledge at level six is usually transmitted by general system-theoretic analysis, or is induced from the patterns of knowledge at level five. Many mystical and consciousness-raising techniques may be seen as attempts to communicate knowledge at this level when formal analysis is impossible. It involves moving outside the framework established at the lower levels. Pope (1984) has given examples of this process in a wide range of cultures.

Knowledge-transfer processes

The action and cultural transmission hierarchies of Figs 6 and 7 may be combined to give an overall framework for knowledge representation, processing, acquisition and transmission as shown in Fig. 8. This gives detailed foundations for the processes of knowledge engineering discussed in the first part of this paper and shown in Fig. 1. The primary learning process is one of direct acquisition based on attempts to solve tasks through interaction with a world, as shown in the lower part of Fig. 8. This is applicable to all three worlds: physical, social and knowledge.

This primary process is the only one open to humanity as a species and is distributed across partially autonomous individuals coordinated by cultural communication processes. From an individual perspective it is these communication processes that dominate knowledge acquisition and transfer.

The right hand side of Fig. 8 shows the cultural communication processes supporting knowledge transfer:

At level one mimicry by watching experts is the main transfer mechanism;

At level two reinforcement by working under expert supervision is the main transfer mechanism;

At level three rational explanation from interviewing experts and books is the main transfer mechanism;

At level four use of analogous models to derive knowledge from similar situations is the main transfer mechanism;

At level five use of basic laws to derive specific models is the main transfer mechanism.

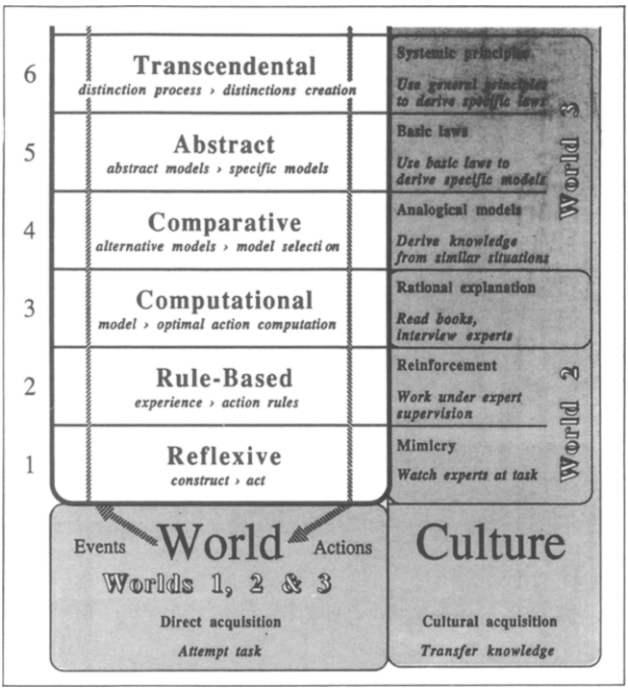

FIG. 8. Knowledge transfer processes in three worlds.

At level six use of system principles to derive specific laws is the main transfer mechanism.

Levels 1 and 2 of the cultural transmission process, mimicry and reinforcement, clearly belong to world 2, the social environment of subjective experience. Levels 4 through 6, analogies, laws and principles, clearly belong to world 3, the knowledge environment of statements in themselves. Level 3, of rational explanation, bridges both worlds.

Knowledge transfer methodologies

One practical consequence of the analysis of knowledge acquisition and transfer given above is a hierarchy of knowledge transfer methodologies based on Fig. 8 which defines the areas of application of the knowledge engineering techniques already in use. Figure 9 shows this hierarchy in relation to the basic architecture of an expert system:

Knowledge base: the central feature is a *knowledge base* of facts and inference rules;

Expert-system shell: the lower oval shows the operational system for applying the knowledge base to knowledge support.

The inference system derives the consequences of facts about a particular situation.

The planning system determines how to use the inference system to satisfy the specified objectives.

The explanation system answers questions about the basis of the inferences made.

Knowledge-acquisition system: the upper oval shows the knowledge transfer processes that may be used to establish the knowledge base.

The knowledge generation system implements knowledge acquisition by the raw induction of models from experience without cultural support as exemplified by ATOM (Gaines, 1977) and AM (Davis & Lenat, 1982).

The expertise modeling system implements knowledge acquisition by mimicking an expert's behavior as exemplified by INDUCE (Michalski & Chilausky, 1980).

The performance reinforcement system implements knowledge acquisition from performance feedback as exemplified by the PERCEPTRON (Rosenblatt, 1958) and STeLLA (Gaines & Andreae, 1966).

The knowledge-elicitation system implements knowledge acquisition by interviewing experts as exemplified by ETS (Boose, 1985), KITTEN (Gaines & Shaw, 1986), MORE (Kahn, Nowlan & McDermott, 1985) and SALT (Marcus, McDermott & Wang, 1985).

The knowledge-structuring system implements knowledge acquisition by analogy as exemplified by TEIRESIAS (Davis & Lenat, 1982) and CYC (Lenat, Prakash & Shepherd, 1986).

The basic laws system implements knowledge acquisition by building physical models as exemplified by simulation languages such as SIMULA (Nygaard & Dahl, 1981).

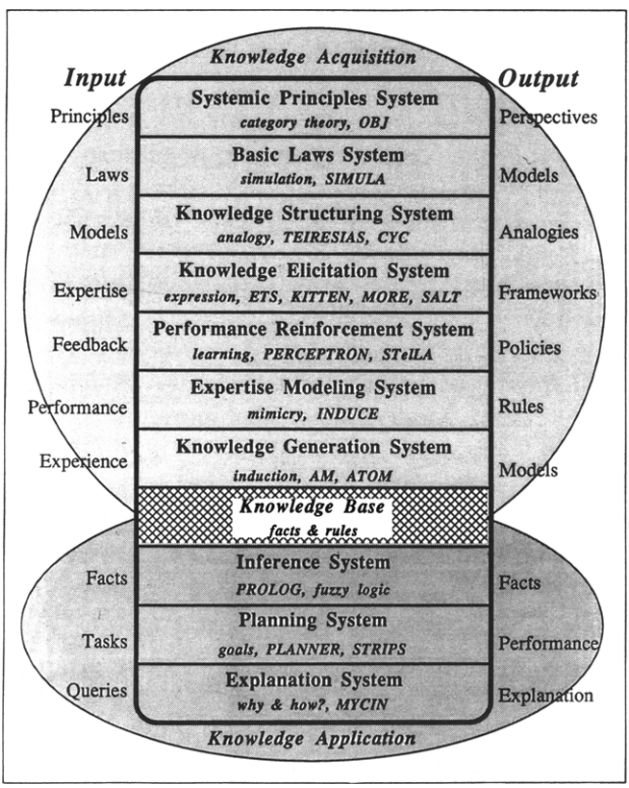

FIG. 9. Knowledge acquisition hierarchy for expert systems.

The systemic principles system implements knowledge acquisition by derivation from abstract principles as exemplified by the category-theoretic language OBJ (Goguen & Meseguer, 1983).

Most current expert-system shells offer a limited knowledge base represented in terms of production rules and frames, have weak planning capabilities limited to deriving consequences and testing assertions, and give simple explanations in terms of the facts and rules used. However, rapid improvements are taking place in knowledge representation, inferencing and explanation capabilities of expert systems (Hayes-Roth, 1985; Michalski & Winston, 1986; Pearl, 1986; Swartout, 1985).

Most current knowledge acquisition is done by manual knowledge elicitation. However, rapid improvements are taking place in the knowledge automatic knowledge-acquisition capabilities of expert systems (Boose, 1985, 1986; Bradshaw & Boose, 1986; Gaines & Shaw, 1986; Kahn *et al.*, 1985; Marcus *et al.*, 1985; Shaw & Gaines, in press; van de Brug, Bachant & McDermott, 1985).

The knowledge base and shell have been incorporated with the acquisition methodologies in Fig. 9 to show the natural relationships between all three. Shell structure and knowledge-acquisition research and development have so far been treated as separate enterprises. Integrating the two is essential to future knowledge-based system applications so that the combined system provides maximal and coherent support to the expert and knowledge engineer.

Conclusions

The distributed anticipatory system formulation of knowledge acquisition and transfer processes outlined in this paper provides scientific foundations for knowledge engineering:

It provides an operational model of the notion of expertise and the role it plays in our society, and hence provides formal foundations for expert systems;

It provides a detailed structure for the cognitive processes involved in knowledge acquisition and transfer;

There are links between the model and neurological processes in the brain;

It suggests that the basic cognitive system that should be considered is a social organization, rather than an individual;

Computational models of inductive inference already developed can be applied directly to the social model;

Cultural knowledge acquisition and transmission processes are explicitly identified and structured within the model;

Linguistic communication coordinates the distributed autonomous units forming the anticipatory system;

It suggests that groups of experts, rather than individuals, should be used in developing expert systems.

It provides a framework within which all the methodologies for knowledge acquisition and transfer currently under developed can be incorporated, analysed and compared.

There are several implications for future developments in Figs 8 and 9. Whereas past knowledge acquisition and transfer systems have been developed using

techniques at one level of the hierarchy, future systems will increasingly follow the example of human apprenticeship and operate at multiple levels. The combination of teaching by example, performance feedback and explicit statement is ubiquitous in human knowledge transmission and will become increasingly so in expertise transfer systems. The combination of rule-based, model-based, analogical and law-based inference is ubiquitous in human knowledge processing and will become increasingly so in expert systems. A companion paper describes the development of an integrated knowledge acquisition and transfer system, KITTEN, based on these concepts (Shaw & Gaines, in press).

The range and variety of expertise transfer systems being developed is very important. They are complementary rather than competitive. We are entering a phase of system integration when the best features of many systems will be melded to provide a new generation of knowledge acquisition and transfer systems.

Financial assistance for this work has been made available by the National Sciences and Engineering Research Council of Canada. Many of the developments described in this paper have resulted from joint research with my colleague Mildred Shaw.

References

BAINBRIDGE, L. (1979). Verbal reports as evidence of the process operator's knowledge. *International Journal of Man–Machine Studies*, **11**, 411–436.

BAINBRIDGE, L. (1986). Asking questions and accessing knowledge. *Future Computing Systems*, **1**, 143–149.

BOOSE, J. H. (1985). A knowledge acquisition program for expert systems based on personal construct psychology. *International Journal of Man–Machine Studies* **20**, 21–43.

BOOSE, J. H. (1986). Rapid acquisition and combination of knowledge from multiple experts in the same domain. *Future Computing Systems* **1**, 191–216.

BRADSHAW, J. M. & BOOSE, J. H. (1986). NeoETS. *Proceedings of the North American Personal Construct Network Second Biennial Conference*, University of Calgary, Department of Computer Science, pp. 27–41.

BROADBENT, D. E., FITZGERALD, P. & BROADBENT, M. H. P. (1986). Implicit and explicit knowledge in the control of complex systems. *British Journal of Psychology*, **77**, 33–50.

VAN DE BRUG, A, BACHANT, J. & MCDERMOTT, J. (1985). Doing R1 with style. *Proceedings of the Second Conference on Artificial Intelligence Applications*, IEEE 85CH2215-2, pp. 244–249.

COLLINS, H. M. (1985). *Changing Order: Replication and Induction in Scientific Practice*: London: Sage.

DAVIS, R. & LENAT, D. B. (1982). *Knowledge-based Systems in Artificial Intelligence*. New York: McGraw-Hill.

DIXON, N. (1981). *Preconscious Processing*. Chichester: Wiley.

FREUD, S. (1914). *Psychopathology of Everyday Life*. London: Benn.

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?. In GAINES, B. R. Ed., *General Systems 1979*, Vol. 24, pp. 1–9. Kentucky: Society for General Systems Research.

GAINES, B. R. & ANDREAE, J. H. (1966). A learning machine in the context of the general control problem. *Proceedings of the 3rd Congress of the International Federation for Automatic Control*.

GAINES, B. R. & SHAW, M. L. G. (1986). Knowledge engineering techniques. *Proceedings of AUTOFACT'86*. pp. 8–79—8–96. Dearborn, Michigan: Society of Manufacturing Engineers.

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

GOODMAN, N. (1973). *Fact, Fiction and Forecast.* Indianapolis: Bobbs–Merrill.
Hall, E. T. (1959). *The Silent Language.* New York: Doubleday.
HAWKINS, D. (1983). An analysis of expert thinking. *International Journal of Man–Machine Studies,* **18,** 1–47.
HAYES-ROTH, B. (1985). A blackboard architecture for control. *Artificial Intelligence,* **26,** 251–321.
HAYES–ROTH, F., WATERMAN, D. A. & LENAT, D. B., Eds. (1983). *Building Expert Systems.* Reading, Massachusetts: Addison–Wesley.
HUME, D. (1739). *A Treatise of Human Nature.* London: John Noon.
KAHN, G., NOWLAN, S. & MCDERMOTT, J. (1985). MORE: an intelligent knowledge acquisition tool. *Proceedings of the Ninth International Joint Conference on Artificial Intelligence,* pp. 581–584.
KELLY, G. A. (1955). *The Psychology of Personal Constructs.* New York: Norton.
KLIR, G. J. (1976). Identification of generative structures in empirical data. *International Journal of General Systems,* **3,** 89–104.
KLIR, G. J. (1985). *Architecture of Systems Problem Solving.* New York: Plenum Press.
LENAT, D., PRAKASH, M. & SHEPHERD, M. (1986). CYC: using common sense knowledge to overcome brittleness and knowledge acquisition bottlenecks. *AI Magazine* **6,** 65–85.
MARCUS, S., MCDERMOTT, J. & WANG, T. (1985). Knowledge acquisition for constructive systems. *Proceedings of the Ninth International Joint Conference on Artificial Intelligence,* pp. 637–639.
MICHALSKI, R. S. & CHILAUSKY, R. L. (1980). Knowledge acquisition by encoding expert rules versus computer induction from examples—a case study involving soyabean pathology. *International Journal of Man–Machine Studies,* **12,** 63–87.
MICHALSKI, R. S. & WINSTON, P. H. (1985). Variable precision logic. *Artificial Intelligence,* **29,** 121–146.
NISBETT, R. E. & WILSON, T. D. (1977). Telling more than we can know: verbal reports on mental processes. *Psychological Review* **84,** 231–259.
NORMAN, D. (1980). Twelve issues for cognitive science. *Cognitive Science,* **4,** 1–32.
NYGAARD, K. & DAHL, O.-J. (1980). The development of the SIMULA languages.
In WEXELBLAT, R. L., Ed. (1981). *History of Programming Languages,* pp. 439–480. New York: Academic Press.
PASK, G. (1975). *Conversation, Cognition and Learning.* Amsterdam: Elsevier.
PASK, G. (1980). Developments in conversation theory—Part I. *International Journal of Man–Machine Studies,* **13,** 357–411 (November).
PEARL, J. (1985). Fusion, propagation and structuring in belief networks. *Artificial Intelligence,* **29,** 241–288.
POPE, S. (1984). Conceptual synthesis: beating at the ivory gate?. In SMITH, W. Ed., *Systems Methodologies and Isomorphies,* pp. 31–40. California: Intersystems.
POPPER, K. R. (1968). Eipstemology without a knowing subject. In VAN ROOTSELAAR, B. Ed., *Logic, Methodology and Philosophy of Science III,* pp. 333–373. Amsterdam, Holland: North-Holland Publishing Co..
POWERS, W. T. (1973). *Behavior: the Control of Perception.* New York: Aldine.
Rasmussen, J. (1983). Skills, rules and knowledge; signals, signs and symbols, and other distinctions in human performance models. *IEEE Transactions on Systems, Man & Cybernetics,* **SMC-13,** 257–266.
RAWLS, J. (1971). *A Theory of Justice.* Cambridge, Massachusetts: Harvard University Press.
ROGERS, C. R. (1967). *On Becoming a Person: a Therapist's View of Psychotherapy.* London: Constable.
ROSENBLATT, F. (1958). The Perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review,* **65,** 386–407.
SHACKLE, G. L. S. (1955). *Uncertainty in Economics.* Cambridge University Press.
SHAW, M. L. G. & GAINES, B. R. (1987). KITTEN: knowledge initiation and transfer tools for experts and novices, *International Journal of Man–Machine Studies.* In press.
SHORTLIFFE, E. H. (1976). *Computer-based Medical Consultations: MYCIN.* New York: Elsevier.

STICH, S. P. & NISBETT, R. E. (1984). Expertise, justification and the psychology of inductive reasoning. In HASKELL, T. L. Ed., *The Authority of Experts,* pp. 226–241. Bloomington, Indiana: Indiana University Press.

SWARTOUT, W. R. (1985). Knowledge needed for expert system explanation. *Proceedings of the 1985 National Computer Conference,* Vol. 54, pp. 93–98.

TUCKER, D. M. & WILLIAMSON, P. A. (1984). Asymmetric neural control systems in human self-regulation. *Psychological Review,* **91,** 185–215.

VANDERBURG, W. H. (1985). *The Growth of Minds and Cultures: a Unified Theory of the Structure of Human Experience.* Toronto: University Press.

WELBANK, M. (1983). *A Review of Knowledge Acquisition Techniques for Expert Systems.* BTRL, Ipswich: Martlesham Consultancy Services.