The Knowledge Engineering Review, No. 7:1, 1992, 55–77

The qualitative representation of physical systems

ENRICO COIERA

Hewlett-Packard Laboratories, Filton Rd., Stoke Gifford, Bristol BS12 6QZ, UK

Abstract

The representation of physical systems using qualitative formalisms is examined in this review, with an emphasis on recent developments in the area. The push to develop reasoning systems incorporating deep knowledge originally focused on naive physical representations, but has now shifted to more formal ones based on qualitative mathematics. The qualitative differential constraint formalism used in systems like QSIM is examined, and current efforts to link this to competing representations like Qualitative Process Theory are noted. Inference and representation are intertwined, and the decision to represent notions like causality explicitly, or infer it from other properties, has shifted as the field has developed. The evolution of causal and functional representations is thus examined. Finally, a growing body of work that allows reasoning systems to utilize multiple representations of a system is identified. Dimensions along which multiple model hierarchies could be constructed are examined, including mode of behaviour, granularity, ontology, and representational depth.

1 Introduction

The need for an artificial intelligence system to explicitly represent knowledge of the physical world is accepted almost as an absolute. However, in recent times this basic assumption has been strongly challenged (Brooks, 1991; Morris, 1991). Indeed, it may be that a large proportion of the cognitive functions performed by humans can be driven by concept-free control mechanisms. Equally, however, there is still strong evidence for a symbol rich component to human cognition—for example, the use of language as a communication mechanism (Kirsch, 1991). It is this aspect of intelligence that we believe can be modelled as the representation and manipulation of symbols, that seems to account for many of the higher functions associated with human intelligence. Thus, without ignoring the importance of a non-representational component to intelligence, it seems clear that explicit representations of the physical world will be needed to produce the range of inferences expected from AI systems. Decision support systems, for example, will need to represent knowledge about physical systems ranging from digital circuits, satellites, and manufacturing plants, to the physiology of the human body. They may also need to represent physical systems that are a part of everyday experience, like the fluid contained within a cup or a ball tossed into the air.

Knowledge of physical systems will often be incomplete, consisting of qualitative descriptions of their composition and behaviours. Sometimes the incompleteness arises because the laws governing the operation of a system are not well understood (although in principle such knowledge might exist). Nevertheless, humans are able to make inferences using just such limited knowledge. For example, consider a ball thrown by a child. The child may not know equations for the motion of a body in space, nor even precise data such as the mass of the ball, how hard it was thrown, or the angle it was thrown at. Nonetheless, by using the information at hand as well as past experience, the child can infer that the ball will rise, that its trajectory will take it in a certain direction, and perhaps that its path will be longer than it is high. The need to base decisions on incomplete knowledge and data are also features of many technical domains such as medicine. In other domains, more detailed knowledge is available, perhaps in the form of equations describing the characteristics of a manufacturing plant or electronic circuit. It is often easier, however, to represent such systems at a coarser qualitative level, because this offers a sufficiently descriptive and computationally cheap level of abstraction for some problem solving activities.

It is the need to design AI systems that can perform well on such under specified problems that has driven the various strands of qualitative reasoning research in AI. Over the last decade or so a considerable body of work has accumulated in this area, and several excellent collections of papers exist (Bobrow, 1984; Hobbs & Moore, 1985, Weld & de Kleer, 1990). There are also several good reviews which introduce the basic representational and inferential concepts used in qualitative reasoning (e.g., Forbus, 1988; Cohn, 1989). This background work will not be covered in detail here. Rather, this paper is particularly concerned with the evolving knowledge representation work in qualitative reasoning, and the commonalities emerging between different representational schools. To assist in the task, the changing views of causality and functionality within physical systems will be examined, followed by a review of recent work exploring the use of multiple models within a single reasoning system.

2 Deep knowledge

A central issue in the design of intelligent agents is the manner in which one can capture the expertise they will need to interact with the real world. In particular, there has been much discussion about the proportion of expertise that needs to be explicitly represented as knowledge, and that which must be modelled as inference procedures. The "Knowledge principle" of Feigenbaum (Lenat & Feigenbaum, 1988) states that the key to intelligent performance of a computer system lies in it containing large amounts of knowledge, rather than sophisticated general reasoning procedures. This has been a cornerstone in the construction of first-generation rule based expert systems. However, these systems only operate successfully when they are restricted to tightly defined domains. Rule based systems are brittle, their competence rapidly decaying at the edge of their expertise (Davis, 1987). In contrast, people exhibit a much more graceful decay from expertise, being cushioned by weaker but more general problem solving capabilities. This "robustness" (Davis, 1987) is a desirable characteristic for expert systems that have to diagnose previously unencountered faults, or failures due to novel fault interactions. Figure 1, taken from Lenat & Feigenbaum (1988), illustrates the hypothesized decay of reasoning performance at the limits of a system's expertise.

In an attempt to deal with these limitations, Lenat's CYC project is trying to amass most of the common sense knowledge people use in their daily lives. The intuition is that expert system brittleness is partly due to a lack of general knowledge about the world. Yet Lenat and Feigenbaum admit that collecting millions of rules is not enough. Intelligent systems need general problem solving methods as well as a breadth of knowledge. They now propose the *breadth hypothesis* to qualify the knowledge principle—intelligent performance often requires the problem solver to fall back on increasingly general knowledge, or to analogize specific knowledge from far flung domains. In other words, if a reasoning system is to deal with problems at the edge of its expertise then it must be equipped with general models and general reasoning procedures to augment its highly specialized knowledge.

CYC has thus seemingly returned to the notion that knowledge of the world can be represented at various depths (Hart, 1982), and that problem solving robustness depends on an ability to represent and reason about a problem at such varying levels of depth. Rule based systems, associating an input pattern with an action, are regarded as shallow representations. Deep systems are those that have some underlying representation of the basic physical laws or mechanisms operating in a domain, and can utilize this knowledge when shallower representations fail. The implication is that the sort of knowledge contained in deep systems allows shallower knowledge to

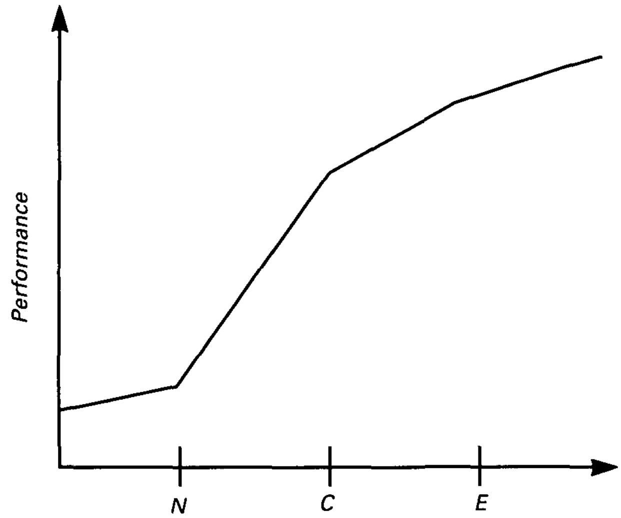

Figure 1 Performance for a system decays at the limits of its expertise. Thresholds are (N)ovice, (C)ompetent, and (E)xpert. The decay between N and C is rapid for shallow systems, and more gradual for humans (after Lenat & Feigenbaum, 1988).

be inferred. Another way of expressing the same relation is to note that one system is deeper relative to another if knowledge implicit in the deep system must be explicitly represented in the shallow (Klein & Finin, 1987). A corollary of this is that deep representations require more complex inference procedures than their shallow counterparts—an example of the trade-off between computational tractability and expressiveness of representational languages (Levesque & Brachman, 1984).

3 Approaches to qualitative representation of the physical world

What truly constitutes a deep model is still an open question (Chandrasekaran, 1983), and there are many ways in which one could represent knowledge of the real world. It is useful to partition knowledge into that which describes a system proper (its structural and functional aspects), and the behaviours exhibited by a system. Traditionally, differential equations have done just that, with the equations capturing the laws that govern a system, and the solutions to those equations representing the possible system behaviours (Kuipers, 1986). Differential equations are thus often seen as deep laws, and qualitative representations as abstractions of such laws at some shallower but still sufficiently deep level (Figure 2). The properties of a system expressed at one level may have limited meaning in another.

Historically, the qualitative reasoning community has been interested in two quite different representational problems. The first is often called *naive physics* (Hayes, 1979) or commonsense reasoning, and is concerned with understanding how humans make inferences about everyday physical systems. Hayes proposed the construction of a formalism to describe a large portion of our common sense knowledge about the everyday physical world. The formalism would handle, for example, the notion of objects, shape, movement, substances, and time. A classic example is the knowledge used by a human to infer the behaviour of fluids without knowing anything about the fluid's material properties or fluid dynamics.

The second broad concern in qualitative reasoning has been termed *qualitative physics.* This work is interested in replicating the sorts of first principles reasoning that an expert might use when reasoning about a difficult problem, e.g., an engineer reasoning about a complex circuit's behaviour for the first time. This form of knowledge centres around notions of the functional and structural relationships that exist within a physical system. Work in this strand is concerned with developing techniques that may complement traditional quantitative analysis tools.

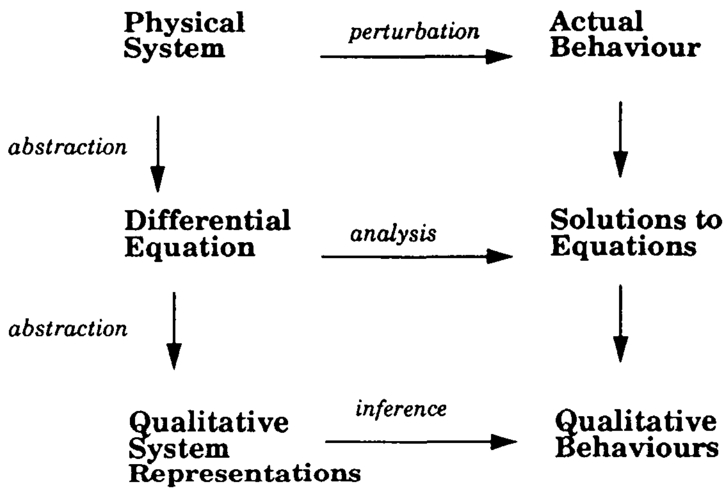

Figure 2 Qualitative representations and differential equations are abstractions of physical systems (after Kuipers, 1986).

The two enterprises of naive and qualitative physics have many common concerns. Both, for example, may use incomplete qualitative descriptions of the world. However, naive physics is everyday physics. It is concerned with the sorts of reasoning and representations that people need to make daily. Qualitative physics, on the other hand, deals with complex device knowledge only held by a few individuals. One would thus expect these two domains to differ in the

- depth of knowledge represented
- amount of knowledge needed
- type of concepts represented
- type of inferences performed

The vast bulk of research in qualitative reasoning has focused on qualitative physics at the expense of naive physics, and these differences have not been explored in any great detail. Early work in the field (e.g., Hobbs & Moore, 1985) did begin with the premise that there was value in capturing aspects of everyday human commonsense reasoning. Subsequent work has often focused on the formal mathematical aspects of qualitative physics, especially for the analysis of models in expert technical domains like electronics.

Fishwick (1989, 1990) criticizes the formal work in qualitative physics, noting that it has until now ignored significant work from other disciplines. For example, the field of econometrics has for years pursued qualitative techniques for the analysis of economic models. There is more than a little truth in this observation. For example, the use of phase portraits as a new global constraint (Lee, 1988; Struss, 1988) are significant contributions to the AI literature on qualitative simulation. Yet the principles behind these techniques can be found in student textbooks of mathematical economics (e.g., Chiang, 1984; Gandolfo, 1981). In fact, the study of the qualitative behaviour of systems, especially complicated systems—non-linear or of degree greater than two—has been proceeding for over a hundred years. Poincare began the modern program for the study of the qualitative properties of ordinary differential equations in the 1880s. Work has proceeded with developments in differential topology and qualitative dynamics. In recent times, catastrophe theory was developed to describe the equilibria of dynamical systems, and chaos theory to describe systems exhibiting complex behaviours. Such work is clearly beyond the scope of the present review.

The move to more mathematically formal representations has, however, had a beneficial effect on qualitative reasoning research. There is clearly an interaction between the representation of qualitative knowledge, and the type of inference that can be performed with it. As we shall see, attributes of the perceived world like causality or functionality can either be explicitly represented, or be inferred. The shifting understanding of this balance offer important insights into the development of this field in AI, and is at least partly a result of the shift to more mathematically formal representational and inferential methods.

4 Causality

A concern at the heart of qualitative reasoning is the notion of causality. Indeed, one of the motivations of early work in qualitative physics was to develop an explicit treatment of causality that was unavailable in traditional physics (de Kleer & Brown, 1984). However, as Bunge (1979) points out, the word "causality" has numerous meanings. At its simplest, causation can denote a connection between two events—event A generated a change in the state of B to produce state C. Causality in the form of the *causal principle* allows the generalization of such connections into a law through the process of induction—C always follows A. At its broadest, causality represents a philosophical position—that everything in the universe happens according to a causal law. This last position at least is on shaky ground with quantum physics—causality here seems to be an artefact of human perception.

Nevertheless, causality is useful, and causal thinking characterizes an important part of everyday and first-principles reasoning. Causality is often associated with explanation and the need to produce a meaningful causal explanation to justify a recommendation is often cited as an important component of an intelligent decision support system (Patil, 1981). The imprecision with which causality has been defined, and the lack of clear agreement about how it might be employed within a reasoning system, has generated much debate within the AI community (Weld & de Kleer, 1990).

4.1 Causal networks

There is a clear early division between systems that attempt to represent causality explicitly, and those that treat it as a property that should be inferred. Much of the later work from the qualitative reasoning community has tended to support the notion that causality is a property that can be inferred from structural and functional knowledge. No explicit causal representation is found, for example, in Kuipers' QSIM (1986). This contrasts with early efforts in model-based reasoning, which attempted to augment rule-based representations with an explicit representation of causality through a probabilistic network. The CASNET system for glaucoma diagnosis (Weiss et al., 1978), and Long's (1983) work in cardiovascular haemodynamics typify systems that represented knowledge as a causal-associational network. Links in CASNET's causal model represented an assertion about the likelihood of one event leading to another, based upon probabilistic estimates. The events linked by causal connections were complex states such as "neural tissue loss" or "cupping of the nerve head".

Although one of the motivations for developing causal networks is to enhance their explanatory power, the mechanisms which underlie the transition from one state to another are still hidden in probabilistic assertions. For example, take a CASNET link between two states involved in the progression of the eye disease glaucoma

> Elevated Pressure Transmitted to Optic Nerve Head → (0.9)
> Decreased Blood Supply to Optic Nerve Head

The process being represented is the compression of blood vessels leading to a reduction in blood flow. The representation tells us nothing about notions such as the effect of pressure on the structure of a hollow tube, or the effect narrowing a tube's cross sectional area produces on the flow of a liquid within it. Yet it is these concepts, even at a very naive level, that a human would grasp when learning about the disease process, or trying to explain it to another. Thus, while causal networks of this type may appear to capture process information, they are ultimately still shallow representations, lacking the means of expressing structural or functional relationships. They do not express *why* one event caused another.

Causality is fundamentally an explanatory notion—it is invoked to describe one's understanding of the chain of events that lead to the existence of a state in a physical system. Causal links are thus acquired as explanations or justifications from experts. Compton (1990) suggests that the context in which such knowledge is acquired is fundamental to understanding its value. A causal link or chain of links may well explain a phenomenon, but only in the context of a particular problem. For example, an expert may offer a causal explanation of a particular disease process to justify the values of a set of blood test results. A modified chain of events may be invoked later to explain why another set of results is slightly different. Compton maintains that the second chain is only a valid representation of knowledge in the context of the failure of the first chain to explain events. Further, Compton reports that experts will happily offer alternate explanations for similar phenomena.

This suggests that a model constructed by aggregating a large number of causal links obtained in different contexts may not have logical internal consistency. Such a model does not represent an expert's true knowledge of a physical system, but an assembly of context dependent explanations of events derived from some deeper knowledge. For a causal model to be constructed, Compton would demand that the links be assembled in the context in which they were acquired to have any validity.

Thus causal links are of their own a shallow representation for a number of reasons. This is an important insight—that the act of simply representing causal connections between events in no way guarantees that we have captured any deep knowledge of a system or its behaviours (Keravnou & Washbrook, 1989; Bylander, 1990).

4.2 Differential equations

AI researchers see differential equations as one important form of deep knowledge, and some have attempted to give causal readings to such equations. For example, using the equation V = IR, one could interpret that an increase in voltage will result in an increase in current when resistance is constant. Forbus & Gentner (1990) further attempt to distinguish between symmetric and asymmetric functional relations, where the difference in symmetry rests solely in the causal interpretation of the equations. Some equations are essentially asymmetric and explicitly represent causal processes. In our example above, one reading of causality from voltage to current might be favoured over others through some external process of identifying voltage as the "independent" variable.

In the general case, however, if an inverse exists for a functional relationship this asymmetry cannot hold. Thus one key aspect of a causal link—that it is essentially asymmetrical—is violated with differential equations. Mathematical forms of their own do not capture causal relations, just the form that the relations will take (Bunge, 1979). When a causal interpretation of a mathematical form is reasonable, it has to be added. Such an interpretation does not belong to the mathematical symbols themselves, but to the semantic system applied to the relations. Thus functions are seen as syntactic forms only, and they cannot of their own replace causal propositions. Causality is thus seen as an explanatory inference performed over a set of equations. The equations themselves represent structural or functional relations in a physical system providing a syntactic form between parameters. Exogenous knowledge of the context within which we place the set of equations is needed to form a causal explanation.

4.2.1 Causal ordering

One way of extracting a causal reading from a set of equations is to identify dependencies between variables in the equation system. To this end, Iwasaki & Simon (1986) advocate the use of causal ordering and the method of comparative statics long used in economics and thermodynamics.

These provide an operational means to both determine a causal relation among variables in a set of equations, and to assess the qualitative effects of a disturbance applied to the system.

A model is formed by a set of simultaneous differential equations. The equations come from an understanding of mechanisms, and each equation in the model must be a *structural equation* standing for a mechanism through which variables influence other variables. Unfortunately there is no easy way to know if an equation is structural (Iwasaki, 1987). If the model involves feedback, an assessment of the system's stability needs to be made before its response over time can be determined.

The causal ordering is an asymmetric relationship among the variables and equations. The ordering is derived in the following way. A system is *self-contained* if it has n equations and n unknowns. The system can have self-contained subsets, and a subset that does not contain a proper self-contained subset is a *minimal complete subset.* The set of minimal subsets of zero order are solved, and these values substituted into the remaining system equations to obtain a new structure. Then the new set of minimally complete subsets of first order is solved. The procedure is repeated until the derived structure of highest order contains no subset that is self-contained. If V_i is the set of variables in the complete subsets of ith order, they are said to be *causally dependent* on the elements in V_{i-1}. Importantly, a causal order cannot be assigned in feedback loops.

Causal ordering also relies on the notion of an *exogenous variable* to assist in making causal assignments, and in particular when there is a need to establish the structural equations for a system. Such variables are considered external to the subsystem they occur in, and their values are set by external processes. A variable that can be manipulated experimentally, for example, would be considered exogenous. This exogenous ordering is similar to Forbus & Gentner's (1990) notion of favoured causal interpretations of equations. Iwasaki and Simon offer a list of methods by which one can determine the exogenous variables in a system, all relying on information external to the system equations. The causal ordering produced by this set of techniques is sometimes counter-intuitive. This usually results when the technique is applied to equations which do not represent an equilibrium system. Iwasaki (1988) extends causal ordering to systems with static and dynamic components to rectify this limitation.

Forbus (1988) argues that causal ordering is limited by its dependence on quantitative equations and knowledge of which variables are exogenous. In particular, he notes that the notion of causal independence based on exogenous parameters limits the system to specific modes of behaviour. What is exogenous in a system may well change when it becomes part of a larger system. This criticism does not allow for the fact that the same set of equations can have different sets of exogenous variables identified for different behaviour modes. Indeed, it should be directed more towards Forbus own work (see section 6.2), which places favoured causal readings on equations, thus fixing the context in which they can be applied.

4.2.2 Bond graphs

Bond graphs are a formal representational language for analysing physical systems developed for system dynamics. There has been increasing interest in the use of bond graphs for qualitative modelling in AI because of their simplicity and representational power (Fishwick, 1989b; Top & Akkermans, 1990; Söderman & Strömberg, 1991).

The modelling process is based upon the conservation of energy, with physical processes being linked in a labelled digraph through energy flows. Basic concepts like effort, flow, inertia, and capacitance are used in modelling, and their generality allow the method to be applied to domains like thermodynamics, rotational and translational mechanics, fluid dynamics and electronics. Once a bond graph has been constructed for a physical system, traditionally one can generate a block diagram representation of the full differential equations that describe the system and then proceed with formal analysis.

The bond graph can however also support a useful set of qualitative inferences, including an analysis of the causal ordering that occurs within a system. In fact, bond graphs will produce identical causal orderings to those obtained using the methods of Iwasaki and Simon described in section 4.2.1, when the underlying mathematical models are equal (Top & Akkermans, 1991). Top and Akkermans go further in their analysis, suggesting that the causal orderings produced with the bond graph method are richer than those produced by the method of Iwasaki and Simon because bond graphs can assign causality in a feedback loop, and have automatic methods for determining whether a system is self-contained and ensuring that a model is maximally resolved.

4.2.3 Causality as constraint propagation

Systems of equations can be seen as a set of constraints, and the way in which values are propagated through these constraints can be used to assign causal interpretations. Many workers have abstracted differential equations into qualitative constraints (see section 6), starting with de Kleer & Brown's (1984) seminal work. Here a system was described by a set of *confluences*, which were constraint equations written in terms of the qualitative derivatives of variables, with possible values of +, −, and 0.

Assuming a system is in an equilibrium state, then a disturbance is superimposed upon the set of qualitative equations or confluences. The progression this produces from one system state to another describes the system's time varying behaviour, and can be given a causal reading. de Kleer and Brown were also interested in the sequence of non-equilibrium states that were formed by the process of constraint propagation, *between* two equilibrium states. This second sequence happened instantaneously or in their terms, in "mythical time". They called the causal interpretation of the sequence a "mythical causality". Mythical causality thus summarized "physical action taking place at a lower level", no conventional time passing between the mythical time instants.

As values were propagated through their constraint network, ambiguities would be encountered. Three heuristics—the component, conduit and confluence heuristics—were used to help resolve such situations. The three are essentially similar, capturing the notion that places in the network untouched by the propagation so far remain unchanged. During the process of causal ordering, Iwasaki and Simon were similarly forced to look to exogenous variables to assist in resolving ambiguities in their causal assignment. In fact, a now famous debate ensued between de Kleer and Brown defending their seemingly *ad hoc* notions of mythicallity and locality of propagation, and Iwasaki and Simon proposing their arguably more formal mechanisms for assigning causality (Iwasaki, 1986; Iwasaki, 1986b; de Kleer, 1986). While causal ordering was a structural analysis of a system, de Kleer and Brown's methods required that their confluences were assigned values. In other words, the equations needed to be solved for a given set of initial conditions. In both cases, knowledge external to the equations is needed to help resolve causal arguments.

When systems with no feedback are considered, the techniques of causal ordering and constraint propagation (specifically the methods of de Kleer and Brown) produce the same ordering (Iwasaki, 1986b). Causal ordering, however, does not assign causality in feedback systems, and requires further examination of extra second order equations using the methods of comparative statics to analyse such systems. The viewpoint from causal ordering is that there is no sense in determining a causal order in feedback loops—a view shared in the earlier work of Bunge (1979). Some have come to the conclusion that, at least for the question of feedback (Weld & de Kleer, 1990), the need to produce an ordering is purely a domain issue. Economists like Simon do not need to do it, but engineers working in circuit diagnosis do. This might be an easy way to resolve the debate, but perhaps will also avoid resolving deeper questions about the many different meanings and uses of causality. While it would be unfair to declare the debate closed, it is clear that thrust towards formalization provided by causal ordering has had a positive effect on the area.

4.2.4 Causality as temporal abstraction

A key problem in the assignment of causality to constraint or equation representations of physical systems is that the need for a asymmetric reading of the functional relations is at variance with their mathematical meaning. The solutions proposed so far are computational rather than representational—they have relied on the use of external knowledge to explicitly assign causal direction when structural dependence cannot be determined, or have relied on an inferred propagation order of values.

An alternative viewpoint sees causality as an emergent property that can be attributed to the varying temporal granularity of system models. One can model a system at varying levels of abstraction, hiding details when they are not appropriate (Patil, 1981). It is clear that time can be used in this manner, abstracting away details of a system's behaviour if they are too fast or slow for current purposes. From a modelling point of view, systems that operate considerably faster than the current scale can be seen as instantaneous, and those that operate significantly more slowly as constant (Kuipers, 1987).

Kuipers (1987) shows that what appears as a structural constraint (or equation) at one level of temporal abstraction, can be seen as an embedded process when seen from a faster level. This view solves the problem of assigning an asymmetric interpretation to functions in some sense, by seeing the problem in modelling terms. A relationship which is temporally asymmetrical at a detailed level is abstracted into a symmetrical constraint by shrinking its temporal extent to zero (Kuipers, 1987b). For example, if a set of equations generate a sequence of states over time and show a change in one function *eventually* resulting in a change in another function, that relationship can be mapped onto a single functional relation between the two variables if we ignore time.

The implication of this is that causality is very much a property of the level of abstraction of our model of the world, and that the causal interpretations drawn must change, depending on the level of model abstraction. Iwasaki (1990b) shows that this is certainly true for causal ordering. Importantly, when one performs a causal ordering, the model used should be of a uniform temporal grain size. Mixing temporal abstractions in the one model could lead to erroneous causal interpretations. Equally, each time a reasoning system changes models to a more appropriate abstraction level, a causal analysis would need to be redone.

5 Device structure and function

When people reason about objects in the physical world, they are often concerned with explaining how things work—their function. Explanations are often based on some understanding of the manner in which objects are put together—their structure. One may thus view objects in the physical world as devices, and attempt to derive a behavioural description of the device's function from a description of its internal structure. A device is usually made up of components, each governed by a set of laws, and the behaviour of the overall device is obtained from those of its parts. Parts in a device model might communicate by conduits or ports, passing values to one another. The contributions of de Kleer & Brown (1984) to the device ontology will be examined in the following section. Other workers like Davis (1984), Pan (1984) and Hamscher (1987) have applied model-based representations of mechanisms to fault localization problems, principally in the area of digital circuit analysis.

5.1 Reasoning with confluences

The early work of de Kleer & Brown (1984) made important contributions both to the development of qualitative simulation techniques, and to the development of qualitative representations. The behaviour of a device component in their work was governed by a set of qualitative differential equations (confluences), and by their connections to other components. The simulation of the device model produced an *envisionment*, a directed graph of all the states the device could reach. Envisionment is probably the first qualitative simulation technique in the AI literature, and was based on the propagation of perturbations around the constraint network created by the confluences.

A goal of their work was to develop representations of device structure that allowed function to be inferred. To assist in this endeavour, de Kleer and Brown proposed the "No Function in Structure" (NFIS) principle, which prohibited model builders from incorporating behavioural information into a device description. In other words, when constructing a model, one should not presuppose the function it is intended for, but simply describe its composition. An advantage of such a method is that it enforces representational modularity—libraries of device component models could be constructed, and then assembled as needed to describe new composite devices.

It has become clear that NFIS cannot be more than a modelling guideline, however, since the context in which a device is used may be needed to infer its purpose. Its structural description may be inadequate to make functional inferences. Keunke & Allemang (1989) offer a simple example of the use of a wooden beam as a device. One may be able to construct a physical model of the beam, but the function of the beam may only be inferrable by examining the larger physical system it is placed in. A wooden beam may be used to form the keel of a boat, act as a stationary wave breaker, or as a support in a roof. Another good example is a propeller fan—its function changes depending on whether it is immersed in water, air or vacuum, and whether it is a passive device or being rotated by an engine.

The second major criticism levelled at the NFIS principle by Keunke and Allemang is that the more closely it is adhered to, the more background knowledge a reasoning system will need to interpret a device. If a battery is described purely in terms of its chemical components—a functionally poor description, a large amount of knowledge about electrochemistry would be needed to decide what the battery was doing in an electrical circuit. If some assumptions can be made about its function, however, then less background knowledge is needed, and the process of inference can be more efficient. It may be that de Kleer and Brown's initial domain of electronic circuits was more suited to the application of the NFIS principle, but it is clear that it fails to generalize for the representation of a wide class of common devices.

Forbus (1988) also criticizes the device representation because not all phenomena can easily be described in it. For example, the device representation has difficulties in representing continuously varying systems. However, one might be forgiven for dismissing this second criticism. A reasoning system should always use the most appropriate representation, and for domains like digital circuits for example, the device representation is most appropriate because of the structured nature of the environment.

5.2 Recording causal and functional context

Causal and functional descriptions of physical systems are explanatory notions. Given a system description, a number of different explanations should be derivable, based upon the context in which it is used. It is interesting to note that despite their similar roles, functionality and causality had quite different initial treatments—causality was represented explicitly and functionality explicitly prohibited from representation.

As we have seen, the value of causal and functional explanations of device operation are dependent on the context in which they are obtained. Many workers are now engaged in developing and using modelling languages that make explicit use of device purpose or function (e.g., Sticklen, 1991). This view is entirely supported by Compton's (1990) work on knowledge acquisition. Rather than attempting to have functionally or causally devoid descriptions of physical systems, Compton would have such knowledge recorded with the functional or causal context in which it was acquired. This view sees the explanations that a model can generate only being valid in the context in which the model was initially constructed. While attempting to represent the structural aspects of a physical system in as general as possible a manner, the assumptions under which the model is created or acquired are also explicitly recorded wherever possible. The manner in which such general models can be constructed will now be examined, along with the ways a reasoning system can use multiple models and select amongst them based upon the assumptions associated with them.

6 Qualitative differential equations

Whether qualitative knowledge represents the manner in which a device is assembled (de Kleer, 1984), or a description of the processes occurring within a system (Forbus, 1984), it is usually modelled as a set of qualitative differential equations. A considerable effort has been devoted to developing qualitative algebras (e.g., Williams, 1988) which are sufficiently powerful to allow useful inferences to be drawn. This is a specialized topic in itself, and will not be dealt with in detail here. However, the general properties of the QDE representation are important, as they form a common language with which to compare the different representational schools in qualitative reasoning.

6.1 QSIM

QSIM (Kuipers, 1986) is a qualitative simulation algorithm that reasons with time varying functions. Its concern is to detail the time-varying relationships between functions, much as classical numeric simulation does. It thus is a relatively low level description language and could be placed within either the device or process centred description languages (Forbus, 1988b). QSIM will thus serve as a general model of a system that utilizes the qualitative differential form. The QSIM algorithm for qualitative simulation makes use of some of the best features of previous simulators such as those proposed in de Kleer (1984) and Forbus (1984). The algorithm and qualitative representation is presented formally by Kuipers, along with a critical analysis. It represents a milestone in the qualitative reasoning literature, presenting for the first time a vehicle that can be subjected to active experimentation and formal analysis by other workers.

6.1.1 The model

QSIM reasons with a constraint model. These constraints are qualitative abstractions of ordinary differential equations (ODE). Thus a model can be considered to be a set of *qualitative differential equations* (QDE). There is no commitment in the inference process to the meaning of the equations—QSIM is not interested in whether they represent some component of a device, or have a causal or functional interpretation. This is because there is no need to use such knowledge in QSIM's simulation procedure. Modelling is seen as the direct transformation of ODEs to QDEs, and the source of the model is considered external to the simulation process. QSIM uses several representational concepts common to many qualitative reasoning systems. In particular, the use of qualitative constraints, a discrete value set for functions, an interval representation of time, and a quantity space for each parameter are almost universal.

6.1.2 Functional form and the quantity space

Within the QSIM formalism, a physical system is characterized as a set of real valued parameters which vary continuously over time. Discontinuous functions are represented by separate models for each continuous section of the function. The values a function may take in the simulation of a qualitative model are usually taken from a restricted set of quantitative values. This reflects the loss of detail in abstracting from ODEs to QDEs. The qualitative description of a function comprises a pair ⟨*qval, qdir*⟩. The *qval* is a real value or real interval, and the *qdir* is the sign of the time derivative for the function at *qval*. The sign is represented by *inc, dec* or *std*, reflecting whether the function is increasing, decreasing or steady at a given time point. Further, each real *qval* is a *landmark value*, which represents a critical point for the function. The simulation is not interested in the values a function takes when it is not at a landmark, the specification of a range between two landmarks being sufficient. The set of landmarks values for a function form its *quantity space*, and this is used to define all the values of interest for a function.

A simulation involves obtaining the qualitative description for each function at each successive qualitative state that the system passes through in time. A progression to another state thus occurs whenever any of the functions in the system model changes its value. The *qualitative behaviour* of a system can now be defined as the temporal sequence of qualitative states for each of its parameters as constrained by their individual functional relationships.

6.1.3 Model constraints

Modelling a system involves representing functional relationships with qualitative constraints. Constraints in QSIM restrict the qualitative values that may be assigned to functions. The basic

QSIM constraints allow relations of addition, multiplication, the time derivative and monotonicity to be expressed qualitatively. For example, the monotonically increasing constraint is

$$M^{+}(f,g) \text{ iff } f(t) = H(g(t)) \text{ where } H'(x) \text{ is strictly monotonically increasing}$$

The monotonic constraints M^{+} and M^{-} map many quantitative functional relationships. For example, both exponential and logarithmic functions map onto them. Periodic functions must be broken up into separate regions that are monotonically increasing or decreasing. This many to one mapping allows QSIM to specify incomplete knowledge about a system, but has the trade-off that the simulation may produce ambiguous results.

6.1.4 Limitations of qualitative representation and simulation

While they are a powerful representational form, QDEs are still limited in some respects, as is qualitative simulation. Some of their major problems will now be discussed, along with examples of work that seek to address these and extend the formalism.

Representational coverage One of the motivating factors behind the QDE representation is its ability to represent a broad range of physical phenomena at a level which allows useful and verifiable inference to be drawn. There are classes of phenomena that may however, not be covered by the QDE formalism. In particular, some relationships between parameters may be best captured in purely probabilistic terms.

Wellman (1990) formalizes the probabilistic semantics of qualitative influences which express monotone relations between functions. Qualitative influences become statements of constraint on the joint probability distributions of parameters. Statements like "A positively influence B" are interpreted as a special case of relative likelihood in Wellman's formulation. Such statements leave open the possibility that a relationship is non-deterministic. A monotonically increasing QDE prevents two functions from ever changing in opposing directions, a positive probabilistic influence merely states that it is unlikely.

Qualitative probabilistic influences are, however, limited to expressing monotone associations, and are shown to be unable to resolve combinations of influences of different signs without recourse to other information. The probabilistic representation is thus complementary to the more powerful QDE formalism.

Ambiguity of the qualitative algebra Often the ambiguity of qualitative relationships prevents a single prediction to be made. For example, given the relationship *add(a,b,c)* when a is specified as having a positive value and b as negative, then c is completely unspecified. This is the trade-off for the power of the qualitative algebra in specifying incompletely characterized systems. One technique to help resolve the problem is to use some quantitative or order of magnitude knowledge. Knowing that $a > b$ in absolute terms for the above example constrains c to be a positive value. One method for using such incomplete quantitative knowledge is proposed in Kuipers & Berleant (1988).

A more rigorous extension to the qualitative algebra is proposed in Williams (1988). He allows a combination of qualitative and real operators. This allows the delay of the abstraction of values into qualitative terms until after algebraic manipulation has occurred. The result of delaying the qualitative abstraction is that more information is preserved. He also completely restores the properties of addition, multiplication, subtraction and division operators. Previous algebras like that used in QSIM focused on the results of using these operators on the signs of functions, not their actual values.

Locality Not all the behaviours generated by the process of qualitative simulation are true, and this is a key observation contributed by Kuipers. This is partly because of the so called "locality" of the simulation process. The transition to the next state in a simulation is determined solely by its predecessors' value. In particular, the generation of transition values is not constrained by any global behaviour of the system. However, subsequent work on the representation and utilization of second derivative constraints (Kuipers, 1987; Chiu, 1988) and phase portraits (Lee & Kuipers, 1988; Struss, 1988) have signficantly reduced the locality problem. It seems that the simulation languages like QSIM are thus not necessarily inherently flawed, but more likely incomplete.

Temporal abstraction Although qualitative simulation is concerned with the time course of function values, time itself plays a secondary role in its determination. The temporal resolution is limited to distinguishing distinct instants and intervals, and to force a total ordering on the state succession. In short the temporal representation for QSIM, and most other simulators, is poor (Weld, 1988). Temporal reasoning forms a distinct AI research area, but there is significant synergy between its goals and those of qualitative reasoning. McDermott (1987) provides a review of some of the recent work in temporal logics.

When QSIM is faced with large models, the number of potential behaviours for a set of initial conditions can grow alarmingly. Kuipers (1987a) proposes the decomposition of large models through the application of temporal abstractions. Briefly, a hierarchy of subsystems in the model is created, based on the relative time scale the subsystems operate in. Slow systems that might take a day or two to respond to changes are separated from much faster subsystems that might operate within seconds. As shown in section 4.2.3, functions in a fast system see slower systems as if they were constant, and slow systems see fast ones as being instantaneous. QSIM is applied repeatedly to each subsystem, shifting attention from slow systems to the faster ones they communicate with, and returning the simulation values obtained to the slower system again. This is a simple extension to QSIM that allows the time scale of constraints to be specified. Without it and the global constraints discussed above, QSIM is restricted to small scale models.

6.2 Qualitative process theory

While QSIM avoids the issue of assigning meaning to its constraints, other workers are more concerned with the modelling process, and the interpretation of such constraints. In an attempt to find a qualitative representation that reflects the way people informally reason about the physical world, Forbus (1984) proposed his qualitative process theory (QPT). His intuition was that people describe changes in terms of *processes* like liquid flow, motion, boiling, bending and compressing. His processes "are the analog of the differential equations used to describe the dynamics of a system".

Forbus (1988) lists among the advantage of the process representation that it can represent discontinuous events for substances, that processes provide a simple notion of causality, and that it allows modelling assumptions to be explicitly represented. These may be true, yet they are not exclusive to QPT. Most qualitative simulation systems, including QSIM, will allow different sets of constraints and relations to be invoked as a system moves into different operating ranges (section 7.1). One of the disadvantages of QPT is that setting a model up is quite laborious. Further, there are no concise algorithmic accounts of the QP simulation mechanism. It was the emphasis on algorithmic definition that allowed Kuipers to make his analysis of some of the limits of qualitative simulation in Kuipers (1986). Forbus (1984) is an exposition of a theory in evolution, not a simulation technique. Forbus (1990) attempts to partially redress this imbalance.

6.2.1 Modelling in QPT

QPT attempts to model continuous physical processes, as well as discontinuities like the creation or destruction of objects, or the transition from heating water to it boiling. Boiling for example, would be a separate process to heat flow, and would be invoked during a heat flow process when the liquid's boiling point was reached. Forbus gives Hayes' notion of the history of an object primacy, and models time as a sequence of intervals and points. Within QPT, physical models are described by a *view* and a set of *influences* on an object. The view is the list of individuals involved in the object and any conditions and relations amongst them. In QSIM this would be the names of functions, their operating ranges, and any other constraints on the values they might take. The influences loosely correspond to QSIM's qualitative constraints. Forbus also uses the notions of quantity space, and corresponding values for monotonic constraints. So, despite the emphasis on process as an organizing concept, much of the information represented in a QSIM model is also representable as a process model. In fact, QPT was an important influence in the devlopment of QSIM, along with de Kleer and Brown's confluence based theory. Bylander (1988) provides a direct comparison of these schools through an example of heat flow into a pan of water, and reinforces some of the similarities of their representations. While such a direct comparison between the bodies of work represented by QPT and QSIM sheds some light, it can also be misleading. As pointed out earlier, one can view QSIM's constraint language as a subset of an ontology like QPT. The implication then is that QPT is attempting to represent and infer about a larger set of phenomena. Indeed, Forbus claims that QPT is the most general account of qualitative reasoning to date.

6.2.2 The qualitative process compiler

An important contribution to this discussion is the work presented by Crawford et al. (1990) on the qualitative process compiler (QPC). They have implemented a system that separates qualitative reasoning into two quite separate tasks—that of building a model, and then of simulating the model to obtain behavioural predictions. They assert that some of the difficulties that have arisen in comparing the different representations lies in the confusion of the model building and simulation tasks, and that the separation of the two allows clearer statements about the values of different representations to be made.

Broadly, QPC first assembles a model of a problem using the view-process representation of QPT, and then transforms that model into a set of QDEs, which are then simulated through QSIM. The transformation process is most illuminating, and there are several aspects to it. Influences in QPT are causally directed statements of relationship—that X will influence Y, all else being equal. We know from the earlier discussion of causality that QSIM constraints are undirected functional relationships. This raises a conflict in semantics. An example will make this clearer. Assume two parameters A and B influence a third C. In QPT if A provides a positive influence and B a negative influence, then C is ambiguously defined. If these are monotonic relations in QSIM, the situation is not ambiguous, but rather is inconsistent. One constraint says C should be increasing, and the other says it should be decreasing—clearly it cannot do both. So the labelling is inconsistent.

QPC attempts to resolve this problem by recognizing that this semantic difference does indeed exist, and attempts to circumvent it. The process takes the following general form (Figure 3). First, QPC applies a closed world assumption when transforming influences. Thus, if it is known that only A and B influence C, one can assume that no other influences exist (even though the influence semantics does not preclude their existence). Next, QPC transforms QPT proportionalities into monotonic relations, and influences into derivative relations. These are then summed using the QSIM *add* constraint. Some recent discussion has focused on whether this genuinely represents an assumption that the functions involved in the influences can be linearly combined, as stated by

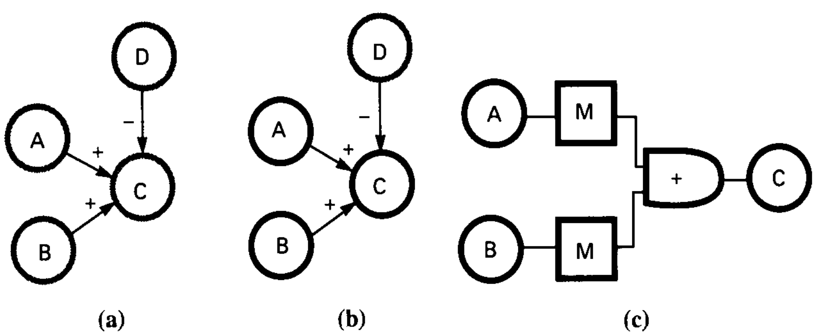

Figure 3 QPC transforms QPT influences into QSIM constraints: (a) There are potentially many influences; (b) Select the actual influences and apply a closed world assumption; (c) Transform the influences into QSIM constraints and combine linearly.

Crawford (1990), or whether the QSIM constraint $add(a,b,c)$ does not in fact require linearity, but only maintains that c is monotonically increasing in a and b.

The issue of the disappearing causality moving from influences to constraints is not addressed in the transformation. One of Bunge's (1979) criticisms of causality is that it assumes the linear independence of causal influences—that each causal influence is independent of the others, and that when necessary one can superimpose causal influences and derive their interacting effect. Thus QPT assumes qualitative linearization when summing influences. There is nothing in QSIM's representation which inherently requires this assumption—it only becomes apparent when the transformation between representations is attempted. It may well be that representational systems like QSIM which treat causality as an inferred property, and those like QPT which represent it explicitly, do model slightly different phenomena. If one takes Kuipers' view (1987), then the causality captured in QPT influences should instead be modelled at a finer level of temporal abstraction as a set of acausal qualitative differential equations.

QPT provides an important organizing point in the qualitative reasoning literature, and has identified key ideas used by subsequent workers. The urge to integrate QPT with other representational systems is an expression of the need to understand whether these represent essentially incompatible world views, or have differences based on misunderstandings through poor specification. It seems clear that the transformation described above does preserve some important aspects of the relations between parameters, and thus establishes commonalities between the two ontologies. Equally, it is clear that the transformation is not perfect. The choice of a process, device or other representation of the physical world at this stage perhaps should not be an article of faith, but a pragmatic choice. The representation which best describes the situation is the most appropriate. Weld (1986), for example, uses the process as the organizing concept for his work in molecular genetics, while Ironi (1990) uses QSIM for modelling human iron metabolism.

7 Multiple representations

There is a growing interest in equipping AI systems with multiple representations and multiple models of a particular domain, and there are a number of motivations for this. While a diagnostic system equipped with deep models may handle a broad range of problems, it will often be inefficient computationally. Applying complex inference to problems that are simple or common is inefficient when they could be adequately dealt with by a less detailed or even shallower representation. Thus, the selection of representation should be appropriate to the problem at hand. Systems may also need to switch between different models of a physical system when one is a more useful representation of the phenomenon at hand—problems may only have a solution after changing perspective. Finally, a representational form or particular model may make assumptions that are invalid in the context of a given problem, and a more appropriate one needs to be selected. Equipping a reasoning system with the capacity to switch between models and representations should make it both robust and efficient.

From a representational viewpoint, the need to switch between models is fascinating because it requires an understanding of the mappings that exist between models of the world. Pragmatically, the mapping allows the definition of limitations and hidden assumptions of a given representation (e.g., QPC's mapping of QPT to QSIM). A number of dimensions can be identified along which models could be distinguished from one another, and their commonalities mapped. These include

- mode of behaviour
- granularity of detail
- depth of representation
- ontological view

While work in this area is still in its early stages, it seems to represent a critical part in the development of systems that need to deal with a wide variety of real world phenomena.

7.1 Mode of behaviour

Depending on the conditions that currently affect it, a system might behave in several quite different modes, each requiring a different model. For example, one might wish to model the behaviour of water as a fluid in normal circumstances, but the model must be changed when temperature drops below freezing. Such behavioural switches require explicit modelling assumptions to be associated with individual models. Facilities to enable model selection are included in QPT. One of the first things done when simulating with QPT is to determine which processes are currently active, based on the initial conditions.

QSIM also allows model selection (Kuipers, 1986). For example, complex or periodic functional relations can be partitioned based upon range descriptions, decomposing a system's behaviour around discontinuities. For the relation x = cos θ, one could specify a model description as a series of tuples of the form FC(θ,x,descript). Each tuple describes a region of the function with a qualitatively different behaviour

$$\text{descript} = \begin{cases} (0,1) \\ ((0,\pi),(1,-1),M-(\theta,x)) \\ (\pi,-1) \\ ((\pi,2\pi),(-1,1),M+(\theta,x)) \\ (2\pi,1) \end{cases}$$

7.2 Granularity of detail

The ability to represent the same physical system at increasingly finer levels of detail was an important component of Patil's (1981) work on causal reasoning in the domain of acid-base disorders. As problems become more complex, the reasoner resorts to more detailed knowledge in an effort to solve them. The cost is computational complexity. The computation required to generate qualitative states from a model when simulating is exponential with the number of variables (Kuipers, 1986). Models that optimize problem coverage and computational complexity thus need to be selected.

Iwasaki (1990b) and Weld (1991) list a number of axes along which one could choose to abstract model detail:

Structural Components can be lumped together if they are structurally related. Components are thus represented at a level that is sufficient to solve the current set of problems. Bylander's work (1987) on consolidation emphasizes a notion of structural containment and composition. Here components can be lumped together and considered as a single more abstract and simpler one.

Functional Collectively, one may wish to model the function of a number of components as a single unit. Recall from section 5 that while it may not be ideal to explicitly represent device function, it is often practically unavoidable and also useful.

Temporal One may choose to reason about behaviours that occur over a certain span of time only—perhaps ignoring very brief or very extended ones (e.g., Kuipers, 1987).

Quantitative While increased coverage of physical phenomena is obtained by abstracting away quantitative details, reasoning power is lost. We may thus wish to include quantitative models of a system to assist with more difficult problems. There are in fact a range of models that can be contemplated, from the purely qualitative, through semi-quantitative representations (Kuipers & Berleant, 1988) to fully quantitative ones. Abstraction from quantitative to qualitative representations can operate both on the values of a parameter or its relations (e.g., ODEs to QDEs).

Approximation assumptions As models become more abstract, insignificant terms can be eliminated, creating coarser approximations to the more detailed model. While an approximate model may be useful in solving a particular problem, its predictions do not necessarily agree with the more detailed model with which it is associated. For example, one may consider the equations of Newtonian physics to be a useful approximation of relativistic physics, with the discrepancies between the two theories only becoming evident in special circumstances. Weld (1991) seeks to derive approximate models which are formulated in response to a particular problem context. The approximate models only have validity in the context of the problem being solved.*

These dimensions, as Iwasaki points out, are not necessarily independent. As we move along one dimension in an abstraction hierarchy, we may move along others as well. Parts that are structurally close also tend to operate in the same time frame and may represent a functional unit.

7.3 Ontological view

A battery can be seen as a set of chemical reactions or as a supplier of electromotive force. The view used depends on problem solving needs. As the discussion of the NFIS principle showed, different ontological views are selected based on the functional context of our problem. Placing a battery in the context of an electronic circuit presupposes that it is viewed as a supplier of electromotive force. Employing a particular language to describe a device's function (e.g., current, voltage) thus commits to a view of the device that supports that language (Liu & Farley, 1990). One would only want to consider the more detailed chemico-physical level if there was a problem with the circuit that could not be solved by the simpler view of the battery, or if one were reasoning about a battery in a context outside that of an electronic circuit. The initial work on ontology began when Hayes (1985) presented two views of liquids—one as stuff that is held within a container, and the other as multiple small pieces of stuff. Collis & Forbus (1987) extend the second ontology to modelling fluids as molecular collections.

Until recently, little has been done to explore the ways in which different views of a system could be utilized, and even less on the relations between such models, or how a reasoner would decide which view was most appropriate. One such attempt is presented by Liu & Farley (1990) who offer a charge-carrier ontology to supplement the standard device ontology for digital electronic circuits. Importantly, they begin to examine the relations between the two ontologies, albeit with small examples. They offer a number of rules for selecting between the device and charge carrier ontology when reasoning about problems with circuits. They first note that there must be some method of linking ontologies through bridging relations. An example they offer is that a perturbation in current (device ontology) can be viewed as a perturbation in charge flow (charge carrier ontology). Perturbations between ontologies can be carried through such bridging relations. Addanki et al. (1989b) have also worked on the problem of linking common parameters between different problem solving contexts.

Ontologies are shifted in Liu and Farley's work through the application of a number of choice rules. For example, if the question at hand interrogates an axiom of our current ontology, then there is no way of deciding the answer. One would then switch to a related ontology in which the question does not address axioms. If in the device ontology, one had the axiom that $V = IR$ and sought to know why current increases when voltage increases across a resistor, then an answer cannot be provided. Switching to the charge carrier ontology, an explanation of the axiom can be generated in terms of the movement of electrical force, cross-sectional area and the movement of the charge carriers through the area.

Liu and Farley posit a structural compatibility principle—that to preserve spatiotemporal continuity between shifts, different ontologies must share a common structural view of the world. The common structure between device and charge carrier views is the notion of a region of space. They state that the principle allows bridging relations to be formulated, and that it ensures the

*This work offers an interesting example of knowledge being formulated in the context of a query—and may offer some insight into the processes discussed in Compton (1990).

continuity of causal propagation when shifting between different ontologies. The second assertion is presented without formal justification. Since causal ordering cannot be preserved when shifting between different levels of abstraction within the same ontology (Iwasaki, 1990b), it seems unlikely that it would be preserved between different ontologies. Clearly much work needs to be done exploring such issues.

7.4 Depth of representation

Deep models may contain knowledge that is required for solving a large number of problems, but the reasoning architectures that support deep representations are comparatively expensive computationally. Some work has gone into developing systems that capture domain knowledge at a deep level, but then extract a large number of shallow relations from these models. Often in the form of rules, the shallow representations is then used for problem solving. The Kardio system (Bratko, 1989) developed detailed qualitative models of cardiac electrophysiology, and then generated large numbers of rules using the model. Rules were not just for single diseases, but included the patterns expected for multiple concurrent diseases (or faults).

The cost of increased efficiency is storage, and problem coverage to some extent—as one moves from deep to shallow computational efficiency is gained at the price of representational efficiency. A deep model is a compact representation when compared to the large number of rules that can be generated from it. Bratko's team solved the storage problem by eliminating large numbers of unlikely or impossible combinations, and compressed the knowledge base through an inductive learning algorithm. Another problem with this general approach is that it is also not feasible to generate rules for all possible disease interactions. An arbitrary number of faults per interaction needs to be selected. Interactions due to a larger number of faults than covered by the rules will need access to the full model to be solved, rare though they may be.

Some systems (e.g., Abu Hanna, 1990; Coiera, 1989; Dvorak, 1989; Ironi, 1990) attempt to use both deep and shallow representations, optimizing in some way the selection of representation and inference task. Dvorak's Mimic uses QSIM models to predict behavioural states, but in a manner similar to Kardio, uses model generated rules to select initial hypotheses. It cannot, however, hypothesize event interactions not coded directly into a recognition rule, and so would potentially miss identifying events that in principle are identifiable from its deep model. Few systems to date actively switch between levels based on an understanding of the problem at hand, selecting the representation that is most appropriate. Further, it is not altogether clear under which conditions one would need to move between these two representational extremes.

7.4.1 Intermediate representations

Rosch's classification theory (1978) proposes that the most useful level of representational abstraction of an object is the most cognitively economic one—which she calls the basic level. Objects in the basic level have the quality that they are *prototypic* of their class. Such prototypic instances contain the attributes that are most representative of items inside, and least representative of items outside a category. A classic example is the concept of a chair, which is considered a prototype. Specializing the description of an object may not give much more information, e.g., a dining chair, and a lot is lost by abstracting to a more general category like furniture. Similarly, representations intermediate between deep and shallow are specializations of the deeper level, and generalizations of the shallow. They may thus share some of the characteristics of a Roschian prototype, optimizing the balance between computational and representational efficiency on the one hand, and problem coverage on the other.

Rosch's work on representation is largely concerned with the classification of objects, rather than processes or events. She suggests, however, that a good candidate for the prototypic representation of events are Schankian scripts describing individual units of action like making a cup of coffee or going to a lecture. The task of diagnosing the time varying behaviour of systems is a concern for many AI systems. Consequently, representations that allow physical processes to be inferred or represented in prototypic form are of particular interest. Little has, however, been done in this area to-date.

Qualitative histories, composed of a time ordered sequence of qualitative state descriptions, are one candidate for a prototypic representation intermediate between qualitative models and shallower representations. Hayes' Naive Physics Manifesto (1979) proposed the history of an object as a key reasoning concept. A history is "a connected piece of space time" describing the occurrence of a process on or through an object. Hayes saw a history as "a basic ontological primitive", but most work since Hayes' paper has focused on the mechanisms by which histories can be generated. Histories have become a by-product of an inference process such as qualitative simulation rather than knowledge worth representing in its own right.

Histories can be used for diagnosis or prediction, and a key requirement for both of these tasks is to be able to represent the behaviour of multiple interacting events. This requirement proves to be a central issue when reasoning with explicit representations of qualitative histories. Theoretical results in (Coiera, 1989, 1992), based on the mathematical properties of qualitative histories, show that histories of individual events can often be added using qualitative superposition to replicate the behaviours of the interacting events. The superposition is legal for all linear systems, and for non-linear systems, occasionally requiring special conditions, such as the dominance of one history over another. Based upon these results, some preliminary conclusions can be drawn about the benefits of using qualitative histories as an intermediate representation. They offer representational efficiency intermediate between recognition rules and full models, since interactions do not need to be generated and stored. Computationally, they are again intermediate—requiring some computation to predict interactions but not a full simulation. Problem coverage also suggests the intermediate nature of histories. One can handle, within the limits of superposition, an arbitrary number of diseases in any interaction thus extending problem coverage from a shallow system. Equally, superposition limits define when a deeper model must be invoked.

The benefit of any representational depth from shallow to deep clearly varies with domain and problem type. Another constraint on choice is more practical than theoretical, and results from the degree to which knowledge has been formalized in a domain. If deep models are unavailable, then problem solving must rely on shallower techniques. An advantage of intermediate representations is that they should allow a system to maximize problem coverage in such poorly specified domains. For example, physiological models of disease may be unavailable, but if diseases can be represented as histories one can make more powerful inferences than would be possible with a shallower representation (Coiera, 1990).

7.5 Switching between views

It is clear from the preceding discussion that there are numerous ways in which one can represent physical phenomena. A key ability for any reasoning system that utilizes such different views of the world is to intelligently switch representation. Switching between views, or indeed selecting amongst a set of candidate views requires an understanding of the limitations of any particular view and its representation, and a means of selecting a more appropriate one.

If there existed some measure of the likely success of a model in answering our problems, then one would have a means of deciding whether it is appropriate or whether it is necessary to switch. The work on ontology switching by Liu & Farley (1990) offers some rules for selecting ontology based on a notion of problem solvability within an ontology. The work on qualitative history superposition (Coiera, 1990) offers another measure of success based upon the assumptions that need to be met before intermediate depth histories can be used in favour of the deeper model-based level. Such a view selection process could also consider computational cost when several alternatives are available.

Other workers have relied on the failure of a current model to drive the selection of an alternative (e.g., Addanki et al., 1989; Weld, 1990, Abu-Hanna & Gold, 1990). Addanki et al. make a particular attempt to quantify model failure, and direct the selection of an alternate model based on a metric of appropriateness. They represent multiple models in a graph, each node being a particular model, and the arcs between models explicitly detailing the assumptions that change in transition from one model to another. For example, a model of a physical system may assume that no friction operates or that energy is conserved during the operation of a device—assumptions which often simplify problem solving. Their system checks to see if the current model prediction matches actual meassurements, and assembles a list of delta vectors which record the mismatches that have been detected. It then searches for a model connected to the current model which would produce the desired changes in parameters associated with a delta vector. The information detailing expected parameter change is recorded in explicit rules, and it is not clear how often such rules could be found to assist in the selection process. Weld (1990) describes another method of linking models which are approximations of one another. If one model is a simpler approximation of another, and a mapping called an approximation reformulation can be found, then one can switch from one to the other. The reformulation is possible when the complex model has an exogenous parameter that allows the behaviours of the two models to become successively closer as it tends towards a limit. It is not clear how often such reformulations are possible between different models, and whether the technique will thus prove general enough to be practicable.

There are clearly a variety of clues that a reasoning system can exploit when selecting an appropriate world view. All of the work described here has focused on a particular context switch—either between ontologies or representational depths or approximations. A more unified approach is perhaps needed to analyse the interrelation between the different strategies, specifying their individual utilities and success rates, and generality.

8 Conclusions

A large body of work has been covered in this paper, and it is clear that the qualitative representation of physical systems is still evolving actively. Early confusion over issues like causality, and the relation of structure and function in a representation, while not entirely resolved, are much clearer. Equally, the relations between the various competing qualitative knowledge representations like QPT or QSIM are now beginning to be formalized profitably.

One of the disappointments of the last few years has been the scant attention paid to the more commonsense aspects of real world representation, in favour of more formal ones of complex physical systems. With the current interest in multiple representations and models, this may begin to be redressed. Commonsense knowledge may be important not just in dealing with physical phenomena at a simple level, but form a part of the knowledge with which one can switch from one model to another. It will also be interesting to see the outcome of the CYC project's attempt at capturing commonsense knowledge.

It seems likely that the construction of truly robust and efficient programs that will reason across a wide range of problems will require the multiple model approach. The growing understanding of the benefits of deep and shallow representations is beginning to allow the design of such systems, and to contemplate intermediate representations that may optimize functionality between the two. At one stage deep representations were favoured because of their perceived benefits over shallow systems. The ability to now actively envisage a synergy between such representational levels suggests that significant progress in the design of robust reasoning systems may be made in the next few years.

Acknowledgements

Early parts of this work were carried out while the author was at the Department of Computer Science at the University of New South Wales, with the support of a New South Wales Government Medical Research Training Scholarship, and a research award from the Australian Medical Engineering Research Organisation. Graham Higgins deserves thanks for introducing me to

Rosch's work on classification, and along with Elpida Keravnou and Chris Preist offered comments on an earlier draft of this paper.

References

Abu-Hanna, A and Gold, Y, 1990. "Adaptive, multilevel diagnosis and modelling of dynamic systems" *International Journal of Expert Systems* **3** 1–30.

Addanki, S, Cremonini, R and Penberthy, SJ, 1989. "Reasoning about assumptions in graphs of models". In: *Proceedings of the 11th International Joint Conference on Artificial Intelligence,* pp 1432–1438.

Addanki, S, Cremonini, R and Penberthy, SJ, 1989. "Contexts: dynamic identification of common parameters in distributed analysis of complex devices". In: *Proceedings of the 11th International Joint Conference on Artificial Intelligence,* pp 104–109.

Bobrow, D. (ed.), 1984. *Qualitative Reasoning About Physical Systems* MIT Press.

Bratko, I, Mozetic, I and Lavrac, N, 1989. *Kardio—A Study in Deep and Qualitative Knowledge for Expert Systems* MIT Press.

Brooks, R, 1991. "Intelligence without representation" *Artificial Intelligence* **47** 139–159.

Bunge, M, 1979. *Causality and Modern Science* Dover Publications.

Bylander, T, 1987. "Using consolidation for reasoning about devices" Technical Report, Laboratory for Artificial Intelligence Department of Computer and Information Science, Ohio State University.

Bylander, T, 1988. "A critique of qualitative simulation from a consolidation viewpoint" *IEEE Transactions on Systems, Man, and Cybernetics* **18** 252–263.

Bylander, T, 1990. "Some causal models are deeper than others" *Artificial Intelligence in Medicine* **2** 123–128.

Chandrasekaran, B, 1983. "Towards a taxonomy of problem solving types" *The AI Magazine* Winter/Spring, pp 9–17.

Chiang, AC, 1984. *Fundamental Methods of Mathematical Economics* McGraw-Hill.

Chiu, C, 1988. "Higher order derivative constraints and a QSIM-based total simulation scheme" Technical Report AITR88-65, The University of Texas at Austin, Artificial Intelligence Laboratory.

Cohn, AG, 1989. "Approaches to qualitative reasoning" *Artificial Intelligence Review* **3** 177–232.

Coiera, EW, 1989. "Reasoning with qualitative disease histories for diagnostic patient monitoring" Ph.D. Thesis, Department of Computer Science, University of New South Wales.

Coiera, EW, 1990. "Monitoring diseases with empirical and model generated histories" *Artificial Intelligence in Medicine* **2** 135–147.

Coiera, EW, 1992. "Qualitative superposition" Artificial Intelligence (to appear).

Collins, JW and Forbus, KD, 1987. "Reasoning about fluids via molecular collections". In: *Proceedings of AAAI-87,* pp 590–595.

Compton, P and Jansen, R, 1990. "A philosophical basis for knowledge acquisition" *Knowledge Acquisition* **2** 241–257.

Crawford, J, Farqhuar, A and Kuipers, B, 1990. "QPC: A compiler from physical models into qualitative differential equations". In: *Proceedings of AAAI-90,* pp 365–372.

Davis, R, 1984. "Diagnostic reasoning based on structure and behaviour" *Artificial Intelligence* **24** 347–410.

Davis, R, 1987. "Robustness and transparency in intelligent systems". In: *Proceedings of the Third Australian Conference on the Applications of Expert Systems,* pp 143–164.

de Kleer, J and Brown, JS, 1984. "A qualitative physics based on confluences" *Artificial Intelligence* **24** 7–83.

de Kleer, J and Brown, JS, 1986. "Theories of causal ordering" *Artificial Intelligence* **29** 33–61.

Fishwick, PA, 1989. "A study of terminology and issues in qualitative simulation" *Simulation* **52** 5–9.

Fishwick, PA, 1989b. "Qualitative methodology in simulation model engineering" *Simulation* **52** 95–101.

Fishwick, PA and Zeigler, B, 1990. "Qualitative physics: Towards the automation of systems problem solving". In: *Proceedings on AI, Simulation and Planning in High Autonomy Systems* IEEE Computer Society Press, pp 118–134.

Forbus, KD, 1984. "Qualitative process theory" *Artificial Intelligence* **24** 85–168.

Forbus, KD, 1988. "Commonsense physics: A review" *Annual Review of Computer Science* **3** 197–232.

Forbus, KD, 1990. "The qualitative process engine". In: D Weld, and J de Kleer (eds.), *Readings in Qualitative Reasoning about Physical Systems* Morgan Kaufmann.

Forbus, KD and Gentner, D, 1990. "Causal reasoning about quantities". In: D Weld and J de Kleer (eds.), *Readings in Qualitative Reasoning about Physical Systems* Morgan Kaufmann, pp 686–677.

Gandolfo, G, 1981. *Methematical Methods and Models in Economics* North-Holland.

Hamscher, W and Davis, R, 1987. "Issues in model based troubleshooting" AI Memo 893, Artificial Intelligence Laboratory, Massachusetts Institute of Technology.

Hart, PE, 1982. "Directions for AI in the eighties" *SIGART 79,* pp 11–16.

Hayes, PJ, 1979. "The naive physics manifesto". In: D Michie (ed.), *Expert Systems in the Microelectronic Age* Edinburgh University Press.
Hayes, PJ, 1985. "Naive physics 1: Ontology for liquids". In: JR Hobbs and RC Moore (eds.), pp 71–89. (Reprinted in Weld, D and de Kleer, J, *ibid.*)
Hobbs, JR and Moore, RC (eds.), 1985. *Formal Theories of the Commonsense World* Ablex.
Ironi, L, Stefanelli, M and Lanzola, G, 1990. "Qualitative models in medical diagnosis" *Artificial Intelligence in Medicine* 2 85–101.
Iwasaki, Y and Simon HA, 1986. "Causality in device behaviour" *Artificial Intelligence* **29** 3–32.
Iwasaki, Y and Simon, HA, 1986. "Theories of causal ordering: reply to de Kleer and Brown" *Artificial Intelligence* **29** 63–72.
Iwasaki, Y, 1987. "Generating behaviour equations from explicit representations of mechanisms" Carnegie Mellon University, Computer Science Department Report CMU-CS-87-131.
Iwasaki, Y, 1988. "Causal ordering in a mixed structure". In: *Proceedings of AAAI-88,* pp 313–318.
Iwasaki, Y, 1990. "Reasoning with Multiple Abstraction Models" Knowledge Systems Laboratory Report No. KSL 90-52, Stanford University, August.
Iwasaki, Y, 1990. "On the relationship between model abstraction and causality: Variance of causal ordering under abstraction operations" Knowledge Systems Laboratory Report No. KSL 90-59, Stanford University, September.
Keravnou, E and Washbrook, J, 1989. "Deep and shallow models in medical expert systems" *Artificial Intelligence in Medicine* **1** 11–28.
Keuneke, A and Allemang, D, 1989. "Exploring the no-function-in-structure principle" *Journal of Experimental and Theoretical Artificial Intelligence* **1** 79–89.
Kirsch, D, 1991. "Today the earwig, tomorrow man?" *Artificial Intelligence* **47** 161–184.
Klein, D and Finin, T, 1987. "What's in a deep model? In: *Proceedings of the 10th International Joint Conference on Artificial Intelligence,* pp 559–562.
Kuipers, B, 1986. "Qualitative simulation" *Artificial Intelligence* **29** 289–338.
Kuipers, B and Chiu, C, 1987. "Taming intractable branching in qualitative simulation". In: *Proceedings of the 10th International Joint Conference on Artificial Intelligence* pp 1079–1085.
Kuipers, B, 1987. "Abstraction by time-scale in qualitative simulation". In: *Proceedings of AAAI-87* pp 621–625.
Kuipers, B, 1987b. "Qualitative simulation as causal explanation" *IEEE Transactions on Systems Man and Cybernetics* **17** 432–444.
Kuipers, B and Berleant, D, 1988. "Using incomplete quantitative knowledge in qualitative reasoning" *Proceedings of AAAI-88.*
Lee, W and Kuipers, B, 1988. "Non-intersection of trajectories in qualitative phase space: A global constraint for qualitative simulation". In: *Proceedings of AAAI-88* pp. 286–296.
Lenat, DB and Feigenbaum, EA, 1988. "On the thresholds of knowledge". In: *Proceedings Fourth Australian Conference on the Applications of Expert Systems,* pp 31–56. (Revised version reprinted in *Artificial Intelligence* **47** (1991) 185–250.)
Levesque, H and Brachman, R, 1984. "A fundamental tradeoff in knowledge representation and reasoning". In: *Proceedings of CSSI-84,* London, Ontario, pp 141–152. (Reprinted in *Readings in Knowledge Representation,* R Brachman and H Levesque (eds.), Morgan Kaufmann 1985) 41–70.
Liu, Z and Farley, A, 1990. "Shifting ontological perspectives in reasoning about physical systems". In: *Proceedings of AAAI-90,* pp 395–400.
Long, W, 1983. "Reasoning about state from causation and time in a medical domain". In: *Proceedings of AAAI-83,* pp 251–254.
Morris, M, 1991. "Why there are no mental representations" *Minds and Machines* **1** 1–30.
McDermott, D, 1987. "Logic, problem solving, and deduction" *Annual Review of Computer Science* **2** 187–229.
Pan, J, 1984. "Qualitative reasoning with deep-level mechanism models for diagnoses of mechanism failures". In: *Proceedings of CAIA-84,* pp 295–301.
Patil, RS, 1981. "Causal representation of patient illness for electrolyte and acid-base diagnosis" Ph.D. Thesis, Massachusetts Institute of Technology.
Rosch, E, 1978. "Principles of classification". In: E Rosch and B Lloyd (eds.), *Cognition and Categorization,* Erlbaum, 27-48. (Reprinted in *Readings in Cognitive Science,* Morgan Kaufmann, 1988), pp 312–322.
Söderman, U and Strömberg, J, 1991. "Combining qualitative and quantitative knowledge to generate models of physical systems". In: *Proceedings of the 12th International Joint Conference on Artificial Intelligence,* pp 1158–1163.
Sticklen, J, 1991. "Functional reasoning and functional modelling" *IEEE Expert* **6** 20–21.
Struss, P, 1988. "Global filters for qualitative behaviours". In: *Proceedings of AAAI-88,* pp 275–279.
Top, J and Akkermans, H, 1990. *Bond-Graph Based Reasoning about Physical Systems,* Working Papers of the 1990 Workshop on Model Based Reasoning, Boston.

Top, J and Akkermans, H, 1991. "Computational and physical causality". In: *Proceedings of the 12th International Joint Conference on Artificial Intelligence,* pp 1171–1176.

Weiss, S, Kulikowski, C et al., 1978. "A model-based method for computer-aided medical decision making" *Artificial Intelligence* **11** 145–172.

Weld, D, 1986. "The use of aggregation in causal simulation" *Artificial Intelligence* **30** 1–34.

Weld, D, 1988. "Comparative analysis" *Artificial Intelligence* **36** 333–373.

Weld, D, 1988. "Exaggeration". In: *Proceedings of AAAI-88,* pp 291–295.

Weld, D and de Kleer, J, 1990. *Readings in Qualitative Reasoning About Physical Systems* Morgan Kaufmann.

Weld, D, 1990. "Approximation reformulations". In: *Proceedings of AAAI-90,* pp 407–412.

Weld, D and Addanki, S, 1991. "Query-directed approximation". In: B Faltings and P Struss (eds.), *Recent Advances in Qualitative Physics,* MIT Press.

Wellman, M, 1990. "Fundamental concepts of qualitative probabilistic networks" *Artificial Intelligence* **44** 257–303.

Williams, B, 1988. "MINIMA—A symbolic approach to qualitative algebraic reasoning". In: *Proceedings of AAAI-88,* pp 264–269.