MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

A.I. Memo No. 664 A

February 1982
Revised, May 1983

Qualitative Process Theory

Kenneth D. Forbus

Abstract

Things move, collide, flow, bend, heat up, cool down, stretch, break, and boil. These and other things that happen to cause changes in objects over time are intuitively characterized as processes. To understand common sense physical reasoning and make machines that interact significantly with the physical world we must understand qualitative reasoning about processes, their effects, and their limits. Qualitative Process theory defines a simple notion of physical process that appears quite useful as a language in which to write physical theories. Reasoning about processes also motivates a new qualitative representation for quantity, the Quantity Space. This paper includes the basic definitions of Qualitative Process theory, describes several different kinds of reasoning that can be performed with them, and discusses its implications for causal reasoning. The use of the theory is illustrated by several examples, including figuring out that a boiler can blow up, that an oscillator with friction will eventually stop, and how to say that you can pull with a string, but not push with it.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research Contract number N00014-80-C-0505.

CONTENTS

FIGURES

1. Introduction

Many kinds of changes occur in physical situations. Things move, collide, flow, bend, heat up, cool down, stretch, break, and boil. These and the other things that happen to cause changes in objects over time are intuitively characterized as *processes*. Much of formal physics consists of characterizations of processes by differential equations that describe how the parameters of objects change over time. But the notion of process is richer and more structured than this. We often reach conclusions about physical processes based on very little information. For example, we know that if we heat water in a sealed container the water can eventually boil, and if we continue to do so the container can explode. To understand common sense physical reasoning we must understand how to reason qualitatively about processes, their effects, and their limits. This paper describes a theory I have been developing, called *Qualitative Process theory*, for this purpose.

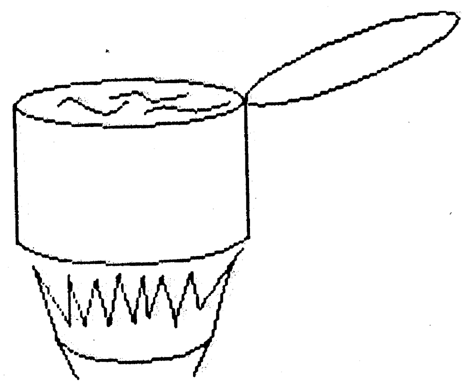
In addition to providing a major part of the representational framework for common sense physical reasoning, I expect Qualitative Process theory to be useful in reasoning about complex physical systems. Programs that explain, repair and operate complex systems such as nuclear power plants and steam machinery will need to draw the kinds of conclusions discussed here. Figure 1 illustrates some of the commonsense conclusions about physical situations that will be discussed in this paper.

Qualitative reasoning about quantities is a problem that plagues AI. Many schemes have been tried, including simple symbolic vocabularies (TALL, VERY TALL, etc.), real numbers, intervals, and fuzzy logic. None are very satisfying. The reason is that none of the above schemes makes distinctions that are relevant to physical reasoning. Reasoning about processes provides a strong constraint on the choice of representation for quantities. Processes usually start and stop when orderings between quantities change (such as unequal temperatures causing a heat flow). In Qualitative Process theory the value of quantities are represented by a partial ordering of other quantities determined by the domain physics called the *Quantity Space*. The Quantity Space representation appears both useful and natural in modeling a wide range of physical phenomena.

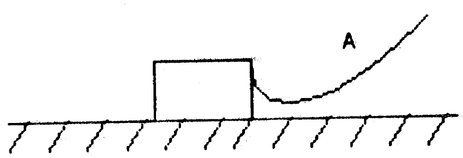
1.1 Motivation

The goal of Naive Physics [Hayes, 1979a] is to represent the commonsense knowledge people have about the physical world. Here we will examine why a theory of processes is needed, what representational burden it will carry in Naive Physics, and the properties such a theory must have.

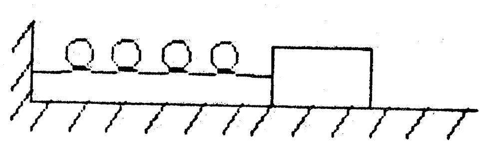
Fig. 1. Some Examples of QP theory Conclusions
Here is a sample of the kinds of conclusions QP theory will be used to draw.



Q: What might happen when the heat source is turned on?
A: The water inside might boil, and if the container is sealed it might blow up.



Q: Can we push the block with A if it is a string?
A: No, but you can pull the block if it is taut.
Q: Assuming A is an elastic band and the block is fixed in position, what might happen if we pull on it?
A: It would stretch and if pulled hard enough would break.



Q: What happens if we release the block?
A: Assuming the spring doesn't collapse, the block will oscillate back and forth. If there is friction it will eventually stop.
Q: What if it gets pumped?
A: If there is no friction the spring will eventually break. If there is friction and the pumping energy is constant then there will be a stable oscillation.

1.1.1 Change, Histories, and Processes

Reasoning about the physical world requires reasoning about the kinds of changes that occur and their effects. The classic problem which arises is the Frame Problem[McCarthy & Hayes, 1969], namely when something happens, how do we tell what facts change and what facts don't? Using the Situational Calculus to represent the changing states of the world requires writing explicit axioms that state what things change and what things remain the same. The number of axioms needed rises as the product of the number of predicates and the number of actions, and so adding a new action requires adding a large number of new axioms. There have been several attempts to fix this problem [Fikes & Nilsson, 1971][Minsky, 1974], but none of them have seemed fruitful. Hayes [Hayes, 1979a] argues cogently that the Situational Calculus is fundamentally impoverished, and has developed the notion of Histories as an alternative.

In Situational Calculus, situations are used to model the world at different instants in time. Temporally each situation is an instant, but is spatially unbounded. Situations are connected by actions, and actions are specified in terms of what facts can be deduced about the situation which results from performing the action. By contrast, Histories are descriptions of objects that are extended through time but always are bounded spatially. Histories are divided into pieces called *episodes*, corresponding to a particular kind of thing happening to the object.

Histories help solve the Frame Problem because objects can interact only when their histories intersect. For example, suppose we are building a clock in our basement. In testing parts of this gadget we look to see what parts touch each other, what parts will touch each other if they move in certain ways, and so on. By doing so we build descriptions of what can happen to the pieces of the clock. We do not usually consider interactions with the furnace sitting in the corner of the basement, because whatever is happening in there is spatially isolated from us (if it is summer it can also be "temporally isolated").

The assumption that things interact only when they touch in some way also permeates "non-Naive" physics - action at a distance is banished, with fields and particle exchanges introduced to prevent its return. It means that spatial and temporal representations bear most of the burden for detecting interactions. While not simple, developing such representations seems far more productive than trying to develop more clever frame axioms.[1] In particular, the qualitative representations of space and time developed in AI have precisely the desired properties for reasoning with histories -- they often allow ruling out interactions even with very little information.

Histories are to qualitative physical reasoning what descriptions of state parameters over time are to classical numerical simulations. Processes are the analog of the differential equations used to describe the dynamics of the system.

While the classical Frame Problem is solved, two new problems arise to take its place.

- The *Local Evolution* problem: How are histories generated? Under what circumstances can they be generated for pieces of a situation independently, and then pieced together to describe the whole situation?

In the basement example above, for instance, we could safely ignore the furnace in the corner and concentrate on figuring out how the pieces of the clock we are building will move. The divisions are only semi-independent, because certain kinds of changes can violate the conditions for isolation. For example, if the internal thermostat of the furnace gets stuck and it explodes, we can no longer safely ignore it.[2]

- The *Intersection/Interaction* problem: Which intersections of histories actually correspond to interactions between the objects?

Dropping a large steel ball through a flame, for example, won't affect its motion even if the flame is hot enough to melt it unless the gases are going fast enough to impart significant momentum. Solving these problems in general requires knowing what kinds of things can happen and how they can affect each other - in other words, a theory of processes.

In classical mechanics *dynamics* describes how forces bring about changes in physical systems.

---

1. For an example of Histories in use, see [Forbus, 1981a] which describes a program called FROB that reasons about motion through space. FROB used a diagram to compute qualitative spatial representations which were used to rule out potential collisions between objects as well as describing possible motions.
2. Unless the physical situation is simulated by some incremental time scheme, the reasoning involved in extending histories will be inherently "non-monotonic"[McDermott & Doyle, 1980]. The reason is that conclusions reached by considering one part of a system may have to be reconsidered in the light of unexpected interactions. In incremental time simulations the changes in the entire system are computed over a very short timespan, and then the system is tested to see if any new interactions occur (such as objects colliding). The timespan is usually chosen to be small enough that interactions during a step can be ignored. The cost is that the work required to simulate a system is a function of the time scale rather than the actual complexity of the system's behavior.

For any particular domain, such as particles or fluids, a dynamics consists of identifying the kinds of forces that act between the classes of objects in the domain and the events that result from these forces. In general, we can view a *qualitative dynamics* as a qualitative theory about the kinds of things that "can happen" in a physical situation. Qualitative Process theory claims that such theories have a common character, in that they are organized around the notion of *physical processes.*

1.1.2 Reasoning Tasks Involving Qualitative Dynamics

Aside from the basic role of dynamics in representing change, there are a number of reasoning tasks involving Naive Physics in which dynamics is central. Each of them can be viewed as a different "style" of reasoning, appropriate for solving different classes of problems. The catalog below, while surely incomplete, seems to cover a large proportion of the cases. Examples of inferences from several of these categories will be presented later.

Determining Activity: Deducing "what is happening" in a situation at a particular time. Besides providing direct answers to a class of questions ("what is happening here?"), it is also a basic operation in the other reasoning tasks.

Prediction: Deducing what will happen in the future of some situation. We usually must work with incomplete information, so we can only generate the possibilities for what might occur. de Kleer's notion of *envisioning* is a powerful theory about this type of deduction.[1]

Postdiction: Deducing how a particular state of affairs might have come about. Hayes [Hayes, 1979b] contains a good example of this kind of deduction. Postdiction is harder than prediction because of the potential necessity of postulating individuals. If we have complete knowledge of a situation and have a complete dynamics, we know what individuals will vanish and appear. But usually there are many ways for any particular situation to have come about. Consider walking back to our basement and finding a small pile of broken glass on the floor. Looking at it we may deduce that a coke bottle was dropped, but we do not know much about its history before that, or about anything else that might have been in the room before we looked. There could have been a troupe of jugglers filling the basement, each manipulating six bottles, and a minor mishap occurred. The simplest explanation is that a single bottle

---

1. Useful as it is, envisioning has certain limitations, especially as a sufficient model of human behavior on this task. See [Forbus, 1983a] for details.

was dropped, but our criteria for simplicity is not due solely to our theories of physics. Postdiction will not be considered further here.

Skeptic: Determing if the description of a physical situation is consistent. An example of this task is evaluating a proposed perpetual motion machine. This kind of reasoning is essential if we want to recover from inconsistent data and discover inadequacies in our theories about the world.

Measurement Interpretation: Given a partial description of the individuals in the situation and some observations of their behavior, inferring what other individuals exist and what else is happening.[1] The first part of a QP-based theory of measurement interpretation is described in [Forbus, 1983b].

Experiment Planning: Given knowledge of what can be observed and what can be manipulated, planning actions that will yield more information about the situation.

Causal Reasoning: Computing a description of behavior that attributes changes to particular parts of the situation and particular other changes. Not all physical reasoning is causal, especially as more expert kinds of deductions are considered.[2] Causality seems mainly a tool for assigning credit to hypotheses for observed or postulated behavior. Thus it is quite useful for generating explanations, measurement interpretation, planning experiments, and learning.

1.1.3 Desiderata for Qualitative Dynamics Theories

There are three properties a theory of dynamics must have if it is to be useful for commonsense physical reasoning. First, a dynamics theory must explicitly *specify direct effects* and *specify the means by which effects are propagated.* Without specifying what can happen and how the things that happen can interact, there is no hope of solving either the Local Evolution or Intersection/Interaction problems. Second, the descriptions the theory provides must be *composable.* It should be possible to describe a very complicated situation by describing its parts and how they relate.[3] This property is especially important as we move towards a more complete Naive Physics that encompasses several "domains". In dealing with

---

1. [Simmons, 1982] explores the the related problem of reconstructing a sequence of events from a static final state.
2. The experienced elegance of constraint arguments in fact argues against their being central in Naive Physics. Usually some kind of animistic explanation is proposed to justify constraint arguments to non-experts ("the particle senses which path has the least action").
3. Producing models with this property is a motivation for the "No Function in Structure" principle [de Kleer & Brown, 1983].

a single kind of reasoning in a particular class of situations an ad hoc domain representation may suffice, but sadly the world does not consist of completely separate domains. Transferring results between several ad hoc representations may be far more complex than developing a useful common form for dynamics theories.[1] Finally, the theory should allow *graceful extension.* First, it should be possible to draw at least the same conclusions with more precise data as can be drawn with weak data. Second, it should be possible to resolve the ambiguities that arise from weak data with more precise information.

These properties are not independent -- for example, specifying direct and indirect effects cleanly is necessary to insure composability. Nevertheless, they are not easy to achieve. Graceful Extension is bound up with the notion of good qualitative representations. Qualitative representations allow the construction of descriptions that include the possibilities inherent in incomplete information. If designed properly, more precise information can be used to decide between these alternatives as well as perform more sophisticated analyses. Representing quantities by symbols like TALL and VERY-TALL or free space by a uniform grid, for instance, does not allow more precise information to be easily integrated. Note also that while qualitative descriptions are approximations, not all approximations are good qualitative descriptions. Changing a value in a qualitative representation should lead to qualitatively distinct behavior. Consider, for example, heating a pan of water on a stove. Suppose that we represent the value of the temperature of the water at any time by an interval, and the initial temperature is represented by the interval [70.0 80.0], indicating that its actual temperature is somewhere between 70 and 80 degrees farenheit. Changing the "value" of its temperature to [75.0 85.0] doesn't change our description of what's happening to it (namely, a heat flow), whereas changing it to [200.0 220.0] changes what we think can be happening to it -- it could be boiling as well. While an interval representation makes certain distinctions, they usually are <u>not</u> distinctions relevant to physical reasoning. By defining a basic theory using qualitative representations, we can later add theories involving more precise information - perhaps such as intervals - to allow more precise conclusions. In other words, we would like extensions to our basic theory to have the logical character of extension theories - more information should result in a wider class of deductions, not changing details of conclusions previously drawn. In this way we can add theories that capture more sophisticated reasoning (such as an engineer would do when estimating circuit parameters or stresses on a bridge) onto a common base.

---

1. An initial exploration of linking results from reasoning within multiple domains is being carried out by [Stanfill, 1982].

1.2 Perspective

The present theory has evolved from several strands of work in Artificial Intelligence. The first strand is the work on *envisioning*, started by de Kleer [de Kleer, 1975](see also [de Kleer, 1979][Forbus, 1981a]). Envisioning is a particular style of qualitative reasoning. Situations are modeled by collections of objects with *qualitative states*, and what happens in a situation is determined by running simulation rules on the initial qualitative states and analyzing the results. The weak nature of the information means the result is a directed graph of qualitative states that corresponds to the set of all possible sequences of events that can occur from the initial qualitative state. This description itself is enough to answer certain simple questions, and more precise information can be used to determine what will actually happen if so desired.

While a powerful idea, the assumptions of envisioning as it has been developed thus far are too restrictive. The qualitative state representation of what is happening to an object is impoverished; the processes that they represent often involve several objects at once in an interdependent fashion. The use of qualitative simulation rules means that the only time information about events consists of local orderings, making new interactions between things happening in the situation ("collisions") hard to detect. Also, simulation rules are a rather opaque way to encode knowledge about how things can happen in a situation. The rules themselves do not explicitly describe the mechanism by which the state transformation is accomplished, thus making it difficult (or impossible) to reason about changes in the assumptions which underly the rules. Qualitative Process theory should provide the basis for building much more flexible systems.

The second strand of work concerns the representation of quantity. Most AI schemes for qualitative reasoning about quantities violate what I call the *relevance principle* of qualitative reasoning - qualitative reasoning about something continuous requires some kind of quantization to form a discrete set of symbols; the distinctions made by the quantization must be relevant to the kind of reasoning being performed.[1] Almost all previous qualitative representations for quantity violate this principle. One exception is the notion of quantity introduced by de Kleer as part of Incremental Qualitative (IQ) analysis [de Kleer, 1979], which represented quantities according to how they changed when a system input was perturbed - increasing, decreasing, constant, or indeterminate. For more general physical reasoning a

1. For an example of this principle applied to spatial reasoning, see [Forbus, 1981a].

richer theory of quantity is necessary. IQ analysis alone does not allow the limits of processes to be deduced. For instance, IQ analysis can deduce that the water in a kettle on a lit stove would heat up, but not that it would boil. IQ analysis does not represent rates, so we could not deduce that if the fire on the stove were turned down the water would take longer to boil (see Differential Qualitative analysis below). The notion of quantity provided by QP theory is useful for a broader range of inferences about physical situations than the IQ notion.

The final strand relevant to the theory is the Naive Physics enterprise initiated by Pat Hayes [Hayes, 1979a]. The goal of Naive Physics is to develop a formalization of our common sense physical knowledge. From the perspective of Naive Physics, Qualitative Process analysis is a cluster - a collection of knowledge and inference procedures that is sensible to consider as a module. The introduction of explicit processes into the ontology of Naive Physics should prove quite useful. For instance, in the axioms for liquids [Hayes, 1979b] information about processes is encoded in a form very much like the qualitative state idea.[1] This makes it difficult to reason about what happens in situations where more than one process is occurring at once - Hayes' example is pouring water into a leaky tin can. In fact, difficulties encountered in trying to implement a program based on the axioms for liquids were a prime motivation for developing Qualitative Process theory.

1.3 Overview of the Paper

This paper is an expanded treatment of the central ideas of Qualitative Process theory [Forbus, 1981b][Forbus, 1982]. While at this writing the theory is incomplete (notably the notation for abstraction and descriptive hierarchies), other workers in Naive Physics have already found its concepts useful. It is hoped that this exposition will stimulate further work in the area.

The next two sections provide the basic definitions for the qualitative representation for objects, quantities and physical processes. Objects and quantities are discussed first because they are required for the process definitions. The basic deductions sanctioned by the theory are discussed as well, including analyzing the net effects of processes and the limits of their activity. These deductions are then illustrated by several extended examples, including modeling a boiler, motion, materials, and an oscillator. Further implications of Qualitative Process theory, including causal reasoning, are then discussed. Finally the

1. See for example axioms 52 through 62.

theory is placed into the perspective of similar work in Artificial Intelligence.

A word on notation. Axioms are used only when they will help the reader interested in the fine details. Although a full axiomatic description might be desirable, there are a host of complex technical details involved, few of which essentially contribute to understanding the ideas. When used, axioms are written in a more or less standard sorted predicate calculus notation. The following notational conventions will be used in axioms: Predicates and relations will be capitalized (e.g., `FluidConnection`), and functions will be in lower case (e.g., `amount-of`, `made-of`). Sorts will be underlined (e.g., <u>`time`</u>). Individuals (often physical objects) will be capitalized (e.g., `WA`) and variables will be in lower case (e.g., `p`). Small finite sets will be enclosed by braces ("{" "}"). When non-standard notation is introduced an effort will be made to show an interpretation of it in terms of logic. This should not necessarily be taken as an endorsement of logic as "the meaning of" the statements.

At this writing, major parts of the theory have been tested via implementation. The basic deductions sanctioned by the theory (see section 3.6) are coded, but few of the more sophisticated analyses that use these deductions. In particular, the examples presented do not represent the results of a currently running program. Further work on implementation will certainly continue to increase the precision of the theory, but at this writing the basic ideas seem quite stable. This paper will not discuss the implementation at all.

2. Objects and Quantities

To talk about change we first must establish some conventions for describing objects and their properties at various times. In this section we will describe the temporal notation used and then develop the representation of quantity and the *Quantity Space* representation for values. *Individual Views* are then introduced to describe both the contingent existence of objects and object properties that change drastically with time. The idea of a *qualitative proportionality* ($\propto_Q$) is then introduced to describe how quantities relate to each other. Finally *histories* are introduced to represent "what happens" to objects over time.

2.1 Time

The assumptions we will make about time are very simple. We will assume there are *instants* and *intervals*, where intervals have some definite duration and instants have zero duration. A function `time` maps from instants to some (implicit) global ordering, so that we can define ordering relations `before`, `after`, and `simultaneous` corresponding to one instant being less than, greater than, or the same in the global ordering. Intervals will have functions `start` and `end`, which map from intervals to instants. The function `during` maps from an interval to the set of instants that are between its start and end, and the function `duration` describes the difference between the times at the start and the end. [1]

We will use the modal operator T to tie the truth of a statement to a particular time.

```
(T <statement> <time>)
```

reads "<statement> is true at(during) <time>", and its semantics are defined in [Moore, 1979][McDermott, 1981]. An example from the Fluids World (which will be described below) is:

```
(T Aligned(Pipe3) During(Filling(c1)))
```

---

1. In addition, we will take the view of [Allen, 1982], that instants can be viewed as intervals, in that they can have a `start` and an `end`. This property doesn't matter unless we are discussing observability (see [Forbus, 1983b]).

2.2 Quantities

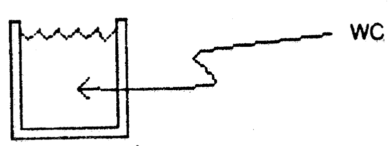
Processes affect objects in various ways. Many of these effects can be modeled by changing <u>parameters</u> of the object, properties whose values are drawn from a continuous range. The representation of a parameter for an object is called a <u>quantity</u>. Examples of parameters that can be represented by quantities include the pressure of a gas inside a container, one dimensional position, the temperature of some fluid, and the magnitude of the net force on an object.

The predicate `Quantity-Type` will be used to indicate that a symbol is used as a function that maps objects to quantities. To say that an object has a quantity we will use the relationship `Has-Quantity`. Figure 2 illustrates some quantities that pertain to the liquid in a cup.

2.3 Parts of Quantities

A quantity consists of two parts, an <u>amount</u> and a <u>derivative</u>. The derivative of a quantity can in turn be the amount of another quantity (for example, the derivative of (one dimensional) position is the amount of (one dimensional) velocity). Amounts and derivatives are <u>numbers</u>, and the functions `A` and `D` map from quantities to amounts and derivatives respectively. Every number has distinguished parts <u>sign</u> and <u>magnitude</u>. The functions `s` and `m` map from numbers to signs and magnitudes respectively. For conciseness, the combinations of these functions that select parts of quantities will be noted as:

Fig. 2. Quantities
Quantities represent continuous parameters of objects. Here are some quantities that are used in representing the liquid in the cup below.




```
Quantity-Type(Amount-of)
Quantity-Type(Level)
Quantity-Type(Pressure)
Quantity-Type(Volume)

Has-Quantity(WC, Amount-of)
Has-Quantity(WC, Level)
Has-Quantity(WC, Pressure)
Has-Quantity(WC, Volume)
```

A_m - "magnitude of the amount"
A_s - "sign of the amount"
D_m - "magnitude of the derivative", or "rate"
D_s - "sign of the derivative"

Numbers, magnitudes, and signs take on values at particular times. When we wish to refer to the value of a number or part of a number, we will write:

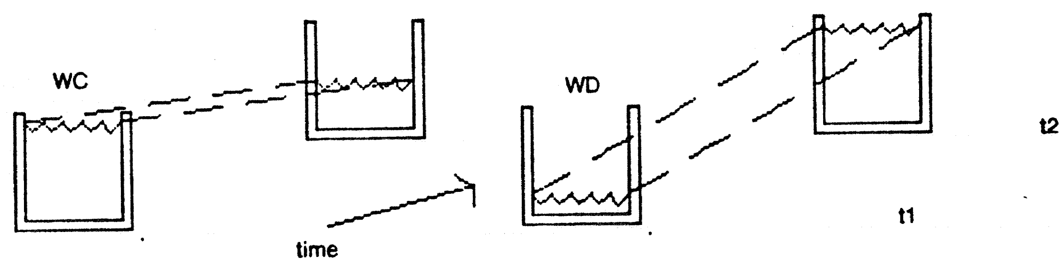
```
(M Q t)
```

which means "the value of Q measured at t". Of course,

```
(T (<quantity or part> <relationship> <quantity or part>) <time>) ⇒
(M <quantity or part> <time>) <relationship> (M <quantity or part> <time>)
```

Signs can take on the values -1, 0, and 1. Figure 3 provides an example of the use of M. We will take elements of ℝ as our model for the values of numbers and magnitudes so that operations of comparison and combination are well defined. Note however that in basic Qualitative Process theory we will never know numerical values. What we do know about values is described next.

Fig. 3. M Describes Values at Different Times
Here are some facts about the two containers below expressed as relationships between their quantities:



(M A[level(WC)] t1) > (M A[level(WC)] t2)

(M A[level(WC)] t1) = (M A[level(WD)] t2)

(M Ds[level(WC)] t1) = -1

(M Ds[level(WD)] t1) = 1

2.4 Quantity Space

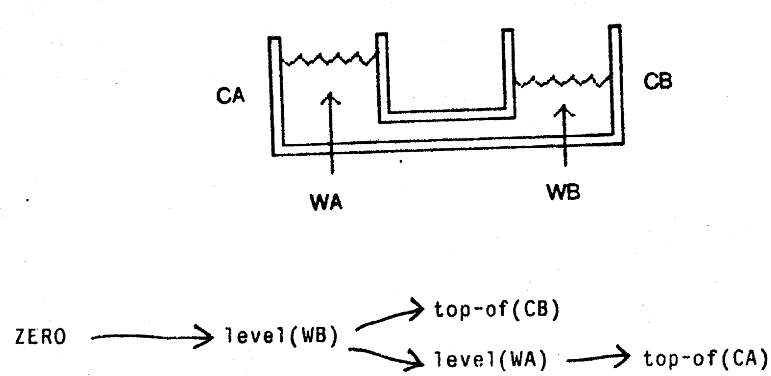
The value of a number is defined in terms of its *Quantity Space.* A Quantity Space is a collection of numbers which form a partial order. (Sometimes we will find it easier to speak loosely of the value of a quantity, rather than of a number. In this case the value is the value of the quantity's amount.) Figure 4 illustrates a Quantity Space for the levels of fluid in two tanks A and B connected by a pipe. Note that the orderings and even the elements of a Quantity Space will not be fixed over time. The elements in a particular Quantity Space are determined by the comparisons needed to establish certain kinds of facts, such as whether or not processes are acting (Shortly, we will see another kind of description that contributes elements to Quantity Spaces). This means there will only be a finite number of elements in any reasonable Quantity Space, hence there are only a finite number of distinguishable values. Thus the Quantity Space is a good symbolic description.

We shall now be a bit more formal about defining Quantity Spaces and the relationships between parts of quantities. Readers who don't wish to be bothered by the details may skip to the next section.

The Quantity Space of a quantity will consist of a set of elements (numbers or magnitudes, often the amounts of quantities) N and a set of orderings. The value of a quantity Q will be the ordering relations between Q and the other elements in the Quantity Space. The value is completely specified if the ordering between Q and every other element in N is known, and is incomplete otherwise. Every

---

Fig. 4. Graphical Notation for a Quantity Space
The arrow indicates that the quantity at the head is greater than the quantity at the tail. As drawn, `Level(WA)` and `Top-of(CB)` are unordered. For simplicity, we ignore temporal references here.



Quantity Space can in principle be completely specified. A collection of inequality statements whose union with the orderings of an incompletely specified Quantity Space results in the Quantity Space being completely specified will be called a *completion* of that Quantity Space.

All quantity spaces have the distinguished element ZERO. ZERO serves to connect the sign of a number with inequality statements, as follows:

$$\forall\, q \in \underline{quantity}\ \forall\, t \in \underline{time}$$
$$(M\ A[q]\ t) > ZERO \leftrightarrow (M\ A_s[q]\ t) = 1$$
$$\wedge\ (M\ (A[q]\ t) = ZERO \leftrightarrow (M\ A_s[q]\ t) = 0$$
$$\wedge\ (M\ A[q]\ t) < ZERO \leftrightarrow (M\ A_s[q]\ t) = -1$$

Note also that the values of magnitudes are related to the values of signs and the number, in that:

$$\forall\, n \in \underline{number}\ \forall\, t \in \underline{time}$$
$$Taxonomy((M\ m[n]\ t) > ZERO,\ (M\ m[n]\ t) = ZERO)$$
$$\wedge\ ((M\ m[n]\ t) = ZERO \leftrightarrow (M\ s[n]\ t) = 0)$$

(Taxonomy is drawn from [Hayes, 1979b] and means that exactly one of its arguments is true.) Thus if the value of D_s for some quantity is 0, then the derivative itself is zero and the quantity is unchanging. We will sometimes need to combine sign values across addition and multiplication. Figure 5 illustrates the algebra used. Combining sign values for derivatives of expressions will be described shortly.

Two points that are ordered and with no points in the ordering known to be between them will be called neighbor points. For the Quantity Space in Figure 1, Level(A) has ZERO, Top-of(A), and Level(B) as neighbors, but not Top-of(B). Distinguishing neighboring points will be important in determining when processes start and stop acting.

Fig. 5. Combining Sign Values
This table specifies how sign values combine across addition and multiplication. The cases marked by notes require additional information to determine the result.

For s[A + B]

A \ B	-1	0	1
-1	-1	-1	N1
0	-1	0	1
1	N1	1	1

```
N1: if m(A) > m(B) then s(A)
    if m(A) < m(B) then s(B)
    if m(A) = m(B) then 0
```

For s[A * B]

A \ B	-1	0	1
-1	1	0	-1
0	0	0	0
1	-1	0	1

2.5 Individual Views

Objects can come and go, and their properties can change dramatically. Water can be poured into a cup and then drunk, for example, and a spring can be stretched so far that it breaks. Some of these changes depend on values of quantities - when the amount of a piece of fluid becomes zero we can consider it gone, and when a spring breaks, it does so at a particular length (which may depend on other continuous parameters such as temperature). To model these kinds of changes we will use *Individual Views.*

An Individual View consists of four parts. It must contain a list of *Individuals,* the objects that must exist before it is applicable. It has *Quantity Conditions,* that are statements about inequalities between quantities of the individuals and statements about whether or not certain other Individual Views hold, and *Preconditions* that are still further conditions that must be true for the view to hold, and a collection of *Relations* that are imposed by that view. Figure 6 illustrates a simple description of the fluid in a cup.

For every collection of objects that satisfies the description of the individuals for a particular type of Individual View, there is an *Individual View Instance*, or IVI, that relates them. Whenever the Preconditions and Quantity Conditions for an IVI hold we will say that it is ACTIVE, and INACTIVE otherwise. Whenever an IVI is active the specified Relations hold between its individuals. An Individual View can be thought of as defining a predicate on (or relation between) the individual(s) in the Individuals field, and we will often write them that way. The Contained Liquid description of the previous figure has been translated into logical notation in figure 7 to illustrate.

The distinction between Preconditions and Quantity Conditions is important. The intuition is to separate changes that can be predicted solely within dynamics (Quantity Conditions) from those which cannot (Preconditions). If we know how a quantity is changing (its D_s value) and its value (specified as a Quantity Space), then we can predict how that value will change (as we will see in the next section). We cannot predict within a purely physical theory that someone will walk by a collection of pipes through which fluid is flowing and turn off a valve. Despite its unpredictability, we still want to be able to reason about the effects of such changes when they do occur, hence any dependence on these facts must be

---

Fig. 6. Individual Views Describe Objects and States of Objects
Here is a simple description of the fluid contained in a cup. This description says that whenever there is a container that contains some liquid substance then there is a piece of that kind of stuff in that container. A more elaborate set of descriptions will be developed later on.

```
;we will take "amount-of-in" to map from substances and
;containers to quantities

Individual View Contained-Liquid
        Individuals:
                con a container
                sub a liquid
        Preconditions:
                Can-Contain-Substance(con, sub)
        QuantityConditions:
                A[Amount-of-in(sub, con)] > ZERO
        Relations:
                There is p ∈ piece-of-stuff
                Amount-of(p) = Amount-of-in(sub, con)
                made-of(p) = sub
                container(p) = con
```

Fig. 7. Translation of Individual View Notation into Logic
Here is the contained liquid description of the previous figure translated into logical notation.

```
∀ c ∈ container ∀ s ∈ liquid
        (∃ IV ∈ individual-view-instance
        ;names of individuals are used as selector functions
         con(IV) = c ∧ sub(IV) = s ∧
        ;logical existence of individual is timeless
          (∃ p ∈ piece-of-stuff
                container(p) = c ∧ made-of(p) = s)
        ∧ (∀ t ∈ time
        ;it is active whenever Preconditions and Quantity Conditions hold
          (T Status(IV, Active) t) ↔
           ↔(T Can-Contain-Substance(con(IV), sub(IV)) t)
                ∧ (T A[Amount-of-in(sub(IV), con(IV))] > ZERO t)
        ;when active, p exists physically and its amount is the
        ;amount of that kind of substance in the container
        ∧(T Status(IV, Active) t) ⇒
                (Exists-In(p, t)
                ∧ (M Amount-of(p) t) = (M Amount-of-in(s, c) t)))
;of course,
∀ IV ∈ individual-view-instance ∀ t ∈ time
        (T Taxonomy(Status(IVI, Active), Status(IVI, Inactive)) t)
```

---

explicitly represented. That is the role of Preconditions.

2.6 Functional Relationships

A key notion of Qualitative Process theory is that the physical processes and individual views in a situation induce functional dependencies between the parameters of a situation. In other words, by knowing the physics you can tell what, if anything, will happen to one parameter when you vary another. In keeping with the exploration of weak information, we define

$$Q_1 \propto_{Q+} Q_2$$

(read "Q_1 is qualitatively proportional to Q_2") to mean "there exists a function that determines Q_1, and is increasing monotonic in its dependence on Q_2. In algebraic notation, we would write

$$Q_1 = f(\ldots, Q_2, \ldots)$$

If the function is decreasing monotonic in its dependence on Q_2, we will say

$$Q_1 \propto_{Q-} Q_2$$

and if we don't wish to specify if it is increasing or decreasing,

$$Q_1 \propto_{Q} Q_2$$

For example, the level of water in a cup increases as the amount of water in the cup increases. We would express this fact by adding into the Relations of the Contained-Liquid description:

```
level(p) ∝Q+ amount-of(p)
```

Often we will leave the function implied by $\propto_Q$ implicit. When it is necessary to name the function, we will say

```
Function-Spec(<id>, <specs>)
```

where `<id>` is the name of the function being defined and `<spec>` is a set of $\propto_Q$ statements and correspondences (see below) that further specify that function. Suppose for example that `level` is expressed in a global coordinate system, so that whenever two open containers have fluid at the same level the pressure the fluid exerts (on the bottom, say) is the same. We might introduce a function `P-L-fun` that relates pressures to levels:

```
Function-Spec(P-L-fun, (pressure(p) ∝Q+ level(p)))
```

Then if c1 and c2 are containers such that

```
(M level(c1) t0) = (M level(c2) t0)
```

then since

```
pressure(c1) = P-L-fun(level(c1))
pressure(c2) = P-L-fun(level(c2))
```

(making a closed world assumption on $\propto_Q$ statements concerning `pressure(c1)` and `pressure(c2)`), by the equalities above we have

```
(M pressure(c1) t0) = (M pressure(c2) t0)
```

Sometimes we will want to express the fact that a function depends on something that is not a quantity. In that case we will say

```
F-dependency(<id>, <thing>)
```

In the contained liquid description, for instance, the level depends on the size and shape of the cup as well as the amount of water. Assuming `shape` and `size` are functions that return something other than quantities, we would write

```
Function-Spec(Level-Function, {Level(p) ∝Q+ Amount-of(p)})
F-dependency(Level-Function, Shape(container(p)),
F-dependency(Level-Function, Size(container(p)))
```

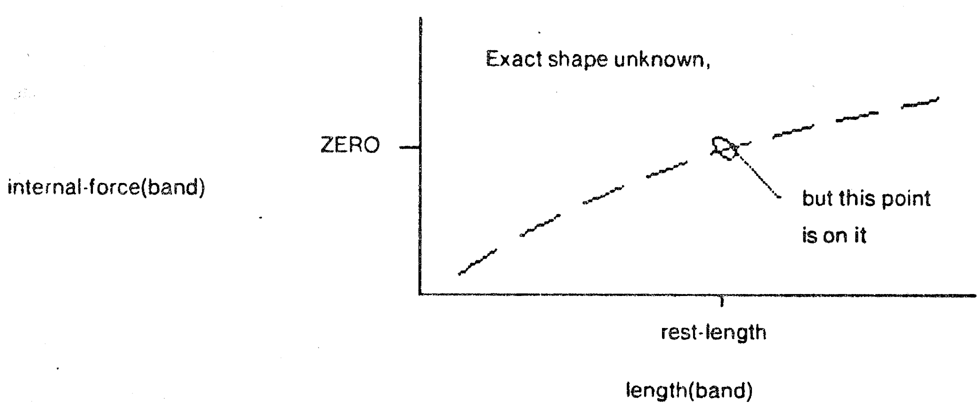
to express this fact. Thus if two containers have the same size and shape, a particular amount of water will result in the same level, but if the size or shape is different we cannot deduce anything about the level that will result.

The definition of $\propto_Q$ is motivated in part by issues involved in causal reasoning, and we will postpone further discussion of its variants until then. There is one other kind of information that can be specified about the function implied by $\propto_Q$'s, and that is a finite set of correspondences it induces between points in the two Quantity Spaces it connects. An example of a correspondence is that the force exerted by an elastic band is zero when it is at rest. This would be written:

```
Correspondence((internal-force(band), ZERO),
               (length(band), rest-length(band)))
```

Correspondences are the means of mapping value information (inequalities) from one Quantity Space to another via $\propto_Q$. For example, if the length of the band described above is greater than its rest length the internal force is greater than zero. Figure 8 illustrates.

Fig. 8. Correspondences Link Quantity Spaces Across $\propto_Q$
A correspondence statement allows information about inequalities to be transferred across qualitative proportionalities ($\propto_Q$'s). The rough shape of the graph below is determined by the $\propto_Q$, the equality between the two points is determined by the correspondence.




```
internal-force(band) ∝Q+ length(band)
Correspondence ((internal-force(band), ZERO)
                (length(band), rest-length(band)))
```

2.7 Histories

To represent how things change through time we will use Hayes' notion of a History. We will assume the concepts introduced in [Hayes, 1979] as our starting point. To summarize, the history of an object is made up of episodes and events. Episodes and events differ in temporal components; events occur at instants and episodes occur over intervals. Each episode will have a `start` and an `end`, events that serve as its boundaries. Each event will have a set of `starts` and `ends` which indicate what episodes it is the start or end of, respectively. We will allow episodes to be either open or closed, according to whether or not what is true at the event is true during the interval it is connected to. This will allow us to say, for example, that the episode of heating water on a stove is ended by event of the water reaching its boiling temperature, yet during the episode the temperature was below the boiling point.

The particular histories Hayes introduced will be called parameter histories, since they are mainly concerned with how a particular parameter of a specific individual changes.[1] Objects can have more than one parameter, and these parameters often can change independently. For example, if we drop a steel ball past a flame, the ball will heat up a bit but the motion won't be affected (unless the combustion gases impart significant momentum to it). Thus the history of an object includes the union of its parameter histories. Figure 9 illustrates the parameter histories for the situation just described. The criteria for individuation, for breaking up time into episodes and events (the spatial component of parameter histories is inherited from the object they are a parameter of) are changes in the values of quantities and their parts. In the previous figure, for example, the events consist of the ball's position reaching `h2` and `h1`, because different values occurred before and after that time. The final component of an object's history are the histories for the processes it participates in, but this will be elaborated later.

Again following Hayes, a slice of a history denotes a piece of an object's history at a particular instant. We will denote the slice of an individual `i` at an instant `t` by

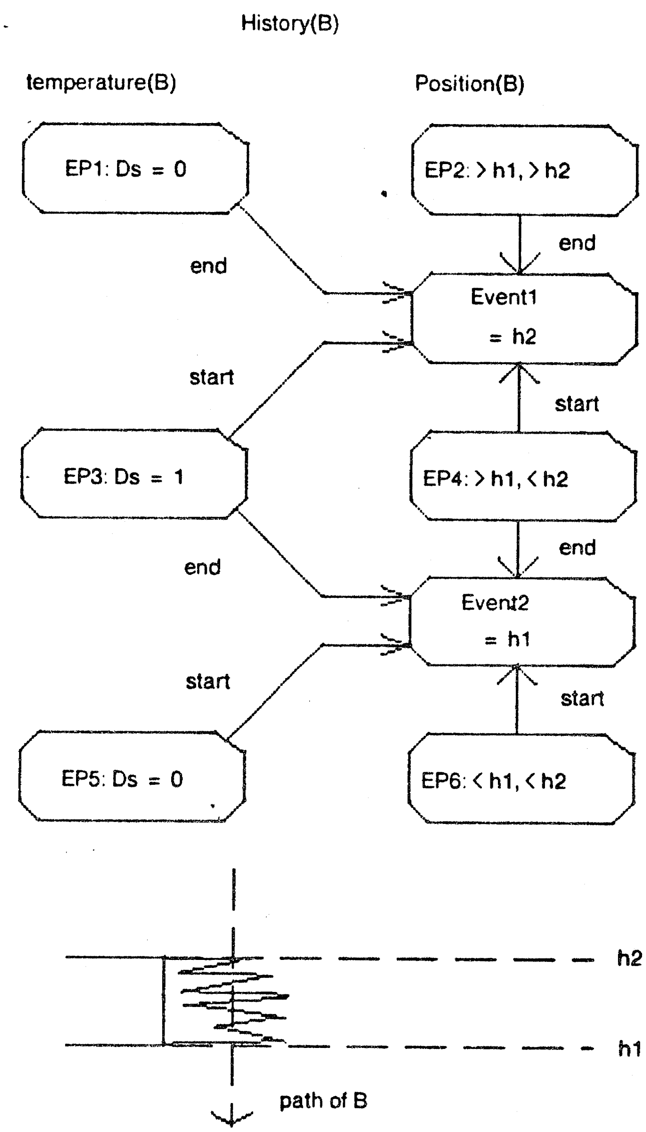
```
at(i, t)
```

If we let all functions, predicates, and relations that apply to objects apply to slices as well, with functions that map from objects to quantities map from slices to values, then we could be rid of `T` and `M` and just talk in terms of slices. For instance, instead of writing

---

1. In fact, Hayes' example are parameter histories for "amount of stuff", representing an object by a subset of $\mathbb{R}^4$.

Fig. 9. Parameter Histories Describe When Values Change
Part of the parameter histories for a ball being dropped through a flame are depicted below. Time runs from top to bottom, and the portion of the history that depicts what is happening (motion and heat flow) are not shown.



```
(T Aligned(P1) t0)
(M amount-of(WC) t0) > (M amount-of(WB) t0)
```

we could write

```
Aligned(at(P1, t0))
amount-of(at(WC, t0)) > amount-of(at(WB, t0))
```

This notation can become quite opaque, so we will resist using it for exposition.

The notion of history so far is "object centered". Since processes will often act between several objects, we need a way of talking about several objects at a particular instant. We will recycle the term situation to mean a collection of slices for a set of objects under consideration at some particular time. As in Situational Calculus, the temporal aspect of a situation will be an instant. The difference is that a situation is now spatially bounded - its spatial extent is that of the slices that comprise it. In formulae where times are required, we will assume a coercion from a situation or event to its time so that we can freely use the names of situations in expressions involving T and M. For the moment the criteria of what constitutes useful situations (the choice of objects to consider together) is unspecified; we will be able to say more about this once we have discussed processes.

3. Processes

A physical situation is usually described in terms of a collection of objects, their properties, and the relationships between them. So far our description of the world has been static - we can describe that things are different from one time to another, but have not provided the means by which changes actually occur. The ways in which things change are intuitively characterized as processes. A physical process is something that acts through time to change the parameters of objects in a situation. Examples of simple processes include fluid and heat flow, boiling, motion, collisions, stretching and compressing.

This section describes what processes are, including how to specify them, and elaborates the notion of *influences*. A catalog of basic deductions involving processes illustrates the kinds of conclusions that can be drawn within QP theory. Histories are extended to include occurances of processes, and the role of processes in specifying a language of behavior is discussed.

3.1 Specifying Processes

A process is specified by five parts:

- The Individuals it applies to,
- A set of Preconditions, that are statements about the individuals and their relationships other than quantity conditions.
- A set of Quantity Conditions, that are either assertions of inequalities between quantities belonging to the individuals (including domain-dependent constants and functions of them) or assertions about the status of processes and individual views.
- A set of Relations the process imposes between the parameters of the individuals, along with new entities that might be created.
- A set of Influences imposed by the process on the parameters of the individuals.

Figure 10 illustrates simple process specifications for heat flow and boiling. (for fans of logic, Figure 11 illustrates how the boiling process would look translated into predicate calculus).

As you can see, a process is just like an individual view - it is a time-dependent thing - except it has something called *influences*. To recapitulate, a process will act between any individuals to which it can apply, exactly whenever both the preconditions and the quantity conditions are true. Preconditions

Fig. 10. Physical Process Definitions

Here are two examples of process specifications from the Fluids World. Heat flow happens between two objects that have heats and are connected via some path through which heat can flow. The predicate `heat-aligned` is true exactly when heat can flow through the path, named for the analogous predicate in fluid flow. Boiling happens to a contained liquid being heated, and creates a gas made of the same stuff as the liquid.

```
process heat-flow

Individuals:
        src an object,HasQuantity(src, heat)
        dst an object,HasQuantity(dst, heat)
        path a HeatPath, HeatPath(path,src,dst)

Preconditions:
        heat-aligned(path)

QuantityConditions:
        A[temperature(src)] > A[temperature(dst)]

Relations:
        Let flow-rate be a quantity
         A[flow-rate] >  ZERO
         flow-rate ∝Q+ (Am[temperature(src)] - Am[temperature(dst)])

Influences:
        I-(Heat(src), A[flow-rate])
        I+(Heat(dst), A[flow-rate])

------------------------------------------------------------
process boiling

Individuals:
        w a contained-liquid
        hf a process-instance, process(hf) = heat-flow
                                 ∧ dst(hf) = w

QuantityConditions:
        Active(hf)
        ∧A[temperature(w)] = A[t-boil(w)]

Relations:
        There is g ∈ piece-of-stuff
        gas(g)
        substance(g) = substance(w)
        temperature[w] = temperature[g]
        Let generation-rate be a quantity
        A[generation-rate] > ZERO
        generation-rate ∝Q+ flow-rate(hf)

Influences:
        I-(heat(w), A[flow-rate(hf)])
        I-(amount-of(w), A[generation-rate])
        I+(amount-of(g), A[generation-rate])
```

Fig. 11. Boiling Expressed as an Axiom

Here is how the boiling description could be written as an axiom. For clarity, the temporal references have been omitted.

```
∀ w ∈ contained-liquid
∀ hf ∈ process-instance
(process(hf) = heat-flow ∧ dst(hf) = w
⇒
[∃ pi ∈ process-instance
 process(pi) = boiling ∧ w(pi) = w ∧ path(pi) = path ∧ hf(pi) = hf
 ∧
 [(Active(hf) ∧ A[temperature(w)] = A[t-boil(w)]) ⇒ Status(pi, Active)
  ∧
    Status(pi, Active) ⇒
        [∃ g ∈ piece-of-stuff
         ∃ generation-rate ∈ quantity
         gas(g) ∧ substance(g) = substance(w) ∧ T[w] = T[g]
         ∧ A[generation-rate] > ZERO
         ∧ generation-rate ∝Q+ flow-rate(hf)
        ∧ A[flow-rate(hf)] ∈ MinusInputs(InfluenceAdder(heat(w)))
∧ A[generation-rate] ∈ MinusInputs(InfluenceAdder(amount-of(w)))
∧ A[generation-rate] ∈ PlusInputs(InfluenceAdder(amount-of(g)))]]]

; A process instance is either active or inactive, corresponding to whether or
;not the process it represents is acting between the particular individuals.

(∀ pi ∈ process-instance
        Active(pi) ≡ Status(pi, Active)
        ∧ Taxonomy(Status(pi, Active), Status(pi, Inactive))
```

are those factors that are outside Qualitative Process theory, such as someone opening or closing a valve to establish a fluid path, but still relevant to whether or not a process occurs. The Quantity Conditions are those statements that can be expressed solely within QP theory, such as requiring the temperature of two bodies to be different for heat flow to occur, or boiling to occur as a prerequisite to generating steam. The set of Relations associated with a process are the relationships it imposes between the objects it is acting on. The Relations component usually describes indirect effects via functional relationships between quantities, such as the flow rate in fluid flow being qualitatively proportional to the difference in the pressures of the contained fluids involved. The Relations also include descriptions of any new individuals created by the process, as for example the steam generated by boiling. We will discuss influences next.

3.2 Influences and Integration

Influences specify the direct effects of the process on the quantities of the objects involved. For example, in a flow process the flow rate will typically correspond to the increase in the amount of "stuff" at the destination and to the decrease in the amount of "stuff" at the source. If the number n is a direct influence on the quantity Q, we will write

```
I+(Q, n)
I-(Q, n)
I±(Q, n)
```

according to whether its influence is positive, negative, or unspecified.

If a quantity is directly influenced then its derivative is the sum of the direct influences. Determining the sign (and possibly the magnitude) of the derivative will be called *resolving* the influences, by analogy with resolving forces in classical mechanics. Resolving direct influences requires combining D_s values; figure 12 illustrates the algebra of signs involved.

Fig. 12. Combining Sign Values in Derivatives
This table specifies how sign values combine across derivatives of addition and multiplication. de Kleer's formulation used the symbol ? to denote the result for the cases that require information about amounts and rates.

```
D[A <op> B] = -1, 0, or 1

where <op> is +, *
```

A \ B	-1	0	1
-1	-1	-1	N1
0	-1	0	1
1	N1	1	1

```
N1: if + then if m(A)>m(B) then s(A)
              if m(A)<m(B) then s(B)
              otherwise 0
    if * then if m(A)*A(B)
                 > m(B)*A(A) then s(A)
              if m(A)*A(B)
                 < m(B)*A(A) then s(B)
              otherwise 0
```

A notion of integrability - the relationship between the derivative of a quantity and its amount - is needed. Essentially, if the derivative is negative then the amount will decrease over an interval, if positive then the amount will increase, and if zero then the amount will be the same:

```
∀ q ∈ quantity ∀ I ∈ interval
(constant-sign(D[q], I) ∧ ¬ Duration(I) = ZERO) ⇒
        (M Ds[q] during(I)) = -1 ↔ (M A[q] end(I)) < (M A[q] start(I))
        ∧ (M Ds[q] during(I)) = 1 ↔ (M A[q] end(I)) > (M A[q] start(I))
        ∧ (M Ds[q] during(I)) = 0 ↔ (M A[q] end(I)) = (M A[q] start(I))
```

where

```
∀ n ∈ numbers, ∀ I ∈ intervals,
constant-sign(n I) ≡ (∀ i1,i2 ∈ during(I) (M s[n] i1) = (M s[n] i2))
```

This statement is very weak compared to our usual notion of integrability.[1] In particular, it does not rest on knowing an explicit function describing the derivative and thus does not require an explicit notion of integrals.

In addition to direct influences, a quantity may be changing because it is a function of some other quantity that is changing. Qualitative proportionalities ($\propto_Q$s) are the means of specifying these indirect effects. If a quantity Q_0 is directly influenced by a process P and Q_1 is qualitatively proportional to Q_0 then we will say that P <u>indirectly influences</u> Q_1.

Importantly, note that at any particular time a quantity will be either directly influenced, indirectly influenced, or not influenced at all. It is further assumed that no quantity is both directly and indirectly influenced at once. The intuition is that those quantities that are directly influenced are in some sense "independent" quantities, and that for causal reasoning this distinction must be preserved. This issue will be discussed fully later on.

3.3 Limit Points

The quantities and constants that are compared to a particular quantity by Quantity Conditions are included as elements in its quantity space. Because they correspond to discontinuous changes in the processes that are occurring, they are called *limit points*. Limit points serve as boundary conditions. For

---

1. If the time involved is an instant (i.e., an interval of duration ZERO), then we will also assume that the quantity "doesn't change very much" during this time. This assumption underlies one of the cases of the Equality Change Law that is discussed below.

example, the temperature Quantity Space for a an object w made of substance s would include the limit points:

```
t-melt(w, s) -> t-boil(w, s)
```

where the object undergoes state changes that result in qualitatively distinct behavior. As we have seen, these different modes of behavior are modeled by Individual Views.

3.4 The Sole Mechanism Assumption and Process Vocabularies

The central assumption of Qualitative Process theory is the *Sole Mechanism* assumption, namely:

All changes in physical systems are caused directly or indirectly by processes.

As a consequence, the physics for a domain must include a vocabulary of processes that occur in that domain. This *Process Vocabulary* can be viewed as specifying the dynamics theory for the domain. A situation, then, is described by a collection of objects, their properties, the relations between them (including Individual Views), and the processes that are occurring. A consequence of the Sole Mechanism assumption is that it allows us to reason by exclusion. If we make the additional assumption that our Process Vocabulary for a domain is complete, then we know what quantities can be directly influenced. If we understand the objects and relationships between them well enough, we know all the ways quantities can be indirectly influenced. Thus we know how the physical world will change. Without these Closed World assumptions, (see [Collins et. al., 1975].[Moore, 1975], [Reiter, 1980]) it is hard to see how a reasoning entity could use, much less debug or extend, its physical knowledge.

3.5 Reprise

Processes should be first class entities in the ontology of Naive Physics. It may be tempting to think that processes are mere abbreviations for "deeper" representations, such as constraint laws. However, they are not. The temptation arises both because constraint laws are often judged to be the most elegant physical descriptions in "non-Naive" physics, and because constraint-based computer models have been fairly successful for analyzing engineered systems ([Stallman & Sussman, 1977], [de Kleer & Sussman, 1978]). However, the aims of Naive Physics are not the same as the aims of physics or engineering analysis. In physics we are trying to construct the simplest models that can make detailed predictions about physical phenomena. When performing an engineering analysis, even a qualitative one, we have chosen a particular point of view on the system and abstracted away certain objects. Unlike these enterprises, Naive Physics attempts to uncover the ideas of physical reality that people actually use in daily life. Thus the notions that physics throws away (objects, processes, causality) for conciseness in its formal theory -- the equations -- are precisely what we must keep.

Qualitative Process theory concerns the form of dynamics theories, not their specific content. For example, the heat flow process illustrated adheres to energy conservation, and does not specify that "stuff" is transferred between the source and destination. The language provided by the theory also allows a heat flow process that violates energy conservation and transfers "caloric fluid" between the source and destination to be written. The assumptions made about the content of dynamics theories are quite weak. Aside from the ability to write a wide variety of physical models, the weakness of its assumptions allow other theories to be written that impose further constraints on the legal vocabularies of processes. For example, conservation of energy can be expressed as a theory about certain types of quantities and the allowable patterns of influences in processes that directly affect those types of quantities (see section 4.5). We do not, however, wish to saddle QP theory with these assumptions.

3.6 Basic Deductions

To be useful, a representation must be computable from other information and in turn sanction other deductions. Several basic deductions involving the constructs of Qualitative Process theory are cataloged below. It may be helpful to skip momentarily to the example in section 4.1, which illustrates these deductions step by step.

3.6.1 Finding Possible Processes

A Process Vocabulary determines the types of processes that can occur. Given a collection of individuals and a Process Vocabulary, the individual specifications from the elements in the Process Vocabulary must be used to find collections of individuals that can participate in each kind of process. These *Process Instances* (PI's) represent the potential processes that can occur between a set of individuals. A similar deduction is used for finding Individual View instances.

3.6.2 Determining Activity

A Process Instance has a status of `Active` or `Inactive` according to whether or not the particular process it represents is acting between its individuals. By determing whether or not the Preconditions and Quantity Conditions are true, a status can be assigned to each Process Instance for a situation.[1] The collection of active process instances is called the *Process Structure* of the situation. The Process Structure represents "what's happening" to the individuals in a particular situation.

3.6.3 Determining Changes

Most of the changes in an individual are represented by the D_s values for its quantities. As stated previously, there are two ways for a quantity to change. A quantity can be directly influenced by a process, or it can be indirectly influenced via $\propto_Q$. (By the Sole Mechanism assumption, if a quantity is uninfluenced its D_s value is 0.) As mentioned above, determining the D_s value for an influenced quantity is called *resolving* its influences, by analogy to resolving forces in classical mechanics.

Resolving a directly influenced quantity involves sorting its influences by sign and determing which collection has the greater magnitude (Of course, we will not always have even that much information). Resolving an indirectly influenced quantity involves gathering the $\propto_Q$ statements that implicitly specify it as a function of other quantities. Because the function is implicit, in many cases indirect influences cannot be resolved within the basic theory. An example will make this point clearer. Suppose we have a quantity Q_0 such that in a particular Process Structure:

$$Q_0 \propto_{Q+} Q_1 \wedge Q_0 \propto_{Q-} Q_2$$

If we also know that

$$D_s[Q_1] = 1 \wedge D_s[Q_2] = 1$$

then we cannot determine $D_s[Q_0]$, because we do not have enough information to determine which

1. This can require searching the completions of the relevant Quantity Spaces.

indirect influence "dominates".[2] However, if we had

$$D_s[Q_1] = 1 \wedge D_s[Q_2] = 0$$

then we can conclude that

$$D_s[Q_0] = 1$$

because Q_1 is now the only active indirect influence.

Domain specific and problem specific knowledge will often play a role in resolving influences. We may know that a certain influence can be ignored, such as when we ignore the heat lost by a kettle on a stove to the air surrounding it while it is being heated to boiling. Our knowledge about particular functions may tell us which way things combine. Suppose for instance that our model of fluid flow included influences to model the changes in heat and temperature that result from mass transfer. In the source and destination both `heat` and `temperature` would be indirectly influenced (via `Amount-of`), and if we knew nothing but the D_s values we could say nothing about how they will change. From physics, however, we know that the temperature of the source is unchanged and the temperature of the destination will rise or fall according to whether the temperature of the source is greater or less than the temperature of the destination.

3.6.4 Limit Analysis

The changes in a situation can result in the Process Structure itself changing. Determining the possible changes in a situation's Process Structure is called *Limit Analysis.* Limit Analysis is carried out by using the D_s values and Quantity Spaces to determine which Quantity Conditions can change.

The first step is to find the neighboring points within the Quantity Spaces of each changing quantity. If there is no neighbor in a direction, then a change in that direction cannot affect any process. The ordering between each neighbor and the current amount of the quantity can be combined with the D_s values of each to determine if the relationship will change (see Figure 13). If the neighbor is a limit

---

2. The "dominates" is in quotes because the word implicitly assumes Q_1 and Q_2 are separate terms whose sum is part of the function that determines Q_0, and this need not be the case. For example,

$$Q_0 = Q_1 * (1 - Q_2)$$

is consistent with the statements above.

Fig. 13. Linking Derivatives with Inequalities

This table summarizes how the ordering relationship between two quantities may change according to the sign of their derivatives over some interval.

For A > B;

	B: -1	B: 0	B: 1
A: -1	N1	=	=
A: 0	>	>	=
A: 1	>	>	N2

N1:
If $D_m[A] > D_m[B]$ then >;
If $D_m[A] < D_m[B]$ then =;
If $D_m[A] = D_m[B]$ then >;

N2:
If $D_m[A] > D_m[B]$ then =;
If $D_m[A] < D_m[B]$ then >;
If $D_m[A] = D_m[B]$ then >;

For A = B;

	B: -1	B: 0	B: 1
A: -1	N3	<	<
A: 0	>	=	<
A: 1	>	>	N4

N3:
If $D_m[A] > D_m[B]$ then >;
If $D_m[A] < D_m[B]$ then <;
If $D_m[A] = D_m[B]$ then =;

N4:
If $D_m[A] > D_m[B]$ then >;
If $D_m[A] < D_m[B]$ then <;
If $D_m[A] = D_m[B]$ then =;

---

point, some process may end there and others begin. Thus the set of possible changes in orderings involving limit points determines the ways the current set of active processes might change. The set of single changes plus consistent conjunctions of changes (corresponding to simultaneous changes in quantity conditions) forms the set of *Limit Hypotheses*. Each Limit Hypothesis determines a situation called an *Alternate Ending* for the Process Structure, and the one which actually occurs is called the *Next*

Structure for that Process Structure.[1]

More than one change is typically possible, as the examples below will illustrate. There are three reasons for this. First, if the ordering within a Quantity Space is not a total order more than one neighbor can exist. Second, a process can influence more than one quantity. Finally more than one process can be occurring simultaneously. The basic theory does not in general allow the determination of which alternative actually occurs. Using Calculus as the model for quantities, the alternative which occurs next is the one for which time to integrate the quantities involved to their limit points is minimal. Since the basic theory does not include explicit integrals, this question typically cannot be decided.

There are some special situations, due to the nature of quantities, where sometimes we can do better. Consider two quantities A and B that are equal, and C and D that are unequal. If all of the quantities are changing (D_s value of -1 or 1), then the finite difference between C and D implies that the change in the equality between A and B occurs first. In fact, we will assume that the change from equality occurs in an instant, while the change to equality usually will take some interval. We will further assume that the only time a change to equality will take an instant is when the change in value was due to a process that acted only for an instant. We will summarize this as the *Equality Change Law:*

> *With two exceptions, a Process Structure lasts over an interval of time. It lasts for an instant only when either*
>
> *(1) A change from equality occurs in a Quantity Condition or*
>
> *(2) A change to equality occurs between quantities in a Quantity Condition that were influenced away from equality for only an instant.*

The first case assumes that the values of numbers aren't "fuzzy", and the second case assumes that the changes wrought by processes are well-behaved (i.e., no impulses).

Now consider the maximal element (in the sense of inclusion) of the set of changes that occur in an instant. The Limit Hypotheses that contain this maximal element are the ones which can occur next, because the duration of an instant is shorter than the duration of an interval. By using the Equality Change Law to identify those Limit Hypotheses that represent changes that occur in an instant, we can

---

1. This assumes that rates are not infinitesimals, so that if a quantity is moving towards some point in its space it will actually reach that value in some finite time. Note that relaxing this assumption would result in only one additional state in the possibilities returned by the limit analysis - that the current set of active processes never changes.

sometimes get a unique result from Limit Analysis within the basic theory (see section 5.1).

For some kinds of tasks just knowing the possible changes is enough (such as envisioning, in [de Kleer, 1975]). If required, knowledge outside the scope of Qualitative Process theory can be used to disambiguate the possibilities. Depending on the domain and the style of reasoning to be performed there are several choices: simulation [Forbus, 1981a], algebraic manipulation [de Kleer, 1975], teleology [de Kleer, 1979], or possibly default assumptions or observations [Forbus, 1983a].

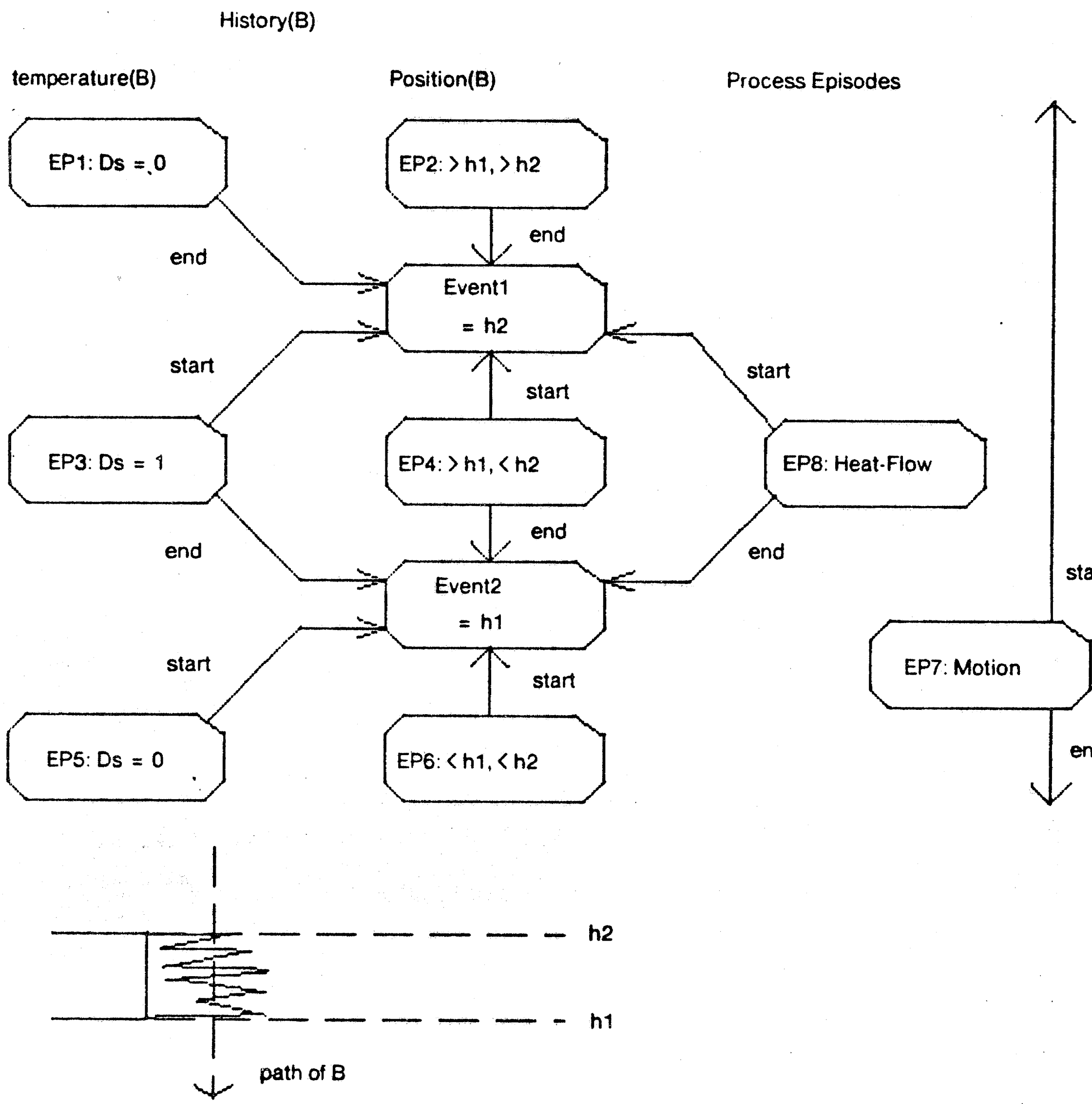
3.7 Processes and Histories

Adding processes to the ontology of Naive Physics allows the History representation of change to be extended. In addition to object histories, we will also use process histories to describe what processes are occurring when. The temporal extent of a process episode is the interval during which the process is active, and the spatial extent is the spatial extent of the individuals involved in it. The events that bound episodes in the process history occur at the instants at which Quantity Conditions or Preconditions change. Process episodes are included in the histories of the objects that participate in the process, and the union of the object's parameter histories and the history of the processes it participates in comprise its total history. Figure 14 illustrates the full history over a small interval for the ball being dropped through a flame discussed previously.

As mentioned previously, two key problems in reasoning with histories are the *Local Evolution* problem (extending the known portion of a an object's history, preferably by carving up the situation into pieces that can be reasoned about semi-independently) and the *Intersection/Interaction* problem. The key to solving them lies in having explicit descriptions of the ways changes are caused.

The processes active in a situation form its Process Structure (we will also include the active individual views, the View Structure, in the Process Structure to simplify discussion). We will define *P-components* as equivalence classes of the Process Structure as follows. A process is in the same P-component as another process (or individual view) if either: (a) it influences a quantity mentioned in the other's Quantity Conditions, (b) it influences a quantity influenced by the other, (c) its Quantity Conditions mention a quantity mentioned in the Quantity Conditions of the other, or (d) the other description contains a $\propto_Q$ that propagates an influence of the process. As long as a particular Process Structure lasts, the P-components can be reasoned about independently. For example, we usually don't worry about getting our feet wet in a basement despite the proximity of flowing water and steam in our plumbing. Changes in the Process Structure can bring about changes in P-components, so the

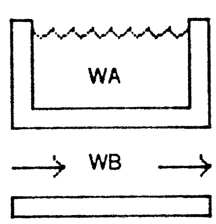
Fig. 14. History for a Ball Dropping Through a Flame
Here is a piece of the history for the ball again, but with process episodes added. As before, EP<n> are episodes, and time runs from top to bottom.



conclusions made in each component may have to be modified depending on how the Process Structure changes. If our plumbing leaks, for instance, there are now ways for our feet to get wet.

The individuals affected by the processes in each P-component define a collection of things that can be reasoned about in isolation, barring changes in Process Structure. Thus we can generate object histories by evolving P-components, and combining the results when the Process Structure changes to get new P-components, and so forth. The Interaction Problem becomes trivial - two episodes interact if and only if the processes that give rise to them are part of the same P-component of a Process Structure on a situation made up of slices from those particular episodes. Figure 15 provides a graphical illustration.

Fig. 15. Determining Interactions
Suppose A is a liquid and B is a fluid flowing through the channel below A's container. Below are the Process Structures that result from different assumptions about the situation, with potential interactions indicated.




```
If shared wall is not a heat path,
        PS: Fluid-Flow(B), no interaction
Otherwise, if A[Temperature(A)] = A[Temperature(B)]
        PS: Fluid-Flow(B), no interaction
        Otherwise, PS: Fluid-Flow(B), a heat flow, so interaction
```

3.8 A Language for Behavior

Qualitative Process theory concerns the structure of qualitative dynamics. We can view it as specifying a language in which certain commonsense physical models can be written. Can this language be extended to form a full language of behavior for physical systems? Although I have not yet done so, I will argue that the answer is yes, and that several advantages would result from the extension.

A language should have primitives, a means of combining these primitives, and some means of abstraction to allow new entities to be defined. Processes and Individual Views are obviously the primitives in this language.[1] There are two sensible kinds of compound processes. The first kind consists of processes that form a P-component, a shared parameter combination. An example of a shared parameter combination is the intake stroke of a four cycle engine, which consists of a flow of air and gas into a cylinder and motion of the piston. The second kind consists of sequences of processes occurring over the same individuals. An example of a sequential combination is the sequence of intake, compression, combustion and exhaust strokes of a four-cycle engine. Treating these combinations as new "things" then allows properties of the system they describe to be reasoned about, as we will see when deducing the conditions under which a pumped oscillator will remain stable.[2]

It should be clear that the shared parameter combination can be treated exactly as a simple process, specified by the union of the properties of the component processes. The sequential combination is <u>not</u> a process, because the same influences and relations do not hold over every time within the occurances of the sequential combination. A sequential combination is really a piece of a history! In particular, it is the history of the Individuals affected by the processes viewed as a system. In honor of this mixed ontological status such descriptions will be called *Encapsulated Histories.* Encapsulated Histories (EH) are quite important for two reasons. First, some phenomena which are

---

1. The choice of what is primitive in any particular domain's vocabulary will of course vary - for example, the description of a gas we will use later is macroscopic. Presumably a richer process vocabulary would contain the "mechanisms" that induce these relations (i.e., the Kinetic theory of gases), but there is no reason to always include such detail. Consider for example a resistor in a circuit that never exceeds its electrical capacity. The detailed mechanics of conduction hinder rather than help when calculating the current that will result from a voltage across it.

2. One use of compound processes would be representing the device models in de Kleer and Brown's theory of machines [de Kleer & Brown, 1983]. The Preconditions and Quantity Conditions of the compound process would correspond to their assumptions about the validity of the device model.

described in that form seem irreducible in terms of processes - collisions, for example. Second, they serve as abstract descriptions for more complex behavior, as in describing the pattern of activity in an oscillator. When writing sequential combinations, we will use most of the structure of a process, in that the combination will have Individuals, Preconditions, and Quantity Conditions. However, there will be no specification of Influences, the Relations component is restricted to holding a description of a piece of the history for the individuals, and the Preconditions and Quantity Conditions are written relative to episodes in the piece of history. If the Preconditions and Quantity Conditions are ever true for a partial history of a collection of objects matching the Individual specifications, then the piece of history described in its Relations is predicted to be part of the history of those objects.[1] Figure 16 illustrates an Encapsulated History describing water being heated to boiling on a stove.

If the entry point is in the beginning of the encapsulated piece of history, then the prediction will hold if the P-component in the objects is the same as the P-component in the objects at the entry slice. For those phenomena which are irreducible, the EH may be the only way to evolve the history of the object past that point. For systems where the EH serves as a summary, an interesting kind of perturbation analysis becomes possible. In performing an energy analysis, for example, the Quantity Conditions are re-written in terms of energy (see section 4.5). Changes to the system are modeled by processes that influence energy (such as friction), and the effects are determined by examining the episodes that comprise the EH.

3.9 Classification and Abstraction

A classification hierarchy is also needed to account for the various kinds of conditions under which processes occur. For example, Hayes has elucidated several distinct conditions under which fluid flow occurs [Hayes, 1979b]. Another example are the kinds of motion - flying, sliding, swinging, and rolling. Sliding and rolling are examples of motion along a surface, and along with swinging form motions involving constant contact with another object. A third example is heat flow, whose influences change according to whether a change of state is involved or not. Each of these conditions has slightly different properties, but they are sufficiently similar in the individuals they involve and the pattern of

---

1. Many of diSessa's "phenomenological primitives"[diSessa, 1983] might be representable as Encapsulated Histories.

Fig. 16. Encapsulated Histories Describe a Pattern of Behavior

Turning on a stove, boiling away some water and turning the stove off when finished are modeled below. While useful as summaries and as descriptions for ill-understood behavior, Encapsulated Histories violate the composability criteria for dynamics theories because they presuppose independence. For simplicity we will assume here that `tboil` as a function from a piece of stuff to the quantity that represents its boiling point is constant over time.

```
Encapsulated History boil-on-stove

Individuals:
        w a contained-liquid, made-of(w) = water
        b a burner
        E an episode
Preconditions:
        (T On(w, b) start(E))
        (T Turned-On(b) start(E))
        (T Stays-On(b) during(E))
        (T Turned-Off(b) end(E))
QuantityConditions:
        (M temperature(w) start(E)) < (M Tboil(w) start(E))
Relations:
        there are EP1,EP2,EP3  ∈ process-episode
        there are EP4, EP5, EP6, EP7, EP8, EP9, EP10 ∈ parameter-episode
        there are PI1, PI2 ∈ process-instance
        there is E1 ∈ event
        ;first what sorts of things can happen-
        process(PI1) = heat-flow, src(PI1) = b, dst(PI1) = w
        process(PI2) = boiling, w(PI2) = w, hf(PI2) = PI1
        ;what the episodes are about
        what(EP1) = PI1, what(EP2) = what(EP3) = PI2
        what(EP4) = heat(w), what(EP5) = what(EP6) = temperature(w)
        what(EP7) = what(EP8) = amount-of(w)
        what(EP9) = what(EP10) = temperature(w)
        ;now the temporal relationships
        start(EP1) = start(E), end(EP1) = end(E) ; the heat flow
        start(EP2) = start(E), end(EP2) = E1 ;not boiling
        start(EP3) = E1, end(EP3) = end(E) ;boiling
        start(EP4) = start(E), end(EP4) = end(E)
        start(EP5) = start(E), end(EP5) = start(EP6) = E1, end(EP6) = end(E)
        start(EP7) = start(E), end(EP7) = start(EP8) = E1, end(EP8) = end(E)
        start(EP9) = start(E), end(EP9) = start(EP10) = E1
        ;we don't specify the end of EP10 because we are not thinking
        ;about how the water cools down
        ;now we describe what is happening in each
        (M A[temperature(w)] time(E1)) = A[tboil(w)]
        (T Status(PI1, Active) time(EP1))
        (T Status(PI2, Inactive) time(EP2))
        (T Status(PI2, Active) time(EP3))
        (M Ds[heat(w)] time(EP4)) = 1
        (M Ds[temperature(w)] time(EP5)) = 1
        (M Ds[temperature(w)] time(EP6)) = 0
        (M Ds[amount-of(w)] time(EP7)) = 0
        (M Ds[amount-of(w)] time(EP8)) = -1
        (M A[temperature(w)] time(EP9)) < A[tboil(w)]
        (M A[temperature(w)] time(EP10)) = A[tboil(w)]
```

influences they engender to be considered the same kind of process. Having explicit abstract descriptions of processes should also be useful because they are often easier to rule out than more detailed descriptions. If, for instance, there is no path between two places through which an object can be moved, it cannot get there by sliding, flying, rolling, or any other kind of motion that might exist.

In theory, disjunctions could be used within a single process description to cover the various cases. Doing so would lead to complicated descriptions that could not easily be reasoned about. Instead, every case will be represented by a different process. We will say that P1 is a case of P2, such as

```
Case-of(Motion, Swinging)
```

The following restrictions hold on cases:

> *Specificity:* There is a subset of the Individuals specified for P1 such that they or individuals whose existence is implied by them match the Individual specifications of P2 The Preconditions and Quantity Conditions for P1 respectively imply the Preconditions and Quantity Conditions for P2.
>
> *Inheritance:* All statements in the Relations and Influences fields of P2 hold for P1 unless explicitly excluded.

Figure 17 illustrates some specializations of the abstract motion process that will be discussed in section 5.3.

Fig. 17. Some Specialized Descriptions of Motion

Cases of motion are organized around constraints on kinematics. The abstract motion process already includes the individuals B (a movable object) and dir, a direction. The abstract motion process will be explained in more detail later. In sliding and rolling there is contact with a surface, but different constraints on the kind of contact. Otherwise the same facts pertain to them as to the abstract version of motion.

```
Process Motion(B,dir)

Individuals:
        B an object, Mobile(B)

Preconditions:
        Free-direction(B, dir)
        Direction-Of(dir, Vel(B))

QuantityConditions:
        Am[Vel(B)] > ZERO

Influences:
        I+(position(B), Vel(B))
```

```
Process Slide

Case-of: Motion

Individuals:
        S a surface

Preconditions:
        Sliding-Contact(B, S)
        AlongSurface(dir, B, S)
```

```
Process Roll

Case-of: Motion

Individuals:
        S a surface

Preconditions:
        Contact(B, S)
        Round(B)
        AlongSurface(dir, B, S)
```

4. Examples

At this point a great deal of representational machinery has been introduced. It is time to illustrate how QP theory can be used in physical reasoning. The examples will be fairly informal for two reasons. First, the formalization of the domains is still underway.[1] Secondly, while Qualitative Process theory provides an important part of any domain's theory, there are several crucial considerations that lie outside of it, such as spatial reasoning. Still, the examples presented are complex enough to provide an indication of the theory's utility. The assumptions about other kinds of knowledge needed will be noted as they occur.

4.1 Modelling Fluids and Fluid Flow

This example will illustrate some of the basic deductions sanctioned by Qualitative Process theory and introduce the representations of fluids that will be used in other examples. These representations are slightly more complex than the contained liquid description we have been using. Consider the two containers illustrated in Figure 18. What will happen here?

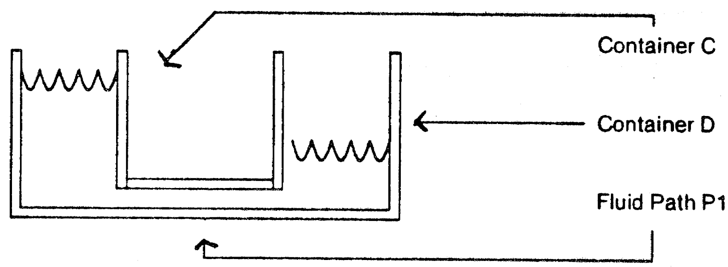
We first introduce descriptions of the fluids. Following Hayes[Hayes, 1979b], we will individuate liquids according to what contains them. Figure 19 describes "pieces of stuff" and contained substances. Any piece of stuff must be in some state, either solid, liquid, or gas. Figure 20 describes the states of substances. The interaction of state and containment is described in Figure 21. Since the containers initially contain some water, we will create individuals corresponding to the water in each container. Call the pieces of stuff in containers `C` and `D` `WC` and `WD` respectively. We will assume their temperatures are such that they are both liquids. For simplicity we will ignore the liquid in the pipe `P1`. We will also ignore the precise definition of fluid paths, except to note that `P1` is one, connecting the two contained fluids.

Suppose our Process Vocabulary consists of fluid flow, whose description is illustrated in Figure 22. This model is very simple, because it ignores the possibility of different kinds of fluids and the details

---

1. At present work is focusing on two domains: the *Mechanism World*, and the *Fluids World*. The Mechanism World includes the BlocksWorld but also more complex shapes and some non-rigid materials. The aim of work in the Mechanism World is to understand devices such as mechanical watches and automobile transmissions. The Fluids World is an attempt to extend Hayes' theory of liquids to include gases and more complex fluid systems such as found in steam plants.

Fig. 18. Two Partially Filled Containers
Containers C and D are connected by a pipe. C contains more water than D. In general an "-in" suffix indicates a function that maps from a container and a substance to a quantity.




```
;structural description

Container(C)
Container(D)
FluidConnected(C, D, P1)

;some substances are in the containers
ContainsSubstance(C, water)
ContainsSubstance(D, water)

;the levels are related
(M Level-in(C, water) Initial) > (M Level-in(D, water) Initial)
```

---

of how fluids move through the fluid paths ([Hayes, 1979b] illustrates some of the distinctions that must be drawn).

With the situation we have so far, there are two Process Instances, one corresponding to flow from C to D and the other corresponding to flow from D to C. To determine if either is active (thus determining the Process Structure) we have to know the relative pressures of WC and WD. Assume we deduce from the relative levels that the pressure in C is greater than the pressure in D. Then the Process Instance representing fluid flow from WC to WD will be active, and the Process Instance representing fluid flow from WD to WC will be inactive. Thus the Process Structure is the set consisting of `Fluid-Flow(WC, WD, P1)`.

To find out what changes are occurring we must resolve the influences. In this situation resolving the influences is simple. The fluid flow from C to D is the only cause of direct influences, changing `amount-of` for WC and WD. Each of them has only one influence, hence

D_s[amount-of(WC)] = -1

Fig. 19. Pieces of Stuff and Contained Substances
Contained-Stuff describes the material of a particular kind inside a container. A piece of stuff is mainly described by several quantities and its location.

```
Individual-View Contained-Stuff

Individuals:
        c a container
        s a substance

Preconditions:
        ContainsSubstance(c, s)

QuantityConditions:
        A[amount-of-in(c, s)] > ZERO

Relations:
        There is p ∈ piece-of-stuff
        amount-of-in(c, s) = amount-of(p)
        s = made-of(p)
        inside(c) = location(p)


;A piece of stuff consists of several quantities, a substance, and
;a location
(∀ p ∈ piece-of-stuff
        HasQuantity(p, amount-of)
        ∧ HasQuantity(p, volume) ∧ HasQuantity(p, pressure)
        ∧ HasQuantity(p, temperature) ∧ HasQuantity(p, heat)
        ∧ Substance(made-of(p)) ∧ Place(location(p))
        ∧ temperature(p) ∝Q+ heat(p) ∧ temperature(p) ∝Q- amount-of(p))

;where
(∀ p ∈ things ∀ q ∈ quantity-type
        HasQuantity(p, q) ↔ Quantity(q(p)))
```

---

and

D_s[amount-of(WD)] = 1

These in turn influence level and pressure, each of which has only one $\propto_Q$ applicable(see figure 21). Thus we can deduce that the level and pressure in C are decreasing, and the level and pressure in D are increasing. All other quantities are uninfluenced, hence unchanging. Limit Analysis is similarly easy. The pressures will eventually be equal, which means the fluid flow will stop. It is also possible that the container C will run out of water, thus ending WC's existence (although it is not possible in the particular drawing shown). These results are summarized in Figure 23. This graph of process structures can be used to generate a history by first creating the appropriate episodes for objects and processes, then selecting one or the other Limit Hypothesis as the end event for that episode. Usually we will just represent the interconnections between possible Process Structures as we have done here. With only a

Fig. 20. States of Matter

The temperatures at which state changes occur are modeled by two functions `t-melt` and `t-boil`. Making them depend on the piece of stuff represents a weaker commitment than depending on the substance and pressure. The Quantity Conditions express the fact that a substance can be in either state at a phase change, but that a particular piece cannot be in both states at once. To determine the state of a piece of stuff at the phase boundary requires either knowing its history or making an assumption.

```
Individual-View Solid(p)

Individuals:
        p a piece-of-stuff

QuantityConditions:
        ¬ A[temperature(p)] > t-melt(p)
        ¬ Liquid(p)
```

```
Individual-View Liquid(p)

Individuals:
        p a piece-of-stuff

QuantityConditions:
¬ A[temperature(p)] < t-melt(p)
¬ A[temperature(p)] > t-boil(p)
¬ Solid(p)
¬ Gas(p)
Relations:
        volume(p) ∝Q+ amount-of(p)
```

```
Individual-View Gas(p)

Individuals:
        p a piece-of-stuff

QuantityConditions:
        ¬ A[temperature(p)] < A[t-boil(p)]
        ¬ Liquid(p)

Relations:
        D[temperature(p)] ∝Q+ D[amount-of(p)]
        D[temperature(p)] ∝Q- D[volume(p)
        D[temperature(p)] ∝Q+ D[heat(p)]
        D[pressure(p)] ∝Q+ D[amount-of(p)]
        D[pressure(p)] ∝Q- D[volume(p)]
        D[pressure(p)] ∝Q+ D[heat(p)]

;Instead of writing a constraint law, we use qualitative proportionalities
;to preserve the direction of physical effect.  The section on Causal
;reasoning explains why.
```

---

single process and simple relationships between quantities, resolving influences and limit analysis are simple and yield unique results. In more complex situations resolving influences and disambiguating the possibilities raised by limit analysis will require more information, as we will see below.

Fig. 21. Effects of State on Containment

```
;Contained stuff has states as well -

(∀ p ∈ piece-of-stuff
        (Contained-Gas(p) ↔ (Contained-Stuff(p) ∧ Gas(p)))
        ∧ (Contained-Liquid(p) ↔ (Contained-Stuff(p) ∧ Liquid(p)))
        ∧ (Contained-Solid(p) ↔ (Contained-Stuff(p) ∧ Solid(p))))

;Contained liquids have levels, which are tied to amounts
;and in turn (assuming an open container) determines pressure

(∀ c ∈ contained-liquid
        HasQuantity(c, level)
        ∧ level(c) ∝Q+ amount-of(c)
        ∧ (OpenContainer(space-of(location(c)))
                ⇒ pressure(c) ∝Q+ level(c)))
```

Fig. 22. A Process Description of Fluid Flow

This simple model does not describe the existence and behavior of the fluid within the fluid path.

```
process fluid-flow

Individuals:
        src a contained-fluid
        dst a contained-fluid
        path a fluid path, FluidPath(path, src, dst)

Preconditions:
        Aligned(path)

Quantityconditions:
        A[pressure(src)] > A[pressure(dst)]

Relations:
        Let flow-rate be a quantity
        flow-rate ∝Q+ (A[pressure(src)] - A[pressure(dst)])

Influences:
        I+(amount-of(dst), flow-rate)
        I-(amount-of(src), flow-rate)   ;


;A fluid path is aligned only if either it has no valves or every valve is open
(∀ p ∈ fluid-path
        ((Number-of-valves(p) = 0) ⇒ Aligned(p))
        ∧ ((number-of-valves(p) > 0) ⇒ (∀ v ∈ valves(p) open(v)) ↔ Aligned(p))
        ∧ ¬ (number-of-valves(p) < 0))
;A heat path is defined in terms of objects in contact, and aligned
;indicates that the contact is unbroken.  Note this ignores heating
;by radiation and by convection.
```

Fig. 23. Resolved Influences and Limit Analysis
The results of resolving influences and limit analysis for the situation involving two containers are summarized below. The individuals in the situation are labeled IS, the Process Structure by PS, and the Limit Hypothesis by LH.

```
Ds[Amount-of(WC)] = -1          Ds[Amount-of(WD)] = 1
Ds[Volume(WC)] = -1             Ds[Volume(WD)] = 1
Ds[Level(WC)] = -1              Ds[Level(WD)] = 1
Ds[Pressure(WC)] = -1           Ds[Pressure(WD)] = 1
Ds[Heat(WC)] = 0                Ds[Heat(WD)] = 0
Ds[Temperature(WC)] = 0         Ds[Temperature(WD)] = 0

Limit Analysis:

                  IS: WC, WD
                  PS: Fluid-Flow(WC, WD, P1)

LH: Pressure(WC) = Pressure(WD)           LH: A[Amount-of(WC)] = ZERO
IS: WC, WD                                IS: WD
PS: none                                  PS: none
```

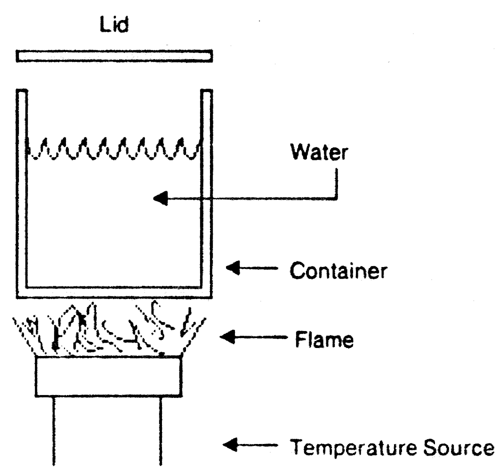
4.2 Modelling a Boiler

Let us consider the possible consequences of the situation shown in figure 24. The situation consists of a container partially filled with water. Initially the lid of the container is open; we will stipulate that if boiling ever occurs, the lid will be closed and sealed. A flame, modelled as a temperature source, is placed so that heat can be conducted to the container and water (i.e., there is an aligned heat path between them). What will happen?

The individuals we need are the contained substances defined in the previous example and the container. We will assume that if the pressure inside the container exceeds a particular pressure PBurst(can), the container will explode. Figure 25 illustrates. In addition to fluid flow, the Process Vocabulary will include heat flow and boiling presented in section 2. The rest of the details, such as the nature of heat and fluid paths and the detailed geometry of containers, will be ignored. As in all of the examples, we will assume domain-dependent results outside of Qualitative Process theory as needed. After all, the aim of this paper is to illustrate the ontological power of Qualitative Process theory rather than develop a full Naive Physics for several domains.

We will start by assuming that before the heat source is turned on that no processes are active; in

Fig. 24. A Simple Boiler



other words that all temperatures, pressures, etc. are equal so there are no flows, and that the temperatures are in the appropriate regions of their Quantity Spaces so that no state changes are occurring. (Note that we are making a closed world assumption both in assuming our Process Vocabulary is complete and that we know all of the Process Instances). Since there is a heat path between the source and the container, if we turn the heat source on and if `T(source) > T(water)` there will be a heat flow from the source to the water. We will ignore the influence of the heat flow on the source by assuming D_s`[T(source)]=0`. The only influence on `T(container)` is that of the heat flow, so D_s`[T(container)]=1`. This in turn will cause a heat flow to the air surrounding the container and to the air and the water inside the container. Since we are only thinking about the container and its contents most of these changes will be ignored, and from now on when we refer to heat flow it will be the flow from the flame to the contained stuff, using the container as the heat path. The temperature Quantity Space that results is illustrated in Figure 26. If `temperature(source) > t-boil(water)` and the process is unimpeded (i.e., the Preconditions for the heat flow remain true), the next Process Structure that will occur includes a boiling.

Suppose the Preconditions for the heat flow continue to be met and boiling occurs. Then by assumption the lid will be sealed, closing all fluid flow paths. The `amount-of` quantity space that results is illustrated in figure 27. The influence of the boiling on `amount-of(water)` moves it towards `ZERO`. So one of the ways the Process Structure might change is that all of the water is converted to steam.

Fig. 25. A Simple Container Model
For clarity we will model a container only as a collection of quantities, some contents, and a condition. The geometric information necessary to determine flow paths and the spatial arrangement of the contents will be ignored.

```
∀ c ∈ container
[HasQuantity(c, Volume) ∧ HasQuantity(c, Pressure)
∧ HasQuantity(c, Temperature) ∧ HasQuantity(c, Heat)
∧ (Rigid(c) ⇒ Ds[Volume(c)] = 0)
∧ (∀ p ∈ Contents(c)
        Pressure(c) = Pressure(p)
        ∧ Temperature(c) = Temperature(p))]
;note we are assuming instantaneous equilibration

Encapsulated History Explode

Individuals:
        c a container, rigid(c)
        e an episode

Preconditions:
        (T ClosedContainer(c) e)

QuantityConditins:
        (M A[pressure(c) end(e)) = (M A[p-burst(c)] end(e))
        (M A[pressure(c) during(e)) < (M A[p-burst(c)] during(e))

Relations:
        Let EV1 be an event
        end(e) = EV1
        Terminates(c, EV1)

;Terminates indicates that the object does not exist past
;that particular event
```

Fig. 26. Quantity Space for Water Temperature
The heat flow is increasing the heat, and thus (via $\propto_{Q+}$) the temperature of the water. The lack of ordering information between the temperature of the source and the boiling temperature leads to uncertainty concerning what will occur next.

```
t-melt(water)→temperature(water)→ temperature(source)
                               → t-boil(water)
```

Fig. 27. Amount-of Quantity Spaces
`initial` is the name for the start of the boiling episode. If we have deduced that no flows can occur and believe in conservation of mass, `amount-of(steam)` will also have `(M amount-of(water) Initial)` as the top of its Quantity Space.

```
ZERO→amount-of(water)→(M A[amount-of(water)} Initial)

ZERO→amount-of(steam)
```

If all the water is converted to steam, the only active process is a heat flow from the heat source to the steam. Thus the sole influence on the heat of the steam is positive, and (because of $\propto_Q$) the temperature also rises. Heat indirectly influences pressure, so the pressure of the steam will also rise. There are two Quantity Conditions that may be reached, namely that the temperature can reach that of the heat source and that the pressure can reach the explosion point. In the first case there are no active processes, and in the second an explosion occurs. We have found one possible disaster, are there more? To find out, we must go back to the boiling episode and check the indirect consequences of the changes in `amount-of(steam)`.

Consider an arbitrary slice within the boiling episode. Because the steam is still in contact with the water their temperatures will be the same. Since we assumed the container would be sealed when boiling began, there are no fluid paths hence no fluid flows. Therefore the only influence on `amount-of(steam)` is from boiling. So D_s`[amount-of(steam)]=1` and D_s`[amount-of(water)]=-1`.

Because steam is a gas, there are several indirect influences on `temperature(steam)` and `pressure(steam)` (see figure 20). In particular,

$$D[\texttt{temperature(steam)}] \propto_{Q+} D[\texttt{amount-of(steam)}]$$
$$D[\texttt{temperature(steam)}] \propto_{Q-} D[\texttt{volume(steam)}]$$
$$D[\texttt{pressure(steam)}] \propto_{Q+} D[\texttt{amount-of(steam)}]$$
$$D[\texttt{pressure(steam)}] \propto_{Q-} D[\texttt{volume(steam)}]$$

Assuming the container is rigid, D_s`[volume(can)]=0`, and since the spaces of the steam and water are separate and fill the container,

```
volume(can) = volume(water) + volume(steam)
```

Since D_s`[volume(water)]=-1`, D_s`[volume(steam)]=1` and D_m`[volume(steam)]=`D_m`[volume(water)]`.

For any particular `D[amount-of(steam)]`, we can find a `D[volume(steam)]` such that D_s`[pressure(steam)]=0` (assuming the underlying multivariate function is continuous). Call that value of `D[volume(steam)]` β. A fact about steam is that, at any particular pressure and temperature, the volume of steam is very much greater than the volume of water it was produced from.[1] In other words,

$$D_s[\texttt{pressure(steam)}]=0 \Rightarrow D_m[\texttt{volume(water)}] \ll D_m[\texttt{volume(steam)}].$$

So in fact

1. At standard temperature and pressure, about 220 times greater in fact.

D_m[volume(steam)]=D_m[volume(water)] $\Rightarrow$ D[volume(steam)]< β.

Then by $\propto_{Q-}$, D_s[pressure(steam)]=1. A similar argument holds for temperature.

So far we have discovered that

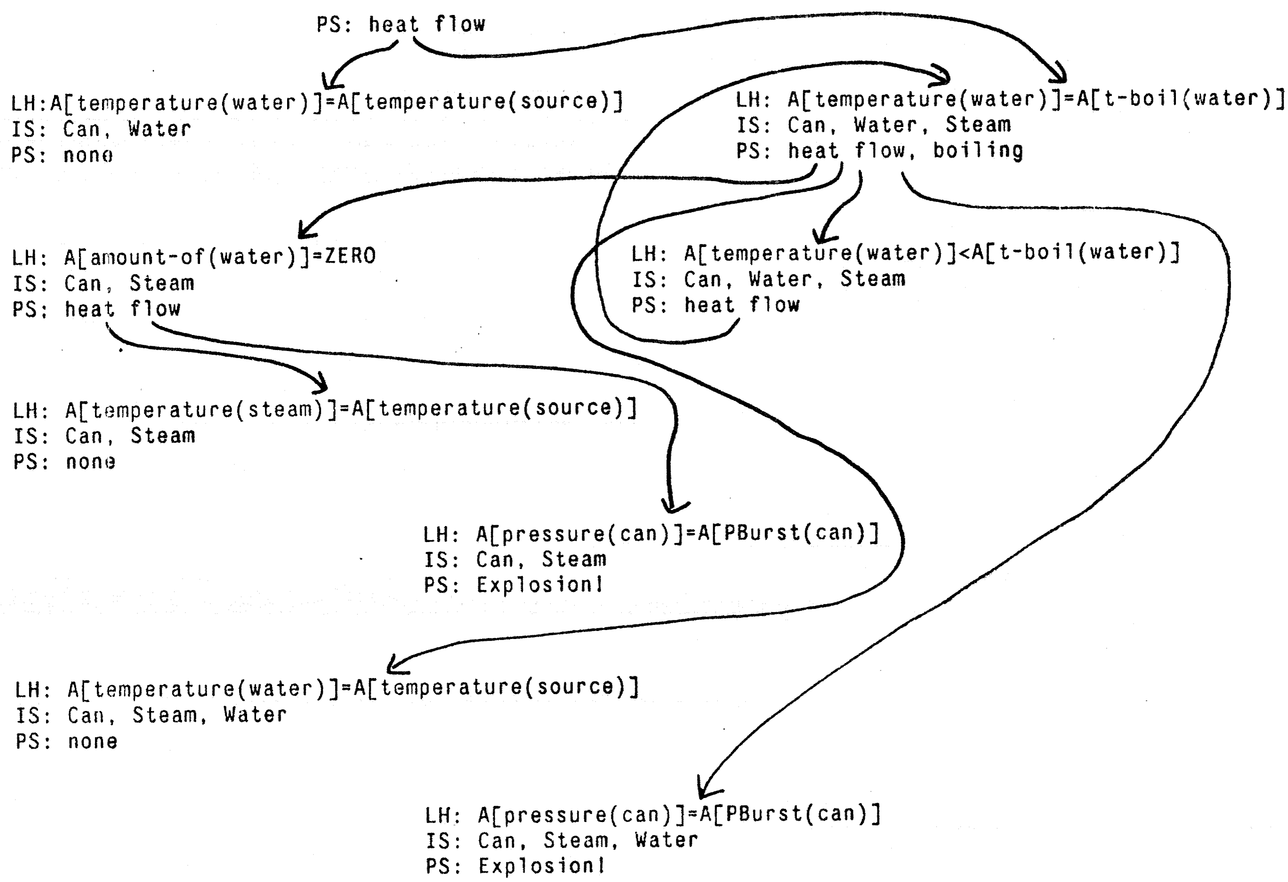
D_s[pressure(steam)] = D_s[temperature(steam)]=1.

Since the water and steam are in contact their pressures will be equal, and since pressure indirectly affects the boiling temperature, the boiling temperature will also rise. The possible relative rates introduce three cases. If the boiling temperature is rising faster (D_m[t-boil(water)] > D_m[temperature(steam)]) then the boiling will stop, the heat flow will increase heat(water) again, the temperature will rise, and the boiling will begin again.[1] In the other two cases the boiling will continue, albeit at a higher temperature and pressure. In any case, the increasing pressure makes A[pressure(can)] = A[PBurst(can)] possible, in which case the container explodes. The alternatives are summarized in figure 28. To actually determine which of these occurs requires more information, but at least we have a warning of potential disaster.

---

1. The astute reader will notice that this situation actually involves a rising equilibrium rather than an oscillation. We will discuss how to use the Equality Change Law to distinguish between these cases later.

Fig. 28. Alternatives for Sealed Container
Here are the Process Structures envisioned for water being heated in a sealed container, generated by repeated limit analysis.



4.3 Modelling Motion

One process we reason about every day is motion. Motion is complex because it is intimately connected with our concepts of space and shape. Since Qualitative Process theory describes the form of qualitative dynamics theories, it can only carry part of the representational burden imposed by motion. After developing a simple motion vocabulary, we will compare the QP descriptions with the earlier Qualitative State representation in order to illustrate the strengths and weaknesses of the Qualitative Process theory model.

4.3.1 A Simple Motion Vocabulary

We will consider the case of a single object moving in one dimension for illustration. The vocabulary will be abstract, in that the particular information regarding the type of motion (FLY, SLIDE, SWING, ROLL) which depends on shape and type of contact with other objects will be ignored. While very weak, such descriptions still have their uses -- we can deduce that if we kick something and it isn't blocked, for instance, then it will move, and if we can rule out the most abstract motion possible we have ruled out all the more specific kinds.

Figure 29 contains the process specifications for motion and acceleration. The motion process occurs when a movable object is free in the direction of its velocity (i.e., no other objects in the way), when that velocity is non-zero. Motion is a positive influence on position of an object, in that if the velocity is positive the position will "increase" and if the velocity is negative the position will "decrease". (We will worry a little about mapping a Quantity Space onto geometric frames of reference in a moment). Acceleration occurs when a movable object has some non-zero net force in a free direction. Acceleration provides a positive influence on velocity, and in fact the influence is qualitatively proportional to the net force and qualitatively inversely proportional to the mass of the object (Newton's Second Law).

While this description is Newtonian, Aristotelian and Impetus theories can also be described.[1] One form of Aristotelian motion, for example, can be written as in Figure 30. Here motion only occurs when an object is being pushed. An Impetus theory is described in figure 31. Aristotelian theory has the problem of explaining what keeps a moving object going once it didn't touch anything else; Impetus theory explains this by the push giving an object a kind of internal force or "impetus". While superficially like momentum, impetus kinematics is very different.[2] Also impetus differs in that it can spontaneously dissipate. Compare the dissipation of impetus with the Newtonian model of sliding friction in Figure 32. Here friction occurs when there is surface contact, and produces a force on the object that is qualitatively proportional to the normal force and acts in a direction opposite that of the

1. [McClosky, 1983].and [Clement, 1983] argue that naive theories of motion in our culture correspond to Impetus theories, rather than Aristotelian theories as previously suggested [diSessa, 1982].
2. In particular, impetus is not a vector quantity. The means of combination vary across subjects; they include things like "The motion is in the direction of the biggest one". There are other oddities as well -- for example, impetus "remembers" not just the direction of the push but some of the previous history of directions, so that leaving a spiral tube will result in spiral motion for a little while. See [McCloskey, 1983].

Fig. 29. Process Descriptions of Motion and Acceleration
Note that we are not assuming that motion is the sole influence on position. This motion vocabulary is quite abstract, ignoring the kind of motion occurring and the complexities of motion in more than one dimension. Assigning quantities to frames of reference will not be discussed here.

```
Process Motion(B,dir)

individuals:
        B an object, Mobile(B)

Preconditions:
        Free-direction(B, dir)
        Direction-Of(dir, Vel(B))

QuantityConditions:
        Am[Vel(B)] > ZERO

Influences:
        I+(position(B), Vel(B))

-------------------------------------------

Process Acceleration(B,dir)

Individuals:
        B an object, Mobile(B)
        dir a direction

Preconditions:
        free-direction(B,dir)
        Direction-Of(dir, Vel(B))

QuantityConditions:
        Am[net-force(B)] > ZERO

Relations
        Let Acc be a quantity
        Acc ∝Q+ net-force(B)
        Acc ∝Q- mass(B)
        ; The basic QP version of F = m * a
        Correspondence((Acc ZERO)
                          (net-force(B) ZERO))
Influences: I+ (Vel(B) Acc)
```

motion (encoded by the sign difference). The effect of friction occurs indirectly, through changing acceleration rather than directly as in the Impetus theory.

Collisions are complicated in any theory. The reason collisions are complicated is that they are usually described in terms of a piece of history. We will use an Encapsulated History, as described in Section 2. The simplest description of a collision just involves a reversal of velocity, as illustrated in figure 33. Here `Direction-Towards(C, B, dir)` asserts that the object is moving in direction `dir` from `C` to `B`, and

Fig. 30. Aristotelian Motion
Aristotle theorized that objects required a constant push to keep them going. Note that `vel` does not have an existence independent of the motion process.

```
Process Motion

Individuals:
        B an object, Mobile(B)
        dir a direction

Preconditions:
        Free-Direction(B, dir)
        Direction-Of(dir, net-force(B))

QuantityConditions:
        Am[net-force(B)] > ZERO

Relations:
        let vel be a quantity
        vel ∝Q+ net-force(B)
        vel ∝Q- mass(B)

Influences:
        I+(position(B), vel)
```

`Contact(B, C)` asserts that `B` and `C` touch. As a simplification we have assumed `C` is immobile so that we won't have to worry about momentum transfer between moving objects and changes of direction in more than one dimension. Even our more complicated models of collisions appear to use such Encapsulated Histories, such as a compound history consisting of contacting the surface, compression, expansion, and finally breaking contact. The type of collision that occurs can be specified by referring to the theory of materials for the objects involved.

4.3.2 Relationship to Qualitative States

The Qualitative State representation for motion [de Kleer, 1975][Forbus, 1981a] is an abstraction of the notion of state in classical mechanics. Certain parts of the classical state are represented abstractly (typically position is represented by a piece of space, and velocity by a symbolic heading) and the type of activity, which classically is implicit in the choice of descriptive equations, is made explicit. Qualitative states are linked by qualitative simulation rules that map a qualitative state into the qualitative states that can occur next. These rules are started from some initial state and run until no new states are generated, producing a description of all the possible states called the *envisionment*. Figure 34 illustrates. The envisionment can be used to answer simple questions directly, assimilate certain global assumptions about

Fig. 31. An Impetus Dynamics for Motion

In impetus theories of motion, a push imparts "impetus" to an object. An object's impetus is an internalized force that keeps on pushing the object, thus causing motion. Motion eventually stops because impetus dissipates with time.

```
Process Motion

Individuals:
        B an object, Mobile(B)
        dir a direction

Preconditions:
        Free-Direction(B, dir)
        Direction-Of(dir, impetus(B))

QuantityConditions:
        Am[impetus(B)] > ZERO

Relations:
        let vel be a quantity
        vel ∝Q+ impetus(B)

Influences:
        I+(position(B), vel)
```

```
Process Impart

Individuals:
        B an object, Mobile(B)
        dir a direction

Preconditions:
        Free-Direction(B, dir)
        Direction-Of(dir, net-force(B))

QuantityConditions:
        Am[net-force(B)] > ZERO

Relations:
        Let acc be a quantity
        acc ∝Q+ net-force(B)
        acc ∝Q- mass(B)

Influences:
        I+(impetus(B), acc)
```

```
Process Dissipate

Individuals:
        B an object, Mobile(B)
        dir a direction

QuantityConditions:
        Am[impetus(B)] > ZERO

Relations:
        Let acc be a quantity
        As[acc] = As[impetus(B)]

Influences:
        I-(impetus(B), acc)
```

---

motion, and plan solutions to more complex questions. By examining the relationship between the Qualitative State representation and the Qualitative Process theory representation we will understand both more clearly.

Suppose motion and acceleration are the only processes that occur. The Limit Analysis for a moving object will include only the possibilities raised by dynamics. To include the possible changes in process caused by kinematics (i.e., being in different places, colliding with other objects) the relevant geometry of the situation must be removed from the Preconditions and mapped into a Quantity Space.

Fig. 32. Moving Friction in Newtonian Sliding

```
Individual View  Moving-Friction

Individuals:
        B an object, Mobile(B)
        S a surface

Preconditions:
        Sliding-Contact(B,S)

QuantityConditions:
        Motion(B,along(S))

Relations:
        Let fr be a quantity
        fr ∝Q+ normal-force(B)
        As[fr] = Opposite(tangent-force(B))
        Friction-force(fr)
        fr ∈ ForcesOn(B)
```

Fig. 33. Collision Specification

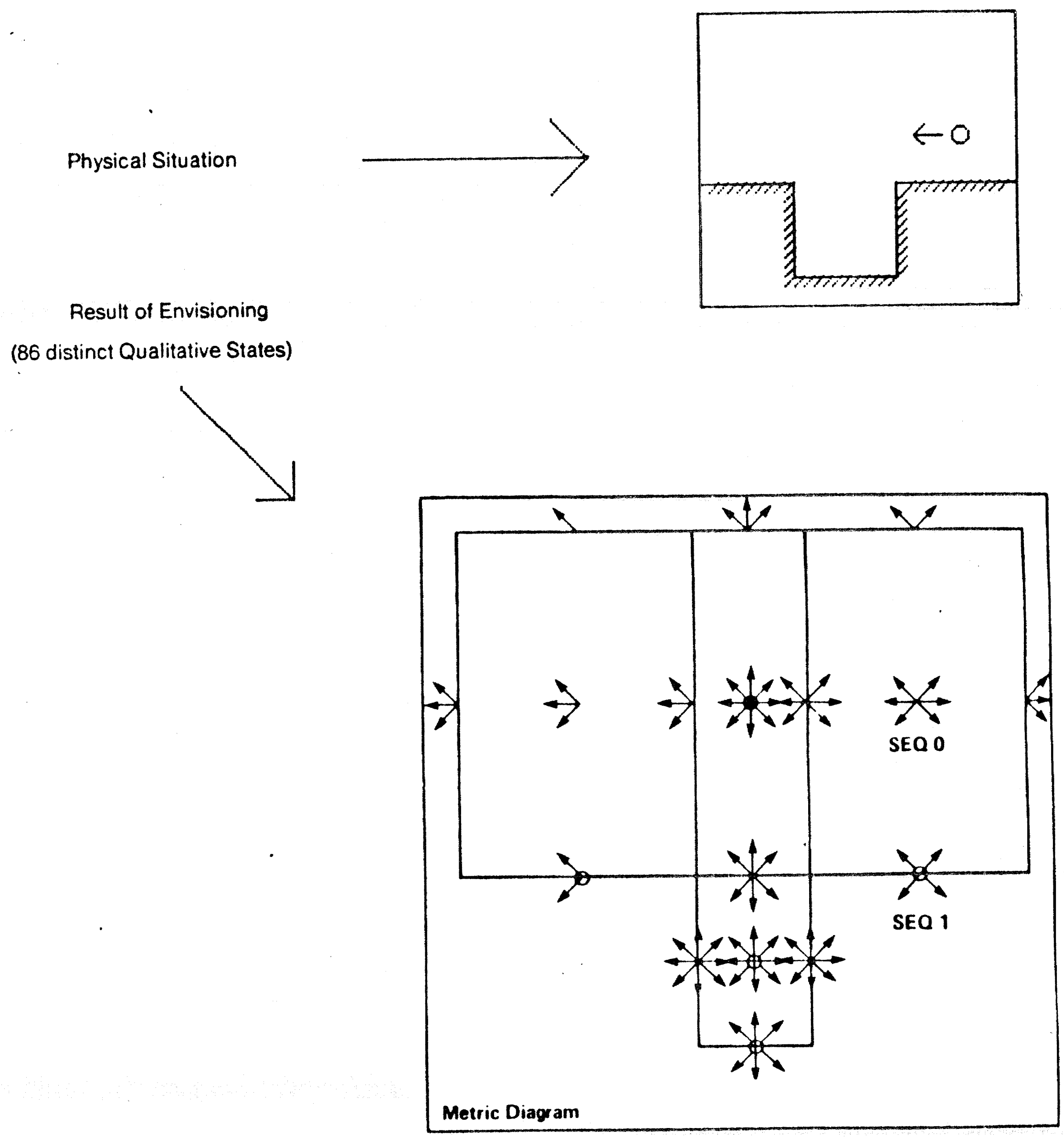
Unless specified otherwise, the Preconditions and Quantity Conditions refer to the start of the episode described by the relations field.

```
Encapsulated-History Collide(B,C,dir)

Individuals:
        B an object, Mobile(B)
        C an object, ¬ Mobile(B)

Preconditions:
        (T contact(B,C) start(E1))
        (T direction-towards(B,C,dir) start(E1))

QuantityCondition:
        (T Motion(B,dir) start(E1))

Relations:
        There is E1 an event
        (M Vel(B) start(E1)) =  -(M Vel(B) end(E1))
        (M Vel(B) during(E1)) =  ZERO
        duration = ZERO
        (T direction-towards(C,B,dir) end(E1))
        (T contact(B,C) end(E1))

; (T <statement> <time>) means "<statement> is true during <time>"
```

Fig. 34. Qualitative State Description of Motion

These Qualitative States consists of a type of motion, a PLACE (a symbolic description of pieces of space, in this case computed from an analog representation), and a symbolic heading. The sample Qualitative States are circled on the depiction of the envisionment below.

```
              SEQ0: FLY Space-Region3 (Left Down)
     leads to
              PASS Segment12 (Left Down)
              COLLIDE Segment3 (Left Down)

              SEQ1: COLLIDE Segment3 (Left Down)
     leads to
              SLIDE/STOP Segment3 (Left Down)
              FLY Segment3 (Left Up)
```



This requires describing space by a *Place Vocabulary*,[1] and using the elements in the Place Vocabulary as the elements in the position Quantity Space. To induce an ordering between the elements for motion in two and three dimensions a direction must also be included in the process description, since partial orders are only well-defined for one dimension. The ambiguity due to dimensionality and symbolic heading can be encoded by the lack of ordering between the Quantity Space elements. This also means the place must be encoded in the motion process, which in turn means that an instance of a motion process in this vocabulary will look like a Qualitative State for the same collection of places and type of motion. The qualitative simulation rules thus correspond to a compilation of the Limit Analysis on this new motion vocabulary.

From this perspective we can see the relative strengths of the two representations. For evolving descriptions of motion the Qualitative State representation seems superior, because kinematic constraints are essential to motion. However, simulation rules are an opaque form of dynamics theory -- they do not contain the assumptions under which they operate. Thus the "compiled" nature makes the Qualitative State representation inappropriate for very simple deductions (where only part of a Qualitative State is known), or for more subtle analyses that involve perturbing a system. In particular, the Qualitative State representations for motion are not easily composable to form descriptions of more complex systems.[2] A later example will illustrate a more subtle analysis of motion made possible by the ontology of Qualitative Process theory.

1. [Forbus, 1981a] describes the principles involved and defines a place vocabulary for motion through space in a simple domain.

2. This is not true for the Qualitative State representations for engineered systems used in de Kleer and Brown [de Kleer and Brown, 1982]. The difference lies in their use of an even weaker notion of quantity than the one provided in Qualitative Process theory to connect descriptions. State is confined to device models, and device models are linked explicitly by shared parameters to form more complicated systems. In this case the descriptions are highly composable - in fact, the scope of their descriptions is exactly the same class of devices as classical Systems Dynamics formalisms. The Qualitative State descriptions for motion, however, have not had this character.

4.4 Modelling Materials

Let us consider what happens when we pull on something. If it doesn't move, then its internal structure is "taking up" the force (this can happen even if it does move - try hitting an egg with a baseball bat - but we will ignore this case). Three things can happen - (1) it can do nothing (rigid behavior), (2) it can stretch (elastic behavior) or (3) it can break. For a push, (2) becomes compression and (3) becomes crushed. We can use the notions of quantity and process provided by Qualitative Process Theory to state these facts. In particular, we can express the changes between these kinds of behavior by creating a Quantity Space for forces on an object.

The concepts involved with elasticity can be thought of in terms of applied force verses internal force. If the magnitude of the applied force is greater than that of the internal force the length of the object will change. The change in length will result in an internal force that will counterbalance the applied force. There will be three individual views for an elastic object, corresponding to whether or not it is stretched, relaxed, or compressed. Figure 35 illustrates the Individual View for elastic objects and their states. To avoid the complications of shape and connectivity, we will only model one-dimensional objects that have one end fixed. By convention, forces into an object (pushes) will be negative and applied forces directed outwards (pulls) will be positive.

An imbalance between internal and applied forces will result in the length changing. Exactly what occurs depends on the state of the elastic object (stretched, relaxed, compressed), the sign of the applied force, and the relative magnitudes of the forces (the dependence on the sign of the internal force is encoded in the state of the object via the $\propto_Q$ and correspondence.). The four possibilities are stretching, compressing, and two kinds of relaxing. These processes are described in Figure 36.

Of course, objects are not perfectly elastic. If the applied force is very small, objects will often behave rigidly. If too much force is applied an object can break or crush. The rigidity under small forces can be modeled by adding another Quantity Condition to stretching and compressing. For a partially elastic object the thresholds for compressing and stretching will be called $f_{compress}$ and $f_{stretch}$ respectively. The conditions under which crushing and breaking can be captured similarly by thresholds f_{crush} and f_{break}, which are functions of both the material and the object (to allow for dependence on the shape). The process descriptions for crushing and breaking are however more complex than compressing and stretching because they involve discontinuous change. This requires statements in the relations field that explicitly mention time, turning the description into an encapsulated history rather

Fig. 35. Descriptions of Elastic Objects
Intuitively an elastic object stores energy in terms of deformations of shape that are more or less reversible. The basic view of an elastic object relates the internal force and length, and the other three views describe the states it can be in.

```
Individual-View Elastic-Object

Individuals:
        B a physical object

Preconditions:
        Elastic-Substance(made-of(B))

Relations:
        Has-Quantity(self, length)
        Has-Quantity(self, InternalForce)
        Has-Quantity(self, rest-length)
        Ds[rest-length(B)] = 0
        internal-force(B) ∝Q+ length(B)
        Correspondence((internal-force(B) ZERO)
                       (length(B) rest-length(B)))
------------------------------------------------

Individual-View Relaxed

Individuals:
        B an elastic object

QuantityConditions:
        A[length(B)] = A[rest-length(B)]
------------------------------------------------

Individual-View Stretched

Individuals:
        B an elastic object

QuantityConditions:
        A[length(B)] > A[rest-length(B)]
------------------------------------------------

Individual-View Compressed

Individuals:
        B an elastic object

QuantityConditions:
        A[length(B)] < A[rest-length(B)]
```

---

than a true process. The statements that must be included in them concern deformation of shape and the transformation of one object into several. As with kinematics, these issues are beyond the scope of Qualitative Process theory.

Fig. 36. Stretching, Compressing, and Relaxing
The continuous changes that can occur to elastic objects are described below. The individual views of stretched, compressed, and relaxed are described in the previous figure.

```
process Stretching

Individuals:
        B an elastic object

QuantityConditions:
        ¬ Compressed(B)
        As[applied-force(B)] = 1
        Am[applied-force(B)]
           > Am[internal-force(B)]

Relations:
        Let SR be a quantity
        SR ∝Q+ (Am[applied-force(B)]
                - Am[internal-force(B)])

Influences:
        I+(length(B), SR)
```

```
process Relaxing-Minus

Individuals:
        B an elastic object

QuantityConditions:
        Stretched(B)
                Am[applied-force(B)]
           < Am[internal-force(B)]

Relations:
        Let SR be a quantity
        SR ∝Q+ (Am[applied-force(B)]
                - Am[internal-force(B)])

Influences:
        I-(length(B), SR)
```

```
process Compressing

Individuals:
        B an elastic object

QuantityConditions:
        ¬ Stretched(B)
        As[applied-force(B)] = -1
        Am[applied-force(B)]
           > Am[internal-force(B)]

Relations:
        Let SR be a quantity
        SR ∝Q+ (Am[applied-force(B)]
                - Am[internal-force(B)])

Influences:
        I-(length(B), SR)
```

```
process Relaxing-plus

Individuals:
        B an elastic object

QuantityConditions:
Compressed(B)
                Am[applied-force(B)]
           < Am[internal-force(B)]

Relations:
        Let SR be a quantity
        SR ∝Q+ (Am[applied-force(B)]
                - Am[internal-force(B)])

Influences:
        I+(length(B), SR)
```

Figure 37 illustrates the force Quantity Spaces that will result for different kinds of materials. In theory a taxonomy such as this one could be used for classifying a material by applying forces to it and seeing what sorts of behavior result. In a richer model of materials forces along different directions could result in different behavior (such as attempting to bend balsa wood against its grain instead of along the grain) and the effects of plastic deformation would be included.

A classic AI conundrum is to be able to express in some usable form that "you can pull with a

Fig. 37. Materials Classified by Quantity Spaces
The distinct kinds of materials give rise to different Quantity Spaces because different kinds of processes can occur. This taxonomy in theory allows a material to be classified by applying forces and observing what kinds of things actually occur.

Rigid:
<no processes affecting length>

Elastic:
<stretching and compressing apply>

Breakable:
ZERO -> f_{break}

Crushable:
f_{crush} -> ZERO

Partially stretchable:
ZERO -> $f_{stretch}$

Partially compressible:
$f_{compress}$ -> ZERO

Brittle:
f_{crush} -> ZERO -> f_{break}

Partially elastic:
$f_{compress}$ -> ZERO -> $f_{stretch}$

Normal:
f_{crush} -> $f_{compress}$ -> ZERO -> $f_{stretch}$ -> f_{break}

---

string, but not push with it".[1] This fact can be succinctly stated, at least to a first approximation, using Qualitative Process theory. First, consider what pushes and pulls are. Both concepts imply one object making contact with another to apply force. If the direction of the applied force is towards the object it is a push, and if the direction is away from the object then it is a pull. Obviously a push can occur with any kind of contact, but pulls cannot occur with an abutting.

Understanding how pushes and pulls are transmitted is fundamental to understanding mechanisms. For a first pass model, consider the notion of *push-transmitters* and *pull-transmitters.* We will say an object is a push transmitter if when it is pushed, it will in turn push an object that is in contact with it, in the direction between the two contact points. Pull transmitters can be similarly defined. This

1. Marvin Minsky, personal communication

particular set of definitions is obviously inadequate for mechanisms,[2] and is only for illustration. Note also that push-transmitters and pull-transmitters need not be reflexive relations. Rigid objects are an exceptional case:

```
∀ B ∈ object
   rigid(B) ⇒   (∀ c1, c2 ∈ contact-points(B)
                    Push-Transmitter(c1, c2)
                  ∧ Push-Transmitter(c2, c1)
                  ∧ Pull-Transmitter(c1, c2)
                  ∧ Pull-Transmitter(c2, c1))
```

Strings, however, are more complicated. A string can never be a push-transmitter:

```
∀ s ∈ string
 (∀ t ∈ time (T (¬ Push-Transmitter(end1(s), end2(s))
                                         ∧ ¬ Push-Transmitter(end2(s), end1(s))) t))
```

But if it is taut it can be a pull transmitter:

```
∀ s ∈ string
(∀ t ∈ time
        (T taut(s) t) ⇒ (Pull-Transmitter(end1(s), end2(s))
                               ∧ Pull-Transmitter(end2(s) end1(s))))
```

Now the problem becomes how to define taut. As a first pass, we can define `Taut` as an individual view:

```
Individual View Taut
        Individuals:
                s a string
        Quantity Conditions:
                ¬ A_m[ends-distance(s)] < A_m[length(s)]
```

This model assumes that only the ends of the string contact other objects - it would fail for a rope hanging over a pulley, for instance. A better model is to divide up the string into segments according to whether or not that part of the string is in contact with a surface. A string is then taut if each segment that is not in contact with a surface is taut:

```
∀ S ∈ string
        (∀ seg ∈ segments(geometry(S)) Free-Segment(seg, S) ⇒ taut(seg))
        ⇒ taut(S)
```

This of course ignores the fact that the non-free segments may not be tight, as say for string lying on the floor. A full definition would also require tension along the entire string, but we have strayed far enough

---

2. Consider for example a rocker arm or an object that is tied to a string by another object. In the first case a push will be transmitted in a different direction, and in the second case it will be transformed into a pull. Kinematic issues are involved, along with resolving forces in more than one dimension.

from dynamics already.

4.5 An Oscillator

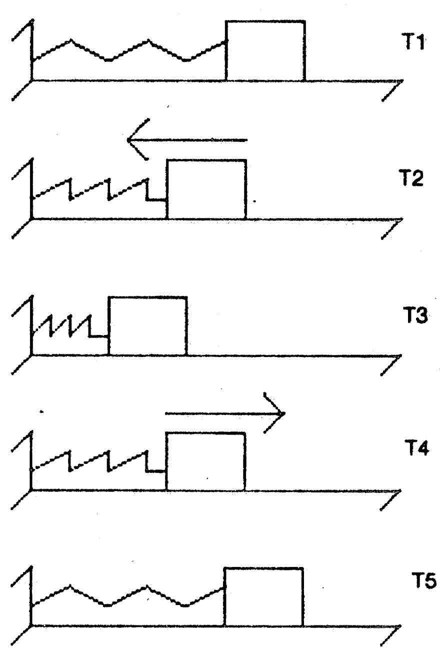
One purpose of reasoning about the physical world is to produce descriptions of how a system behaves and summaries of its eventual disposition(c.f. [Forbus, 1981a]). In classical physics these analyses are often concerned with stability. Here we will examine a simple situation involving motion and materials, ascertain that it is an oscillator, and perturb it to figure out under what conditions it will remain stable.

Consider the block B connected to the spring S in figure 38. Suppose the block is pulled back so that the spring is extended. Initially we will also assume that the contact between the block and the floor is frictionless. What will happen?

We will model the spring S as device satisfying Hooke's Law (see figure 35). Since

---

Fig. 38. A Sliding Block
Here is a system we will analyze to determine what it does and how different factors, such as whether or not there is friction, affect its behavior.

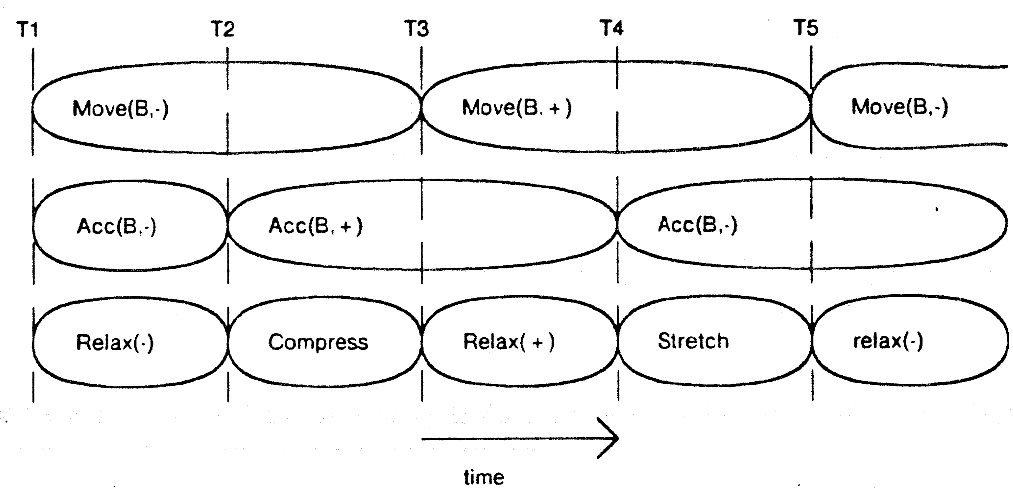


displacement(S) is greater than ZERO, the spring will exert a force. Because the block is rigidly connected to the spring, the net force on the block will be negative and since the block is free to move in the direction of the force, an acceleration will occur. The acceleration will in turn cause the velocity to move from ZERO, which will in turn cause D_s[position(B)]=-1. By rigid contact, D_s[length(S)]=-1 and by the $\propto_Q$ relation with displacement, D_s[net-force(S)]=1. The processes occurring are Motion(B, -1), Relaxing-Minus(S), Acceleration(B, -1). The next process limit occurs when length(S)=rest-length(S), ending the relaxing. The correspondence tells us the force on the block becomes ZERO, so the acceleration will end as well. However, the motion does not. Setting aside the details, the next set of processes are Motion(B, -1), Compressing(S), and Acceleration(B, +1). The only limit point in the Quantity Spaces that are changing is the zero velocity point (assuming the spring is unbreakable), so the motion will continue until the velocity is zero. The conclusion that the next set of processes are Motion(B, +1), Relaxing-Plus(S), Acceleration(B, +1) and then Motion(B, +1), Stretching(S), Acceleration(B, -1) follows in the same way. At the end event of the last set of processes, the orderings on the quantity spaces and the processes evoked are the same as the initial instant. Thus we can conclude that an oscillation is occurring. Note that the active processes need to be the same, because the Preconditions can change. Figure 39 illustrates the process history for the oscillator.

Some of the assumptions made in producing the process history can now be perturbed to

---

Fig. 39. Process History for the Oscillator



examine the effects of different physical models. For instance, suppose the spring is crushable and breakable, as defined previously. Then there are limit points around `rest-length(S)` that correspond to the occurance of crushing or breaking. It seems crushing must be ruled out by assumption, since the machinery we have developed so far does not allow us to rule it out via contradiction. We can however deduce that the spring won't break under the conditions above.

If we can prove that the block will go out no further than when it started then we can claim that it won't break because it didn't break in the first place. This requires an energy argument. The energy theory we will use is very simple. There are certain types of quantities that are energy-quantities, which are qualitatively proportional to certain other quantities and exist whenever they do. Two kinds of energy are kinetic energy and "spring" energy. For every object there is a total energy, which is the sum of its energy quantities (Figure 40 describes systems and energy quantities more formally, and and figure 41 describes sources, sinks, and conservation laws.)

Here the system is the mass and spring combination. At time `t1` the block is still but the spring is stretched, i.e.,

```
(M A[Velocity(B)] t1) = ZERO
(M Disp(S) t1) > ZERO
```

which means that

```
(M total-energy(system) t1) > ZERO
```

If energy is conserved and there is no influx of energy (i.e., a pattern of influences involving energy quantities where one such quantity inside the system is positively influenced and another such quantity outside the system is negatively influenced), then we know

```
∀ t ∈ time
        After(t, t1) ⇒ ¬ (M total-energy(system) t) > (M total-energy(system) t1)
```

This means that the block can only go out as far as it was at `t1`, since if it ever went out farther we would contradict the previous statement.

Fig. 40. A Simple Energy Theory - Energy & Systems

The predicate `EnergyQuantity` asserts that its argument is a quantity representing a kind of energy. Energy quantities occur as a consequence of other quantities in particular kinds of objects. The energy of a system is the sum of the energy quantities for its parts.

```
EnergyQuantity(kinetic-energy)

;motion gives rise to kinetic energy

∀ B ∈ object
        HasQuantity(B, velocity) ⇒
                (HasQuantity(B, kinetic-energy)
                 ∧ kinetic-energy(B) ∝Q+ velocity(B)
                 ∧ Corrspondence((kinetic-energy(B) ZERO),
                                 (velocity(B) ZERO)))

EnergyQuantity(SpringEnergy)

;an internal force gives rise to "spring" energy

∀ B ∈ object
        HasQuantity(B, internal-force) ⇒
                (HasQuantity(B, spring-energy)
                 ∧ spring-energy(B) ∝Q+ internal-force(B)
                 ∧ Correspondence ((spring-energy(B) ZERO),
                                   (internal-force(B) ZERO)))

;the total energy of an object is the sum of its energy quantities

∀ B ∈ object
        HasQuantity(B, TotalEnergy) ∧ Set(energy-quantities(B))
        ∧ ∀ q ∈ quantities(B) EnergyQuantity(q) ⇒ q(B) ∈ energy-quantities(B)
        ∧ TotalEnergy(B) = sum-over(energy-quantities(B))

;the energy of a system is the sum of the energy in its objects

∀ sys ∈ system Set(objects(sys))
        ∧ (∀ b ∈ objects(sys) Physob(B) ∨ System(B))
        ∧ HasQuantity(sys, total-energy)
        ∧ Set(energy-quantities(sys))
        ∧ (∀ B ∈ objects(sys)
                Subset(energy-quantities(B), energy-quantities(sys)))
        ;ignore converse case for now (all members must be from some part)
        ∧ TotalEnergy(sys) = sum-over(energy-quantities(sys))
```

Fig. 41. A Simple Energy Theory - Sources, Sinks, and Conservation

There are several forms of energy conservation, some stronger than others. The weakest says that if there is no inflow then at a later time the energy is always the same or smaller than an earlier time. The strongest says that in a closed system the energy is always the same.

```
;processes can be sources and sinks w.r.t. a system

∀ pi ∈ process-instance ∀ sys ∈ system ∀ q ∈ quantity-type
        Source(pi, sys, q) ≡
        (∃ B ∈ objects(sys) Influences(pi, Q(B), +1))
        ∧ ¬(∃ B ∈ objects(sys) Influences(pi, Q(B), -1))
;define sinks similarly, and CROSS-FLOW where there
;is both a negative and positive influence

∀ pi ∈ process-instance ∀ sys ∈ system
        EnergySource(pi, sys) ≡ (∃ q ∈ quantity EnergyQuantity(q) ∧ Source(pi, sys, q))
                ∧ (∀ q ∈ quantity EnergyQuantity(q) ⇒ ¬ Sink(pi, sys, q))
;energy sinks and cross-flows are defined analogously

;simple form of conservation:

;Processes locally conserve energy
∀ pi ∈ process-instance
        (∃ q1 ∈ quantity
                EnergyQuantity(q1) ∧ Influences(pi, q1, -1))
        ↔ (∃ q2 ∈ quantity
                EnergyQuantity(q2) ∧ Influences(pi, q2, 1))

;if you don't kick it it won't get any higher..

∀ sys ∈ system ∀ i ∈ interval
(T ¬ (∃ pi ∈ process-instance EnergySource(pi, sys)) i) ⇒
        ¬ (M total-energy(sys) end(i)) > (M total-energy(sys) start(i))

;more complex version:

∀ sys ∈ system ∀ i ∈ interval
[(T ¬ (∃ pi ∈ process-instance EnergySource(pi, sys)) i)
∧ (T ¬ (∃ pi ∈ process-instance EnergySink(pi, sys)) i)
∧ (T ¬ (∃ pi ∈ process-instance EnergyCrossFlow(pi, sys)) i)]
        ⇒ (M A[TotalEnergy(sys)] start(i)) = (M A[TotalEnergy(sys)] end(i))
```

4.5.1 Stability Analysis

To further analyze this system, we must treat the processes that occur as a compound process. We can start by writing an Encapsulated History, but we will also need to define properties of the objects over the piece of history so defined. These will include the total energy of the system, the energy lost during a cycle (one occurance of the piece of history), and the maximum displacement during this cycle. The rationale for choosing these parameters of the collection lies in the energy theory presented above. We will assume the Relations for the compound process include:

```
MaxDisp(S) ∝Q+ total-energy(system)
correspondence((MaxDisp(S), ZERO), (total-energy(system) ZERO))
```

To perform an energy analysis we will also want to re-write any inequalities in the Quantity Conditions in terms of energy, to wit:

```
QuantityConditions:
        A[total-energy(system)] > ZERO
```

Thus if the total energy of the system ever reaches ZERO during an occurance of the compound process it will no longer be active. Note that the Quantity Condition is no longer tied to a particular episode of the Encapsulated History. This means that, unlike the Encapsulated Histories previously encountered, we cannot use this one for simulation. Instead, we will use it to analyze global properties of the system's behavior.

We can use the basic QP deductions on this new description to determine the consequences of perturbing the situation in various ways. Each perturbation is represented by a process that influences one of the parameters that determines the energy of the system. For example, if friction were introduced(i.e., D_s[total-energy(system)]=-1), by the relationship above we can deduce (via Limit Analysis) that the oscillation process will eventually stop, and that if the system is pumped so that its energy increases (i.e., D_s[total-energy(system)]=1), that the materials involved in the oscillator may break in some way.[1] Suppose for example the oscillator is subject to friction, but we pump it with some fixed amount of energy per cycle, as would happen in a mechanism such as a clock. Is such a system stable? The only things we will assume about the friction process in the system is that

```
Relations:
        e-loss(sys) ∝Q+ total-energy(system)
        correspondence((e-loss(sys), ZERO), (total-energy(system), ZERO))

Influences:
        I-(total-energy(system), e-loss(sys))
```

where e-loss(sys) is the net energy lost due to friction over a cycle of the oscillator process. The loss being qualitatively proportional to the energy is based on the fact that the energy lost by friction is proportional to the distance traveled, which in turn is proportional to the maximum displacement, which itself is qualitatively proportional to the energy of the system, as stated above.

The lower bound for the energy of the system is ZERO, and an upper bound for energy is implicit

---

1. The Tacoma Narrows bridge phenomena, for example, something every engineer should know about.

in the possibility of the parts breaking. The result, via the $\propto_Q$ statement above, is a set of limits on the Quantity Space for `e-loss(sys)`. If we assume `e-pump(sys)`, the energy that is added to the system over a cycle, is within this boundary then there will be a value for `total-energy(system)`, call it `e-stable(sys)`, such that:

```
∀ t ∈ intervals
(M A[total-energy(system)] t) = (M A[e-stable(sys)] t))
        ⇒ (M A[e-loss(sys)] t) = (M A[e-pump(sys)] t)
```

Note that `e-stable(sys)` is unique because $\propto_Q$ is monotonic. If the energy of the system is at this point, the influences of friction and pumping will cancel and the system will stay at this energy. Suppose

```
(M A[total-energy(system)] t) > (M A[e-stable(sys)] t)
```

over some cycle. Then because the loss is qualitatively proportional to the energy, the energy loss will be greater than the energy gained by pumping, i.e., D_s`[total-energy(system)]=-1`, and the energy will drop until it reaches `e-stable(sys)`. Similarly, if `total-energy(system)` is less than `e-stable(sys)` the influence of friction on the energy will be less than that of the pumping, thus D_s`[total-energy(system)]=1`. This will continue until the energy of the system is again equal to `e-stable(sys)`. Therefore for any particular pumping energy there will be a stable oscillation point. This is a qualitative version of the proof of the existence and stability of limit cycles in the solution of differential equations.

5. Further Entailments

The constructs of Qualitative Process theory provide a representational framework for a certain class of deductions about the physical world. In this section we examine the consequences of this framework for several "higher-level" issues in common sense physical reasoning. Several of these issues arise in reasoning about designed systems, while others cover more general topics.

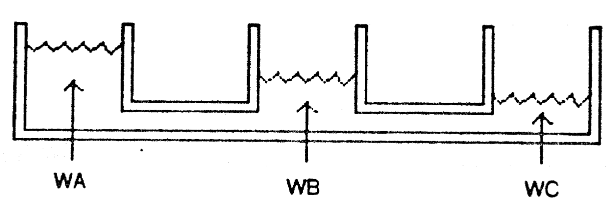
5.1 Distinguishing Oscillation from Stutter

Envisioning - generating all the possible behaviors of a system - can be performed by repeated Limit Analysis. The result will be a linked graph of Process Structures, which can be transversed to form any of the possible histories for the objects that comprise the system. In walking this graph we may find a terminal state (either because we do not know how to evolve a history past a certain kind of event or because we simply haven't bothered) or we might find a loop. A loop must be summarized if we are to get a finite description of the system's behavior. There are several ways to produce such summaries. In some systems the major regularity is spatial, in which case we would produce descriptions like "is bouncing around inside the well"[Forbus, 1981a]. Another concise summary is when a system is oscillating, since there is a pattern of activity that occurs over and over again.

While oscillation in the physical system results in loops in the envisionment, there are other circumstances that give rise to loops as well. In part this is due to the qualitative nature of the descriptions used. Consider the situation illustrated in Figure 42. Initially there will be two flows, one from A to B and the other from B to C. What can happen? Limit Analysis reveals three alternatives, corresponding to each of the flows stopping individually and to both ending simultaneously (see Figure

---

Fig. 42. Three Container Example
Suppose we have three containers partially filled with water and connected by pipes, as shown below. If we assume the water moves slowly, what can happen?

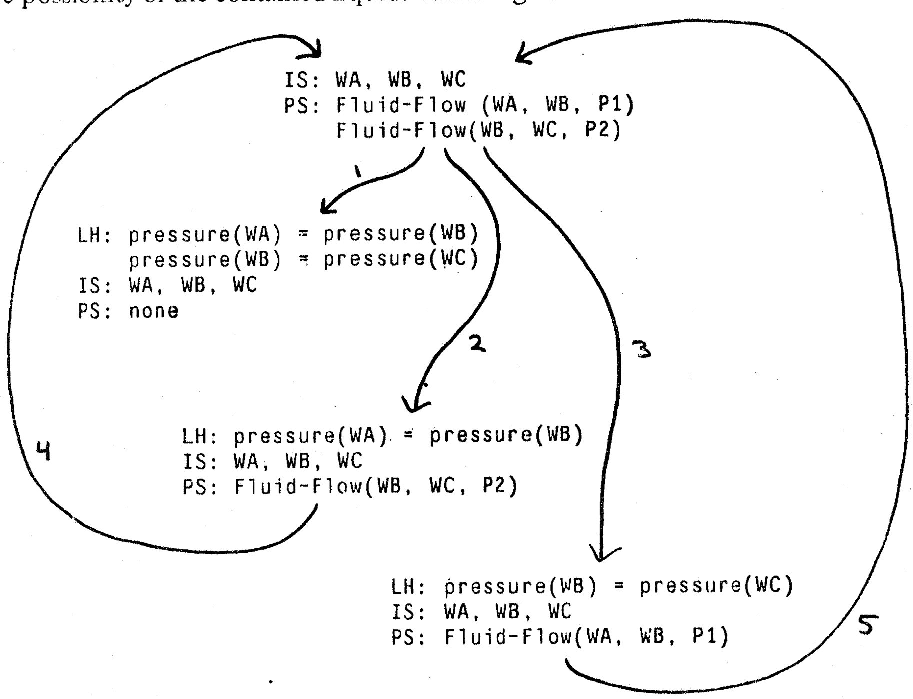


43). In the cases where one flow stops before the other, the flow that continues will decrease the amount, and hence pressure, so the flow will start again. This is only an oscillation in some degenerate sense. A more natural description is that the change in level "follows" the other change. In other words, we have a decaying equilibrium.[1] We will call the behavior represented by the degenerate cycles *stutter*. How can we distinguish stutter from true oscillation?

Physically an oscillation requires that the system have some form of inertia or hysteresis. Thus whenever the thing causing the change stops acting, the change will continue for a while afterwards.[2] A real oscillation will therefore consist of process episodes that last over an interval, whereas stutter will

---

Fig. 43. Stutter in Fluid Flow
This graph of transitions between Process Structures formed by repeated Limit Analysis contains two cycles, neither of which correspond to true oscillations in the physical world. For simplicity, we ignore the possibility of the contained liquids vanishing as a result of the flows.



---

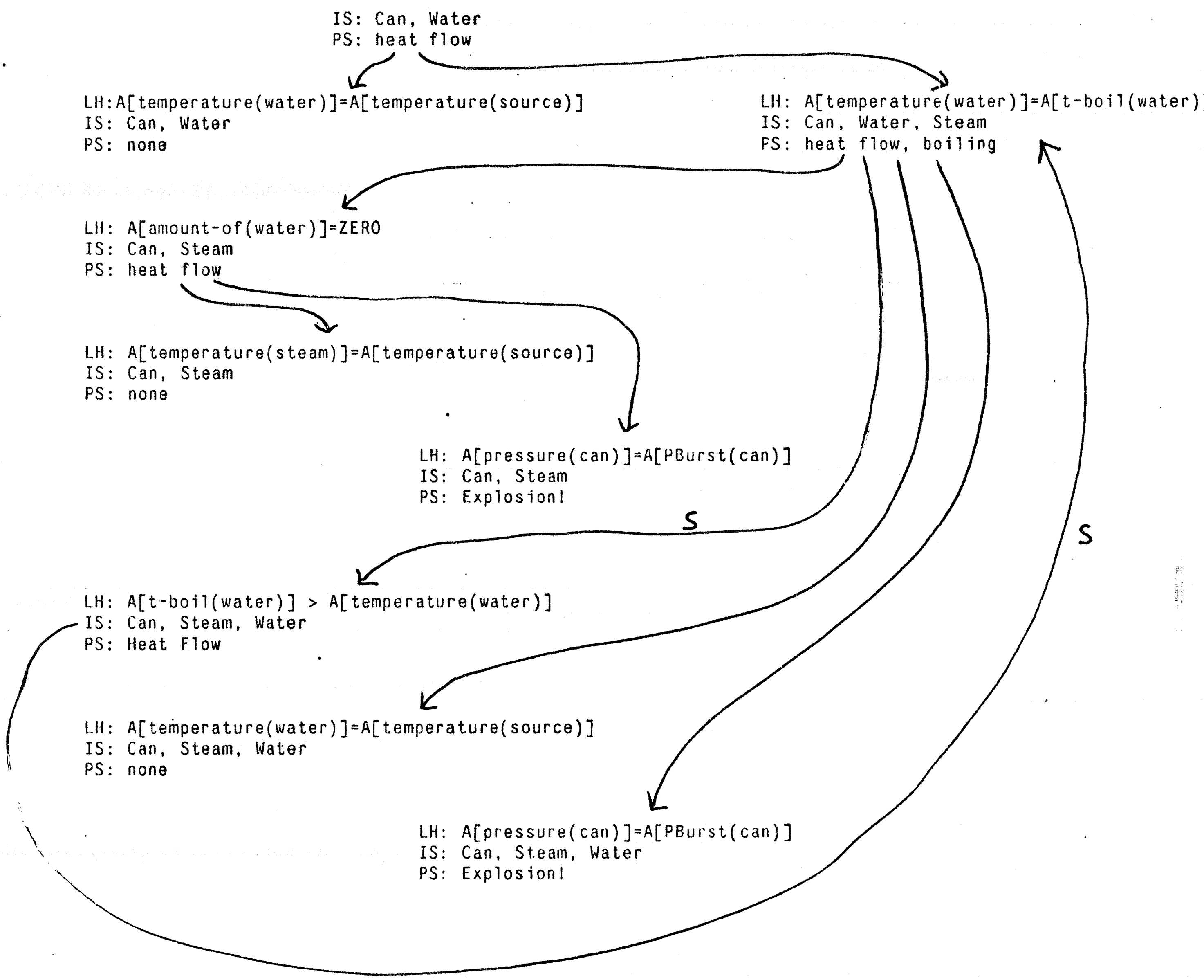
1. Brian Williams (personal communication) has noted the same phenomena in reasoning about VLSI circuits, particularly in charging up capacitors.
2. de Kleer and Brown have noted that some concept like momentum is needed to build reasonable qualitative models, but do not explain how to detect when something like momentum is operating [de Kleer & Brown, 1983].

result in process episodes that last but an instant.

Fortunately the Equality Change Law provides a way of distinguishing these cases. In the previous transition diagram, for example, transitions 4 and 5 take place in an instant. Therefore we have two cases of stutter, corresponding to the two fluids participating in a decreasing equilibrium.

A similar phenomena occurred in the Boiler model presented earlier(Section 5.2). Figure 44 illustrates the full transition diagram. Note that if `t-boil(water)` rises faster than `temperature(water)`

Fig. 44. Stutter in the Boiler Model
The temperature and pressure will be continuously increasing in the boiler, but unless the changes in the links marked "s" are recognized as something special, the system will appear to be oscillating.



the boiling will stop. Since this is a change from equality, by the Equality Change law it will occur in an instant. This in turn means that `t-boil(water)` was only influenced for an instant. When the boiling stops only the heat flow is acting, so `temperature(water)` will rise, and by the Equality Change law the change to equality will occur in an instant. Therefore this cycle is an instance of stutter as well, corresponding to a rising equilibrium.

Being able to distinguish stutter from oscillation means we can write rules that summarize the Process History. For example, when stutter occurs we can note the D_s values for the quantities involved and assert that one kind of change is "following" another, a decaying or rising equilibrium. Informal observations suggest that novices in a domain often confuse stutter and oscillation, and even experts who describe the situation as decaying or rising equilibrium are able to reconstruct the view of stutter as an oscillation. These informal observations need to be examined in the light of empirical data, but if true it may be useful in testing Qualitative Process theory as a psychological model.

5.2 Causal Reasoning

We use causality to impose order upon the world. When we think that "A causes B", we believe that if we want B to happen we should bring about A, and that if we see B happening then A might be the reason for it. Causal reasoning is especially important for understanding physical systems, as de Kleer has noted[de Kleer, 1977]. Exactly what underlies our notion of causation in physical systems is still something of a mystery.

Consider the representations used in physics. Typically equations are used to express constraints that hold between physical parameters. A salient feature of equations is that they can be used in several different ways. For example, if we know "X = A + B", then if we have A and B we can compute X, but also if we have X and A we can compute B. It has been noted that in causal reasoning people do <u>not</u> use equations in all possible ways [diSessa, 1983][Reily, 1981]. Only certain directions of information flow intuitively correspond to causal changes. I propose the following *Causality is Functional Hypothesis:*

> *Changes in physical situations that are perceived as causal are due either to direct changes caused by processes or propagation of those direct effects through functional dependencies.*

This section will attempt to justify that hypothesis.

First, note that causality requires some notion of mechanism. Imagine an abstract rectangle of a particular length and width. If we imagine a rectangle that is longer, it will have greater area. There is no sense of causality in the change from one to the other. If however we imagine the rectangle to be made of some elastic material and we bring about the increased length by stretching it, then we are comfortable with saying "the increased length causes the area to increase". Qualitative Process theory asserts that processes are the mechanism that directly cause changes. The quantities that can be directly influenced by processes are in some sense independent parameters, because they are what can be directly affected. All other quantities are dependent, in the sense that to affect them some independent parameter or set of independent parameters must be changed. This suggests representing the relationships between parameters in terms of functions rather than constraint relations.

Some examples will make this clearer, as well as emphasizing that the point is not academic. In generating explanations of physical systems, it is often useful to characterize how the system responds to some kind of change (this variety of qualitative perturbation analysis was invented by de Kleer, who calls it *Incremental Qualitative Analysis*). One way to perform IQ analysis is to model the system by a constraint network, in which the relationships are modeled by "devices" that contain local rules that enforce the desired semantics.[1] The values of quantities are modeled by the sign of their change - increasing, decreasing, or constant. An analysis is performed by placing a value in a cell of the network and using the rules to propagate the effects of that value. The dependency relationship between cells is interpreted as causality. Figure 45 illustrates fragments from two different models.[2] The top fragment states that heat is the product of the temperature of the "stuff" and the amount of the "stuff", and the bottom fragment is the definition of sodium concentration in a solution.

In building a causal argument an impasse can be reached - a quantity receives a value, but no further values can be computed unless an assumption is made. The safest assumption is that, unless you know otherwise, a quantity doesn't change. The problem is determining which quantity to make the

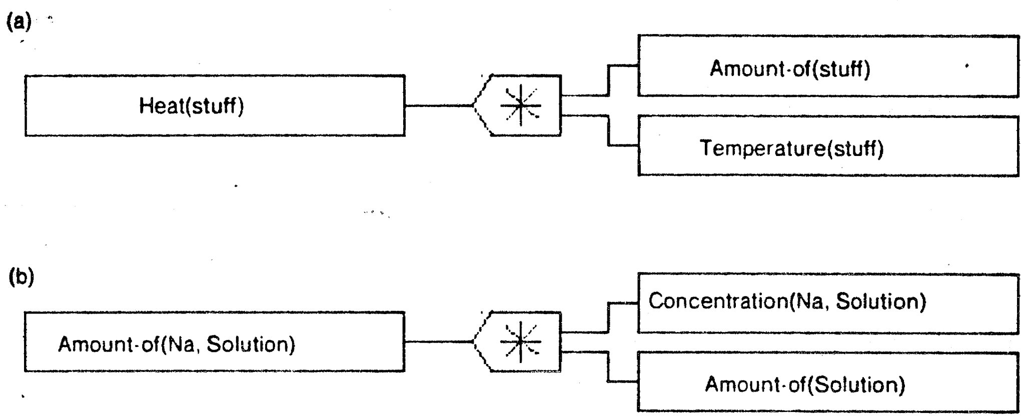
---

1. These examples are drawn from systems implemented in CONLAN [Forbus, 1980], a constraint language. The notation is similar to that of logic diagrams, except that the terminals are given explicit names and the devices are multi-functional. The technique described here is a simplification of de Kleer's algorithms, which are more subtle. de Kleer's models use directional rules rather than constraint laws, although his theory does not provide any criteria for selecting which direction in a constraint law is appropriate.

2. The problem was observed in implementing the model of a student's understanding of a heat exchanger described in [Williams, et. al., 1983], in my own work on understanding Automatic Boiler Control systems, and in an early version of the kidney model described in [Asbell, 1982].

Fig. 45. Constraint Representation of Relationships
(a) is drawn from the model for a piece of "stuff" used in an effort to represent a student's understanding of heat exchangers.
(b) is drawn from an IQ model of a kidney to be used in explaining the syndrome of inappropriate secretion of anti-diarretic hormone (SIADH).

```
Correct Causal argument: "The increasing heat causes the temperature to rise"
Incorrect Causal argument: "The increasing heat causes the amount of fluid to rise"
```



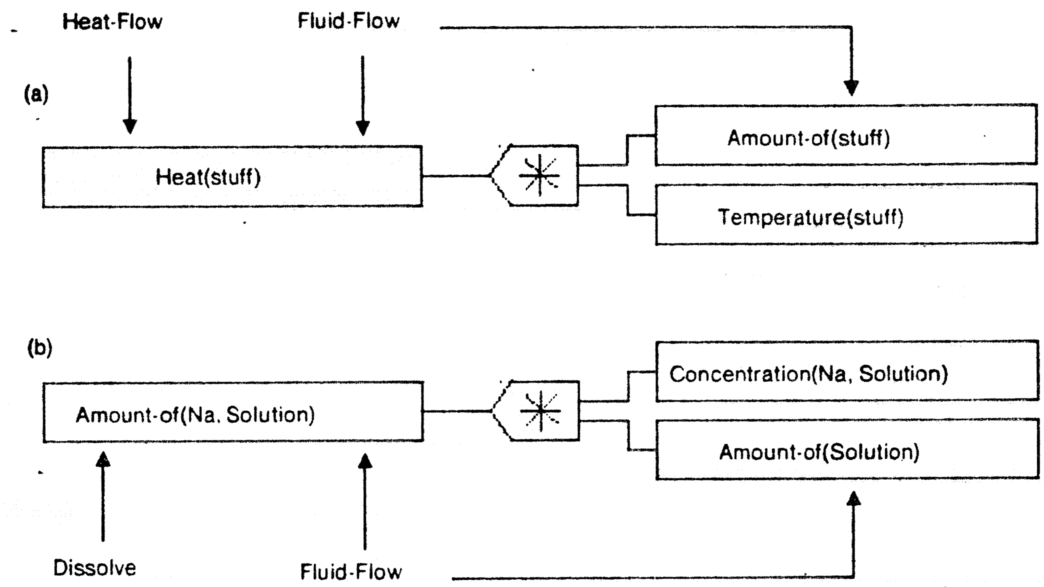
---

assumption about. If we assume that the amount of stuff is constant, for example, we get an increase in temperature, which makes sense. However, if we assume that the temperature remains constant, we are left with the conclusion that a rise in heat <u>causes</u> the amount of stuff to decrease! Barring state changes, this does not correspond to our ideas of what can cause what. In the second fragment the problem is more serious - increasing sodium will cause the amount of water to increase, <u>if</u> the rest of the kidney is working as it should! To do this requires a complicated feedback mechanism that is triggered by detecting an increased sodium concentration, not by the definition of concentration itself!

The problem lies in the ontological impoverishment of the constraint representation. Temperature and concentration are not directly influenced by processes (at least in most people's Naive Physics) - physically they are *dependent* variables, and thus are not proper subjects of assumptions. Amount of stuff, on the other hand, can be directly affected, so assuming it does not change will avoid generating ill-formed causal arguments. Figure 46 illustrates.

Of course, the proper assumptions to make concern what processes are active and how indirect influences are resolved. If we do not represent processes, we can only assume facts about quantities. If

Fig. 46. Model Fragments with Possible Proceses
Here are the models from the previous figure with the quantities annotated with the (likely) processes that might affect them. Note that certain quantities (temperature, concentration) cannot be directly changed. These are dependent quantities, and should not be the subject of assumptions in building causal arguments.



we assume a quantity is constant and later discover that assumption is wrong, we are left in the dark about why that assumption was wrong. For example, if the amount of stuff turns out not to be constant, we can look for fluid flows or state changes to explain why it isn't. Since processes tend to have more than one effect, there is some chance that the contradiction can lead to discovering more about the system rather than just being a bad guess.

5.3 Qualitative Proportionalities Revisited

The previous section proposed that functional dependence is central to the kind of "incremental" causality that people find useful in reasoning about the physical world. As discussed previously, one goal of Naive Physics should be to develop a theory of observation. One use of observation is to interpret measurements in terms of theories ([Forbus, 1983b]), but another role for observation is in developing physical theories. While this problem has been studied before (c.f. [Langley, 1979]), the target representations have been equations. As a result the learning procedure has relied on numerical data and cannot build theories around weaker information. Learning constraint laws also differs from learning causal connections. As noted above, an equation carries only part of what we know about a domain. Constructing a learning theory for physical domains will require ways to learn process descriptions and causal connections.

One way to learn about a system is to "poke" it and see what it does. The observed behavior can be used to make conjectures about causal connections between the parts of the system, and further experiments can be made to test the conjectures. This requires some notation to express the local causal connections conjectured on the basis of these simple observations. This requirement helped motivate the definition of $\propto_Q$ (see Section 2), which asserts that a functional dependence exists between two quantities. If whenever parameter A in a system is poked parameter B changes, the result can be expressed as `(B` $\propto_Q$ `A)`. A physical explanation for the dependence comes from writing the $\propto_Q$ within the scope of an Individual View or Process.

More powerful statements about a system will require extensions of $\propto_Q$. To see what is involved, consider the analogous situation of learning how an old-fashioned typewriter works.[1] If the space bar is pushed, the carriage will move to the left. This is analogous to the kind of statement that can be made with $\propto_Q$. But lots of other things can happen to move the carriage, namely all of the letter keys and a few more. Thus it would be useful to be able to state that we know <u>all</u> of the influences (at least, within the current grasp of the situation) on some particular parameter. Suppose also that we just wanted to move the paper up without changing anything else. The return bar would move the paper up, but before doing so would return the carriage to the right. Being able to say there are no (known) intervening parameters is then also a useful ability.

To see how these notions can be expressed, consider the collection of $\propto_Q$ relations that hold at some instant in time. For any quantity, the $\propto_Q$ statements relevant to it can be thought of as a tree with the dependent quantity at the root and the "independent" quantities at the leaves.[2] A plus or minus denotes the sense of the connection (whether or not it will reverse the sign of the change in the input). $(Q_1 \propto_Q Q_0)$, then, only specifies that Q_1 is on some branch "above" Q_0.

---

1. This is not proposed as a serious example because the quantity definitions and $\propto_Q$ would apply only in some very abstract sense.
2. Actually a directed graph with cycles can be formed, as for instance in a control system.

Figure 47 illustrates such a dependency tree. Suppose we are trying to cause Q_0 to change. If we don't want to change Q_2, then Q_3 or Q_1 are our only choices. We need a way to express that (at least within our knowledge of the situation) there are no intervening parameters. To say this, we use

$\propto_Q$-direct(Q_0, Q_1)

which can be modified by + or - as before. $\propto_Q$-direct adds a single link to the tree of dependencies. Another problem is to find all the ways to bring a change about, or to prove that changing one thing won't cause a change in some other quantity of interest. We do this by stating that a particular collection of quantities together "closes off" the tree - there will be exactly one quantity for each branch. Our notation will be

$\propto_Q$-all(<quantity>, <plus-set>, <minus-set>)

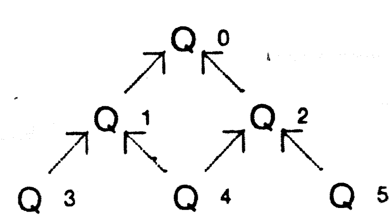
which means that there is a function induced by a process that determines the quantity, and that relies on the quantities in the two sets solely. If a quantity is not mentioned in a $\propto_Q$-all statement, then either it is irrelevant to the quantity of interest, it depends on some quantity in the $\propto_Q$-all statement (above the slice of the tree it induces), or some quantity in the $\propto_Q$-all statement depends on it. By ruling out the other two possibilities, independence can be established.

As a rule $\propto_Q$ statements will not hold for all time. In the typewriter analogy, imagine the carriage at the end of its travel - Hitting the space bar will no longer result in movement. More to the point, consider Q_0 given by:

$$Q_0 = (a - b*Q_2)*Q_1$$

Note that:

Fig. 47. A Tree of Functional Dependencies



```
if a > b*Q2, Q0 ∝Q+ Q1
   a = b*Q2, ¬ Q0 ∝Q Q1
   a < b*Q2, Q0 ∝Q- Q1
```

In the case of equality, Q_0 and Q_1 are not related at all, and in the other two cases the sign of the function connecting them is different. Thus the collection of $\propto_Q$ statements that are true for a system can vary as a function of the values of the quantities, which is why they usually appear within some Individual View or process. The collection of $\propto_Q$ statements that holds for some class of situations will define a mode of the system being described. Multi-mode systems include four stroke engines and automobile transmissions.

5.4 Differential Qualitative Analysis

The idea of a comparison in Incremental Qualitative (IQ) analysis [de Kleer, 1979] suggests a complementary qualitative reasoning technique. IQ analysis concerns the relationship between two situations, one of which is the direct result of things happening in the other. Another case of interest concerns situations that are just slightly different from one another. For instance, we often have an idea of the different consequences that would result if something were changing a bit faster - if we put the heat up on the stove the water in the kettle would boil sooner, and if our arm were quicker the serve would have been returned. Such inferences are essential in debugging and monitoring execution of plans that involve physical action. The language needed to express such conclusions is in part the same as that used in IQ analysis - amounts are either the same, increased, decreased, or indeterminate as compared with the old situation. Answering these kinds of questions will be called *Differential Qualitative* analysis.

Let us consider a situation A. If we get a new situation B by changing some ordering in A or by changing a single process in A, we will call B an alternative to A. There are two kinds of changes that may occur as a result of perturbing A. The first kind are changes in quantities, as noted above. Secondly, the process history for the situation itself may change, apart from any changes made to define B in the first place. An example would be punching a hole in the bottom of a kettle, which could let all the water drain out before a boiling occurs. Even changes in orderings can lead to historical consequences - if we reduce the intensity of a flame while still agreeing that it will be turned off in five minutes, boiling may again be prevented.

Let DQ(q, A, B) for some quantity q be the sign of the difference between two situations A and B that are alternatives. Then the inequality order between them defines DQ values, as follows:

```
(M q A) > (M q B)  DQ(q, A, B) = 1
(M q A) < (M q B)  DQ(q, A, B) = -1
```

$(M\ q\ A) = (M\ q\ B)\quad DQ(q, A, B) = 0$

The inequality orderings for instants must of course be extended to apply over intervals. For equality this is simple:

$$\forall\ q_1, q_2 \in \underline{quantity}\ \forall\ i \in \underline{interval}$$
$$(M\ q_1\ i) = (M\ q_2\ i) \equiv \forall\ i_1 \in during(i)\ (M\ q_1\ i_1) = (M\ q_2\ i_1)$$

For the other cases the choice is less clear. The strongest version of greater-than is having it hold over every instant in the interval:

$$\forall q_1, q_2 \in quantities,\ i \in intervals$$
$$(M\ q_1\ i) > (M\ q_2\ i) \equiv$$
$$(\forall\ i_1 \in during(i)\ (M\ q_1\ i_1) > (M\ q_2\ i_1))$$

but for extending our notion of integrability, the following will also suffice:

$$\forall q_1, q_2 \in quantities,\ i \in intervals$$
$$(M\ q_1\ i) > (M\ q_2\ i)$$
$$\wedge\ (\exists\ i_1 \in during(i)\ (M\ q_1\ i_1) > (M\ q_2\ i_1))$$
$$\wedge\ (\forall\ i_1 \in during(i)\ \neg\ (M\ q_1\ i_1) < (M\ q_2\ i_1))$$

A version of < for intervals may be similarly defined.

An episode in a parameter history has several numbers associated with it. The relationships between these numbers allows new DQ values to be determined. The first number is rate, e.g., the D_m of the quantity the parameter history is about. The second number is the duration of the interval associated with the episode. The third number is the difference in the value measured at the beginning and end of the interval, which we will call the distance.

How are these numbers related? Intuitively we know that if the rate increases or decreases, the duration of time will decrease or increase, or the distance the value moves will increase or decrease for the same duration. Implicit in this simple intuition is the restriction that the rate is constant during the interval, i.e., that the function defining the change of the quantity is linear and time invariant. This often is not the case, so we must require that either the beginning or the end of the two episodes being compared are the same. If we apply DQ analysis only to alternative situations this restriction will be satisfied.

With these assumptions, the DQ value of the distance is just the product of the DQ values of the rate and duration. Using the algebra of signs introduced previously, we can draw conclusions such as "If the rate is higher then over the same time the distance traveled will be greater." and "If the duration is shorter and the rate is the same then the distance traveled will be less." These inferences are illustrated in figure 48.

The direct historical consequences of these changes can be characterized by their effects on

Fig. 48. Differential Qualitative Analysis
Differential Qualitative analysis uses the algebra of signs from Incremental Qualitative analysis to answer questions about how a situation would change if parts of it are perturbed.

```
Suppose we have alternate situations A and B, with a
quantity Q in both of them.

∀ S ∈ episode Distance(Q) = (M Q end(S)) - (M Q start(S))

DQ[distance(Q, S), A, B] = DQ[rate(Q), A, B] * DQ[duration(Q), A, B]

Then assuming time(start(A)) = time(start(B)),

        DQ[rate(Q), A, B] = 1 ∧ DQ[duration(Q), A, B] = 0
                ⇒ DQ[distance(Q), A, B] = 1

        DQ[rate(Q), A, B] = -1 ∧ DQ[distance(Q), A, B] = 0
                ⇒ DQ[duration(Q), A, B] = 1

i.e., "If Q is going faster then it will get farther in the same time" and
"If Q is going slower it will take longer to go the same distance."
```

Limit Analysis. Consider a collection of Limit Hypotheses for a P-component. Recall that each hypothesis concerns a possible change in the Process Structure, brought about by changes in quantities that cause changes in Quantity Conditions. Suppose a particular Limit Hypothesis is chosen as representing what actually occurs. This means the change it stands for happens before the changes represented by the other hypotheses. If in an alternate situation this hypothesis has an increased duration (a DQ value of 1) or one of the other Limit Hypotheses has a decreased duration (a DQ value of -1), then in fact a different change could occur. Once again, the weak nature of our information prevents us from actually deciding if the change would occur - but we at least know that such a change is possible in these circumstances.

6. Discussion

This paper has presented a new theory, Qualitative Process theory, that models aspects of common sense physical reasoning. To summarize:

- Our theories about how things change in the physical world have a common character. Physical processes are the mechanisms by which change occurs. Reasoning about processes, their effects and limits, form an important part of our commonsense physical reasoning.

- A process is specified by the Individuals it occurs among, the Preconditions and Quantity Conditions that must be true for the process to be active, the Relations it imposes among those individuals, and the Influences it imposes on their quantities.

- Reasoning about processes provides constraint on qualitative representations of quantity. The Quantity Space representation, which describes the value of a quantity in terms of ordering information with other quantities, is appropriate since processes usually start and stop when inequalities between particular quantities change.

- Several kinds of qualitative conclusions can be drawn using the constructs of QP theory, including reasoning about the effects of combined processes, the limits of processes, and alternative situations.

- Interesting phenomena in common sense reasoning appear to be described reasonably well by QP theory, including flows, state changes, motion, materials, and oscillation.

- QP theory provides a highly constrained account of physical causality (all changes are due to a finite vocabulary of processes) and a useful notation for expressing causal connections ($\propto_Q$).

- QP theory provides a highly constrained role for the use of experiential and default knowledge in physical reasoning - for example, in resolving influences and choosing or ruling out alternatives in Limit Analysis.

- Processes can provide a language for writing physical dynamics theories. In particular, the primitives are simple processes and individual views, the means of combination are sequentiality and shared parameters, and the means of abstraction are naming these combinations, including encapsulating a piece of the process history.

6.1 Application Areas

While designed to be a theory about Naive Physics, Qualitative Process theory has several other potential applications. Two are discussed below.

6.1.1 Reasoning about Engineered Systems

Many engineered devices are implemented as physical systems, and thus are subject to physical laws. A qualitative understanding of such systems involves our common sense physical knowledge. I have been applying Qualitative Process theory to reasoning about the physics of steam plants as part of the STEAMER project[Stevens, et. al., 1981].[1] The qualitative nature of its descriptions appear similar to those often used in understanding and explaining complex physical systems, suggesting QP theory would be useful as a representation language for ICAI systems. In addition, the notions of quantity and functional dependence have been useful in thinking about more abstract functional descriptions (such as COMPARATOR and FEEDBACK-LOOP), because signals in a large class of engineered systems are continuous.

Applications other than teaching are imaginable. If extension theories were provided to interface the basic QP theory descriptions with quantitative descriptions of what is actually happening in a system, several new possibilities arise. Controlling systems should ultimately be possible, using Individual Views to express desired and undesired operational characteristics. More immediately feasible would be an interpretation module, that would gather data from instruments and hypothesize what processes generated that data (e.g., [Simmons, 1982][Forbus, 1983b]). Such a module could be used as part of a diagnosis program or as a "hypothesizer" that could serve as a devil's advocate during the operation of a complex system. For example, the incident at the Three Mile Island reactor probably wouldn't have happened if the operators had thought of the alternate explanation for the overpressure in the reactor vessel - that instead of being too high, the level of cooling water was too low, thus causing a boiling that raised the pressure.[2]

---

1. STEAMER is a joint project of NPRDC and BBN, to develop intelligent computer aided instruction techniques to train propulsion plant officers for the Navy.
2. [Pew et.al., 1982] hypothesizes premature commitment by operators to a particular theory about the state of the plant as a common source of human errors in power plant operation.

6.1.2 Economic Modelling and Support Systems

Many non-physical systems are often modeled with continuous parameters and processes, notably economies. A theory of physical reasoning might provide useful leverage in understanding such systems in several ways. First, physical limitations often constrain such systems (storage capacities, transportation capacities, time required for physical processes involved in crop growth or manufacture, etc.). Second, economic systems are often described by analogy with physical systems (Samuelson, for instance, cites the aphorism "the central bank can pull on a string (to curb booms), but it can't push on a string (to reverse deep slumps)"[Samuelson, 1973]). Third, non-physical processes themselves might be usefully described using a theory like the present one.

Several caveats are in order. First, unlike physical systems, there is no real agreement on what are valid process descriptions in domains like economics. Second, changes in circumstances may dictate changing Process Vocabularies (tax laws can change, for instance). This means that the set of possible influences is essentially unbounded. These application areas are therefore much harder than physical reasoning.

6.2 Other Work

The first attempt to formalize processes was [Hendrix, 1973]. While a significant advance over the models of action available at the time, the importance of qualitative descriptions had not yet been discovered. For example, the values of numbers were known real numbers, and relationships between parameters were expressed as constraint equations. The process descriptions were used in simulation, solving simultaneous equations in the process descriptions to determine when the collection of active processes would change. Since the goal was to model general processes (non-physical as well as physical), add lists and delete lists were also used to specify effects. Qualitative Process theory, by using qualitative descriptions and focusing on physical processes only, can be used in several other kinds of deductions in addition to simulation, often with less information.

Recently several attempts have been made to model temporal reasoning, including [Allen, 1982], [McDermott, 1981]. Allen's model is the one assumed here, mainly because modeling instants as "very short" intervals makes formalizing certain facts about derivatives easier. McDermott's axioms for time contain several interesting ideas, including the chronicle representation of possible futures and its links with planning. Unfortunately, too much is expected of the temporal logic. For example, the notion of a "lifetime" (how long a fact is true, provided by fiat - "The senses actually tell you about persistences.") is needed precisely because the logic is developed independently from a theory of dynamics. Given a dynamics (and closed world assumptions about individuals and relationships), we can deduce what will and will not change. If an estimate of how long something will remain true is needed, we can compute such an estimate by figuring out how long it is likely to be before something that can change it occurs. To use McDermott's example, if you look at a boulder you might be able to estimate that if you came back in 50 years it would still be there (a weaker conclusion than implied by the notion of lifetimes, but it will do). However, if I tell you that there is dynamite underneath, your estimate will be considerably different. In either case, if you came back the next day and discovered the boulder was some distance from its original location, you would have some theory about why, not just the feeling that your senses had lied to you. Similarly, using "average rate" instead of derivatives means many of the dynamical conclusions described here (such as distinguishing oscillation from stutter) cannot be drawn.

6.3 Current Directions

Since the original publication of Qualitative Process theory, several projects have adopted or extended some of its ideas. In particular,

> Ben Kuipers has analyzed protocols of causal reasoning in medicine, including an implementation of rules to reason about changes within a Process Structure[Kuipers, 1982].
>
> Reid Simmons has developed process representations for geologic map interpretation by qualitative simulation, including the use of a diagram. He also has developed an extension theory for quantities, describing them in terms of intervals[Simmons, 1982].
>
> Brian Williams is studying qualitative time domain analysis of VLSI circuits (personal communication).
>
> Al Stevens, Kathy Larkin, and Albert Boulanger are using Qualitative Process theory in constructing a theory of explanations for machines.
>
> Alan Collins and Dedre Gentner are using Qualitative Process theory to express theories of evaporation in order to understand how to shift from one level of description to another. Also, we are using QP theory in developing a psychological theory of learning for physical domains.

6.4 Acknowledgments

Mike Brady, John Seely Brown, Alan Collins, Johan de Kleer, Dedre Gentner, Pat Hayes, David McAllester, Bruce Roberts, Al Stevens, Gerald Sussman, Dan Weld, and Patrick Winston have all influenced the development of this theory in various ways. Portions of this work were supported by NPRDC under the STEAMER project.

7. Bibliography

Allen, J. "A General Model of Action and Time" University of Rochester Computer Science Department TR-97, November, 1981

Asbell, Irwin "A Constraint Representation and Explanation Facility for Renal Physiology", MIT SM Thesis, August 1982

Collins, A., Warnock, E. Aiello, N. and Miller, M. "Reasoning from Incomplete Knowledge", in Representation and Understanding, D. Bobrow and A. Collins, editors. New York, Academic Press, inc.

de Kleer, J. "Qualitative and Quantitative Knowledge in Classical Mechanics" TR-352, MIT AI Lab, Cambridge, Massachusetts, 1975

de Kleer. J. "Causal and Teleological Reasoning in Circuit Recognition" TR-529, MIT AI Lab, Cambridge, Massachusetts, September 1979

de Kleer, J. and Brown, J. "Assumptions and Ambiguities in Mechanistic Mental Models" in Mental Models, D. Gentner and A. Stevens, LEA, Inc, April 1983

de Kleer, J. and Sussman, G. "Propagation of Constraints Applied to Circuit Synthesis" MIT AI Lab Memo No. 485, Cambridge, Massachusetts, 1978

diSessa, A. "Unlearning Aristotelian Physics: A Study of Knowledge-Based Learning", Cognitive Science, Vol. 6, No. 1, Jan-March 1982

diSessa, A. "Phenomenology and the Evolution of Intuition" in Mental Models, D. Gentner and A. Stevens, LEA, Inc, April 1983

Forbus, K. "A Conlan Primer" BBN Technical Report No. 4491, prepared for Navy Personnel Research and Development Center, 1980.

Forbus, K. "Using Qualitative Simulation to Generate Explanations" BBN Technical Report No. 4490, prepared for Navy Personnel Research and Development Center, 1980.

Forbus, K. "A Study of Qualitative and Geometric Knowledge in Reasoning about Motion" TR-615, MIT AI Lab, Cambridge, Massachusetts, February, 1981

Forbus, K. "Qualitative Reasoning about Physical Processes" Proceedings of IJCAI-7, 1981

Forbus, K. "Qualitative Reasoning about Space and Motion" in Mental Models, D. Gentner and A. Stevens, LEA, Inc, April 1983

Forbus, K. "Measurement Interpretation in Qualitative Process theory" Proceedings of IJCAI-8, 1983

Hayes, Patrick J. "The Naive Physics Manifesto" in Expert Systems in the Micro-Electronic Age, edited by D. Michie, Edinburgh University press, May 1979

Hayes, Patrick J. "Naive Physics 1 - Ontology for Liquids" Memo, Centre pour les etudes Semantiques et Cognitives, Geneva, 1979

Hendrix, G. "Modeling Simultaneous Actions and Continuous Processes" Artificial Intelligence, Volume 4, 1973, pp145-180

Kuipers, B. "Getting the Envisionment Right", Proceedings of the National Conference on Artificial Intelligence, 1982

Langley, P. "Rediscovering Physics with BACON.3" Proceedings of IJCAI-6, 1979

Mcclosky, M. "Naive Theories of Motion" in Mental Models, D. Gentner and A. Stevens, LEA, Inc, April 1983

McCarthy, J. and Hayes, P. "Some Philosophical Problems from the Standpoint of Artificial Intelligence", Machine Intelligence 4, Edinburgh University Press

McDermott, D. "A Temporal Logic for Reasoning about Processes and Plans" Cognitive Science, Vol. 6, No. 2, April-June, 1982

McDermott, D. and Doyle, J "Non-Monotonic Logic I", Artificial Intelligence, Volume 13, Number 1, April 1980

Minsky, M. "A Framework for Representing Knowledge" MIT AI Lab Memo No. 306, June 1974

Moore, R. "Reasoning from Incomplete Knowledge in a Procedural Deduction System" MIT AI-TR-347, Cambridge, Massachusetts, December 1975

Moore, R. "Reasoning about Knowledge and Action" MIT PhD thesis, February, 1979.

Pew, R., Miller, D. and Feeher, C. "Evaluation of Proposed Control Room Improvements Through Analysis of Critical Operator Decisions" Electric Power Research Institute Report NP-1982, August, 1982

Reiter, R. "A logic for Default Reasoning" Artificial Intelligence, Vol. 13, 1980, pp 81-132.

Riley, M. Bee, N. and Mokwa, J. "Representations in Early Learning: The acquisition of Problem Solving Strategies in Basic Electricity/Electronics" University of Pittsburgh Learning Research and Development Center, 1981

Samuelson, P. "Economics", McGraw-Hill, N.Y., 1973

Simmons, R. "Spatial and Temporal Reasoning in Geologic Map Interpretation" Proceedings of the National Conference on Artificial Intelligence, 1982

Stallman, R. and Sussman, G. "Forward Reasoning and Dependency-Directed Backtracking in a System for Computer-Aided Circuit Analysis" Artificial Intelligence, Volume 9, pp 135-196, 1977

Stanfill, C. "Partitioned Constraint Networks: A Prescription for Representing Machines", University of Maryland, Computer Science Technical report No. 1037, 1981

Stansfield, J. "Conclusions from the Commodity Expert Project" MIT AI Lab Memo 601, November, 1980

Stevens, A. et. al. "Steamer:Advanced Computer Aided Instruction in Propulsion Engineering" BBN

Technical Report No. 4702, July 1981

Williams, M., Hollan, J. and Stevens, A. "Human Reasoning about a Simple Physical System", in Mental Models, D. Gentner and A. Stevens, LEA, Inc, April 1983