MASSACHUSETTS INSTITUTE OF TECHNOLOGY
ARTIFICIAL INTELLIGENCE LABORATORY

Working Paper 243 January 1983

Representing Change
for Common-Sense Physical Reasoning

Richard J. Doyle

Abstract: Change pervades every moment of our lives. Much of our success in dealing with a constantly changing world is based in common-sense physical reasoning about processes and physical systems. Processes are the way quantities interact over time. Physical systems can be described as a set of quantities and the processes that operate on them. Representations for causality, time, and quantity are needed to fully characterize change in this domain. Several ideas for these representations are examined and synthesized in this paper towards the goal of constructing a framework to support understanding of, reasoning about, and learning how things work.

CONTENTS

FIGURES

TABLES

1. INTRODUCTION

The world is a dynamic place. A rock thrown into the air rises to some height, falls back to the ground and smashes a window. Inflation drops, and unemployment rises. People are born, they mature, grow old, and die. The seasons return every year, in the same order. Turn on a faucet, and a sink fills with water. Pull the stopper and the water flows out. To reference a cliché, the one constant in all of our lives is change.

But of course, change is hardly a constant. The ways in which things change are as rich and varied as the different ways things are because of change. We have many ways of describing change. Sometimes a change is sudden, discrete, isolated (that rock breaking a window) and it is characterized as an *event*. Other times a change is gradual (a child becomes an adult) and it is more useful to specify an *interval* over which the change occurs. A change or a set of changes may repeat again and again, and become predictable, specifying a *cycle*.

But change does not mean chaos. Change is part of the way things naturally are. The notion of *quantity* captures how a change can occur while coherence is maintained at some level. A quantity represents a host of values subsumed under a single concept. Examples are a person's age, the height of water in a sink, a rock's position as it flies through the air, the percentage of the work force out of jobs, the time of the year.

The idea of quantity is very familiar but common-sense reasoning about change treats quantities qualitatively rather than quantitatively. In other words, values for a quantity do not have to specified explicitly in order for an understanding to be obtained of how a quantity can change. Common-sense notions such as increase, decrease, and equilibrium do not need access to an explicit measuring scale to have meaning. More, less, and the same are very qualitative terms.

Causality, or in the case of quantities, *functional dependencies*, describe how quantities can interact. There is a functional dependency between unemployment and inflation. When one drops, the other rises, and vice versa. Understanding causal relations allows predictions to be made about the consequences of some change.

The myriad ways in which things change over time - growing, breaking, peaking, oscillating - are collectively characterized as *processes*. A set of quantities, their interactions, and the processes which operate on them comprise a *physical system*. Examples of physical systems are a sink, an automobile, and (alas a poorly understood one!) the economy.

Underlying all processes, all change, is *time*. In the above paragraphs, many common-sense notions about processes have been introduced. The goal of this paper is to investigate how these notions should be incorporated into representations for causality, time, and quantity in the context of describing, understanding, and figuring out how things - physical systems - work.

2. PLANNING: An Aside

There is a major area of AI research which always has a need for at least an implicit representation of time, and that area is planning. (When time is represented explicitly and quantitatively, it is called scheduling). This paper will not examine what is needed in a good representation of time for planning in depth, but will only briefly mention the issues here.

Planning is specifying a sequence of actions to obtain some goal [FiN71]. Typically this is done by forward chaining or backward chaining or both. In the basic method, states of the world, including the initial state and the goal state, are represented as a set of conditions. Operators applied to a state produce a new state. Usually, an operator specifies a set of *preconditions* which must be true in order to apply the operator and a set of *postconditions*: an add and a delete list which indicate conditions which are respectively, now true and no longer true, once the operator has been applied. Forward chaining is matching the preconditions of available operators to the postconditions of the last operator applied (i.e. the conditions of the current state). An operator is then chosen and applied to make progress toward the goal state. Backward chaining is matching the postconditions of available operators to the preconditions of an operator which is already (or proposed to be) in the plan (i.e. the conditions of the goal or some intermediate goal state). In this complementary way to construct plans, progress is made backwards to the initial state.

Clearly the ordering of actions or operators in a plan is critical. Locally, each operator's postconditions must satisfy the preconditions of the next operator in the sequence. Globally, an entire plan can be treated as a meta-operator whose preconditions are the initial state and whose postconditions are the goal state. Time appears implicitly in the use of pre- and postconditions which really mean *before* and *after*. For a plan to be successful, the actions it specifies must be executed in the order specified.

The ordering problem is actually more severe in most cases. Before and after constraints can sometimes apply between operators which are not adjacent in a plan. Care must be taken to make sure that preconditions established for some later operator are not undone by some intermediate operator. Reordering sometimes is necessary and satisfying all the before and after relations within a plan can be quite difficult [Sac77].

Planning requires information about simple temporal relations. The requirements for a representation of time needed for common-sense physical reasoning are more severe. Before dealing with these issues, let us consider more closely two concepts which appear implicitly in the above discussion of planning. These are enablement and causality.

3. ENABLEMENT

In planning, enablement is embodied in the notion of preconditions. In general, if some condition or a set of conditions must be true in order for something to take place, then those conditions, when true, *enable* the action or event in question. Enabling conditions often are present in physical systems. A gun will not fire until the safety lock is released. The water in a bathtub will not flow out until the drain plug is pulled.

Multiple preconditions are usually treated as a conjunction, i.e. every one of the preconditions must be true before the action can take place. To be completely general, disjunctions of preconditions, as well as conjunctions of disjunctions, disjunctions of conjunctions, etc. should be allowed. A child soon learns that getting permission from one of his parents is usually a lot easier than getting permission from both.

Once the concept of enablement is represented, it is useful to distinguish between *tendencies*, which by definition must occur whenever their preconditions are satisfied, and *actions*, which require an actor and a decision in addition to satisfied preconditions [Rie76]. Once the safety lock on a gun is released, the gun is ready to fire, but will not until the trigger is pulled. In contrast, once the pin is pulled on a hand grenade, the grenade will explode, with no further action required.

Tendencies are always lurking about. Satisfy the preconditions of an action, and it *may* happen. Satisfy the preconditions of a tendency, and it *will* happen.

A better example of a tendency is gravity. When an object is unsupported, it will fall, immediately, period. A grenade will explode after its pin is pulled, but not immediately. The ancillary concept here is *imminency*. Later, in the context of rates, intervals, and thresholds, I will discuss how imminency can act as a demon and precipitate actions (like heaving the grenade as fast and as far as possible).

4. CAUSALITY

Causality is a more powerful concept than enablement. Enabling conditions specify how to *allow* things to happen. Causal relations specify how to *make* things happen. Our understanding of the world is permeated with cause-effect relationships, and our ability to interact with the environment is based in these representations.

The philosopher David Hume argued that no matter how well a cause-effect relationship is thought to be understood, all such relationships are ultimately and only empirically formed. The best that can ever be stated about a cause-effect relationship is that two events have been observed always to be conjoined, but there is never anything necessary or compelling about that connection. This statement may well be true, but its usefulness in the domain of common sense is nil, especially if you are the guy holding the hand grenade.

Causal relations, at the least, reflect the way we perceive the universe to behave, and are therefore perfectly suited to aiding us in understanding, predicting, and manipulating our world. Whatever they say, or fail to say, about the real nature of the universe is irrelevant in the domain of common sense. Identifying causal relations is our most powerful way to understand how things work.

Patrick Winston recognizes the importance of causal relations for reasoning by analogy [Win81]. Analogizing means finding a partial match between two concepts, situations, etc. and then mapping additional information from one to the other. Matching is an integral part of Winston's analogy system which compares situations represented primarily by agents, actions, and objects. The matching process is subject to combinatorial explosion, but Winston controls it by matching only causal relations. He argues, and rightly so, that causal relations are the most important relations to match, and that an analogy must succeed there if it is to work at all.

It is interesting to note that causal relations are almost always described temporally, even if they do not have a temporal aspect. A causes B is almost synomomous with A is before B. Especially with mechanical devices, causal relations often apply to simultaneous events, but the desire to order them in time is very compelling.[1]

Sometimes a temporal cause-effect relationship is turned around and the existence of a cause is inferred from the observation of an effect. This is an unusual form of reasoning at first glance, but not an uncommon one. The whole meaning of the word *evidence* is based on the idea of tracing a causal link backwards.

Casusality is a pervasive concept. And it is not a simple one. What would be a good representation of causality? Chuck Rieger has addressed this question and constructed a representation of cause-effect relationships which identifies different flavors of causality [Rie76]. He distinguishes causal links which have preconditions, and those that do not. He further notes the difference between discrete and continuous causes. A light only has to be turned on once, but the poor boy in Holland with his finger in the dike wall had to stay there indefinitely. Finally, Rieger distinguishes direct and byproduct causes. Sometimes, an effect cannot be achieved directly, but only as a side-effect. This is different from indirect causes which follow a chain of causality. In all, there are eight flavors of causality, constructed from all combinations of enabled/free, discrete/continuous, and direct/byproduct (see **Fig. 1**).

Causal relations in physical systems take the form of functional dependencies between quantities. Ken Forbus has formalized the notion of a qualitative functional dependency [For82]. The weakest form states only:

When x changes, y changes (x and y are quantities).

The two stronger forms specify the sign of the dependency. A positive qualitative functional dependency

---

1. Strictly speaking, no effect can propagate faster than the fundamental limiting velocity - the speed of light - according to Einstein's theory of relativity. If you push on one end of a structure that is large enough and rigid enough, the other end cannot start moving before light could get there. Of course, this knowledge is about as applicable to everyday happenings as Hume's observations.

Fig. 1. Flavors of Causality (after [Rie76]).
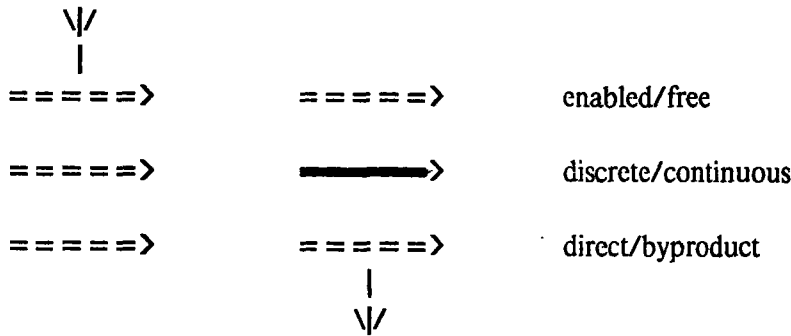
means:

> When x increases, y increases and
> when x decreases, y decreases.

And of course, a negative qualitative functional dependency means:

> When x increases, y decreases and
> when x decreases, y increases (see **Fig. 2**).

The descriptive power of these relations is enormous, although only the sign of the dependency ever is specified; the actual values of the quantities are superfluous. The well-known General Gas Law, relating temperature, pressure, and volume may be expressed as in **Fig. 3**.

This simple representation, which specifies only directions of change, encodes common-sense understanding of processes such as: blowing up a balloon, a teapot whistling on a burner, an aerosol can exploding when heated.

Note that the functional dependencies in this description are bidirectional, i.e. there is no distinction between dependent and independent variables. This is not always the case, of course.

Another well-known functional dependency is illustrated in **Fig. 4**. This is a very simple model. Inflation and unemployment are not quantities which can be modified directly, or in isolation. A better model of the economy might be as shown in **Fig. 5**.

This model is not intended to be taken seriously; it is full of inconsistencies (although the individual

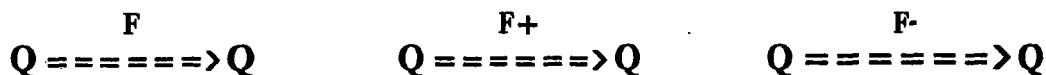**Fig. 2. Qualitative Functional Dependencies (after [For82]).**
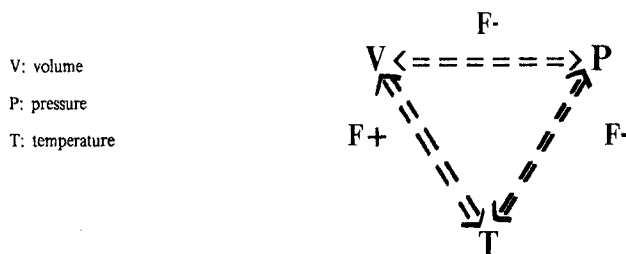
Fig. 3. The General Gas Law.

V: volume

P: pressure

T: temperature

$$V \Longleftarrow\!\!=\!=\!=\!=\!=\!\Longrightarrow P$$

with the triangle labeled F- (top), F+ (left), F- (right), and vertices V, P, T.

Fig. 4. Joe Schmoe's model of the economy.

I: inflation

U: unemployment

$$I \Longleftarrow\!\!=\!=\!=\!=\!=\!\Longrightarrow U$$

(labeled F- above the arrow)

Fig. 5. An economist's model of the economy.

C: consumption

G: Gross National Product

I: inflation

L: lending (interest) rates

M: money supply

N: investment

P: productivity

S: government spending

T: taxes

U: unemployment

W: wages
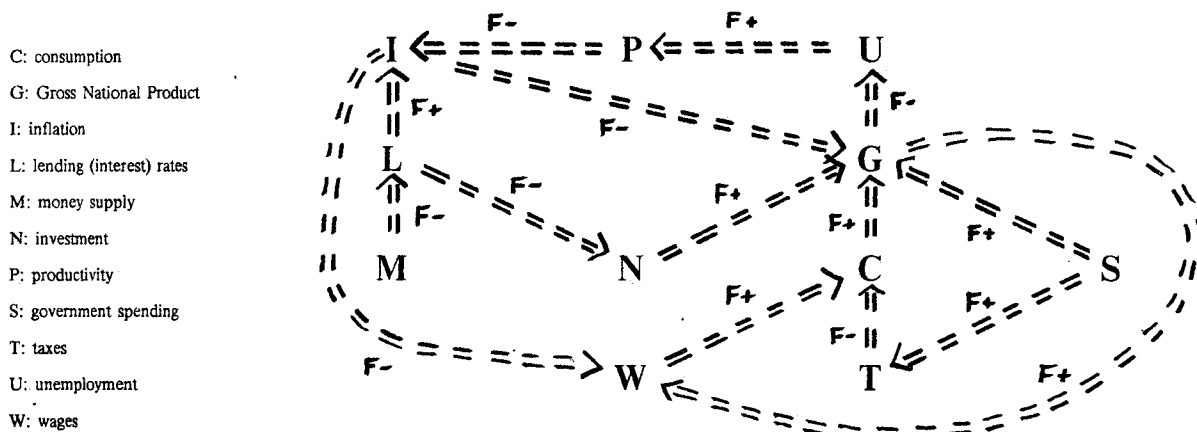
dependencies seem to make some sense). The only purpose of this example is to illustrate *abstraction*. An abstraction is a concept or a representation which distills out the most important aspect(s) of a host of other concepts or representations which it subsumes. An abstraction trades off compactness and accessibility with accuracy and explanatory power. For the same amount of effort, dealing with abstractions can get more work done, but because detail is lost, there is always the danger of missing some point or becoming stymied for lack of information.

Concerning causal relations and functional dependencies, any such relation may in fact subsume an entire chain of causality (see Fig. 6). Our simple model of the economy does not indicate how to change the level of inflation or unemployment. The more detailed model reveals a quantity - the money supply - which is settable by the Federal Reserve Board and provides a handle into the physical system of the economy. So one way to decrease unemployment might be to increase the money supply. Then interest rates (or the cost of

Fig. 6. A Causal Chain.

$$Q \overset{F}{==/==>} Q \qquad \text{is in fact} \qquad Q \overset{F}{====>} Q \overset{F}{====>} \ldots \overset{F}{====>} Q \overset{F}{====>} Q$$
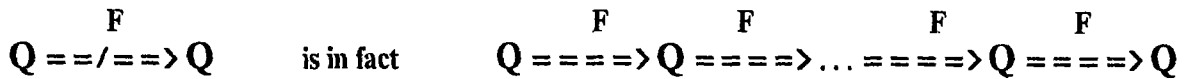
money) will drop, encouraging investment. The gross national product will rise, creating more jobs. Sound easy? See **Fig. 7.**

It is natural and useful to employ hierarchical descriptions of things, whether they be physical systems, or stories, or whatever. The details can be often glossed over - blind reductionism wastes resources, and is boring - but the information is available when it is needed. For example, a light switch turns on a light, and that is sufficient knowledge - until something breaks. The switch may still move up and down and the bulb may be sound, so now what? The next level of description is needed. In particular, a description of the circuit of which the switch and the light bulb are a part.[2]

Fig. 7. The real economy.
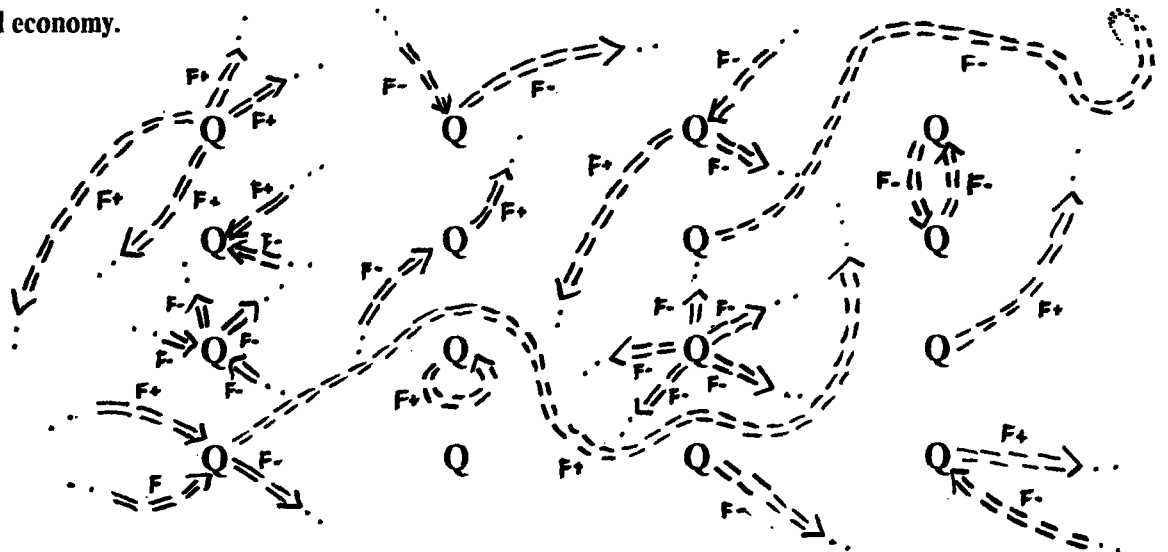
2. For physical systems, it is extremely useful to have both structural and functional descriptions available, but that is another paper.

5. TIME

Characterizing time is a problem which has plagued both philosophers and physicists. Philosophers worry about such issues as the source of the unidirectionality of time. Physicists wonder if time is really continuous.

At the level of interactions between primitive atomic particles, there may be a resolution at which time can be treated as discrete without any loss of explanatory power. In other words, no interaction can take place in a shorter time interval. This quantum unit of time might be the time needed for light to cross the smallest particle known, for instance. Physicists even have a name for this unit - the chronon (analog of the photon or graviton).

Whatever may be true at this most basic level of the universe, there is plenty of evidence that our own sensory perceptions of the world (especially vision) are strobed, rather than continuous. Apparently, there are buffers which hold information and make the resolution deliberately fuzzy, so that there is an appearance of continuity. If this were not the case, movies and television could never work. Both of these media rely on the fact that visual resolution in time is no finer than about one-thirtieth of a second.

What about time's direction? We seem always to be moving inexorably toward the future. What is the source of this unidirectionality? At the quantum level once again, time does not appear to have a preferred direction. Interactions can be played through in either direction; both ways make sense.

But at the level of everyday happenings, time certainly does have a direction. This need not be as mysterious as it seems. The source of this unidirectionality may be based simply in information, memory, and learning. What distinguishes the past from the future? One is known, the other is unknown. Information about the past is accumulated and stored in memory through learning, but no *immutable* information ever is available about the future. At first this observation seems rather obvious and may even be thought to beg the question but I think it is actually quite profound. Would an entity with no learning capacity, i.e. no means of accumulating information, perceive time to have a direction?

The senses also play an important role in constructing a rich representation of time. Until an entity can interact with processes throught its sensory equipment, it cannot have much understanding of "real-time" or of how time can vary subjectively. It is the interaction between perceptions and mental processes that is important. Artificially intelligent agents with a general learning ability and sophisticated sensory devices may one day be quite able to construct their own representations of time. For now, we must worry about constructing that representation.

Now we turn to the question of what would be a good representation of time - one that can capture some common-sense notions of time and which would be useful in the domain of naive physical reasoning.

James Allen has a good general purpose representation of time [All81]. Allen argues that temporal knowledge is almost always concerned with *intervals*, rather than *instants* of time and he has developed accordingly an interval-based representation. He defines five relations that can exist between intervals: *before, equal, overlap, meet, during* (see **Fig. 8**). Unfortunately, as Forbus points out, in the domain of common-sense physical reasoning, it will be necessary to treat instants of time explicitly as well, because

Fig. 8. Relations between intervals (after [All81]).

```
= = = = = =
              = = = = = =    before


= = = = = =
= = = = = =              equal


= = = = = =
       = = = = = =          overlap


= = = = = =
              = = = = = =    meet


= =
= = = = = =              during
```
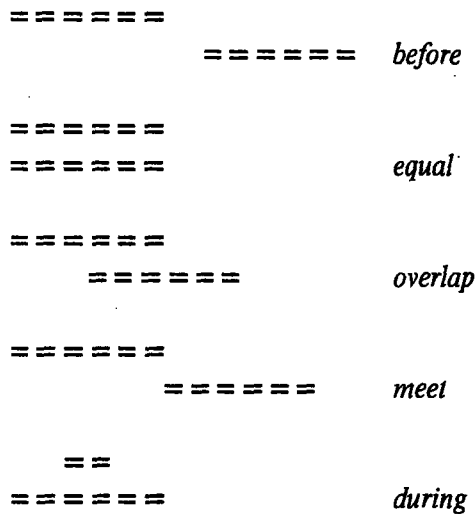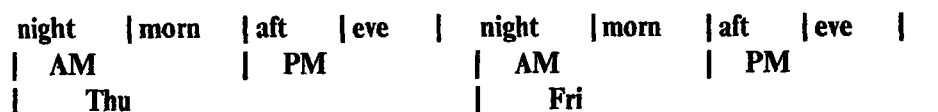
continuous processes often undergo interesting changes at points in time. Allen's model defines the endpoints of intervals, but never treats these points explicitly. Consider a rock thrown into the air. A common-sense representation of this process should reveal the point at which the rock reaches its peak height. Just as important are the intervals during which the rock is rising and falling, but if the representation is going to have a handle on continuity, the event of the rock being stationary must be made explicit also.

Earlier it was mentioned how hierarchical descriptions of causality are very natural and permit a physical system to be treated at several different levels of abstraction. Similar comments apply to the *during hierarchy* which Allen defines. The *during* relation is transitive, and intervals can be naturally arranged in a hierarchy (see Fig. 9), e.g. today is in the month of December in the year 1982 and I know therefore that today is in the year 1982.

One of the powers of the during hierarchy is that *before* relations can be inherited. If some day in the year 1981 is under consideration, the fact that that day is before today can be inferred immediately by noting that that day is *during* 1981, today is *during* 1982, and 1981 is *before* 1982. The intervening months, weeks, days, and their relations need not be considered.

Cycles can be easily represented in Allen's representation by making use of the during hierarchy. At some level of the hierarchy, a set of intervals (or a single interval) is repeated. The individual intervals which

Fig. 9. A During Hierarchy (after [All81]).

```
night   | morn  | aft  | eve  |  night    | morn  | aft  | eve  |
| AM              | PM           | AM              | PM            |
|   Thu                          |   Fri                          |
```

match across different cycles are naturally disambiguated by virtue of having *during* relations (of arbitrarily deep nesting) with disjoint intervals somewhere above in the hierarchy (e.g. December 1982 is not confused with December 1983).

Allen's model can deal with the concept of *now*, the dynamic present, in a computationally efficient manner. If *now* is treated as an interval at the finest level of the during hierarchy, then only relations at that level need to be updated routinely. Other relations can continue to be inherited from the hierarchy. Only when an update results in the crossing of an interval boundary higher in the hierarchy do relations at that level need to be updated. By treating now as an interval rather than a point, Allen avoids the problem of updating the present continuously. Of course, each *now* interval itself spans a continuum. Allen's solution is reasonable however; no one knows how to deal with continuity directly.

Allen also investigates the notion of *persistence*. When some condition is known to be true, it is normal to continue to expect it to be true until it is learned otherwise. Within Allen's description, it is easy to define an interval over which something is believed to be true. The interval can be initially open-ended. When faced with the question of whether something is true at some given time, the during hierarchy can be consulted to provide an answer with the current knowledge. As we will see, in the domain of common-sense physical reasoning, the intervals that should be represented are those during which the signs of quantities and signs of rates are persistent.

6. QUANTITY

To understand processes, we need to know how quantities can interact and change over time. Besides causality and time, *quantity* needs a representation in order for change - in the domain of physical systems and processes - to be fully characterized.

Ken Forbus has done the pioneering work in this area in developing his Qualitative Process Theory [For82]. Part of this theory is a rich representation for quantity. Here I will reproduce the important aspects of that representation, augmented with a few observations of my own.

A quantity is actually two quantities, an *amount* and a *rate*, each of which consists of a *magnitude* and a *sign*. The amount of one quantity can be the rate of another (e.g. the amount of velocity is the rate of position).

In common-sense physical reasoning, of primary concern are the signs of quantities and their relative magnitudes; these provide powerful information needed to understand how things change.

A quantity represents a continuous range of values, but with the preceding comment in mind, explicit access to each and every one of these values is not desired. Yet there are particular values which are of interest - these are the points at which process changes occur, where the interactions between quantites in a physical system change. Thus Forbus defines the *Quantity Space*, the continuous range from which values for

the amount or rate of a quantity are drawn, and he gives it structure by specifying *limits* and *distinguished points*, the points at which process changes occur. A limit is, of course, some maximum or minimum value that a quantity can attain, while a distinguished point is any other important value which is made explicit. Fig. 10 illustrates the Q-spaces needed to describe a rock's flight through the air. Note that a single point can qualify as a Quantity Space also. Fig. 11 shows the Q-space for the height of water in a sink. Note that a distinguished point is included which corresponds to the height of the safety drain. At this point, the water may stop rising, although the tap has not been turned off. This is an example of a process change.

A distinction may be made between *known* points and *existent* points in a Quantity Space. A known point is essentially a constant; its value does not change (although that value need never be assigned a number). An example would be the height of a rock when at rest on the ground. An existent point, on the other hand, has a variable value which is functionally dependent on the values of other quantities. For instance, the peak height of the rock's flight depends on the rock's initial velocity. Nevertheless, there will exist always a peak height for the rock's flight, whatever its value, and the Quantity Space describing the height of any rock in flight must include this point as one of the limits.

It is useful to specify *correspondences* between points in Quantity Spaces which are related by a functional dependency. In other words, whenever the cause quantity has a certain distinguished value the effect quantity is guaranteed to have some other certain distinguished value. An example would be the peak height of a rock corresponding to zero velocity.

There are two ways that a quantity can change. One is that some other quantity changes on which the given quantity is functionally dependent. The other is that the quantity has a rate which is non-zero. There

---

Fig. 10. Quantity Spaces (after [For82]) for the flight of a rock.


```
= = = = =   peak         = = = = =   + max
    |                        |
    |                    ------------   zero      = = = = =   -gravity
    |                        |
= = = = =   ground (zero)  = = = = =   -max
   Height                  Velocity                Acceleration
```
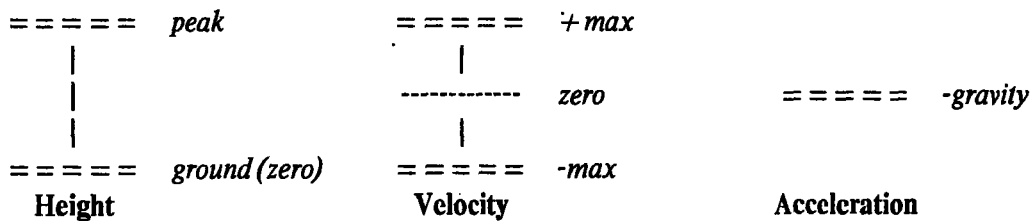
Fig. 11. Quantity Space for the height of water in a sink.

```
= = = = =   full
    |
------------   safety
    |
    |
    |
    |
= = = = =   empty
```
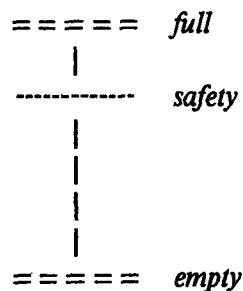
may be any number of contributions, which Forbus calls *influences*, to the rate of a quantity. For instance, the height of water in a sink can be affected by the flow-in of the tap, the flow-out of the drain, the flow-out of the safety drain, or the flow-out onto the floor. The sum of the individual influences tells how the quantity will actually change.

As is the case with functional dependencies, there are three forms of influences. The weakest form specifies simply:

q changes (*q* is a quantity).

The stronger forms, positive and negative influences specify, respectively:

q increases,
q decreases (see **Fig. 12**).

Functional dependencies and influences are similar in that they both tell how a quantity can change. The difference is that an influence need only be non-zero to effect a change while a quantity which is the cause in a functional dependency relation must itself be changing, i.e. must itself have a non-zero rate, to effect a change.

Hierarchical descriptions of causality and time were seen to be very natural, and the same observation applies to quantity. For any quantity, there are always several scales available, corresponding to different levels of abstraction. And as is true with any abstraction, sometimes detail is needed and sometimes it is not. If I live in Brookline and wish to travel to Pasadena, I first worry about getting from Boston to Los Angeles. Later I consider distance at successively smaller scales until I reach a certain apartment in a complex on a street two blocks from Caltech.

Quantities in the description of a physical system can receive additional structure through semantics which tell the different ways the system may be used. For example, the Q-Space for the height of water in a sink might contain additional points which correspond to the water level needed to shave, to wash gym socks, etc.

Common-sense, qualitative representations for quantities always can contain certain default points which are existent, rather than known. Examples are halfway, near one limit, near the other limit.

Fig. 12. Influences (after [For82]).

```
           I                         I+                        I-
Q = = = = = =>Q            Q = = = = = =>Q            Q = = = = = =>Q
```
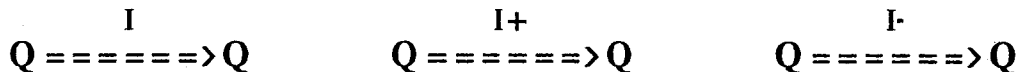
7. PROCESSES

Now we are ready to synthesize the representations for causality, time, and quantity discussed in the previous sections to describe processes, or how quantities interact and change over time.

For the moment, let us make no distinction between functional dependencies and influences and consider only the *net* change that results in a quantity due to any number of contributions of these two types. If we distinguish between positive and negative contributions, then we can define the useful abstractions in **Fig. 13** which describe the alternate underlying processes which result, on the surface, in an increasing, decreasing, or static quantity. For instance, an increasing quantity may have only positive contributions affecting it, or there may be both positive and negative contributions - but the positive ones are greater. What is not being made explicit is information about magnitudes, only the relative values of the sums of positive and negative contributions are indicated.

Of particular interest is the abstraction for equilibrium. Here positive and negative contributions exactly cancel, producing a static quantity hiding some very dynamic process.

To support reasoning about how processes themselves can change, information about how one of the abstract states in **Fig. 13** can be transformed into another is needed. The transition table in **Table 1** provides this information. There are three operators which can be applied to quantities which are already increasing, decreasing, or static: 1) add to the contributions of a particular sign (this could mean increasing a single contribution of that sign, or adding another contribution, or any combination of these actions), 2) subtract from the contributions of a particular sign, and 3) delete all the contributions of a particular sign. The table shows what new state results from applying one of these operators to a given state. For instance, adding some negative contribution to a state of equilibrium produces a state of negative tradeoff.

Two interesting observations can be made from this table. One is that sometimes it is unclear what state will result without any information about magnitudes. As an example, adding some positive contribution to a state of decrease may result in positive tradeoff, negative tradeoff, or equilibrium.

The other observation is that there is no way to achieve equilibrium. This is not very satisfactory. Equilibrium would appear to be a terribly unstable state, requiring positive and negative contributions to

---

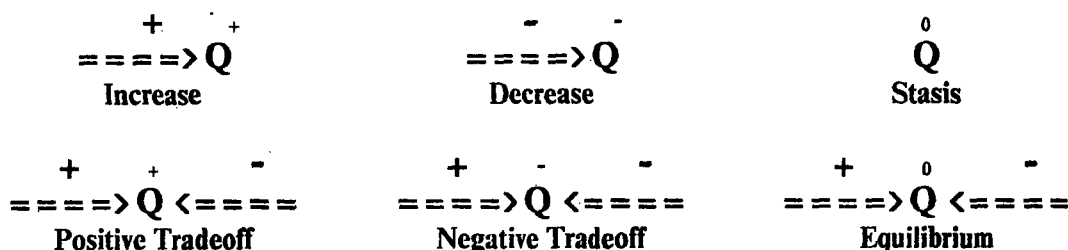Fig. 13. Abstractions for 1st-Order Change.

Table I. 2nd-order Change.

		Add		Subtract		Delete	
		+	-	+	-	+	-
I: increase	I	I	?	I		S	
D: decrease	D	?	D		D		S
S: stasis	S	I	D				
P: positive tradeoff	P	P	?	?	P	D	I
N: negative tradeoff	N	?	N	N	?	D	I
E: equilibrium	E	P	N	N	P	D	I

exactly balance.

What is missing is the concept of a feedback loop (see **Fig. 14**). When a quantity moves away from equilibrium, a causal chain of functional dependencies of the opposite sign is activated which brings the quantity back to the stable value. Once the water in a sink rises to the level of the safety drain, a negative feedback loop is enabled. As soon as the water rises above the critical height, the flow-out of the safety drain begins and the water will rise no further. When trying to understand a physical system with quantities in equilibrium, an intelligent program should look for feedback loops. Also, an intelligent agent interacting with a physical system should be able to conjecture feedback loops of which its own actions are a part. Imagine a household robot drawing a bath. Getting the temperature just right involves playing with the cold water and hot water faucets.

The above discussion relates causality and quantity; time is mentioned implicitly through rates. Now time will be treated more concisely.

If a quantity is changing, there is an interval over which it is changing. For any *persistent* interval, during which the sign of a quantity's rate does not change, the amount of the quantity at the end of the interval will be *more, less,* or *the same* for respectively, a *positive, negative,* or *zero* rate.

As Forbus points out, this is a very weak notion of integration over time. Rieger's concept of *thresholding* is a bit more powerful in that it means continuation toward some value, without requiring the appropriate rate to be persistent within the interval. But, as with any abstraction, information needed (to predict whether a value actually will be attained) may not be available.

The important relations between rate, amount, and time are shown in **Fig. 15.** The qualitative

Fig. 14. Feedback Loops.

Fig. 15. Rate = Amount/Time.

A: amount
T: time
R: rate

```
            F+
  A <= = = = = => T
   ↖             ↗
 F+  ↘         ↗  F-
      ↘       ↗
       ↘    ↗
        ↘ ↙
          R
```

---
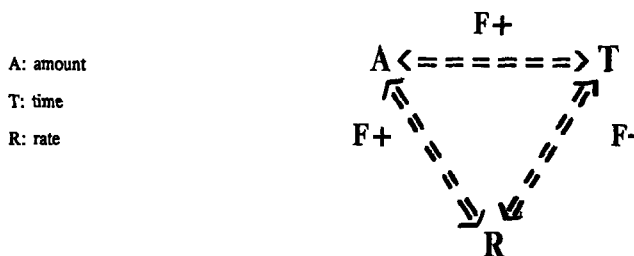
statements that can be lifted from this diagram are:

In the same interval of time, a greater rate results in a greater change in an amount.

For a constant rate, a longer interval results in a greater change in an amount.

The same change in an amount can be achieved with a greater rate and a shorter interval.

The important point here is that time itself can be treated as a quantity. This is a very necessary observation if we wish an intelligent program to be able to perform time-domain analyses of processes or to construct time simulations of physical systems. Earlier, the distinction between functional dependencies and influences as sources of change in a quantity was blurred. Now that distinction can be formalized. An influence (or collectively, a rate) can be treated as a functional dependency on time. To perform a simulation, time can be represented as a quantity with an unchanging positive rate (the unidirectionality of time!).

Consider the flight of a rock once again. **Fig. 11** shows the quantities which are needed. There are correspondences between zero velocity and the peak height of the rock and between top upward and top downward velocity and the ground (zero) height of the rock. Velocity is an influence on height and gravitational acceleration is an influence on velocity (see **Fig. 16**).

Note that it is unclear when the influence of velocity on height will be positive and when it will be negative without identifying persistent intervals. However, if influences are treated as functional dependencies on time (see **Fig. 17**), persistent intervals can be identified and it is easy to figure out what will happen.

The rock's flight starts with ground (zero) height and maximum positive velocity. Acceleration is always negative, so all of time is a persistent interval in which velocity will decrease. Because of the persistence, the fact that the zero velocity point will be reached can be inferred. The range from maximum positive to zero velocity defines a persistent interval in which height is increasing. Zero velocity corresponds to peak height.
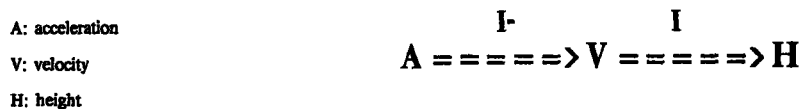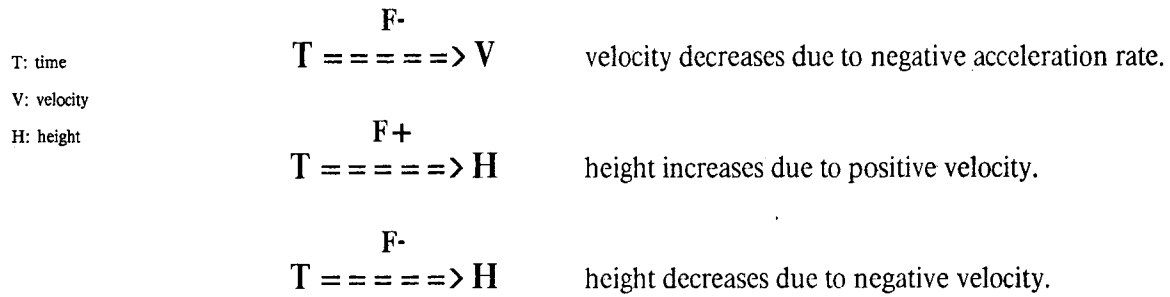

Fig. 16. Influences become...

A: acceleration
V: velocity
H: height

$$A = = = = = => V = = = = = => H$$

Fig. 17. ...functional dependencies on time.

T: time

V: velocity

H: height

$$\begin{array}{c} \text{F-} \\ T ===== \text{>} V \end{array} \qquad \text{velocity decreases due to negative acceleration rate.}$$

$$\begin{array}{c} \text{F+} \\ T ===== \text{>} H \end{array} \qquad \text{height increases due to positive velocity.}$$

$$\begin{array}{c} \text{F-} \\ T ===== \text{>} H \end{array} \qquad \text{height decreases due to negative velocity.}$$
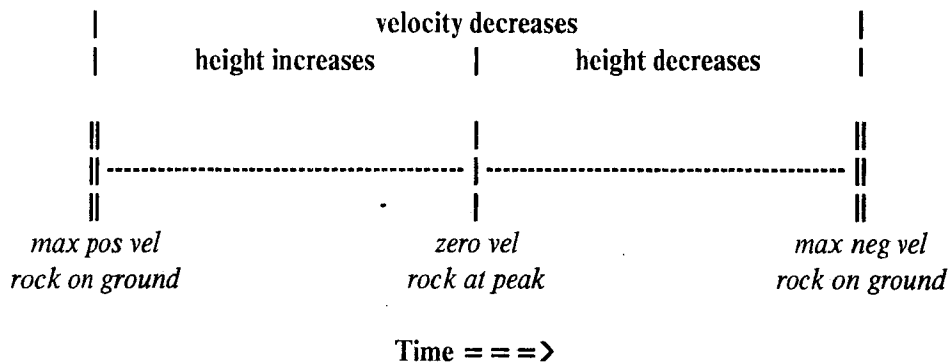
Time is always increasing.

If we define *episodes* and *events* as the analogues of intervals and instants in the time simulation of a process, then the final result of all these inferences and observations is a timeline of episodes and events characterizing the flight of a rock (see **Fig. 18**).

Notice how the structures of the Q-spaces of height, velocity, and acceleration map onto this timeline. Height maps onto the height-increasing and height-decreasing episodes. Velocity maps onto the entire velocity-decreasing episode. Acceleration, which is free of any functional dependency on time, maps onto the entire time continuum. Notice also how the episodes in this timeline form a during hierarchy.

This discussion of persistent intervals and timelines leads naturally to Patrick Hayes' concept of *histories* [Hay79]. A history describes more than the sequence of values that a quantity attains over some time interval. A history is a closure of causal relations within a bounded chunk of space-time. Hayes gives the example of an ordinary room to illustrate a causal closure. The room is an enclosure which neatly divides space into an *inside* and an *outside*. The space inside the room becomes a history by being extrapolated over a time interval, say an afternoon. Processes occurring on the outside of the history can be ignored effectively. This amounts to saying that none of the effects of those processes will be observed inside the room during that afternoon. It is only when two histories intersect that causes from one history can produce effects in another (e.g. a meteor

---

Fig. 18. Timeline of the flight of a rock.

```
        |              velocity decreases              |
        |    height increases    |   height decreases  |

      ||                         |                      ||
      ||-------------------------|----------------------||
      ||               ·         |                      ||
   max pos vel               zero vel               max neg vel
   rock on ground            rock at peak           rock on ground

                        Time = = =>
```

crashes through the ceiling of the room).

A history is more general than a trajectory; it describes the coupling of chunks of space and time rather than a succession of point-instant pairs. Histories can describe whatever trajectories can (such as a bullet successfully shattering a clay pigeon) and more (such as the failure of a surprise birthday party when the birthday boy or girl does not show up). The concept of a history appears in the stereotype murder trial when the prosecuting attorney queries the defendant: "Where were you on the night of August the 18th..."

Earlier, it was shown how additional structure can be imposed on a Quantity Space when rich semantics are brought to bear on different ways to think of a quantity. Similarly, when time is treated as a quantity, additional, existent instants can be overlaid on an interval to give it more structure. For example, nearness to an endpoint becomes *imminency* when the quantity in question is time. Recognizing instants at which some process change becomes imminent can act as a demon to precipitate some other parallel process or action which is to be completed before the original process change. As an example, a person drawing a bath may read the newspaper or otherwise idle away time while the tub is filling. However, at the point where the time to disrobe is judged to be equal to the time remaining to fill the tub, the person will put the paper aside and begin preparing for the bath.

8. SUMMARY

This paper has synthesized several ideas and introduced a few more about how to characterize change in the context of common-sense physical reasoning. In particular, representations for causality, time, and quantity were treated. The result is a rich framework for describing processes and physical systems (see example in **Fig. 19**). This framework is intended to provide support for any of several research issues, including describing and reasoning about processes and physical systems, troubleshooting of engineered systems, and learning how such systems work.[3]

---

3. The last, if anyone is interested, is my motivation.

Fig. 19. Description of a Sink.

1. Quantity-Spaces

```
=== open          === open          === max
 |                                    |                    === full
=== closed        === closed        === zero                |
Faucet Setting    Stopper Setting   Tap Flow-In           ------- safety
                                                            |
=== max           === max           === max              === empty
 |                 |                 |               Water Height
=== zero          === zero          === zero
Drain Flow-Out    Safety Flow-Out   Top Flow-Out
```
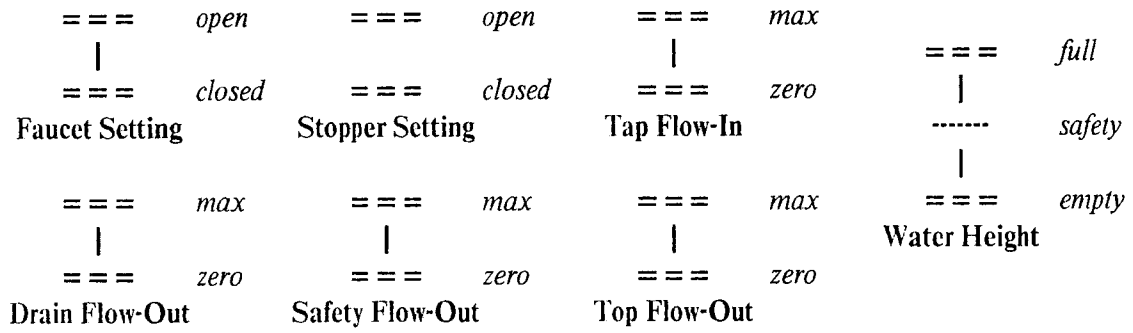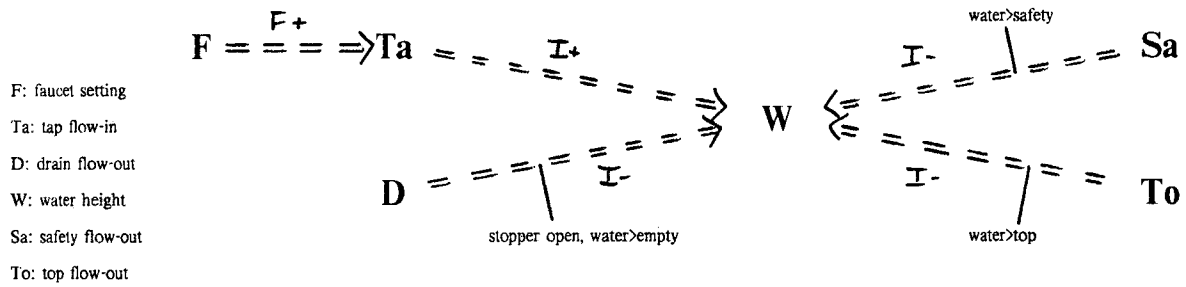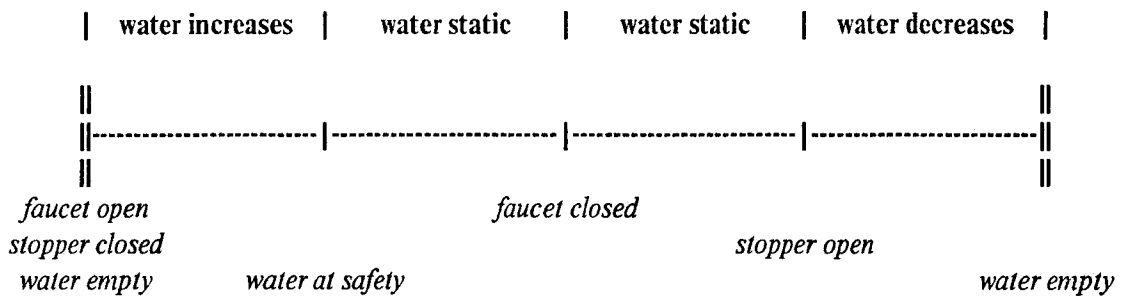
correspondence between Faucet *open* and Tap Flow-In *max*.
correspondence between Faucet *closed* and Tap Flow-In *zero*.

2. Causal Relations

F: faucet setting
Ta: tap flow-in
D: drain flow-out
W: water height
Sa: safety flow-out
To: top flow-out

$$F \overset{F+}{===} \Rightarrow Ta \overset{I+}{=} \cdots \Rightarrow W \overset{I-}{\Leftarrow} \cdots = Sa \quad (water{>}safety)$$

$$D \overset{I-}{=} \cdots \Rightarrow \quad (stopper\ open,\ water{>}empty) \qquad \overset{I-}{=} \cdots = To \quad (water{>}top)$$

3. A possible timeline.

```
| water increases | water static | water static | water decreases |

||                                                              ||
||------------------|------------------|------------------|------||
||                                                              ||
faucet open                                                   
stopper closed                           faucet closed
water empty        water at safety          stopper open
                                                          water empty
```

9. ACKNOWLEDGEMENTS

10. REFERENCES

[All81] James F. Allen, "An Interval-Based Representation of Temporal Knowledge", *7th International Joint Conference on Artificial Intelligence*, 221-226, 1981.

[deK75] Johan de Kleer, "Qualitative and Quantitative Knowledge in Classical Mechanics", *MIT Artificial Intelligence Lab* Technical Report 352, December 1975.

[deK79] Johan de Kleer, "Causal and Teleological Reasoning in Circuit Recognition", *MIT Artificial Intelligence Lab* Technical Report 529, September 1979.

[FiN71] Richard E. Fikes and Nils J. Nilsson, "STRIPS: A New Approach to the Application of Theorom Proving to Problem Solving", *Artificial Intelligence 2*, 189-208, 1971.

[For81] Kenneth D. Forbus, "A Study of Qualitative and Geometric Knowledge in Reasoning about Motion", *MIT Artificial Intelligence Lab* Technical Report 615, February 1981.

[For82] Kenneth D. Forbus, "Qualitative Process Theory", *MIT Artificial Intelligence Lab* Memo 664, February 1982.

[Hay79] Patrick J. Hayes, "The Naive Physics Manifesto", in *Expert Systems in the Micro-Electronic Age*, ed. D. Michie, Edinburgh: Edinburgh University Press, 1979.

[Len81] Wendy G. Lenhert, "Plot Units and Narrative Summarization", *Cognitive Science 4*, 293-331, 1981.

[McD81] Drew McDermott, "A Temporal Logic for Reasoning about Processes and Plans", *Yale Univeristy Dept. of Computer Science*, Research Report 196, 1981.

[Rie76] Chuck Rieger, "An Organization of Knowledge for Problem Solving and Language Comprehension", *Artificial Intelligence 7*, 89-127, 1976.

[Sac77] Earl D. Sacerdoti, *A Structure for Plans and Behavior*, New York: Elsevier North-Holland, 1977.

[Sta80] James L. Stansfield, "Conclusions from the Commodity Expert Project", *MIT Artificial Intelligence Lab* Memo 601, November 1980.

[Win81] Patrick H. Winston, "Learning New Principles from Precedents and Examples", *MIT Artificial Intelligence Lab* Memo 632, May 1981.