Beyond "Qualitative" Reasoning

Extended Abstract Submitted to the Organizing Committee of the **1987 Workshop on Qualitative Physics**

Brian C. Williams
MIT Artificial Intelligence Laboratory
545 Technology Square
Cambridge, MA 02139
(Arpanet: williams%mit-oz@mit-mc.arpa)
(Phone: (617) 253-6292)

 This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support is provided in part by an Analog Devices Fellowship, and in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-80-C-0505.

1 Introduction

Current research in qualitative reasoning has been devoted to developing techniques for analyzing the behavior of physical systems in terms of a restricted representation for quantities (i.e., sets of partial orders or regions of state space) that is called "qualitative". The success of this line of research is a result of focusing on a well defined representation followed by a careful analysis of the representation's properties, and its use in predicting behavior.

Although the "qualitative" representations studied are important, they can be used to capture only a fraction of those descriptions that we intuitively consider to be qualitative. The current state of the art is at a point where we are ready to ask the question: "What directions must we push our research in order to reason about the broad spectrum of behaviors that we consider qualitative?"

In this abstract we consider one such direction, specifically: "What temporal representations and reasoning techniques are required to analyze the time varying behavior of a complex system, while permitting a wide variety of value representations to be used?" We begin with a summary of the current state of the art in qualitative analysis. Next we discuss the limitations of the representations used for values and the broader set of value representations that is required. We then examine several fundamental limitations with the current representations used for behavior. Finally we briefly sketch out a new representation for behavior that incorporates many important elements of history-based and state-based descriptions, and thereby avoids these limitations. This new representation can be used to incorporate a wide spectrum of value representations, in a description of a system's behavior. A detailed description of this representation and its corresponding analysis technique (*partial-state envisionment*) is not included, but will be the topic of the extended version of this abstract and the corresponding talk.

2 The Qualitative Analysis of Time-Varying Behavior

Given the model of a complex device and a set of time-varying inputs, qualitative analysis produces 1) a description of the device's behavior over time in terms of a set of properties of interest, and 2) a causal explanation of why this behavior came about.[1] Both aspects of qualitative analysis play an essential role in a problem solving task such as design. Generating the device's behavior is crucial to verifying the design, while the causal explanation is key to debugging any design flaws. It is essential that a qualitative analysis technique be able to generate clear concise explanations and descriptions of behavior for a broad class of systems.

While current work on qualitative analysis provides a starting point, these techniques have limitations that prevent them from generating descriptions like the one shown in figure 1. Two of these limitations are: 1) the qualitative representations for values are

[1] The term *simulation* is generally used to refer to the process of generating a description of what a system's behavior is. We use the term *analysis* in order to emphasize the additional process of generating a rationale for how that behavior came about.

overly restrictive, and 2) the use of global state descriptions to represent temporal behavior introduces irrelevant details.

3 Incorporating a Spectrum of Qualitative Representations into Analysis

Earlier work in qualitative reasoning has focused on the properties of particular qualitative representations for values that have proven useful in engineering practice and common sense experience. Representations explored recently include the direction of change in a quantity's value (e.g., increasing, decreasing or constant), the relative position of a quantity with respect to a set of distinguished values (e.g., for the mass and spring: compressed, at rest position, extended) and contiguous regions of state space where the quantity lies (e.g., for an MOS transistor: on, saturated, unsaturated and off). In addition the interaction during the analysis process between these representations and physical principles (e.g., continuity and feedback) have been explored. The utility of these representations have been demonstrated for a broad set of domains including analog bipolar circuits, performance digital MOS circuits, mechanical and fluid systems and common sense reasoning.

Each of these representations has demonstrated its importance in developing a theory of qualitative reasoning; however, they are not enough alone. Instead we must adopt a broader perspective on qualitative reasoning, simultaneously pushing our representations in two directions. In one direction we must explore more abstract representations. For example, in order to understand the operation of a complex device, such as the MOS clock driver in figure 1, we need to combine discrete and dense representations for values (e.g., representing voltages both as discrete boolean and real numbered values), and then use these representations to reason about a mixture of discontinuous and continuous behaviors. In the opposite direction, we must explore more detailed representations for values. For example rather than just describing a spring and mass as moving back and forth, our problem might require the information that the oscillation is sinusoidal with a specific frequency. Thus our qualitative representations must incorporate more complex notions, such as frequency, amplitude and the shape of the waveform. In addition our analysis must include a quantitative component, allowing it to resolve situations left ambiguous based on the qualitative representations.

The need for a more general set of value representations has been understood by many researchers for quite a while, and many of these representations have been or are beginning to be studied. Our point here is that the new behavior representations and analysis techniques developed must be able to incorporate this variety of value representations.

4 Verbose Behavioral Descriptions are Parasitic

As we will see in a moment, the most serious limitation of existing qualitative analysis techniques is a consequence of the behavioral representations they are based upon. Given one or more qualitative representations for values, we must be able to use it to generate clear, concise descriptions of a system's behavior.

The remainder of this abstract focuses on the temporal aspects of representing behavior. A concise description of behavior, like any good representation, extracts only the properties of interest from the device's behavior, eliminating information considered irrelevant. The behavior of a device is described in terms of a set of features observed over time (e.g., the features of a fluid system might include direction of fluid flow and valve position). A concise behavioral description makes explicit 1) for every feature, a sequence consisting of only those points (i.e., *events*) where that feature changes, and 2) for every event e, a set containing exactly those constraints between other events that are necessary for e to occur. A concise explanation of this behavior has analogous properties discussed later. As we will see in a moment, descriptions generated by earlier qualitative analysis systems are verbose – they introduce additional events and constraints beyond those contained in a concise description. This has serious consequences that are detrimental to the performance of an analysis technique.

Earlier work on qualitative analysis has focused on the representation for values, adopting a simple state-based description of behavior. In this approach the behavior of a device is viewed in terms of a qualitative state diagram, where each state describes the qualitative value of every parameter of the device. The behavior of the device over time is then described as a particular path through this state diagram. Each state along this path represents an interval of time over which the device's parameters maintain their values.

The traditional use of state descriptions in qualitative reasoning is *global*. As a consequence, relationships must be specified between events that don't interact. This has severe consequences that become evident even in very simple systems like the dual spring/mass system shown in figure 2. The two mass/spring sets are completely isolated from each other by placing them in separate boxes that are separated by an ocean.

A description of the system's behavior using global states is also shown in figure 2. This description has several serious problems – each a consequence of the description not being concise. First, the description introduces several irrelevant distinctions. Because the description is global, we must specify the value of *every* variable at every point in time. For example, figure 2 corresponds to the case where the left mass-spring system is oscillating at exactly four times the frequency of the right system and where both systems began in the same initial state. However, this information was not in the original problem statement. Furthermore, the information about frequency and initial state is completely irrelevant; since the masses can't interact, frequency and initial state of one spring/mass system does not affect the other system's behavior.

Second, the introduction of irrelevant orderings produces a profusion of states. We could have concisely described the behavior of the two masses, using two state diagrams similar to the one shown in figure 3. This would result in a total of 16 states. Instead the global state description required 48 states. This situation becomes significantly worse as the systems under analysis become more complex.

Third, if a particular ordering cannot be determined (as it is often the case in qualitative reasoning), we must split cases, quickly leading to an explosion in the number of possible descriptions for the system's behavior. For example, there is an *infinite* number of global state descriptions for the two spring mass systems, since their frequency of oscillation could be anything. Thus a small number of irrelevant orderings can incur a tremendous cost.

5 Value and Justification Episodes: the Core of Concise Descriptions

Given the limitations described above, our first goal is to develop a behavioral representation that doesn't force us to introduce needless details through irrelevant constraints. Instead the representation should allow us to generate concise descriptions AND concise explanations by only introducing constraints on a piece of behavior when the behavior is a consequence of those constraints. Our second goal is to develop analysis techniques that predict a system's behavior in terms of these representations. Like, earlier techniques based on global state descriptions our new approach must generate two types of behavioral descriptions. The first is a description of the system in a particular situation and is the analog of a sequence of states in the earlier approach. The second is a complete description of a system's behavior and is analogous to a state diagram (which can be viewed as a compact description of all possible state sequences).

The root of our problem is at the core of the state-based descriptions — the notion of a state as global. Thus the key to developing a concise behavioral representation is to replace the notion of a global state with one that avoids these limitations. Recall that the source of our problem is that a global state requires us to specify the value of every property in the system. The system is then in a particular state during the interval over which nothing changes. Instead we only want to partially specify the state of the system. To accomplish this we begin with a more primitive notion, called an *episode*. There are two types of episodes: 1) a *value episode* which is used to construct a description of a system's behavior, and 2) a *justification episode* which is used to explain how the behavior came about.

A value episode, specifies the value of only one property in the system. The property is then in a particular episode during the interval over which its value doesn't change. More specifically, an episode of a property, specifies a value of that property and an interval (called the *extent*) over which that value holds. The end points of the extent denote points where the value changes, thus the extent of the episode is a maximal interval of uniform behavior. For example the statement "The cup is above the floor from time t1 to t2 (but not immediately before or after)" is captured by an episode with property CUP-POSITION, value ABOVE-FLOOR, and extent (t1,t2).

The second component of a behavioral description is an explanation for how each part of the system's behavior came about. This is accomplished by constructing a rationale for each value episode. The rationale justifies a property having a particular value throughout the episode. The reason why a property has some value may change during the extent of an episode, (even though the value doesn't). For example, if I pick the cup up off the floor and put it on the table, then part of the time the reason the cup is off the floor is that I'm holding it and part of the time it is because the cup is supported by the table. In addition there may be several reasons for the property having that value. For example, when I'm setting the cup on the table, it is supported both by my hand and the table. The rationale for a value episode is composed of a set of justification episodes, where each justification episode corresponds to one reason explaining the value episode for part of its extent. Each justification episode has an extent corresponding to the largest interval over which the reason holds. This extent is a subset of the value episode's extent being justified. In addition, the extents of a set of justification episodes together cover the extent of the value episode which they justify. Thus the behavior described by each value episode is justified throughout its extent.

Justification episodes also provide the link that interrelates the behaviors of different properties; this link provides the relevant orderings between events. Like most other systems which record explanations, each justification includes a consequent, informant and a set of antecedents. For a justification episode, the consequent is the value episode it supports, the informant is often a rule of inference or an equation used to deduce the consequent and the antecedents is a set of exactly those value episodes necessary for the deduction to follow. Typically a reason holds as long as all its antecedents; i.e., the extent of a justification episode is constrained to be the intersection of the antecedent value episode extents. Of course this intersection must be non-empty in order for the justification to hold at all. A justification episode provides the means of specifying the relevant temporal relationships between value episodes, since the extent of a justification episode is related to both its antecedents and consequents. The relationships are relevant since they are necessary for either the consequent piece of behavior to occur or for that piece of behavior to have a particular rationale. To summarize, value and justification episodes provide us with the necessary primitives to construct pieces of behaviors and explanations, while only introducing relevant interactions.

6 Concise Descriptions of Behavior

Next we provide a whirlwind tour of the behavioral representations composed from these primitives, and the analysis techniques that use them. The presentation is an extremely simplified version of the approach being developed.

The behavior of a system in a particular situation is described using episodes in a manner analogous to a state sequence. A linear history is constructed for each of the system's properties, where a linear history consists of a contiguous, non-overlapping sequence of value episodes. The points where each pair of episodes meet defines the events (i.e., changes in values), while the justification episodes define the relevant partial ordering between these events. The complete description of a system's behavior using episodes is described in a manner analogous to state diagrams. Histories are generalized by allowing them to conditionally branch (e.g., as in a branching future), reconverge and loop. For lack of a better term we call these *non-linear histories.* Each of these representations provides the desired behavioral descriptions and explanations, avoids the limitations of global states discussed earlier, and will support any of a variety of value representations.

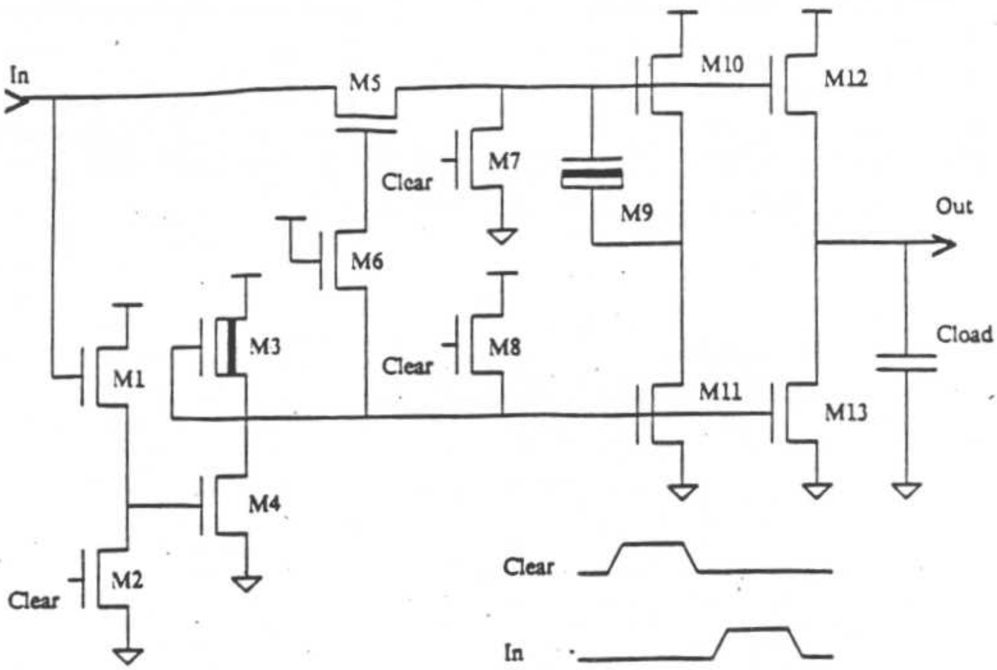
7 Predicting Behavior

A system is described by a set of structural equations on the system's parameters that reflect the local interactions of the system under analysis. These equations are typically constructed from a process or device-centered description of the system. The behavior of a system is predicted based on the system's structural equations and a set of inputs. The equations are used to define constraints between time varying parameters in the system and need not be algebraic. The inputs can be episodes or histories, both linear and non-linear.

Consider the process of predicting behavior when the inputs are single episodes. An equation is used to determine new value episodes from those already known, where all the episodes involved are related to the parameters in the equation. This occurs in two stages; first an equation is applied to a set of value episodes and a resulting justification episode is constructed. Next the justification episode is incorporated into a value episode (possibly with other justification episodes). This process is repeated with the newly deduced value episodes and can be viewed as a propagation of episodes through constraints.

If the inputs are histories (linear or non-linear) then we can view the prediction task as applying equations to construct justification histories from a set of value histories, and then summarizing these justification histories into new value histories. This process, called partial-state envisionment involves a number of subtleties; for example, applying an equation to a set of linear histories may result in non-linear histories, and feedback in the system may produce cycles in a set of justifications. These subtleties are beyond the scope of this abstract; however, many of those related to linear histories are described in a previous publication by the author.

The process of generating behavioral descriptions in terms of non-linear histories involves a sophisticated problem solver including an equality system, context mechanism and a system for solving systems of inequalities with a restricted form of disjunction. This problem solver will be the focus of the extended version of this paper.



- Initially CLEAR produces a step which precharges the gates of M11 and M13 high, turning them on and holding the output at ground.

- When IN starts to rise, it charges capacitor M9 through M5 and starts to turn M10 and M12 on. M6 isolates node 18, which allows that node to bootstrap with M5's gate capacitance, keeping M5 turned on hard. M1 and M4 form a comparator that notices when IN has gone above 2 threshold drops. When this happens, M4 turns on and pulls nodes 12 and 18 down to ground.

- When node 18 discharges, M5 turns off, isolating node 16. Also, when node 13 discharges, M11 and M13, which had been holding down nodes 13 and 14, now turn off, letting those nodes rise.

- Capacitor M9 then bootstraps node 16 (which was isolated by M5 when M5's gate fell), turning M10 and M12 on hard. M12 pulls the output node voltage up. IN can now fall without affecting the rest of the circuit because M5 is off.

- When CLEAR rises again, M11 and M13 turn on and M10 and M12 turn off, forcing the output low and resetting the circuit. The bootstrapping capacitor M9 is driven from node 13 and not from node 14 to get more gate drive on M12 which significantly improves the output rise time.

Figure 1: Generating explanations of complex behavior, like the above explanation for an MOS clockdriver, requires a fundamentally new type of qualitative analysis.

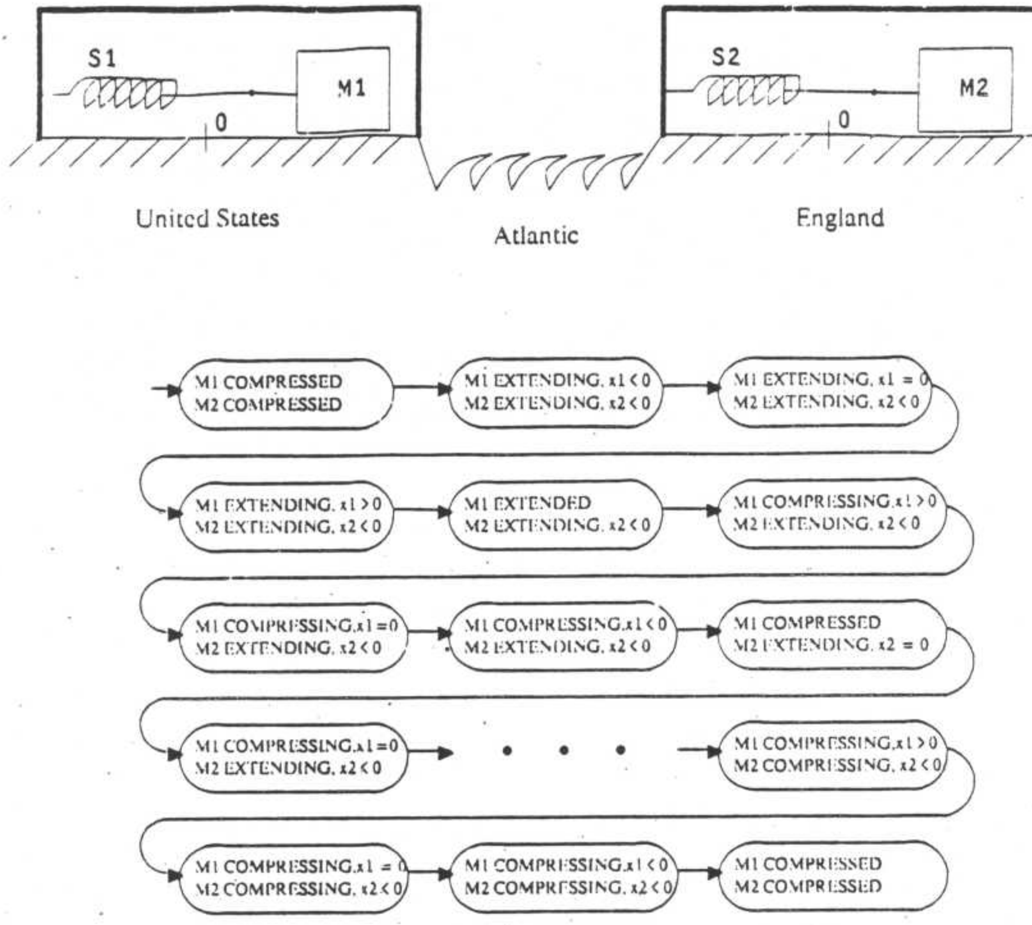


Figure 2: Using global states to describe behavior introduces many irrelevant details. This has serious consequences even for very simple devices like a dual spring/mass system.

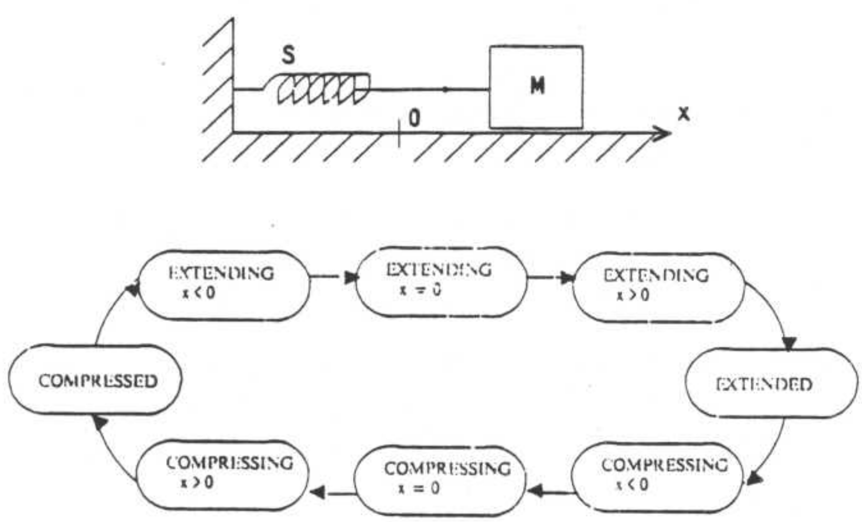


Figure 3: An explosion in the number of states is avoided by describing the dual mass/spring system with two independent partial state diagrams like the one shown above.