Modeling Practical Reasoning

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

Knowledge modeling perspectives on knowledge acquisition suggest that it is reasonable to analyze knowledge bases as collections of models. This article focuses on those parts of the knowledge base that model the practical reasoning processes of human experts, and asks what properties those models might be expected to have. It surveys the general notion of a model and its connotations in information systems science. It analyzes the structure of practical reasoning as a control process with information flows involving essential uncertainty and adaptivity, where robustness is more significant than optimality. Systemic considerations suggest that models of the relevant knowledge will consist of a collection of isolated productions, many of which will be concerned with avoidance rather than goal achievement. The relations of such models to role-limiting methods, generic tasks, and deep knowledge are discussed. Finally, the implications of the knowledge modeling perspectives for developments in knowledge acquisition are discussed. © 1993 John Wiley & Sons, Inc.

I. INTRODUCTION

This article addresses the modeling of knowledge acquisition processes whereby the knowledge engineer models the knowledge processes of the expert, which may themselves be viewed as involving modeling processes. Each level of this triple involves presuppositions, systemic principles, and scope for a plurality of definitions and analyses. A science of knowledge engineering has to develop a theory of the relations between these levels of modeling. In particular, it has to characterize those aspects of human practical reasoning that do not lend themselves to the imputation of deep scientific models, but rather reflect the pragmatic basis of achieving objectives and avoiding problems in a dynamic environment using limited information flows and uncertain data. What kind of models will the knowledge engineer form of such practical modes of knowing?

The notion that what is being done in the development of an expert system is a modeling activity has become a major theme in the literature. The KADS methodology is presented as one concerned with "developing a knowledge-level model of expert reasoning".[1] Clancey raises the question "How do expert systems differ from conventional programs?" and answers it[2] by:

*e-mail: gaines@cpsc.ucalgary.ca.

INTERNATIONAL JOURNAL OF INTELLIGENT SYSTEMS, VOL. 8, 51–70 (1993)
© 1993 John Wiley & Sons, Inc. CCC 0884-8173/93/010051-20

> . . . expert systems contain qualitative world models . . . Briefly put, qualitative models describe systems in the world in terms of causal, compositional, or subtypical relationships among objects and events. . . . Knowledge engineering is not just a new kind of programming. It is a new methodology for modeling systems so that we can assemble, modify, and control them automatically. We are not so much programmers as engineers, scientists, and even philosophers.

The *knowledge modeling* perspective on knowledge acquisition is attractive relative to previous characterizations in terms of "expertise transfer" because it involves no supposition of preexistent knowledge to be transferred. The knowledge engineer and expert together are involved in developing a model of the basis for the expert's expertise, and that model may never have previously existed.

In particular, the modeling processes involved in knowledge engineering, and the models developed, are clearly distinguished from any modeling processes and models that they impute to the expert in his or her roles as expert, as contrasted with his or her role in modeling that expertise. It is not presupposed that the expert has privileged access to existing internal knowledge, nor even that he or she has a specially privileged role in self-observation. This does not prescribe a behaviorist methodology that offers no scope for introspection. To the contrary, it legitimates introspection, but as a modeling technique rather than as direct access to an existent model. It also raises the question of the sources that the expert uses in modeling, both for the purposes of performing tasks, and for communicating the basis of doing so. For example, many of these sources may be situational rather than part of a store of internal "knowledge."

However, there is a danger in placing significant emphasis on the concept of modeling in explicating knowledge acquisition processes in that the term "model" is not in itself well defined. There is a great diversity of literature in which the term is used, often with substantial disagreement as to its application even in a particular discipline. There is a complex set of connotations associated with the term "model" and these are not well defined even in the systems literature. If the term does capture an important component of the knowledge acquisition process then it is reasonable to expect it to be complex in its connotations. However, we should also be aware that the use of the term without substantial additional explanation conveys little information.

Both the references cited above give a detailed exposition of their use of the term "model." This article complements them through a more general analysis of the term, its usage in the literature, its connotations, and its application to knowledge acquisition. In particular, it notes that the modeling involved in expertise has a major component of "control" associated with it, and hence that the "models" we acquire directly from experts will have a quite different structure from those that we might expect from scientific literature concerned with the domain of expertise. This is used to analyze the concept of generic tasks and to develop further the analysis of the knowledge acquisition processes involved in building expert systems.

II. THE NATURE OF MODELS

The term "model" has widely understood colloquial connotations, in such usage as "model toy," and is also commonly used in many artistic and scientific communities. Thus Clancey's prescription to look at a knowledge base as a model, or collection of models, is one that most people would accept as feasible, and carry out, with little hesitation—they would "know what they were doing." However, to analyze some of the issues noted above in greater depth requires somewhat more overt knowledge about what might be meant by the term "model," all the time being careful that we do not define away connotations that are important to its usage in the knowledge acquisition context.

The *Encyclopedia of Philosophy*[3] in its section on "Models and Analogy in Science" commences:

> The term "model" has become fashionable in the literature and philosophy of science, with the result that the many different senses of the term need to be distinguished before the philosophical problems connected with models in the sciences can be understood.[3]

Webster's dictionary has twenty-one definitions, of which the following is closest to the usage in the knowledge acquisition literature:

> a representation, generally in miniature, to show the construction or serve as a copy of something.

In practice, all of the other definitions and the more technical ones in philosophy, science, and engineering, can be seen as specializations of this—that a model is a representation that serves some purpose in relation to the system that is represented. Thus a mere representation is not enough, and neither is one that serves a purpose largely unrelated to what is represented; for example, a painting such as Picasso's *Guernica* does not become a model though it serves a clear political purpose.

There is also a prerequisite to the use of the term "model" which is not made explicit in the above definition, and that is that the representation should be reasonably minimal in serving its function. Extraneous components in the model should be clearly classifiable as such so that: its function of showing the construction is not undermined by superfluous constructions that have no referents in the original system, or are irrelevant to the purpose for which the model is being created; and its function of serving as a copy is not undermined by interference from material that does not copy anything in the original. This is why computer code usually does not serve as a good model of a system—there is too much extraneous material concerned with the representation on, and control of, the machine on which the code will execute. At an abstract level, an "analog" computer is one in which the bulk of the internal representation is in correspondence with an external world, such that when one examines the internal structure of a module it still corresponds to a structure in the world.[4]

Where both colloquial and technical usages of the term "model" show great variety is in the range of functions that a model may serve, and, in particular, whether the function requires the model to be operational, that is generating behavior in addition to representing structure. Differing functional requirements also vary the precision of representation necessary for something to be accepted as a model. In mathematical logic, a model for an axiom set has to conform to a precise set of formal constraints. In engineering and science, modeling involves approximation, and the notions associated with approximation are themselves subject to definition. In poetry and parables, a model can involve a remote metaphor where both the representational schema and the purpose are not subject to univocal definition, but may be indisputably present. The lack of unicity in the function of a poetic metaphor or a religious parable is also a common phenomenon in colloquial discourse, and it is paralleled by the plurality of functions that may be equipresent in a scientific, engineering, or computational model.

The study of models has been undertaken in many scientific communities: the computation of models from data in the simulation and control communities; the role of models in science in the philosophical community; the role of models in cognition in the psychology and cognitive science communities. The theoretical foundations of models and modeling processes have been developed in greatest depth in the cybernetics and general systems literature, both in terms of mathematical modeling and system identification theory, and also in terms of applications to biological and social systems. The following quotations capture the key issues relating to the purpose of models in this literature.

Weinberg, in his Introduction to *General Systems Thinking*,[5] emphasizes the role of models as tools to aid human understanding through the representation of the unknown in terms of the known:

> Every model is ultimately the expression of one thing we think we hope to understand in terms of another that we think we do understand. The chain of reasoning may be a hundred logical steps or a single analogical leap, but always ends in terms of some primitives that we agree among ourselves not to question further. For a science to have explanatory "power," this set of primitives must be neither too large nor too small.

Eykhoff, in his overview of practical modeling techniques, *System Identification*,[6] emphasizes that usability is as important as accuracy:

> The knowledge should be presented in usable form. This is an essential aspect since the model must provide the background for further decisions. If the model is too complex its usefulness becomes questionable. The relative simplicity is a dominant feature of model construction. A model is a representation of reality with reduced complexity.

Kelly states this particularly strongly in the context of computer simulation of human personality[7]:

> To be candid about it, there is considerable virtue in the fact that computers do not simulate persons perfectly. If they did, they would produce some pretty complex results. As long as the task of science is to reduce matters to terms our pedestrian minds can grasp, the instruments of science will have to select, distort, and simplify the vast complexities of nature.

Checkland's book on his widely used general modeling techniques of soft systems theory[8] has a glossary in which the term "model" is defined with emphasis on not failing to remember that the model is only representation:

> An intellectual construct, descriptive of an entity in which at least one observer has an interest. The observer may wish to relate his model and, if appropriate, its mechanisms, to observables in the world. When this is done it frequently leads—understandably, but not accurately—to descriptions of the world couched in terms of models, as if the world were identical with models of it.

Rosen points to the difficulty of modeling *natural systems* in terms of *formal systems*[9]:

> . . . the entities to be related are fundamentally different in kind. A natural system is essentially a bundle of linked qualities or observables, coded or named by the specific percepts which they generate, and by the relations which the mind creates to organize them. As such a natural system is always incompletely known; we continually learn about such a system by watching its effects on other systems with which it interacts, and attempting to include the observables rendered perceptible thereby into the scheme of linkages established previously. A formal system, on the other hand, is entirely a creation of the mind, possessing no properties beyond those that enter into its definition and their implications. We thus do not "learn" about a formal system, beyond establishing the consequences of our definitions through the application of conventional rules of inference, and sometimes by modifying or enlarging the initial definitions in particular ways.

Shariff, in his introduction to a book on system models for decision making,[10] gives a useful taxonomy of types of model as shown in Fig. 1, and an overall summary of the characteristics of models:

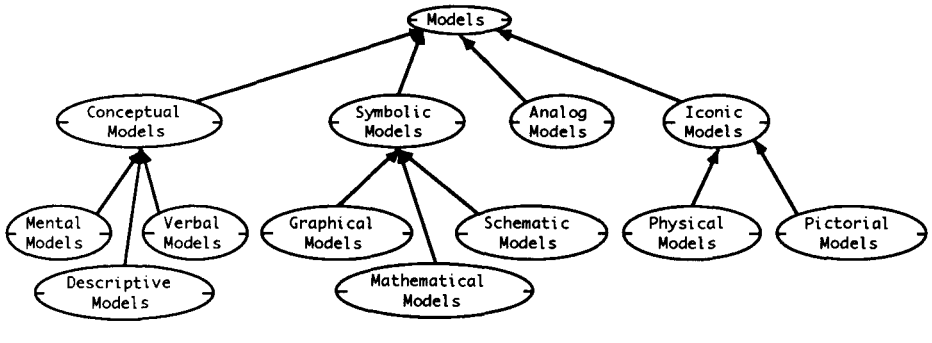

Figure 1. Sharif's taxonomy of types of models.

> —Models are attempts to imitate systems. They try to capture the major components and interactions of a system. By referring to models one can obtain valuable insights into the behavior of a system. The development of models is a sophisticated combination of art and science.
> —A model is a representation, an abstraction of an actual object or situation. They show the relations and interrelations of action and reaction of cause and effect, in operational situations.
> —A model can be conceptually regarded as a substitute for the real system. Thus, instead of investigating and experimenting with the real system, we can interrogate the model, usually with less risk, less time and less money.
> —Models are, for the most part, caricatures of reality, but if they are good, then, like good caricatures, they portray, though perhaps in a distorted manner, some of the features of the real world. The main role of model is not so much to explain and predict, as to polarize thinking and to pose sharp questions.

His taxonomy, in particular, substantiates Clancey's emphasis on the variety of structures that may be reasonably termed a model, and the value of going beyond the notion that some knowledge structure is a model to questions as to its type, purpose and adequacy.

Synopsis

1. A model is a minimal representation that serves some purpose in relation to the system that is represented.
2. The purpose of creating models of a phenomenon include prediction, explanation, and increased understanding.
3. The representations and representation schema used in models should be comprehensible.
4. Modeling involves essential simplification and approximation.
5. The model should not be reified as actually *being* the phenomenon.
6. In modeling natural systems the relationship between the system and its model is essentially uncertain and dynamic.

III. EXPERTISE AS PRACTICAL KNOWLEDGE INVOLVING CONTROL

The original "expertise transfer" notion of knowledge acquisition was based on the assumption that experts have mental models of the basis of their skill and that techniques and tools for knowledge acquisition are concerned with accessing these models. This assumption has gone out of favor in recent years as generative perspectives on human behavior that do not reify the hidden mental variables of cognitive models have become prevalent:

> The new perspective, often called *situated cognition*, claims that all processes of behaving, including speech, problem-solving, and physical skills, are generated on the spot, not by mechanical application of scripts or rules previously stored in the brain. Knowledge can be represented, but it cannot be exhaustively inventoried by statements of belief or scripts for behaving. Knowledge is a capacity to behave adaptively within an environment; it cannot be reduced to representations of behavior or the environment.[2]

However, the knowledge engineer still models the expert's skills, even if the models are not longer imputed to exist within the expert, and the problem remains of understanding what kind of *conceptual models*[11] the knowledge engineer will form of the expert's behavior. A companion article[12] analyzes the full range of conceptual models used by the knowledge engineer, their types and origins, while this article focuses on the expert's skilled performance as a knowledge source.

In examining the basis for modeling a skill it is important to take into account that what one is observing is human practical reasoning, what in education has been termed "practical modes of knowing":

> We define practical knowledge as procedural information that is useful in one's everyday life. In proposing this definition, we are clearly imposing two critical restrictions on the domain of knowledge that we are willing to call practical, namely, that the knowledge be procedural rather than declarative and that the knowledge be relevant to one's everyday life. We require practical knowledge to be procedural because of our view that practical knowledge is of and for use. We view practical knowledge as stored in the form of productions, or condition-action sequences that implement actions when certain preconditions are met.[13]

The leap in this quotation from a definition of terms to that of a mechanism obviously requires detailed justification, and the basis for this link will be further analyzed in systemic terms. However, it is also important to note that the spirit of what is being said exactly captures that of the origins of expert system development. We go to experts when we do not have overt domain and problem-solving models, precisely because they are the relevant practical reasoners in the domain, using their practical knowledge as part of their everyday professional life to achieve the practical objectives of interest in system development. This is the essential difference between expert systems development and that based on classical systems analysis, that we model the human operator in a domain when, for some reason, we cannot adequately model the domain itself.[14]

The emphasis of practicality is a classification of the mode of reasoning involved, not the domain. A skilled mathematician is a problem solver in a domain that may be highly theoretical but where his or her problem-solving skills are examples of the practical reasoning described above. The mathematician's everyday life as a mathematician is lived in the world of mathematics and he or she develops skilled behavior to navigate and manipulate that world.[15] Even though we may have highly overt models of the domain they do not automatically provide us with overt knowledge of the practical problem-solving skills in the domain. The knowledge of an expert mathematician is the capacity to "behave adaptively" in a mathematical environment.

A systemic analysis of the models that might be imputed to a practical reasoner has to take into account that these models are not being constructed through a scientific or engineering procedure based on the identification of some system through observations of its behavior, but rather as through a pragmatic procedure concerned with the achievement of objectives. Practical reasoning has as many connotations of *control* as it does of modeling, and the knowledge

that we might expect to impute to an expert will be concerned with, and situated within the context of, such control.

Synopsis

7. The expert's skilled performance is an example of human practical reasoning.
8. Practical reasoning is focused on the achievement of objectives and is situated in control activities.

IV. MODELING A CONTROLLER

There is a widely quoted early paper in the cybernetics literature that seems to justify in very general terms the notion that a domain model might reasonably be imputed to an expert. Its title is "Every good regulator of a system must be a model of that system,[16] but unfortunately that title is an overgeneralization of what is actually demonstrated. What the authors would have liked to show is that the behavior of an optimal controller can be imputed to its possessing a model of the system being controlled. The intuitive logic is that to be optimal the system doing the controlling must have some representation of the system being controlled. However, this turns out to be unjustified in general.

Under more restricted assumptions of linear, continuous models, a related result can be derived. To the extent that the operator of a linear dynamic system achieves good regulation under disturbances that cover the bandwidth of the system, that operator can be modeled with low error as another linear system.[17] This is the explanation of why linear "describing function" techniques were so effective in modeling aircraft pilots even though human motor skills are known to be nonlinear and discontinuous.

When linearity and optimality conditions are dropped, however, the reflection of the system in its controller is no longer derivable. For example, it can be shown that variants of the "combination lock" problem for finite automata can be solved effectively by a simple two-state stochastic automata, whereas the optimum solution, even for a finite-state task, requires a recursive automaton with a complex algorithm and a large number of states.[18] The simple solution would show no evidence whatsoever of any model of the task, although it is interesting to note that if one attempted to model it as if it were the optimal deterministic automation one would obtain an extremely complex, but meaningless, model.[19]

These systemic results are very insightful when applied to practical reasoning and models of expert behavior. First, to the extent that the expert is dealing with a "natural system" in Rosen's terminology above, he or she is facing one that is "always incompletely known." The expert has limited information flows from the system and is using these to achieve his or her practical objectives rather than to model the system in itself. A good control strategy under low information flow conditions is the stochastic one cited above in which behavior

is emitted at random until a desired end is achieved, for example, jiggling a key until it slips into a lock. The main defect of this strategy is that it will not avoid catastrophic terminal states if they exist. Thus the knowledge necessary to its effective use is that of conditions under which certain actions should *not* be taken.

Thus on purely systemic grounds one would expect that the practical reasoning of experts faced with natural systems, limited information flows, and irreversibility, would be found to have a major component of knowledge concerned with activities to avoid certain situations rather than with goal-directed achievement. As "models" of the domain this knowledge will generally be a collection of isolated and discontinuous material. The conditions in these productions indicate where avoidance actions are necessary. They indicate problems that have a priority over objective-oriented action, situations in which Piaget's "l'object se laisse faire"[20] breaks down—at the interface where reality bounds constructivism in that the object does *not* allow itself to be treated in the way we might chose.

This avoidance knowledge corresponds to the "aversion conditioning" of classical psychology, of taking action to avoid certain situations that may or may not arise if the action is not taken. Again on systemic grounds, avoidance knowledge once in place is very difficult to change because the evidence necessary to do so is intrinsically not available. We do not walk under the ladder to avoid a superstitious problem, and the problem does not occur so our behavior is validated. If a problem does occur then we can ascribe it to something else and our behavior is not necessarily invalidated. A substantial part of the practical knowledge of an expert might reasonably be expected to be avoidance knowledge that may no longer apply.

Clearly, one does not expect expertise to be based only on avoidance and randomness, but similar arguments apply to goal-directed behavior. The goal directed actions will also be based on limited information flows and uncertainty making elaborate planning inappropriate. What might be expected is actions generally oriented to improving the chances of achieving the goal by taking advantage of significant opportunities, that is again, not an overall coherent model but rather a set of isolated and discontinuous productions.

Thus the quote above viewing practical knowledge as "stored in the form of productions, or condition-action sequences that implement actions when certain preconditions are met" has *a priori* systemic validation apart from any empirical findings. The conceptual models that we form of expert behavior, to the extent that they accurately represent the practical reasoning may be expected to consist of a set of isolated, discontinuous condition-action productions. They will be "models" in the strict sense of the term, but they will not be scientific, principled, coherent models of either the domain or the problem-solving skill involved. The production rules that were so successful in the development of MYCIN are a natural representation of the basis of skilled behavior. That they do not in themselves provide a suitable basis for the teaching of the medical science involved from an academic viewpoint is totally unsurpris-

ing. At most they can act as isolated checks on the body of knowledge necessary for GUIDON, being consistent with it and having rationalizations within it but certainly not representing it.[21]

When they were first developed, further rationalizations of production rules were proposed other than that they are a natural representation of skilled behavior. Production rules were said to offer the advantage over normal programming that they were modular and hence more easily developed and modified. Quite the contrary is true, but this seems to have taken a long time to disseminate—papers are still appearing warning of the software engineering problems of system development based on production rules.[22] Production rules are models of practical reasoning, not in themselves a good software engineering technique.

We have also to remember that production rules are only models of the behavior observed in practical reasoning. Wittgenstein's arguments apply that we have no basis for the assertion that human behavior is generated through following rules.[23] We may find condition-action productions a useful representational mechanism for modeling practical reasoning, but even this representation is a high-level abstraction compared with the underlying neurological phenomena. For the purposes of emulating some aspects of the behavior this may not be a major issue, but when for example we attempt to emulate the "acquisition of rules" we may discover difficulties in doing so that stem from the approximation involved in this abstraction.

Synopsis

9. Models of practical reasoning will consist of a set of isolated, discontinuous condition-action productions.
10. Goals may be achieved through a succession of productions under feedback control.
11. Many productions will be concerned with the avoidance of problems rather than the achievement of objectives.
12. Avoidance knowledge is not easily disconfirmed, and some of that modeled may not be functional.
13. Condition-action productions are only models of some phenomena of practical reasoning, and cannot be assumed to generalize to all aspects of it.

V. THE ROLE OF DEEP KNOWLEDGE

One may examine the role of deep knowledge on this analysis by considering what we might expect if the presuppositions of a natural environment and limited information flows do not apply. Consider a formal deterministic environment with total state information available. In this case algorithms can be developed for problem-solving on a theorem-proving basis, and if the system is simple enough they can be applied realistically. Thus human artefacts may give scope for formal reasoning. For example, a set of regulations or a legal code is a formal artefact whose logical structure is well defined. Practical reasoning and theoretical reasoning may coincide in the analysis of the implica-

tions of the regulations or code.[24] However, it is rare to find that the interpretation of the regulations or code is as well defined, particularly if it is intended to apply to natural systems such as people and societies. Hence, we may expect practical reasoning associated with interpretation of the code to have some of the characteristics outlined above.

Expert system theory has from the early days involved considerations about the interplay of "deep" scientific knowledge and "compiled" practical knowledge.[25] There is clearly nothing wrong in developing rationalizations of practical knowledge in terms of theory, provided we do not assume that the derivation of the practical knowledge validates the theory, or provides a basis for future derivations of practical knowledge from the theory. That is, there may be many theories from which the same knowledge derives, only some of which extrapolate to the future, and hence the practical knowledge may remain constant when the current theory says it should change, or it may need to change in a different way from that predicted. We may well be wrong to value deep knowledge more than practical knowledge in many domains. The scientific and software engineering push to "clean up" expert systems may well be counterproductive.

We also have to consider the sensitivity of practical knowledge derived from theoretical knowledge, both in terms of minimal effect of discrepancies between the theory and practice, and in terms of minimal information requirements. A control strategy that is suboptimal but robust is usually very much to be preferred over one that is highly sensitive to correct presuppositions and data. In particular, a control strategy that allows recovery in the event of unexpected errors and events is very much to be preferred to one that does not—"defensive driving" through the problem environment is a major component of expertise.

The condition-action productions that characterize practical reasoning may mimic access to deep knowledge when observed in use in a particular situation.[26] That is, we may find it useful to impute such knowledge to an intelligent agent as a short form of explanation even when we know that it is not present. Situated autonomous agents can appear to be "planning" because they are goal-directed and show regular trajectories of behavior. However, the resultant "deep knowledge" need not conform to any scientific theory and still be useful both for the agent and for communication to other agents.[27] It is common in everyday human life for what are scientifically false theories to guide sound practical reasoning. My reasons why I do something may be incorrect at some level of depth but they may guide me correctly and also help others in the same situation. The golfer's view that the "follow-through" affects the flight of the ball may be physical nonsense, but the muscular preparation to achieve an excellent follow-through may be highly significant to the earlier impact of the ball with the club—a physically acceptable explanation but not more practically useful than the nonsensical one. Naive theories of motion[28] and medicine[29] raise educational problems because we have some view of what is "scientifically respectable" rather than because they adversely affect practice.

Control engineering provides a case history of the problems that may arise

in switching from practical knowledge to highly principled deep knowledge requiring system information that is unavailable. In the 1960s there was a move towards basing industrial control on the newly developed "optimal control theory."[30] However, while this approach proved very effective in the control of satellites and space vehicles, it was highly ineffective in industrial applications. The problem was that the optimum control strategy was highly sensitive to the accuracy of the model of the plant being controlled and this model had to be identified from limited information flows through sensors subject to inaccuracy and noise. Discrepancies between the model and the plant could lead to greater suboptimality than that through much simpler suboptimal control strategies. The need to continually update the model also led to the philosophically open-loop, optimal controller becoming very similar to the classical feedback controller. Theoretical results were obtained showing that the old "proportional-integral-derivative" feedback controllers were not very suboptimal and much more robust against discrepancies in design assumptions than complex mathematically optimal controllers.[31] Naturally, these controllers had always also had limit stops and alarm shut-down conditions corresponding to avoidance knowledge.

None of these considerations is negative towards scientific, principled models. To the extent that they can be realistically derived and are properly applied they are invaluable in reducing design time in system development and the need for pragmatic "fixes" through applications experience. However, it was precisely for situations where scientific, principled models are not usable that the expert systems paradigm was developed. It is not an improvement to classical system design in general, but a price we have to pay when more principled design based on deep knowledge is not possible. Thus we should not use it unnecessarily.

It is to be expected as the expert systems approach becomes better integrated with conventional system development that major components of the "knowledge base" become based on structured models rather than unstructured production rules. Clancey's (1989) paper argues that the expert is only one source for the knowledge base, not the sole source, and all practical system development is based on a combination of modeling techniques using many different sources.[12] However, we should not assume that structured, objective models are a potential replacement for all components of expert systems, and we should continue to apply effort to greater understanding of the basic expert system paradigm of modeling human practical knowledge.

Clancey's characterization of all models in expert systems as "qualitative" makes sense in terms of their origins in practical knowledge. Even when deeper knowledge is added from other sources it wil necessarily have many of the characteristics of condition-action productions in representing parts of the domain or task through discrete and discontinuous structures. The frames and taxonomies of more structured knowledge bases are still qualitative in nature. Global, algorithmic models are rarely available or appropriate in the types of tasks that have been allocated to knowledge based system designs.

Synopsis

14. Deep knowledge does not uniquely determine practical knowledge.
15. Limited information flows may make it impossible to estimate adequately the parameters of deep models.
16. Practical knowledge may be derived from, and consistent with, deep knowledge, but it can not always be replaced with deep knowledge.

VI. ROLE-LIMITING METHODS AND GENERIC TASKS

It is useful to view some of the research on generic approaches to knowledge acquisition as being ways of characterizing practical knowledge at a metalevel. McDermott's "role-limiting methods" are based on the abstraction of control knowledge from a family of related tasks, and the use of this to classify knowledge requirements and usage[32]:

> A method whose control knowledge is not very task-specific will use task-specific knowledge for identification and implementation that is highly regular, each piece of task-specific knowledge will play one of a small number of roles; thus the method will provide strong guidance with respect to what knowledge is required to perform a particular task and how that knowledge is encoded.

We need not read anything deeply ontological or psychological into this practical approach to knowledge engineering. When we generate a classification of tasks it is reasonable to expect that it will restrict the types of condition-action productions involved, and that this further classification can be linked to one of knowledge sources. The utility of this approach comes from its application to task classifications that already exist in industrial situations, and derive from the way in which we structure system design for human operators and the associated personnel selection and training.

It may be that there are psychological factors in the transfer of human experience between tasks that support industrial task design practices. It may be that there are role-limiting aspects to all human skills similar to those that may be supposed to genetically determine innate linguistic abilities.[33] These are open scientific issues of *modularity* in cognition.[34] However, we do not need such stereotypes to exist in the person to justify the arguments for role-limiting methods. It is sufficient if they exist in the practices of tasks design and allocation in a range of significant tasks that we wish to reallocate to the computer using human expertise as a model. Empirical research on the transfer of practical knowledge[35,36] is confounded by the stereotypical classifications that we have of cognitive tasks, particularly when we bring them into a laboratory setting.[37] It is reasonable that we should design the built environment to maximize the transfer of practical learning, and that this design behavior is sufficient to make classifications in terms of role-limiting methods practically useful.

As we identify role-limiting methods we may come to rationalize them and develop alternative approaches that are more principled. This is the basis of Chandrasekaran's analysis of *generic tasks*[38]:

> . . . we are only pointing out that systems perform these tasks, but not necessarily in the manner that we propose that the tasks be performed. Our claim will be that once we understand the knowledge requirements and the inference strategies for each of the tasks, we can use methods that are more natural for the tasks.

This paradigm is similar to that which has been applied to naive physics and other aspects of commonsense reasoning,[39,40] to take the problem and the activity and rationally reconstruct a model that conforms with physical and system science. It is an approach that underlies all scientific and technical development, and a legitimate strategy for the development of a science of practical reasoning.

However, the considerations of the preceding section clearly apply. Even when a rationalization provides a correct model of a situation it does not imply a unique control policy for achieving goals. There is a multitude of such policies with differing information requirements and differing sensitivities to the correctness of state parameters. The design of *robust* policies as a basis for generic task strategies is what is required to emulate human practical reasoning. However, we are more commonly used in computing to developing optimal policies under assumptions of complete and correct information flows. The generic task approach is highly attractive both theoretically and practically, but it needs new approaches to algorithm design.

Synopsis

17. Classifications affecting task design practice induce classifications of practical knowledge in terms of role-limiting methods and generic tasks.
18. In deriving condition-action productions from deep knowledge in terms of generic tasks it is more important to derive robust policies insensitive to incorrect models than optimal policies.

VII. THE ADAPTIVE COMPONENT OF KNOWLEDGE

The greatest danger in thinking that we are modeling "expertise" in an expert system is to forget that "knowledge is a capacity to behave adaptively within an environment." Schön has emphasized the need for professionals in a discipine to recognize this themselves, defining a *reflective practitioner* to be an expert who remains aware of the dynamic nature of his or her expertise. He illustrates this through case histories and through some example differentiations, such as[41]:

> Expert: I am presumed to know, and must claim to do so, regardless of my own uncertainty.
> Reflective Practitioner: I am presumed to know, but I am not the only one in the situation to have relevant and important knowledge. My uncertainties may be a source of learning for me and for them.

The notion of true expertise as reflective practice is corroborated in Hawkins's observations of mineral exploration expert systems. He characterizes the ex-

pert's advice giving procedure as a modeling cycle in which: the expert elicits data about the problem from the client; develops a minimal model that accounts for the data provided; and generates advice based on the model and feeds this back to the client; the client may accept the advice, or query it and, possibly, the model; the queries lead to further data elicitation, and repeat of the cycle. Thus, in Hawkins's model, the client plays an active role in further developing the model by providing more data until he or she is satisfied with the model and consequent advice. Expert advice giving and taking is part of a cycle of negotiation around a process of model formation.[42]

Gaines has shown how this model extends naturally to the learning processes of the expert over multiple clients, which in turn leads to a model of the formation of expertise in a community of learning systems.[43] Any inclination on the part of clients to take problems to those with a track record of effective problem solving generates a positive feedback loop which tends to focus expertise in a few individuals. He relates this and Schön's arguments to Goodman's model of induction as a process of reflective equilibrium[44]:

> 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

and, in particular, to Stich and Nisbett's extension in which they propose that the management of this equilibrium is a social process[45]:

> 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.

He recommends changes to the architecture of the expert systems to capture as much as possible of the dynamics of reflective practice[46]:

> The advice giving system should contain rules which indicate problem-solving situations that are beyond the boundaries of its capabilities and refer these to a human expert (in the same way that human experts refer difficult problems to others). It should log actual case histories of advice giving for later review by human experts so that these rules can be validated. It should make provision for feedback as to the results of advice so that learning opportunities are not lost.

However, we have to beware of overintellectualizing adaptive capabilities, as if they were a high-level design feature rather than an intrinsic feature of the low-level processes of practical reasoning. A system of productions that avoids disasters and is weakly goal directed can exhibit remarkable "adaptivity" in that it is insensitive to major changes in the domain in which it is operating. The addition of quite simple means to add new productions can further enhance this adaptivity, and the detection and communication of anomalies can allow the system to take advantage of social processes. There are a few expert systems that already exhibit such adaptivity and the capability to detect anomalies. A good example is Compton's "ripple-down rule" system, which generates a

network of rules with exceptions that is readily adapted to new cases without effect on previous performance.[47] The adaptation can be either manual or through machine induction and anomaly detection is a by-product of the structure of the system.

In general, knowledge acquisition has to become seen as an ongoing, long-term feature of a knowledge-based system's routine use, rather than as a rapid prototyping technique. The social aspects of the reflective process have to be a fundamental design consideration—experts, clients, knowledge engineers, and knowledge-based systems are partners in an enterprise, all mutually supportive of the knowledge processes of the others.

Synopsis

19. Adaptivity arises naturally out of practical reasoning that is insensitive to environmental change.
20. Adaptivity can be enhanced through the detection of anomalies and their communication in a social context.

VIII. CONCLUSIONS

The knowledge modeling perspective that views knowledge bases as collections of models is valuable and productive. This article has surveyed the general notion of a model and its connotations in information systems science. It has focused in particular on those parts of the knowledge base that model the practical reasoning processes of human experts, and asked what properties those models might be expected to have. It has analyzed the structure of practical reasoning as a control process with information flows involving essential uncertainty and adaptivity, where robustness is more significant than optimality. Systemic considerations suggest that models of the relevant knowledge will consist of a collection of isolated productions, many of which will be concerned with avoidance rather than goal achievement. The relations of such models to role limiting methods, generic tasks, and deep knowledge have been discussed.

Early approaches to the development of "expert systems" equated expertise with the possession of knowledge, assumed that this knowledge was accessible through interaction with the expert, and focused knowledge acquisition methodologies on improving the processes of "expertise transfer." This simplistic viewpoint has undergone successive deconstructions, each one at first sight apparently providing a deep analysis of the fundamental nature of expertise. This article is not intended to be yet another such analysis, although it may have given every appearance of being so. The ultimate position is that there is no ultimate position—the Hegelian spiral takes another turn—the Foucaultian excavation finds the foundation under the foundations. Derridanian deconstruction is the first truly Sisyphian enterprise of humankind. Human knowledge processes are situated—in a physical, social, and mental environment more vast than any we can ever know. They are reflexive and reflective, partially

self-generated and partially grounded in the universe in which our species survives.

However, the lack of absolutes, and of the possibility of absolutes, does not mean that the pragmatics of knowledge acquisition methodologies and tools cannot be improved. We can build better knowledge support systems by viewing experts not only as high-performance autonomous agents but also as highly interdependent components of a social and environmental infrastructure. How did the expertise come into being, how is it sustained, how is it supported, within what contexts does it manifest? An appreciation of the complex dynamics of human expertise and its reliance on subsymbolic processes also enables us to place the knowledge level in perspective—as a basis of communicative rationality, applicable and achievable when the universe is algorithmic. This rationality in some subuniverse may arise either because we created it so, or because, by the miracle that so impressed Newton and Milton, it is in some small, but significant part actually algorithmic.

We have far to go in our comprehension of the nature of human practical reasoning, and it is important that we go on developing our understanding together with our capability to operationalize it in knowledge-based systems. There is a danger that we may fail to recognize the importance of the new information processing paradigms available to us because we attempt to subsume them within our existing paradigms of deep scientific knowledge and optimal algorithms. The robust and adaptive strategies possible through the heuristic application of collections of weak methods are an important addition to our repertoire of information system methodologies, arguably *the* most important contribution that artificial intelligence research has made. We have to incorporate the basis for robustness and adaptivity into the structure and application of generic domain and task modeling—to bring knowledge acquisition as an essential component of expertise within the scope of our knowledge acquisition methodologies that attempt to model human expertise.

Overall Synopsis

1. A model is a minimal representation that serves some purpose in relation to the system that is represented.
2. The purpose of creating models of a phenomenon include prediction, explanation, and increased understanding.
3. The representations and representation schema used in models should be comprehensible.
4. Modeling involves essential simplification and approximation.
5. The model should not be reified as actually *being* the phenomenon.
6. In modeling natural systems the relationship between the system and its model is essentially uncertain and dynamic.
7. The expert's skilled performance is an example of human practical reasoning.
8. Practical reasoning is focused on the achievement of objectives and is situated in control activities.
9. Models of practical reasoning will consist of a set of isolated, discontinuous condition-action productions.
10. Goals may be achieved through a succession of productions under feedback control.

11. Many productions will be concerned with the avoidance of problems rather than the achievement of objectives.
12. Avoidance knowledge is not easily disconfirmed, and some of that modeled may not be functional.
13. Condition-action productions are only models of some phenomena of practical reasoning, and cannot be assumed to generalize to all aspects of it.
14. Deep knowledge does not uniquely determine practical knowledge.
15. Limited information flows may make it impossible to estimate adequately the parameters of deep models.
16. Practical knowledge may be derived from, and consistent with, deep knowledge, but it can not always be replaced with deep knowledge.
17. Classifications affecting task design practice induce classifications of practical knowledge in terms of role-limiting methods and generic tasks.
18. In deriving condition-action productions from deep knowledge in terms of generic tasks it is more important to derive robust policies insensitive to incorrect models than optimal policies.
19. Adaptivity arises naturally out of practical reasoning that is insensitive to environmental change.
20. Adaptivity can be enhanced through the detection of anomalies and their communication in a social context.

This work was funded in part by the Natural Sciences and Engineering Research Council of Canada. I am grateful to Bill Clancey for access to his own research in this area, to Mildred Shaw for critical comments on earlier drafts of this article, and to colleagues at knowledge acquisition workshops worldwide for discussions which have raised many of the issues in this article.

References

1. H. Akkermans, F. van Harmelen, G. Shreiber, and B. Wielinga, "A formalisation of knowledge-level models for knowledge acquisition," *Int. J. Intell. Syst.* **8** (1993). Also in *Knowledge Acquisition as Modeling*, K.M. Ford and J.M. Bradshaw, Eds., Wiley, NY, 1993.
2. W.J. Clancey, "Viewing knowledge bases as qualitative models," *IEEE Expert* **4,** 9–23 (1989).
3. M. Hesse, "Models and analogy in science," in *The Encyclopedia of Philosophy*, P. Edwards, Ed., Collier Macmillan, New York, 1967, pp. 354–359.
4. B.R. Gaines. "Varieties of computer—their applications and inter-relationships, in *Proceedings of IFAC Symposium on Pulse Rate and Pulse Number Signals in Automatic Control, Budapest, 1968,* Butterworths, London, 1968, pp. 1–16.
5. G.M. Weinberg, *An Introduction to General Systems Thinking,* Wiley, New York, 1975.
6. P. Eykhoff, *System Identification*, Wiley, London, 1974.
7. G.A. Kelly, "Aldous: the personable computer," in *Computer Simulation of Personality*, S.S. Tomkins and S. Messick, Eds., Wiley, New York, 1963, pp. 221–229.
8. P. Checkland, *Systems Thinking, Systems Practice*, Wiley, Chichester, 1981.
9. R. Rosen, *Anticipatory Systems*, Pergamon, Oxford, 1985.
10. N. Sharif, "An introduction to modeling," in *Systems Modeling Making*, N. Sharif and P. Adulbhan, Eds., Asian Institute of Technology, Bangkok, 1978 (distributed by Pergamon, Oxford).
11. D.A. Norman, "Some observations on mental models," in *Mental Models*, D. Gentner and A. Stevens, Eds., Erlbaum, Hillsdale, NJ, 1983.
12. B.R. Gaines, M.L.G. Shaw, and J.B. Woodward, "Modeling as a framework for knowledge acquisition methodologies and tools," *Int. J. Intell. Syst.* **8** (1993). Also

in *Knowledge Acquisition as Modeling*, K.M. Ford and J.M. Bradshaw, Eds., Wiley, NY, 1993.
13. R.J. Sternberg and D.R. Caruso, "Practical modes of knowing," in *Learning and Teaching the Ways of Knowing*, E. Eisner, Ed., University of Chicago Press, Chicago, 1985, pp. 133–158.
14. B.R. Gaines and M.L.G. Shaw, "From fuzzy sets to expert systems," *Fuzzy Sets Syst.* **36,** 5–16 (1985).
15. K.H. Wolff, *Surrender and Catch: Experience and Enquiry Today*, Reidel, Holland, 1976.
16. R.C. Conant and W.R. Ashby, "Every good regulator of a system must be a model of that system," *Int. J. Sys. Sci.* **1,** 89–97 (1970).
17. B.R. Gaines, "Linear and nonlinear models of the human controller," *Int. J. Man-Mach. Stud.* **1,** 333–360 (1969).
18. B.R. Gaines, "Memory minimization in control with stochastic automata," *Electron. Lett.* **7,** 710–711 (1971).
19. B.R. Gaines, "On the complexity of causal models," *IEEE Trans. Syst. Man Cybern.* **SMC-6,** 56–59 (1976).
20. J. Piaget, *Le Structuralisme*, Presses Universitaires de France, Paris, 1970.
21. W.J. Clancey, *Knowledge-Based Tutoring: The GUIDON Program.* MIT Press, Cambridge, 1987.
22. X. Li, "What's so bad about rule-based programming?" *IEEE Softw.* **8,** 103–105 (1991).
23. L. Wittgenstein, *Philosophical Investigations*, Blackwell, Oxford, 1953.
24. M. Sergot, "Representing legislation as logic programs," in *Machine Intelligence 11*, J.E. Hayes, D. Michie, and J. Richards, Eds., Clarendon, Oxford, 1988, pp. 209–260.
25. B. Chandrasekaran and S. Mittal, "Deep versus compiled knowledge approaches to diagnostic problem solving," *Int. J. Man-Mach. Stud.* **19,** 425–436 (1983).
26. P. Maes, Ed., *Designing Autonomous Agents*, MIT Press, Cambridge, 1990.
27. B. Rogoff and J. Lave, Eds., *Everyday Cognition*, Harvard University Press, Cambridge, 1984.
28. M. McCloskey, "Naive theories of motion," in *Mental Models*, D. Gentner and A. Stevens, Eds., Erlbaum, Hillsdale, NJ, 1983, pp. 299–324.
29. P.J. Feltovich, R.J. Spiro, and R.L. Coulson, "The nature of conceptual understanding in biomedicine: The deep structures of complex ideas and the development of misconceptions," in *Cognitive Science in Medicine*, D.A. Evans and V.L. Patel, Eds., MIT Press, Cambridge, 1989, pp. 113–172.
30. P. de la Barrière, *Optimal Control Theory*, Saunders, Philadelphia, 1967.
31. A.T. Fuller, "Linear control of non-linear systems," *Int. J. Control* **5,** 197–243 (1967).
32. J. McDermott, "Preliminary steps toward a taxonomy of problem solving methods," in *Automating Knowledge Acquisition for Expert Systems*, S. Marcus, Ed., Kluwer, Boston, 1988, pp. 225–256.
33. N. Chomsky, "Rules and representations," *Behav. Brain Sci.* **3,** 1–15 (1980).
34. J.A. Fodor, *The Modularity of Mind*, MIT Press, Cambridge, 1983.
35. M.K. Singley and J.R. Anderson, *The Transfer of Cognitive Skill*, Harvard University Press, Cambridge, 1989.
36. S.M. Cormier and J.D. Hagman, Eds., *Transfer of Learning: Contemporary Research and Applications*, Academic, San Diego, 1987.
37. J. Lave, *Cognition in Practice*, Cambridge University Press, Cambridge, UK, 1988.
38. B. Chandrasekaran, "Generic tasks as building blocks for knowledge-based systems: The diagnosis and routine design examples," *Knowl. Eng. Rev.* **3,** 183–211 (1988).
39. D.G. Bobrow, Ed., *Qualitative Reasoning about Physical Systems*, MIT Press, Cambridge, 1985.

40. J.R. Hobbs and R.C. Moore, Eds., *Formal Theories of the Commonsense World*, Ablex, Norwood, NJ, 1985.
41. D.A. Schön, *The Reflective Practitioner*, Basic Books, New York, 1983.
42. D. Hawkins, "An analysis of expert thinking," *Int. J. Man-Mach. Stud.* **18,** 1–47 (1983).
43. B.R. Gaines, "Social and cognitive processes in knowledge acquisition," *Knowl. Acquis.* **1,** 251–280 (1989).
44. N. Goodman, *Fact, Fiction and Forecast*, Bobbs-Merrill, Indianapolis, 1973.
45. S.P. Stich and R.E. Nisbett, "Expertise, justification and the psychology of inductive reasoning," in *The Authority of Experts*, T.L. Haskell, Ed., Indiana University Press, Bloomington, 1984, pp. 226–241.
46. B.R. Gaines, "Positive feedback processes underlying the formation of expertise," *IEEE Trans. Syst. Man Cybern.* **SMC-18,** 1016–1020 (1988).
47. P. Compton and R. Jansen, "A philosophical basis for knowledge acquisition," *Knowl. Acquis.* **2,** 241–258 (1990).