Qualitative Simulation*

Benjamin Kuipers
Department of Computer Sciences, University of Texas at Austin, Austin, TX 78712, U.S.A.

Recommended by Johan de Kleer

ABSTRACT

Qualitative simulation is a key inference process in qualitative causal reasoning. However, the precise meaning of the different proposals and their relation with differential equations is often unclear. In this paper, we present a precise definition of qualitative structure and behavior descriptions as abstractions of differential equations and continuously differentiable functions. We present a new algorithm for qualitative simulation that generalizes the best features of existing algorithms, and allows direct comparisons among alternate approaches. Starting with a set of constraints abstracted from a differential equation, we prove that the QSIM algorithm is guaranteed to produce a qualitative behavior corresponding to any solution to the original equation. We also show that any qualitative simulation algorithm will sometimes produce spurious qualitative behaviors: ones which do not correspond to any mechanism satisfying the given constraints. These observations suggest specific types of care that must be taken in designing applications of qualitative causal reasoning systems, and in constructing and validating a knowledge base of mechanism descriptions.

1. Introduction

An expert system is often a "shallow model" of its application domain, in the sense that conclusions are drawn directly from observable features of the presented situation. Researchers have long felt that genuinely expert performance must also rest on knowledge of "deep models," in which an underlying mechanism, whose state variables may not be directly observable, accounts for the observable facts [13].

One major line of research toward the representation of deep models is the study of qualitative causal models [3–20, 24, 25]. Research on qualitative causal models differs from more general work on deep models in focusing on qualitative descriptions of the deep mechanism, capable of representing incomplete knowledge of the structure and behavior of the mechanism. Symbolic

* This research was supported in part by the National Library of Medicine through NIH Grants LM 03603, LM 04125, and LM 04374, and by the National Science Foundation through grants MCS-8303640 and DCR-8417934.

Artificial Intelligence **29** (1986) 289–338
0004-3702/86/$3.50

manipulation of qualitative descriptions also appears to be a plausible model of human expertise [18, 19].

Qualitative causal reasoning consists of a number of different operations. A set of constraint equations describing the relevant structural relationships in a system may be derived by examination of its physical structure. The possible behaviors of the system may be predicted by qualitative simulation from the constraint equations and an initial state. The behavioral description may be used to explain a set of observations or the way a mechanism produces its behavior.

Researchers working in different problem domains have taken very different approaches to the derivation of constraint equations from physical structure. De Kleer and Brown [8] and Williams [25] describe a physical system in terms of components and connections. Constraint equations are derived from the component models and from the interaction paths provided by the connections. This point of view has led them to propose principles of good form, such as "no-function-in-structure," which states that component models must be formulated independently of the device contexts in which they will appear [8]. Studying naive physics reasoning about everyday physical situations, Forbus [12] determines the current set of active *processes*. The constraint equations are derived from the complete set of currently active processes. Working primarily in medical physiology, Kuipers [16, 19] treats constraint equations as given, either by textbook or experimental learning, but outside the scope of immediate causal problem solving.

The central inference within all of these approaches is qualitative simulation: derivation of a description of the behavior of a mechanism from the qualitative constraint equations. Differential equations provide a useful analogy (Fig. 1). A differential equation describes a physical system in terms of a set of state variables and constraints. The solution to the equation may be a function representing the behavior of the system over time. A description of structure in terms of constraint equations is a further abstraction of the same system, and qualitative simulation is intended to yield a corresponding abstraction of its behavior.

The goal of this paper is to clarify and formalize the qualitative mathematics behind the prediction of behavior from qualitative constraint equations. The results presented here apply across approaches to qualitative physics. The representation for constraints and behavior was originally described in [16]. The QSIM algorithm presented here replaces the previous ENV qualitative simulation algorithm. QSIM is both more efficient and more amenable to clear mathematical proof of correctness and limitations.

A theory and algorithm for qualitative reasoning must address several issues, which provides a framework for comparing the proposals of different researchers, and the contribution of this paper:

- how quantities are described qualitatively,

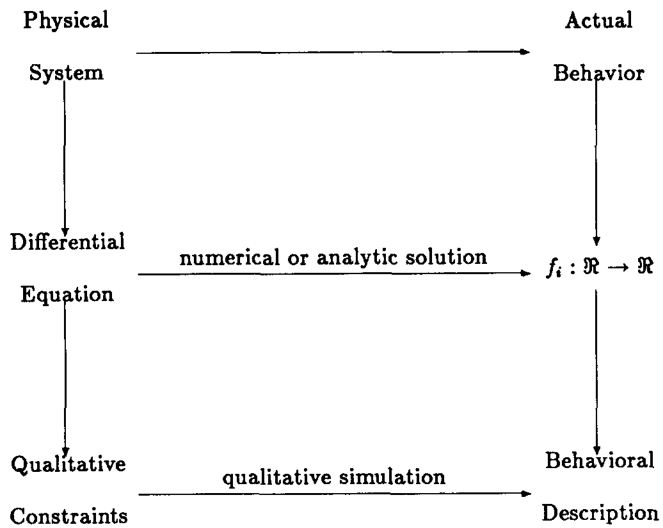

FIG. 1. Qualitative simulation and differential equations are both abstractions of actual behavior.

- how state transitions are selected,
- whether quantities correspond to standard mathematical analysis,
- whether qualitative simulation produces all and only valid behaviors.

All qualitative simulation systems describe quantities in terms of their ordinal relations with a small set of landmark values. De Kleer, Bobrow, and Brown [6, 8] and Williams [24, 25] normally take the only landmark to be zero, and thus define three *qualitative values*, $\{+, 0, -\}$. While they allow for the possibility of more complex quantity spaces, the definitions of addition and multiplication as operators over qualitative values do not extend usefully to the more complex situation, and all of their results use the "$\{+, 0, -\}$ semantics." A nonzero landmark a of a quantity x can be accommodated by defining an auxiliary quantity $y = x - a$ whose zero refers to the value $x = a$. Values defining operating region boundaries may also be used, but they are not part of the qualitative addition and multiplication operations.

Forbus [10, 12] and Kuipers [16] define a *quantity space* as a partially ordered set of landmark values, so that a quantity is described in terms of its ordinal relations with the landmarks. The Kuipers [16] approach is different from the others in allowing new landmarks to be discovered during the qualitative simulation, and used to define new qualitative distinctions. The QSIM algorithm presented here describes quantities in terms of a *linearly ordered* set of landmarks, but still allowing new landmarks to be discovered and inserted. In this paper, we demonstrate that without discovering and using new landmark values, important qualitative distinctions can be missed, such as the distinction between increasing, decreasing, and stable oscillation.

Different qualitative simulation systems take different positions on whether quantities should be an abstraction of the *standard* mathematical notion of real numbers—in which case $\mathbb{R}$ is described as an alternating sequence of points and open intervals—or whether a nonstandard model should be used, allowing two points to be infinitesimally separated. Both Forbus [10, 12] and de Kleer and Brown [8] adopt nonstandard models of time in which "mythical" or "infinitesimal" time may separate qualitative states that correspond to the same physical point in time. Such sequences of states appear to be required when the computational inference cycle must run more than once to generate a state corresponding to the next physical state. De Kleer and Bobrow [6] adopt the standard model for quantities, but appear less committed to alternating points and intervals in the time domain. Kuipers [16] and Williams [24, 25] follow the standard model. As Williams' work and this paper demonstrate, the standard model makes it possible to state and prove useful theorems about the validity of the predictions made by qualitative simulation.

All qualitative simulation systems produce the set of possible behaviors by generating and filtering the set of possible transitions from one qualitative state description to its successors. Most systems simulate forward, by generating all possible successors of the current state; de Kleer, Brown, and Bobrow [6, 8] generate all possible qualitative states, then determine the valid transitions among them. De Kleer's approach can only succeed if there is a fixed set of qualitative values, so that the set of possible states can be generated in advance. In both cases, the filtering criteria are local: they depend on the quantities in the two state descriptions, and on the structural constraints.

An important class of filtering criteria are *transition-ordering* rules [6, 16, 25]. For example, if $A + B = C$ with $A, B, C > 0$, and B and C are approaching zero, then B must reach zero before C. A large number of these rules can be formulated, corresponding to different signs, directions of approach, and combinations of quantities approaching limits. In designing a system, it is difficult to be sure that all possible such rules have been captured; in implementing it, it is difficult to check that they have been written correctly. As described in Appendix B, all of the transition-ordering rules can be recognized as special cases of a simple test of valid relationships between the current values of a set of quantities and a set of corresponding values. These tests, applying to the ADD, MULT, M^+, and M^- constraints, capture all single-constraint transition-ordering criteria of this type, can be implemented efficiently, and most importantly, can be straightforwardly proven correct.

All qualitative simulation systems predict multiple possible behaviors given certain sets of qualitative constraints and initial conditions. Researchers in this area (myself included) have hoped to prove that the predicted behaviors include all and only the possible behaviors of real mechanisms satisfying the given constraints. Half of this is correct: we prove below that qualitative simulation cannot miss any actual behavior. However, because of the local nature of its decision criteria, qualitative simulation *can* predict behaviors that are not possible for any real mechanism satisfying the given description, and we construct a counterexample. We discuss the implications of these results for the construction of a qualitative causal reasoning system.

Qualitative simulation systems vary widely in speed.[1] In order to be useful as part of an expert problem solver, a qualitative simulation system must be efficient. The QSIM algorithm is very fast. Furthermore, experiments with semantic variants (e.g. the $\{+, 0, -\}$ semantics) can be made easily by changing the entries in a table of possible state transitions. It has been implemented in LISP on the Symbolics 3600, and all examples in this paper have been run, as well as numerous others in elementary physics and in medical physiology [20].

1.1. Overview

This section provides an overview of qualitative simulation and the QSIM algorithm. The concepts presented here are defined more formally below.

Qualitative simulation of a system starts with a description of the known structure of the system, and an initial state, and produces a directed graph consisting of the possible future states of the system and the "immediate successor" relation between states. The possible behaviors of the system are the paths from the initial state through the graph. After defining terminology, the next section discusses the constraints and behavior describing a simple mechanism in both informal and formal terms.

The structure of a system is described by a set of symbols representing the *physical parameters* of the system (continuously differentiable real-valued functions), and a set of *constraint equations* describing how those parameters may be related to each other. The constraints are two- or three-place relations on physical parameters. Some specify familiar mathematical relationships: DERIV(vel,acc), ADD(net,out,in), MULT(mass,acc,force), MINUS(fwd,rev). Others assert qualitatively that there is a functional relationship between two physical parameters, but only specify that the relationship is monotonically increasing or decreasing: M^+(price,power) and M^-(mph,mpg). The constraints are designed to permit a large class of differential equations to be mapped straightforwardly into qualitative constraint equations.

Each physical parameter is a continuously differentiable real-valued function of time. Its value at any given point in time is specified qualitatively in terms of its relationship with a totally ordered set of *landmark values*. The landmark values may be either numerical (e.g. zero) or symbolic; their ordinal relationships are their essential properties. As the qualitative simulation proceeds, it can discover and add new landmark values to the sequence. The *qualitative*

[1] De Kleer, personal communication; Forbus, personal communication.

state of a parameter consists of its ordinal relations with the landmark values and its direction of change.

Time, similarly, is represented as a totally ordered set of symbolic *distinguished time-points*. The current time is either at or between distinguished time-points. All of the time-points are generated as a result of the qualitative simulation process.

At a distinguished time-point, if several physical parameters linked by a single constraint are equal to landmark values, they are said to have *corresponding values* which can be discovered and used by the qualitative simulation. The special case of a monotonic function constraint with corresponding values $(0, 0)$ is sufficiently common that it is signified by the constraints M_0^+ and M_0^-.

A set of constraints on the physical parameters of the system is only valid in some *operating region*, defined by the legal ranges of values that some parameters may take on. The *legal range* of a parameter is a closed interval whose endpoints are landmark values of that parameter. These endpoints may be associated with transitions to other operating regions where a different set of constraints apply.

The *initial state* of the system is defined by the operating region and a set of qualitative values for the physical parameters. The qualitative simulation proceeds by determining all of the possible changes in qualitative value permitted to each parameter, then filtering the combinations by applying progressively broader constraints. If more than one qualitative change is possible, the current state has multiple successors, and the simulation branches.

Two qualitative states in the same operating region are *identical* if all parameters are equal to the same landmark values, and all the directions of change are the same. If one of the successors to a given state is identical to a direct predecessor, a cyclic behavior can be created, resulting in a graph of states.

1.2. Example: The U-tube

A U-tube, consisting of two partially filled tanks connected at the bottom by a thin tube, is in equilibrium when an increment of water is added to one side (Fig. 2). The system reaches a new equilibrium with the level in both tanks higher than before.

Constraints

- Each tank has a pressure which depends on its amount of water.
- The rate of flow through the tube depends on the difference between the pressures.
- A flow through the tube increases the amount in one tank and decreases the other.

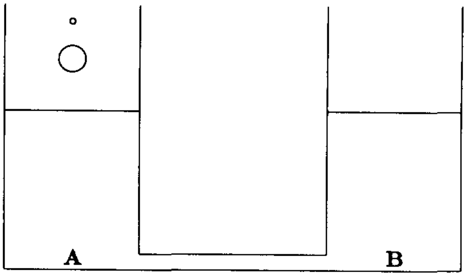

FIG. 2. The U-tube in equilibrium receives an increment of water.

Behavior

- After the increment of water, the amount and pressure are increased in tank A, leading to a flow from A to B.
- Water flows from A to B. The level in A falls, the level in B rises, and the pressure difference and rate of flow approach zero.
- The pressure difference and rate of flow become zero as the U-tube reaches equilibrium with the level in both tanks higher than before the increment.

Qualitative simulation determines the essentially different regions of the system's behavior. It need not be given the initial levels nor the amount of the increment. It does not determine how high the level in A is increased, what its final position is, or how long the equilibration process requires. It does guarantee, however, that the level in A falls, the level in B rises, that neither tank returns to its initial level, and that a new equilibrium is reached. The parameters and constraints describing the structure of the U-tube are shown graphically in Fig. 3. The qualitative behaviors of the six parameters, responding to an increment of water, is shown by the six "qualitative plots" in Fig 4.

A *qualitative plot* graphically describes the qualitative behavior description of a parameter. The vertical axis represents the set of landmark values for that parameter; the only meaningful vertical positions are at, or midway between, landmark values. The horizontal axis is slightly more complex. Known states are shown first for reference, followed by the sequence of time-points. To reduce visual clutter in the plot, time-points are not labeled, though distinguished time-points are indicated by tick-marks on the axis, and other time-points are plotted midway between adjacent ticks. In Fig. 4, the four positions on the horizontal axis represent the known state NORMAL and the three time-points t_0, (t_0, t_1), and t_1. Although the plot of level(A) is horizontal from t_0 to (t_0, t_1), its value is not constant. Rather, the qualitative state description of level(A) remains the same (i.e. between two landmarks and decreasing) while the underlying real value changes. The graphical conventions for qualitative behavior plots are somewhat unfamiliar, but the graphical output improves the comprehensibility of the output of qualitative simulation, and thus facilitates development and debugging of sets of constraints.

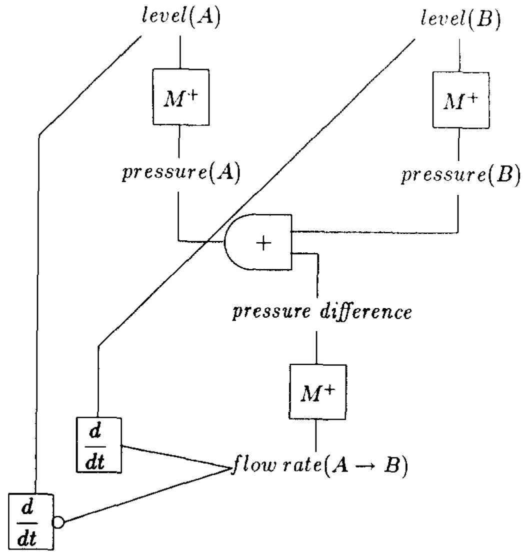

FIG. 3. The constraints describing the structure of the U-tube.

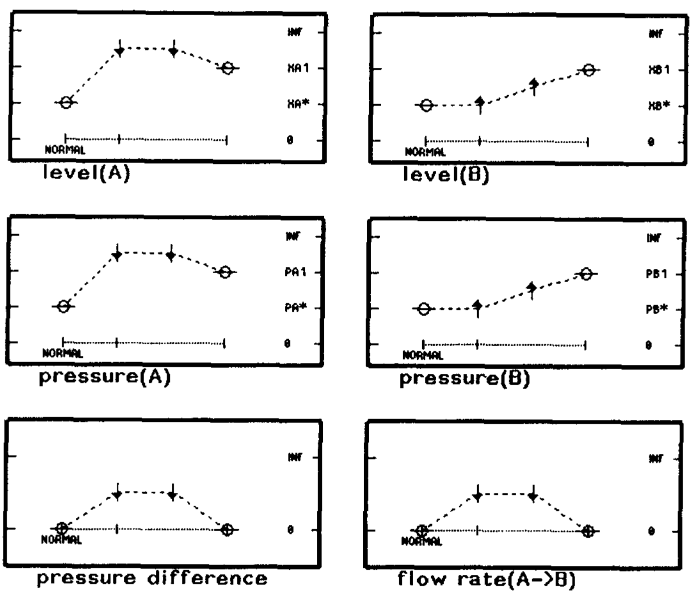

FIG. 4. The behavior of the U-tube given an increment in tank A.

2. Qualitative Behavior

In the following sections, we present a more rigorous definition of qualitative simulation, leading up to a definition of the QSIM algorithm and the proof of several theorems characterizing its strengths and limitations. The validity proofs for several steps of the algorithm are contained in the appendices.

A physical system is characterized by a number of real-valued parameters, which vary continuously over time. We consider each physical parameter to be a function $f:[a, b] \rightarrow \mathbb{R}^*$, where $\mathbb{R}^* = [-\infty, \infty]$, the extended real number line. The domain and range of a function f are both closed intervals in the extended reals, $\mathbb{R}^*$. We use $\mathbb{R}^*$ instead of $\mathbb{R}$, treating ∞ as a genuine landmark value, because it is useful (though not essential) to have the invariant that t and $f(t)$ are always bounded by explicitly stated landmark values in the domain and range of f. The function $f:[0, \infty] \rightarrow \mathbb{R}^*$ is defined to be continuous at ∞ exactly if $\lim_{t \rightarrow \infty} f(t)$ exists. For example, both e^{-t} and e^{t} are continuously differentiable on $[0, \infty]$, but $\sin t$ is not. This allows us to express asymptotic approach as a move to a limit, where the limit is reached at $t = \infty$.

2.1. Behavior of a single function

We will define the qualitative behavior description first for a single, continuously differentiable function $f:[a, b] \rightarrow \mathbb{R}^*$.

Definition 2.1. For $[a, b] \subseteq \mathbb{R}^*$, define $f:[a, b] \rightarrow \mathbb{R}^*$ to be a *reasonable function* if

(1) f is continuous on $[a, b]$,
(2) f is continuously differentiable on (a, b),
(3) f has only finitely many critical points in any bounded interval,
(4) $\lim_{t \downarrow a} f'(t)$ and $\lim_{t \uparrow b} f'(t)$ exist in $\mathbb{R}^*$; define $f'(a)$ and $f'(b)$ to be equal to these limits.

The restriction to finitely many critical points in any bounded interval excludes examples like $f(t) = t^2 \sin 1/t$ that are continuously differentiable, but whose behavior changes infinitely quickly around $t = 0$. Without the fourth restriction, $f'(t)$ can still behave pathologically around the endpoints of the interval, even without crossing zero.

Definition 2.2. Every reasonable function $f:[a, b] \rightarrow \mathbb{R}^*$ has associated with it a finite set of *landmark values*. The landmark values must include 0, $f(a)$, $f(b)$, and the value of $f(t)$ at each of its critical points, and may include any number of additional values.

Definition 2.3. Where f is a reasonable function, $t \in [a, b]$ is a *distinguished time-point* of f if t is a boundary element of the set $\{t \in [a, b] \mid f(t) = x$, where x is a landmark value of $f\}$.

That is, the distinguished time-points are those points where something important happens to the value of f, such as passing a landmark value or reaching an extremum. The restriction to boundary elements handles the case where f becomes constant over an interval: only the endpoints of the interval are distinguished time-points. De Kleer and Bobrow [6] eliminate this case by assuming that parameters have derivatives of all orders, in which case any function which is constant over an interval is constant everywhere.

All functions mentioned in the rest of this paper should be presumed reasonable unless specified otherwise. A reasonable function $f:[a,b]\to \mathbb{R}^*$ has the finite set of distinguished time-points:

$$a = t_0 < t_1 < \cdots < t_n = b\ ,$$

and the finite set of landmark values:

$$l_1 < l_2 < \cdots < l_k\ .$$

We can now define the qualitative state of f at t in terms of its ordinal relations with its landmarks, and its direction of change.

We reluctantly contribute to the proliferation of notations for qualitative description of continuous functions. The advantages of the notation used here are that it (1) naturally allows for an arbitrary and changing set of landmark values, (2) uses a single term for the qualitative description of a function's magnitude and derivative, and (3) emphasizes that the qualitative description of the derivative is of low and fixed resolution, while qualitative description of magnitude is of higher and possibly changing resolution.

Definition 2.4. Let $l_1 < \cdots < l_k$ be the landmark values of $f:[a,b]\to \mathbb{R}^*$. For any $t \in [a,b]$, QS(f, t), the *qualitative state of f at t*, is a pair $\langle$qval,qdir$\rangle$, defined as follows:

(1)

$$\text{qval} = \begin{cases} l_j, & \text{if } f(t) = l_j \text{, a landmark value,} \\ (l_j, l_{j+1}), & \text{if } f(t) \in (l_j, l_{j+1}); \end{cases}$$

(2)

$$\text{qdir} = \begin{cases} \text{inc}, & \text{if } f'(t) > 0\ , \\ \text{std}, & \text{if } f'(t) = 0\ , \\ \text{dec}, & \text{if } f'(t) < 0\ . \end{cases}$$

For example, QS(water-temp, now) = $\langle$(32°F, 212°F), inc$\rangle$.

Proposition 2.5. *Where $a = t_0 < \cdots < t_n = b$ are the distinguished time-points of f, consider $s, t \in (a, b)$ such that $t_i < s < t < t_{i+1}$ for some i. Then* $\mathrm{QS}(f, s) = \mathrm{QS}(f, t)$.

Proof. By the Intermediate Value Theorem, since f is continuously differentiable, $f(t)$ cannot pass a landmark value, and $f'(t)$ cannot change signs between adjacent distinguished time-points. □

This justifies our basic intuition that the qualitative state of the function is constant over intervals between landmarks. Hence, we may make the following definitions.

Definition 2.6. For adjacent distinguished time-points t_i and t_{i+1}, define $\mathrm{QS}(f, t_i, t_{i+1})$, the *qualitative state of f on (t_i, t_{i+1})*, to be $\mathrm{QS}(f, t)$ for any $t \in (t_i, t_{i+1})$.

Definition 2.7. The *qualitative behavior* of f on $[a, b]$ is the sequence of qualitative states of f:

$$\mathrm{QS}(f, t_0), \mathrm{QS}(f, t_0, t_1), \mathrm{QS}(f, t_1), \ldots, \mathrm{QS}(f, t_{n-1}, t_n), \mathrm{QS}(f, t_n)$$

alternating between qualitative states at distinguished time-points, and qualitative states on intervals between distinguished time-points.

2.2. Systems of functions

Definition 2.8. A *system* is a set $F = \{f_1, \ldots, f_m\}$ of reasonable functions $f_i : [a, b] \to \mathbb{R}^*$, each with its own set of landmarks and distinguished time-points. The *distinguished time-points of a system F* are the union of the distinguished time-points of the individual functions $f_i \in F$. The *qualitative state* of a system F of m functions is the m-tuple of individual qualitative states:

$$\mathrm{QS}(F, t_i) = [\mathrm{QS}(f_1, t_i), \ldots, \mathrm{QS}(f_m, t_i)],$$
$$\mathrm{QS}(F, t_i, t_{i+1}) = [\mathrm{QS}(f_1, t_i, t_{i+1}), \ldots, \mathrm{QS}(f_m, t_i, t_{i+1})].$$

If t_i, and/or t_{i+1} are not distinguished time-points of a particular f_j, then t_i and the interval (t_i, t_{i+1}) must be between two distinguished time-points of f_j, say t_k and t_{k+1}. Then $\mathrm{QS}(f_j, t_i)$ and $\mathrm{QS}(f_j, t_i, t_{i+1})$ are defined to be the same as the containing $\mathrm{QS}(f_j, t_k, t_{k+1})$. The *qualitative behavior* of F is the sequence of qualitative states of F:

$$\mathrm{QS}(F, t_0), \mathrm{QS}(F, t_0, t_1), \mathrm{QS}(F, t_1), \ldots, \mathrm{QS}(F, t_n).$$

These definitions give us a precise semantics for the qualitative description of continuous functions, and clarifies the concept of the "next state." Every state has a qualitative description QS(F, t), but that description changes only at discrete distinguished time-points, and remains constant on the open intervals between them. Thus the "next state" of a mechanism is more properly called the next distinct qualitative state description of the mechanism.

2.3. Qualitative state transitions

Since a reasonable function f is continuously differentiable, the Intermediate Value Theorem and the Mean Value Theorem restrict the way it can change from one qualitative state to the next. There are two types of qualitative state transitions: *P-transitions*, moving from a point to an interval, and *I-transitions*, moving from an interval to a point.

Definition 2.9. Where t_i, is a distinguished time-point, a *P-transition* of f is a pair of adjacent qualitative states of f,

$$\mathrm{QS}(f, t_i) \Rightarrow \mathrm{QS}(f, t_i, t_{i+1}),$$

whose first state is the qualitative state at a distinguished time-point. An *I-transition* is a pair of adjacent qualitative states of f,

$$\mathrm{QS}(f, t_{i-1}, t_i) \Rightarrow \mathrm{QS}(f, t_i),$$

TABLE 1. The possible transitions
(A reasonable function $f : [a, b] \to \mathbb{R}^*$ is restricted to the following set of possible transitions from one qualitative state to the next. The contents of this table are justified by Propositions A.1, A.2, A.4, and A.5)

P-transitions	QS(f, t_i)	⇒QS(f, t_i, t_{i+1})	I-transitions	QS(f, t_i, t_{i+1})	⇒QS(f, t_{i+1})
P1	$\langle l_j, \text{std}\rangle$	$\langle l_j, \text{std}\rangle$	I1	$\langle l_j, \text{std}\rangle$	$\langle l_j, \text{std}\rangle$
P2	$\langle l_j, \text{std}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	I2	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l_{j+1}, \text{std}\rangle$
P3	$\langle l_j, \text{std}\rangle$	$\langle (l_{j-1}, l_j), \text{dec}\rangle$	I3	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l_{j+1}, \text{inc}\rangle$
P4	$\langle l_j, \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	I4	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$
P5	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	I5	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l_j, \text{std}\rangle$
P6	$\langle l_j, \text{dec}\rangle$	$\langle (l_{j-1}, l_j\rangle, \text{dec}\rangle$	I6	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l_j, \text{dec}\rangle$
P7	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	I7	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{dec}\rangle$
			I8	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l^*, \text{std}\rangle$
			I9	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l^*, \text{std}\rangle$

In cases I8 and I9, f becomes std at l^*, a new landmark value such that $l_j < l^* < l_{j+1}$. In these cases, a previously unknown landmark value is discovered because other constraints force $f'(t)$ to become zero.

whose first state is the qualitative state on the interval between distinguished time-points.

Table 1 specifies the set of possible transitions that can take place in the qualitative behavior of a single function. The validity of this table is proved by the propositions in Appendix A.

3. Qualitative Structure

The structure of a mechanism may be described by a set of qualitative constraint equations applied to the parameters that represent the state of the mechanism. Simulation attempts to assign behaviors to the parameters. Constraints holding between parameters in the structural description serve to limit the possible combinations of qualitative behavior. The constraint notation used here has the advantage, like de Kleer's confluences, of having a clear correspondence with differential equations by making explicit all the functions and operators in the equation.

Constraints are expressed as predicates rather than as functions for two reasons. First, they will be *used* as predicates in the QSIM algorithm to test the consistency of sets of qualitative values. Second, if a constraint were to be treated as a function, it is unclear how to define precisely the function's range. On the other hand, while keeping these semantic considerations in mind, the reader will probably find it clearer to read constraints as functions: mpg = M^-(mph) rather than M^-(mph,mpg).

3.1. Arithmetic constraints

Constraints corresponding to the basic arithmetic and differential operators are fundamental to a structural description.

Definition 3.1. ADD(f, g, h) is a three-place predicate on reasonable functions $f, g, h : [a, b] \rightarrow \mathbb{R}^*$ which holds iff $f(t) + g(t) = h(t)$ for every $t \in [a, b]$.

Definition 3.2. MULT(f, g, h) is a three-place predicate on reasonable functions $f, g, h : [a, b] \rightarrow \mathbb{R}^*$ which holds iff $f(t) \cdot g(t) = h(t)$ for every $t \in [a, b]$.

Definition 3.3. MINUS(f, g) is a two-place predicate on reasonable functions $f, g : [a, b] \rightarrow \mathbb{R}^*$ which holds iff $f(t) = -g(t)$ for every $t \in [a, b]$.

Since addition and multiplication are commutative,

$$\text{ADD}(f, g, h) \Leftrightarrow \text{ADD}(g, f, h),$$
$$\text{MULT}(f, g, h) \Leftrightarrow \text{MULT}(g, f, h),$$
$$\text{MINUS}(f, g) \Leftrightarrow \text{MINUS}(g, f).$$

Definition 3.4. DERIV(f, g) is a two-place predicate on reasonable functions $f, g:[a, b] \rightarrow \mathbb{R}^*$ which holds iff $f'(t) = g(t)$ for every $t \in [a, b]$.

3.2. Qualitative function constraints

In describing the qualitative structure of a mechanism, one might need to state that one physical parameter is a function of another, without specifying the function completely. Rather, the relationship should be described qualitatively in terms of regions of monotonic increase or decrease, and landmark values passed through.

The most common and important cases are functional relationships that are strictly monotonic everywhere. The monotonic function constraint M^+ applies in the situation when the function is strictly monotonically increasing, and M^- when it is decreasing. In fact, the definition is slightly more restrictive: the derivative of the function must be nonzero, except possibly at the endpoints of the domain.

Definition 3.5. M^+ is a two-place predicate on reasonable functions $f, g:[a, b] \rightarrow \mathbb{R}^*$. $M^+(f, g)$ is true iff $f(t) = H(g(t))$ for all $t \in [a, b]$, where H is a function with domain $g([a, b])$ and range $f([a, b])$, differentiable and with $H'(x) > 0$ for all x in the interior of the domain. M^- is defined similarly, except that $H'(x) < 0$.

The restrictions on H are motivated by two requirements. First, the critical points of f and g must match across an $M^+(f, g)$ constraint. Second, it must be possible to break a function such as $\sin x$ "at the joints" into regions of monotonic increase and decrease, so $H'(x) = 0$ must be allowed at the boundary of the domain.

Clearly, $M^+(f, g) \Leftrightarrow M^+(g, f)$, and $M^-(f, g) \Leftrightarrow M^-(g, f)$.

Note that $M^+(f, g)$ does not imply that f and g are monotonic functions on $[a, b]$. For example, $M^+(2 \sin t, \sin t)$ holds on $[0, 2\pi]$, where $H(x) = 2x$.

Proposition 3.6. *Consider two continuously differentiable functions* f, g: $[a, b] \rightarrow \mathbb{R}^*$, *where* $M^+(f, g)$. *Then for all* $t \in (a, b)$,

$$\begin{aligned} f'(t) > 0 \quad &\textit{iff} \quad g'(t) > 0 \,, \\ f'(t) = 0 \quad &\textit{iff} \quad g'(t) = 0 \,, \\ f'(t) < 0 \quad &\textit{iff} \quad g'(t) < 0 \,. \end{aligned}$$

Proof. $M^+(f, g)$ means that $f(t) = H(g(t))$, so $f'(t) = H'(g(t)) \cdot g'(t)$. Since $H'(x) > 0$, $g'(t) = 0$ if and only if $f'(t) = 0$. The two strict inequalities follow from the monotonicity of H. □

Thus, the sets of distinguished time-points may not correspond precisely across f and g, but the subset of distinguished time-points which are critical points, and hence the regions of constant directions-of-change, are identical.

Although constraints for expressing monotonic relationships are used by all researchers on qualitative simulation, the precise definition of the constraint varies significantly. Kuipers [16] expresses an increasing monotonic relationship between X and Y as $Y = M^+(X)$ or $M^+(X, Y)$, meaning that there is some well-defined but unspecified function f with the desired properties, such that $Y = f(X)$. Forbus [12] uses the notation $X \propto_{Q+} Y$, meaning that $Y = f(\ldots, X, \ldots)$ where the dependence of Y on X is monotonically increasing, *all else held equal*. The reason for this is that constraints are associated with processes, and only when the complete process configuration is known can the set of constraints be closed and the final dependencies computed. Thus, these qualitative proportionality constraints are implicitly added once the complete set is known. De Kleer, Brown, and Bobrow [6, 8] express the same relationship with the confluence $\partial X = \partial Y$, meaning $\operatorname{sgn} dX/dt = \operatorname{sgn} dY/dt$. Strictly speaking, $X(t)$ and $Y(t)$ need not have any functional relationship at all, as long as the signs of their derivatives remain identical. The implications of these different definitions have not yet been fully explored. However, the correctness proof for the QSIM algorithm presented in this paper relies on the strong definition of the monotonic constraints M^+ and M^-.

A qualitative functional relationship need not be strictly monotonic if it can be divided into sections that are alternately increasing or decreasing monotonic, with critical points at the joints between sections. For example, suppose that $x = \cos\theta$ for $x \in [-1, 1]$ and $\theta \in [0, 2\pi]$. We may say that $\mathrm{FC}(\theta, x, \text{descrip})$, where

$$\text{descrip} = \begin{cases} (0, 1) \\ ((0, \pi), (-1, 1), M^-) \\ (\pi, -1) \\ ((\pi, 2\pi), (-1, 1), M^+) \\ (2\pi, 1) \end{cases}$$

That is, if $\theta = 0$, $x = 1$; when $\theta \in (0, \pi)$, then $x \in (-1, 1)$ and $M^-(\theta, x)$; and so on. At the joints between monotonic sections, the restrictions on permissible combinations of directions of change are weakened. In particular, it is possible for one parameter to have direction of change std while the other is inc or dec. Between the joints, qualitative simulation can treat an FC constraint exactly like the specified M^+ or M^-.

In a similar spirit, we may define S^+ and S^- constraints, which behave like monotonic function constraints in the interval between two sets of corresponding values, and leave one parameter constant while the other is unconstrained outside that interval. It is not difficult to imagine other extensions such as "log-like," "polynomial-like," and "exponential-like" monotonic function constraints which support inferences about relative asymptotic magnitudes. These topics are beyond the scope of this paper.

3.3. Qualitative differential equations

The constraint definitions now allow us to define precisely the abstraction relation between qualitative constraint equations and ordinary differential equations (ODEs). If a mechanism can be described by an ODE meeting certain restrictions there is a corresponding but weaker set of qualitative constraint equations for the same mechanism. "Weaker" in this context means that any behavior that satisfies the ODE must satisfy the constraints, but not necessarily vice versa. Thus the constraints constitute a form of qualitative differential equation.

Starting with a suitable ODE, we can decompose it into an equivalent set of simultaneous equations by introducing terms for each subexpression of the original ODE. When this process is complete, each equation can be mapped to a qualitative constraint. For example, consider the ODE:

$$\frac{d^2u}{dt^2} - \frac{du}{dt} + \arctan ku = 0\,. \tag{1}$$

The simultaneous equations (a) and the qualitative constraints (b) are derived as follows:

(a)	(b)
$f_1 = du/dt$,	DERIV(u, f_1) ,
$f_2 = df_1/dt$,	DERIV(f_1, f_2) ,
$f_3 = ku$,	MULT(k, u, f_3) ,
$f_4 = \arctan f_3$,	$M^+(f_3, f_4)$,
$f_2 - f_1 + f_4 = 0$,	ADD(f_2, f_4, f_1) .

Any solution $u(t)$ to equation (1) uniquely determines the auxiliary functions $f_1, \ldots, f_4$, and so defines a solution to the simultaneous equations (a). Each constraint in (b) is mathematically equivalent to the corresponding equation, with the exception of the M^+, which is less restrictive. Thus the solution to the ODE (1) must also satisfy the constraints (b). The introduction of the monotonic function constraint, of course, requires that the corresponding function from the ODE have nonzero derivative, except possibly at the endpoints of the domain.

We may summarize this discussion as the following theorem.

Theorem 3.7. *Let*

$$F[u(t), u'(t), \ldots, u^{(n)}(t)] = 0 \tag{2}$$

be an ordinary differential equation of order n, to be satisfied by a function $u:[a, b] \to \mathbb{R}$, where F is defined only in terms of the arithmetic operations addition, multiplication, and negation, along with functions of continuous and strictly nonzero derivative. Then a set of parameters and constraints can be defined, corresponding with (2), *such that any reasonable function $u: \mathbb{R} \to \mathbb{R}$ which satisfies* (2) *also satisfies the set of constraints.*

The procedure for decomposing an ODE into simultaneous equations can easily be specified so that each ODE generates a unique set of constraints. However, different functions may be mapped to the same M^+ constraint so a given qualitative differential equation can be the abstraction of multiple ODEs.

4. Qualitative Simulation

This section describes the QSIM qualitative simulation algorithm, and refers to the proofs of the various steps, appearing in the appendices.

4.1. Input and output

The qualitative simulation algorithm is given the following description of a mechanism.

(1) A set $\{f_1, \ldots, f_m\}$ of symbols representing the functions in the system.

(2) A set of constraints applied to the function symbols: $M^+(f, g)$, $M^-(f, g)$, ADD(f, g, h), MULT(f, g, h), MINUS(f, g), or DERIV(f, g). Each constraint may have associated corresponding values for its functions.

(3) Each function is associated with a totally ordered set of symbols representing landmark values; each function has at least the basic set of landmarks $\{-\infty, 0, \infty\}$.

(4) Each function may have upper and lower range limits, which are landmark values beyond which the current set of constraints no longer apply. A range limit may be associated with a new operating region which has its own constraints and range limits.

(5) An initial time-point symbol, t_0, and qualitative values for each of the f_i at t_0 are given.

The result of the qualitative simulation is one or more qualitative behavior descriptions for the function symbols given. Each qualitative behavior description consists of the following:

(1) A sequence $\{t_0, \ldots, t_n\}$ of symbols represents the distinguished time-points of the system's behavior.

(2) Each function f_i has a totally ordered set of landmark values, possibly extending the originally given set.

(3) Each function has at each distinguished time-point or interval between adjacent time-points, a qualitative state description expressed in terms of the landmark values of that function.

4.2. The algorithm QSIM

The qualitative simulation algorithm, QSIM, repeatedly takes an active state and generates all possible successor states, filtering out states that violate some consistency criterion. Because it may not be able to determine the next state uniquely, QSIM builds a tree of states representing the possible behaviors of the mechanism.

Place the initial state on the list ACTIVE of states whose successors need to be determined. Repeat the following steps until ACTIVE becomes empty or a resource limit is exceeded.

Step 1. Select a qualitative state from ACTIVE.

Step 2. For each function, determine (from Table 1) the set of transitions possible from the current qualitative state.

Step 3. For each constraint, generate the set of tuples (pairs or triples) of transitions of its arguments. Filter for consistency with that constraint.

Step 4. Perform pairwise consistency filtering on the sets of tuples associated with the constraints in the system, applying the consistency criterion that adjacent constraints must agree on the transition assigned to the shared parameter.

Step 5. Generate all possible global interpretations from the remaining tuples. If there are none, mark the behavior as inconsistent. Create new qualitative states resulting from each interpretation, and make them successors of the current state.

Step 6. Apply global filtering rules to the new qualitative states, and place any remaining states on ACTIVE.

After an example, the individual steps of the algorithm are discussed in detail.

4.3. Example: The Ball system

To illustrate one cycle of the QSIM algorithm, consider a very simple system consisting of a ball thrown upward in a constant gravitational field. This section will demonstrate the derivation from its predecessor of the third qualitative state ($t = t_1$), where the ball reaches its maximum height. The constraints are:

$$\text{DERIV}(Y, V)\,, \qquad \text{DERIV}(V, A)\,, \qquad A(t) = g < 0\,.$$

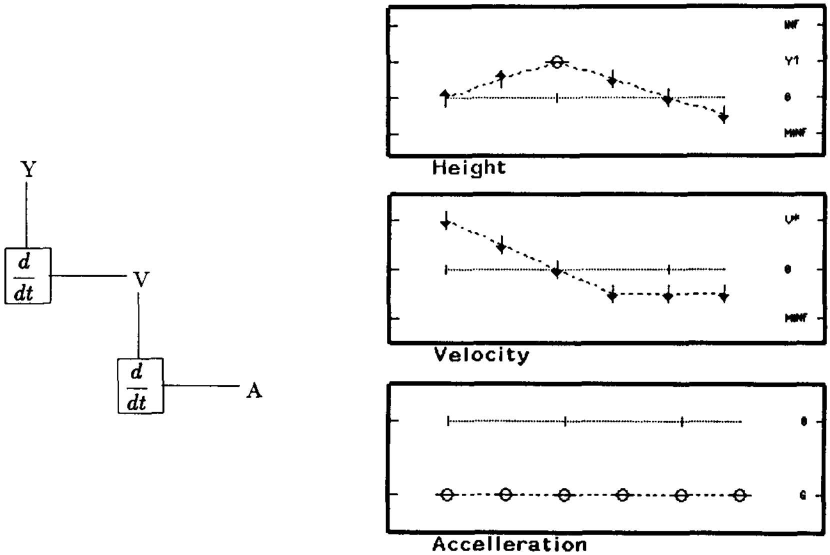

FIG. 5. The ball thrown upwards: constraints and behavior.

Figure 5 shows a graphical representation of the constraints and a qualitative plot of the behavior of $Y(t)$.

We start with an active state, $t = (t_0, t_1)$, whose description is:

$$\begin{aligned} QS(A, t_0, t_1) &= \langle g, \text{std} \rangle , \\ QS(V, t_0, t_1) &= \langle (0, \infty), \text{dec} \rangle , \\ QS(Y, t_0, t_1) &= \langle (0, \infty), \text{inc} \rangle . \end{aligned}$$

For each function, retrieve from Table 1 the set of possible qualitative state transitions from the current state of that function. Since the current state represents the time-interval (t_0, t_1), only I-transitions are applicable. For simplicity, we exclude the possibility that $Y(t_1) = \infty$, and so exclude transitions I2 and I3 from Y's list. Even without this assumption, the methods in Appendix A.2 would exclude these behaviors.

$$\begin{array}{lll} A & \text{I1:} & \langle g, \text{std} \rangle \Rightarrow \langle g, \text{std} \rangle ; \\[6pt] V & \text{I5:} & \langle (0,\infty), \text{dec} \rangle \Rightarrow \langle 0, \text{std} \rangle , \\ & \text{I6:} & \langle (0,\infty), \text{dec} \rangle \Rightarrow \langle 0, \text{dec} \rangle , \\ & \text{I7:} & \langle (0,\infty), \text{dec} \rangle \Rightarrow \langle (0,\infty), \text{dec} \rangle , \\ & \text{I9:} & \langle (0,\infty), \text{dec} \rangle \Rightarrow \langle L^*, \text{std} \rangle ; \\[6pt] Y & \text{I4:} & \langle (0,\infty), \text{inc} \rangle \Rightarrow \langle (0,\infty), \text{inc} \rangle , \\ & \text{I8:} & \langle (0,\infty), \text{inc} \rangle \Rightarrow \langle L^*, \text{std} \rangle . \end{array}$$

Next, each constraint forms a set of transition tuples. Those marked with c are eliminated by constraint consistency filtering. For example, tuple (I4, I5) would require Y to continue to increase while $V = 0$, an obvious inconsistency. Then those marked with w are eliminated by pairwise consistency filtering. The tuple (I4, I9) finds no remaining tuple associated with DERIV(V, A) that can agree that V might take transition I9.

DERIV(Y, V)		DERIV(V, A)	
(I4, I5)	c	(I5, I1)	c
(I4, I6)	c	(I6, I1)	
(I4, I7)		(I7, I1)	
(I4, I9)	w	(I9, I1)	c
(I8, I5)	w		
(I8, I6)			
(I8, I7)	c		
(I8, I9)	c		

The remaining tuples can be formed into the following two global interpretations:

Y	V	A
I4	I7	I1
I8	I6	I1

The first of these interpretations would leave states (t_0, t_1), t_1 and (t_1, t_2) with identical qualitative state descriptions, and the problem of determining the state at t_2 would be precisely what we have just done. This possibility is already adequately described by the qualitative state over (t_0, t_1), so we need not generate a successor state for it.

Every time-interval must have an endpoint (see Appendix A.2), so the only remaining possibility becomes the unique successor, defining a new landmark value for Y.

$$\mathrm{QS}(A, t_1) = \langle g, \mathrm{std} \rangle ,$$
$$\mathrm{QS}(V, t_1) = \langle 0, \mathrm{dec} \rangle ,$$
$$\mathrm{QS}(Y, t_1) = \langle Y_{\mathrm{max}}, \mathrm{inc} \rangle ,$$

The following sections explain the steps of the QSIM algorithm in detail and discuss the proofs of their validity.

4.4. Function consistency

The possible transitions that a single parameter can take from one qualitative state to the next are given in Table 1. In Step 2, the current state of each function is used to retrieve the set of applicable transition patterns from Table 1. Constraints between neighboring functions are not considered until Step 3. Transitions are also checked against invariant assertions at this stage, to eliminate impossible transitions for functions that are (e.g.) always finite or never negative.

For any particular qualitative state, Table 1 provides at most 4 possible transitions. Thus, if there are n functions in the system, the possible next states are to be found within a product space of at most 4^n points. At this stage, however, we do not explicitly generate this product space, so we need create at most $4n$ individual transitions.

Appendix A presents the proofs that justify the possible transitions given in Table 1. It also discusses the handling of divergence to ∞ and asymptotic approach to limiting values.

4.5. Constraint consistency

Step 3 of the QSIM algorithm aggregates the individual transitions into 2-tuples and 3-tuples corresponding to the arguments of individual constraints. These tuples can then be checked for consistency according to two criteria local to individual constraints (see Appendix B).

(1) The directions-of-change tuple must be consistent with the constraint in the state resulting from the transition.

(2) The result of the transition tuple can be compared with corresponding values of the arguments to that constraint.

Definition 4.1. Landmark values p and q are *corresponding values* of f and g if there is some $t \in [a, b]$ such that $f(t) = p$ and $g(t) = q$.

$M_0^+(f, g)$ and $M_0^-(f, g)$ are abbreviations for $M^+(f, g)$ and $M^-(f, g)$, respectively, with corresponding values $(0, 0)$.

Definition 4.2. Suppose $QS(f, t_i, t_{i+1}) = \langle (l_k, l_{k+1}), \text{inc} \rangle$. Then l_{k+1} is the *limit of f during* (t_i, t_{i+1}). If $f(t_{i+1}) = l_{k+1}$, we say that f has moved to its limit. Otherwise, $f(t_{i+1}) < l_{k+1}$, and we say that f has moved toward, but not reached, its limit. Similarly if f is decreasing during $QS(f, t_i, t_{i+1})$.

Informally, if f is between landmark values but moving toward a limit, it may or may not reach that limit by the next distinguished time-point. If several functions are moving toward limits, constraints between functions limit the space of possibilities. For example, if $M^+(f, g)$ is true, and f and g are moving toward corresponding limit values, then either both will reach their limits, or neither will. Similarly, if $f + g = h$, and two functions are moving toward corresponding limits, while one is bounded away from the third corresponding value, the possible transitions can be filtered.

These constraint-based consistency criteria generalize the *transition-ordering* rules of Williams [24, 26] to quantity spaces which may contain nonzero landmark values, and where sets of corresponding landmark values play a significant role. Appendix B discusses the generalization to quantity spaces with nonzero landmarks, and justifies the comparison of proposed transition tuples with known corresponding values.

4.6. Pairwise consistency filtering

Two constraints are *adjacent* if they share an argument. At this point, each constraint has an associated set of transition tuples, consistent with that individual constraint. A tuple is a proposed assignment of transitions to the functions in that constraint. To be pairwise consistent, tuples on adjacent constraints must assign the same transition to the function they share. For certain tuples, there may be no opposite number to make such a consistent pair. If so, that tuple may be deleted.

Waltz [23] developed this local consistency filtering algorithm to converge quickly on a small set of possible labelings for a graph representing the edges, vertices, and regions of a visual scene. (Mackworth and Freuder [21, 22] present this and a related class of algorithms, and assess their relative complexities.) A key step in the development of the QSIM algorithm was the observation that if *transitions*, rather than *qualitative states* are taken as the analog of edge labels, the Waltz algorithm could be applied directly.

Filtering on transitions rather than states simplifies several steps of the algorithm. The possibility of creating new landmarks can be considered without actually creating landmarks that might have to be retracted. The pairwise and global consistency filtering can match atomic transition names rather than much more expensive structure-matching on the predicted next state. Finally, some of the global filters (Section 4.8) depend on the sequence of transitions leading up to a proposed state, and would be more difficult to express in terms of state descriptions.

The Waltz algorithm visits each constraint in turn, looking at all the adjacent constraints and the function joining the pair. It applies the following rule to each transition tuple associated with the constraint it is visiting.

> **if** that tuple assigns a transition to the function which is not assigned by any tuple associated with the other constraint,
> **then** delete that tuple.

The algorithm then visits each constraint adjacent to a constraint at which a tuple was deleted, and terminates when no more filtering is possible. This process is important to the efficiency of the QSIM algorithm, since deleting a single tuple eliminates an entire region of the cross-product space of global interpretations.

4.7. Generating global interpretations

A global interpretation is an assignment of a transition to each function in the system. The result of Waltz filtering is a reduced set of tuples associated with each constraint. Not all combinations of these tuples are possible global interpretations. Suppose, for example, that we have the following constraints and associated transitions tuples:

$M^+(f, g)$	$M^+(g, h)$
(I2, I2)	(I2, I2)
(I3, I3)	(I3, I3)

Clearly, although no further local consistency filtering is possible, there are only two possible assignments of transitions to (f, g, h), namely (I2, I2, I2) and (I3, I3, I3). This pruning takes place as the global interpretations are created.

Global interpretations are built one at a time by a depth-first traversal of the space of assignments of tuples to constraints. An attempted interpretation fails if the next tuple cannot be assigned without conflicting with transitions assigned to functions by previous tuples. In case all possible next states are eliminated, the current state must be the endpoint of the domain.

A global interpretation is then used to construct a new qualitative state description, which is added to the tree of state descriptions as a successor to the current state. At this point, if all functions in a constraint are equal to landmark values, the constraint records them as a set of corresponding values.

4.8. Global filters

The completed qualitative state descriptions are mathematically plausible successors to the current state. There are, however, several global filters that can be applied before a new state is added to ACTIVE.

The mathematically valid filters applied at this stage are the following.

– *No Change*. Delete the new state if all transitions are in the set {I1, I4, I7}, because the new state description would be identical to its immediate predecessor, which therefore already adequately describes its qualitative behavior. In other words, *something* must reach a limit point for an I-transition to take place.

– *Cycle*. If the new state is identical to one of its predecessors (all functions have identical *landmark* values, and all directions of change are the same), then mark the behavior as cyclic, install a pointer to the identical predecessor, and do not add the new state to ACTIVE.

– *Divergence*. If any function takes on the value ∞ or $-\infty$, the current time-point must be the endpoint of the domain, so the new state does not go onto ACTIVE.

The first filter does not reduce the number of behaviors described, but only eliminates a redundant description. The second detects when all the consequences of a particular state have already been determined, and need not be explored anew. The third determines when a state must be at the endpoint of the domain, and thus can have no successors.

We refer to the qualitative simulation algorithm described here as the *pure* QSIM algorithm. For a particular application, additional heuristic filters may be added.[2]

4.9. Complexity

The process of formalizing qualitative simulation led to the improved QSIM algorithm, which turned out to be 30 to 60 times faster than its predecessor ENV [15, 16] on a variety of examples ranging from 3 parameters and 2 constraints (the ball) up to 16 parameters and 14 constraints (the Starling equilibrium [18, 19]). We can estimate the algorithmic complexity of QSIM as follows. Suppose there are p parameters in the system, c constraints, and the longest behavior has length t. (t is then, on average, log of the total number of qualitative states.) Since a constraint can have no more than three parameters, $p = o(c)$.

– A set of possible state transitions is assigned to each parameter from a fixed-length table, and no more than 4 transitions can be assigned to any parameter. This defines a search space of 4^p state transitions, but only $4p$ transitions need actually be created, requiring $o(p)$ time.

– A constraint can have no more than $4^3 = 64$ transition tuples. Filtering a tuple against the direction-of-change tables (Appendix B) takes constant time, but the number of corresponding values grows linearly (though slowly) with the length t of the behavior. Thus constraint filtering requires $o(ct)$ time.

[2] Some possible heuristics include:

– *Quiescence*. If all functions have derivative zero, conclude that the system is quiescent, the new time-point is the endpoint of the domain (possibly $t = \infty$), and do not place the new state on ACTIVE.

– *No Divergence*. In physical systems, eliminate transitions in which any state goes to ∞ or $-\infty$. A more accurate description of the system would include an operating region change corresponding to some component breaking.

– Waltz filtering visits each constraint at least once, but beyond that visits only neighbors of constraints where it was able to delete a tuple. Thus, the number of constraints visited is proportional to the total number of tuples, which is linear in the number of constraints. Each visit takes bounded time. Thus, Waltz filtering takes o(c) time [22].

– Generating the global interpretations explicitly constructs the remaining parts of the product space. Typically, the remaining space is small, but unfortunately there are pathological cases which yield 2^p possible successor states.

– The most expensive of the global filters is the check for previous identical states, which requires o(pt) time.

Mackworth and Freuder [22] show that in a sparse graph such as this, an interpretation satisfying the constraints may be generated in linear time. However, the *number* of global interpretations may be exponential in the number of parameters, and the QSIM algorithm generates them all. An example of this pathological case can be constructed easily. Consider a system with three parameters f, g, and h, and two constraints, DERIV(f, g) and DERIV(g, h), in a state where f, g, and h are all positive and increasing. Then the possible tuples are:

DERIV(f, g)	DERIV(g, h)
(I3, I3)	(I3, I3)
(I3, I4)	(I3, I4)
(I4, I3)	(I4, I3)
(I4, I4)	(I4, I4)

Neither local consistency filtering nor the formation of global interpretations eliminate any of the possible assignments, so for p parameters linked by a chain of DERIV constraints, there are 2^p interpretations.

f	g	h
I3	I3	I3
I3	I3	I4
I3	I4	I3
I3	I4	I4
I4	I3	I3
I4	I3	I4
I4	I4	I3
I4	I4	I4

In practice, creation of the global interpretations significantly reduces the number of compatible assignments. While this estimates the complexity of a single cycle of QSIM, the algorithm need not halt, and can continue forever producing longer and longer behaviors, each of which satisfies the qualitative constraints.

Although the QSIM algorithm is exponential in the worst case, in practice generating the successors of a given state appears to be approximately $o(ct)$. A rough sense of the effective speed of QSIM on the Symbolics 3600 can be seen from the following examples.

Example 4.3. The Spring example (3 parameters, 3 constraints) produces a three-way branching behavior of length 8 with 11 states, simulation halting after one branch is identified as a cycle (see Fig. 7). Run time: approximately 0.4 seconds.

Example 4.4. The Starling mechanism (16 parameters, 14 constraints) [18, 19] produces a single unbranching behavior of 3 states in response to a perturbation from equilibrium, halting on reaching a new equilibrium state. Run time: approximately 1.0 seconds.

5. Questions and Answers

Now that we have defined the QSIM algorithm, with a clear structure and mathematically accessible properties, we can examine it to answer some of our questions about the utility of qualitative simulation as a reasoning method. We can also compare different approaches to qualitative simulation by changing the table of permissible transitions.

5.1. Should simulation create landmarks?

The most important semantic difference between QSIM and other approaches to qualitative simulation is that QSIM can create new landmark values during the simulation, while the other algorithms require all landmarks to be specified when the structure is defined. In this section, we show that the inability to create new landmark values makes it impossible to express certain important qualitative distinctions, such as that between increasing, decreasing, and stable oscillation.

The fixed-landmark assumption is particularly deeply embedded in the de Kleer, Brown and Bobrow approach [6, 8], which depends on arithmetic operators defined over a fixed set of *qualitative values*, $\{+, 0, -\}$. A change in landmarks would change the qualitative values, and thus require the operators to be redefined. Such a redefinition is not always possible.

The structure of QSIM makes it possible to experiment with $\{+, 0, -\}$ semantics for qualitative simulation simply by replacing Table 1 with an alternate table of legal transitions (Table 2).

TABLE 2. Possible transitions under $\{+, 0, -\}$ semantics

P-transitions	$QS(f, t_i)$	$\Rightarrow QS(f, t_i, t_{i+1})$	I-transitions	$QS(f, t_i, t_{i+1})$	$\Rightarrow QS(f, t_{i+1})$
P1	$\langle l_j, \text{std}\rangle$	$\langle l_j, \text{std}\rangle$	I1	$\langle l_j, \text{std}\rangle$	$\langle l_j, \text{std}\rangle$
P2	$\langle l_j, \text{std}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	I2	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l_{j+1}, \text{std}\rangle$
P3	$\langle l_j, \text{std}\rangle$	$\langle (l_{j-1}, l_j), \text{dec}\rangle$	I3	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l_{j+1}, \text{inc}\rangle$
P4	$\langle l_j, \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	I4	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$
P5	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	I5	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l_j, \text{std}\rangle$
P6	$\langle l_j, \text{dec}\rangle$	$\langle (l_{j-1}, l_j), \text{dec}\rangle$	I6	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l_j, \text{dec}\rangle$
P7	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	I7	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{dec}\rangle$
Q8	$\langle (l_j, l_{j+1}), \text{std}\rangle$	$\langle (l_j, l_{j+1}), \text{std}\rangle$	J8	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{std}\rangle$
Q9	$\langle (l_j, l_{j+1}), \text{std}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	J9	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{std}\rangle$
Q10	$\langle (l_{j-1}, l_j), \text{std}\rangle$	$\langle (l_{j-1}, l_j), \text{dec}\rangle$	J10	$\langle (l_j, l_{j+1}), \text{std}\rangle$	$\langle (l_j, l_{j+1}), \text{std}\rangle$

The landmarks are fixed as $\{-\infty, 0, \infty\}$. The transitions that create new landmarks (I8 and I9 from Table 1) are eliminated, and new transitions are added (with Q and J names) to permit direction of change std between landmarks.

Figure 6 shows the behavior of the spring system under the $\{+, 0, -\}$ semantics. This behavior can be considered a cycle only if two functions are allowed to match between landmark values. That is, only if we may conclude from this simulation that $V(t_4) = V(t_0)$. A match between the states t_4 and t_0 in this behavior suppresses the distinction between increasing, stable, or decreasing amplitude (see Fig. 7). De Kleer and Bobrow [6] present an example of a

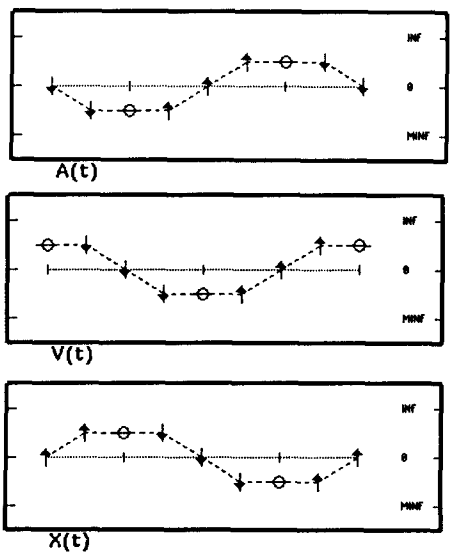

FIG. 6. The spring behavior with $\{+, 0, -\}$ semantics. In this behavior, we have $QS(V, t_0) = \langle (0, \infty), \text{std}\rangle = QS(V, t_4)$, but not necessarily $V(t_0) = V(t_4)$.

spring with frictional damping, whose actual behavior is a decreasing oscillation. The behavioral description they present is cyclic, and similar to that given in Fig. 6, with the addition of terms for the frictional force. Their description accurately captures the repetitive series of increase and decrease in the different parameters, but since it does not express a distinction between increasing, decreasing and steady amplitude, it cannot even ask which qualitative behavior is correct.

The heart of the problem is the inability to create new landmarks, or equivalently, to give *names* to newly discovered critical values. Without representing the initial value (or subsequent critical values) of a parameter in a way that permits ordinal comparison, it is not possible to ask whether the next repetition of a cycle leaves that parameter increased, decreased, or stable. If, in addition, states can be matched between landmark values, three very distinct types of behavior can be collapsed into a single, apparently cyclic, behavior. Thus, we argue that the $\{+, 0, -\}$ semantics, and in fact any semantics with a fixed set of landmarks, can collapse importantly distinct behaviors.

5.2. Is the real behavior found?

In this section, we show that all actual behaviors of a mechanism are predicted by its qualitative simulation. We take as our "gold standard" the solutions to the ordinary differential equation describing the mechanism.

We say that a real-valued function *satisfies* a given qualitative behavior description if the qualitative description of the function matches the given qualitative behavior. We then prove that any solution to a differential equation satisfies some qualitative behavior produced by the corresponding constraint equations. The proof is straightforward, since the bulk of the work has already been done in validating the individual steps of the QSIM algorithm. The algorithm generates a space including all possible behaviors of a given set of functions and constraints, and then discards only behaviors which are internally inconsistent. Thus, the remaining behaviors necessarily include all of the actual behaviors of the mechanism.

Definition 5.1. Suppose we have a reasonable function $u:[a, b] \rightarrow \mathbb{R}$ and a qualitative behavior description of the function symbol f,

$$\mathrm{QS}(f, t_0), \mathrm{QS}(f, t_0, t_1), \ldots, \mathrm{QS}(f, t_{n-1}, t_n), \mathrm{QS}(f, t_n)$$

with distinguished time-points $\{t_0, \ldots, t_n\}$ and landmarks $\{l_1, \ldots, l_k\}$. We say that u *satisfies* the behavior description if there is an order-preserving mapping m of $\{t_0, \ldots, t_n\}$ into $[a, b]$ with $m(t_0) = a$ and $m(t_n) = b$, and an order-preserving mapping of $\{l_1, \ldots, l_k\}$ into $\mathbb{R}$, such that, for all distinguished time-points t_i, $\mathrm{QS}(u, m(t_i))$ matches $\mathrm{QS}(f, t_i)$ and $\mathrm{QS}(u, m(t_i), m(t_{i+1}))$ matches $\mathrm{QS}(f, t_i, t_{i+1})$.

Theorem 5.2. *Let*

$$F[u(t), u'(t), \ldots, u^{(n)}(t)] = 0 \tag{3}$$

be an ordinary differential equation of order n, and let $\{u(t_0) = y_0, u'(t_0) = y_1, \ldots, u^{(n)}(t_0) = y_n\}$ *be the initial conditions on the solution to* (3). *Suppose that* (3) *and its initial conditions are satisfied by a reasonable function* $u : [a, b] \to \mathbb{R}$. *Let C be the set of functions and constraints derived from* (3) *by the methods of Section* 3.3, *and let* $\mathrm{QS}(F, t_0)$ *be the qualitative state description derived from the given set of initial conditions. Let T be the tree of qualitative state descriptions derived from C and* $\mathrm{QS}(F, t_0)$ *by the pure* QSIM *algorithm. Then the function u and the subexpression functions derived from it satisfy some behavioral description in T.*

Proof. QSIM works by progressively restricting the region of a space of qualitative behaviors that it is considering. By showing that any actual solution u is initially in the space, and that no filtering operation can eliminate a genuine solution, we conclude that u and its derived functions must satisfy some behavior in T.

The function u satisfies the initial state description $\mathrm{QS}(F, t_0)$ because it is a qualitative abstraction of the initial conditions to equation (3). Step 2 in QSIM generates all possible qualitative state transitions for the functions in C from a given qualitative state, using Table 1 which is justified by Propositions A.1, A.2, A.4, and A.5. Thus, any change in qualitative state of the system must be included in the possibilities generated. Step 3 of QSIM filters out combinations of transitions whose result is a state which fails to satisfy individual constraints. Inconsistent sets of directions of change are detected by comparison with tables in Appendix A. The proper implications of sets of corresponding values are checked against Propositions B.1–B.3, and B.9. The pairwise consistency filtering of Step 4 simply eliminates from consideration transitions tuples which are inconsistent with all neighboring tuples, and thus could not contribute to a global interpretation. Step 5, similarly, eliminates combinations of tuples which do not make consistent assignments of state transitions to particular functions. Finally, the global filters included in the pure QSIM algorithm are discussed in Section 4.8 and shown not to eliminate possible behaviors of the system. Thus, at each stage of the simulation, all possible successors to the current qualitative state lie in the space generated, and no genuinely possible successor is eliminated. □

5.3. Are all the behaviors real?

In this section, we show that the QSIM algorithm, and local qualitative simulation algorithms in general, cannot be guaranteed against producing spurious behaviors: behaviors which are not actual behaviors for *any* physical system satisfying the constraint equations.

A qualitative differential equation may provide few constraints, and thus predict many possible behaviors. However, the constraints are also consistent with many possible ODEs, and we might hope that each qualitative behavior corresponds to the solution to *some* ODE corresponding to the constraints. Although this is often the case, and has been conjectured to be universally true, there are cases where spurious behaviors are generated. Thus, if several behaviors are generated, some of them may not be possible behaviors of the mechanism.

One of the attractive applications of qualitative simulation is to predict possible future states, particularly to warn of surprising or disastrous events. Theorem 5.2 guarantees that there can be no false negatives: every actual behavior is predicted. However, if a valid description of the mechanism can produce invalid predictions (false positives), its usefulness is limited. As we discuss below, the problem is not fatal, but requires substantial care in the construction and use of a problem solver.

Theorem 5.3. *Let C be a set of function symbols and qualitative constraints, and let $\mathrm{QS}(F, t_0)$ be the initial qualitative state description. Let T be the tree of qualitative state descriptions derived from C and $\mathrm{QS}(F, t_0)$ by the pure* QSIM *algorithm. For some C and $\mathrm{QS}(F, t_0)$ there are behaviors in T which do not correspond to any solution $u:[a, b]\rightarrow \mathbb{R}$ to any differential equation and initial condition corresponding to C and $\mathrm{QS}(F, t_0)$.*

Proof. Consider a mass on a spring, oscillating on a frictionless surface. The constraints for this system are

$$\mathrm{DERIV}(X, V), \qquad \mathrm{DERIV}(V, A), \qquad \mathrm{M}_0^-(A, X), \tag{4}$$

which might also be written in the form of a second-order differential equation:

$$\frac{\mathrm{d}^2X}{\mathrm{d}t^2} = -\mathrm{M}_0^+(X). \tag{5}$$

With initial state $X(t_0) = 0$, $V(t_0) = V_{\mathrm{init}}$, $A(t_0) = 0$, this system is periodic for any function $A = -\mathrm{M}_0^+(X)$, because if we define total energy as

$$\mathrm{TE}(x, v) = \int_0^x \mathrm{M}_0^+(y)\,\mathrm{d}y + \tfrac{1}{2}v^2,$$

then (5) implies that $\mathrm{dTE}/\mathrm{d}t = 0$.

The local inference methods of QSIM are not able to determine, at the end of one cycle, whether the oscillation of the system is periodic, or increases or decreases in magnitude. It does, however, branch to express all three behaviors.

Figure 7 shows the behavioral description produced by qualitative simulation of the spring system. The simulation proceeds without branching through the cycle, predicting and creating new landmarks for the extrema of X, V, and A until they approach 0, V^* and 0, respectively. X and A must reach their limits together, but the simulation branches according to whether V reaches its limit at the same time (behavior 1), later (behavior 2), or earlier (behavior 3). In the first case, the state at t_4 matches the state at t_0, so the behavior is stable and periodic. In the second, the oscillation is decreasing with a new critical point less than V^*. And in the third case, motion continues past V^* to a different new critical point greater than V^*. Furthermore, having taken this branch, there is no way to represent the decision as a permanent selection of divergence, convergence, or stable oscillation. The same choice recurs at approaches to other landmarks.

Only the stable periodic behavior is an actual behavior possible for the constraints, but the local inference methods of QSIM cannot prove this fact. Thus, there are behaviors produced by the qualitative simulation algorithm which do not correspond to the behavior of any system satisfying the qualitative constraints. □

The problem also occurs with the algorithms of de Kleer and Forbus, even without creating new landmarks, if we can describe the initial state completely in terms of landmark values. In Forbus' case, we introduce a landmark value initial-length(S) for the initial displacement of the spring mass, such that A[initial-length(S)] > A[rest-length(S)] [12, pp. 144–146]. In de Kleer and Brown's case we may define a translated variable $W(t) = V(t) - V^*$ so that $W(t) = 0$ corresponds to $V(t) = V^*$ [8]. In both cases, when the system is approaching its initial state both position and velocity are approaching limits, with no way to determine which arrives first. Without the translated variable, neither approach expresses the distinction between increasing, steady, and decreasing amplitudes [17].

The underlying problem is the combination of locality with qualitative description. Both numerical and qualitative simulation are inherently local: the transition to a state is derived from its immediate predecessor. In numerical simulation, excluding truncation errors, the numerical values of the parameters implicitly encode invariant relations such as energy conservation that might be derivable from the equation. A numerical simulation of the oscillating spring will thus identify the single periodic behavior. In qualitative simulation, however, the qualitative state description of the spring is compatible with a variety of states, not all of which satisfy the invariant. There is simply not enough information in the previous state, or even the complete history of

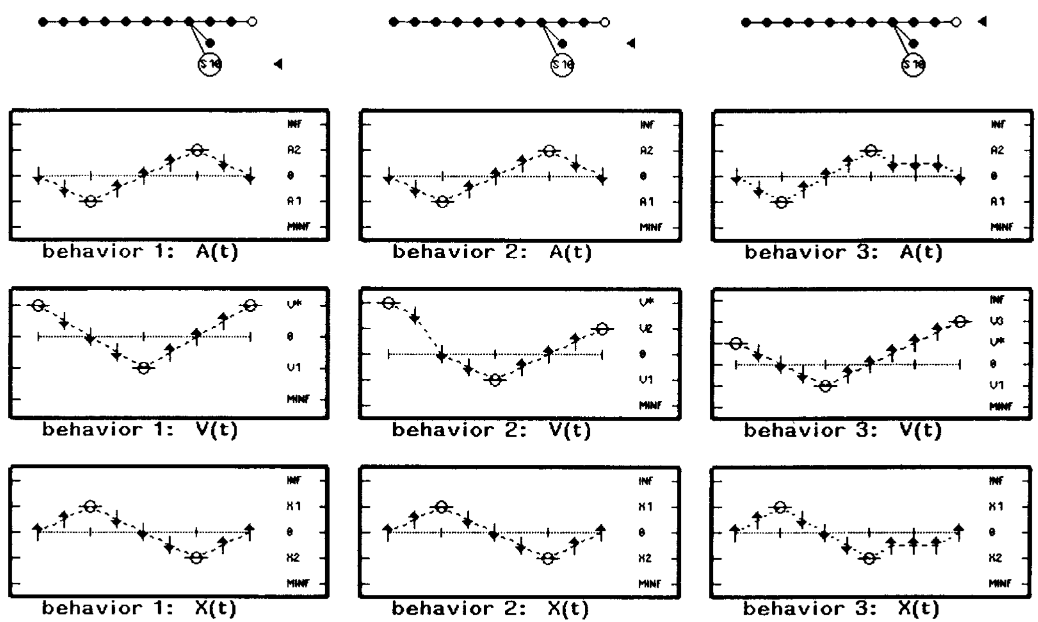

FIG. 7. The spring simulation. As the simulation approaches $t = t_4$, there is no local rule that can determine whether V reaches V^* before, after, or at the same time as X and A reach zero. The simulation branches three ways, even though only one behavior is valid.

states, to exclude all impossible behaviors. This combination of locality with qualitative state descriptions leads to the problem of spurious predictions. Thus none of the qualitative simulation algorithms can avoid this problem.

If we explicitly add the constraints representing conservation of energy to the oscillating spring constraint equations, the single correct behavior is found. However, although the additional constraints are derivable from the original equations, it is not at all clear how to do such a derivation automatically for an arbitrary mechanism.

These observations yield some important warnings about the proper use of qualitative descriptions of mechanisms, and the result of their simulation.

– Theorems 5.2 and 5.3 have a corollary that highlights their implications for knowledge engineering.

Corollary 5.4. *If a set of constraints is consistent, and if* QSIM *predicts a single behavior, then that behavior represents the actual behavior of the mechanism.*

– The consistency of a set of constraints must be demonstrated, to ensure that the qualitative simulation includes at least one genuine behavior (Theorem 5.2). When the constraints are constructed by hand, this can be done by exhibiting the ODE that they abstract. However if the set of constraints is to be derived automatically from the current process structure [12], guaranteeing consistency may be more difficult.

– If qualitative simulation yields several possible behaviors, further analysis is required before concluding that they represent possible futures.

Qualitative simulation is an important step in the process of qualitative reasoning about the behavior of mechanisms, and QSIM is a particularly complete, efficient implementation of it. However, like all tools, it has important limitations. The formal analysis we have used in this paper is valuable both for the design of the QSIM algorithm and for determining the strengths and limitations of qualitative simulation in general.

5.4. What next?

Two directions for further research appear promising for more accurate qualitative predictions of behavior. First, the *dynamical systems* approach to qualitative analysis of differential equations (e.g. [1]) has greater expressive and inferential power than local qualitative simulation methods. By describing the behaviors of the spring as trajectories through phase space rather than temporal sequences of qualitative states, it is possible to take a single branch between increasing, decreasing, and stable oscillation, rather than repeating the choice at each move toward limits. The theory of dynamical systems also includes global classification theorems delimiting the possible qualitatively distinct behaviors. Further study is needed to determine how practical problems can be stated and solved, and how the solutions can be applied.

Second, if one structural description of a mechanism has spurious behaviors, a different description might not. By changing the problem to take into account the conservation of total energy, an expanded view of the spring mechanism allows QSIM to determine that there is a single, periodic behavior. A physicist can look at equation (5) and recognize or derive the fact that it represents an energy conserving system, and therefore that the behavior must be periodic. Part of this knowledge is the ability to recognize the physical system described by a set of constraints, and to know that there is a better structural description for it; one which adds parameters and constraints (e.g. energy) that illuminate the actual behavior. This approach takes us outside the realm of qualitative simulation, and into the realm of problem formulation. Chi, Feltovich, and Glaser [2] have shown that expert causal reasoning involves the use of domain-specific knowledge to select the correct formulation of a problem, leading to its direct solution.

Returning to the larger problem of qualitative causal reasoning about mechanisms, an important problem is to *formulate* a suitable set of constraints given a physical situation, using the device-topology approach of de Kleer, Brown and Bobrow [6, 8], the process-based approach of Forbus [12], or some approach yet to be discovered.

Appendix A. The Qualitative State Transitions

This appendix applies the Intermediate Value and Mean Value Theorems to prove the validity of the transition rules in Table 1 that restrict the possible qualitative behaviors of a single function.

Let $f:[a, b] \rightarrow \mathbb{R}$ be a reasonable function with distinguished time-points

$$a = t_0 < \cdots < t_n = b ,$$

and landmark values

$$l_1 < \cdots < l_k .$$

We repeat Definition 2.9.

Definition. Where t_i is a distinguished time-point, a *P-transition* of f is a pair of adjacent qualitative states of f,

$$\mathrm{QS}(f, t_i) \Rightarrow \mathrm{QS}(f, t_i, t_{i+1}) ,$$

whose first state is the qualitative state at a distinguished time-point. An *I-transition* is a pair of adjacent qualitative states of f,

$$\mathrm{QS}(f, t_{i-1}, t_i) \Rightarrow \mathrm{QS}(f, t_i)$$

whose first state is the qualitative state on the interval between distinguished time-points.

Proposition A.1. *Let* $\mathrm{QS}(f, t_i)$ *and* $\mathrm{QS}(f, t_i, t_{i+1})$ *be adjacent qualitative states of f. Then there is some landmark value* l_j *such that* $f(t_i) = l_j$, *and the only possible P-transitions of f are given by the table below*:

	$\mathrm{QS}(f, t_i)$	$\Rightarrow$	$\mathrm{QS}(f, t_i, t_{i+1})$
P1.	$\langle l_j, \mathrm{std}\rangle$		$\langle l_j, \mathrm{std}\rangle$
P2.	$\langle l_j, \mathrm{std}\rangle$		$\langle (l_j, l_{j+1}), \mathrm{inc}\rangle$
P3.	$\langle l_j, \mathrm{std}\rangle$		$\langle (l_{j-1}, l_j), \mathrm{dec}\rangle$
P4.	$\langle l_j, \mathrm{inc}\rangle$		$\langle (l_j, l_{j+1}), \mathrm{inc}\rangle$
P6.	$\langle l_j, \mathrm{dec}\rangle$		$\langle (l_{j-1}, l_j), \mathrm{dec}\rangle$

Proof. If t_i is a distinguished time-point, then by definition there must be a landmark value l_j such that $f(t_i) = l_j$. In cases P1–P3, there are reasonable functions f with $\mathrm{QS}(f, t_i) = \langle l_j, \mathrm{std}\rangle$, and with direction of change std, inc, or dec in $\mathrm{QS}(f, t_i, t_{i+1})$, so no subsequent direction of change can be excluded. In these cases, by the Mean Value Theorem, $f(t)$ must be equal to, greater than, or less than $f(t_i) = l_j$, respectively, on the interval (t_i, t_{i+1}). By Proposition 2.5 and the Intermediate Value Theorem, $f(t)$ must be within (l_j, l_{j+1}) if it is increasing, or (l_{j-1}, l_j) if it is decreasing. In case P4, if the direction of change is inc, then $f'(t_i) > 0$. Since the derivative is continuous, there is an interval around $t = t_i$ in which $f'(t) > 0$. By Proposition 2.5, since there are points within (t_i, t_{i+1}) where the direction of change is inc, it must be inc throughout (t_i, t_{i+1}), so $f(t)$ must be within (l_j, l_{j+1}). Case P6 is similar. □

The three P-transitions from the state $\langle l_j, \mathrm{std}\rangle$ handle the case where a higher-order derivative, which is not explicitly represented by QSIM, determines the direction of motion. In such a situation, all three transitions are generated, and other constraints filter out the impossible cases. De Kleer and Bobrow [6] determine and use higher-order derivatives explicitly to make that decision, at least for linear equations. Their approach determines the order of the structural description, and thus knows how many higher-order derivatives to compute. The advantages of the current approach are that it places much weaker conditions on the differentiability of the parameters, and that it is not restricted to linear equations.

Proposition A.2. *Let* $\mathrm{QS}(f, t_i, t_{i+1})$ *and* $\mathrm{QS}(f, t_{i+1})$ *be adjacent qualitative states of f. Then there are landmark values* l_j *and* l_{j+1} *such that the only possible I-transitions are given by the table below*:

	$\text{QS}(f, t_i, t_{i+1})$	$\Rightarrow$ $\text{QS}(f, t_{i+1})$
I1.	$\langle l_j, \text{std}\rangle$	$\langle l_j, \text{std}\rangle$
I2.	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l_{j+1}, \text{std}\rangle$
I3.	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle l_{j+1}, \text{inc}\rangle$
I5.	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l_j, \text{std}\rangle$
I6.	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle l_j, \text{dec}\rangle$

Proof. Similar to the proof of Proposition A.1 Cases I1, I2, I3, I5, and I6 correspond to cases P1, P3, P6, P2 and P4, respectively. □

When considering f in the context of a system F, we need transitions appropriate to a time-point t which is distinguished for the system but not for the individual function. In other words, t might reach a distinguished time-point before $f(t)$ reaches its next landmark value.

Proposition A.3. *Suppose that an additional time-point $t^* \in [a, b]$ is introduced into the set of distinguished time-points of f:*

$$a = t_0 < \cdots < t_k < t^* < t_{k+1} < \cdots < t_n = b\ .$$

Then

$$\text{QS}(f, t_k, t^*) = \text{QS}(f, t^*) = \text{QS}(f, t^*, t_{k+1}) = \text{QS}(f, t_k, t_{k+1})\ .$$

Proof. Since, for all $t \in (t_k, t_{k+1})$, $\text{QS}(f, t) = \text{QS}(f, t_k, t_{k+1})$ by definition, we conclude that $\text{QS}(f, t_k, t_{k+1})$ is identical to each of $\text{QS}(f, t_k, t^*)$, $\text{QS}(f, t^*)$, and $\text{QS}(f, t^*, t_{k+1})$. □

Proposition A.4. *Let $f: [a, b] \to \mathbb{R}$ be a reasonable function, and let $a = t_0 < \cdots < t_n = b$ be a set of time-points including all the distinguished time-points of f, but possibly additional points in $[a, b]$. Then the possible transitions of f consist of those listed in Propositions* A.1 *and* A.2, *plus the following I-transitions and P-transitions*:

	$\text{QS}(f, t_i, t_{i+1})$	$\Rightarrow$ $\text{QS}(f, t_{i+1})$
I4.	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$
I7.	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{dec}\rangle$
	$\text{QS}(f, t_i)$	$\Rightarrow$ $\text{QS}(f, t_i, t_{i+1})$
P5.	$\langle (l_j, l_{j+1}), \text{inc}\rangle$	$\langle (l_j, l_{j+1}), \text{inc}\rangle$
P7.	$\langle (l_j, l_{j+1}), \text{dec}\rangle$	$\langle (l_j, l_{j+1}), \text{dec}\rangle$

Proof. Consider an I-transition beginning with $QS(f, t_i, t_{i+1})$. If t_{i+1} is a distinguished time-point of f, the transition is specified by Proposition A.2. If t_{i+1} is not a distinguished time-point of f, Proposition A.3 shows that the qualitative state remains constant across the transition, providing the above alternatives as the only additional possibilities. If the transition is a P-transition beginning with $QS(f, t_i)$ the proof is similar, depending on whether t_i is a distinguished time-point of f. □

A.1. Discovering a new landmark

Suppose that $QS(f, t_i, t_{i+1}) = \langle (l_j, l_{j+1}), \text{inc} \rangle$. Since there is no landmark in (l_j, l_{j+1}), we can exclude the transition to

$$QS(f, t_{i+1}) = \langle (l_j, l_{j+1}), \text{std} \rangle$$

because that would imply that $f'(t_{i+1}) = 0$, making $f(t_{i+1})$ a landmark value, by definition. However, if $l_1 < \cdots < l_k$ is only a *partial* set of the landmarks of f, then the above transition is possible, but only when $f(t_{i+1}) = l^*$, a landmark value of f such that $l_j < l^* < l_{j+1}$.

In this case, the following partial behavior is possible:

qualitative state	known landmarks
$QS(f, t_i, t_{i+1}) = \langle (l_j, l_{j+1}), \text{inc} \rangle$	$l_j < l_{j+1}$
$QS(f, t_{i+1}) = \langle l^*, \text{std} \rangle$	$l_j < l^* < l_{j+1}$

$QS(f, t_i, t_{i+1})$ is now seen to be syntactically incorrect, given subsequent information acquired about the true set of landmarks. It should be revised to be $QS(f, t_i, t_{i+1}) = \langle (l_j, l^*), \text{inc} \rangle$. Furthermore, it is possible for f to move across l^* several times before encountering the critical point that reveals its existence as a landmark. However, the modifications needed to correct the behavioral description are straightforward and locally computable. In the ball system example discussed above, the maximum height of the ball is such a new landmark, discovered when $V(t)$, and therefore $Y'(t)$, become zero. We summarize this discussion in:

Proposition A.5. *Suppose that $l_1 < \cdots < l_k$ are all the known landmarks of a reasonable function f, which may have other landmarks as yet unknown. Then, in addition to the transitions listed in Propositions* A.1, A.2, *and* A.4 *the following I-transitions are possible*:

	$QS(f, t_i, t_{i+1})$	$\Rightarrow$ $QS(f, t_{i+1})$
I8.	$\langle (l_j, l_{j+1}), \text{inc} \rangle$	$\langle l^*, \text{std} \rangle$
I9.	$\langle (l_j, l_{j+1}), \text{dec} \rangle$	$\langle l^*, \text{std} \rangle$

In case one of these transitions is followed, the set of landmark values in $\mathrm{QS}(f, t_{i+1})$ *is augmented by* $l_j < l^* < l_{j+1}$. *Note that the total ordering on the set of landmarks is preserved.*

Table 1 collects and names the transitions permitted by Propositions A.1, A.2, A.4, and A.5 for use in the QSIM algorithm.

A.2. Infinity and asymptotic approach

I-transitions express the possible consequences of a changing parameter reaching a limiting landmark value. But what if a parameter approaches its limit asymptotically? By allowing both domain and range to include $+\infty$ and $-\infty$ as endpoints, we can express asymptotic approach as reaching the limit point at $t = \infty$. The same method allows us to treat divergence to infinite values as a possible behavior. Thus, every time-interval has an endpoint, but some distinguished time-points (e.g. $t = \infty$, or t such that $f(t) = \infty$) may have no successor states. There are two constraints on these types of behavior.

First, at $t = \infty$, every function in the system must be equal to some landmark value and must, if that landmark is finite, have derivative zero (i.e. direction of change std). Recall that oscillatory systems are handled with a finite domain and repeated states, rather than with an infinite domain.

Proposition A.6. *Let* $f : [a, \infty] \rightarrow \mathbb{R}^*$ *be a reasonable function. If the limit of* $f(t)$ *as* $t \rightarrow \infty$ *is finite, then* $\lim_{t\rightarrow\infty} f'(t) = 0$.

Proof. If $\lim_{t\rightarrow\infty} f'(t) > 0$, then for some interval (c, ∞), $f'(t)$ must be bounded away from zero. In this interval, $f(t) = f(c) + f'(t^*) \cdot (t - c)$ for some $t^* \in (c, \infty)$, by the Mean Value Theorem. Thus, $\lim_{t\rightarrow\infty} f(t) = \infty$. Similarly in case $\lim_{t\rightarrow\infty} f'(t) < 0$, so the limit must be zero. □

Second, if $f(t) = \infty$, then $t = b$, the right-hand endpoint of the domain, since a function cannot be continuously differentiable across ∞. If $b < \infty$, then the direction of change must be inc.

Proposition A.7. *Let* $f : [a, b] \rightarrow \mathbb{R}^*$ *be a reasonable function such that* $\lim_{t\rightarrow b} f(t) = \infty$, *where* b *is finite. Then* $\lim_{t\rightarrow b} f'(t) = \infty$.

Proof. Suppose that $\lim_{t\rightarrow b} f'(t)$ has a finite limit $M > 0$. Then for some interval $(b - \delta, b)$, $f'(t) \in (M - \varepsilon, M + \varepsilon)$, which implies that

$$f(b - \delta) + \delta \cdot (M - \varepsilon) < f(b) < f(b - \delta) + \delta \cdot (M + \varepsilon)$$

which contradicts $f(b) = \infty$. □

Using these propositions, we can test whether a distinguished time-point can match $t = \infty$, and test moves to ∞ for consistency. With these observations, the extended reals $[-\infty, \infty]$ can be treated like any other closed interval, and asymptotic approach is handled.

Appendix B. Constraint Consistency

This appendix specifies the rules by which each type of constraint tests a tuple of qualitative state transitions for consistency. There are separate tests for consistency of the qualitative magnitudes and the directions of change.

B.1. Qualitative magnitude consistency

This appendix defines and justifies the evaluation of M^+, M^-, ADD, or MULT constraints when applied to particular qualitative values.

The magnitude of a quantity is described qualitatively in terms of its ordinal relations with a set of landmark values. The validity of a particular application of an M^+, M^-, ADD, or MULT predicate is tested using not only the signs of the arguments, but also their relations with other sets of corresponding values. For example, if we know that ADD(p, q, r) is true, then (p, q, r) is a set of corresponding values for this ADD constraint, and if $p' < p$ and $r' > r$, we can determine that ADD(p', q, r') must be false. In the QSIM algorithm, these predicates are evaluated in order to test the validity of a possible tuple of transitions at a particular constraint.

The criteria below compare transition tuples with known sets of corresponding values. Each criterion checks whether the ordinal relations between the current values of the functions and the corresponding values will be consistent with the constraint after the proposed transition. The qualitative state before the transition is presumed to be consistent.

These criteria generalize the transition-ordering rules of Williams [24, 25], which evaluate combinations of transitions to zero across various constraints. The presence of nonzero landmark values means that, for example, a particular function could be moving toward zero, but be separated from it by an intervening landmark. Thus, transition-ordering rules must take into account not only of directions of change, but also relative position in an ordering. Furthermore, a problem with transition-ordering rules in general is that their proofs, while not deep, often consist of numerous cases and are difficult to check. The simple and transparently valid criteria of Propositions B.3 and B.9 subsume the other rules and are more efficiently applied.

B.1.1. *Monotonic function constraints*

If two function f and g, related by $M^+(f, g)$, are approaching corresponding limits, we know that either both reach their limits together, or neither does.

Proposition B.1. *Suppose* $M^+(f, g)$, *with corresponding values* (p, q), *and*

$$QS(f, t_1, t_2) = \langle (p, p'), \text{dec} \rangle ,$$
$$QS(g, t_1, t_2) = \langle (q, q'), \text{dec} \rangle .$$

Then one of the following two possibilities must be true at t_2:
(1) $f(t_2) = p,\ g(t_2) = q$;
(2) $f(t_2) > p,\ g(t_2) > q$.

Proof. Since $M^+(f, g)$ is true, there is a strictly monotonic function H such that $f(t) = H(g(t))$ for all $t \in [a, b]$. In particular, since p and q are corresponding values, $p = H(q)$. Thus, if $g(t_2) = q$, we know that $f(t_2) = p$, and conversely by the symmetry of $M^+(f, g)$. □

If f and g are approaching limits, but only one of the limit points belongs to a corresponding value pair, then only the *other* limit point is possibly reachable in the next state.

Proposition B.2. *Suppose* $M^+(f, g)$, *with corresponding values* (p, q) *and*

$$QS(f, t_1, t_2) = \langle (p, p'), \text{dec} \rangle ,$$
$$QS(g, t_1, t_2) = \langle (q'', q'), \text{dec} \rangle ,$$

where $q'' \neq q$. *Then one of the following two possibilities must be true at* t_2:
(1) $f(t_2) > p,\ g(t_2) = q''$,
(2) $f(t_2) > p,\ g(t_2) > q''$.

Proof. Since (p, q) is a corresponding value pair, there must be some $t^* \in [a, b]$ such that $f(t^*) = p$ and $g(t^*) = q$. It is not possible for $q' \leq q$, because then $g(t) < g(t^*)$ while $f(t) > f(t^*)$ for $t \in (t_1, t_2)$, which contradicts $M^+(f, g)$. Thus $q < q'' < q'$.

f cannot reach p without g simultaneously reaching q, as shown in the previous proposition. Thus the only possibilities are that g reaches q'', or that neither reaches its limit. □

By symmetry, it is clear that analogous propositions hold whether the constraint is M^+ or M^-, or whether the corresponding limits are approached from above, below, or one from each side.

B.1.2. *Addition constraint*

We can use exactly the same technique to prove similar consistency criteria for ADD(f, g, h) and MULT(f, g, h), in cases where three, two or only one of the limit points belongs to a known set of corresponding values. The complexity of determining, implementing, and verifying all such tests is formidable. Fortunately, there is a general relationship that captures all possible such criteria.

Proposition B.3. *Let p, q, and r be corresponding values of f, g, and h, where* ADD(f, g, h). *Then, for any $t \in [a, b]$, the following holds:*

$$(f(t) - p) + (g(t) - q) = (h(t) - r). \tag{B.1}$$

Proof. $f(t) + g(t) = h(t)$ and $p + q = r$. □

If we know that ADD(f, g, h), and if the limits approached by f, g, and h intersect with the corresponding values $p + q = r$, then (B.1) may be applied. A proposed transition tuple is checked for consistency by seeing whether its result would change the sign of any of the terms in (B.1). The signs of the three terms can be determined directly from the ordinal relations among function values and landmarks, and the resulting state is checked for consistency with the ADD relation by table lookup.

The following propositions demonstrate the effect of this filter on cases where the limits of f, g, and h share three, two, or only one value with a particular correspondence.

Proposition B.4. *Suppose* ADD(f, g, h), *with corresponding values $p + q = r$, and*

$$\begin{aligned}
\text{QS}(f, t_1, t_2) &= \langle (p, p'), \text{dec} \rangle\ ,\\
\text{QS}(g, t_1, t_2) &= \langle (q, q'), \text{dec} \rangle\ ,\\
\text{QS}(h, t_1, t_2) &= \langle (r, r'), \text{dec} \rangle\ .
\end{aligned}$$

Then exactly one of the following four possibilities must be true at t_2:

(1) $f(t_2) = p$, $g(t_2) = q$, $h(t_2) = r$;
(2) $f(t_2) = p$, $g(t_2) > q$, $h(t_2) > r$;
(3) $f(t_2) > p$, $g(t_2) = q$, $h(t_2) > r$;
(4) $f(t_2) > p$, $g(t_2) > q$, $h(t_2) > r$.

Proof. For $t \in [t_1, t_2]$, $f(t) - p \geq 0$, $g(t) - q \geq 0$, and $h(t) - r \geq 0$. The three terms of (B.1) must have compatible signs, and cannot change discontinuously from their state in (t_1, t_2), so (1)–(4) are the only possibilities. □

In case we are not so fortunate as to have f, g, and h approaching corresponding limits, we may have two of the functions approaching corresponding limits, and know where the corresponding value of the third function is with respect to its limit. This allows us further to constrain the set of possible next states.

Proposition B.5. *Suppose* ADD(f, g, h), *with corresponding values* $p + q = r$, *and*

$$\begin{aligned} \mathrm{QS}(f, t_1, t_2) &= \langle (p, p'), \mathrm{dec} \rangle , \\ \mathrm{QS}(g, t_1, t_2) &= \langle (q, q'), \mathrm{dec} \rangle , \\ \mathrm{QS}(h, t_1, t_2) &= \langle (r'', r'), \mathrm{dec} \rangle , \end{aligned}$$

where $r'' \neq r$. *Then it is not possible to have both* $f(t_2) = p$ *and* $g(t_2) = q$.

Proof. Consider (B.1). If $r \geqslant r'$, then the term $h(t) - r$ is negative for $t \in (t_1, t_2)$, while the other two terms are positive, which is a contradiction. Thus $r < r''$.

All terms of (B.1) are positive on (t_1, t_2), and $h(t) - r$ must be strictly positive at $t = t_2$, so at most one of the other terms can be zero at t_2. □

Proposition B.6. *Suppose* ADD(f, g, h), *with corresponding values* $p + q = r$, *and*

$$\begin{aligned} \mathrm{QS}(f, t_1, t_2) &= \langle (p, p'), \mathrm{dec} \rangle , \\ \mathrm{QS}(g, t_1, t_2) &= \langle (q'', q'), \mathrm{dec} \rangle , \\ \mathrm{QS}(h, t_1, t_2) &= \langle (r, r'), \mathrm{dec} \rangle , \end{aligned}$$

where $q'' \neq q$. *Then it is not possible to have both* $f(t_2) = p$ *and* $h(t_2) = r$. *If* $q < q''$, *it is impossible to have* $h(t_2) = r$. *If* $q > q''$, *it is impossible for* $f(t_2) = p$.

Proof. If $q < q''$, the middle term of equation (B.1) is strictly positive on $[t_1, t_2]$, so only the first term can possibly become zero at t_2, so only f and g can possibly reach their limits. If $q > q''$, then the second term of (B.1) is strictly negative on $(t_1, t_2]$, so only the third term can possibly become zero at t_2, so only g and h can possibly reach their limits. □

The same technique can be used in case only one of a set of corresponding values appears in the current set of limits.

Proposition B.7. *Suppose* ADD(f, g, h), *with corresponding values* $p + q = r$, *and*

$$\begin{aligned} \mathrm{QS}(f, t_1, t_2) &= \langle (p, p'), \mathrm{dec} \rangle , \\ \mathrm{QS}(g, t_1, t_2) &= \langle (q'', q'), \mathrm{dec} \rangle , \\ \mathrm{QS}(h, t_1, t_2) &= \langle (r'', r'), \mathrm{dec} \rangle , \end{aligned}$$

where $q'' \neq q$ and $r'' \neq r$. Then if $q'' > q$ and $r'' < r$, or if $q'' < q$ and $r'' > r$, it is impossible for $f(t_2) = p$.

Proof. Examination of (B.1) shows that these cases would result in the first term being zero, while the other two have opposite signs, which is impossible. □

Proposition B.8. *Suppose* ADD(f, g, h), *with corresponding values $p + q = r$, and*

$$\mathrm{QS}(f, t_1, t_2) = \langle (p'', p'), \mathrm{dec} \rangle ,$$
$$\mathrm{QS}(g, t_1, t_2) = \langle (q'', q'), \mathrm{dec} \rangle ,$$
$$\mathrm{QS}(h, t_1, t_2) = \langle (r, r'), \mathrm{dec} \rangle ,$$

where $p'' \neq p$ and $q'' \neq q$. Then if $p'' > p$ and $q'' > q$, or if $p'' < p$ and $q'' < q$, it is impossible for $h(t_2) = r$.

Proof. Examination of (B.1) shows that these cases would result in the last term being zero, while the other two have the same signs, which is impossible. □

By symmetry, similar propositions hold in the cases where f, g, and h are approaching their limits from various combinations of directions, not only when all are decreasing. Fortunately, equation (B.1) makes it unnecessary to implement checks based directly on Propositions B.4–B.8.

B.1.3. *Multiplication constraint*

If the three functions in a MULT constraint are approaching related limits, we can constrain the possible results, similarly to what we did with ADD constraints in the previous section. A separate consistency test checks for legal combinations of signs $(+, 0, -)$ at a multiplication constraint.

Proposition B.9. *Let p, q, and r be nonzero corresponding values of the functions f, g, and h, respectively, where* MULT(f, g, h). *Then, for any $t \in [a, b]$, the following holds*:

$$\left(\frac{f(t)}{p}\right) \cdot \left(\frac{g(t)}{q}\right) = \left(\frac{h(t)}{r}\right) . \tag{B.2}$$

Proof. $f(t) \cdot g(t) = h(t)$ and $p \cdot q = r$. □

As with the addition constraint, QSIM uses equation (B.2) directly to test the consistency of various combinations of f, g, and h reaching their limit values, in comparison with a corresponding set of values, $p \cdot q = r$. When $f(t)$ and p have the same sign, we can retrieve their ordinal relations to classify the term $f(t)/p$ as greater than, less than, or equal to 1. With respect to this classification, the legal combinations of A, B, and C, where MULT(A, B, C) are given by the following table:

A \ B	<1	=1	>1
<1	<1	<1	any
=1	<1	=1	>1
>1	any	>1	>1

Note that it is not necessary for the value 1 to be a landmark value of any of the functions involved. The table is a guide to the implementation of a consistency test for MULT(f, g, h), rather than representing an inference that QSIM makes explicitly. The consistency test applies when the ordinal relation between $f(t)$ and p changes as f reaches its limit.

Propositions can be proved to demonstrate the degree of filtering possible with different sets of corresponding values, similar to Propositions B.4–B.8 for addition, but they are omitted here.

B.1.4. *Landmark values as a representation for critical points*

There are several ways to represent the critical values of a parameter, but they are not identical in power and cost. QSIM creates landmark values in the quantity space of a parameter to represent critical values, where the derivative of the parameter becomes zero. The $\{+, 0, -\}$ semantics can express nonzero landmarks by appending a new, translated parameter and constraint to the set of constraints. Thus, given a parameter $V(t)$, if we wish to represent a new landmark value $V_{max} > 0$, we define a translated parameter $W(t)$ and the constraint $W(t) = V(t) - V_{max}$.

The translated parameter technique captures many, but not all, of the properties of landmark values. In particular, each constraint in a QSIM model records sets of corresponding landmark values that are known to satisfy it. These corresponding values are very useful for limiting the possible next states predicted by QSIM, by applying the qualitative magnitude filtering methods of this appendix. The translated parameter technique does not support this filtering method in any straightforward way, if at all.

Suppose we have parameters x, y, and z, such that $x + y = z$. At some time-point, all three are observed to be equal to landmark values, so the addition constraint stores the corresponding values (p, q, r). Now suppose that we are in a state where $x > p$, $y > q$, and $z > r$, all decreasing toward their landmark values. In the QSIM representation, the methods of Propositions B.3 and B.4 let us conclude that only four of the eight possible resulting qualitative states are consistent.

Using the technique of translated parameters, in order to represent the landmark values p, q, and r, we need to define additional parameters and redescribe the current state, as follows:

$$\begin{aligned} x' &= x - p\,, & x' &> 0\,,\\ y' &= y - q\,, & y' &> 0\,,\\ z' &= z - r\,, & z' &> 0\,. \end{aligned}$$

Existing techniques, such as Williams' transition-ordering rules [24, 25] would then be able to reason about the order that x', y', and z' would reach zero, except that we do not necessarily have the constraint that $x' + y' = z'$!

The event that led QSIM to record a set of corresponding values would be represented using translated parameters as $x' = 0$, $y' = 0$, and $z' = 0$. Short of deriving all algebraic consequences of all combinations of constraints, it is not clear how a qualitative simulation algorithm would focus its attention sufficiently to derive the relation $x' + y' = z'$ as a way of capturing the correspondence (p, q, r). There is at least no straightforward translation of corresponding values into the translated parameter technique.

Without the corresponding values, the translated parameter approach must predict eight possible successor states for the three parameters, while QSIM can restrict the set to four.

B.2. Direction-of-change consistency

An important difference between QSIM and the algorithms used by de Kleer and Forbus is that quantities are represented by qualitative *descriptions* rather than qualitative *values*. Thus, rather than being a partial function that sometimes fails to compute a result, ADD is a three-place relation evaluating to true or false according to whether its arguments satisfy the addition constraint. This appendix specifies the tables of acceptable directions of change for ADD and MULT. The corresponding tables for M^+ and M^- are obvious.

B.2.1. ADD(f, g, h)

The following table summarizes the combinations of directions of change that satisfy the ADD(f, g, h) constraint.

f \ g	inc	std	dec
inc	inc	inc	any
std	inc	std	dec
dec	any	dec	dec

B.2.2. MULT(f, g, h)

The combinations of directions of change that satisfy the MULT constraint depend on the signs of f, g, and h, as shown in the following tables, derived from the identity $h' = f'g + fg'$.

(1) If $f > 0$, $g > 0$, $h > 0$,

f \ g	inc	std	dec
inc	inc	inc	any
std	inc	std	dec
dec	any	dec	dec

(2) If $f < 0$, $g < 0$, $h > 0$,

f \ g	inc	std	dec
inc	dec	dec	any
std	dec	std	inc
dec	any	inc	inc

(3) If $f > 0$, $g < 0$, $h < 0$,

f \ g	inc	std	dec
inc	any	dec	dec
std	inc	std	dec
dec	inc	inc	any

In case $f < 0$ and $g > 0$, the table is transposed.

(4) If $f > 0$, $g = 0$, $h = 0$,

f \ g	inc	std	dec
inc	inc	std	dec
std	inc	std	dec
dec	inc	std	dec

In case $f < 0$, the table remains the same, but with the signs reversed. If $f = 0$ and $g \neq 0$, the table is transposed.

(5) If $f = 0$, $g = 0$, $h = 0$,

f \ g	inc	std	dec
inc	std	std	std
std	std	std	std
dec	std	std	std

Appendix C. The QSIM Program and its Output

This appendix provides the constraints, the initialization, and a trace of one cycle of the output of QSIM on the spring mechanism

```
(define-structure spring
  (functions a v x)
  (landmarks (v (minf 0 v* inf)))
  (constraints (d//dt v a) (d//dt x v) (m0- a x))
  (invariants
    (a ((minf inf) nil))
    (v ((minf inf) nil))
    (x ((minf inf) nil))))

(defun initialize-spring ()
  (make-initialization spring (generate-time-point)
                       '((x (0 inc))
                         (v (v* std))
                         (a (0 dec)))))
```

With appropriate trace switches set, the QSIM program will print out information about the quantitative progress of the filtering algorithm. The following fragment shows the spring reaching its first extremum. The sections of the trace describe:

- the initial qualitative state description;
- the number of qualitative state transitions assigned to each individual function;
- the decrease in number of transition tuples as each constraint applies its filters;
- the effect of Waltz filtering when consideration of an adjacent constraint decreases the number of tuples;
- the effect of global filters;
- the assignment(s) of transition rules that will create the successor state(s).

```
Predicting successors of S2 in region SPRING.
  0 < V[DEC] < V*
  0 < X[INC] < INF
  MINF < A[DEC] < 0
Function A has 2 transitions.
Function V has 4 transitions.
Function X has 2 transitions.
Constraint M⁻(A X) has (4)→(2) tuples.
Constraint D//DT(X V) has (8)→(4) tuples.
Constraint D//DT(V A) has (8)→(4) tuples.
Waltz filter of D//DT(X V): (4)→(3).
Waltz filter of D//DT(X V): (3)→(2).
Global interpretations: (2)→(1).
Predicting:
  V: I6 ((M1 M2) DEC)⇒(M1 DEC)
  X: I8 ((M1 M2) INC)⇒(M* STD)
  A: I9 ((M1 M2) DEC)⇒(M* STD)
⇒1 successors.

Predicting successors of S3 in region SPRING.
  0 = V [DEC]
  X1 = X [STD]
  A1 = A [STD]
```

ACKNOWLEDGMENT

The environment of the Clinical Decision Making Group at the MIT Laboratory for Computer Science provided valuable support and stimulation for this research. Christopher Eliot has been important to this work in many ways. He first suggested the idea of filtering sets of local qualitative state transitions. He is also responsible for the implementation of the QSIM algorithm with a very comfortable interface, and for a number of valuable technical contributions. I am also grateful to Daniel Bobrow, Johan de Kleer, Ken Forbus, Walter Hamscher, Ramesh Patil, Elisha Sacks, Mike Wellman, and Brian Williams for helpful discussions and observations on qualitative simulation in general and QSIM in particular. I would also like to thank the two hard-working referees for this paper.

AVAILABILITY OF THE QSIM PROGRAM

The QSIM program is available to researchers interested in qualitative simulation, in order to encourage detailed exploration and evaluation of these ideas and their possible applications beyond what is possible in a published paper. The current implementation runs in ZETALISP on the Symbolics 3600. It is not advertised or warranted as a software product, and any commercial rights to the program are retained. Please contact me if you are interested.

REFERENCES

1. Arnold, V.I., *Ordinary Differential Equations* (MIT Press, Cambridge, MA, 1973) (translated and edited by R.A. Silverman).
2. Chi, M.T.H., Feltovich, P.J. and Glaser, R., Categorization and representation of physics problems by experts and novices, *Cognitive Sci.* **5** (1982) 121–152.
3. de Kleer, J., Multiple representations of knowledge in a mechanics problem-solver, in: *Proceedings Fifth International Joint Conference on Artificial Intelligence*, Cambridge, MA, 1977.
4. de Kleer, J., Causal and teleological reasoning in circuit recognition, MIT Artificial Intelligence Laboratory TR-529, Cambridge, MA, 1978.
5. de Kleer, J., The origin and resolution of ambiguities in causal arguments, in: *Proceedings Sixth International Joint Conference on Artificial Intelligence*, Tokyo, Japan, 1979.
6. de Kleer, J., and Bobrow, D.G., Qualitative reasoning with higher-order derivatives, in: *Proceedings Fourth National Conference on Artificial Intelligence*, Austin, TX (William Kaufman, Los Altos, CA, 1984).
7. de Kleer, J. and Brown, J.S., The origin, form, and logic of qualitative physical laws, in: *Proceedings Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, F.R.G. (William Kaufman, Los Altos, CA, 1983).
8. de Kleer, J. and Brown, J.S., A qualitative physics based on confluences, *Artificial Intelligence* **24** (1984) 7–83.
9. Forbus, K.D., Qualitative reasoning about physical processes, in: *Proceedings Seventh International Joint Conference on Artificial Intelligence*, Vancouver, BC (William Kaufman, Los Altos, CA, 1981).
10. Forbus, K.D., Qualitative process theory, MIT Artificial Intelligence Laboratory Memo 664, Cambridge, MA, 1982.
11. Forbus, K.D., Measurement interpretation in qualitative process theory, in: *Proceedings Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, F.R.G. (William Kaufman, Los Altos, CA, 1983).
12. Forbus, K.D., Qualitative process theory, *Artificial Intelligence* **24** (1984) 85–168.
13. Gentner, D. and Stevens, A., (Eds.), *Mental Models* (Erlbaum, Hillsdale, NJ, 1983).
14. Hayes, P.J., The naive physics manifesto, in: D. Michie (Ed.), *Expert Systems in the Micro-electronic Age* (Edinburgh University Press, Edinburgh, 1979).
15. Kuipers, B.J., Getting the envisionment right, in: *Proceedings Second National Conference on Artificial Intelligence*, Pittsburgh, PA, 1982.
16. Kuipers, B.J., Commonsense reasoning about causality: Deriving behavior from structure, *Artificial Intelligence* **24** (1984) 169–203.
17. Kuipers, B.J., The limits of qualitative simulation, in: *Proceedings Ninth International Joint Conference on Artificial Intelligence*, Los Angeles, CA (William Kaufman, Los Altos, CA, 1985).
18. Kuipers, B.J. and Kassirer, J.P., How to discover a knowledge representation for causal reasoning by studying an expert physician, in: *Proceedings Eighth International Joint Conference on Artificial Intelligence*, Karlsruhe, F.R.G. (William Kaufman, Los Altos, CA, 1983).

19. Kuipers, B.J. and Kassirer, J.P., Causal reasoning in medicine: Analysis of a protocol, *Cognitive Sci.* **8** (1984) 363–385.
20. Kuipers, B.J. and Kassirer, J.P., Qualitative simulation in medical physiology: A progress report, MIT Laboratory for Computer Science TM-280, Cambridge, MA, 1985.
21. Mackworth, A.K., Consistency in networks of relations, *Artificial Intelligence* **8** (1977) 99–118.
22. Mackworth, A.K. and Freuder, E.C., The complexity of some polynomial network consistency algorithms for constraint satisfaction problems, *Artificial Intelligence* **25** (1985) 65–74.
23. Waltz, D., Understanding line drawings of scenes with shadows, in: P.H. Winston (Ed.), *The Psychology of Computer Vision* (McGraw-Hill, New York, 1975).
24. Williams, B.C., The use of continuity in a qualitative physics, in: *Proceedings Fourth National Conference on Artificial Intelligence*, Austin, TX (William Kaufman, Los Altos, CA, 1984).
25. Williams, B.C., Qualitative analysis of MOS circuits, *Artificial Intelligence* **24** (1984) 281–346.

Received December 1984, revised version received January 1986