KUIPERS

Pushing the edge of the (QP) envelope

by

Kenneth D. Forbus

Qualitative Reasoning Group
Beckman Institute, University of Illinois
405 N. Mathews Avenue
Urbana, Illinois, 61801, USA
Internet: forbus@p.cs.uiuc.edu

Abstract

Qualitative Process theory has been successfully used to model a variety of domains and in exploring several qualitative reasoning tasks. However, there has been little explicit analysis of where and how it breaks down. The purpose of this paper is to look at several ways in which QP theory (or implementations of it) currently fail in some way. Two concepts for classifying QP models are introduced and used to clarify some assumptions underlying all existing qualitative simulators. The relationship between QP theory, QPE, and human intuition is shown to be more complicated than previously suspected. Finally, some limitations of QP theory relative to other systems of qualitative physics are explored. While the analysis focuses on QP theory, many of the issues and suggested research directions are applicable to other systems of qualitative physics as well.

Submitted to QPW-89

1 Introduction

Qualitative Process theory [9,10] was developed to represent commonsense knowledge of the physical world, from the models of the person on the street to the tacit knowledge that guides engineers and scientists in more precise analyses. It has been used to model a variety of physical phenomena [10,3,2], and to study qualitative simulation[13], measurement interpretation [12,4], planning [17,14], learning [7], and textbook problem solving [23]. These studies have given us some idea of what QP theory can be used to model and the kinds of reasoning it can be used to perform. However, boundaries are more crisply drawn by failures than successes. This paper attempts to highlight some areas where QP theory (or current implementations of it) fails. Some of these failures are easily repaired, some aren't.

This analysis attempts to shed light on the following important questions:

1. *Theory versus simulation:* How closely can simulators implement QP theory? Kuipers has provided an elegant analysis of this question for QSIM [19], and a similar analysis for QP theory would help those using it.

2. *Power and limitations:* Most analyses of qualitative physics have focused on mathematical aspects [19,24,22]. Yet ontological issues are just as important, albeit more difficult to analyze. Where does QP theory fail, or provide less plausible models than alternative theories?

3. *Modeling physical intuition:* One goal of qualitative physics is to provide a formal language for human *mental models* [15,18]. How well does the space of models expressible in QP theory fit the space of human mental models?

This paper provides at least some initial insights into these issues. While clearly more remains to be done, there are already some interesting results that merit discussion.

Section 2 defines two concepts for classifying QP models, and uses them to clarify some implicit assumptions underlying qualitative simulation. Section 3 explores the relationship between QP theory, simulators based on it, and human intuition. Ideally, one might hope that any model written in a qualitative physics would exactly match some human intuition, or at least be a subset of human intuition. I demonstrate that neither of these is the case for QP theory. In fact, even the relationship of QP theory to QPE, an envisioner for QP theory, is shown not to be a simple subset-superset relationship. Section 4 explores some limitations of QP theory compared to other systems of qualitative physics. Finally, Section 5 suggests some new directions for qualitative representation and reasoning based on these analyses.

2 A classification of QP models

Recall that in QP theory a modeler creates a *domain model*, which describes classes of objects, relationships, and processes that characterize the dynamics of some class of physical systems. The model for a particular system (a *scenario model*) is constructed automatically by the QP simulator, by instantiating the constructs of the domain model on the scenario's structural description. Much of the power of QP theory comes from putting more of the scenario modeling burden on the QP simulator. However, the use of explicit quantification (in the logical sense) in domain descriptions makes the domain modeler's job more complex. For example, as shown below QP theory allows scenario models which no existing qualitative simulator can run. Here we introduce some concepts for characterizing domain models in order to make these limitations explicit.

(A preliminary assumption: I assume the domain model consists only of processes and views. Most real QP models also include some implementation-specific rules, which install relationships that universally hold. All such rules could be rewritten as views, albeit with some loss in efficiency, so for this analysis they will be ignored.)

2.1 Groundedness

A crucial step in analyzing any scenario model is determining what processes and views are active. Recall that the collection of processes active at some time is the *process structure* of the scenario model at that time, and similarly the collection of active views is its *view structure*. The process structure indicates "what is happening", while the view structure indicates which time-varying relationships (such as object existence and connectivity) hold. The QP simulator builds the scenario model by building view and process *instances*, based on the generic descriptions of views and processes in the domain model. Ascertaining whether or not these instances are `active` or `inactive` (i.e., determining their *status*) requires knowing whether or not their *preconditions* and *quantity conditions* hold. An instance is active when both the preconditions and quantity conditions are satisfied, and inactive otherwise. Quantity conditions are a conjunction of inequalities or statuses of other instances. Preconditions are a conjunction of statements other than inequalities or status instances.

Computing the status of instances is a basic inference required of QP systems. The *groundedness* of a model pertains to what kinds of information are required to fix the status of each process and view instance. By analogy with linear algebra, this information can be thought of as the "basis set" for the qualitative state space. Let us then define the *basis set* of a scenario model to be a set of statements whose truth values always suffice to fix the statuses of all view and process instances. Any qualitative simulator makes assumptions about what this basis set contains. In an ATMS-based envisioner, for example, elements of the basis set are generally assumptions, combined using interpretation construction to explicitly build the state space. Unlike linear algebra, the basis sets for qualitative state spaces are rarely independent. The constraints of the physics define relationships between these assumptions, often subtly, which can drastically reduce their "span" (in this case, the size of the state space). Let a *minimal basis set* be a basis set such that no subset of it is also a basis set. While a minimal basis typically cannot be computed in advance, it is to one's advantage in simulation to pick the smallest adequate basis set, since the amount of computational work depends strongly on its size. There could in fact be several minimal basis sets, due to interrelationships between assumptions.

(One complication is that the basis set as defined above does not generate the entire state space – it must be extended as analysis proceeds to include ordering relationships required to resolve ambiguities (see [13] for details). Nevertheless, this notion of basis set is useful for making certain distinctions involving domain models.)

A scenario model is *inequality-grounded* if it has a minimal basis set consisting only of ordering relationships between pairs of numbers. This is the simplest case, corresponding to a purely dynamical model. A scenario model is *precondition-grounded* if it has a minimal basis set consisting only of statements from the preconditions of the view and process instances, and is not inequality-grounded. The last stipulation is necessary because statements used as preconditions can also appear in the `Relations` fields of views and processes. Without this stipulation, one could construct QP models that were both inequality-grounded and precondition-grounded, by creating a one-to-one mapping between a minimal inequality basis and a minimal precondition basis.

Figure 1: An ungrounded QP model

```
defView Yin(?x)
 Individuals: ?x, World(?x)
              ?y, view-instance(?y) ∧ instance-of(?y,Yang)
 QuantityConditions: ¬ Active(?y)

defView Yang(x)
 Individuals: ?x, World(?x)
              ?y, view-instance(?y) ∧ instance-of(?y,Yin)
 QuantityConditions: ¬ Active(?y)
```

Precondition-grounded models can still have dynamical transitions, since Ds values can change, but the view and process structures are fixed. Both kinds of models are easily expressed in QP theory. A scenario model is *mixed-grounded* if every minimal basis set includes both ordering relationships between pairs of numbers and some statements from the preconditions of its view and process instances. Most QP models are mixed-grounded.

It is possible to write QP domain models which are *ungrounded*: That is, for some scenario, no combination of inequalities or preconditions suffices to determine the view and process structures. Figure 1 illustrates. This is a consequence of allowing statuses of view and process instances to be used as quantity conditions. This choice made sense because (1) in typical domain models inequalities suffices to determine the status of most instances, and (2) it substantially increases modularity. For example, one can represent the thermal effects of fluid mixing as a separate process, predicated on both the desire to model thermal properties and a fluid flow being active, without being forced to duplicate every precondition and quantity condition of the flow process in the mixing process. This ability is vital to control the instantiation of large-scale domain models [8].

Every QP implementation I have seen has assumed that scenario models are at worst mixed-grounded. To simulate ungrounded models requires changing the definition of precondition, so that status assumptions used as preconditions could be added to the basis set. That by itself isn't complicated, but unfortunately it doesn't address the real issue. In all existing QP simulators, neither `Yin` nor `Yang` of Figure 1 will be instantiated. The reason is that processes and views cannot be instantiated with unknown individuals. Suppose we have a scenario description consisting of the statement `World(W)`. Then there are three models (in the model-theoretic, as opposed to the engineering, sense of models):

`World(W)`	`World(W)`	`World(W)`
	`Active(Yin(W))`	`¬ Active(Yin(W))`
	`¬ Active(Yang(W))`	`Active(Yang(W))`

Existing qualitative simulators will only find the first of these.

The implicit assumption in all QP implementations is that the views and processes of a domain are "well-founded", in some sense. We can define this more clearly as follows. Consider a scenario model constructed using the domain model. Let *Individuals*(pv) be the set of entities filling the roles defined in the `Individuals` field of the process or view of which pv is an instance. Furthermore, let *Individuals*(i) $= i$ for i other than view or process instances. Let *Isupport*(i) be the transitive closure of *Individuals*(i). The scenario model is grounded exactly when $i \notin$ *Isupport*(i) for all of its individuals i.

How can we tell if a domain model is grounded, given this definition for grounding of scenario models? What we would like is a method for constructing scenario models from the domain model itself, such that if these scenario models are grounded, then every scenario model will be. The following approach looks promising. Let a *Herbrand scenario* for a domain model consist of a set of skolem constants and statements about them such that the scenario model resulting from expanding the statements using the domain model includes at least one instantiation of every process and view from the domain model. For example, if a domain model included the typical definition of heat flow a Herbrand scenario for it might include the skolem constants `hf-src`, `hf-dst`, and `hf-path`, with the following facts known about them:

```
Quantity(Heat(hf-src))
Quantity(Heat(hf-dst))
Heat-Path(hf-path)
Heat-Connection(hf-path,hf-src,hf-dst)
```

Call a Herbrand scenario *minimal* when no proper subset of it is itself a Herbrand scenario. It seems clear (but remains to be proven) that if all minimal Herbrand scenarios for a domain model are grounded, then every scenario model that can be generated from that domain model will also be grounded.

If this conjecture is true, then an algorithm for detecting ungrounded aspects of domain models could be built as follows. Generate an initial Herbrand scenario by creating a skolem constant for each role of every view and process instance in the domain model, using the restrictions in the `Individuals` field to generate the necessary facts relating them. Then search for the minimal Herbrand models by finding the maximal sets of equality relations between the skolem constants (i.e., how many of them in fact could be standing for a single entity).

2.2 Creativity

A distinguishing feature of QP theory from other systems of qualitative physics is its ability to reason about changes in existence. It does this by declaring the existence of individuals within the scope of the `Relations` field for some view or process C. When C is active the individual will exist, and C becoming inactive heralds the end of the individual's existence (unless some other view or process instance maintains it). One can tie existence to a view directly, to indicate that an individual exists only when that view is active. This is accomplished by the predicate `There-is-unique`, which is defined as:

$$\forall I \in \textit{individuals}\ \forall C \in \textit{processviewinstances}$$
$$\texttt{There-is-unique}(I) \in \texttt{Relations}(C) \Rightarrow$$
$$\forall \texttt{t} \in \textit{times}\ \texttt{T[Exists}(I)\texttt{,t]} \Leftrightarrow \texttt{T[Active}(C)\texttt{,t]}$$

Figure 2 illustrates a typical example. We call domain models which contain such views or processes *creative.*

Recall that a scenario model is constructed by instantiating views and processes on some initial scenario description. We assume the initial scenario description explicitly describes only a finite number of individuals. We can further subdivide creative vocabularies into *bounded* versus *unbounded* models. In domain models of bounded creativity every scenario model contains only a finite number of individuals. In domain models of unbounded creativity there can be scenario models containing an infinite number of individuals. As

Figure 2: A creative QP description

This individual view defines the conditions for the existence of contained stuff (a generalization of Hayes' contained liquid ontology). A domain model which contains views or processes which define the existence of new individuals is called *creative*.

```
defview Contained-Stuff-Existence(?sub,?st,?c)
  Individuals: ?sub, Substance(?sub)
               ?st, State(?st)
               ?c, Container(?c)
  Preconditions: Can-Contain-Stuff(?sub,?st,?c)
  Quantity Conditions: A[Amount-of-in(?sub,?st,?c)] > ZERO
  Relations: There-is-unique(C-S(?sub,?st,?c))
             Contained-Stuff(C-S(?sub,?st,?c))
             Amount-of(C-S(?sub,?st,?c)) = Amount-of-in(C-S(?sub,?st,?c))
```

might be expected, qualitative simulators require models of bounded creativity, since they explicitly generate all objects. Figure 3 illustrates a view exhibiting unbounded creativity.

Some models with unbounded creativity represent bugs in the domain model. Consider a detailed model of forces, as might be found in a qualitative dynamics for motion. Each action, as every schoolchild knows, results in an equal and opposite reaction. We can encode this principle by creating a *reflection force* for every applied force. Notice that if this reflection force is itself considered an applied force, we are in trouble. (I have seen physics students get stuck this way sometimes, although they tend to recover better than current qualitative simulators.)

Unfortunately, not all such models represent blunders. Many phenomena are most naturally described by models with unbounded creativity. Consider a crumbling substance, such as chalk, which can decompose into several pieces, each of which in turn are pieces of a crumbling substance. Or water waves in a tank, where collisions with the sides of the tank cause yet more waves. A radioactive source can generate an arbitrary number of alpha particles. In physical terms of course none of these cases yield an infinite number of individuals – crumble a substance finely enough and one gets atoms, waves in a tank disperse, and the radioactive source transmutes into a more stable combination of elements. This does not mean that we can ignore unbounded creativity in models: many intuitive models have this character, and even if there are limits, the number of individuals involved

Figure 3: A view exhibiting unbounded creativity

```
defview Has-Constituent-Parts(?obj)
    Individuals: ?obj, Physob(?obj)
    Preconditions: Divisible-Object(?obj)
    Relations: Part-of(Part1(?obj),?obj)
               Part-of(Part2(?obj),?obj)
               Physob(Part1(?obj))
               Physob(Part2(?obj))
               Divisible-Object(Part1(?obj))
               Divisible-Object(Part2(?obj))
               Amount-of(?obj) = Amount-of(Part1(?obj)) + Amount-of(Part2(?obj))
```

quickly becomes unmanageable.

A subtle point: Instances of views and processes are themselves considered to be individuals. Consequently, instantiating views and processes creates new individuals, which were not part of the initial scenario description. The definition of creativity ignores these individuals, since then any non-trivial QP model would be creative. Accidentally specifying domain models that create unbounded numbers of instances constitutes a problem for modelers. I have not yet found an example where models involving only unbounded view and/or process instances and a finite number of other types of individuals are intuitively plausible.

2.3 Implications for qualitative simulation

All existing qualitative simulators, not just implementations of QP theory, assume grounded, bounded creativity models. So far only QP theory allows creative models at all; QSIM does not explicitly represent individuals, and device-centered models start with a fixed schematic. It is not clear whether or not reasoning with ungrounded models is desirable, since the ones that crop up while trying to model physical phenomena typically have other undesirable features as well. However, automatically detecting ungrounded aspects of domain models will become more important as we attempt to scale up to higher-fidelity domain models.

Moving to unbounded creative models is very hard (Section 4.1 shows just how hard). The only existing reasoning technique which might be applied to this problem is aggregation [25], assuming one could abstract away from individuals to a set of continuous parameters in each case. While aggregation can probably work in cases where individuals are generated incrementally over time (such as figuring out why champagne goes flat), it offers no help for cases where an infinite number of individuals exist all at once (such as divisible objects, which are divisible even if no one is doing the dividing).

I believe in such cases the answer is to eschew simulation, or at least combine simulation with a style of qualitative analysis that works directly on quantified descriptions. Some combination of consequent reasoning and resource limitations (such as ONTIC's bound on the number of binding premises [21]) seems to me to be a promising approach for controlling inference in such models. Efficient reasoning about quantified knowledge invariably requires a scheme for controlling computational resources. Most resource-limitation schemes, such as depth of inference or number of nodes, or time, do not map naturally onto the constraints of intuitive arguments. McAllester argues that mathematical arguments are restricted to a limited number of individuals at any time, but allow extremely complex reasoning about those individuals. I suggest the same is true of intuitive arguments about physical phenomena. Intuitive arguments appear to only require a finite number of explicitly named individuals at a time. The number can be kept small by using multiple ontologies and abstractions to control the level of detail used to represent each aspect of a problem. So computational schemes which express resource bounds as limitations on the number of introduced individuals seem to be a reasonable substrate for qualitative physics.

3 Relating QP theory, QPE, and intuition

The relationship between a theory and programs which implement it is rarely straightforward. The numerous engineering decisions required to create a working artifact often compromise the fidelity with which an implementation follows a theory. Evaluating a theory requires testing how well it explains and predicts its subject phenomena, and when some of these predictions are made by an implementation this relationship takes on new importance. Part of the motivation behind QP theory was to provide a formal language for expressing mental models. Detailed psychological studies are of course required to ascertain how well QP models can explain human reasoning. However, a simple theoretical analysis can reveal some interesting limitations.

Consider the following three sets of models:

Intuit : { Models corresponding to our physical intuitions }

Theory : { Models expressible in QP theory}

Sim : { Models that can be simulated using QPE}

This is one instantiation of a general scientific schemata. What a theorist hopes for is $Theory = Intuit$, and feels satisfied if $Theory \subset Intuit$, assuming *Theory* covers "enough of" *Intuit*. Similarly, implementers are happy to the degree that $Sim = Theory$. Often an implementation might have limitations which could be overcome with larger machines, or simply more coding effort (such as QPE's current inability to handle user-defined explicit sets), but we ignore such accidents here.

Unfortunately, for QP theory it can be demonstrated that none of *Intuit*, *Theory*, or *Sim* subsumes the other. We proceed by showing for each pair of sets that one contains something the other doesn't.

$Intuit \not\subset Theory$ This is the easiest. Mental models research indicates that animism and discrete processes often play roles in intuitive models of physical phenomena, both of which lie outside the bailiwick of QP theory.

$Theory \not\subset Intuit$ Some cynics might think this case is difficult, because at first glance it seems one can find people who believe just about anything. Here is an example, however, which I claim no human being would find intuitive for very long:

```
defview Universal-Flow-Rates(?flow1,?flow2)
   Individuals: ?flow1, Process-Instance(?flow1), Instance-of(?flow1,Liquid-Flow)
                ?flow2, Process-Instance(?flow2), Instance-of(?flow2,Liquid-Flow)
   Quantity Conditions: Active(?flow1) ∧ Active(?flow2)
   Relations: Am[Flow-Rate(?flow1)] = Am[Flow-Rate(?flow2)]
```

This description says that every two liquid flows involve flow rates of the same magnitude. That is, the rate of water flowing out of my faucet is the same as water leaking from a cracked class, which is the same as the rate at which a hummingbird sucks nectar from a flower, which is the same rate as water pours over Niagara Falls, since they are all active at the same time. Such a model grossly violates our notions of locality and causality, but it is a perfectly legal QP model.

$Intuit \not\subset Sim$ QPE cannot reason about the motivations of agents.

$Sim \not\subset Intuit$ The universal flow rate model isn't creative and has mixed grounding, and hence is easily simulated using QPE.

$Theory \not\subset Sim$ Unbounded creative models, as described in Section 2, cannot be simulated by QPE.

Sim⊄theory Programs can take on a life of their own. In my work on action-augmented envisionments [14], I define how actions can be integrated into a qualitative physics, thus providing a superset of QP theory. In implementing this superset theory, QPE was endowed with the ability to simulate STRIPS-style actions. This turned out to be surprisingly simple, involving only a few changes in the temporal inheritance algorithm. The result is that the modeling language supported by QPE is now powerful enough to encode and simulate standard blocksworld problems, hardly a domain suitable for pure QP theory[1].

This analysis suggests two conclusions. The first, that new qualitative reasoning techniques are required to handle models with unbounded creativity, echos a theme of the previous section. The second is that to fit human intuitions, QP models must satisfy some other set of constraints, which have yet to be formalized. It is not clear yet whether or not this body of constraints consists of just a handful of simple laws or is itself a deep, systematic theory. The only qualitative physics research I know of which sheds some light on these issues is Doyle's exploration of the role of connectivity in causality [6] and Bylander's use of paths in consolidation [1]. Each is attempting to formalize some aspect of these deep physical intuitions, although their intent is to produce a formalism that can be used without qualitative mathematics. I doubt that some form of qualitative mathematics can be avoided given the need to compose effects. Consequently these approaches seem unlikely to by themselves yield adequate accounts of qualitative physics. However, it would be interesting to see if their ideas could be extended and re-cast as constraints on the formulation of QP models. They might provide exactly the restrictions needed to produce only intuitively compelling models.

4 Some QP Conundrums

How powerful is QP theory? One way to answer this question is to look for ultimate limits in what can be expressed, and another is to compare it to other systems of qualitative physics. This section makes some observations about both.

4.1 A limit on qualitative simulation

Qualitative simulation tends to be thought of as exponential in the worst case, since each state can have more than one potential transition to other states. Envisionments are finite by definition, while behavior generation[2], because of its ability to arbitrarily increase resolution on numerical values, can result in infinite behaviors. These results are true for theories other than QP theory, and hold for QP models of bounded creativity. Unfortunately, models with unbounded creativity throw a monkey wrench into the works.

It is generally assumed that all systems of qualitative physics are powerful enough, in principle, to implement detailed models of digital computers. Although no one has done so, this assumption appears very reasonable. This means, of course, that we could build a Turing machine. In other systems of qualitative physics making an infinite tape is something of a problem, since it is assumed (often implicitly) that the structural description is finite. It is easy to build an infinite tape in QP theory, as Figure 4 illustrates. Given a single square of a tape to start with, this description implies the existence of an infinite tape. And now we are in trouble. Simulation with QP models of unbounded creativity is in

[1] Winslett [27] has in fact recently proposed an abstract formulation for reasoning about actions which in essence generalizes this temporal inheritance algorithm.

[2] Formerly "history generation". Dan Weld suggested this change in terminology, since histories necessarily imply a spatial aspect, while some qualitative simulators (e.g., QSIM) don't require this.

Figure 4: Building an infinite tape in QP theory
The view **Right-Square** creates a square to the right of any tape square. The view **Extended-Left** helps rule out finite models.

```
defView Right-Square(?S,Right-of(?S))
 Individuals: ?S,Tape-Square(?S)
 Relations: Tape-Square(Right-of(?S))
            Left-of(Right-of(?S)) = ?S
            Right-of(?S) ≠ ?S
            Left(?S,Right-of(?S))

defView Extended-Left(?S1,?S2)
 Individuals: ?S1,Tape-Square(?S1)
              ?S2,Tape-Square(?S2)
              ?S3,Tape-Square(?S3), Left(?S1,?S3), Left(?S3,?S2)
 Relations: Left(?S1,?S2)
            ?S1 ≠ ?S2
```

the worst case undecidable. If it wasn't, we could solve the halting problem by simulating the Turing machine, for any input, in finite time.

Just as the existence of the halting problem has not stopped work on automated program understanding, this result does not mean that qualitative simulation is "doomed". It just shows that qualitative simulation of unbounded models is fundamentally harder than on bounded models, and thus we are unlikely to find a simple universally applicable solution anytime soon.

4.2 Modeling causality accurately

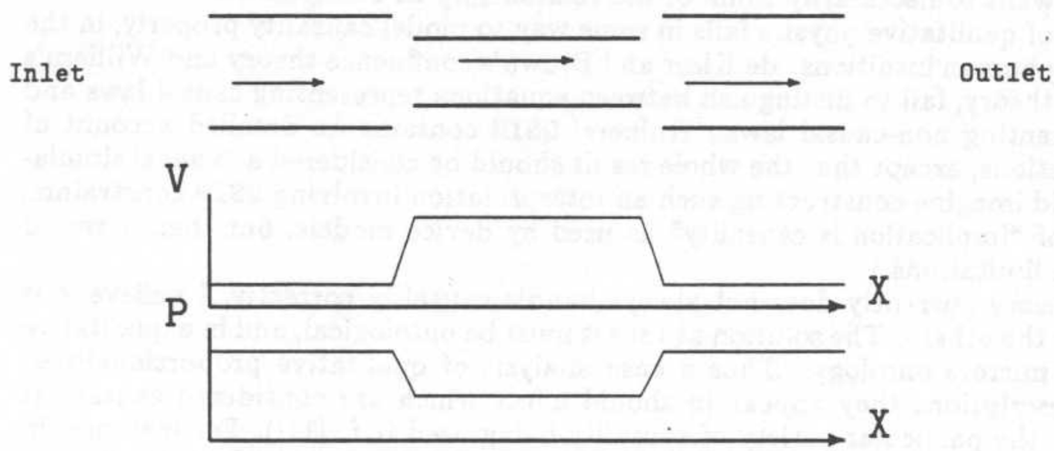
Consider a liquid being pumped around a fluid system in steady state. As it enters a constriction, its velocity increases and its pressure decreases (see Figure 5). This example of the Bernoulli effect can be understood in terms of the following argument. Since the system is in steady-state, the amount of fluid passing by each slice of the system must remain constant, and since liquids are incompressible, it must move faster in the constriction to maintain the requisite volume flow. Since kinetic energy depends on velocity, and energy is conserved in this system (ignoring friction), then some other component of energy must drop, and in fact pressure drops.

Conservation arguments are powerful, of course, but there is a deeply-held intuition that for every physical phenomena there must be some causal explanation. I have not come across one for the Bernoulli effect that is completely satisfactory, nor to my knowledge have others. Yet in the sense of causality defined by QP theory there is a very simple causal model, shown in Figure 6.

This view defines the relationships between the parameters of a liquid flowing in a section of pipe. One may quibble about whether or not contained liquids can have a velocity (one could use molecular collections as easily [3]), but the important part is the **Relations** field. Qualitative proportionalities force a causal interpretation (c.f. Section 3.2 of [10]), so by stipulation we have constructed a causal model of the Bernoulli effect.

This model is certainly correct, in the sense that using it in a comparative analysis would yield correct results (i.e., that liquid will speed up in a constriction while the pressure will drop, and that both effects are caused by the difference in the diameter of the pipe). Why am I unsatisfied? Informal observations (and some protocol analyses) by Dedre Gentner

Figure 5: What is a causal explanation for the Bernoulli Effect?
This diagram shows a piping section from a fluid system, within which liquid is flowing in steady-state. As the graphs indicate, velocity increases in a constriction while pressure drops. The causal explanations usually stipulated by qualitative physics appear unintuitive for this example.



indicate that when naive subjects are first introduced to the Bernoulli effect they find it extremely counterintuitive. We do not know if this is due to lack of familiarity with this causal model or lack of acceptance of it. Sophisticated subjects have different problems with this model. Given that they know macroscopic properties are derived from microscopic, they tend to view the microscopic as "more real", and thus insist on a causal story in molecular terms. One can feel fairly comfortable with this model by viewing qualitative proportionalities as summaries of causal arguments expressed in finer-grained ontologies (in this case, perhaps, involving statistics on the behavior of molecular collections), much as Kuipers uses M to summarize the effects of a fast process when reasoning about a slower one [20]. However, carrying through this derivation remains something of a challenge[3]

[3] Gregg Collins (personal communication) uses it as an interesting puzzle.

Figure 6: A causal model for the Bernoulli Effect

```
DefView Fluid-Properties-Flowing-In-Pipe(?pipe,?w,?lf)
   Individuals: ?pipe, Pipe-Section(?pipe)
                ?w, Contained-Liquid(?w) ∧ Container(?w) = ?pipe
                ?lf, Process-Instance(?lf)
                     ∧ Instance-of(?lf,Liquid-Flow)
                     ∧ ?pipe ∈ Path(?lf)
Quantity Conditions: Active(?lf)
Relations: Velocity(?w) ∝Q- diameter(?pipe)
           Pressure(?w) ∝Q+ diameter(?pipe)
```

This example points up a broader issue, of how causality is modeled in qualitative physics. There are cases where one really wants to separate causality from equations. It is certainly the case that

```
Area(?rectangle) ∝Q+  Length(?rectangle)
Area(?rectangle) ∝Q+  Width(?rectangle)
```

but one doesn't want to necessarily think of the relationship as being causal.

Each system of qualitative physics fails in some way to model causality properly, in the sense of violating human intuitions. de Kleer and Brown's confluence theory and William's system, like QP theory, fail to distinguish between equations representing causal laws and equations representing non-causal laws. Kuipers' QSIM contains no detailed account of causal interpretations, except that the whole result should be considered a "causal simulation". (One could imagine constructing such an interpretation involving QSIM constraints, along the lines of "implication is causality" as used by device models, but then it would inherit the same limitations.)

While QP theory currently does not always handle causality correctly, I believe it is easier to fix than the others. The solution as I see it must be ontological, and in a qualitative physics, syntax mirrors ontology. Thus a case analysis of qualitative proportionalities, based on the descriptions they appear in should affect which are considered causal. It must also reflect the particular variety of causality being used (c.f. [11]). For instance, in continuous causality direct influences are interpreted causally, and only those qualitative proportionalities propagating the effects of direct influences are considered causal. For differential causality it seems safe to attribute causality to all qualitative proportionalities. Whether or not such a case analysis can be consistently completed, and whether it fits with human intuition, is an open question at this writing.

4.3 Analyzing complex structural descriptions

Complexity in physical systems can arise from several sources. The components themselves can be complex, as in automobiles. Complexity can also arise from the way simple parts are connected together, as in VLSI circuits. Representing systems with complex connectivity remains somewhat problematic in QP theory.

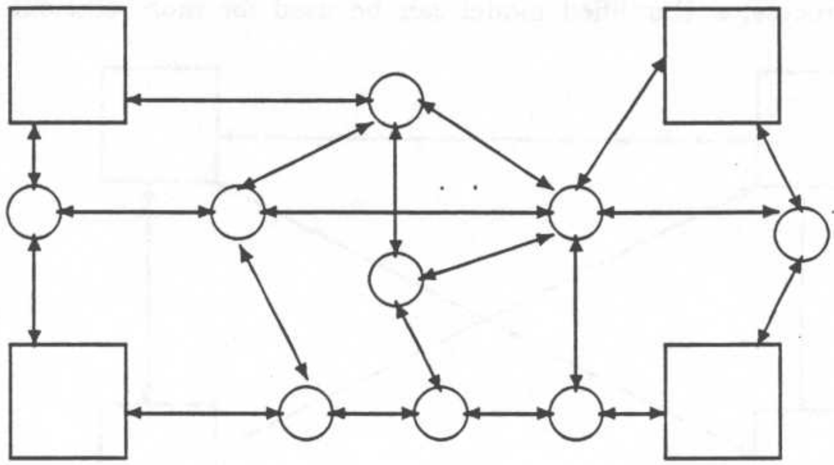
Device-centered theories handle complex connections reasonably well. Confluences representing compatibility (e.g., Kirchoff's Voltage Law) are introduced for every connected triple of conduits, and a confluence representing continuity of stuff (i.e., Kirchoff's Current Law) is introduced for every component [5]. These local laws suffice for simple networks. For more complex networks containing redundant nodes (such as Figure 7), extra continuity laws are sometimes required in addition to purely local laws in order to rule out impossible behaviors. In ENVISION the addition of these extra confluences was controlled by a switch. Few device models used in practice actually contain redundant nodes, though, since the modelers typically simplify the structural description by hand. (One case where redundant nodes have been used is for studying diagnosis, to model shorted components[4].)

Some aspects of these constraints are easy to enforce in QP theory. In models where pressures at nodes are explicitly represented and compared to ascertain flow, compatibility constraints are automatically enforced by ruling out violations of transitivity[5]. The real problem concerns how nodes should be represented. Device models tend to be sketchy when representing the "stuff" that flows through the network. Aside from a generalized

[4] J. de Kleer, personal communication

[5] Barry Smith, unpublished manuscript.

Figure 7: Complex networks can require extra analysis
In this diagram the boxes correspond to components containing significant fluid volumes, while the small circles indicate nodes in the piping system which connects



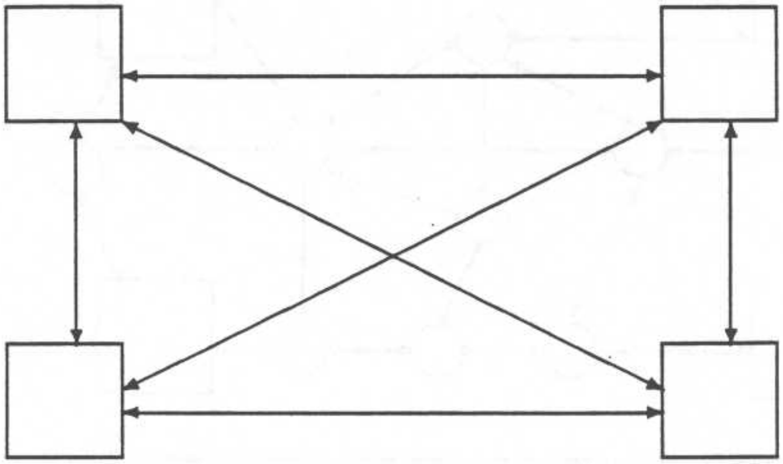
them.

pressure and current, other physical properties, including existence, are not made explicit. Adding the kinds of details that are desirable for many purposes (such as thermal properties of fluids) raises new problems. Suppose we represent nodes as little containers, holding whatever kind of stuff is in the parts they are connected to. With this representation enforcing material continuity is straightforward; one simply constrains the sum of the flow rates involving the node to be zero. The easiest way to do this is simply to assert that the amount of stuff in the node never changes. Call this the *fixed node* strategy.

The fixed node strategy is fine for systems where the stuff flowing is uniform and exists in all parts of the system for all time. But what about systems which can be emptied, like steam plants? Moving an operating steam plant to "cold iron" involves emptying almost every container and piping system. An intuitive representation of liquids requires that there be *some* stuff at a node, and thus when emptying the system that stuff has to vanish. Yet it cannot if we model its amount as constant! Conversely, when we fire up the steam plant we must "fill up" each and every node in the piping system, introducing many extra state transitions in the simulation. The more realistic the model, the worse the situation gets. Modeling thermal properties of fluids, for instance, requires capturing the thermal effects of mixing. To be consistent we must then consider how the temperature changes in each node as it fills, and so forth. Taking flows seriously as being about physical "stuffs" seems to require a fundamentally different approach.

One way Falkenhainer and I have been exploring to overcome this limitation is to automate the simplification of the structural description, much as a good engineer would. The system in Figure 7, for example, could be re-represented as a set of simple fluid paths between the components as in Figure 8. The properties of these fluid paths must be computed with regard to the properties of their constituents, of course – they are aligned only when some path in the original network between the components is aligned, and the fluid resistance must be calculated for the aligned part using the laws of fluid networks. Call this the *node removal* strategy. (This strategy is one example of the problem of translating from structural descriptions to *structural abstractions*, which is a generally ignored but

Figure 8: Structural transformations can simplify simulation
By making the translation of structural descriptions (i.e., what you might see on an engineering diagram) to structural abstractions (i.e., the constructs of the physics) part of the modeling process, a simplified model can be used for more efficient simulation.



critical aspect of the modeling process [8].)

How effective is the node removal strategy? Consider a network with no redundant nodes. In general, if a node connects K components, then it can be replaced by $\frac{K(K-1)}{2}$ bidirectional fluid paths. If there are M nodes in the network, all connecting K components, then we introduce at worst $M\frac{K(K-1)}{2}$ fluid paths, and no new contained stuffs. With the fixed node strategy, we still have MK fluid paths, since each connection to a node corresponds to a fluid path. But the fixed node strategy also entails M new fluid individuals, with associated parameters and dynamical transitions. Depending on how loosely coupled the parts of the system are, these extra transitions can grow the envisionment by a multiplicative factor (by the arguments in [16,26]). By contrast, the extra analysis of the node removal strategy is done once, in the initial construction of the scenario model. Thus the overall complexity of the node removal strategy appears better for this case.

When the possibility of redundant nodes is considered, node removal looks even better. Consider N components connected by a set of M nodes, each of connectivity K, such that at least one path exists between each pair of components. In the fixed-node strategy, there will be M new individuals corresponding to the fluids in the nodes as well as between $\frac{MK-N}{2}+N$ and MK fluid paths linking them. In the node removal strategy only $N\frac{(N-1)}{2}$ paths are needed, since the network is equivalent to one with a single node, and it takes that many paths to replace an N-terminal node. While the complexity of the initial analysis depends on both M and K (since we have to find what paths exist between components and record their dependencies), the size of the final result is independent of them. As M and K increase the node removal strategy looks better and better, since the number of individuals (and hence the number of envisionment states) remains unchanged, while it continues to grow for the fixed-node strategy.

I have not yet implemented the analysis required for the node removal strategy, but it seems straightforward. For explanation generation and diagnosis it is important to maintain dependencies between the original structural description and the structural abstraction(s) arising from it. Clearly, maintaining knowledge of valves and alignment is always important, as is knowing what kind of stuff is flowing through each pipe (in order to understand leaks) But just how much information should be kept for different tasks is an unexplored issue at this point. For example, Pat Hayes has pointed out[6] that the node removal strategy ignores the possibility of signficant flows within cycles in the fluid paths themselves, which can cause problems in real life (e.g., plumbing systems). But suppose we keep track of cycles discovered during the computation of the simplified flow network. Given the qualitative states computed for the simplified diagram, we could then analyze each cycle to see what flow patterns could arise within them (using the molecular collection ontology).

It is important to note that device-centered theories could be extended to take advantage of these same techniques. By adapting QP-like descriptions of individuals, preconditions, and quantity conditions, for example, the conventions for structuring large-scale models worked out by Falkenhainer and I could be applied to the device ontology. Network simplification could (and should) be done automatically for complex systems in both ontologies.

5 Summary and afterthoughts

The issues discussed here arose during the investigations our group has made using QP theory over the last five years. They constitute tricky terrain, relatively unexplored territory where the correct path is not understood. Some are peculiar to QP theory, others haunt everyone in qualitative physics. This paper has attempted to clarify the dividing line between what we can now do and what we can't, by "pushing the edge of the envelope", taking QP theory into places it wasn't clearly meant to go. This section suggests some directions for future research based on these analyses.

The distinctions concerning domain models point out a fundamental limitation of today's qualitative simulation technology (i.e., requiring grounded models of bounded creativity). The potential undecidability of such models suggests searching for efficient heuristic methods to incrementally control the introduction of new individuals. Qualitative analysis tools that focus on analyses more sophisticated than simulation need to be explored.

I also believe we must find better ways to test our domain models. Knowing that a domain model works on one's standard set of cases is both necessary and reassuring. But it is still a hit-or-miss proposition: One can make the catalog large and varied enough that it is likely to catch many problems with a domain model, but cannot ever know for sure that one has caught them all. I suspect that the Herbrand scenario construction outlined here, where the domain model itself is used to generate scenarios to be analyzed, could be useful for several aspects of verifying domain models.

The fact that QP theory can build unintuitive models points out the need for additional theoretical constraints when using it as a formal representation for mental models. The most obvious constraints are those concerning connectivity – things which aren't connected physically shouldn't be causally connected. This restriction is implicit in the specifications found in the `Individuals` field of reasonable QP models, but it needs to be formalized. How many other implicit constraints are there? Right now the only way of finding out is to search for bizarre cases and finding constraints that rule them out. The consistent attribution of causality is crucial for many applications as well as psychological modeling, since bizarre explanations are unlikely to be believed. This may require breaking the

[6] Personal communication.

definitional link between qualitative proportionalities and causality, equating the two only when justified by ontological analysis.

Finally, the analysis of structural descriptions suggests that a crucial part of the next generation of qualitative analysis tools should be the ability to control the translation from structural descriptions to structural abstractions. Our qualitative analysis systems should take advantage of structural re-writes and equivalences just as engineers do. Whether device-centered or process-centered, explicit use of structural transformations and equivalences will be vital to scaling up qualitative physics to handle real systems.

6 Acknowledgements

This paper benefited from discussions with Brian Falkenhainer, Pat Hayes, Johan de Kleer, and C.L. Liu. It was made more readable by feedback from John Collins, Dennis Decoste, and Gordon Skorstad. This research was supported by the Office of Naval Research, Contract No. N00014-85-K-0225, an NSF Presidential Young Investigator award, and an equipment grant from IBM.

References

[1] Bylander, T. and Chandrasekaren, B. "Understanding behavior using consolidation" Proceedings of IJCAI-85, August, 1985.

[2] Collins, J. "Building qualitative models of thermodynamic processes", Proceedings of QPW-89, August, 1989.

[3] Collins, J. and Forbus, K. "Reasoning about fluids via molecular collections", Proceedings of AAAI-87, July, 1987.

[4] Decoste, D. "Dynamic across-time measurement interpretation", Proc. of the 3rd Qualitative Physics Workshop, Palo Alto, August, 1989.

[5] de Kleer, J. and Brown, J. "A qualitative physics based on confluences", *Artificial Intelligence*, **24**, 1984

[6] Doyle, R. "Constructing and refining causal explanations from an inconsistent domain theory" Proceedings of AAAI-86, August, 1986.

[7] Falkenhainer, B. "An examination of the third state in the analogy process: Verification-based analogical learning" Proceedings of IJCAI-87, August, 1987.

[8] Falkenhainer, B., and Forbus, K. "Setting up large-scale qualitative models", Proceedings of AAAI-88, August, 1988.

[9] Forbus, K. "Qualitative reasoning about physical processes" IJCAI-7, Vancouver, B.C., August, 1981.

[10] Forbus, K. "Qualitative process theory" *Artificial Intelligence*, **24**, 1984

[11] Forbus, K. and Gentner, D. "Causal reasoning about quantities", Proceedings of the Eighth annual conference of the Cognitive Science Society, Amherst, Mass., August, 1986

[12] Forbus, K. "Interpreting observations of physical systems", *IEEE transactions on systems, man, and cybernetics*, Vol. SMC-17, No. 3, May/June 1987.

[13] Forbus, K. "QPE: A study in assumption-based truth maintenance" *International Journal of Artificial Intelligence in Engineering*, October, 1988.

[14] Forbus, K. "Introducing Actions into Qualitative Simulation", Report No. UIUCDCS-R-88-1452, Department of Computer Science, University of Illinois, August, 1988.

[15] Gentner, D., Stevens, A. (Eds.), *Mental Models*, LEA Associates, Inc., 1983.

[16] Hayes, P. "The naive physics manifesto" in *Expert systems in the micro-electronic age*, D. Michie (Ed.), Edinburgh University Press, 1979.

[17] Hogge, J. "Compiling plan operators from domains expressed in Qualitative Process theory", Proceedings of AAAI-87, July, 1987.

[18] Kuipers, B. and Kassirer, J. "Causal reasoning in medicine: Analysis of a protocol" *Cognitive Science*, **8**, pp 363-385, 1984.

[19] Kuipers, B. "Qualitative Simulation", *Artificial Intelligence*, **29**, September, 1986.

[20] Kuipers, B. "Abstraction by time-scale in qualitative simulation", Proceedings of AAAI-87, July, 1987.

[21] McAllester, D. ONTIC: *A knowledge representation system for mathematics*, MIT Press, Cambridge, Mass. 1989, in press.

[22] Sacks, E. "Qualitative Mathematical Reasoning", Proceedings of IJCAI-85, August, 1985.

[23] Skorstad, G. and Forbus, K. "Qualitative and quantitative reasoning about thermodynamics", Proceedings of CogSci-89, August, 1989.

[24] Struss, Peter "The limitations of qualitative mathematics", presented at the first Qualitative Physics Workshop, Urbana, Illinois, May, 1987.

[25] Weld, D. "The use of aggregation in qualitative simulation", *Artificial Intelligence*, 30(1), October 1986.

[26] Williams, B. "Doing time: Putting qualitative reasoning on firmer ground", Proceedings of AAAI-86, August, 1986

[27] Winslett, M. "Reasoning about action using a possible models approach" Proceedings of AAAI-88, August, 1988.