Part I

Reconciling Natural and Mental Philosophy

1

Mechanical intelligence

What do you think when someone claims that people are mechanical?

Some people find this claim offensive, as likening their own thoughtful behavior to the unthinking behavior of the machine, as in Skinner's famous aphorism "The real question is not whether machines think but whether men do" (Skinner 1969, p. 288). Even though passing time has changed the prototypical machine from the pulley in the well to the steam locomotive to the automobile to the home computer, a comparison to machines represents one common form of insult ("Dali, pfui. He paints like a machine."). In this view, depicted in Figure 1.1, claiming people to be mechanical brings people down to a lower level.

Students of artificial intelligence seeking to construct intelligent machines often share the underlying revulsion against comparing people to washing machines and other "dumb" appliances, but usually take a broader view of machines that includes ones not yet constructed, and paint a picture in which one

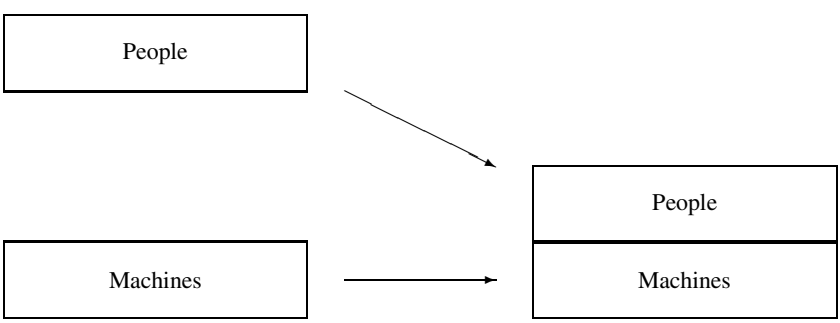

Fig. 1.1. Bringing people down to the level of machines.

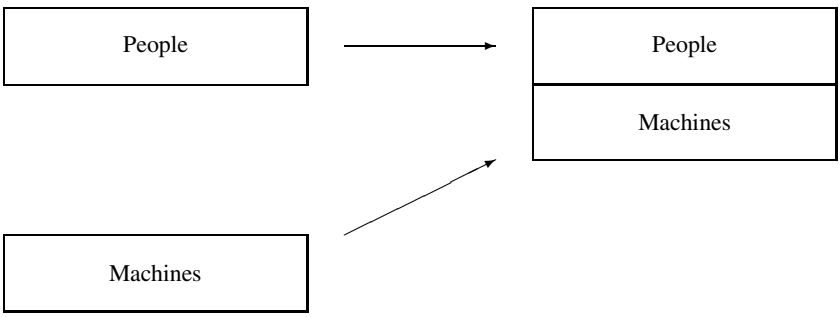

Fig. 1.2. Bringing machines up to the level of people.

endeavors to raise machines up to the level of humans, in stages of approximation if not all at once, as depicted in Figure 1.2. For example, throughout history clothes have been washed by people, and it is not demeaning to compare a person with a person who washes clothes. If only we could make washing machines as smart and capable and, well, as personable as people, say students of artificial intelligence, surely comparison to such machines need not be offensive.

This response does not placate those who view the hypothetical assumption as an offensive impossibility, or who think people have a character that no machine, no matter how intelligent, could possess. For example, many think that people have a spiritual character that sets them apart from merely material devices like machines, which only contain what their human designers put into them. Churchill famously advanced such a concern in his own inimitable way:

> The destiny of man is not decided by material computation. When great causes are on the move in the world, stirring all mens' souls, drawing them from their firesides, casting aside comfort, amusement, wealth, and the pursuit of happiness, in response to impulses at once awe-striking and irresistible, then it is that we learn that we are spirits, not animals, and that something is going on in space and time, and beyond space and time, which, whether we like it or not, spells duty. (Churchill 1941)

Making this distinction lets one offer complimentary comparisons to machines ("I can't believe how much she gets done. She just keeps going like some sort of machine!") without diminishing the sense of separation between our kind and theirs, and without legitimizing claims that people are just complicated machines.

In the following, I approach this dispute about the nature of people from a different direction.

1.1 Mechanical philosophy

Behold the man!
John 19:5 (KJV)

Thoughtful ancients saw mysteries everywhere, in space, in time, in matter, and in themselves. Some of the greatest puzzles turned on the multiple natures of things; a beach might look as smooth as the water shaping it, but close inspection reveals an array of grains of sand as seemingly numerous as the more obviously discrete stars in the heavens. The ancients recognized themselves as providing the greatest of such puzzles: once dead, composed of matter suitable for the worms of the soil; in life, manifesting both the smooth motions of the gymnasium and distinct decisions of judge and marketplace, in body bound to a location changing only a few kilometers a day, in mind free to range unbounded through worlds real and fanciful, past, present, and future, all within minutes, and for many, contemplating a life for the soul unhindered by the death of the body.

The new sciences developing across the centuries explored evidence that man's body exists subject to various regularities or laws of physics, chemistry, biology, and physiology, with some of these laws explaining much more than the human body. The array of scientific theories yielded by these explanations have greatly increased understanding of the world, and have supported powerful technologies affecting almost all areas of life: labor, transportation, communication, agriculture, medicine, manufacturing, trade, and war. We share species with the ancients, but increase in knowledge has transformed the environment of life in fundamental ways.

The advance of transformative science proceeded slowly before a dramatic acceleration in the seventeenth and eighteenth centuries, when discoveries in mathematics and rational mechanics altered the character of natural philosophy in fundamental ways.

The term *rational mechanics* has fallen out of general use, but it remains the traditional name given to the conceptual or mathematical investigation of mechanical concepts (Truesdell 1958). The term has persisted from the time of Newton to the revival of rational mechanics by Truesdell and others in the past century. Although today the scientific term *rational* is closely tied to the concept of rational decision and action in much of the literature, the term *rational mechanics* itself in no way refers to rational action as studied in psychology and economics.

Rational mechanics, as developed by Newton, Euler, and others, reworked natural philosophy into the modern sciences we know today. Earlier natural philosophy was dominated by informal, largely philosophical debate and

observation. Rational mechanics, and the mathematical viewpoint more generally, focused on technical investigation, on explicit models of the evolution of physical systems in accordance with specific mathematical equations, and on explicit calculation from specific hypothesized initial conditions to observable and unobservable properties of physical systems. These mathematical models enabled scientists to refine physical theories, and enabled engineers to construct complicated physical systems to meet precise specifications. The advance in understanding changed perspectives so much that Leibniz claimed sufficiently great calculating abilities and a full description of conditions at some initial time would permit determination of the entire future of the world subsequent to the initial time.

1.2 The great divorce

The optimism expressed by the natural philosophers did not bear out in the contemporaneous early stages of the human sciences of psychology and economics. In contrast to the progress seen in understanding the physical world, understanding the mind has proven very difficult. We understand much today compared with past centuries, but in honest appraisal this represents comparison of infinitesimals.

Why did the advance of science scant mental philosophy even while enriching natural philosophy? Part of the explanation might lie in the limited applicability of the new conceptual tools.

Recall that the seventeenth century also saw Descartes' promulgation of a dualistic theory of mind, in which a mental substance of the mind accompanied the physical substance of the body. Discourse at the time also spoke of forces on minds and bodies, just as it does today. In spite of such conceptions in which mind and body consisted of substances acted upon by forces, the mathematical tools of the new mechanics did not apply to Cartesian minds, for their mental substances lacked physical position, meaning that mental actions lacked description in terms of the physical motion treated by mechanics. The new mechanics thus offered no way to apply its developing formal concepts to understanding the relation of the mind to the body or the nature of forces acting on minds.

The study of the mind did not stagnate, however, and mathematical theories of psychology and economics emerged later from nonmechanical theories of logic, probability, and utility. These theories gave central place to the notion of rational action, eventually understood as action chosen so as to maximize the expected utility of action. The principle of rational action provided the study of the mind with a formal framework for investigation and analysis comparable with the formal framework that the central mechanical notions of force, mass,

and motion provided for physics. This difference with respect to mathematical formalism produced an increasingly wide separation of the mental and the physical sciences, between those based on the concept of rationality and those based on the concept of force.

The scientific import of this divorce of mental and physical sciences became clearer later as psychology began to explore computational characterizations of reasoning and behavior, and as economics began to cast about for theories that match human capabilities better than its foundational theory of ideal rational choice. Computational formalizations of psychological theories involved motion in "spaces" of mental states that, though very different than physical space, at least proved susceptible to mathematical formalization. Realistic economists grew appreciative of the hard work involved in making choices and of the slowness of the mind to change when subjected to new information or other influences. Popular discourse still spoke of mental forces, work, and inertia to reflect these concerns ("I had to force myself to concentrate"), much as in the days of Descartes and Newton. People also came to use mechanical concepts of inertia, force, energy, and pressure informally in describing economic markets and behavior ("Market forces are putting increasing pressure on oil prices"). In spite of the continuing application of seemingly similar concepts, the divorce of the mental and physical sciences impoverished the mental sciences when compared with the physical sciences by abandoning to the purely physical realm mechanical concepts of force and inertia that proved fruitful in analyzing physical behaviors. Study a physical problem, and one has recourse to physics, chemistry, and biology, as well as differential equations and mathematical theorems that aid in analysis and prediction. Study a mental system, and one lacks almost all of this intellectual heritage, for the traditional conceptual tools do not apply.

1.3 The awaiting reconciliation

The scientific separation of mental and physical need not stand. In the following, I bridge the gap between matter and mind with mechanics, and explore the possibility that people are indeed mechanical, in both mind and body, but are not necessarily machines or material machines. I do this, as depicted in Figure 1.3, by understanding "mechanical" in the sense of the science of mechanics, and show how one can rework the traditional mechanics one learns in high school or college physics classes to cover reasoning and other mental phenomena in a natural way.

Specifically, I show that the mathematical concepts of modern axiomatic rational mechanics apply more broadly than generally recognized. The quiet progress of mechanics in recent years provides formal concepts of force, mass,

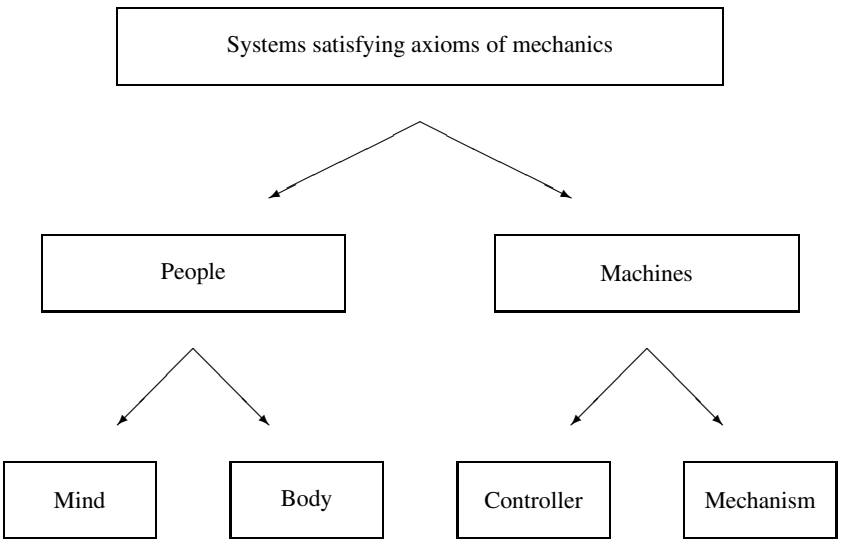

Fig. 1.3. Understanding people and machines as mechanical systems. Not all machines have nontrivial controllers.

momentum, and work that enable one to transform some heretofore metaphorical uses of these terms into meaningful, true or false, nonmetaphorical statements about psychological and economic systems within the axiomatic framework of modern rational mechanics.

In psychology, applying the mechanical perspective to mental inertia and mental forces helps one understand and formalize the difficulty of changing one's mind, of learning, of maintaining a focus of attention in the presence of distractions, and of overcoming habitual behaviors. Mechanics helps one understand the different characters of people and types of people.

In economics, nonphysical applications of mechanical axioms provide new means for characterizing more realistic notions of economic rationality and limits on reasoning abilities, and translate studies of different types of psychological and economic agents into studies of new types of mechanical materials.

In artificial intelligence, mechanics provides new concepts for analyzing the structure of artificial agents, new terms with which to specify desired characteristics of agents, and new paths for implementing agents efficiently.

The mechanical perspective provides these benefits without requiring one to give up nonmechanical perspectives. It instead provides an additional perspective offering clearer paths to some familiar apprehensions than those offered by traditional perspectives.

The mechanical perspective does not demote people to the level of machines. One might think of the human body as a machine of magnificent design, or one might not, for even traditional mechanics appears to transcend standard conceptions of "machine," especially the notion crystallized by Turing and popularized in today's digital computers. In a similar way, one might think of the human mind as a machine, or as something more, because mechanics itself does not say what sorts of forces exist in the world, nor from whence they issue.

2

Why mechanics?

The mechanical understanding of mind bridges both the gap between the mental and the physical and the gap between the rational and the dynamical. In addition to seeking a better understanding of the relation of mind to body, one specific motivation in pursuing this understanding stems from an interest in finding new means with which to characterize and analyze limits to rationality, a central interest common to psychology, economics, and artificial intelligence. Pursuing this motivation requires facing philosophical problems that have puzzled people for millennia.

Although science has answered some of these philosophical questions about nature and mind, it has left others unanswered. For example, one ancient question concerns determinism, or more generally, lawfulness. Many views hold the mind to exhibit essential freedoms not enjoyed by matter; other views hold the mind subject to various laws of psychology, economics, sociology, and anthropology, and argue about the precedence of these competing regulations. Though scientific progress has inspired some of the competing variants and the development of quantum theories has complicated the stark alternatives contemplated by earlier generations, scientific evidence has done less than one might expect to support or weaken the cases for the fundamental alternatives. The liberty or lawfulness of the mind remains controversial.

Unresolved questions do not represent failures of science. They represent the human condition. Given the long lifetime of fundamental questions, one measures the contribution of science not so much in terms of how many questions it has answered, but in terms of how many problems it made amenable to technical and experimental investigation. Truesdell, as usual, states the issue beautifully:

> Now a mathematician has a matchless advantage over general scientists, historians, politicians, and exponents of other professions: He can be wrong. *A fortiori*, he can also be right. (Truesdell 1968b, p. 140)

The quickest way to tire the lay observer is with what appears to be philosophical debate, for philosophical debate has the reputation, perhaps deserved, of never resolving anything. The mathematical and theoretical advances of modern natural science have left some fundamental questions unanswered, but they have shown how to remove others from the domain of opinion into the domain of knowledge.

The continuing lack of consensus on fundamental characteristics of mind illustrates the paucity of progress in converting the questions of mental philosophy into subjects for technical and experimental investigation. Accordingly, I believe the primary immediate benefit provided by the reconciliation of the mental and the physical comes not in providing immediate answers to long-standing questions but in opening some long-standing philosophical problems to serious mathematical investigation. The more one removes technical limitations that handicap the human sciences relative to the physical sciences, the more one improves prospects for rich and effective mental sciences.

We cannot yet see all the ramifications of the mechanical perspective. Nevertheless, it seems likely that augmentation of the existing technical conceptions of logic, economics, and computational intelligence with the formal concepts of mechanics will permit construction of mechanical theories of the interaction of mind and body and of limits on ideal economic rationality. These mechanical theories in turn seem likely to offer improvements in techniques used in engineering artificial agents. The remainder of this chapter sketches elements of such potential benefits. Later chapters return to the ideas to provide more details.

2.1 Rethinking materialism

The ancient question of materialism, as regards psychology, asks whether people have minds or spirits distinct from their body, or whether these are mere by-products of brain and body. Philosophers have speculated for centuries about possible relations between mind and body, with theories ranging from nonexistence of mind to nonexistence of body, and from complete disconnection of mind from body to complete correspondence of mind and body.

Although Descartes viewed mind and body as somewhat separate entities acting on each other, dualistic theories fell into disrepute for at least two reasons. First, proponents of dualistic theories could not supply any formal model for or rules governing either mental motions or the proposed interactions between mind and body. Science was just beginning to understand physical forces in mathematical terms, but not in a way that applied to understanding interactions of mind and body. Second, even setting aside the lack of a formal