Chapter 7

Physics

I know people who would not deposit a nickel and a dime in a cigarette-vending machine and push the lever even if a diamond necklace came out. I know dozens who would not climb into an aeroplane even if it didn't move off the ground. In none of these people have I discerned what I would call a neurosis, an "exaggerated" fear; I have discerned only a natural caution in a world made up of gadgets that whir and whine and whiz and shriek and sometimes explode.

James Thurber, "Sex ex Machina," in *Let Your Mind Alone!*

To act effectively and flexibly, to take advantage of opportunities and to avoid dangers, an intelligent creature must understand the behavior of the physical world. In particular, it must understand how its own actions will affect the world.

The kind of knowledge required for sensible behavior can be quite different from scientific theories. In general, it is not necessary nor even useful to incorporate the most complete theories of modern physics into a robotics program. On the one hand, these theories deal with phenomena far outside the scope of ordinary experience. On the other hand, though it is probably true that any valid statement about the commonsense physical world is, in principle, a consequence of these underlying theories, deriving a useful commonsense inference from fundamental physical laws, or even stating boundary conditions for a commonsense problem in terms of fundamental physical properties, is wholly impractical. A commonsense physical reasoning system should deal with concepts more or less at the level of everyday discourse.

The theories that we will study in this chapter are all grounded in scientific theories; they are approximations to scientific truth. The physical theories that can be elicited from the man on the street are, apparently, substantially different. (See, for example [McCloskey 1983].) We have chosen to ignore these, first, because correct theories are presumably more useful, and, second, because correct theories are more uniform, better known, more easily specified, and less prone to internal inconsistency. Also, since the man on the street can, generally, accept physical reality without continual surprise, the predictions of the true physical laws must be largely compatible with his beliefs in most situations. (There are exceptions, such as gyroscopes, even among simple physical situations.)

In their daily lives, human beings deal with myriads of different types of physical substances and interactions. AI theories of physical reasoning do not yet begin to reflect this range of phenomena; so far, they have studied only a rather small number of different areas. Often the choice of an area for study has been based on a perceived potential for practical application, rather than on centrality or interest for commonsense reasoning. This chapter, necessarily, follows the existing research in its choice of topics.

7.1 The Component Model

A divide-and-conquer strategy that is often useful in analyzing physical systems, particularly man-made devices, is to view the system as a whole as composed of separate components connected together. The behavior of the system can then be analyzed by studying the behaviors of the components, each of which is presumably simpler than the overall system, and determining how these behaviors interact. This kind of analysis is easiest and most effective if the assemblage has the following properties:

- The device is assembled out of a set of components, which are connected together. What components are used, and how they are connected together, is invariant over time.
- The instantaneous state of the assemblage can be characterized by the values of a number of one-dimensional *parameters*. (We rule out devices where two- or three-dimensional motion is important, except where each dimension of motion can be handled independently.)
- Each component has a fixed number of *ports*. Components interact only by being connected at ports.

- Each parameter is associated with one port of one component. The behavior of a component is entirely characterized in terms of constraints that it imposes on values of parameters at its various ports. The behavior of a connection is entirely characterized in terms of constraints that it imposes on the ports that it joins.

The best examples of such systems are electronic devices, which, however, are at best marginally objects of a commonsense understanding. The model may also be applied, less well, to other types of devices such as hydraulic systems, heat-transfer systems, and simple mechanical devices. Our description of the component model is based on the well-known ENVISION program [de Kleer and Brown 1985].

An important objective in analyzing such systems is the principle of "No function in structure" [de Kleer and Brown 1985]. That is, the component and connection descriptions, which constitute the input description of the system (structure), should be given in a form that is independent of properties of the overall system. This objective is rarely fully achievable in a component analysis, except for electronic devices, but it can be partially achieved if the specified descriptions apply across a large range of devices that use the component. For example, a description of a switch that specifies that current flows through a switch just if the switch is closed would violate "No function in structure" badly, since there are many closed switches with no current flow, such as a switch in a circuit with no power source, or a closed switch in series with an open switch. A description that specifies that a switch prohibits current flow if it is open and prohibits voltage difference if it is closed would be valid for most standard uses of a switch, and so would observe the "no function in structure" principle. The account will not serve for nonstandard uses of a switch — e.g., as a paperweight, or to cast a shadow, or to create an electric arc by placing it almost closed. No description of an object in terms of constraints among ports will cover all of its possible physical behaviors.

As an example of component-based analysis, consider the simple scale shown in Figure 7.1. In this scale, the height of the needle varies with the mass in the pan. The relation is controlled by the two springs. For the purposes of this example, we will assume that, within the operating range of the scales, the slope of the lever is small enough that the horizontal displacement of the ends can be ignored.

We divide the scales into five components: A weight with two ports, one for the gravitational force and one for the support; a lever with two ports; a base with one port; and two springs, each with two ports. Each port has two parameters associated: its height and the vertical force exerted on the component at the port. (In a substantial violation

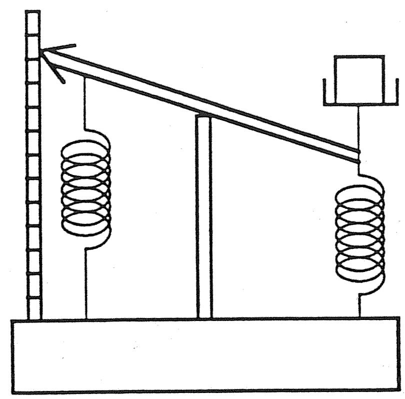


Figure 7.1 Scale

of "no function in structure," our models of springs and lever presume that the springs are vertical and that the lever is nearly horizontal.) The component models are as follows: The port of the base is at a constant height. The force on the two ends of the lever are equal. The heights of the two ends of the lever are constrained by the fact that their midpoint is a fixed fulcrum of constant height, so that the sum of their heights is constant. The forces exerted on the spring at each of its ports are equal in magnitude and opposite in direction, and the difference in height between the two ports is an increasing function of the difference in the forces. (The sign is not reversed, because we are considering the force exerted on the spring rather than the force exerted by the spring.) The weight obeys Newton's laws: The total force on the weight is equal to its mass times its vertical acceleration. The total force on the weight is equal to the gravitational force, which is considered one port, plus the force exerted on the weight by its support, which is considered another port.

The following rules apply to connections:

- All ports at a connection must have equal heights.
- The sum of the forces exerted on all the ports at a connection must be zero.

The system is at rest (in equilibrium) when all parameter values are constant. In particular, the acceleration of the weight must be zero.

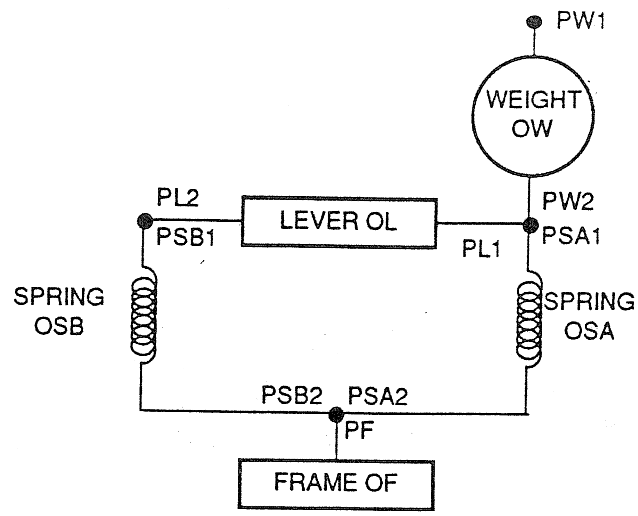


Figure 7.2 Schematic of scale

Figure 7.2 shows an abstract component diagram of the scale. Table 7.1 shows the primitives used in representing the scale. Table 7.2 gives an axiomatization of the physical laws involved. Table 7.3 shows a specific description of the scale. Table 7.4 shows the equations that govern the scale; these can be derived straightforwardly from the axioms and problem statement.

If we restrict attention to situations where the system is at rest, so that all derivatives are zero, then the last equation in Table 7.4 becomes, "$f_w = mg$."

There are a number of ways that these equations can be used for physical inference:

Exact solution: If the properties of the springs are stated exactly, rather than merely specifying that the expansion of the spring is an increasing function of the force applied, then the equations can be solved exactly. For example, suppose it is specified that the springs are linear; that is, they are governed by the rule

$$\text{spring}(O, P1, P2) \Rightarrow$$
$$(1/2) \times (\text{force\_on}(P1) - \text{force\_on}(P2)) =$$
$$\text{spring\_const}(O) \times (\text{height}(P1) - \text{height}(P2) - \text{rest\_length}(O))$$

Table 7.1 Primitives for Scales

Sorts: Ports, components, heights, forces, situations, and fluents.

Atemporal:

grav_acc	—	Constant. Acceleration of gravity.
mass(O)	—	Function. Mass of weight O.
base_height(O)	—	Function. Height of frame O.
center_height(O)	—	Function. Central height of lever O.
frame(O, P)	—	Predicate. Object O is a frame with port P.
weight($O, P1, P2$)	—	Predicate. Object O is a weight with ports $P1$ and $P2$.
lever($O, P1, P2$)	—	Predicate. Object O is a lever with ports $P1$ and $P2$.
spring($O, P1, P2$)	—	Predicate. Object O is a spring with ports $P1$ and $P2$.
connection($P1, P2 \ldots Pk$)	—	Predicate. Ports $P1 \ldots Pk$ are connected.

Fluents:

height(P)	—	height of port P.
force_on(P)	—	force exerted on port P.

Let k_a = spring_const(osa); k_b = spring_const(osb); l_a = rest_length(osa); and l_b = rest_length(osb). The above constraint gives us the equations

$$\tfrac{1}{2}(\mathrm{f}_{sa1} - \mathrm{f}_{sa2}) = \mathrm{k}_a \cdot (\mathrm{x}_w - \mathrm{x}_f - \mathrm{l}_a)$$

$$\tfrac{1}{2}(\mathrm{f}_{sb1} - \mathrm{f}_{sb2}) = \mathrm{k}_b \cdot (\mathrm{x}_b - \mathrm{x}_f - \mathrm{l}_b)$$

Solving the above equations algebraically, together with those of Table 7.4, we can derive that the height of the arrow obeys the equation

$$\ddot{\mathrm{x}}_b = \mathrm{g} - \frac{1}{\mathrm{m}}((\mathrm{k}_a + \mathrm{k}_b)\mathrm{x}_b - (\mathrm{k}_a + \mathrm{k}_b)\mathrm{h}_f + \mathrm{k}_a\mathrm{l}_a - \mathrm{k}_b\mathrm{l}_b - 2\mathrm{k}_a\mathrm{h}_{lf})$$

If the system is at rest, the arrow is at height

$$\mathrm{x}_b = \mathrm{x}_{rest} = \mathrm{h}_f + \frac{1}{\mathrm{k}_a + \mathrm{k}_b}(mg + \mathrm{k}_b\mathrm{l}_b + 2\mathrm{k}_a\mathrm{h}_{lf} - \mathrm{k}_a\mathrm{l}_a)$$

Otherwise, the arrow executes a motion of form

$$\mathrm{x}_b(t) = \mathrm{x}_{rest} + \mathrm{a}\ \sin(\omega t + \mathrm{t}_0)$$

Table 7.2 Axioms for Scales

Component Rules

SC1. weight($O, P1, P2$) $\Rightarrow$
value_in(S,force_on($P1$)) = mass(O) $\times$ grav_acc.

SC2. weight($O, P1, P2$) $\Rightarrow$
mass(O) $\times$ deriv(deriv(height($P2$))) =
force_on($P1$) + force_on($P2$).

SC3. spring($O, P1, P2$) $\Rightarrow$
monotonic(height($P1$) − height($P2$),
force_on($P1$) − force_on($P2$), pos).

SC4. spring($O, P1, P2$) $\Rightarrow$ force_on($P1$) = −force_on($P2$).

SC5. frame(O, P) $\Rightarrow$ value_in(S,height(P)) = base_height(O).

SC6. lever($O, P1, P2$) $\Rightarrow$
value_in(S,height($P1$)) − center_height(O) =
center_height(O) − value_in(S,height($P2$)).

SC7. lever($O, P1, P2$) $\Rightarrow$ force_on($P1$) = force_on($P2$).
(The lever has arms of equal length.)

Connection Rules

SC8. connection($P1 \ldots Pk$) $\Rightarrow$ height($P1$) = ... = height(Pk).

SC9. connection($P1 \ldots Pk$) $\Rightarrow$ force_on($P1$) + ...+ force_on(Pk)=
0.

Table 7.3 Problem Description of Example Scale

weight(ow,pw1,pw2).
lever(ol,pl1,pl2).
frame(of,pf).
spring(osa,psa1,psa2).
spring(osb,psb1,psb2).
connection(pw2,pl1,psa1).
connection(pl2,psb1).
connection(psa2,psb2,pf).

Table 7.4 Equations of Scales

Define the following constants and fluents:
Constants: (See Figure 7.3)

m = mass(ow).
g = grav_acc.
h_l = center_height(ol).
h_f = base_height(of).
$h_{lf} = h_l - h_f$. (Height of the lever above the frame.)

Fluents:

x_w = height(pw2) = height(psa1) = height(pl1).
x_f = height(pf) = height(psa2) = height(psb2).
x_b = height(pl2) = height(psb2).
f_w = force_on(pw2).
f_{sa1} = force_on(psa1).
f_{sa2} = force_on(psa2).
f_{sb1} = force_on(psb1).
f_{sb2} = force_on(psb2).
f_l = force_on(pl1) = force_on(pl2).
f_f = force_on(pf).

Then we have the following relations:

$f_l + f_{sa1} + f_w = 0$.
$f_l + f_{sb1} = 0$.
$f_{sa2} + f_{sb2} + f_f = 0$.
$f_{sa1} = -f_{sa2}$.
$f_{sb1} = -f_{sb2}$.
$x_f = h_f$.
$x_b - h_l = h_l - x_w$.
monotonic($x_w - x_f, f_{sa1} - f_{sa2}$, pos).
monotonic($x_b - x_f, f_{sb1} - f_{sb2}$, pos).
$m\ddot{x}_w = f_w - mg$.

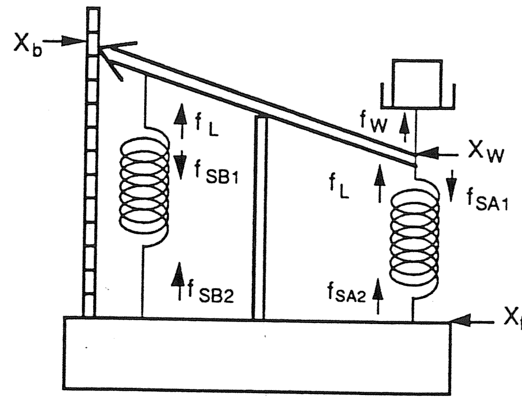


Figure 7.3 Constants of scale

where $\omega = \sqrt{(k_a + k_b)/m}$ and "a" and "t0" depend on the initial values of motion.

Perturbation of equilibrium: If the exact properties of the springs are not known, or if they obey a complex nonlinear equation, then it may be impossible or impractical to find closed-form solutions like those above, or even to find numerical values. However, it is still possible to extract various types of qualitative information. One type of qualitative inference is to calculate how the equilibrium state changes with a change to the constant parameters of the problem. For example, we can calculate how the rest position of the arrow is affected by a change in the mass. Suppose that the mass of the weight increases, and that all the other parameters of the system — the spring constants and lengths, the heights of the base and the lever, and the gravitational field — remain constant. Assuming the equilibrium state, where $\ddot{x}_w = 0$, and applying axioms SGN.1–SGN.3 of Table 4.6 governing the Δ operation, we obtain the following relations:

$$\Delta f_l + \Delta f_{sa1} + \Delta f_w \sim 0.$$
$$\Delta f_l + \Delta f_{sb1} \sim 0.$$
$$\Delta f_{sa2} + \Delta f_{sb2} + \Delta f_f \sim 0.$$
$$\Delta f_{sa1} \sim -\Delta f_{sa2}.$$

$\Delta f_{sb1} \sim -\Delta f_{sb2}$.
$\Delta x_f \sim 0$.
$\Delta x_b + \Delta x_w \sim 0$.
$\Delta x_w - \Delta x_f \sim \Delta f_{sa1} - \Delta f_{sa2}$.
$\Delta x_b - \Delta x_f \sim \Delta f_{sb1} - \Delta f_{sb2}$.
$0 \sim \Delta f_w - \Delta m$

The last equation above relies on the facts that the gravitational acceleration constant, grav_acc, is positive. Given these equations, it is easy to show that, if Δm is positive, then Δx_b is likewise positive. (Exercise 1)

Qualitative dynamic equations: It is also possible to use these partial constraints to derive a qualitative description of the dynamic behavior of the system, for fixed values of the constant parameters. Differentiating and applying sign operations to the equations in Table 7.4, we obtain the following qualitative equations:

$\text{sign}(f_l) + \text{sign}(f_{sa1}) + \text{sign}(f_w) \sim 0$.
$\text{sign}(f_l) + \text{sign}(f_{sb1}) \sim 0$.
$\text{sign}(f_{sa2}) + \text{sign}(f_{sb2}) + \text{sign}(f_f) \sim 0$.
$\text{sign}(f_{sa1}) \sim -\text{sign}(f_{sa2})$.
$\text{sign}(f_{sb1}) \sim -\text{sign}(f_{sb2})$.
$\partial x_f \sim 0$.
$\partial x_b + \partial x_w \sim 0$.
$\partial x_w - \partial x_f \sim \partial f_{sa1} - \partial f_{sa2}$.
$\partial x_b - \partial x_f \sim \partial f_{sb1} - \partial f_{sb2}$.
$\partial x_w \sim \text{sign}(f_w)$.

We may use algebraic techniques to reduce the above relations to the following:

$\partial^2 x_b \sim -\text{sign}(f_w)$.
$\partial f_w \sim \partial x_b$.

These equations thus have the same form as those analyzed in Section 4.9. Using the techniques discussed there, we can show that the height of the arrow follows the state transition illustrated in Figure 4.11; that is, it oscillates, regardless of the exact properties of the springs involved. (We have discussed some limitations on this conclusion in Section 4.9)

The ENVISION program carries out qualitative analyses of equilibrium perturbation and of dynamic behavior; it does not find exact quantitative solutions.

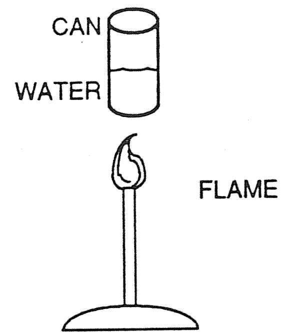


Figure 7.4 Boiling water in a can

7.2 Qualitative-Process Theory

As we have seen, the component model analyzes a complex physical system in terms of the questions "What are the pieces of the system?," "How does each piece constrain the parameters associated with it?," and "How are the pieces connected?" An alternative method, called *qualitative process* (QP) theory [Forbus 1985], focuses instead on the questions "What processes take place in the system?," "How do these processes influence system parameters?," and "How do the processes interact?" Consider, for example, a closed can of water suspended over a lit Bunsen burner (Figure 7.4). We would like to determine that heat flows from the burner through the can to the water, and that the water first becomes hotter until reaching its boiling point, then it boils and turns to vapor, then it continues to get hotter until reaching the temperature of the flame (or bursting the can). To analyze this system in terms of components constraining parameters at ports would be unnatural. Also, by focusing exclusively on the relations between parameters, the component model would entirely avoid facts such as the boiling of the water, which seem to be central to a commonsense understanding of the system behavior.

The central concept in Forbus's analysis[1] is that of a *process*. An individual process is a state token of a particular type. The process is said to be *active* during the time interval in which it takes place. The significant properties of a process are its *preconditions* and its *influences*. The preconditions of a process are states that must hold

[1] We simplify Forbus's theory and depart from his terminology in some respects.

if the process is to be active. (In some circumstances, it is useful to distinguish between the initiating conditions, which must hold for the process to start, and the sustaining conditions, which must hold for the process to continue. For example, a fire requires an external spark as an initiating condition, but not as a sustaining condition.) The influences of a process are the effects it has on parameters.

For example, consider the process of a heat flow from object A to object B. A sufficient precondition for heat flow is that A and B are thermally connected, and the temperature of A is greater than the temperature of B. Necessary preconditions are that A and B are thermally connected and that the temperature of A is greater than or equal to the temperature of B. (Heat can flow from one object to another of equal temperature if there is an external heat flow into the first. Consider, for example, flow between the two levels of a double boiler when both levels are boiling.) The influences of the heat flow are to reduce the heat of A and to increase the heat of B.

Besides characterizing processes, a QP theory must describe the connections between parameters. Parameters are divided into two types. Parameters of the first type are directly influenced by processes. The time derivative of such a parameter is equal to the sum of the influences of all processes on the parameter. In our example, the parameter "heat of A" is directly influenced by the heat flows into and out of A. Parameters of the second type are affected directly by other parameters; they are affected by processes only indirectly, through other parameters. In our example, the parameter "temperature of A" is directly affected by the parameter "heat of A" and only indirectly by the heat-flow process. QP theory expresses relations between parameters using statements of qualitative proportionality. Parameter $R1$ is qualitatively proportional to parameter $R2$, written $R1 \propto_{Q+} R2$, if an increase in $R2$ will cause an increase in $R1$, other things being equal. $R1$ is negatively qualitatively proportional to $R2$, written $R1 \propto_{Q-} R2$, if an increase in $R2$ will cause a decrease in $R1$, other things being equal. In our example, QP theory would specify that temperature(A) is qualitatively proportional to heat(A).

(Neither of the ideas in the previous paragraph — distinguishing between directly and indirectly affected parameters, and the definition of qualitative proportionality — have sound foundations in actual physics. Consider a piston containing gas. In such a system it is possible to control either the volume, by fixing the position of the piston, or the pressure, by putting a weight on the piston, or some function of the volume and pressure, by attaching a spring to the piston. Similarly, one can control either the temperature, by bringing it into thermal equilibrium with a heat reservoir of fixed temperature, or the heat, by

Table 7.5 Nonlogical symbols for Qualitative-Process Theory

Sorts: Situations (S), processes (P), process types (A), parameters (Q). Formally, processes and process types are just special cases of state tokens and state types. A parameter is a fluent into a quantity space.

Nonlogical symbols:

active(S, P)	—	Predicate. Process P is active in situation S. Equivalent to "$S \in$ time_of(P)."
process(P, A)	—	Predicate. P is a process of type A. Equivalent to "token_of(P, A)."
influence(P, Q)	—	Function. Influence of process P on parameter Q. A fluent onto the differential space, "Units of Q per unit time."
$P1 \propto_{Q+} P2$	—	Predicate (at least syntactically). $P1$ is qualitatively proportional to $P2$.
$P1 \propto_{Q-} P2$	—	Predicate. $P1$ is negatively qualitatively proportional to $P2$.

insulating the system. Thus, which parameters are directly affected and which are indirectly affected depends on circumstances. The relations between parameters similarly depends on circumstances. For example, if the pressure is held constant, then temperature and volume increase together, while if heat is held constant (adiabatic expansion), then temperature increases as volume decreases. Asking whether the temperature is an increasing or decreasing function of volume is somewhat like asking whether the area of a rectangle is an increasing or decreasing function of the length of its longer side. If the length of the shorter side is held constant, then the area is increasing; if the perimeter of the rectangle is held constant, then the area is decreasing.)

We now have all the representational equipment to express the example of the water in the can. Table 7.5 shows nonlogical symbols and axioms for QP theory. Tables 7.6 and 7.7 give the axioms neeeded for the particular example. Table 7.8 gives the particular problem statement.

Table 7.6 Nonlogical Symbols for Heat Flow Example

Sorts: Situations (S), processes (P), process types, states, fluents, temperature, heat, mass, objects (O). Temperature, heat, and mass are parameters. Objects, for the purpose of this example, are rather artificial constructs; they are collections of stuff, of uniform composition and uniform temperature, which at any instant may be gaseous, liquid, or solid, depending on the temperature.

Atemporal properties:		
boiling_point(O)	—	Function. Boiling temperature of object O.
freezing_point(O)	—	Function. Freezing temperature of object O.
mass(O)	—	Function. Mass of object O.
Parameters:		
temperature(O)	—	Function. Fluent of temperature of object O.
heat(O)	—	Function. Fluent of heat of object O.
solid_mass(O)	—	Function. Fluent of the mass of the solid part of O.
liquid_mass(O)	—	Function. Fluent of the mass of the liquid part of O.
gas_mass(O)	—	Function. Fluent of the mass of the gaseous part of O.
State type:		
thermally_connected($O1, O2$).	—	Function. State of $O1$ being thermally connected to $O2$.
Processes types:		
heat_flow($O1, O2$)	—	Function. Process type of a heat flow from $O1$ to $O2$.
boiling(O)	—	Function. Process type of object O boiling.

Table 7.7 Axioms for Heat-Flow Example

Process Definition for Heat Flow

HF1. [true_in(S,thermally_connected(OS, OD)) $\wedge$
value_in(S,temperature(OS)) >
value_in(S,temperature(OD))] $\Rightarrow$
$\exists_P$ process(P,heat_flow(OS, OD)) $\wedge$ active(S, P).
(Sufficient precondition for heat flow: If source OS is thermally connected to destination OD and OS is hotter than OD, then heat will flow from OS to OD.)

HF2. [process(P,heat_flow(OS, OD)) $\wedge$
active(S, P)] $\Rightarrow$
[$OS \neq OD$ $\wedge$ true_in(S,thermally_connected(OS, OD)) $\wedge$
value_in(S,temperature(OS)) $\geq$
value_in(S,temperature(OD))].
(Necessary preconditions for heat flow: For heat to flow directly from OS to OD, they must be thermally connected and OS must be at least as hot as OD.)

HF3. [token_of(P,heat_flow(OS, OD)) $\wedge$ active(S, P)] $\Rightarrow$
[value_in(S,influence(P,heat(OS))) < 0 $\wedge$
value_in(S,influence(P,heat(OD))) > 0]
(Influences of heat flow.)

Process Definition for Boiling

HF4. [value_in(S,liquid_mass(OB)) > 0 $\wedge$
value_in(S,temperature(OB)) = boiling_point(OB) $\wedge$
process($P2$,heat_flow(OS, OB)) $\wedge$ active($S, P2$)]$\Leftrightarrow$
$\exists_P$ process(P,boiling(OB)) $\wedge$ active(S, P).
(Preconditions for boiling: An object OB will boil if it is partially liquid and is at its boiling point and there is heat flow into OB.)

HF5. [process(P,boiling(OB)) $\wedge$ active(S, P)] $\Rightarrow$
[influence(P,liquid_mass(OB)) < 0 $\wedge$
influence(P,gas_mass(OB)) > 0.]
(Influences of boiling: It reduces the liquid mass of OB and increases its gaseous mass.)

Qualitative Proportionality

HF6. temperature(O) $\propto_{Q+}$ heat(O)

Table 7.7: Axioms for Heat-Flow Example (Continued)

State Coherence Axioms

HF7. value_in(S,solid_mass(O)) $> 0 \Rightarrow$
value_in(S,temperature(O)) $\leq$ freezing_point(O).
(An object can exist in solid form only below its freezing point.)

HF8. value_in(S,liquid_mass(O)) $> 0 \Rightarrow$
freezing_point(O) $\leq$ value_in(S,temperature(O)) $\leq$ boiling_point(O).
(An object can exist in liquid form only between its freezing and boiling points.)

HF9. value_in(S,gas_mass(O)) $> 0 \Rightarrow$
boiling_point(O) $\leq$ value_in(S,temperature(O)).
(An object can exist in gaseous form only above its boiling point.)

HF10. value_in(S,solid_mass(O)) + value_in(S,liquid_mass(O)) +
value_in(S,gas_mass(O)) = mass(O) > 0.
(The sum of the solid, liquid, and gaseous parts of an object is equal to the total mass of the substance, which is positive and time-invariant.)

HF11. value_in(S,solid_mass(O)) $\geq 0 \wedge$
value_in(S,liquid_mass(O)) $\geq 0 \wedge$
value_in(S,gas_mass(O)) ≥ 0.
(No masses are less than zero.)

Table 7.8 Problem Specification for Heat-Flow Example

Constants: oflame, owater, s0, i0.

infinite_on_right(i0) $\wedge$ start(i0) = s0.
(i0 is an infinite interval beginning with s0.)

$\forall_{S \in \mathrm{i0}}$ true_in(S, thermally_connected(oflame,owater)).
(The water is always thermally connected to the flame.)

$\forall_{S \in \mathrm{i0}}$ value_in(S,temperature(oflame)) > boiling_point(owater).
(The flame is always above the boiling point of water.)

freezing_point(owater) < value_in(s0, temperature(owater)) <
boiling_point(owater).

(The water temperature starts between its freezing point and its boiling point.)

Given process characterizations and qualitative proportionalities, as in Table 7.7, and the boundary conditions of a physical situation, as in Table 7.8, it is possible to predict the behavior of the system. This involves the following algorithm:

Algorithm 7.1: Prediction in QP Theory

1. Determine what processes have their preconditions satisfied and are therefore active. If no processes are active, then halt.
2. Determine the influences of these processes.
3. Determine the direction of change of the parameters. First, determine the change in each directly influenced parameter by summing the influences of the active processes. Then determine the change in indirectly influenced parameters by using the qualitative proportionalities in topological sort order. If the direction of change is ambiguous, then consider all combinations of possibilities disjunctively.
4. Extrapolate the changes to the parameters to predict which quantity conditions will change first, causing the beginning or end of a process. If it is ambiguous which change occurs first, then con-

sider all possibilities disjunctively. If no such change can occur, then halt.

5. Compute the state of the world at the transition point computed in (4). Go to 1.

Algorithm 7.1 above is very similar to the techniques discussed in Section 4.9 for solving qualitative differential equations. However, it involves three types of nonmonotonic inference. First, a closed-world assumption is applied to the predicate "active(S, P)"; that is, we assume that the only processes that are active in a given situation S are those that we have proven. Second, the use of a qualitative proportionality $P1 \propto_Q P2$ in step (3) assumes that any parameter relevant to $P1$ whose change has not been calculated is constant. Third, the algorithm makes the assumption, discussed in Section 4.9, that if a quantity-valued fluent approaches a value, then it will eventually attain the value.

Table 7.9 shows how this algorithm behaves on the heat-flow example.

7.3 Rigid Solid Objects

Perhaps the most common area of commonsense physical reasoning is the interactions of solid objects. Solid objects are involved in most terrestrial natural phenomena and in nearly every man-made artifact. They are familiar to every human from infancy and fairly well understood by childhood. So natural and familiar is solid-object behavior that it was long believed to be the fundamental type of physical behavior; hence, the efforts of physicists from the Greek atomists through the beginning of this century to explain physical phenomena of all sorts in mechanistic terms.

In reasoning about solid objects, it is often possible to assume that the class of solid objects and the shape of each solid object remains fixed. That is, the solid objects involved are not created, destroyed, broken, bent, or worn down. If so, we can define a time-invariant function, "shape(O)," mapping an object O to the region that it occupies in some standard position. The fluent "place(O)" is the region occupied by O in each situation; since O is rigid, place(O) is always congruent to shape(O). The fluent "position(O)" is the rigid mapping that maps shape(O) to place(O).

Table 7.9 Prediction in QP Theory

First iteration:

1. Active process: From HF1, infer that there is a heat flow ph1 from oflame to owater active in s0.
2. From HF3, infer that ph1 has a positive influence on the heat of owater and a negative influence on the heat of oflame.
3. Applying the closed-world assumption on influences, infer that the heat of owater increases and the heat of oflame decreases. From HF6, infer that the temperature of owater increases and the temperature of oflame decreases. (This is not true, of course, of an actual flame, but it will not affect the remainder of the prediction. A richer theory of flames could block this nonmonotonic inference (see Exercise 2). The inference would be valid if we immersed the can in a high-temperature bath, rather than putting it over the flame.)
4. By extrapolating the changes, we can see two possible changes in the quantity conditions. First, the temperature of the flame and of the water could become equal, which might bring an end to the heat flow. Second, the temperature of the water might reach its boiling point, which would cause a boiling process to begin. However, due to the boundary condition that the temperature of the flame is always above the boiling point of water, we can determine that the second transition always takes place before the first. Hence, we can define a new situation s1, in which the temperature of owater is at its boiling point.
5. In s1, the temperature of owater is equal to the boiling point of water. All other relations remain the same as in owater.

Table 7.9: Prediction in QP Theory (Continued)

Second iteration:

1. Active processes: From HF1, infer that the heat flow ph1 from oflame to owater continues. (See Exercise 3 for a discussion of this inference.) From HF4, infer that there is a boiling process pb1 active in s1.

2. From HF3, the heat flow ph1 has a positive influence on the heat of owater and a negative influence on the heat of oflame. From HF5, the boiling has a positive influence on the gaseous mass of owater and a negative influence on its liquid mass.

3. Applying the closed-world assumption on influences, we infer that the heat of owater increases, the heat of oflame decreases, the liquid mass of owater decreases, and the gaseous mass decreases. Applying HF6, we infer that the temperature of oflame decreases. We would also infer that the temperature of owater increases, except that we can show that it cannot, since, by HF8, its temperature cannot exceeed its boiling point unless its liquid mass is zero. Therefore, we conclude that, due to the change from liquid to gas, HF6 does not apply. (To make the blocking of this inference more direct, it may be advisable to add the qualitative proportionality "temperature(O) $\propto_{Q-}$ gaseous_mass(O)." Note that this rule will only be invoked during a boiling (or condensing) process, since those are the only times that the gaseous mass changes.)

4. By extrapolating the changes we can see two possible changes. Either the temperature of the flame will reduce to the point where it is no longer higher than the water, bringing the heat flow to an end, or the liquid mass of the water will reach zero, bringing the boiling to an end. Again, the boundary conditions rule out the first event. We can thus predict that at the next transition situation, s2, the liquid mass of owater will be zero.

5. In s2, the liquid mass of owater is zero and (by HF7 and HF10) the gaseous mass is equal to the total mass. All other relations remain as in s1.

Table 7.9: Prediction in QP Theory (Continued)

Third iteration:

1. In s2, the only active process is the heat flow ph1.
2. The heat flow hf1 has a positive influence on the heat of owater and a negative influence on the heat of oflame (HF1).
3. As above, we infer nonmonotonically that the heat of owater and its temperature increase, while the heat and temperature of oflame decrease. (HF6.)
4. The only possible transition is that the temperature of the water and the flame become equal, possibly bringing the heat flow from oflame to owater to an end.
5. In s3, the temperatures of oflame and owater become equal. All other relations are as in s2.

Fourth iteration:

1. In s3, the heat flow ph1 may or may not continue (HF1, HF2). If it does not, then the system has reached stasis. If it does, we continue as follows:
2. The influence of the heat flow is to raise the heat of owater and decrease the heat of the flame.
3. Applying the closed-world assumption, we infer that the heat of owater increases and the heat of oflame decreases. From HF6, we would infer that the temperature of owater increases and the temperature of oflame decreases. If this happened, however, the heat flow would immediately halt, by HF2. Therefore, we can infer that as long as the heat flow continues, the temperatures of owater and oflame must be equal. If the heat flow ever stops, then since there are no active processes, the temperatures will still remain the same.
4. No further transitions are possible.

The basic constraints on solid objects are that they are rigid, they move continuously, and two objects do not overlap. We further assume that all objects are regular and connected. Axioms SO.1–SO.5 express these formally.

SO.1. value_in(S,place(O)) =
image(value_in(S,position(O)),shape(O)).
(Relation of position and place.)

SO.2. rigid_mapping(value_in(S,position(O))).
(Objects are rigid.)

SO.3. continuous(position(O)).
(Objects move continuously over time.)

SO.4. $O1 \neq O2 \Rightarrow$
$\neg$overlap_reg(value_in(S,place($O1$)), value_in(S,place($O2$))).
(Two distinct objects do not overlap in the same situation.)

SO.5. regular(shape(O)) $\wedge$ connected(shape(O)).
(Objects are regular and connected.)

A *kinematic* analysis of a system of solid objects is one that uses only the above properties. Thus, it considers only how the geometry of the solid objects involved causes them to block each other's motion or to push one another into position, and ignores issues such as mass, forces, energies, friction, and so on. Kinematic analysis by itself is sufficient to establish inferences like the following: (Figure 7.5)

A. If the topmost gear in Figure 7.5A is rotated clockwise, the other objects will move as shown. (We assume that the gears are pinned at their centers to a fixed frame.)

B. The gear in Figure 7.5B cannot move.

C. An object cannot go from inside to outside a closed box.

D. A hanger on a pole cannot move directly downwards.

Formalizing statements such as those above involves only enumerating the objects involved and specifying their shapes, their positions, and the constraints on their motions. Kinematic analysis is then just determining the class of motions consistent with these boundary conditions and the constraints SO.1–SO.5. We have discussed a number of techniques for such analysis, including topological reasoning and the use of configuration spaces, in Sections 6.1.6 and 6.2.5. For example, the statement C can be formalized as follows:

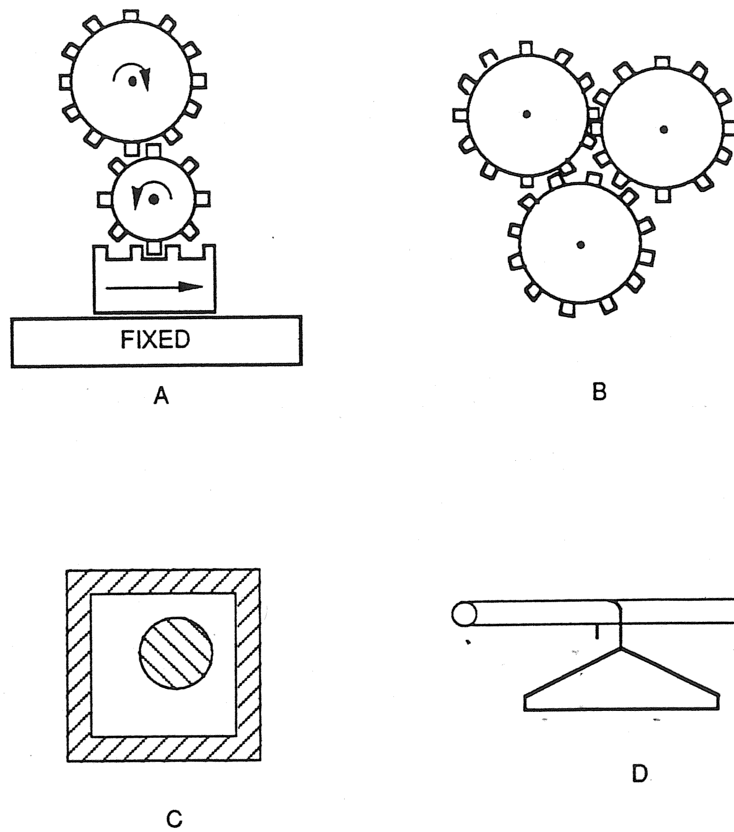


Figure 7.5 Kinematic systems

[precedes($S, S1$) ∧
is_inside(value_in(S,place($O2$)), value_in(S,place($O1$)))] ⇒
is_inside(value_in($S1$,place($O2$)), value_in($S1$,place($O1$)))]

This follows directly from the geometric rule that, if a continuous fluent over regular regions goes from inside a box to outside it, then it must at some point overlap the box.

Another example: We can show that gear B will rotate if gear A rotates in Figure 7.5A, by constructing the configuration space for the two gears and establishing that, for any path P through the permitted region of the space, if A rotates on P, then B rotates as well.

There are, however, many aspects of solid-object behavior that are not captured in a kinematic analysis. For example, consider the following inferences (Figure 7.6):

E. A block sitting on a table will remain motionless. A block dropped on a table will come to rest near where it was dropped.

F. If you lift a table at one end, it will rotate around the farther legs.

G. A round wheel will roll easily; a square wheel will not.

H. A garden rake placed in an umbrella stand is likely to knock it over. A rake placed against a piano may fall, but it will not knock the piano over.

Kinematic analysis supports none of these conclusions; any other continuous rigid motion that does not bring objects into collision is equally consistent with the constraints of kinematics. These inferences require a *dynamic* analysis, involving masses, forces, and related concepts.

The standard theory of solid-object dynamics is Newtonian mechanics. The central new concept is that of a force. In most problems, three types of forces are needed: the gravitational force, which may be taken, within a local terrestrial environment, to be uniformly downward; constraint forces, which enforce the constraint that solid objects do not overlap; and frictive forces, which act to reduce the sliding between objects. Force is related to motion by two differential equations: Force equals mass times linear acceleration, and torque equals the moment of inertia times angular acceleration. Collisions must be handled separately by conservation of energy and momentum.

In order to avoid a theory in which all objects are in free fall, it is necessary to define certain objects as being fixed in space. We introduce the predicate "fixed(O)," satisfying the axiom that fixed objects do not move.

$$\text{fixed}(O) \Rightarrow \exists_P \forall_S \text{ value\_in}(S,\text{position}(O)) = P.$$

It is often useful to introduce the *ground* as a special fixed object occupying an infinite surface below all other objects.

We now present Newtonian mechanics for a very restricted domain, in which there is only one mobile point object moving among fixed obstacles. This radical simplification enables us to avoid the messy issues involved in variable orientations and the associated issues of angular velocity, angular momentum, and torque; extended contact between objects; and collisions among collections of mobile objects.

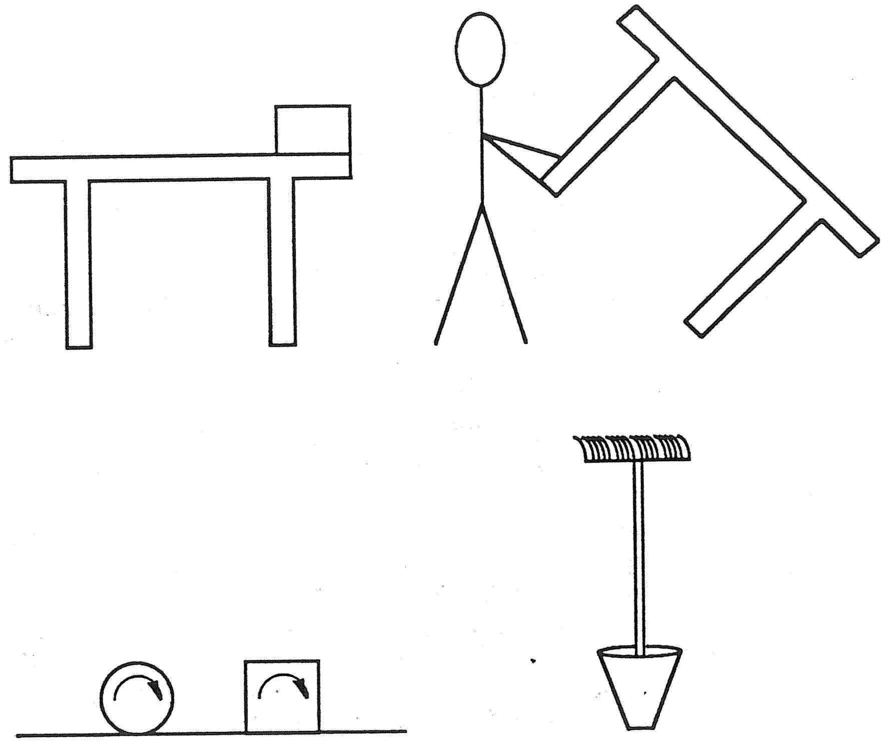


Figure 7.6 Dynamic systems

(More on these below.) We denote the space occupied by the obstacle as "obstacle," and we posit that it is regular and smooth. Table 7.10 shows the nonlogical symbols involved in the theory. (The physical symbols are all constants, owing to the restricted nature of the domain. We also introduce some new geometric symbols.) Table 7.11 gives the axioms that govern the behavior of the object.

From the rules in Table 7.11 it is possible (though not easy) to prove the following two useful rules: (i) Reduction of energy: The mechanical energy of the object never increases. (ii) Attainment of rest: The object eventually reaches a state of rest, unless there is an infinitely deep hole, or the coefficient of friction is zero.

PO.17. energyp = massp × grav_acc × height(placep) +
1/2 massp × velocityp · velocityp.
(Definition of the fluent "energyp", mechanical energy of the object.)

Table 7.10 Nonlogical Symbols for Point-Object Dynamics

Sorts: Mass (a quantity space), spatial sorts, states, situations, forces (vectors with dimension mass times distance over time squared).

Physical:

obstacle	—	Spatial region occupied by the obstacle.
massp	—	Mass of the point object.
coeff_fricp	—	Coefficient of friction of the object against the obstacle (dimensionless quantity).
restitutionp	—	Coefficient of restitution of the object colliding with the obstacle (dimensionless quantity).
placep	—	Place of the object (fluent).
velocityp	—	Velocity of the object (fluent).
contact	—	State of contact between the object and the obstacle.
collision	—	State of a collision between the object and the obstacle.
forcep	—	Total force exerted on object (fluent).
fricp	—	Frictive force exerted on object (fluent).
normalp	—	Normal force exerted on object (fluent).
grav_acc	—	Acceleration of gravity.

Geometric and quantitative:

surf_norm(**RR**, **P**)	—	Function. Surface normal out of **RR** at boundary point **P**.
smooth(**RR**)	—	Predicate. Region **RR** has a smooth boundary.
mag(V)	—	Function. Magnitude of vector V.
low_limit(F)	—	Function. Let F be a fluent. low_limit(F) is likewise a fluent. The quantity value_in(S,low_limit(F)) is the limit of F as time approaches S from below (previous times).
high_limit(F)	—	Function. Fluent analogous to low_limit.
perp_comp($V1, V2$)	—	Function. Component of vector $V1$ in direction perpendicular to $V2$. perp_comp($V1, V2$) = $V1 - (V1 \cdot V2)V2/\text{mag}(V2)$.

Table 7.11 Axioms for the Dynamics of a Point Object

Kinematics: PO.1–PO.3 are analogous to SO.3–SO.5 but modified to suit the requirements of a point object.

PO.1. continuous(placep).
(The object moves continuously over time.)

PO.2. ¬value_in(S,placep) ∈ interior(obstacle).
(The object is never inside the obstacle.)

PO.3. regular(obstacle) ∧ smooth(obstacle).
(The obstacle is regular and smooth.)

PO.4. velocityp = deriv(placep).
(Definition of velocity.)

Dynamics:

PO.5. ¬true_in(S,collision) ⇒ value_in(S,forcep) = massp × value_in(S,deriv(deriv(placep))).
(Newton's second law: F=ma.)

PO.6. forcep = −grav_acc × massp × $\hat{k}$ + fricp + normalp.
(The forces are gravity, friction, and the normal force.)

PO.7. true_in(S,contact) ⇔ value_in(S,placep) ∈ boundary(obstacle).
(Definition of contact.)

PO.10. ¬true_in(S,contact) ⇒
value_in(S,fricp) = value_in(S,normalp) = 0.
(Frictive and normal forces only apply when the object is in contact with the obstacle.)

PO.11. parallel(value_in(S,normalp),
surf_norm(obstacle,value_in(S,placep))).
(The normal force is always parallel to the surface normal at the contact point.)

PO.12. perpendicular(value_in(S,fricp),
surf_norm(obstacle,value_in(S,placep))).
(The frictive force is always perpendicular to the surface normal at the contact point.)

PO.13. mag(value_in(S,fricp)) ≤ coeff_fricp × mag(value_in(S,normalp)).
(The magnitude of the frictive force is always less than or equal to the magnitude of the normal force times the coefficient of friction.)

PO.14. [value_in(S,velocityp) ≠ 0 ∧ true_in(S,contact)] ⇒
[mag(value_in(S,fricp)) =
coeff_fricp × mag(value_in(S,normalp)) ∧
parallel(value_in(S,fricp), −value_in(S,velocityp))].
(If the object is moving against an obstacle, then the magnitude of the friction is equal to the magnitude of the normal force times the coefficient of friction, and it is directed opposite to the velocity.)

7.11 Axioms for the Dynamics of a Point Object (Continued)

PO.15 true_in(S,collision) $\Leftrightarrow$
[true_in(S,contact) $\land$
$\neg$perpendicular(value_in(S,low_limit(velocityp)),
surf_norm(obstacle,value_in(S,placep))].
(A collision is taking place just if the object is in contact with the obstacle, and the limit of velocity from before is not tangent to the obstacle.)

PO.16. [true_in(S,collision) $\land$
N=surf_norm(obstacle,value_in(S,placep))] $\Rightarrow$
[value_in(S,high_limit(velocityp)) $\cdot$ N =
$-$restitutionp $\cdot$ value_in(S,low_limit(velocityp)) $\cdot$ N $\land$
perp_comp(value_in(S,high_limit(velocityp)),N) =
perp_comp(value_in(S,high_limit(velocityp)),N)].
(Rule of collisions: The velocity in the direction of the surface normal is reversed and reduced by the coefficient of restitution. The velocity tangential to the surface is unchanged.)

PO.18. precedes($S1, S2$) $\Rightarrow$
value_in($S1$,energyp) $\geq$ value_in($S2$,energyp).
(Energy never increases.)

PO.19. [infinite_on_right(I) $\land$ coeff_fricp $\neq 0$ $\land$
[$\exists_H \forall_P$ height(**P**) $< H \Rightarrow$ **P** $\in$ obstacle]] $\Rightarrow$
$\exists_{S \in I}\ \forall_{S2 \in I}$ precedes($S, S2$) $\Rightarrow$
value_in(S,placep) = value_in($S2$,placep).
(The object eventually reaches a state of rest.)

These rules can be used to solve a variety of problems. A simple problem is shown in Figure 7.7: The object is dropped from within a steep funnel. We wish to show that it must eventually exit the bottom of the funnel. The proof is as follows: Geometrically, since the object cannot enter the interior of the obstacle (PO.2), it must either exit the top, exit the bottom, or stay inside forever. To exit the top, it would have to go higher than it is at the start, which (by PO.17) would mean that it has more energy than at the start, which is ruled out by (PO.18). If it stays inside forever, then, by (PO.19), it must eventually come to a state of rest. By (PO.15) and (PO.5), the net forces on it then must be zero. By (PO.6) this means that the frictive and normal forces must counteract the gravitational force. However, using (PO.11), (PO.12), and (PO.13), which assert that the normal force is parallel to the surface normal, the frictive force is perpendicular, and

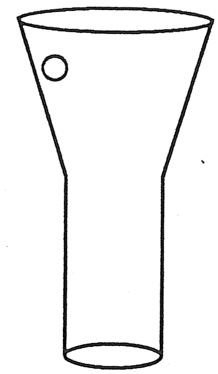


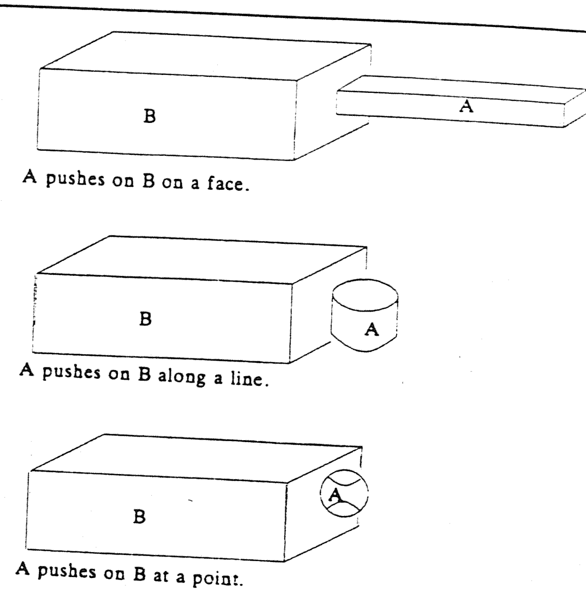
Figure 7.7 Point object in a funnel

the frictive force is at most the coefficient of friction times the normal force, we can show geometrically that this balancing of forces can only hold if the slope of the surface is less than the coefficient of friction. Thus, by showing from the geometric specifications of the funnel that, at any point of the surface that the object can reach, the slope is too great, we can conclude that the object cannot come to rest inside. The only remaining possibility is that it exits the bottom.

Another example of the use of the above theory is in deriving the rules used for a point object on a track in the NEWTON program, discussed in Section 6.2.6. We will illustrate with the derivation of the following rule: The object cannot stay on a curve (or surface) at a point where the curve is convex and has a normal with a downward component. The proof runs as follows: We wish first to show that the object must undergo an acceleration with a positive component in the direction of the surface normal. By (PO.11), the normal force on the object (if any) is parallel to the surface normal; by (PO.12), the frictive force is perpendicular to the surface normal, so that it is irrelevant; by hypothesis, the surface normal is pointing downward, so the gravitational force, which is nonzero and points downward, also has a positive component. By (PO.6) the net force is the sum of these, and therefore has a positive component in the direction of the surface normal. By (PO.5), therefore, the acceleration has a positive component in the direction of the surface normal. However, it is possible to show geometrically that a motion on a curve can have an acceleration with a positive component in the direction of the surface normal only if the curve is concave at that point. Since the curve is convex, the object cannot stay on it.

The full theory of Newtonian mechanics extends the above toy theory to handle multiple extended objects. The complete theory has the strengths that it is correct (to the extent that rigid objects exist at all) and that it is complete: It gives a prediction for any situation. Moreover, no alternative is available; no one has described a commonsense dynamic theory of solid objects that is anywhere near complete. However, Newtonian mechanics has substantial problems as a theory of commonsense reasoning. First, it is in several respects contradictory to a naive understanding of solid objects. This is partly because some actual behaviors of solid objects, such as gyroscopic motion, are strongly counterintuitive even when they are directly perceived, and partly because inertial motion, which is the default in Newtonian theory, is very much the exception in terrestrial environments, owing to the ubiquity of friction, air resistance, and other dissipative forces.

Second, axiomatizing the complete theory of Newtonian dynamics, including friction and collisions, for extended objects is surprisingly complex (see [Kilmister and Reeve 1966; Davis 1988a]). A number of sticky issues arise, including the following: (i) Forces between objects may be applied at a point, on a curve, or across a face (Figure 7.8). The dimensionality of the force at a point depends on which of these holds: (ii) The analogue to rule (PO.19) above, that an isolated system with no source of energy will eventually reach a state of rest, is frequently useful in commonsense reasoning. Unfortunately, this rule is not true for extended objects, unless we include forces such as air resistance, which are hard to quantify. (iii) Newtonian physics is generally deterministic; that is, given the shapes and material properties of the objects involved and their positions and velocities in a starting situation, all future events are determined. There are, however, a variety of circumstances, involving friction or collisions, in which the theory becomes nondeterministic, allowing more than one possible behavior. (iv) Collisions are particularly difficult to analyze within the idealization of inelastic objects, particularly in cases where the collision involves more than two objects, or where the objects collide over some extended surface. The standard rules given in mechanics textbooks for approximating the effects of a collision using the coefficients of restitution of the objects involved are hard to apply to extended objects and, in any case, are adopted more for reasons of theoretical elegance than of close approximation to the truth [Kilmister and Reeves 1966, p. 189]. (By contrast, Coulomb's law of friction, though also an approximation, is very nearly valid across a large range of circumstances.)



(From [Davis 1988a].)

Figure 7.8 Varying dimensionality of force

Thirdly, though Newtonian physics is, no doubt, in principle sufficient to support any correct commonsensical inference, in practice it is often necessary to augment Newton's laws with more specialized heuristics. Consider the following statement: "If object O is dropped from rest onto a flat horizontal surface from height h, then O will come to rest on the surface not far from the vertical projection of its release point, unless either O is very elastic or O rolls well." It seems likely that, once the vague terms in this statement have been suitably tightened, this statement is true and a consequence of Newtonian dynamics. However, proving it as a theorem from Newton's laws seems to be very difficult. (Usual techniques for dealing with differential equations are not applicable because of the collisions involved.) Given a particular object description, starting position, and starting height, it is possible to simulate its behavior using the laws of Newtonian physics by predicting each state of collision, rolling and sliding in sequence; but the general result cannot be established by such simulations. It is currently an open question how to formulate such rules effectively for commonsense reasoning.

Thus, there is currently no qualitative theory of solid-object dynamics that supports efficient inference of commonsensically obvious facts. The Reference section at the end of the chapter lists some preliminary studies.

7.4 Liquids

It is even harder to develop a commonsense theory of liquids than a theory of solids. First, the scientific theory of fluid mechanics is further from commonsense understanding, mathematically enormously more difficult, and scientifically much less complete than Newtonian mechanics. (Basic questions, such as the nature and cause of turbulence are still not wholly solved.) Second, the ontology of liquids is much less clear than the ontology of solids, since liquids are not divided into discrete objects, but combine and divide freely. Conversely, many persistent liquid entities such as the Mississippi River do not consist of a constant body of substance but are constantly depleted and refilled.

The foundations of a commonsense theory of liquids were developed by Pat Hayes [1979]. Our discussion here derives from Hayes's, though it is more limited and differs in details.[2] The central idea is that, rather than think about individual pieces of liquid that move around, as we do with solids, we think about regions of space, and ask how much liquid they contain. Thus, we introduce the fluent, "liquid_in(**RR**)", which is the volume of liquid in region **RR** in each situation. (We will assume throughout that we are dealing with only one type of liquid, so as to avoid the difficult problems of treating mixtures.) We also define the predicate "solid(O)" meaning that O is a solid object. This fluent obeys the following axioms:

LI.1. (Additivity) **RR1** $\cap$ **RR2** $= \emptyset \Rightarrow$
liquid_in(**RR1** $\cup$ **RR2**) = liquid_in(**RR1**) + liquid_in(**RR2**).

LI.2. (Bounds) $0 \leq$ value_in(S,liquid_in(**RR**)) $\leq$ volume(**RR**).

LI.3. (Nonmixing with solids) solid(O) $\Rightarrow$
value_in(S,liquid_in(value_in(S,place(O)))) = 0.

[2] There are two major differences: (i) Hayes uses a nonstandard spatial topology, while we use Euclidean geometry. (ii) Hayes studies transitions between different states of liquids using *histories*, which are chunks of space-time. We will not consider such transitions, and therefore will not use histories.

The amount of liquid in a region **RR** changes as a result of flows into and out of **RR** and of phase changes — melting, freezing, evaporation, and condensation — of material inside **RR**. In our discussion here, we will ignore phase changes and treat only flows. Under this assumption, a conservation law states that the change to the amount of liquid in **RR** is equal to the total flow into **RR**. We can express this as follows: Let "flow_through(**FF**)" be a fluent representing the rate of liquid flow outward through directed surface **FF** in each situation. Let the function "dboundary(**RR**)" give the outward-directed boundary of region **RR**. We can state the basic properties of flow as follows.

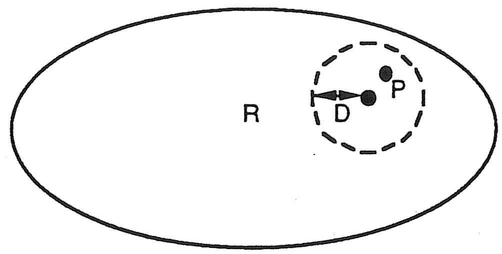
LI.4. deriv(liquid_in(**RR**)) = −flow_through(dboundary(**RR**)).
(Conservation of mass: The change in liquid contained is equal to minus the total outward flow.)

LI.5. **FF1** ∩ **FF2** = ∅ ⇒
flow_through(**FF1** ∪ **FF2**) =
flow_through(**FF1**) + flow_through(**FF2**).
(Flow is additive over faces.)

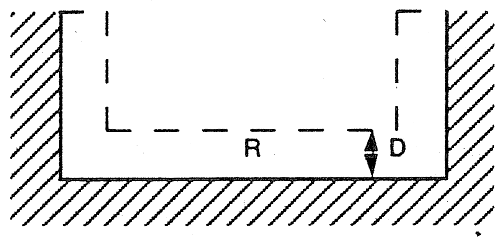
LI.6. [solid(*O*) ∧ **FF** ⊂value_in(*S*,place(*O*))] ⇒
value_in(*S*,flow_through(**FF**)) = 0.
(There is no flow through a solid face.)

It follows directly from these axioms that a region with solid boundaries always has a constant amount of liquid, and that the change of liquid of a region with a number of openings is equal to the net flow through the openings.

In order to express the fundamental dynamic rules governing the behavior of liquids, we need to define some additional geometric and physical concepts. We introduce the following nonlogical symbols: The constant length "thin" is the maximum thickness of liquid wetting a surface. (This distance is, in reality, dependent on a number of factors, particularly the material of the liquid. We ignore this.) The predicate "bulk(**RR**,$\tilde{D}$)", read "Region **RR** is bulk with thickness $\tilde{D}$," holds if, for each point **P** in **RR**, there is a sphere **SS** of radius $\tilde{D}$, such that **P** is in **SS** and **SS** is a subset of **RR** (Figure 7.9). The function "filled_liquid(**RR**)" is the state of region **RR** being filled with liquid. The function "empty(**RR**)" is the state of region **RR** being empty of either liquid or solids. The function "liquid_at_rest(**RR**)" is the state of all liquid in **RR** being motionless and having zero acceleration (to exclude the case of a liquid at the top of a fountain.) The function "solid_coating(**RR**,$\tilde{D}$)" is the state of all points in **RR** being within distance $\tilde{D}$ of a solid object. Axioms LI.7–LI.11 give the formal definition of these symbols.



R is bulk with thickness D



R is solid coating of thickness D

Figure 7.9 Predicates "bulk" and state "solid_coating"

LI.7. bulk(**RR**,$\tilde{D}$) ⇔ $\forall_{P \in RR} \exists_{Q \in RR}$ **P** ∈ sphere(**Q**,$\tilde{D}$) ⊂ **RR**.
(Definition of bulk.)

LI.8. true_in(S,filled_liquid(**RR**)) ⇔
value_in(S,liquid_in(**RR**)) = volume(**RR**).
(Definition of filled_liquid.)

LI.9. true_in(S,empty(**RR**)) ⇔
[value_in(S,liquid_in(**RR**)) = 0 ∧
[$\forall_O$ solid(O) ⇒ **RR** ∩ value_in(S,place(O)) = ∅]] .
(Definition of empty.)

LI.10. true_in(S,liquid_at_rest(**RR**)) ⇒
$\forall_{FF \subset RR}$ value_in(S,flow_through(**FF**)) = 0.
(If the liquid in **RR** is at rest, then there is no flow through any face contained in **RR**.)

LI.11. true_in(S,solid_coating(**RR**,$\tilde{D}$)) ⇔
$\forall_{P \in RR}$ $\exists_O$ solid(O) ∧ distance(**P**,value_in(S,place(O))) < $\tilde{D}$.
(Definition of solid coating.)

We can now express some simple laws of the dynamics of liquids. We will consider only liquid in bulk and liquid wetting a surface; we exclude such other states as absorbed in an absorbent material or spread about as mist. The following rules are then plausible:

LI.12. Liquid at rest must either be in bulk or wetting a surface.
[true_in(S,filled_liquid(**RR**)) $\wedge$
true_in(S,liquid_at_rest(**RR**)] $\Rightarrow$
$\exists_{RRB,RRC}$ **RR=RRB** $\cup$ **RRC** $\wedge$ bulk(**RRB**,thin) $\wedge$
true_in(S,solid_coating(**RRC**,thin)).

LI.13. A liquid in bulk at rest can border the air only at a horizontal surface.
[bulk(**RR**,thin) $\wedge$ true_in(S,filled_liquid(**RR**)) $\wedge$
true_in(S,liquid_at_rest(**RR**)) $\wedge$
true_in(S,empty(**RR2**)) $\wedge$ abut(**RR**,**RR2**,**FF**)] $\Rightarrow$
$\forall_{P \in FF}$ surf_norm(**RR**,**P**) = $\hat{k}$.

LI.14. If, after time $S0$, there is no flow into or out of region **RR**, and no solid object moves in **RR**, then the liquid in **RR** will eventually come to rest. (We introduce the state function, "motionless(O).")
[[$\forall_{S>S0}$ [$\forall_F$ **FF** $\subseteq$dboundary(**RR**) $\Rightarrow$
value_in(S,flow_through(**FF**)) = 0] $\wedge$
[[$\forall_O$ solid(O) $\wedge$ value_in(S,place(O)) $\cap$ **RR** $\neq \emptyset$] $\Rightarrow$
true_in(S,motionless(O))]] $\Rightarrow$
$\exists_{S1>S0}$ $\forall_{S>S1}$ true_in(S,liquid_at_rest(**RR**)).

These rules are not strictly true. For example, the surface of a contained liquid is not perfectly horizontal, due to surface tension; in the scenario of the last rule, the liquid could be scattered as a mist if it encounters sufficiently violent collisions, like a waterfall on rocks. However, they are reasonable initial approximations.

Using the above axioms we can establish results such as the following: Let O be a solid object with an internal spherical cavity of radius $\tilde{R}$ containing liquid of volume V. Let τ =thin. Then, if O is motionless and $V > 4\pi\tilde{R}^2\tau$, then after sufficient time, there will be a puddle of volume at least $V - 4\pi\tilde{R}^2\tau$ at the bottom of the cavity. (Figure 7.10)

The proof is as follows: Since the cavity is the inside of O, which is a closed box, any face of the cavity must border on O. Therefore, by LI.6, there can be no liquid flow into or out of the cavity, and by axiom SO.4, no solid object can come inside the cavity. By LI.4, the quantity of liquid in the cavity must remain constant. Morover, the antecedents of LI.14 are satisfied, so the liquid must attain a state of rest inside the cavity. By LI.12, when the liquid is in a state of

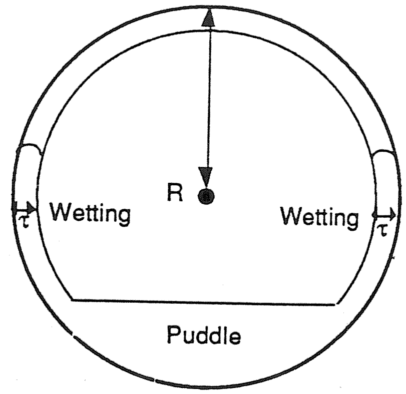


Figure 7.10 Liquid in a cavity

rest, it must all be either wetting a surface or in bulk. Since the solid surface bounding the cavity has area $4\pi R^2$, the maximum quantity that can be wetting the surface is $4\pi R^2\tau$. Therefore, the remainder of the liquid, of quantity $V - 4\pi R^2\tau$, must be in bulk. However, by LI.13, liquid at rest in bulk can border empty space only below a horizontal surface. By a geometrical argument, this remaining liquid must lie in a puddle at the bottom of the cavity.

7.5 Physical Agents

So far, our theories have considered only the behaviors of inanimate objects. In many important applications, however, it is necessary to reason about physical interactions involving intelligent agents. In particular, an intelligent creature will generally be interested in how it can affect the world, and how the world can affect it. To address the first question, a physical theory must supply a language for describing the *actions* of an autonomous agent and a theory of how these action affect the world. We will not here address the question of reasoning about the impact of the world on an agent, though we will touch on it indirectly in Sections 8.7 and 9.3.5, which deal with perception and the adoption of goals by an agent. (The conceptual dependency (CD) representation of human actions [Schank 1975] will be considered separately in Appendix 10.A, since it combines physical, mental, and interpersonal reasoning.)

Given any physical theory of inanimate objects, such as those we have considered so far in this chapter, it is possible to extend the theory to include agents and actions as physical entities and physical events that are only weakly constrained by physical law. For example, in a component theory, we could model an agent as a component with ports, and an action as a relation among the port parameters that is enforced at the will of the agent. Thus, in the scales example, we could add an agent as a component with a port attached to the weight side of the lever, and specify that it executes actions such as "Move the lever down at a constant rate" or "Apply a force of 50 pounds until the height of the port reaches 2 feet; thereafter, hold the height constant there." It is then possible to use the component model to predict the behavior of the system, given this action. Similarly, in a process model, we could model an agent as a process and an action as a trace of its influences over time. In a kinematic theory of solid objects, we could model an agent as a solid object, and an action as a trace of its position over time. In a dynamic theory of solid objects, we could model an agent as a solid object, and an action as a time-varying relation between its position and the forces it applies to objects in contact.

7.6 References

General: [Weld and de Kleer 1989] is a very comprehensive collection of papers on qualitative physics, including most of the papers cited below. Hayes's paper "The Naive Physics Manifesto" [1978] was the first extensive discussion of commonsense physical reasoning. It advocates the development of logical theories that express commonsense physical intuition. [Gentner and Stevens 1983] is a collection of papers on psychological and historical aspects of physical reasoning.

Component model: Our description of the component model is based on the well-known paper [de Kleer and Brown 1985], which describes the ENVISION program. The "No function in structure" principle is discussed there. Also of interest are the papers by Kuipers [1985] and Williams [1985]. Component-based analyses for electronic systems have been studied in [Sussman and Steele 1980] and [Davis, R. 1983]; both papers discuss the use of a hierarchical analysis of systems. An alternative approach to component analysis, based on combining small components together into larger components of known behavior, is considered in [Bylander and Chandrasekaran 1985]. [Rieger and Grinberg 1977] was a relatively early attempt at a component analysis of a complex device (a flush toilet); however, the underlying theory was never very clearly developed, and the component descriptions did not satisfy the "No function in structure" principle.

Qualitative process theory: Qualitative process theory was first developed by Forbus [1985]. [Forbus 1989] discusses an improved version of the program. [Forbus 1986] applies QP theory to the problem of interpreting a given time sequence of measurements. Also of interest is the discussion of processes in [McDermott 1982a].

Solid objects kinematics: [Shoham 1985a] considers the motions possible for an object abutting obstacles. [Faltings 1987a] and [Faltings 1987b] analyze in detail the kinematics of two-dimensional mechanisms composed of parts each with one degree of freedom. [Joskowicz 1987] studies the kinematics of a system that has few degrees of freedom by virtue of the interaction of its components. [Joskowicz and Addanki 1988] uses a kinematic analysis to solve problems of design. [Gelsey 1987] proposes a number of special-case heuristics to increase the efficiency of the analysis. The possibility of a qualitative kinematics is denied in [Forbus et. al. 1987]; this is disproved, at least in part, by [Faltings, et. al. 1989], which exhibits such a qualitative characterization of kinematics, based on the topology of the configuration space. [Gelsey 1989] discusses the construction of kinematic models of varying degrees of detail of a physical system from its geometrical specifications, and the use of the kinematic model in prediction. Closely related is the work on the "piano-movers" problem; citations are given in the Reference section for Chapter 6.

Solid object dynamics: Though, as mentioned in the text, there is no thorough treatment of solid-object dynamics, there are many partial studies. [Fahlman 1974] gives an algorithm for determining the stability of a collection of polyhedral blocks. NEWTON [de Kleer 1975], discussed in Section 6.2.6, predicts the behavior of a point object sliding on a constraint. MECHO [Bundy 1978] uses force analysis and conservation laws to solve a variety of problems in closed form. FROB [Forbus 1979] extends NEWTON by predicting the behavior of a point object flying among fixed constraints. WHISPER [Funt 1980] simulates dynamical systems using an occupancy-array representation. Davis [1988a] argues that formulations of dynamics in terms of differential equations are not adequate for many commonsense problems, and presents a first-order theory for dynamics that, in some cases, avoids the use of differential equations. Nielsen [1988] presents a system that does qualitative dynamical reasoning. Also of interest are works in robotics (e.g., [Mason 1986; Peshkin and Sanderson 1987; Wang 1986]) that study exact solutions to restricted classes of dynamical problems.

Liquids: Our discussion of reasoning about liquids derives from [Hayes 1979]. Other papers of interest include [Schmolze 1986] and [Collins and Forbus 1987]; both of these combine a theory of liquids with a theory of processes.

Other domains: [Doyle 1989] applies a rich component model to the problem of guessing the structure of a device, such as a pressure gauge for tire, from observations of its behaviors. [Bunt 1985] discusses the problems in giving semantics to "mass" nouns; since these are generally associated with physical substances, many of the issues discussed are important for physical reasoning.

Causality: We have deliberately avoided discussing causality in the text, because there is no consensus in the AI community as to whether, where, and how causality should enter into physical theories. Various approches to causality can be found in [Rieger and Grinberg 1977; McDermott 1982a; Allen 1984; de Kleer and Brown 1985; Iwasaki and Simon 1986; de Kleer and Brown 1986; Shoham 1988; and Pearl 1988b]. [Shoham 1988] also has an extensive review of the philosophical literature on the subject.

7.7 Exercises

(Starred problems are more difficult.)

1. Verify the algebraic manipulations used to derive the various solutions to the scale equations in Section 7.1.

2. * Construct a QP model of the process of burning. Use your model to determine the behavior of the sample system of a can of water over a candle flame.

3. * In our sample trace of the qualitative process algorithm we wrote (second iteration, step 1), "From HF1, infer that the heat flow hf1 from oflame to owater continues." Actually, the axiom HF1 only supports the conclusion that there is some heat flow, not that it is necessarily the same heat flow. In fact, it is consistent with the axioms that there should be more than one heat flow, or that the identity of the heat flow(s) changes every instant.

(a) Show that the prediction we derive for the example does not depend on any of these unwarranted assumptions and, in fact, can be derived from the axioms.

(b) In view of part a, it might seem that it would be more natural not to try to distinguish process tokens, and just use process types, such as "the state of heat flowing from oflame to owater." Rewrite the theory in this way.

(c) Under what circumstances will it be useful to use process tokens rather than process types?

(d) Augment the original theory with axioms that guarantee that, in the heat-flow example, there is a single heat-flow process and a single boiling process.

4. Show that the remaining rules used in NEWTON (Section 6.2.6) can be derived from our axiomatization of point-object dynamics, together with the suitable geometric theorems.