Chapter 2
Prehistory and Emergence

> For one good deed leads to another good deed, and one transgression leads to another transgression.
>
> *Pirke About* (*The Wisdom of the Fathers*)

The concept of the feedback loop in twentieth-century social science is a blend of intuitions and ideas from at least six intellectual traditions: engineering, economics, biology, mathematical models of biological and social systems, formal logic, and classical social science literature itself. To understand the evolution of the feedback idea in the social sciences, we must investigate the development and use of loop concepts underlying feedback and circular causality in these other areas. We shall find that ideas from all six traditions are influential, but not equally so across all areas and authors in the social sciences. One feedback thread is more directly linked, both sociologically and methodologically, with engineering servomechanisms and mathematical models in biology (sections 2.1 and 2.2). The other is more influenced by homeostatic mechanisms in biology and ideas from formal logic (sections 2.3 and 2.4). Both threads select from the engineering and the social science literature, but in subtly different ways. At the close of this chapter, we will be in a position to see the beginnings of these different feedback threads in the social sciences.

2.1 Engineering Servomechanisms and Control Theory

The engineers' contribution to our understanding of feedback is often thought to originate with the dramatic proliferation of feedback devices in the late eighteenth century. The classic example, frequently cited as the first conscious use of a feedback device, is James Watt's centrifugal governor for the steam engine (1788) (Figure 2.1). The negative loop nature of this controller is clear from the description in the figure. A change in the speed of the engine produces forces in the governor that counteract the change and return the engine to its normal operating speed.

unstable material and subjected continually to disturbing conditions, constancy is in itself evidence that agencies are acting, or ready to act, to maintain this constancy.

- If a state remains steady it does so because any tendency towards change is automatically met by increased effectiveness of the factor or factors which resist the change.
- The regulating system which determines a homeostatic state may comprise a number of cooperating factors brought into action at the same time or successively.
- When a factor is known which can shift a homeostatic state in one direction it is reasonable to look for automatic control of that factor, or for a factor or factors having an opposing effect (1932, pp. 281–282).

He repeatedly emphasized the second principle, in which change is met by "increased effectiveness" of the factors that resist the change. It is tantamount to a modern definition of a negative feedback loop. The first principle thus essentially asserts that stability in a dynamic system necessarily implies the existence of one or more negative feedback loops. Cannon himself, however, did not expose the loop nature inherent in the regulatory processes he investigated. The linking of homeostasis with the feedback loop came about only later, through the collaboration of Cannon's colleague Auturo Rosenbleuth with Norbert Wiener (see section 2.5).

Such generalizations suggested to Cannon that the concept of homeostasis had application beyond the physiological organization of single living organisms and could be applied meaningfully to societies. He ended *The Wisdom of the Body* with a chapter on "social homeostasis." Central to his thinking was the principle that a degree of constancy in a system was evidence of the necessary existence of homeostatic mechanisms striving to maintain that constancy. Society showed the beginnings of such tendencies:

> A display of conservatism excites a radical revolt and that in turn is followed by a return to conservatism. Loose government and its consequences bring the reformers into power, but their tight reins soon provoke restiveness and the desire for release. The noble enthusiasms and sacrifices of war are succeeded by more apathy and orgies of self-indulgence. Hardly any strong tendency in a nation continues to the stage of disaster; before that extreme is reached corrective forces arise which check the tendency and they commonly prevail to such an excessive degree as themselves to cause a reaction. A study of the nature of these social swings and their reversal might lead to valuable understanding and possibly to means of more narrowly limiting the disturbances. At this point, however, we merely note that the disturbances are roughly limited, and that this limitation suggests, perhaps, the early stages of social homeostasis (1932, pp. 293–294).

Cannon's descriptions in this paragraph also strongly call to mind the problem of oscillations in engineering feedback control systems that Maxwell had addressed some sixty years before. Yet, as noted above, Cannon did not make the link. Instead, he focused on the still primitive nature of societal homeostasis. He concluded that the mechanisms necessary to maintain the stability of society's "fluid matrix" or "milieu interne" have not yet developed sufficiently. To him it was a matter of societal evolution. The higher an organism is on the evolutionary ladder, the more developed its homeostatic capabilities, and the freer its individual cells and organs are to specialize. Cannon suggested that society would follow a similar pattern. The evolutionary goal for society was the development of automatic societal stabilization mechanisms, ultimately operating below the level of conscious control (pp. 302–306). To Cannon it was an exciting goal, full of potential for releasing the "highest activities of the nervous system for adventure and achievement" (p. 305).

The Impact of the Notion of Homeostasis

The central theme of the idea of homeostasis is control. In feedback terms, it is control through the operation of negative feedback loops. Self-reinforcing positive loops have no role in homeostatic mechanisms. The clear link between Cannon's development of the concept of homeostasis and the work of control engineers led some to overemphasize negative feedback and control in socio-economic applications of the feedback idea. As we shall see, Cannon's observations about mechanisms of social homeostasis also reinforced the urge to add on societal feedback control mechanisms, rather than to strive to perceive those feedback structures already contributing to observed societal dynamics.

2.4 Logic Loops

The modern concept of a feedback loop is also influenced by loop notions arising from formal logic and computing. Three kinds of logic loops have been important, in different ways, in shaping the way various modern authors work with the feedback concept: the vicious circle, the self-referring argument, and the modern DO-loop from computing.[10] The former is a very old notion that is completely compatible with the engineer's concept of feedback. The two ideas eventually merge smoothly, with the vicious circle being seen as merely another name for a positive feedback loop. The self-referring argument and the iterative DO-loop, however, are quite distinct ideas. They

are exceptionally powerful concepts with important implications for the social sciences, but they do not derive their power from the concept of positive and negative feedback. Nonetheless, in the writings of some social scientists they interweave with the concept of feedback, with mixed results.

The Vicious Circle

Today we interpret the phrase "vicious circle" to refer to a bad situation that leads to its own worsening. My children, for example, are sometimes "so tired that they can't get to sleep." But the phrase and the concept actually have their origins in formal logic.

In Elizabethan times, the word "vicious" meant, among other things, "flawed" or "faulty" (OED 1933). Any fallacious logical argument was "vicious." One particular form of vicious argument well known in the 1600s (and much earlier) was circular reasoning—basing an argument on the very proposition to be proved. Such an argument naturally came to be known as a vicious circle (*Encyclopædia Brittannica* 1792, cited in OED 1933), meaning simply a process of reasoning that is faulty because it is circular.

But "vicious" also meant evil, harmful, and threatening. By the mid-1800s the concept of the vicious circle had evolved from its narrow meaning of flawed logic to a more general notion of circular causality in which bad leads to worse. The following is an early example from the French:

> I'd need rest to refresh my brain, and to get rest, it's necessary to travel, and, to travel, one must have money, and, in order to get money you have to work, create, etc.: I am in a vicious circle [*cercle vicieux*], from which it is impossible to escape (Balzac 1850, p. 32).

Here the circle of propositions describes a closed sequence of causes and effects, not a logical fallacy. The result is an explicitly circular process, perceived as characteristically self-perpetuating and self-reinforcing—a positive feedback loop. A sketch of the loop underlying Balzac's statements is shown in Figure 2.16.

The concept of a vicious circle in this sense is now so universal that it would be hard to trace the extent of its application or to document its very first appearance. We shall see it developed into a serious working tool for the social sciences in the work of Gunnar Myrdal (see section 2.5). We should note, however, that such a process never occurs by itself in reality. There are always constraints that prevent a self-reinforcing process from expanding itself beyond all bounds. And

since such processes dominate at some times and not at others, there must be influences that shift loop dominance between positive and negative loop processes. That is, there must be structural changes, nonlinearities, or external influences that affect the interplay between self-reinforcing and self-opposing loops. However, in the early appearances of self-reinforcing vicious circles, such as Balzac's statement above, only the positive loop is described.

Self-Referring Statements

The second logic loop idea eventually to influence some modern scholars' views of feedback is associated with classic paradoxes. Perhaps the earliest version is the so-called paradox of Epimenides, or the "liar's paradox." Epimenides is supposed to have said,

> All Cretans are liars.

Not so bad by itself, but a disturbing paradox appears with the realization that Epimenides himself was a Cretan. His statement is akin to saying "This statement is false." The paradox is that we cannot decide whether "This statement is false" is true or false. Either option leads to its own contradiction. If the statement is true, then it must be false, and if it is false then it must be true. In Epimenides's version, if it is true that all Cretans are liars, then Epimenides is a liar and is presumably lying

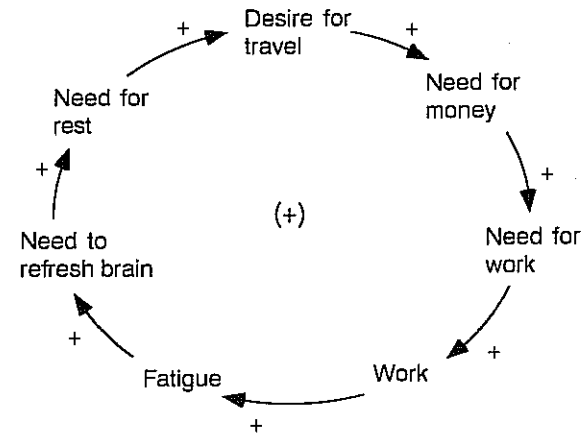

FIGURE 2.16: The vicious circle contained in Balzac's (1850) statement, represented as a positive feedback loop.

when he made his statement, so his statement would be false—a contradiction of what we supposed in the first place.

The problem in such statements is that they refer to themselves. Because of that self-reference, Hofstadter (1979) has called such contradictory statements and arguments "strange loops." Their strangeness results from the fact that they form loops of assertions that close back upon themselves and in the process contradict themselves. In a sense, the paradox of Epimenides is the opposite of a vicious circle. A vicious circle argument supports itself and therefore can be self-consistent and yet be completely false. The Epimenides paradox undermines itself. It is completely self-inconsistent and can be neither true nor false.

Since the time of Epimenides there have been many variations of his paradox, all deriving their paradoxical nature from the fact that they refer to themselves. "The barber of Seville," it is said, "shaves everyone in Seville who doesn't shave himself. Who shaves the barber of Seville?" Such logical conundrums would probably have remained at the level of children's puzzles were it not for the development of set theory in mathematics and a logical paradox of deep significance posed within it by Bertrand Russell.

Russell's paradox grows out of an exquisite argument used by the mathematician Georg Cantor (c. 1871) to prove that no set can be put into a one-to-one correspondence with all of its subsets. The conclusion from Cantor's theorem is that there are somehow "more" subsets of integers than there are integers themselves. There are "more" intervals on a line than there are points on it. There are, in short, different "orders" of infinity—a conclusion that is deeply disturbing to many. The method of proof was almost as disturbing as the conclusion itself. It was an indirect argument in which the contradiction that proved the theorem took the form "statement p is true if and only if statement p is not true." The mathematical world was stunned by the theorem and its proof, and many argued that Cantor's contradiction argument was somehow seriously flawed. Some argued that the fault lay in Cantor's concept of "set" itself, which he had characterized neatly as "a multitude conceived of as a one." Perhaps, some said, mathematicians should not be allowed to conceive of such a thing.

Russell's Paradox

Russell took Cantor's idea and turned it into a paradox casting doubt on the very foundations of mathematical thought. Russell merely said, consider the set R defined to be "the set of all sets in the universe that

are not members of themselves." Now it is hard to imagine a set that *is* a member of itself, so R ought to be a fine, healthy set with a lot of perfectly sensible members. Russell then asked, "Is R a member of itself?" It is, as the reader can check, if and only if it is not, and it is not if and only if it is—a thorough paradox.

Notwithstanding the deep difficulties Russell exposed with this paradox, Cantor's argument as he used it is not flawed. There is no internal contradiction in different orders of infinity in the sense of one-to-one correspondences that Cantor developed. Cantor's conception is a sound mathematical creation, and has had fruitful consequences for later mathematicians. Russell's is a genuine paradox, however, and it acquires its paradoxical nature from self-reference.

It was enormously influential. It forced a reconsideration within mathematics of the concept of a "set" and how it can legitimately be used. Outside of mathematics it led to reconsiderations of what are allowable kinds of statements. More than Epimenides or the barber of Seville, Russell's paradox caused people to question the limits of logic and language. Russell's own solution to the dilemmas revealed by the paradox was the theory of "logical types," that in essence barred the use of self-referential statements in logic. Without self-reference, the paradoxical settings of Epimenides, the barber of Seville, and Russell's own paradox can not even be stated, so the paradoxes can not appear.

The grandest and most astonishing result in this line of reasoning about self-referential statements is a theorem and proof in formal logic by Kurt Gödel (1931). Gödel proved that it is not possible to construct an arithmetic proof that shows that arithmetic itself is logically consistent (Nagel and Newman 1956). The result actually applies to any logical system that is at least as logically complex as arithmetic. It is equivalent to saying that if the logical system is rich enough, it will admit true statements that it cannot prove and false statements that it cannot disprove. While enormously significant in mathematical logic, Gödel's conclusion is less important to us than his method of proof. The proof is based on the rigorous construction of a self-referential logical statement that asserts, in essence, that it itself is not provable.[11]

Implications

These self-referential arguments acquire both their contradictory tendencies and their potential logical power from the meanings attributed to their component parts. They form linguistic loops. They connect to lines of thinking in the social sciences through the notion of the dialectic, which is also a linguistic, closed-loop process focusing on mean-

ing—a generalized conversation. Although it may be hard to see in the mathematical formalisms, there are elements of Russell and Gödel in Hegel and Marx.

Furthermore, there are direct links from these self-referential linguistic processes to modern day notions of computer programs that can create copies of themselves or rewrite their own structure—examples of so-called "self-reproducing automata." These are self-referring structures that perceive, act on, and even rewrite their own structure and meanings. Seeing that societies can apparently do these things, some social scientists place the highest significance on the linguistic loop notions of self-reference, self-reflection, self-transformation, and even self-creation.[12]

Such self-referential systems are perilously close to our concept of feedback, particularly if one views feedback, as Wiener phrased it, as "the transmission and return of information." Yet it is important to realize that they are not the same as the positive and negative feedback loops that are the central focus of this study. Circular causal, feedback processes as we have characterized them do not have the potential to be self-contradictory. Loops whose elements are statements or messages can be paradoxical in that fashion, but the primary focus of this investigation is on loops containing variables that can be interpreted as quantities that increase or decrease over time. If an increase in a quantity in such a loop feeds back eventually to produce a decrease in that quantity, we have a negative loop, but not a self-contradictory loop.

We shall see, however, that self-referring, self-contradictory "message loops" influenced how some social scientists interpreted and worked with the feedback concept. In the evolution of the feedback concept traced in this book, I shall take note of that influence when it occurs, but such "message loops" must remain something of a diversion from my main focus on causal loops with positive and negative polarity. In fact, in the evolution of the concept of positive and negative circular causal processes, such potentially contradictory message loops *are* a diversion. As we shall soon see, the idea draws some feedback thinkers away from self-reinforcing and self-correcting loop structures and leads them into discussions of meta-social structure analogous to the meta-language structures that follow from Russell's theory of types. I do not mean to denigrate these ideas. This sort of "diversion" in our investigation may turn out to be the main stream for subsequent social science. While a good treatment of the significance of self-referring linguistic loops is beyond the scope of this study, interested readers could well start with Hofstadter (1979).

2.5 The Loop Concept in Social Science Literature Before 1945

The social sciences themselves contribute two patterns of thinking to the modern conception of the feedback loop. In one pattern, containing relatives of biological homeostasis and goal-seeking negative feedback, the loop idea is only implicit. In the other, positive loops play the major role, and there is frequent explicit reference to "circular processes" and loops of "mutual causality." The two patterns of thinking appear to have remained largely separated until the 1940s when social scientists became aware of the engineer's concept of feedback.

The Loop Concept Implicit in Social Science Literature

It is hard to resist drawing parallels between the self-equilibrating economy of Adam Smith's *Wealth of Nations* (1776) and the burgeoning diversity of feedback devices of the late 1700s. Indeed, Otto Mayr, an authority on the early history of the feedback concept, has investigated the connections in detail and concludes that Smith's work "implies a conception of the closed causal loop that is in principle the same as the feedback loop" (Mayr 1970, p. 129).

Before investigating this claim, we should note that the notion of a self-regulating system appeared even earlier in the social sciences in the writings of David Hume, one of Smith's closest friends. In his essay "On the Balance of Trade" (1752), Hume argued that there is a basic law operating to keep international trade in equilibrium. His argument was phrased in the following thought experiment, surprisingly similar to an engineer's test of a step input to a control system:

> Suppose four-fifths of all the money of Great Britain to be annihilated in one night, and the nation reduced to the same condition, with regard to specie, as in the reigns of the Harry's and Edward's, what would be the consequence? Must not the price of all labour and commodities sink in proportion, and everything be sold as cheap as they were in those ages? What nation could then dispute with us in any foreign market, or pretend to navigate or to sell manufactures at the same price, which to us would afford sufficient profit? In how little time, therefore, must this bring back the money which we had lost, and raise us to the level of all the neighbouring nations? Where, after we have arrived, we immediately lose the advantage of the cheapness of labour and commodities; and the farther flowing in of money is stopped by our fullness and repletion (Hume 1752, cited in Mayr 1971).

The negative feedback loop is this balancing argument is vivid to a modern reader; see Figure 2.17. Hume, however, made use of a physi-

cal analogy to explain the self-regulation. "All water," he said, "wherever it communicates, remains always at a level . . . [W]ere it to be raised in any one place, the superior gravity of that part not being balanced, must depress it, till it meet a counterpoise." He therefore concluded that "it is impossible to heap up money, more than any fluid, beyond its proper level" (Hume 1752, cited in Mayr 1971).

The Wealth of Nations

It is not clear whether Adam Smith thought in terms of the same fluid balancing metaphor or whether he had some other self-regulating mechanism in mind. Yet in his hands the concept of a self-regulating socioeconomic system became a real working tool, applied to a wide variety of situations. Mayr (1971) cites three particularly outstanding examples, each of which can be easily translated into a negative feedback loop structure. First is Smith's famous argument that the contributions and rewards of all occupations in an economy must be the same:

> The whole of the advantages and disadvantages of the different employments of labour and stock must, in the same neighborhood, be either perfectly equal or continually tending to equality. If in the same neighborhood, there was any employment evidently either more or less advantageous than the rest, so many people would crowd into it in the one case, and so many would desert it in the other, that its advantages would soon return to the level of other employments. This at least would be the case in a society where things were left to follow their natural course, where there was perfect liberty, and where every man was perfectly free both to chuse what occupation he thought proper, and to change

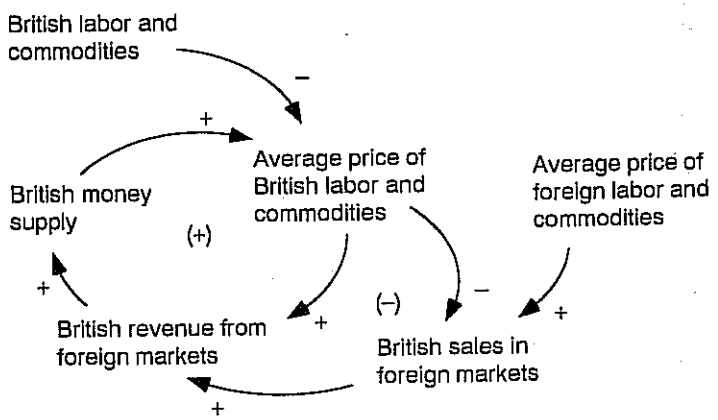

FIGURE 2.17: Negative feedback loop implicit in David Hume's "On the Balance of Trade" (1752).

it as often as he thought proper. Every man's interest would prompt him to seek the advantageous, and to shun the disadvantageous employment (Smith 1776, p. 99).

Smith thus argued that the "relative attractiveness" of all occupations ought to tend to be the same. (The term is from Forrester's *Urban Dynamics* (1969), appearing almost 200 years after *The Wealth of Nations* and applied to land areas not occupations, but the concept is the same.) Smith intended the attractiveness of a given occupation—the "whole of its advantages and disadvantages"—to be defined broadly, including risk, cost of training, and spiritual as well as material rewards. The result is an expression of the tendency of supply and demand for labor in a given occupation to be in equilibrium. The underlying loop structure of the argument is shown in Figure 2.18.

Smith gave several other examples of the self-equilibrating tendency of supply and demand, and then stated an abstract, general argument that places the market price at the center of the equilibrating mechanism:

> When the quantity of any commodity which is brought to market falls short of the effectual demand, all those who are willing to pay the whole value . . . cannot be supplied with the quantity which they want. Rather than want it altogether, some of them will be willing to give more. A competition will immediately begin among them, and the market price will rise more or less above the natural price (Smith 1776, p. 56).

Smith argued similarly in the opposition situation: when supply exceeds demand, the price must go down. Thus the market price is a

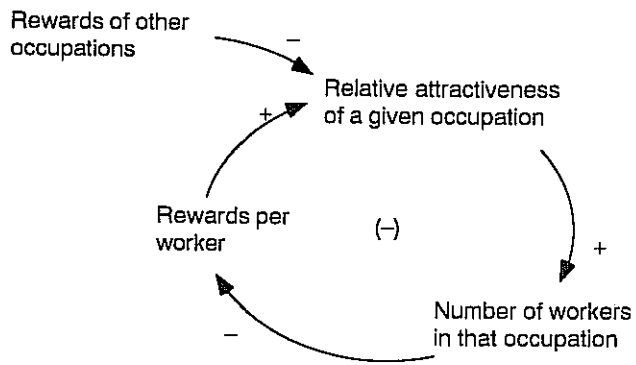

FIGURE 2.18: Negative feedback loop implicit in Adam Smith's argument that the rewards in all occupations must tend to be the same.

function of supply and demand. One could formulate this functional relationship in terms of a negative feedback loop, but the most obvious loop implicit in Smith's argument emerges from the effect of price on production:

> If . . . the quantity brought to market should at any time fall short of the effectual demand, some of the component parts of its price must rise above their natural rate. If it is rent, the interest of all other landlords will naturally prompt them to prepare more land for the raising of this commodity; if it is wages or profits, the interest of all other labourers and dealers will soon prompt them to employ more labour and stock in preparing and bringing it to market. The quantity brought thither will soon be sufficient to supply the effectual demand. All the different parts of its price will soon sink to their natural rate, and whole price to its natural price (Smith 1776, p. 57).

Similarly, he argued that production would decrease if supply exceeded demand. The result is the negative loop shown in Figure 2.19: any persistent discrepancy between the current market price and the "natural price" would generate pressures that would push production in the direction that would decrease the discrepancy. Smith concluded that in the absence of disturbances the system should equilibrate with the market price equal to the natural price of the commodity:

> The natural price, therefore, is, as it were, the central price to which the prices of all commodities are continually gravitating. Different accidents may sometimes keep them suspended a good deal above it, and sometimes force them down even somewhat below it. But whatever may be the obstacles which hinder them from settling in this center of repose and continuance, they are constantly tending towards it (Smith 1776, p. 58).

Smith then observed that disturbances could combine with the equilibrating tendency of the system to produce "occasional and temporary fluctuations in the market price of any commodity" (Smith 1776, p. 59). It is very tempting to believe that this conclusion came from similarities Smith observed between the behavior of prices and the oscillatory "hunting" behavior of some of the regulated machines of his day. Smith may have perceived intuitively that both are in some similar sense controlled, and that controls can produce oscillations. (Maxwell, as we have seen, later shows mathematically why such oscillatory behavior can arise in a negative feedback system.)

One of Smith's applications of his general theory of the self-regulating nature of supply and demand concerns the size of the working population itself.

> The demand for men, like that for any other commodity, necessarily regulates the production of men; quickens when it goes on too slowly, and stops when it advances too fast (Smith 1776, p. 80).

The mechanism of population adjustment for Smith was not the voluntary control of birth rates, but rather involuntary control through infant mortality. If the supply of workers exceeds demand, wages would be low, he argued. Poor living conditions would cause infant mortality to rise and reduce the growth rate of the working class. Eventually—it would presumably take a generation or two—the supply of workers would be reduced to the demand. On the other hand, if workers were in high demand, wages would tend to remain high. Conditions would favor the survivability of infants and children, and the growth rate of the worker population would increase. Some twenty years later, Thomas Malthus (1798) published his famous pessimistic essay addressing the same subject from a very similar point of view (see below). Whether the argument is considered right or wrong is less interesting for us than its clear relationship to the feedback concept.

The feedback loop is thus implicitly but strongly present in Smith's thinking. It is natural to wonder to what extent he was influenced by the growing diversity of mechanical control devices that flourished toward the end of the eighteenth century. Certainly, he saw some connections between human systems and machines. In an early philosophical work, for example, he had produced the following delightful definition of a "system" as an "imaginary machine":

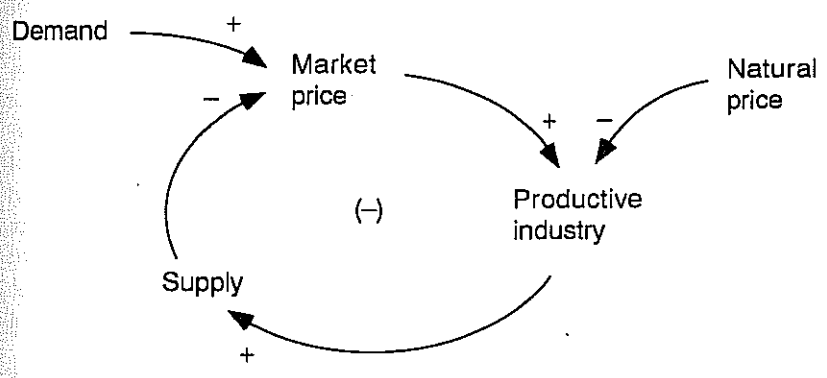

FIGURE 2.19: Negative feedback loop structure of Adam Smith's general theory of the equilibration of supply to demand.

Systems in many respects resemble machines. A machine is a little system, created to perform, as well as to connect together, in reality, those different movements and effects which the artist has occasion for. A system is an imaginary machine invented to connect together in the fancy those different movements and effects which are already in reality performed (Smith, cited in Mayr 1971, p. 17).

Furthermore, the notions of "living machines" and "living automata" trace back at least as far as the "Monadology" of Leibniz (1714).

However, to my knowledge, in *The Wealth of Nations* Smith never used an explicit analogy between his mechanisms of socioeconomic self-regulation and the automatic control devices of his day. Comparing the spread of the ideas of economic liberalism and feedback control of machines, Mayr concludes:

It would be wrong to interpret one of these developments as a direct consequence of the other. A theoretician like Adam Smith may have observed feedback devices in operation, and this may have made his analysis sharper and his formulations more concrete, but the beginnings of economic liberalism lie in the early 18th century, antedating the breakthrough of the feedback concept in technology. On the other hand it is also highly unlikely that 18th-century inventors should have obtained their conceptions by materializing abstract theories of political economy (Mayr 1971, p. 129).

The most that I can say is that self-regulation appears to have been simply a part of the spirit of the times in late-eighteenth-century Britain.

Self-regulation in the Federalist papers

Some support for that conclusion can be seen in other writings of the time. Platt (1966) argues persuasively that the "checks and balances" assiduously written into the Constitution of the United States were a conscious effort to design a system of "stabilization feedbacks." Evidence for the claim can be found in the *Federalist* papers, a series of 85 newspaper articles written by Alexander Hamilton, John Jay, and James Madison to persuade the public to favor the new Constitution.

One of the concerns in these papers, for example, is instability. Madison argued that pure democracy of the Athenian sort is prone to turbulence:

A common passion or interest will, in almost every case, be felt by a majority of the whole; a communication and concert results from the form of government itself; and there is nothing to check the inducements to sacrifice the weaker party or an obnoxious individual. Hence it is that such democracies have ever been spectacles of turbulence and contention; have ever been found incompatible with personal security or the rights of property; and have in general been as short in their lives as they have been violent in their deaths (*The Federalist*, cited in Platt 1966, p. 112).

"Nothing to check the inducements to sacrifice the weaker party or an obnoxious individual" sounds like an uncontrolled positive feedback loop. The solution was not to abandon the goal of democratic government, but to design it so that its tendencies toward instability were countered by the form of government itself. The key to that design was the self-interest of the participants:

The great security against a gradual concentration of the several powers in the same department, consists in giving to those who administer each department the necessary constitutional means and personal motives to resist the encroachments of the others. . . . Ambition must be made to counteract ambition. The interest of the man must be connected with the constitutional rights of the place (*The Federalist*, cited in Platt 1966, pp. 112–113).

Smith found socioeconomic self-regulation in the dictum that "Every man's interest would prompt him to seek the advantageous, and to shun the disadvantageous." The writers of the United States Constitution sought to use similar, powerful self-interests to assure the self-regulation of government. Again the argument sounds vaguely loop-like: governments exist to control the self-interests of people, while the same sort of personal self-interests control government.

Maxwell (1868) showed, and others before him knew intuitively, that controls do not guarantee stability and, in fact, can generate their own instabilities. Platt (1966, pp. 114–115) observes that the Federalists had a sophisticated view of the requirements for control with stability. In modern terminology, they emphasized different "time constants" of response to different sorts of disturbances. Executive actions required "energy and dispatch." The House of Representatives, with its two-year terms, was designed for moderately rapid response to the changing will of the people and assigned duties accordingly, such as the initiation of taxation and spending legislation. The Senate, with its six-year terms, was deliberately focused on longer term issues and adjustment. The Supreme Court, with its lifetime appointments, was designed for very long-term stability and independence from disturbance. Finally, the amendment process itself had to be designed to avoid "that extreme facility which would render the Constitution too mutable; and that extreme difficulty which might perpetuate its discovered faults."

The drive for a governmental structure that naturally regulates itself

is evident in the *Federalist* papers. Yet the implicit feedback-loop nature of Adam Smith's arguments is even less apparent here. Regulation may have been in the spirit of the times, but the concept of the feedback loop remained hidden in the spirit's shadow.

Thomas Malthus

In his essays on population, the gloomy parson Thomas Malthus (1798) came remarkably close to an explicit feedback view of population dynamics. It may appear dangerous to say so, for Malthus is generally regarded as having been "disproved" by subsequent history. Nonetheless, in the structure of his arguments Malthus correctly illuminated the potential for exponential growth contained in what we would now call a positive feedback loop. In addition, he correctly identified several plausible negative loops that strive ultimately to control that potentially runaway loop. Whatever predictions he is considered to have made that have been judged false, his implicit feedback view is fundamentally sound.

Malthus was addressing himself to a philosophical discussion of his day, on the "perfectability of man and of society," and he saw a problem.

> Population, when unchecked, increases in a geometrical ratio. Subsistence increases only in an arithmetical ratio. . . . By that law of our nature which makes food necessary to the life of man, the effects of these two unequal powers must be kept equal. This implies a strong and constantly operating check on population from the difficulty of subsistence. This difficulty must fall somewhere; and must necessarily be severely felt by a large portion of mankind (Malthus 1798, p. 14).

The specter of this difficulty made it hard for Malthus to accept promises of utopias that some in his time were describing, "where all narrow luxuries would be contemned; where [people] would be employed only in collecting the necessities of life; and where, consequently, each man's share of labour would be light, and his portion of leisure ample" (1798, p. 11).

Why does population grow geometrically, or, as we might say, exponentially? The mechanism, which Malthus thought too obvious or too graphic to state blatantly, is that babies tend to become parents:

> Population, could it be supplied with food, would go on with unexhausted vigour, and the increase of one period would furnish the power of a greater increase the next, and this without any limit (p. 107).

The positive feedback loop from births to population and back to births is almost explicit. Births add to population and "furnish the power" for even more births and more population, potentially ad infinitum. Food production does not have this self-reinforcing property, however. Acreage can be added into the land cultivated for food production, but the increase of one period does not "furnish the power of a greater increase the next." Indeed, there would be less remaining land available, and probably of lesser productivity. This is the basis of Malthus's claim that food production tends to grow "arithmetically," while population grows "geometrically." An additional acre is just a one-time increase in potential food production. An additional person increases the potential for future increases in population. Malthus is completely correct in these observations.

If population growth has this self-reinforcing character, why has it not long since surpassed the ability of the land to support it? Malthus proposed two types of mechanisms, which act continually to control population growth:

> a foresight of the difficulties attending the rearing of a family, acts as a preventative check; and the actual distress of some of the lower classes, by which they are disabled from giving the proper food and attention to their children, acts as a positive check, to the natural increase of population (pp. 62–63).

Preventative checks are societal responses, voluntary in some sense, even if imposed by culture and tradition; positive checks are all those "vices" and "miseries" that shorten life. Preventative checks therefore work on the birth rate; positive checks work on the death rate. These checks regulate population, keeping it in line with the food production. It is tempting to add "much like Watt's governor regulated the speed of a steam engine," but Malthus, like Adam Smith, did not make such a connection explicit.

It is also tempting to sketch the feedback structure of Malthus's argument, as shown in Figure 2.20. We must be careful about rewriting history here, however. Malthus himself did not mention the closed-loop nature of his argument. He came close to an explicit statement of a feedback loop in his description of the cause of exponential population growth, cited above, but the closed causal loops in his preventative and positive checks were hidden in the language of control and regulation common to his time.

Malthus argued that the structure he described operated continually to hold down the potentially explosive growth rate of population. He suggested that the preventative and positive checks imposed should result in oscillatory ebbs and flows of population growth:

The constant effort towards population, which is found to act even in the most vicious societies, increases the number of people before the means of subsistence are increased. The food therefore which before supported seven millions, must now be divided among seven millions and a half or eight millions. The poor consequently must live much worse, and many of them be reduced to severe distress. The number of labourers also being above the proportion of the work in the market, the price of labour must tend toward a decrease; while the price of provisions would at the same time tend to rise. The labourer therefore must work harder to earn the same as he did before. During this season of distress, the discouragements to marriage, and the difficulty of rearing a family are so great, that population is at a stand. In the mean time the cheapness of labour, the plenty of labourers, and the necessity of an increased industry amongst them, encourage cultivators to employ more labour upon their land; to turn up fresh soil, and to manure and improve more completely what is already in tillage; till ultimately the means of subsistence become in the same proportion to the population as at the period from which we set out. The situation of the labourer being then again tolerably comfortable, the restraints to population are in some degree loosened; and the same retrograde and progressive movements with respect to happiness are repeated (pp. 63–64).

Two things about this argument are interesting from our point of view. First, it is a feedback argument: it begins and ends with population and contains in between a number of closed causal loops, as shown in Figure 2.21. Second, it argues for the potential for oscillations to arise from the controlling influences on population. It calls to mind the phenomenon of "hunting" that troubled eighteenth-century designers

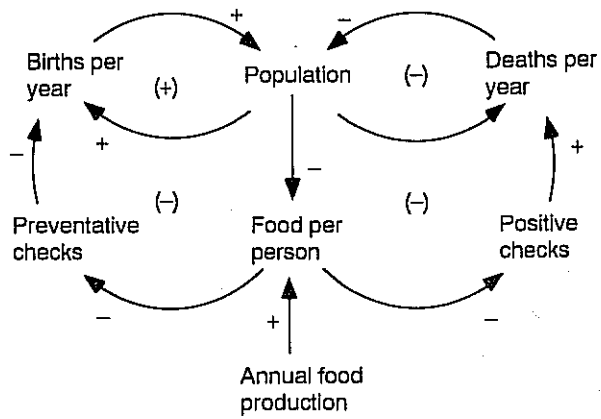

FIGURE 2.20: Implicit feedback structure of Malthus's theory of population growth.

of governors for steam engines. Although it seems a very likely connection, we can, however, again only speculate on whether Malthus was influenced by the feedback devices of his day.

As noted, the arguments of Malthus reflect some of the thinking of Adam Smith. Both held that low wages would suppress population growth and bring it more in line with subsistence. Both, in fact, pessimistically assumed that control would tend to come predominantly through positive checks, such as rising infant mortality, rather than voluntary preventative checks such as late marriages or birth control. Malthus carried the argument further, however, by exposing the engine of population growth, the self-reinforcing positive feedback loop that exists in the nature of the parenting process. Smith focused on the interrelationships adjusting the supply of labor to demand. His theories contain only negative loops. Malthus focused on the mechanisms that could conceivably control runaway growth caused by an unavoidable positive loop. Both argued from the same fundamental philosophical view, however: socioeconomic systems are, by their nature, self-regulated in something approaching an automatic sense, with or without the conscious action or acquiescence of people. In the work of both Smith and Malthus, the closed-loop nature of their arguments and their relationship to automatic control in machines remained implicit.

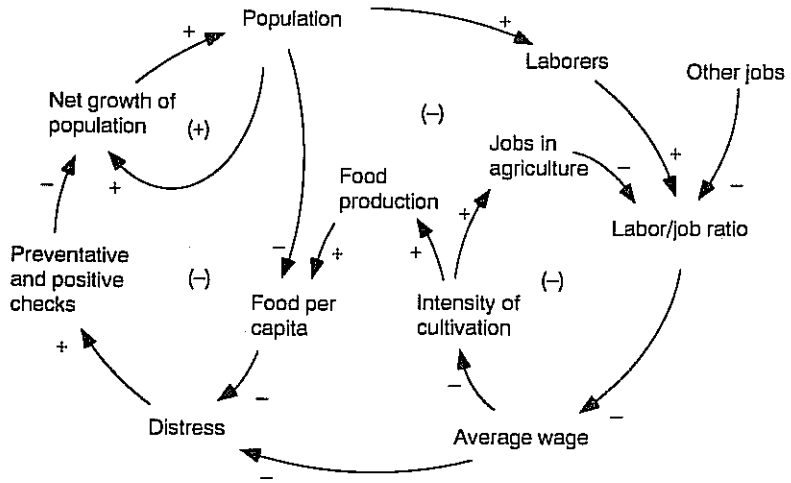

FIGURE 2.21: Closed-loop structure of Malthus's argument suggesting oscillatory tendencies in population growth.

The first unequivocal recognition, that I know of, of a connection between automatic control in engineering devices and self-regulation in living systems is contained in a famous essay sent to Darwin by Alfred Russel Wallace, the "other man" most responsible for the theory of evolution. In it, Wallace writes:

> The action of this principle [the struggle for existence] is exactly like that of the steam engine, which checks and corrects any irregularities almost before they become evident; and in like manner no unbalanced deficiency in the animal kingdom can ever reach any conspicuous magnitude, because it would make itself felt at the very first step, by rendering existence difficult and extinction almost sure to follow (Wallace, cited in Bateson 1972, p. 428).

Wallace's analogy here is something like the discovery of the Americas by Norsemen: only much later did other discoverers open the way for widespread exploitation of the idea.

By the 1900s, analogies between the regulation of machines and socioeconomic systems were made more frequently, and there were attempts to derive substantive implications. Albert Aftalion (1909, 1927), for example, suggested that business cycles stemmed from actions of entrepreneurs that were analogous to attempts to maintain manually the temperature of a furnace. When the temperature is very low, there is a tendency to put in more coal than is necessary in steady state, thereby causing the temperature to rise too high. High temperatures, on the other hand, tend to cause people to wait too long to add more coal, and the temperature drops too far before the need for more coal is perceived. Similar under- and over-reactions in production and ordering decisions, Aftalion argued, could cause economic cycles (Goodwin 1951a; see section 3.3).

The Wealth of Nations was instrumental in changing the way people thought about economic phenomena. It helped smooth the way for "the system of natural liberty" (modern capitalism), and it dramatically advanced economics as a field of scholarly study. Most importantly for us, subsequent economic theoreticians adopted the style of Smith's analyses, including arguments phrased implicitly or explicitly in terms of causal loops. Mayr (1971) observes that the work of David Ricardo, for example, contains numerous loops and is easily translated into the mathematics of feedback systems. (We shall see other examples in the work of other economists in section 5.3.) As in Smith's work, the feedback loops implicitly deal with regulation and are almost always negative.

Other implicit examples of the feedback loop concept

In the generic negative feedback loop, a discrepancy between the actual state of a system and some desired condition results in action designed to reduce the discrepancy. In one form or another this pattern has appeared throughout the history of the social sciences. If one sets out to use the feedback perspective as a kind of lens through which to view older theories, one can invariably see them as implicit feedback theories. I will briefly sketch some additional examples and then comment on what we can learn from such an exercise.

The dialectic of Hegel, and Marx's variation on that theme, contain "discrepancies between desired and actual conditions." The contradictions between thesis and antithesis set up pressures that eventually force a new state of affairs, the synthesis. The common mental picture of the phenomenon, shown in Figure 2.22a, contains no loops. The closed-loop version shown in Figure 2.22b is an attempt to capture the idea that the synthesis emerges from a restructuring of thesis and antithesis. The "desired condition" is synthesis, the elimination of contradiction and conflict between thesis and antithesis. The loops shown in Figure 2.22 are both negative: conflict between thesis and antithesis bring about a restructuring that reduces or eliminates (negates) the conflict.

The loop nature implicit in Marx's views has been pointed out more rigorously by Stinchcombe (1968). In his study of patterns of explanation in the social sciences, Stinchcombe argued that two related kinds of

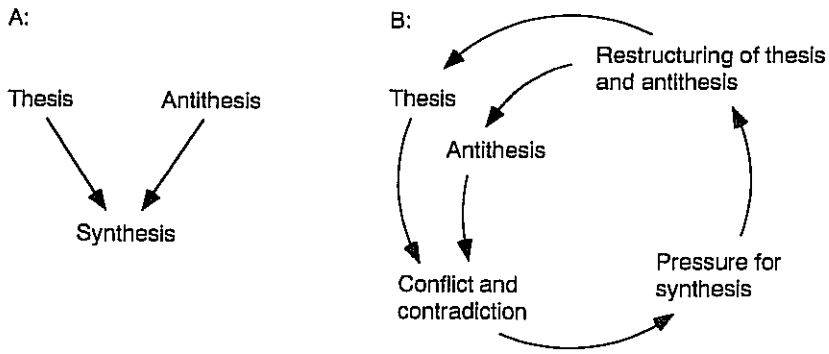

FIGURE 2.22: Dialectic of Hegel and Marx viewed as a feedback structure. a) open-loop view; b) closed-loop view.

causal imagery having a closed-loop character have frequently been employed in the social sciences. In the writings of Marx, the functional anthropologists such as Malinowski and Radcliffe-Brown, and the work of Robert King Merton (among others), he found a circular causal view that he labeled functional causal imagery.

Stinchcombe defined a "functional explanation" to be one in which

> the consequences of some behavior or social arrangement are essential elements of the causes of that behavior (Stinchcombe 1968, p. 80).

He sketched the general causal loop structure of such a functional explanation as shown in Figure 2.23a, and interpreted a particular argument of Marx in those terms with the diagram shown in Figure 2.23b.

In Figure 2.23a, the symbols are defined as

- H: the homeostatic variable, defined as "the consequence or end which tends to be maintained, which in turn functions indirectly as a cause of the behavior or structure to be explained";
- S: social structure or behavior that has a causal impact on H;
- T: tensions or difficulties or other causal forces that tend to disturb or prevent H.

Applied to Marx's argument about power in "bourgeois democracy," the symbols in Figure 2.23b become

- S: parliamentary republicanism as a form of government (Marx's "bourgeois democracy");
- Hn, Hp, Hw: the consequences of parliamentary democracy for the nobles, the bourgeoisie, and the proletariat, respectively;
- Pn, Pb, Pw: the power of the nobles, the bourgeoisie, and the proletariat, respectively.

Figure 2.23b tries to capture the Marxist view that parliamentary republicanism has positive effects on the bourgeoisie and negative effects on the nobles and the proletariat. Each group therefore applies its power either for or against the establishment and maintenance of that form of government. Their efforts determine the evolution of

governmental structure, which is in turn a further cause of their continuing efforts. The closed-loop structure of the argument is clear. Marx expected that the growing power of the proletariat would lead to a workers' revolution, altering the structure of government and the future interplay of power among groups.

From this point of view, then, Marx becomes something of a feedback thinker. Stinchcombe refrained from using the word, however, and may in fact have seen some differences between the closed causal loops of functional explanations and the feedback writings he would have been familiar with in 1968.

Other modern authors have noted that several classical theories of motivation in psychology have a similar negative-loop character. In the drive-reduction theory of C.S. Hull and E.C. Tolman, drive can be thought of as the pressure to move an organism from its current state toward a preferred state. A person is hungry. The hunger drive leads to finding food and eating, which reduces the hunger drive, at least until digestive and metabolic activity exhaust the food consumed. Less physical processes were phrased in similar terms in the theory of tension reduction appearing in psychoanalysis and Gestalt psychology. The loop nature of these processes is clearly discernible, as drives and tensions both affect and are affected by the state of the organism. The generic negative feedback loop is evident from the tendency of such systems to counteract and try to eliminate disturbances from some goal state.

In the conceptual scheme for motivation developed by John Dewey and George H. Mead, stages in the process were labeled impulse,

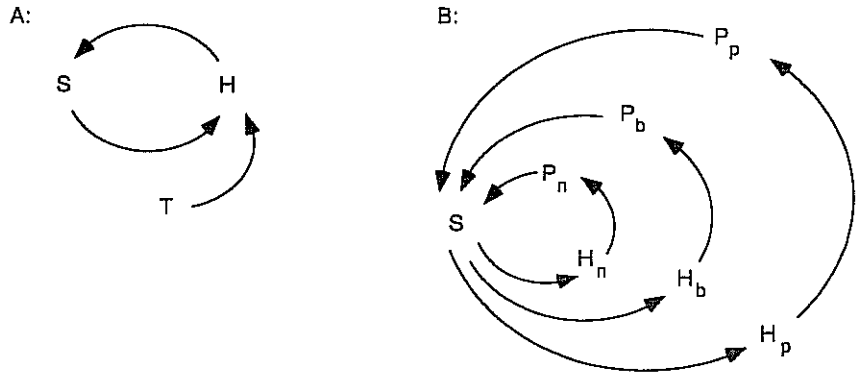

FIGURE 2.23: Circular causal views of functionalist explanations. a) elementary causal structure of a complete functional explanation; b) example of the causal loop structure of Marxian functionalism. Source: Stinchcombe (1968, pp. 89, 94). See text for explanation of symbols.

perception, manipulation, and consummation. Looking in retrospect on their work, one author observed:

> An impulse is a disturbance, any lack of adjustment between an organism and its milieu—pique over an imagined slight, hunger pangs, concern over the whereabouts of a friend who is overdue. An organism is set into motion by a disruption of its steady state; any discomfort leads to attempts to eliminate it. The key principle is that an act, once under way, tends to persist until the discomfort is removed (Shibutani 1968, p. 332).

Again the generic negative loop structure is apparent. A disturbance from steady state leads to impulses, which are perceived by the individual and which lead to manipulating the situation until "consummation," that is, until the impulses signaling the disturbance are no longer perceived. The "manipulations" are what we can observe. Mead emphasized that manipulations give rise to new impulses and new perceptions, which serve to guide the continuing action. It is this emphasis that makes Mead's conception very much a feedback theory of motivation (Shibutani 1968, p. 333).

Finally, in the stimulus-response-reinforcement theories as they developed from Pavlov to Skinner and beyond, modern authors have found implicit suggestions of loop processes. Reinforcement—what follows the response to a stimulus—affects future responses. However, there is a good deal of confusion, some of it semantic and some substantive, about whether positive reinforcement reflects positive or negative feedback. Kramer (1968), for example, considers three rather similar, classic reinforcement situations—a cat repeatedly put in a box and repeatedly scrabbling to get out, Pavlov's dog learning to salivate at the sound of a buzzer, and a rat learning to press a bar to get food pellets. He concludes that the cat's behavior is an example of negative feedback, and the rat's reflects positive feedback up to the point of satiation, while the dog's behavior does not involve any feedback loop at all. I consider the attempt to label an entire complex behavior pattern as exemplifying either positive or negative feedback to be a misguided use of the feedback concept, but nonetheless, Kramer's comments here are a serious attempt to interpret reinforcement theory in feedback terms.

These observations of feedback-like concepts in the social sciences have been repeatedly made by scholars since the emergence of the engineer's concept of feedback in the 1940s. The observations tell us two things. First, they suggest that the feedback concept in one form or another has been implicitly present in the thinking of some of our most respected and influential social scientists, and it underlies a number of the most significant social science concepts and theories. Second, they imply that little is accomplished if all we do is relabel theories with the feedback stamp of approval.

However, even in the most pessimistic view, relabeling social science ideas as feedback theories serves to expose a common and powerful pattern of thinking. I side with Stinchcombe who observed that

> Explicit knowledge accessible to intelligent beginners is obviously more efficient for a science than knowledge perceived by the intuition of its geniuses (Stinchcombe 1968, p. 148).

If loop thinking characterizes the insights of great social scientists, then exposing that loop structure for others to adopt is reason enough to relabel theories as feedback views. In Chapters 3, 4, and 5 we shall be concerned with what people perceive the additional promise of the feedback concept is in the social sciences, once it has been made explicit and well-understood.

Feedback loops in the thinking of John Dewey

One final precursor of the feedback notion in the social sciences falls midway between the implicit loop notions of this section and the explicit loops of the following. In a remarkable ten-page essay published in 1896, John Dewey displayed a conception of the feedback loop in psychological processes that was far ahead of its time (Dewey 1896; Slack 1955). The essay has also been credited with foreshadowing the functionalist school in sociology and anthropology, the Gestalt point of view in psychology, and modern criticisms of behaviorism (Dennis 1948, p. 355). More than sixty years after its publication it contributed significantly to the development in psychology of the feedback idea called the TOTE unit (Miller, Galanter, and Pribram 1960; see section 4.3).

Dewey was disturbed by the rising influence of the reflex arc in psychology. The concept originated in neurophysiology and refers to the neurological path linking a stimulus to a response: stimulus → receptor → afferent nerve → connective fibers → efferent nerve → effector → response (Miller, Galanter and Pribram 1960, p. 22). In Dewey's time, the reflex arc was being groomed as the organizing principle for all psychological phenomena. His primary criticism was that the reflex arc did not solve the old dualism between sensation and idea:

> The sensory stimulus is one thing, the central activity, standing for the idea, is another thing, and the motor discharge, standing for the act proper, is a third. As a result, the reflex arc is not a comprehensive, or organic unity, but a patchwork of disjoined parts (Dewey 1896, in Dennis 1948, pp. 355–356).

What Dewey desired for the organizing principle for psychology was the reality that tied the disjoined parts of the reflex arc together. Taking a cue from physical activity, he termed that reality a "coordination." He envisioned stimulus and response as phases of a coordination. Attention is stimulated by "a conflict within the coordination," an uncertainty about how to complete it. A stimulus is that phase "requiring attention." Motion, as response, is "whatever will serve to complete the disintegrating coordination" (p. 363). The "conflict in the coordination" sounds suspiciously like the feedback engineer's "error signal," or the discrepancy between desired and actual conditions in a negative feedback loop. Indeed, Dewey explicitly argued for a reinterpretation of the reflex arc as a circuit.

To exemplify his arguments, Dewey analyzed two classic situations, William James's example of a child reaching for a candle flame and Baldwin's discussion of a person hearing a loud unexpected sound and reacting by running away. He concluded that the beginning of the candle sequence is not the sensation of light, but the act of seeing, which is an optical-ocular, sensorimotor activity in which movement is the primary element and sensation secondary (p. 356). Then followed a distinctly feedback-like description of the connection between seeing the flame and reaching for it:

> Now if this act, the seeing, stimulates another act, the reaching, it is because both of these acts fall within a larger coordination; because seeing and grasping have been so often bound together to reinforce each other, to help each other out, that each may be considered practically a subordinate member of a bigger coordination. More specifically, the ability of the hand to do its work will depend, either directly or indirectly, upon its control, as well as its stimulation, by the act of vision. If the sight did not inhibit as well as excite the reaching, the latter would be purely indeterminate, it would be for anything or nothing, not for the particular object seen. The reaching, in turn, must both stimulate and control the seeing. The eye must be kept upon the candle if the arm is to do its work; let it wander and the arm takes up another task (p. 356; emphasis added).

The eye stimulates and controls the hand, while the hand stimulates and controls the eye—a clear conception of a closed-loop process involving mutual causality and information feedback.

In analyzing the situation of a person hearing a loud, unexpected sound and running from it, Dewey exposed the circuit he believed to be present.

> Just as the "response" is necessary to constitute the stimulus, to determine it as sound and as this kind of sound, of wild beast or robber, so the sound experience must persist as a value in the running, to keep it up, to control it. The motor reaction involved in the running is, once more, into, not merely to, the sound. It occurs to change the sound, to get rid of it. The resulting quale, whatever it may be, has its meaning wholly determined by reference to the hearing of the sound. . . . What we have is a circuit, not an arc or broken segment of a circle. This circuit is more truly termed organic than reflex, because the motor response determines the stimulus, just as truly as sensory stimulus determines movement (p. 359).

The concept of the reflex arc does not capture this mediating circuit. Consequently, to Dewey it was an inadequate organizing principle for psychology. It is clear that what he wanted was a closed-loop concept closely analogous, if not identical, to the feedback loop. Moreover, he advocated a continuous interplay between elements in the loop. One of his criticisms of the notions of stimulus and response in the reflex arc was their discrete, discontinuous character:

> In its failure to see that the arc of which it talks is virtually a circuit, a continual reconstitution, [the reflex arc] breaks continuity and leaves us nothing but a series of jerks, the origin of each jerk to be sought outside the process of experience itself, in either an external pressure of "environment," or else in an unaccountable spontaneous variation from within the "soul" or the "organism" (pp. 357–358).

In Dewey's view stimulus and response do not follow discretely one after the other in time, but are contemporaneous phases of one and the same coordination (p. 365). They are elements in a closed-loop process in which behavior is continuously controlled by conditions and conditions are in turn continuously altered by behavior.

Dewey thus concluded that the appropriate foundation for psychology was to be found in a "circular coordination" that we would now recognize as the feedback loop:

> It is the coordination which unifies that which the reflex arc concept gives us only in disjointed fragments. It is the circuit within which fall distinctions of stimulus and response as functional phases of its own mediation or completion (p. 365).

Not until 1960 were these suggestions pursued seriously by psychologists (Miller, Galanter, and Pribram 1960; Powers, Clark, and McFarland 1960; see sections 4.3 and 4.7).

Circular Processes and Loops of Mutual Causality

The previous discussion focused on a set of ideas in the social sciences in which the feedback concept is largely implicit. In contrast, another pattern of thinking in the social sciences that contributes to the modern concept of a feedback system makes explicit reference to "circular

processes" and loops of "mutual causality." Explicit loop-like processes appeared in the social sciences as early as the 1840s, as exemplified by the following vivid description of speculation by John Stuart Mill:

> When there is a general impression that the price of some commodity is likely to rise, from an extra demand, a short crop, obstructions to importation, or any other cause, there is a disposition among dealers to increase their stocks, in order to profit by the expected rise. This disposition tends in itself to produce the effect which it looks forward to, a rise of price: and if the rise is considerable and progressive, other speculators are attracted, who, so long as the price has not begun to fall, are willing to believe that it will continue rising. These, by further purchases, produce a further advance: and thus a rise of price for which there were originally some rational grounds, is often heightened by merely speculative purchases, until it greatly exceeds what the original grounds will justify. After a time this begins to be perceived; the price ceases to rise, and the holders, thinking it time to realize their gains, are anxious to sell. Then the price begins to decline: the holders rush into the market to avoid a still greater loss, and, few being willing to buy in a falling market, the price falls much more suddenly than it rose (Mill 1848).

Speculation, as Mill saw it, is a self-reinforcing process—a positive feedback loop (see Figure 2.24). A tendency for the price to rise feeds back to produce a still greater tendency for the price to rise. Mill's description of it is particularly striking because it indicates that he was clearly aware of the closed-loop nature of the phenomenon. In addition, he traces the implications of the closed causal loop in both the rise and the fall of a speculative market. The same self-reinforcing loop operating through trends and expectations tends to exacerbate both rising and falling prices. Finally, Mill's description even contains a hint of the feedback structure that could act to halt the rise of speculative behavior and start its collapse. When the price greatly exceeds the rationally justified price, the lack of support for the high price begins to be perceived. Speculators come to think the price will stop rising, so they start to sell, and indeed the price stops rising and starts to fall. This assumption is captured in the negative loop in Figure 2.24 passing through "perceived speculation."

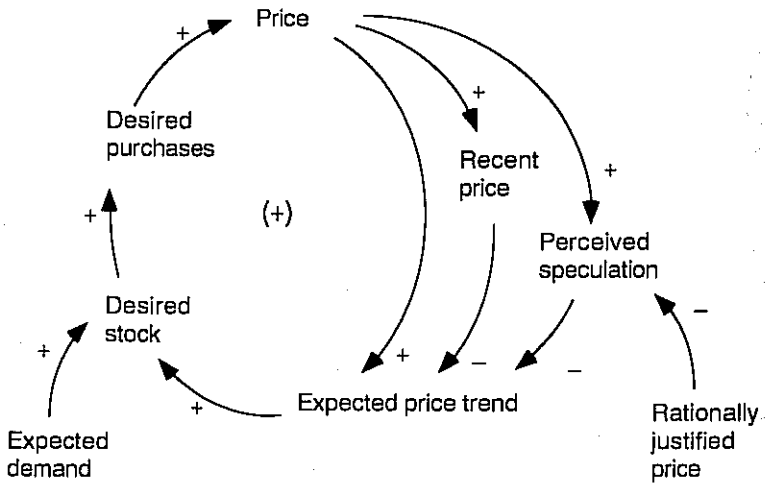

FIGURE 2.24: Mutual causal loops in John Stuart Mill's (1848) description of speculation.

The vicious circle

By the 1900s, observations of circular processes exhibiting loops of mutual or circular causality had become commonplace in the social sciences. There is the vicious circle, for example, in which a bad situation leads to its own worsening. As we have seen (section 2.4), the term actually had its origins in formal logic. Starting from the notion of flawed, circular reasoning, the concept has come to represent an explicitly circular causal process, perceived as characteristically self-perpetuating and self-reinforcing—a positive feedback loop. In that form it has entered the folklore of common conversation.

Myrdal's "principle of cumulation"

In the social sciences, the concept of the vicious circle undoubtedly reached its greatest development in the work of Gunnar Myrdal (1939, 1944, 1957). He re-elevated it from the level of folklore to serious social science and gave it a new name: the "principle of circular and cumulative causation," or more simply, the "principle of cumulation" (Myrdal 1944, p. 75; 1957, p. 23). He preferred the new names because, as he repeatedly observed, the phenomenon can work two ways—in a beneficial sense, as well as in a harmful sense. He recognized that vicious circles could become virtuous. Indeed, the goal of his analyses of American race relations (Myrdal 1944) and the persistent gap between rich and poor nations (Myrdal 1957) was to determine how vicious circles could be turned around.

In the course of the 1,500 pages of his monumental *An American Dilemma* (1944), Myrdal made use of his loop "principle" in no fewer than twenty distinct contexts. At the outset, he stated its central role in his thinking:

A deeper reason for the unity of the Negro problem will be apparent when we now try to formulate our hypothesis concerning its dynamic causation. The mechanism that operates here is the "principle of cumulation," also called the "vicious circle." This principle has a much wider application in social relations. It is, or should be developed into, a main theoretical tool in studying social change (Myrdal 1944, p. 75).

He devoted a section of a chapter and an entire appendix to the development of the idea. Each application of it expressed some aspect of the following general pattern:

Throughout this inquiry, we shall assume a general interdependence between all the factors in the Negro problem. White prejudice and discrimination keep the Negro low in standards of living, health, education, manners and morals. This, in turn, gives support to white prejudice. White prejudice and Negro standards thus mutually "cause" each other (pp. 75–76).

Anticipating later developments, he connected the closed-loop nature of this mutual causality to dynamic behavior (change over time), and in the process came within a hair of inventing on his own the word "feedback":

If, for example, we assume that for some reason white prejudice could be decreased and discrimination mitigated, this is likely to cause a rise in Negro standards, which may decrease white prejudice still a little more, which would again allow Negro standards to rise, and so on through mutual interaction. If, instead, discrimination should become intensified, we should see the vicious circle spiralling downward. The original change can as easily be a change of Negro standards upward or downward. The effects would, in a similar manner, run back and forth in the interlocking system of interdependent causation. In any case, the initial change would be supported by consecutive waves of back-effects from the reactions of the other factor (p. 76).

Myrdal outlined an approach to societal problems that strikingly foreshadowed the work of Forrester and others following in the field that became known as system dynamics. I will quote him at length, because I view his words to be most significant in the evolution of the feedback concept:

It was during this study [*An American Dilemma*] that I first came to realise the inadequacy of the equilibrium approach, and to understand that the essence of a social problem is that it concerns a complex of interlocking, circular and cumulative changes (Myrdal 1957, p. 26).

The main scientific task is, however, to analyse the causal interrelations within the system itself as it moves under the influence of outside pushes and pulls and the momentum of its own internal processes. . . . The scientific ideal is not only to split the factors into their elements and to arrange them in this way, but to give for each of the elements quantitative measures of its ability to influence each of the others, and to be influenced itself by changes in other elements within the system or by changes in exogenous forces (1957, p. 30; also 1944, p. 1068).

Ideally the scientific solution of a problem like the Negro problem should thus be postulated in the form of an interconnected set of quantitative equations, describing the movement—and the internal changes—of the system studied under the various influences which are at work. That this complete, quantitative and truly scientific formulation is far beyond the horizon does not need to be pointed out; but in principle it could be made, and I submit that the working out of such a complete and quantitative solution should be the aim of our research endeavors even when they have to stop far away from the ideal (1957, p. 31; also 1944, p. 1069).

It was not as far off as Myrdal thought. In 1944, when he originally made these observations in *An American Dilemma,* the prospects looked bleak indeed. But he reiterated the comments, slightly revised in the form shown here, in *Rich Lands and Poor* (1957). Just one year later, Forrester's "Industrial Dynamics: a Major Breakthrough for Decision Makers" appeared (see section 3.3), demonstrating the practicality of an approach very much akin to the one Myrdal outlined. In two rather separated parts of the social sciences, and unaware of each other's work, these two authors apparently echoed almost precisely each other's thoughts. Myrdal could easily be considered the grandfather, or perhaps stepfather, of system dynamics, as the following further generalizations about his principle of cumulation suggest:

It is useless to look for one predominant factor, a "basic factor" such as the "economic factor" . . . as everything is cause to everything else in an interlocking circular manner (1957, p. 31; also 1944, p. 1069).

If the hypothesis of cumulative causation is justified, an upward movement of the entire system can be effected by measures applied to one or the other of several points in the system; but this certainly does not imply that from a practical and political point of view it is a matter of indifference where and how a development problem is tackled. The more we know about the way in which the different factors are inter-related . . . the better we shall be able to establish how to maximise the effects of a given policy effort designed to move and change the social system.

Nevertheless, it is unlikely that a rational policy will work by changing only one factor. Thus, though this theoretical approach is bound to suggest the impossibility, in the practical sphere, of all panaceas, it is, on the other hand, equally bound to encourage the reformer. The principle of cumulation—insofar as it holds true—promises final effects of very much greater magnitude than the efforts and costs of the reforms themselves (1957, p. 32).

These remarks are significant in our story for several reasons. First, they show a concept, the vicious circle, that had long since passed from formal logic into folk wisdom, transformed into a serious and powerful analytical tool for the social sciences. Second, they place the concept of the feedback loop at the foundation of social system dynamics. Third, they urge a connection, which we have seen prepared in sections 2.1 and 2.2, between the loop concept of mutual causality and the formulation of mathematical models. They reveal Myrdal's perception of the need for formal models embodying the feedback point of view in policy analysis. Fourth, they advocate a dynamic view, not only of the movement of a system but also its "internal changes." As we have seen in section 2.2, such internal changes can be captured in formal models by nonlinearities.

It seems reasonable that Myrdal, as an economist, was thinking of the econometric tools being developed in the 1930s and 1940s, and he was proposing their use in more general social science policy analyses. If so, then he had clearly made the connection between the loop concept of mutual causality and the structure of mathematical models of the sort described in section 2.2, an important step apparently taken independently of developments of engineering ideas in the social sciences in the 1940s. Presumably, he thought of capturing the "internal changes" of a system with exogenous influences common in econometrics, rather than the nonlinearities appearing in biological models. Yet his prescriptions describe Forrester's independent efforts very closely, as we shall see in section 3.3, particularly if nonlinearities are substituted for exogenous influences. One wonders if Myrdal really had in mind endogenous structural change. The impossibility of solving nonlinear problems at the time of *An American Dilemma* could account for his belief that quantitative models of phenomena like discrimination were "far beyond the horizon." At the very least, the similarity of ideas here from different corners of the social sciences suggest that intellectual developments are as much the product of general currents as they are the genius of any one individual. Zeitgeist again.

The bandwagon effect

Another positive loop concept embedded in the folklore of the social sciences is the "bandwagon effect," meaning the tendency of a movement to gain supporters simply because of its growing popularity. Originally, a bandwagon was the first vehicle in a circus parade, trumpeting the arrival of Barnum and company. The concept came to be used figuratively as a conveyance for a "band" of successful political leaders: "When I once became sure of one majority they tumbled over each other to get aboard the band wagon" (T. Roosevelt 1899, in OED 1972).

Eventually, the bandwagon effect was adopted seriously as an economic concept, defined as "the extent to which the demand for a commodity is increased due to the fact that others are also consuming the same commodity" (Leibenstein 1950, p. 189). As with the vicious circle, it is the self-reinforcing, closed-loop character of the bandwagon effect that gives the concept its appeal. That is to say, its perceived significance is directly due to the fact that it is a generic positive feedback loop.

The self-fulfilling prophecy

In the 1900s before engineering control concepts surfaced in the social sciences, the positive feedback loop continued to be rediscovered in a variety of guises. Still close to the level of social science folk wisdom is the notion of the "self-fulfilling prophecy." Robert King Merton (1936, 1948) is responsible for labeling the idea, but as he pointed out it has a long history. He traced it to Marx and Freud and a host of others, singling out particularly the sociologist W.I. Thomas and a famous theorem attributed to him: "If men define situations as real, they are real in their consequences" (Merton 1948). Merton translated the principle concisely as "social belief fathers social reality." In these expres-

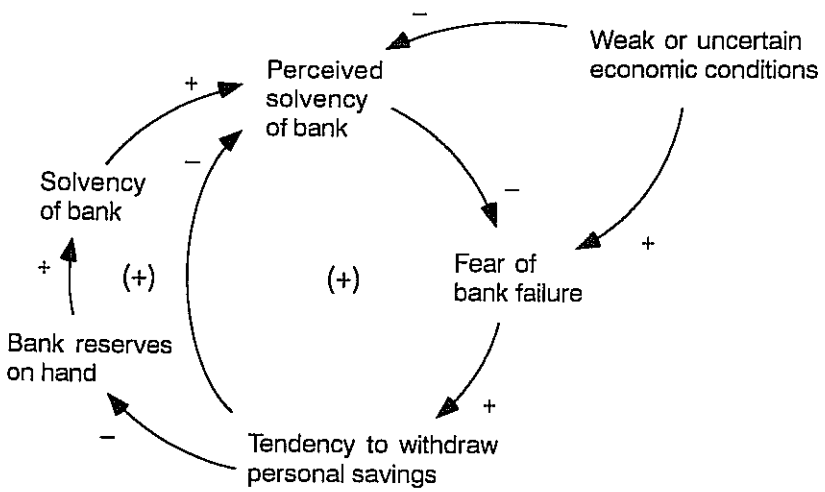

FIGURE 2.25: The self-fulfilling prophecy of a run on a bank in the depression, viewed as a positive feedback loop.

sions, however, it is perhaps hard to see the closed-loop nature of the principle.

As Merton defined it, a self-fulfilling prophecy is an initially false perception of a situation that evokes new behavior that makes the originally false conception come true (Merton 1948). Fearing a collapse of a bank during the depression, people rush to draw out their savings and thereby bring about the very thing they feared. A student, worried sick about his inability to take exams, devotes more time to worrying than studying and consequently fails, just as he knew he would. Two nations, each fearing the war-like moves of the other, stockpile armaments to the point that, as Merton said, "the anticipation of war helps to create the reality." Any of these situations can be viewed as a closed causal loop with a distinctly self-reinforcing character. Figure 2.25 shows the case of a run on a bank as an example. The self-reinforcing property of self-fulfilling prophecies makes the polarity of all such feedback loops positive.

The idea for the self-fulfilling prophecy had occurred to Merton much earlier and appeared briefly in an article addressing the intriguing topic "The Unanticipated Consequences of Purposive Social Action" (Merton 1936). At several points in this essay, Merton thinks in loops. For example, one of the reasons he cites for unanticipated consequences of an action are its ramifications, which feed back upon the system:

> Action in accordance with a dominant set of values tends to be focussed upon that particular value-area. But with the complex interaction which constitutes society, action ramifies, its consequences are not restricted to the specific area in which they were initially intended to center, they occur in interrelated fields explicitly ignored at the time of action. Yet it is because these fields are in fact interrelated that the futher consequences in adjacent areas tend to react upon the fundamental value-system (Merton 1936, p. 903).

Merton was generalizing here from Weber's conclusion about the Protestant ethic, namely that "active asceticism paradoxically leads to its own decline through the accumulation of wealth and possessions." The negative feedback loop in Weber's conclusion and Merton's generalization are obvious, even if implicit.

The self-fulfilling prophecy made its brief appearance in this article as the alter ego of the "suicidal prophecy," a phrase Merton attributed to John Venn (1888), of Venn diagram fame. Marx, for example, had predicted that capitalism would progressively concentrate wealth and increase the misery of the masses. Merton observed that Marx's own prediction, by stimulating the spread of the organization of labor, had a hand in slowing up or eliminating the phenomenon he predicted

(Merton 1936, p. 904). Merton linked the phenomenon to a circumstance that he said was "peculiar to human conduct":

> Public predictions of future social developments are frequently not sustained precisely because the prediction has become a new element in the concrete situation, thus tending to change the initial course of developments (1936, p. 904).

That is, as we would say, human systems are feedback systems! Merton's thoughts in this essay, emphasizing "self-negative prophecies," eventually led him to the positive loop version of the phenomenon.

Merton (1948) used the self-fulfilling prophecy primarily as a framework in which to view prejudice and discrimination. His brief but perceptive article put a different label on the same self-reinforcing circular causal structure used extensively by Myrdal (1944). Throughout, Merton emphasized that the insidiousness of the phenomenon was that the "facts" of a self-fulfilling prophecy indeed support the circular logic. The results are vicious circles in both of the Elizabethan senses of the phrase: they are harmful, and they are logically flawed.

> When the gentleman from Mississippi (a state which spends five times as much on the average white pupil as on the average Negro pupil) proclaims the essential inferiority of the Negro by pointing to the per capita ratio of physicians among Negroes as less than one-fourth that of whites, we are impressed more by his scrambled logic than by his profound prejudices (1948, p. 185).

The mutual causal process such logic sets up, however, contains no contradiction. Instead, its circular, self-reinforcing character creates a very real spiral of discrimination, poverty, and prejudice.

What breaks the vicious circle of an ongoing self-fulfilling prophecy? Merton's answer was "controls":

> The self-fulfilling prophecy, whereby fears are translated into reality, operates only in the absence of deliberate institutional controls (Merton 1948, p. 194).

Had he linked the control idea to negative feedback, he would have brought the two patterns of loop thinking in the social sciences together. As it was, even though he wrote "The Self-Fulfilling Prophecy" in 1948, several years after the emergence of the feedback concept in the social sciences, he did not make the connection.

Gregory Bateson's "schismogenesis"

Among the most interesting loop concepts in the social sciences prior to the emergence of feedback is Gregory Bateson's anthropological

notion of "schismogenesis," literally, the generation of schism (Bateson 1935, 1936, 1949). One of the reasons it is interesting is that Bateson, through the cybernetics movement, eventually came to a deep understanding of the connections between his thinking and the engineer's concept of feedback. He is a transitional figure who wholeheartedly embraced the new cybernetic notions in the social sciences.

Bateson arrived at his concept of schismogenesis by thinking about possible consequences of two different cultures coming in direct contact with each other. They could fuse together over time into one single culture; one of the cultures (or perhaps both) could be completely destroyed; or both groups could come to persist over time as one large community in a kind of dynamic equilibrium (1935, pp. 64–67). Bateson argued that anthropologists could learn the most from studying this third category. What maintains the cultural differences between nation states in Europe or clans, social classes, castes, or age groups in a society? For that matter, what continually maintains "cultural" differences between the sexes?

Bateson's answer was schismogenesis, which he defined as "progressive differentiation" between cultural groups (1935, p. 68, 1936, p. 175). The word "progressive" is used here in the sense of "self-reinforcing," in much the same way that Myrdal (1944) used the word "cumulative":

> If, for example, one of the patterns of cultural behaviour, considered appropriate for individual A, is culturally labeled as an assertive pattern, while B is expected to reply to this with what is culturally regarded as submission, it is likely that this submission will encourage a further assertion, and that this assertion will demand still further submission. We have thus a progressive state of affairs, and unless other factors are present to restrain the excesses of assertive and submissive behaviour, A must necessarily become more and more assertive, while B will become more and more submissive (Bateson 1936, p. 176).

Schismogenesis is thus a sophisticated label for a class of positive feedback loops.

Bateson identified two main types of schismogenesis, "complementary" and "symmetric." Complimentary schismogenesis is exemplified in the previous quotation linking assertive and submissive behavior in a positive loop.

> But there is another pattern of relationships between individuals or groups of individuals which equally contains the germs of progressive change. If, for example, we find boasting as the cultural pattern of behaviour in one group, and that the other group replies to this with boasting, a competitive situation may develop in which boasting leads to more boasting, and so on. This type of progressive change we may call symmetrical schismogenesis (1936, pp. 176–177).

If one grants that submissive behavior is essentially the opposite of assertive behavior, then the symmetric situation can be thought of as a feedback loop composed of two positive links, while the complementary case contains two negatives (see Figure 2.26). In the latter case, an increase in assertive behavior in one actor produces a decrease in assertive behavior in the other actor (i.e., an increase in submissive behavior), which in turn reinforces the first to exhibit still more assertive behavior. In both cases, of course, the loop polarity is positive, and the process is self-reinforcing and disequilibrating.

Bateson was originally led to the concept of schismogenesis by reflecting on the profound contrast he uncovered between the ethos of men and women of the Iatmul tribe in New Guinea (1936). Yet he found examples of the phenomenon in everything from suburban marriages to politics. One of the common pathologies in modern marriages in Western culture, he noted, is the tendency of one of the partners increasingly to cast the other in the role of parent. The relationship between mother and child is initially completely complementary: fostering on the part of the mother, feebleness on the part of the child. As the child grows up, the pattern of fostering and feebleness may persist, or the mother may come to take vicarious pride in the accomplishments of the son, or the relationship may evolve towards assertiveness and submissiveness with either person playing either role. In any case, Bateson asserted, the mother/son relationship is almost always complementary, and if carried over into the son's mar-

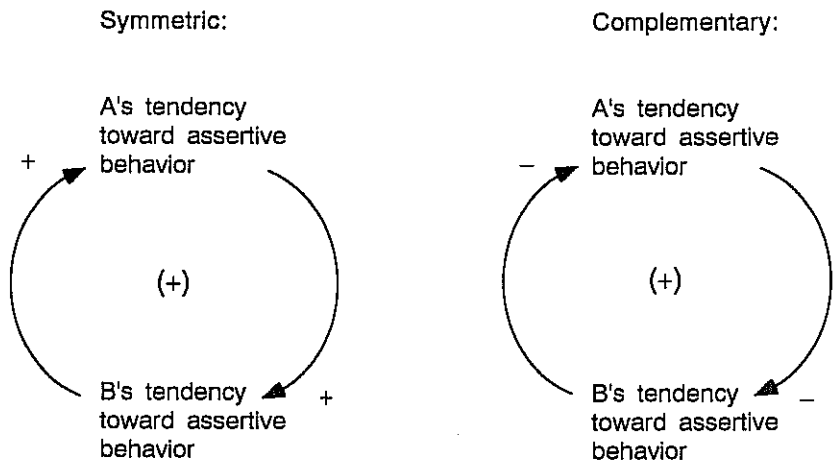

FIGURE 2.26: Symmetric and complementary schismogenesis as positive feedback loops.

riage it can become "the starting point of a schismogenesis which will wreck the marriage" (1936, p. 179). He concluded that viewing such a marital breakdown as cumulative, self-reinforcing, "schismogenic" process would help to explain how the marriage could have been initially satisfactory and yet the breakdown can come to appear inevitable to the participants.

In the area of psychiatry, Bateson hypothesized that the progressive deterioration of a schizophrenic might be due in no small part to the schismogenic character of the person's interactions with others (1936, pp. 179–183). This idea percolated and later emerged, colored by other notions from the field of cybernetics, as the "double bind" theory of schizophrenia (Bateson et al. 1956; see also Reusch and Bateson 1951).

For Bateson politics also has its schismogenic processes (1936, pp. 186–187). He saw symmetric schismogenesis in the intense international rivalries then rife in Europe. He also noted that "class-war" is of the complementary type. Even the megalomaniac dictator and his followers are caught up in a schismogenic process in which the dictator can be pushed closer and closer to an almost psychotic, paranoid state, as his people respond to his behavior in ways that push him toward greater excess. Finally, Bateson suggested that it would be worth investigating the extent to which politicians react to the reactions of their opponents rather than to the more or less objective societal conditions they are supposedly trying to affect.

Controls

Positive feedback loops do not exist without negative loops striving to constrain their tendency toward runaway behavior. Although Bateson did not have a loop concept corresponding to negative feedback, he realized the need for some concept of control or restraint. Cultural differentiation does not usually progress ad infinitum, but reaches a rather steady state. Marriages do not always fail, and even those beset by pathologies of the sort Bateson described sometimes reverse their trend and become healthy. Many schizoid individuals do not experience a complete disintegration of the personality, but maintain a kind of status quo for long periods.

> We must therefore think of schismogenesis, not as a process which goes inevitably forward, but rather as a process of change which is in some cases either controlled or continually counteracted by inverse processes (1936, p. 190).

Bateson thus knew that there if schismogenesis were called a "positive" process, there had to be corresponding "negative" processes.

However, it appears that he did not have a loop view of such inverse processes. The closest he came to identifying constraints as circular processes was in his suggestion that symmetrical and complementary schismogenesis probably never occur in purely separate forms. Every symmetric relationship probably has elements of the complementary in it, and vice versa:

> For example, the Squire is in a predominantly complementary and not always comfortable relationship with his villagers, but if he participates in village cricket (symmetric rivalry) but once a year, this may have a curiously disproportionate effect in easing the schismogenic strain in the relationship (1936, p. 193).

Thus Bateson, using other terms, suggested that any schismogenic situation is a mix of the two loops shown in Figure 2.26. Putting the two loops together results in some circuits that have negative polarity (one negative link) and may act to control the positive, schismogenic loops. A similar structure emerges from his perception of what he called "diagonal relations" between uncles and nephews in the Iatmul. He saw them as a stabilizing mechanism in an otherwise schismogenetic system, but he did not name these negative loop processes. The positive loop remained his focus.

Bateson suggested that schismogenic processes might be controlled by analogous self-reinforcing processes that simply operate in the opposite direction:

> These processes are, like schismogenesis, cumulative results of each individual's reactions to the reactions of members of the other group, but the inverse process differs from schismogenesis in the direction of the change. Instead of leading to an increase in mutual hostility, the inverse process leads rather in the direction of mutual love (1936, p. 197).

Here we have a vicious positive loop counteracted by a virtuous positive loop.

The importance of circular processes

Generalizing from his analyses of schismogenesis in *Naven* (1936), Bateson came to advocate a new definition of the field of social psychology or social anthropology:

> When our discipline is defined in terms of *the reactions of an individual to the reactions of other individuals,* it is at once apparent that we must regard the relationship between two individuals as liable to alter from time to time, even without disturbance from outside. We have to consider, not only A's reactions to B's behaviour, but we must go on to consider how these affect B's later behaviour and the effect of this on A (1936, pp. 175–176, italics in original).

These fields, Bateson thus concluded, should be *founded* on the concept of feedback. He came to this conclusion apparently without any knowledge of developments of the feedback concept in engineering and some eight to ten years prior to its emergence in the social sciences. In slight contrast to previous examples, it is a feedback concept with a discrete character. The behavior of individuals is the concern, and it is viewed much like moves in a game, one behavior following in response to another. Yet Bateson goes beyond a theory merely focusing on the "knowledge of results" of a move. The polarity of the resulting loop of information is critical. Schismogenesis is not merely a circular process; it is a self-reinforcing circular process. It is the association of loop polarity with the knowledge of results of behavior that ties Bateson's work to the feedback concept.

Summary

The loop concept underlying feedback and mutual causality has been present in the social sciences for at least the past 200 years. The frequency and significance of its use has been growing. And it appears in the writing of some of our most distinguished scholars. Negative loops are present implicitly in a great range of explanations in the social sciences. Positive loops appear much more explicitly, as circular processes of self-reinforcing phenomena. With the advent of the engineer's concept of the feedback loop in the social sciences in the 1940s, these two patterns of loop thinking could come together.

In Chapter 3 we shall see how these social sciences ideas interact with notions from engineering, mathematics, formal logic, and biology to produce initially two threads of feedback thinking in the social sciences.

Notes

1. Verhulst (1838), cited in Kormondy (1969, p. 69), and Lotka (1925/1956, p. 66); Pearl and Reed (1920, 1921) and Pearl (1922), cited in Lotka (1925/1956, p. 66). For extensive development and applications see Lotka (1925/1956, pp. 64–76).
2. The dominant polarity can be defined rigorously as the sign of dP'/dP, where $P' = dP/dt$. Here $P' = aP - bP^2$, so $dP'/dP = a - 2bP$, which is positive for $P < a/2b$ and negative for $P > a/2b$. Thus the dominance in Verhulst's population equation shifts from the positive loop to the negative loop when P reaches half its maximum (a/b). See Richardson (1984).
3. See, for example, the citations in H. A. Simon (1957).
4. Expanded versions of these early works appeared in Richardson (1947) and (1960). Richardson (1960) is the source of information for this section.

5. The version given here is from Samuelson (1939), with terminology taken from Baumol (1970, p. 170).
6. The direction of the arrow from Y_t to Y_{t-1} is explained by the observation that as time passes the current value of Y becomes the past value, so Y_{t-1} acquires its value from Y_t.
7. See Graham (1977, pp. 221).
8. Tinbergen's purpose and statistical methods make these circularities a real problem. He wanted to estimate parameters statistically. He had computed the two regressions separately, using ordinary least squares. Without the lag he would have had simultaneous equations, which are now known to require more sophisticated statistical machinery. See Pindyck and Rubinfeld (1976), pp. 126–151.
9. For examples see any econometric text, e.g., Pindyck and Rubinfeld (1976, pp. 266–416).
10. The iterative DO-CONTINUE loop from computer programming can be reasonably termed a logic loop because it can be rephrased in terms of IF-THEN statements.
11. Gödel's argument is a rigorous development that is similar to the paradox of "non-self-descriptive adjectives." Imagine placing all the adjectives in the English language in two lists: those that are "self-descriptive" and those that are "non-self-descriptive." The word "short" describes itself, so it goes in the first list. The word "diminutive" does not describe itself, so it goes in the second list. Now in which list does the adjective "non-self-descriptive" belong? If it belongs in the first list, then it describes itself, so it is non-self-descriptive and therefore belongs in the second list. But if it belongs in the second list, then it is non-self-descriptive so it does describe itself and therefore should go in the first list! The paradox is in the form of the contradiction in Cantor's proof, Russell's paradox, and Gödel's proof: "statement p is true if and only if statement p is not true."
12. See Alker (1981) for a strong argument for such a point of view, as well as extensive references to related social science literature.