The Power of Physical Representations

Varol Akman, Paul J. W. ten Hagen

Leibniz's (1984) *An Introduction to a Secret Encyclopedia* includes the following marginal note:

> Principle of Physical Certainty: Everything which men have experienced always and in many ways will still happen: for example that iron sinks in water (Leibniz 1984).

In our daily lives, we routinely use this principle. Thus, we know that we can pull with a string but not push with it; that a flower pot dropped from our balcony falls to the ground and breaks; that when we place a container of water on fire, water might boil after a while and overflow the container.

The origin of such knowledge is a matter of constant debate. It is clear that we learn a great deal about the physical world as we grow up. However, even philosophers were tricked by the mechanisms that achieve this knowledge; for example I. Kant said:

> All our knowledge begins with sense, proceeds thence to understanding, and ends with reason, beyond which nothing higher can be discovered in the human mind for elaborating the matter of intuition and subjecting it to the highest unity of thought. At this stage of our inquiry it is my duty to give an explanation of this, the highest faculty of cognition, and I confess I find myself here in some difficulty.

Commonsense reasoning about the physical world, as exemplified by "Iron sinks in water" or "If a ball is dropped it gains speed," will be indispensable in future programs. We argue that to make such predictions (namely, envisioning), programs should use abstract entities (such as the gravitational field), principles (such as the principle of superposition), and laws (such as the conservation of energy) of physics for representation and reasoning. These arguments are in accord with a recent study in physics instruction where expert problem solving is related to the construction of physical representations that contain fictitious, imagined entities such as forces and momenta (Larkin 1983). We give several examples showing the power of physical representations.

In this article, we argue that some difficulties regarding commonsense reasoning about the physical world can be overcome by using fictitious entities, laws, and principles of physics. Our argument is along the lines of a recent study in physics instruction where expert problem solving is attributed to the construction of physical representations that contain imagined entities (Larkin 1983).

After a short overview of the role of mental models, we present the motivation for this research. In the next section, we study the nature of physics and a simple abstraction mechanism based on essential attributes and influences. Envisioning and naive versus physical representations are treated next. A brief account of the content of mechanics follows, inspired by, yet basically incomparable to, Hayes's (1983) classification of the fluids domain in his seminal paper. A set of problems showing the power of physical representations can be found in Some Examples. Some interesting research efforts covering the related areas of naive physics and qualitative reasoning are examined in Other Related Research. Finally, the last section summarizes our points and suggests some directions.

Mental Models

Why do we advocate that a theory of commonsense reasoning about the physical world
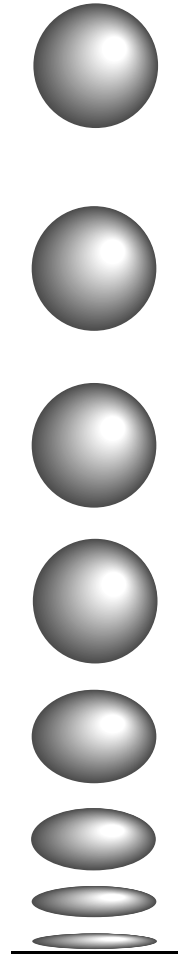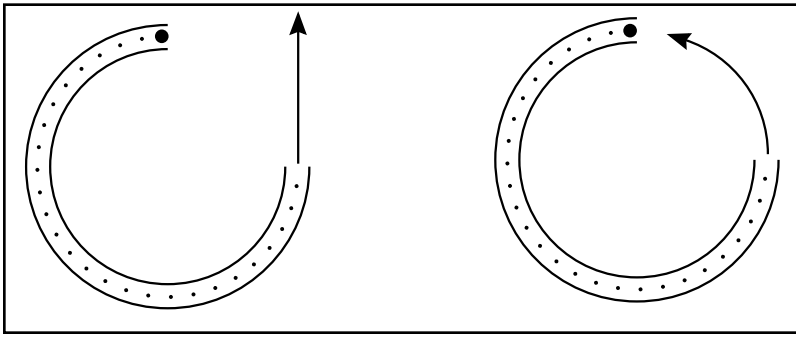
Figure 1. Correct and Incorrect Responses for the Spiral Tube Problem.

should be firmly based on physics? Although we see the psychological literature on this subject as a valuable source of information, studies in learning show, maybe expectedly, that human subjects have fuzzy and even wrong ideas about the physics of everyday life. DiSessa (1982) found out that a group of elementary school students learning to control a computer-simulated Newtonian object invariably had the wrong Aristotelian expectation that bodies must move in the direction they are last pushed. Another similar study by McCloskey (1983) reports that assumptions of the naive theories of motion are quite consistent across college students. It turns out that the theories developed by different individuals are best described as different forms of the same basic theory. What is striking is that this basic theory is highly inconsistent with the fundamental principles of classical physics. McCloskey shows that this naive theory of motion is similar to a pre-Newtonian physical theory—the medieval impetus theory—that the act of setting a body in motion imprints in the object a force, or impetus, which keeps the object in motion. Figure 1 illustrates a case that is similar to the one examined by McCloskey. Asked about how a metal ball put into the end of the tube and shot out of the other end at high speed would behave, more than half the subjects drew the incorrect path.

Motivation

Why is commonsense physical knowledge useful? In this article, we are concerned with mechanics. Therefore, we might answer this question from a viewpoint focusing on mechanical design. We identify two aspects.

First is that mechanical design results in physical objects. After production, a designed object is left in the physical world. From this moment on, it interacts with an environment where physical notions such as force, motion, collision, and so on, are in existence. If we

are interested in reliable products that continue to correctly function under various disturbances, we must take heed of these notions and their effects on our designed objects. Thus, during the design stage, we might want to know what happens to a nuclear reactor when a pressure regulator starts to malfunction. Depending on the outcome of such simulations, we can embed more security checks and redundancies in our designs.

Second is that all mechanical inventions are firmly based on a deep understanding of the physical world and its laws. If we want to automatically design new devices, we need a design system that has an appreciation for physical phenomena. For example, even a simple can opener is a device unifying diverse physical notions such as friction, force, rotation, and cutting. When human designers invent a new device, they use their physics knowledge in a fundamental way to reason about the functioning of the device under consideration. (The reader is referred to Veth [1987]; Akman, ten Hagen, and Tomiyama [1987]; and Akman et al. [1988] for detailed accounts of why commonsense physical knowledge might be crucial for realizing intelligent computer-aided design systems).

The Nature of Physics

Observing that all our ideas in physics require a certain amount of common sense in their application, we see that they are not pure mathematical or abstract ideas. In fact, nearly every page of Feynman, Leighton, and Sands (1966) has a caveat to this effect, as the following excerpts show:

> This system [a system of discourse about the real world] is quite unlike the case of mathematics, in which everything can be defined, and then we do not *know* what we are talking about. In fact, the glory of mathematics is that *we do not have to say what we are talking about*. The glory is that the laws, the arguments, and the logic are independent of what "it" is. If we have any other set of objects that obey the same system of axioms as Euclid's geometry, then if we make definitions and follow them out with correct logic, all the consequences will be correct, and it makes no difference what the subject was.
>
> . . [W]e cannot just call *F = ma* a definition, deduce everything purely mathematically, and make mechanics a mathematical theory, when mechanics is

a description of nature. By establishing suitable postulates it is always possible to make a system of mathematics, just as Euclid did, but we cannot make a mathematics of the world, because sooner or later we have to find out whether the axioms are valid for the objects of nature. Thus we immediately get involved with these complicated and "dirty" objects of nature, but with approximations ever increasing in accuracy (pp. 12-2–12-3).

Let us now study in detail why we cannot make mechanics a mathematical theory.

The Meaning of Physical Laws

Centuries of scientific activity gave rise to an enormous body of physical knowledge that can be found in textbooks. The aim is to provide an account of how the physical world behaves. Theoreticians realized that mathematics is an excellent tool for physics because all laws can be written in symbolic form with absolute clarity and economy. Yet physical formulas by themselves do not provide enough insights.

Although it is interesting and worthwhile to study the physical laws simply because they help us to understand and to use nature, one ought to stop every once in a while and think, "What do they really mean? The meaning of any statement is a subject that has interested and troubled philosophers from time immemorial, and the meaning of physical laws is even more interesting, because it is generally believed that these laws represent some kind of real knowledge. The meaning of knowledge is a deep problem in philosophy, and it is always important to ask, "What does it mean?"

Let us ask, "What is the meaning of the physical laws of Newton, which we write as *F = ma?* What is the meaning of force, mass, and acceleration?" Well, we can intuitively sense the meaning of mass, and we can *define* acceleration if we know the meaning of position and time. We shall not discuss these meanings, but shall concentrate on the new concept of *force.* The answer is equally simple: "If a body is accelerating, then there is force on it." That is what Newton's laws say, so the most precise and beautiful definition of force imaginable might simply be to say that force is the mass of an object times the acceleration. Suppose we have a law which says that the conservation of

momentum is valid if the sum of all external forces is zero; then the question arises, "What does it *mean,* that the sum of all the external forces is zero?" A pleasant way to define that statement would be: "When the total momentum is a constant, then the sum of the external forces is zero." There must be something wrong with that, because it is just not saying anything new. If we discovered a fundamental law, which asserts that the force is equal to the mass times the acceleration, and then *define* the force to be the mass times the acceleration, we have found out nothing. We could also define force to mean that a moving object with no force acting on it continues to move with constant velocity in a straight line. If we then observe an object *not* moving in a straight line with a constant velocity, we might say that there is a force on it. Now such things certainly cannot be the content of physics, because they are definitions going in a circle. The Newtonian statement above, however, seems to be a most precise definition of force, and one that appeals to the mathematician; nevertheless, it is completely useless, because no prediction whatsoever can be made from a definition. One might sit in an armchair all day long and define words at will, but to find out what happens when two balls push against each other, or when a weight is hung on a spring, is another matter altogether, because the way the bodies behave is something completely outside any choice of definitions (Feynman, Leighton, and Sands 1966, p. 12-1).

For instance, the important thing about force is that it has a material origin. If a physical body is not present, then a force is taken to be zero. If we discover a force, then we also try to find something in the surroundings that is a source of the force. Another rule about force is that it obeys Newton's Third Law: the forces between interacting objects are equal in magnitude and opposite in direction (see Laws of Mechanics). These concepts we have about force, in addition to a mathematical rule such as *F = ma*, are the key elements in solving physics problems. It is only through combining the mathematical equations with these concepts that experts attempt to solve physics problems (Larkin 1983).

Consider, for example, the use of constraints to model physical laws. A constraint such as $f(x_1, \ldots x_n) = 0$ can be used to determine any x_i if all the others are given a value. However, as de Kleer (1975) points out:

A theory of commonsense reasoning about the physical world should be firmly based on physics.

All our ideas in physics require a certain amount of common sense in their application.

There is much more information in an expression! If all we are interested in is solving a set of equations, looking at constraint expressions may be a valid perspective. However, [if] we are solving a physical problem in which a duality exists between the mathematical structure of the equations and the actual physical situation we have thrown away most of the information. To the sophisticated student this duality is very clear and the mathematical equation is far more than a constraint expression. For him, the expression encodes a great deal of qualitative knowledge and every mathematical manipulation of the expressions reflects some feature of the physical situation (pp. 110-111).

For example, we see in NEWTON that whenever we calculate an imaginary number from a velocity equation, we decide the body under consideration is not able to reach a certain point. Similarly, we treat vector equations in a careful manner so that signs of the involved quantities make sense. If a force is trying to prevent the motion of a body, its direction is taken opposite the direction of a movement. The issue of what signs should be assigned to the quantities under consideration is the first sign of a correct attempt to solve a physics problem.

However, a more important kind of knowledge is encoded in a physics formula. Again, with $F = ma$, F and m should be placed during problem solving at the same conceptual layer because they can be independently determined from any other quantities (Shoham 1986). However, a is placed at a higher layer because it is determined by the quantities in the previous layer. Quantities such as F and m are only allowed to be manipulated externally (exogenous variables), whereas a is derived (endogenous). The only intended meaning of $F = ma$ is that at any given moment, the net force on a body is equal to the product of its mass and acceleration.

The notion of causality is of key importance in physics although this notion is somewhat against Bertrand Russell's famous advice (Shoham 1986): "The law of causality, I believe, like much that passes to muster among philosophers, is a relic of a bygone age, surviving, like the monarchy, only because it is erroneously supposed to do no harm" (p. 143). We tend to believe, as Shoham aptly stated, that people decide what variables are to be exogenous and then define causal rules correspondingly. The causal rules that are constructed depend on the theory we are dealing with. For example, in the case of $F = ma$, we can choose to write down two rules: one describing the influence of m on a and the other describing the influence of F on a. Consequently, we omit a causal rule describing the effect of a on F; this information is, nevertheless, present in the original formula.

Essential Attributes and Influences

Let us think of a simplified way of looking at physics problems. We usually have various objects under consideration and state several facts about them. We can use predicates to talk about objects, which follows the advice of Hayes (1985). Thus, $p(a)$ means that an object a has an attribute p. To denote that object a has no attribute p, we write $\bar{p}(a)$. Predicates can be one-, two-, . . . place depending on how many objects are required to form them. Thus, the predicate "x is frictionless" is one-place, "x is longer than y" is two-place, "z is faster than x but slower than y" is three-place, and so on. For the preceding examples, we write in predicate notation frictionless(x), longer(x,y), and inbetween(z,x,y). In general, $p(a_1, . . . a_n)$ means that an n-tuple of objects $a_1, . . . a_n$ has an attribute p, in other words, $a_1, . . . a_n$ such that p holds.

By abstraction, we signify the mapping of a set of objects to a new set of objects that have some of their attributes ignored. For instance, we can take a block and map it to a point mass by simply ignoring its spatial dimension. Similarly, we can map a surface to a frictionless surface, a gas to an ideal gas, and a fluid to an incompressible fluid. By doing so, we are creating predicates that are unverifiable; for example, frictionless(s), ideal(g), incompressible(f) denote attributes that are not possible to observe. The objects denoted by abstract terms do not empirically exist because of the way they are constructed (Zinoviev 1983); we decide to take certain objects and neglect some of their attributes.

Our view is that this abstraction process is how physics experts start solving problems. That is, the mapping of empirical objects to abstract objects is an important ingredient in solving a physics problem because the physical laws in general are stated over abstract entities and as state transformations. With the earlier attribute viewpoint, it is easy to write down transformations. Assume that s_1 is a state of an object at a certain moment, and s_2 is a state at a later moment. This transformation is denoted by $\rightarrow$. To denote that object a exists, we can write $E(a)$. Thus, $\bar{E}(a) \rightarrow E(a)$ denotes the creation of object a, whereas $E(a) \rightarrow \bar{E}(a)$ denotes the death of a. We can write similar statements about the
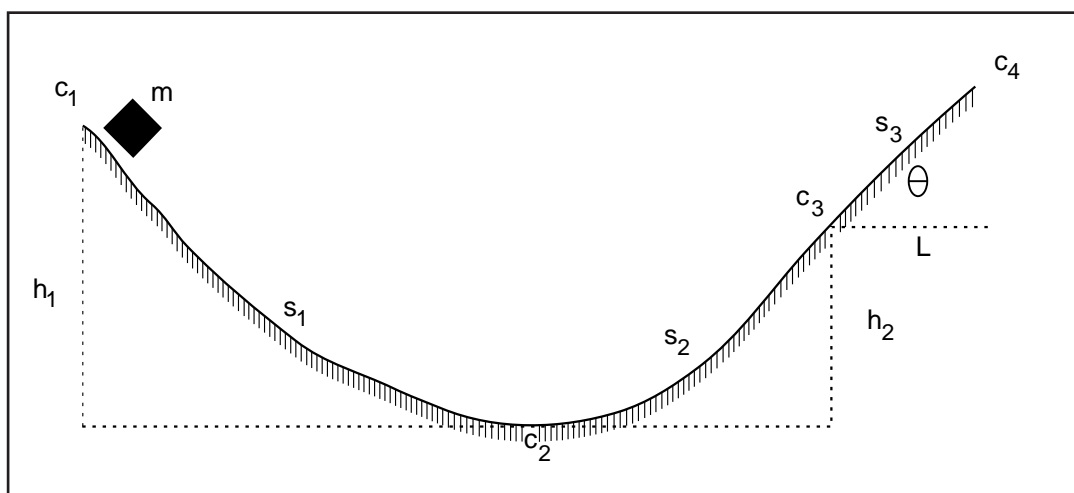
Figure 2. The Sliding Block Problem: Will the Block Reach c_4?

attributes. Thus, $\overline{p}(a) \rightarrow p(a)$ stands for the acquisition of an attribute, and $p(a) \rightarrow \overline{p}(a)$ stands for the loss of an attribute by object *a*. It is messy to tell exactly when an object is created or annihilated. If heat is given to a drop of water, the drop acquires the attribute hot and loses the attribute cold. However, if enough heat is applied, the drop turns to steam. Two possibilities exist: this is the same individual that was liquid before, or this is a new individual.

By *essential influences* we mean those influences which play a central role in a situation. Consider the example of a pendulum clock (Feynman, Leighton, and Sands 1966). If a pendulum clock is standing upright, it works as expected, but if it is tilted, nothing happens. Something else, something outside the machinery of the pendulum clock, is involved in the operation of the clock: the earth is pulling on the pendulum. Once we encode this effect through the vector of gravitational force, we can take the earth out of the picture and consider our pendulum embedded in a field of gravitation—the essential influence for this problem. Consequently, it is an easy matter to deduce that the pendulum will tick with a different rate on the moon.

Shape is a difficult attribute to deal with. With a rectangular block sliding along an inclined plane, one can use the abstraction that the block is a point mass. However, if a ball is rolling down on an inclined plane, a point mass viewpoint can lead to an incorrect perspective because, for example, the angular momentum as a result of rotation is no longer accountable. For the sliding block problem, we take the mass of the block as a constant because we can reasonably assume that it is

going at a low speed (compared to the speed of light), and its mass is independent of time. Similarly, for a block attached to a spring with stiffness constant *k*, we normally think that if the spring is stretched by *x*, then we can find the force on the spring using $F = -kx$. This relation is based on the assumption that for small *x*, *k* is constant. Note it is not implicit in this formula that the spring can break if the force is exceedingly large.

Another way of looking at physical influences is to regard them as functions. Thus, by *(fx)*, let us denote that function *f* takes object *x* and gives *fx* as the new object. This mapping is an intensional definition of a function as opposed to the usual, extensional definition. Functional equivalence then becomes *(fx) = (ghx)*; that is, apply *g* on the result of applying *h* to *x* to obtain a summary effect equivalent to *(fx)*. This notation helps one formulate state transformations. Consider object *x*, which is a block of ice. Now *(fx)*, where *f* is heat, renders a new object *fx*, which is equal to water. Let us call this object $\tilde{x}$. Applying *f* on $\tilde{x}$ gives a new object that is equal to steam. (Thus self-application of functions is sometimes meaningful.)

Envisioning

In envisioning, we deliberately ignore the values of the problem variables; we let them take any value, allowing us to discover all the possibilities for a given problem. Only one of these possible answers is observable for a given set of initial conditions. The idea of envisionment is clearer with a pioneering program called NEWTON.

Newton's use of quantitative knowledge to disambiguate ambiguities remains an important contribution.

NEWTON

The first attempt at predicting the interaction of moving objects in a simple, idealized world was made by Johan de Kleer (1975) with a program called NEWTON. NEWTON works in a two-dimensional world called the roller coaster. We start with a classical problem used by de Kleer. In figure 2, a small sliding block (idealized as a point) of mass m starts at point c_1 along a frictionless surface consisting of three parts: s_1 and s_2, which are concave, and s_3, which is straight. We want to know if the block will reach point c_4.

In NEWTON , a production-rule system looks at the local geometric (topological) features to predict what might happen next. The left-hand side of a production describes the features of the environment (for example, the shape of the paths, the velocity of the block), and the right-hand side lists the consequences (for example, sliding, falling). A closed-world assumption is in effect: The actions on the left-hand side never produce a change in the features (for example, a block sliding on a segment does not cause a change in the shape of the segment).[1]

We can summarize the reasoning suggested by de Kleer as follows: Without falling off or changing direction, m starts to slide down the surface. After reaching the bottom c_2, it starts going up. It still will not fall off, but it might start sliding back. If m ever reaches the straight section (that is, the segment s_3), it still will not fall off, but it might change the direction of its movement. To determine whether m reaches c_4, we must study its velocity. The velocity at the bottom can be computed using the conservation of energy: $v_{c_2} = \sqrt{2gh_1}$. Similarly, using v_{c2} and the conservation of energy, we can set up an equation that can be solved for the velocity v_{c3} at the start of the straight section:

$\frac{1}{2} mv_{c_3}^2 = \frac{1}{2} mv_{c_2}^2 - mgh_2$. If v_{c3} is imaginary, we know that the straight segment is never reached. At the straight section, we would use kinematics to find out whether the block ever reaches c_4. The acceleration of the block

along s_3 must be $a = g \sin\theta$. The length of the straight segment is $\frac{L}{\cos\theta}$; so, using the kinematic equation relating acceleration, distance, and velocities, we get $v_{c_4}^2 = v_{c_3}^2 - 2gLtan\theta$. Again, if v_{c4} is imaginary, c_4 is not reachable. (This argument is simpler and more elegant if you use the same law of energy conservation to explain the behavior of the block on the straight section; however, de Kleer uses the argument given here.)

This line of reasoning can be conveniently depicted in the envisionment tree in figure 3. The branches of the tree correspond to the different situations that can arise depending on the solutions of the equations mentioned earlier. Basically, we have a set of worlds W and a set of times T. There is a linear ordering < (meaning prior to) over T. We can then think of a two-dimensional space where we have differing worlds as we move along one axis and differing times as we move along the other axis. Any particular point in this space can be thought of as a pair of coordinates (w,t) for some $w \varepsilon W$ and $t \varepsilon T$. This index denotes a point whose location is determined by which world it is on the one hand and what time it is on the other hand. In left-to-right order, the possible worlds in figure 3 are as follows:

w_1: Slide until at most c_3, slide backwards, oscillate around c_2.

w_2: Slide until at most c_3, slide backwards, fall off from c_1.

w_3: Slide until after c_3 but before c_4, slide backwards, oscillate around c_2.

w_4: Slide until after c_3 but before c_4, slide backwards, fall off from c_1.

w_5: Slide until c_4, fall off from c_4.

For example, denoting the moments that m passes through the points c_1 by t_i, it is seen that m is at c_2 at t_2, at some point of s_2 for any t such that $t_2 < t < t_3$, and then for all $t > t_3$ on s_1 or s_2 if it follows the prediction denoted by w_1.

At each fork of the envisionment tree, several possibilities exist. A block would always take one of the branches and go down the tree ending up in a leaf. The forks correspond to the points where we have to disambiguate the ambiguities, whereas leaves correspond to particular behaviors. Disambiguation is achieved by solving some equations. Looking at the formula giving v_{c2}, we see that this velocity can never be imaginary. Thus, the block would always reach c_2; it would never stop there because at this point, it would have a positive velocity. Let us look at the

next equation giving the value of v_{c3}. After manipulation, it can be rewritten as $v_{c_3}^2 = 2g(h_1 - h_2)$. This equation has a simple message. If c_3 is located higher than c_1, m cannot reach c_3. Until now, we did not really need any numeric knowledge other than $h_1 > h_2$ to decide that m arrives at c_3. However, after c_3, we need to know the values of L and θ to discover that c_4 is reachable. Assume that m has barely made it to c_4, that is, $v_{c4} \geq 0$. Then the inequality $h_1 - h_2 \geq L \tan\theta$ should be satisfied.

Naive Versus Physical Representation

NEWTON has been a cornerstone in the search for knowledge representation and reasoning methodology for physical domains. A later work on mechanics problem solving in the style of NEWTON is MECHO (Bundy 1978). This program tries to improve the envisioning process by making it more goal directed. The main improvement is that MECHO does not generate the envisionment tree, only the parts needed to answer a question.[2] Thus, for a suitable set of values of the earlier problem parameters, MECHO would only generate the real path that would be taken by the block, say the leftmost.

Larkin (1983) points out that NEWTON has a naive way of representing physical knowledge. In particular, it has an internal representation that contains direct representations of the visible entities mentioned in the problem description. As a result, it performs simulation—NEWTON's inferencing by way of an envisionment tree follows the direction of time flow. However, the lack of deeper physics principles in NEWTON is more serious. For instance, Larkin (1983) notes that a physicist encountering the earlier problem about the sliding block might reason as follows (it should probably be said that the following protocol is hypothetical and does not originate from an interview with an expert physicist): The energy at c_1 consists of kinetic energy, zero because m is at rest, and potential energy determined by h_1. At c_2, the potential energy is zero because the block is at the bottom, and the kinetic energy is unknown because the speed is unknown. At c_3, the potential energy is determined by h_2, and the kinetic energy is still unknown. At some point c (which might be above or below c_4), the block momentarily stops, and the kinetic energy is again zero. Writing down the equation $0 + mgh_1 = 0 + mg(h_2 + X \sin\theta)$, where X is the distance m travels along s_3 after point c_3, immediately leads to a solution. (It is noted that in the last equation, inside the parentheses, Larkin writes $\frac{x}{sin\theta}$, which is wrong.)
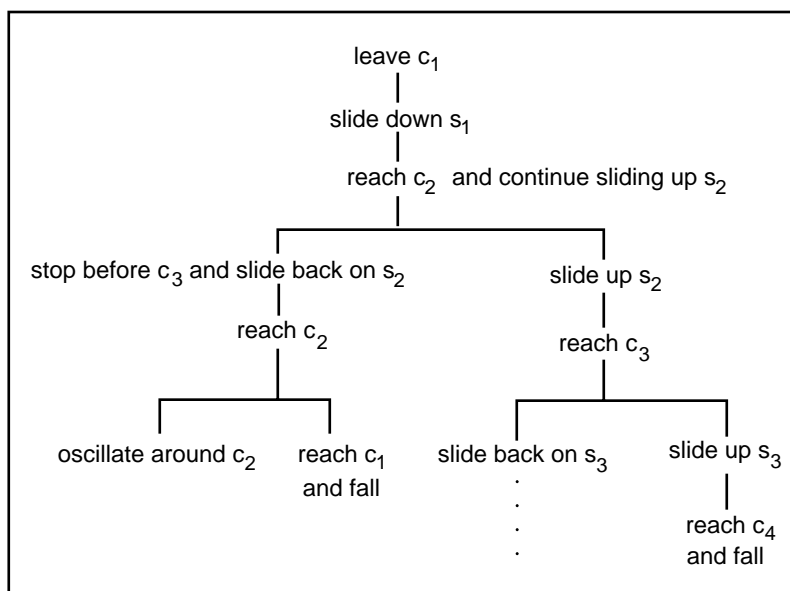

Figure 3. The Envisionment Tree for the Sliding Block Problem.

Briefly, if $X > \frac{L}{cos\theta}$, then m reaches c_4 and falls off.

Compared to NEWTON's solution, there are good insights in this solution. The fact that a body has speed is identified with its having kinetic energy. That a body is at a given height is identified with its having potential energy. In writing down the preceding equation, the absence of friction (and hence no loss of energy as a result of heat dissipation) is used to state a simpler law of energy conservation. The fact that potential energy and kinetic energy are convertible to each other is implicitly used. The expert is also aware that m, which is initialy at rest, is brought to motion by the gravitational field of the earth, which is constant and equal to g at distances not too far away from the surface of the earth. As another pointer to the use of deep knowledge, consider the decision of the expert to select c_2 as a standard point for potential energy calculation. If the expert had used any other point, say c_3 instead of c_2, it is obvious that the potential energy is changed only by the addition of a constant. Because the energy conservation law cares only about changes, it does not matter if a constant is added to the potential energy.

NEWTON's use of quantitative knowledge (represented as frames) to disambiguate ambiguities remains an important contribution. Frames in NEWTON are not procedures but describe dependencies among variables. They are similar to Minsky's (1975) frames in that they are used to chunk physical formulas of the same nature. Because many different

> *A major weakness of the current envisioners is their inability to switch between macroworlds and microworlds.*

equations apply to a problem, grouping equations in frames helps isolate the relevant ones with less effort. Dependencies among variables are then searched for a solution (the goal variable). For example, a kinematics frame[3] knows about the usual equations of motion (de Kleer 1977):

frame kinematics of *object, surface, t_1, t_2*
variables: (v_1: velocity of *object* at time t_1
                v_2: velocity of *object* at time t_2
                 d: distance of *surface*
                 t: time between t_1 and t_2
                 a: acceleration of *object*)
equations:
($v_2 = v_1 + at$, $v_2{}^2 = v_1{}^2 + 2ad$, $d = v_1 t + \frac{1}{2} at^2$)

Frames use two kinds of variables: the names of the objects and their essential attributes (such as velocity or acceleration). Each equation of a frame referencing the goal variable is a possible way to determine this variable. However, unknown variables in equations referencing the goal variable must be given a value (possibly using other frames or asking the user) because every unknown in the equation must be determined before achieving the goal.

The Content of Mechanics

This section outlines the building blocks of classical mechanics. It is neither complete nor uniform; however, it should give an idea about what kind of knowledge should be formalized for a deep coverage of mechanics. It is noted that we are not yet concerned with a concrete knowledge representation scheme. This decision is thought to be in agreement with Hayes (1985), where the building of minitheories (clusters) is given priority over favorite notations: "Initially, the formalizations need to be little more than carefully worded English sentences. One can make considerable progress on ontological issues, for example, without actually formalizing anything, just by being *very* careful what you say" (p. 483).

Fictitious Entities

We begin with a set of the basic fictitious entities of classical mechanics.

Description of Motion. Speed as a derivative   ($v = \lim_{Dt \text{æ} o} \frac{Ds}{Dt}$)   , distance as an integral ($s = \lim_{Dt \text{æ} o} Â , v(t_1) Dt)$, and acceleration as the derivative of speed are the basic notions. We distinguish velocity from speed, which is the magnitude of velocity. (We sometimes denote vectors with bold letters.)

Pseudovectors. Ordinary vectors are, for example, coordinate position, force, momentum, or velocity. Vectors that are obtained as a cross-product behave differently with respect to rotation, for example, torque τ, angular velocity ω, and angular momentum **L**. The corresponding equations are $\tau = \mathbf{r} \times \mathbf{F}$, $v = \omega \times r$, and $\mathbf{L} = \mathbf{r} \times \mathbf{p}$, respectively, where $\mathbf{p}$ is the momentum, and **r** is the radius.

Work. If a body is moving along a curved path, then the change in kinetic energy as it moves from one point (1) to another (2) is equal to \i(1,2,**F**.d**s**).

Torque. Torque bears the same relationship to rotation as force does to linear movement; that is, a torque is that which makes something rotate or twist.

Energy. In mechanics problems, energy basically takes the following forms: gravitational, kinetic, heat, and elastic.

Power. Power equals work done per second. In other words, the rate of change of kinetic energy of a body is equal to the power used by the forces acting on it.

Field. We need two kinds of laws for a field. The first law gives the response to a field, that is, the equations of motion. The second law gives how strong the field is, that is, the field equation. In other words, "One part says that something produces a field. The other part says that something is acted on by the field. By allowing us to look at the two parts independently, this separation of the analysis simplifies the calculation of a problem in many situations" (Feynman, Leighton, and Sands 1966, p. 12-8).

The idea of a field is closely associated with potential energy. We note that the gravitational force on a body is written as mass times a vector, which is dependent only on the position of the body: $\mathbf{F} = m\mathbf{C}$. One can thus analyze gravitation by imagining a vector **C** is at every position that affects any mass placed there. Because the potential energy can be written as $U = m\int(\mathbf{field}).d\mathbf{s}$, we find that the potential energy of a body in space can be written as mass times a function ψ, the potential. To get the force from the potential energy, we use $F_x = \frac{\partial U}{\partial x}$ (and simi-

larly for other directions). To get the field from the potential, we do the same: $C_x = -\frac{\partial y}{\partial x}$ (and similarly for other directions). Succinctly, $\boldsymbol{F} = -\nabla U$, and $\boldsymbol{C} = -\nabla\psi$.

Pseudoforce.   A well-known pseudoforce is centrifugal force. If we are in a rotating coordinate system, we experience a force throwing us outward. Using the pseudoforce, we can explain several interesting problems. In figure 4, adapted from Feynman, Leighton, and Sands (1966), a container of water is pushed along a table, with acceleration *a*. The gravitational force acts downward, but in addition, a pseudoforce is acting horizontally. The latter is in a direction opposite to *a*. As a result, the surface of the water is inclined at an angle with the table, as shown in the left part of the figure. If we stop pushing, the container slows down because of friction, and the pseudoforce reverses its direction, causing the water to stand higher in the forward side of the container. The normal to the surface is parallel to the vector sum of *a* and *g*, that is, to the net force on a surface element, which keeps each point on the water's surface at the same potential energy.

Centrifugal force is not the only force developed in a rotating system. Another force is called the Coriolis force. With this force, when we move a body in a rotating system, the body seems to be pushed sidewise. If we want to move something radially in a rotating system, we must also push it sidewise with force $\boldsymbol{F_c} = 2m\omega v_r$. Here ω is the angular velocity, and $\boldsymbol{v_r}$ is the speed at which the body is moving along the radius.

Conservative Force.   If the integral of the force times the distance in traveling from one point to another is the same (regardless of the shape of the curve connecting them), then the force is conservative (for example, gravity).

Center of Mass.   Given a rigid body, a certain point is such that the net external force produces an acceleration of this point, as if the whole mass is concentrated there. The point does not have to be in the material of the body; it can lie outside.

Laws of Mechanics

The second set of building blocks comprise the laws of mechanics as follows.

Newton's First Law (the principle of inertia).   If an object is left alone (not disturbed), it continues to move with a constant velocity in a straight line if it was originally moving, or it continues to stand still if it was standing still.

Newton's Second Law.   The motion of an object is changed by forces in this way: The
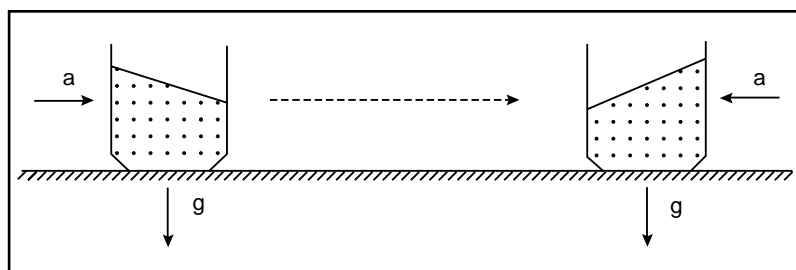

Figure 4. Pseudoforce Acting on a Container of Water.

time rate of change of momentum is proportional to the force.

Newton's Third Law (the principle of action equal reaction).   Suppose we have two small bodies, and suppose that the first one exerts a force on the second one, pushing it with a certain force. Then simultaneously, the second body pushes on the first with an equal force in the opposite direction. These forces effectively act in the same line.

Conservation of Linear Momentum.   If there is a force between two bodies, and we calculate the sum of the two momenta, both before and after the force acts, the results should be equal.

Conservation of Angular Momentum.   If no external torques act on a system of particles, the angular momentum remains constant. Also, the external torque on a system is the rate of change of the total angular momentum: $\tau_{ext} = \dfrac{d\boldsymbol{L}_{tot}}{dt}$

Conservation of Energy.   An abstract quantity exists that does not change in all the natural phenomena which the world undergoes: energy.

(Galilean) Relativity Principle.   The laws of physics look the same whether we are standing still or moving with a uniform speed in a straight line.

Work Done by a Force.   If the force is in one direction, and the object on which the force is applied is displaced in a certain direction, then only the component of force in the direction of the displacement does any work; that is, physical work is expressed as $\int\boldsymbol{F}\boldsymbol{\cdot}\boldsymbol{ds}$.

Work Done by Gravity.   The work done in going around a path in a gravitational field is zero, implying that we cannot make perpetual motion in a gravitational field.

Hooke's Law (the law of elasticity).   The force in a body that tries to restore the body to its original condition when it is disturbed is proportional to the distortion. This law holds true if the distortion is small. If it gets too large, the body is torn apart or crushed.
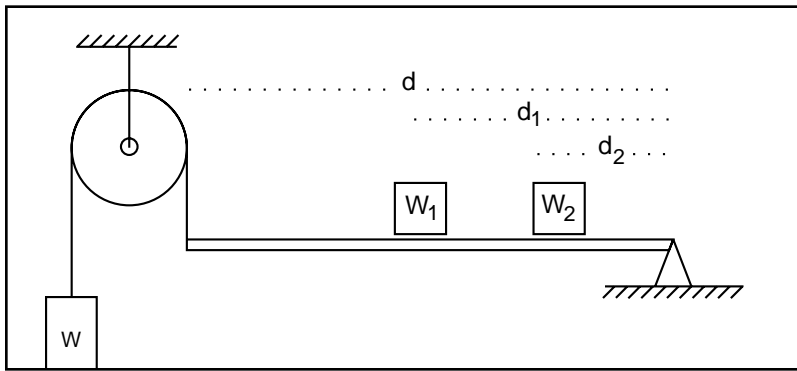
Figure 5. A System of Blocks in Equilibrium.

Principles

Although they are classified under different headings in this article, it should be noted that the boundary between physics laws and physics principles is not well defined. We suggest that by principles, we should refer to the problem-solving elements, for example, the principle of virtual work (see below). In any case, this section, in its current form, makes no claim to completeness.

Superposition. The total field resulting from all the sources is the sum of the fields resulting from each source. Suppose that we have a force F_1 and have solved for the forced motion. Then we find out another force F_2 exists and solve for the other forced motion. Using the superposition of solutions, we can now predict what would happen if we had F_1 and F_2 acting together. The solution is $x_1 + x_2$ if x_i's are the individual solutions for forces F_i. In general, a complicated force can be divided into a set of separate forces, each of which is simple (in the sense that we can solve for the forced motion they cause).

Equivalence of Simple Harmonic Motion and Uniform Circular Motion. Uniform motion of a body in a circle is closely related to oscillatory up-and-down motion. Although the distance y means nothing in the oscillator case, it can still be artificially given to model the oscillation in terms of circular motion.

Virtual Work. We imagine a structure moves a little—even though it is not really moving or even movable. We use this small imagined move to apply the law of energy conservation. This principle is especially useful in problems of the sort depicted in figure 5, where we are asked to find the value of weight W such that the system is in equilibrium. Noting that a small move of W toward the bottom should be counteracted by weights W_1 and W_2, we find

$$hW \quad = \quad \frac{d_1}{d} hW_1 \quad + \quad \frac{d_2}{d} hW_2 \quad , \qquad or$$

$$W = \frac{w_1 d_1 + w_2 d_2}{d} \ .$$

Analogs. We have already seen examples of analogs when we talked about force versus torque, linear momentum versus angular momentum, and so on, in Fictitious Entities. In general, *analogy* refers to relating two domains that are at first sight dissimilar and using the tools of one domain to solve the problems of the other. The analogy between mechanics and electricity is well known, as table 1 partially illustrates (Feynman, Leighton, and Sands 1966). As a result, not only in problem solving but also in design, analogy has found its deserved place. The following excerpt is especially illuminating in this respect:

> Suppose we have designed an automobile, and want to know how much it is going to shake when it goes over a certain kind of bumpy road. We build an electrical circuit with inductances to represent the inertia of the wheels, spring constants as capacitances to represent the springs of the wheels, and resistors to represent the shock absorbers, and so on for parts of the automobile. Then we need a bumpy road. All right, we apply a *voltage* from a generator, which represents such and such a kind of bump, and then look at how the left wheel jiggles by measuring the charge on some capacitor. Having measured it (it is easy to do), we find that it is bumping too much. Do we need more shock absorber, or less shock absorber? With a complicated thing like the automobile, do we actually change the shock absorber, and solve it all over again? No!, we simply turn a dial; dial number ten is shock absorber number three, so we put in more shock absorber. The bumps are worse—all right, we try less. The bumps are still worse; we change the stiffness of the spring (dial 17), and we adjust all these things *electrically*, with merely a turn of the knob (Feynman, Leighton, and Sands 1966, p. 25-8).

Some Examples

In this section, we give some example problems and their solutions. Our aim is to demonstrate the use of physical representations. We regard the following problems as difficult problems for envisioners and expect that they will constitute part of a test set for future programs.

General Characteristic	Mechanical Property	Electrical Property
independent variable	time *(t)*	time *(t)*
dependent variable	position *(x)*	charge *(q)*
inertia	mass *(m)*	inductance *(L)*
force	force *(F=ma)*	voltage *(V)*
velocity	velocity *(v)*	current *(I)*
resistance	drag coefficient *(c=γm)*	resistance *(R=γL)*
stiffness	stiffness *(k)*	(capacitance)$^{-1}$ *(1/C)*
period	$t_0=2\pi \sqrt{m/k}$	$t_0=2\pi \sqrt{LC}$

Table 1. Analogs.

Rocket Problems

These problems are taken from Feynman, Leighton, and Sands (1966). How fast do we have to send a rocket away from the earth so that it leaves the earth? The problem can be stated as a functional requirement: The rocket must leave the earth. We are now asked to find an attribute (the speed) of the rocket. Notice that many other attributes have been left out, and a pair of objects—the earth and the rocket—have been identified. The essential influence is the earth's gravitational field.

The kinetic energy plus the potential energy of the rocket must be a constant. Let us exaggerate and imagine the rocket in two extreme positions. When it is far away from the earth, it has zero potential energy. Its kinetic energy is also zero because we can assume that it barely left the earth. However, initially it has the total energy

$\frac{1}{2}mv^2 - G\frac{mM}{R}$ where *m* is the rocket's mass,

M is the earth's mass, and *R* is the earth's radius. The conservation of energy gives $v=\sqrt{2gR}$, where $g=\frac{GM}{R^2}$.

With a small change to the problem, at what speed should a satellite travel to keep going around the earth? It turns out that energy conservation is not the best approach to this problem. A force approach written as an equality between centrifugal and gravitational forces, $\frac{mv^2}{R}=\frac{GMm}{R^2}$, gives $v=\sqrt{gR}$. The reason we thought that a force expression is more convenient here results from the nature of the problem; there is a rotating object, and this object guides our search among the mathematical expressions applicable to the problem. This line of reasoning should be compared with de Kleer's

(1975) search strategy.

Now let us take a look at the following problem. A rocket of large mass *M* ejects a small piece of mass *m* with a velocity *V* relative to the rocket. With the assumption it was standing still, the rocket now gains a velocity *v*. With the law of momentum conservation, this velocity is seen to be $v=\frac{m}{M}V$. Thus, rocket propulsion is essentially the same as the recoil of a gun. It does not need air to push against (Feynman, Leighton, and Sands 1966).

Let us suppose that the two objects are exactly the same, and then we have a little explosion between them. After the explosion, one is moving, say, toward the right, with a velocity *v*. Then it appears reasonable that the other body is moving toward the left with a velocity *v* because if the objects are alike, no reason exists for right or left to be preferred; so, the bodies would behave symmetrically.

The Ball with Strings

Consider figure 6, which is taken from den Hartog (1961). A heavy ball of weight *W* is suspended by a thin thread and has an identical thread hanging from it. When we start pulling down on the lower string, which of the two strings will break first?
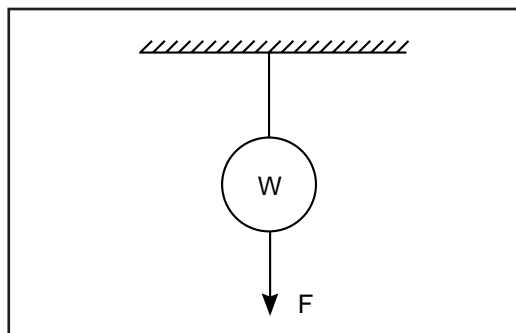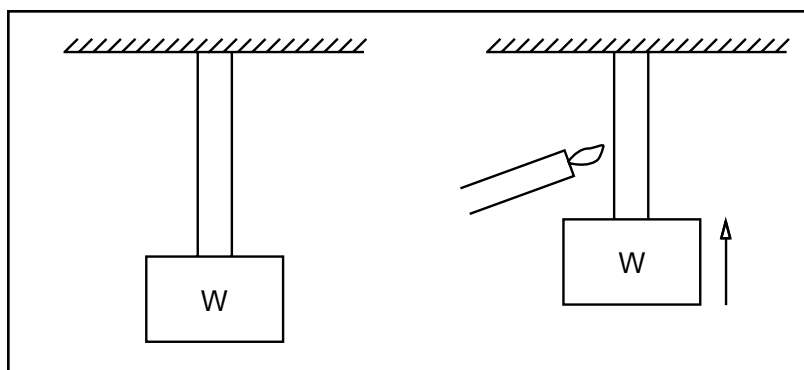

Figure 6. The Ball with Strings.

Figure 7. The Heated Rubber Band.

We assume that we apply force *F*. The bottom string experiences *F*, but the upper string experiences *W* + *F*. Thus, the upper string will break first. According to den Hartog, this effect happens only if we pull down slowly. If we instead give a sudden, sharp pull to the lower string, it breaks, and the ball remains suspended. This effect has to do with the fact that the threads are elastic and have a certain elongation associated with the force sustained by them. By giving a quick pull to the lower thread, the force in the lower thread can be made quite large, and this force accelerates the ball downward. However, the movement of the ball takes some time, and before any appreciable downward displacement is observed, the string is broken.

Den Hartog relates a similar experiment. Consider a ball, with a single string attached to it laying on a table. By a slow pull, one can drag the ball on the table with a uniform speed. In this case, the applied force is equal to the friction force between the ball and the table. A quick pull, however, breaks the thread in an instant. For a short time, the ball is subjected to a large force that subjects it to acceleration. However, the ball hardly moves because the time interval is small. Instead, its velocity is destroyed by the retarding action of the friction force.

The Open Water Jar

A major weakness of the current envisioners is their inability to switch between macroworlds and microworlds. In other words, the individuals that an envisioner knows about are either in the world as we see it or are underlying the world (atomic processes). (In fact, Schmolze's [1987] thesis is the only attempt, albeit a rather restricted one, to deal with atomic processes.) To our best knowledge, no current envisioner can

predict, for example, that water would slowly evaporate if it is left in an open jar at room temperature. This weakness results from the fact that no change is immediately observable at a macrolevel, the realm of discourse of the envisioner.

All things are made of atoms, particles that move around in perpetual motion, attracting or repelling each other depending on the distance between them. Physicists think that our knowledge of atoms gives us useful information (Feynman, Leighton, and Sands 1966). The atomic hypothesis describes processes well. For example, let us observe the jar with some water in it. What happens as time passes? The water molecules are constantly moving around. From time to time, one on the surface is hit strong enough so that it flies out. Thus, molecule by molecule, the water evaporates.

If we cover the jar with a lid, we observe no change because the number of molecules leaving the surface is equal to the number of those coming back. Thus, in the long run, nothing happens. Note, however, that we should find a large number of water molecules among the air molecules. If we take the lid away and replace the moist air with dry air, water again evaporates. The number of molecules leaving the surface is still the same, but not as many are coming back. Interestingly, relating the frequency with which the atoms collide to the temperature, one can also understand such phenomena as why evaporation resumes in a covered jar if you heat the water.

The concept of equilibrium is an important one for envisionment and one that provides an excellent macrodescription of a microconcept. Water molecules are constantly leaving the fluid and entering the air and vice versa. If the rates of entry and exit are equal, then the air is saturated, and there is no evaporation (putting a lid on the jar causes the air to saturate after some time interval and stops the evaporation process).

The Heated Rubber Band

This problem is taken from Feynman, Leighton, and Sands (1966) and is another good example of the convenience of thinking in terms of atomic processes. If we apply a gas flame to a rubber band holding a weight, the band contracts abruptly, as shown in figure 7. How can we explain this phenomenon?

A molecular explanation would be as follows: A rubber band consists of a tangle of long chains of molecules, not unlike a molecular spaghetti with cross-links. When such a tangle is pulled out, some of the chains line

De Kleer and Brown's principle of no function in structure requires that the laws of the parts of the device cannot presume the functioning of the entire device.

up along the direction of pull. At the same time, chains are hitting each other continually because of thermal motion. Thus, if a chain is stretched, it does not remain stretched by itself; it would be hit from the sides by other chains that would tend to unstretch it again. When we heat the rubber band, we increase the amount of bombardment on the sides of the chains.

With thermodynamics, this explanation can be made quantitative. When heat ΔQ is delivered, the internal energy is changed by ΔU, and some work is done. The work done by the rubber band is $-F\Delta L$, where F is the force on the band, and L is the length of the band. Note that F can be considered a function of temperature T and length L, just as gas pressure can be expressed as a function of T and volume V. From the first law of thermodynamics, we get $\Delta Q = \Delta U - F\Delta L$. The term $-F\Delta L$, the work done by the rubber band, is analogous to $P\Delta V$, the work done by a gas. Just as the gas expands when heated and does work on its surroundings, so does the rubber band work and lift the weight attached to it. The second law applies as well; from it, one can calculate how much the force exerted by the rubber band increases (with increasing temperature) in terms of the heat needed to keep the temperature constant when the band is stretched a little bit (Feynman, Leighton, and Sands 1966). Feynman extends the analogy: "If one can write $\Delta U = \Delta Q + A\Delta B$ where A and B represent different quantities, force and length, pressure and volume, etc., one can apply results obtained for a gas by substituting A and B for P and V" (Feynman, Leighton, and Sands 1966, p.45-4). (He then illustrates this point by considering a battery, where A = voltage and B = charge.)

Other Related Research

A discussion of naive physics and qualitative physics is given in Schmolze's (1987) dissertation. Here, we offer a shorter discussion for completeness and refer the reader to Schmolze for an analysis of these areas. (N.B. We omit a discussion of Forbus's [1984] qualitative process theory, nonwithstanding its importance). The articles by de Kleer and

Brown (1984), Kuipers (1984, 1986), Forbus (1984), Forbus and Gentner (1986), and Hayes (1985) form the basis of naive and qualitative physics. Incidentally, it is not easy to delineate the areas covered by these terms. We propose that naive physics should be understood as the construction of knowledge representation methods, and qualitative physics should cover reasoning techniques.

With the term naive physics, Patrick Hayes (1985) proposed the construction of a formal, sizable portion of commonsense knowledge about the physical world. This corpus should include knowledge about objects, shape, space, movement, substances, time, and so on. As for the knowledge representation language to be used, Hayes is not specific; a collection of assertions in logic, for example, might be sufficient (see Essential Attributes and Influences). At this preliminary stage, he is not interested in the efficiency of reasoning with this body of knowledge. What he is really after is "the extent to which it [a naive physics theory] provides a vocabulary of tokens which allows a wide range of intuitive concepts to be expressed, to which it then supports conclusions mirroring those which we find correct or reasonable" (Hayes 1985, p. 470).

Qualitative physics provides an account of behavior in the physical world. Unlike conventional physics, qualitative physics predicts and explains behavior in qualitative terms. Although the behavior of a physical system can be given by the precise values of its variables (temperatures, velocities, forces, and so on) at every moment, such a description fails to provide insights into how the system works. Important concepts causing change in physical systems are concepts such as momentum, force, and feedback that can be intuitively understood (de Kleer and Brown 1984). In conventional physics, they are embedded in a framework of continuous differential equations. In qualitative physics, each measurable property such as the speed of a ball is represented in two parts: a quantity and its rate of change. The representation is qualitative because the quantity values are selected from a discrete quantity space. For example, the quantity space for water can only have two values: the freezing and the
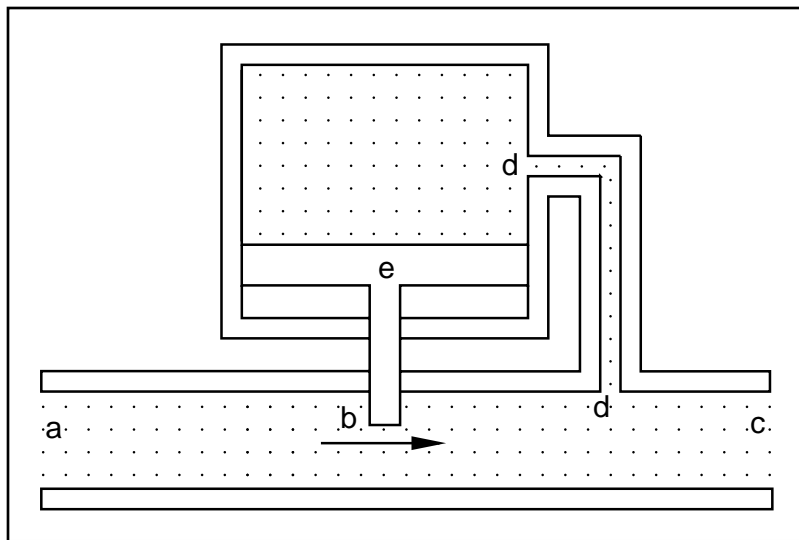
Figure 8. A Pressure Regulator.

boiling temperatures. In general, rates of change are limited to three cases: increasing, decreasing, and constant.

ENVISION

De Kleer and Brown (1984) introduced qualitative differential equations (confluences) and implemented them in a program called ENVISION. To obtain confluences, we let continuous variables take discrete values from a quantity space, such as {*0,+,–*}. Let [*x*] denote the qualitative value of an expression *x* with respect to the quantity space. Then the proposition "*x* is increasing" is written as $\frac{dx}{dt}$ = +, and we can define arithmetic to deal with [*x*] + [*y*] or [*x*][*y*], although cases exist where ambiguities have to be resolved using numeric values (for example, when the operation is addition and [*x*] = + and [*y*] = –). Let ∂x denote $\left[\frac{dx}{dt}\right]$. Using confluences, we can

explain as follows: Consider the pressure regulator in figure 8, which is adapted from de Kleer and Brown (1984). A confluence such as $\partial P + \partial A - \partial Q = 0$, where *P* is the pressure across the valve, *A* is the area available for flow, and *Q* is the flow throughout the valve, describes this device in qualitative terms. An increase in pressure at *a* is seen to cause an increase in pressure at point *b*, which generates more flow through *b*. As a result, pressure at *c* increases and is felt at *d*. Then the diaphragm *e* presses downward, causing the valve to close somewhat. As a result, constant pressure is maintained at *c*, although the pressure at *a* is fluctuating. Because a single confluence cannot characterize the behavior

of a component over its entire range of operation, the range is divided into subregions, each characterized by a different component state where different confluences apply. When the valve is closed, the correct confluence should read $\partial Q = 0$; we simply do not say anything about *P*. Similarly, when the valve is open, the confluence becomes $\partial P = 0$.

ENVISION has a component library where the components relevant to the reasoning domain are stored. Another module holds the topology of the device under consideration. The input to ENVISION is the input signals and boundary conditions. The output is a behavioral prediction along with a so-called causal explanation. De Kleer and Brown's principle of no function in structure requires that the laws of the parts of the device cannot presume the functioning of the entire device. Various classwide assumptions are made to avoid tricky situations. For example, in fluid flow, enough particles always exist in a pipe so that macroscopic laws hold; the mean free path of the particles is small compared to the distances over which the pressure appreciably changes; dimensions of an electric circuit are small compared to the wavelength associated with the highest frequency; and so on.

QSIM and CA

Qualitative simulation can be defined as the derivation of a behavioral description of a mechanism from a qualitative description of its structure. Kuipers (1986) believes that causality can be taken as identical to value propagation with constraints. Hence, Kuipers (1984) has a constraint-centered ontology. (However, de Kleer and Brown [1984] and Forbus [1984] have device- and process-centered ontologies, respectively.) Kuipers' qualitative simulation framework is a symbolic system that solves a set of constraints obtained from differential equations. His QSIM algorithm is guaranteed to produce a qualitative behavior corresponding to any solution to the original equation. He also shows that in some cases, a qualitative description of structure is consistent with an intractably large number of behavioral predictions. A couple of techniques representing different trade-offs between generality and power have been proposed for taming this intractable branching (Kuipers and Chiu 1987).

As an example of how QSIM can be used, consider the problem of throwing an object vertically into the air from some height. This problem is first described to QSIM using a problem-description language. In addition to the initial values of the problem variables, the description includes the physical con-

straints (lawlike knowledge) acting on the ball: $\frac{ds}{dt} = v$, $\frac{dv}{dt} = a$, and $a = g < 0$, where s, v, and a stand for displacement, velocity, and acceleration, respectively. The user enters these constraints into the system. (In the future, it might be possible to extract the relevant portion of a physics knowledge base to use it in a problem.[4]) From this initial information, QSIM reasons in an intuitive way and finds that the ball rises to a particular height, stops, and then falls to the earth again. The landmark point where the velocity becomes zero is discovered by QSIM. Landmark values are important in identifying the different regions of behavior for a mechanism.

In our view, the main problem with QSIM lies in its essence. Obviously, QSIM has no knowledge of physics whatsoever. In this example, QSIM cannot possibly know that the object might leave the earth to orbit around it or even go to infinity if it is given enough initial speed (see Rocket Problems). QSIM assumes for all cases that a is constant; it does not know that the value g for a is applicable only near the surface of the earth and that a decreases inversely with the square of distance. Thus, by submitting a problem to QSIM, it is assumed that all relevant knowledge is given; any physics law left unspecified cannot be used. It is only fair to note that this condemnation of QSIM is not total. The simple yet mathematically elegant framework of QSIM has much to offer. We think that one can teach QSIM the necessary domain-specific knowledge (in our case, classical mechanics but in another case, for example, economics laws) and, thus, employ it in a more useful manner. It remains to be seen whether QSIM's problem-description style is sufficient enough to codify laws such as energy conservation. (The answer is probably no.)

An interesting extension of qualitative simulation is Weld's (1987) comparative analysis which was implemented in a program called CA. CA deals with the problem of predicting how a system reacts to perturbations in its parameters and why. For example, comparative analysis explains why a ball goes up higher if it is thrown with a greater speed. Weld also studies exaggeration. Consider a question such as "What happens to the period of oscillation of a block attached to a spring on a frictionless table as the mass of the block is increased?" Exaggeration suggests that if the mass were infinite, then the block would hardly move, and thus, the period would be infinite. Therefore, had the mass increased a little, the period would increase as well. (For another example of exaggeration, see Rocket Problems).

Summary

Contrary to their public image, expert systems are admittedly not true experts in their fields (McCarthy 1987). This observation has contributed to the emergence of dichotomies in AI such as deep versus shallow knowledge. Although problems exist with using these adjectives, it is understandable that an expert system is not a deep model of its domain of expertise. An expert system's if-then rules can only capture the superficial characteristics of a domain. However, a real expert has deeper knowledge about his or her domain of expertise. Several criteria have been suggested for true expertise; we only give a representative list.

First, experts can explain their line of reasoning in a logical way. Their explanations differ markedly from those generated by current expert systems. When asked about how a certain result was obtained, an expert gives information related both to the real world and its abstract models, not just a list of the rules used.

Second, experts use multiple models of a domain to classify and solve problems in an efficient manner. They solve difficult problems using difficult methods. However, for easy problems, they either reduce the complicated methods to simpler versions or use already available simple methods. Furthermore, experts can predict. Before they delve into a problem, they have some general idea about what the solution should look like.

Third, experts can discover the inconsistencies in ill-defined problems. They can judge and eliminate irrelevant or contradictory information. Presented with incomplete problem statements, they make reasonable (default) assumptions. As a result, they can work in a nonmonotonic mode, occasionally revising their beliefs during the problem-solving process.

In this article, we argue that an expert theory of envisioning should be based mainly on physics knowledge. This theory satisfies the previous criteria. The fictitious entities of physics such as energy, work, and force make up the elements that the envisioner reasons about. Abstract principles such as superposition and laws such as Newton's are used as the essential tools for reasoning. Simplifying techniques such as essential attributes and influences are advocated as preliminary aids to problem solving. Our work is in the precise spirit of Larkin's (1983) study in the expert-novice difference in problem-solving performance and is based on her distinction between naive versus physical representations.

Kuipers' QSIM algorithm is guaranteed to produce a qualitative behavior corresponding to any solution to the original (differential) equation.

Many problems remain to be solved. Central among them is choosing a concrete knowledge representation method. We think frames are appropriate for this purpose (Minsky 1975). Another problem is to define a core subset of physics that can be used effectively for envisioning in different areas. Until now, we have mainly studied the domain of mechanics, following the AI tradition. Yet another project is to study the limits of reasoning without recourse to lawlike knowledge. To invert a remark of Larkin (1983), "Why are people good at predicting the outcome of physical interactions in the world around them and so bad at physics, even the branch of physics (mechanics) that deals with interaction of everyday objects?" Are there other, fundamentally different ways of looking at the physical world? For example, John McCarthy (1987) writes that a commonsense knowledge "database would contain what a robot would need to know about the effects of moving objects around, what a person can be expected to know about his family, and the facts about buying and selling" (p. 1030) in addition to other information. Here, the problem lies in integrating these diverse (and, obviously, not all physical) domains of knowledge using a base language, a problem McCarthy calls generality in AI. McCarthy (1987) adds, "This does not depend on whether the knowledge is expressed in a logical language or in some other formalism" (p. 1030). We agree and note that although these other perspectives might indeed be necessary, they are outside the domain of this discourse and do not diminish the need for representing physics knowledge in an intelligent reasoner. ∎

Acknowledgments

References

Akman, V.; ten Hagen, P.; Rogier, J.; and Veerkamp, P. 1988. Knowledge Engineering in Design. *Knowledge-Based Systems* 1(2): 67-77.

Akman, V.; ten Hagen, P.; and Tomiyama T. 1987. Design as a Formal Knowledge Engineered Activity, Technical Report, CS-R8744, Center for Mathematics and Computer Science.

Bundy, A. 1978. Will It Reach the Top? Prediction in the Mechanics World. *Artificial Intelligence* 10:129-146.

Davis, E. 1986. A Logical Framework for Solid Object Physics, Technical Report, 245, Dept. of Computer Science, Courant Institute of Mathematical Sciences, New York University.

de Kleer, J. 1977. Multiple Representations of Knowledge in a Mechanics Problem Solver. In Proceedings of the Fifth International Joint Conference on Artificial Intelligence, 299-304. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

de Kleer, J. 1975. Qualitative and Quantitative Knowledge in Classical Mechanics, Technical Report, No. AI-TR-352, Artificial Intelligence Laboratory, Massachusetts Institute of Technology.

de Kleer, J., and Brown, J. S. 1984. A Qualitative Physics Based on Confluences. *Artificial Intelligence* 24:7-83.

den Hartog, J. P. 1961. *Mechanics.* New York: Dover.

diSessa, A. A. 1982. Unlearning Aristotelian Physics: A Study of Knowledge-Based Learning. *Cognitive Science* 6:37-75.

Feynman, R.; Leighton, R. B.; and Sands, M. 1966. *The Feynman Lectures on Physics, Volume I: Mainly Mechanics, Radiation, and Heat.* Reading, Mass.: Addison-Wesley.

Forbus, K. D. 1984. Qualitative Process Theory. *Artificial Intelligence* 24:85-168.

Forbus, K. D., and Gentner, D. 1986. Learning Physical Domains: Toward a Theoretical Framework, Technical Report, UIUCDCS-R-86-1247, Dept. of Computer Science, Univ. of Illinois at Urbana-Champaign.

Hayes, P. J. 1985. The Second Naive Physics Manifesto. In *Readings in Knowledge Representation,* ed. R. J. Brachman and H. J. Levesque, 467-485. San Mateo, Calif.: Morgan Kaufmann.

Kant, I. 1984. *Critique of Pure Reason.* London: Everyman's Library.

Kuipers, B. 1986. Qualitative Simulation. *Artificial Intelligence* 29:289-338.

Kuipers, B. 1984. Commonsense Reasoning about Causality: Deriving Behavior from Structure. *Artificial Intelligence* 24:169-203.

Kuipers, B., and Chiu, C. 1987. Taming Intractable Branching in Qualitative Simulation. In Proceedings of the Tenth International Joint Conference on Artificial Intelligence, 1079-1085. Menlo Park, Calif.: International Joint Conferences on Artificial Intelligence.

Larkin, J. H. 1983. The Role of Problem Representation in Physics. In *Mental Models,* ed. D. Gentner and A. L. Stevens, 75-97. Hillsdale, N.J.: Lawrence Erlbaum.

Leibniz, G. W. 1984. *Philosophical Writings.* London: Everyman's Library.

McCarthy, J. 1987. Generality in Artificial Intelligence. *Communications of the ACM* 30(12): 1030-1035.

McCloskey, M. 1983. Naive Theories of Motion. In *Mental Models,* ed. D. Gentner and A. L. Stevens, 299-323. Hillsdale, N.J.: Lawrence Erlbaum.

Minsky, M. 1975. A Framework for Representing Knowledge. In *The Psychology of Computer Vision,* ed. P. Winston, 211-277. New York: McGraw-Hill.

Schmolze, J. G. 1987. Physics for Robots, Technical Report, 6222, BBN Laboratories.

Shoham, Y. 1986. Reasoning about Change: Time and Causation from the Standpoint of Artificial Intelligence. Ph.D. diss., Dept. of Computer Science, Yale Univ.

Veth, B. 1987. An Integrated Data Description Language for Coding Design Knowledge. In *Intelligent CAD Systems I: Theoretical and Methodological Aspects,* ed. P. ten Hagen and T. Tomiyama, 295-313. Heidelberg: Springer-Verlag.

Weld, D. S. 1987. Comparative Analysis, Technical Report, AIM-951, Artificial Intelligence Laboratory, Massachusetts Inst. of Technology.

Zinoviev, A. A. 1983. *Logical Physics.* Dordrecht, Holland: Reidel.

Notes

1. In a related work, Davis (1986) gives the following justification for the closed-world assumption: "The 'frame' or 'persistence' problem of determining what remains true over time requires no special treatment in our logic. We can avoid this problem, not by virtue of any special cleverness on our part, but by virtue of the structure of the domain. The predicates in the domain are divided into two classes. The first class includes predicates which depend on position and velocity of objects. These are not assumed to remain constant over any interval unless they can be proven to be so. The second class includes structural predicates, depending only on the shape and other material properties of the objects. These are defined to be constant over the problem, and so are defined atemporally. Similar considerations would seem to apply to any closed world, complete physical theory; it is not clear why the frame problem should ever create trouble in such a domain."

2. It is arguable that an envisioner should generate only the relevant parts of an envisionment tree. Goal-directed envisioning is fine as long as questions such as "What happens next?" are avoided.

3. De Kleer published a revised, shorter version of de Kleer (1975) in de Kleer (1977). An interesting decision was made in the second version to call the restricted-access local consequent methods (RALCMs) of the first version FRAMEs.

4. An important issue then is the composition of the mechanism under consideration: A system's behavior should be deducible from its structure (for example, components and their connections), as mentioned in the case of envision. Analogs of this principle are used in various domains. In linguistics, it is assumed that the semantic value of any expression is a function of the semantic values of its syntactic constituents. Thus, the semantic rules compute the semantic values of increasingly longer parts of a statement. Admittedly, this view is sometimes dangerously mechanistic.

Varol Akman is an assistant professor in the Department of Computer Engineering and Information Science, Bilkent University, Ankara, Turkey. His current research is concentrated in various theoretical areas of AI, especially from the angle of intelligent computer-aided design systems, for example, common-sense reasoning, naive physics, knowledge representation, and mathematical logic.

Akman received a bachelor of science in electrical engineering in 1979 and a master of science in computer engineering in 1980 from the Middle East Technical University. He received his Ph.D. in computer systems and engineering from Rensselaer Polytechnic Institute in New York.

Akman is an editorial board member of the journal *Computers and Graphics* and the *Bilkent University Lecture Series in Engineering,* and he is on the program committee for the Eurographics Workshops on Intelligent CAD.

Paul J. W. ten Hagen is head of the Department of Interactive Systems, Center for Mathematics and Computer Science, Amsterdam, The Netherlands. As a member of IFIP WG 5.2, he has participated and contributed to several conferences on computer-aided design (CAD) and is co-chairman of a series of Eurographics workshops on intelligent CAD systems. IFIP WG 5.2 is currently involved in knowledge engineering for CAD and aspects of pictorial databases. Both of these topics are within the scope of ten Hagen's research plans for the near future.

Ten Hagen is cofounder of the European Association for Computer Graphics and the Computer Graphics Forum. He was also a convener of ISO/TC97/SC5/WG2 on Graphics Standardization. This group produced the international standard for graphics, GKS.