In G. E. Lasker (Ed.), *Applied systems and cybernetics, Vol. 2*. New York: Pergamon Press, 712–719. 1981.

069

Feedback, Induction, and Epistemology

One of the most successful notions in control theory has been the principle of negative feedback. As Otto Mayr shows in his delightful book *The origins of feedback control* (1970), practical implementations of the principle go back to the third century B.C., explicitly documented in the case of oil lamps that regulate the flow of oil according to the amount they burn. Today we have thermostats, automatic pilots, and guided missiles. Though these devices differ in structure and material, they have one thing in common: within certain limits they are able to carry out activities that formerly required a human agents attention, discrimination, and judgment. All control mechanisms were designed to free someone's hands or mind for a more important task or, perhaps, just for a more entertaining activity. From the very beginning, their purpose was to maintain or create some state which the designer or user deemed desirable in his or her experiential world — to keep a lamp burning after the slaves were sent to bed, to keep the room at an even temperature regardless of the weather, and so on. All this is taken for granted today, and that is one reason why we are prone to overlook some basic aspects of the phenomenon. As Powers (1978) demonstrated, the embeddedness in the users goal structures has led to a serious misinterpretation of how feedback mechanisms actually function.

> The features I want to focus on here are, first, that "control systems... control input, not output" and, second, that they "cannot be organized around objective effects of their behavior in an external world" (Powers, 1978, p.418).

A control system acts when there is a discrepancy between what it senses (sensory signal) and what it is supposed to sense or would like to sense (reference). The connections that matter are those of certain activities in the system's repertoire with the changes they provoke in certain sensory perturbations. A mechanical feedback device that replaces us in a given task is a crystallized piece of experiential learning. It is the materialization of an if-then rule that has been inductively derived from experience by the designer.

Let us, for a moment entertain the fanciful assumption that the thermostat of an air conditioning system were miraculously imbued with awareness and some cognitive functions so that it could think about and organize its experiential world. It would be a very simple world. The only perceptual discriminations the thermostat could make would be between signals that fall short of the reference value, signals that match it, and signals that are in excess of it. There could be no other perceptual data. On the proprioceptive side, i.e., the system's kinesthetic feedback generated by its own acting, the activity of heating could be discriminated from the activity of cooling. In other words, all the thermostat could come to know in its experiential world would be that it feels too hot or too cold and whether it is at the moment exercising its heating or its cooling activity. The connections between the two kinds of perceptual perturbation and the activities are fixed. These connections are similar in that respect to those implied by reflexes or fixed action patterns in living organisms. Neither in the thermostatic control device nor in the organismic reflex did those connections require learning on the part of the individual system that manifests them. They are wired in, by the designing engineer in the case of the device, and by evolution, through the processes of variation and selection, in the case of the organism.

In a more complex system, however, the connections may be the result of learning. Kenneth Craik, a precursor of cybernetic thinking in the early 1940s, suggested how an elementary form of learning could be mechanized (Craik, 1966). It requires two things: on the one hand, something like a memory, a place where sequences of signals could be recorded to be read at some later point in the experiential flow; on the other, the ability to compare past signals to present ones or to a goal-signal that constitutes a reference value. Once that dual capability is there, the preconditions of inductive learning are satisfied. On this initial level, induction is as simple as it was described by David Hume, more than two hundred years ago (Hume, 1742). All that is needed is the disposition or rule that leads the system to repeat actions that were recorded as successful in its past experience. That is to say, in each occurrence of a perturbation, the system will select the activity that reduced or eliminated that specific kind of perturbation in the past. Implicitly or explicitly, there must be the belief that connections that turned out to be successful, will be successful also in the future. For, as Hume said:

> if there be any Suspicion, that the Course of Nature may change, and that the past may be no Rule for the future, all Experience becomes useless, and can give rise to no Inferences or Conclusions. (Hume, 1742; Essay II, Part 2)

No matter how sophisticated the cognitive functions we hypothetically attribute to an imaginary learning thermostat, it could never do more than establish regularities concerning specific connections between its activities and the subsequently experienced changes of sensory signals. It could not discover that by activating its heating machinery it changes the temperature in the environment. All it could learn would be that its heating activity reduces the sensation of cold and the cooling activity the sensation of heat. It could learn to control its perceptions. That there is an external connection could be specified only by an observer, because from an observer's point of view both the organism and its environment are segments of actual experience. From the organism's perspective, whatever connections are made and whatever regularities are found, are always connections and regularities of its own internal signals.

Cognitive Development

The theory of cognitive development that was proposed and elaborated by Piaget has deep biological roots and builds on presuppositions that are intended to apply to all forms of life. Perhaps the most important among these presuppositions is the assumption that what differentiates living organisms from the rest of the universe is their concern with an inner milieu and their relative ability actively to maintain internal states in equilibrium in spite of external perturbations. All activity — and thus also cognitive activity — is considered adaptive in the specific sense that it serves the purpose of self-regulation (e.g. Piaget, 1967a).

The biological organism does not begin life as a *tabula rasa*. We need not claim that it starts out with god-given Platonic ideas or with genetically transmitted knowledge of an outside world. It is sufficient to assume that the organism has a tendency to act in the face of perturbation. Piaget's key to development, i.e., the increase of internal organization, is the concept of *scheme*. Regardless of whether a scheme is implemented in a reflex or a sophisticated arrangement of cognitive structures, it consists of three parts. First, for instance, there is a pattern of sensory signals which, from an observers point of view, may be considered the effect of an external stimulus; second, there is an activity, triggered by the particular pattern of sensory signals and which an observer may consider a response; third, subsequent to the activity, the organism's experiences some change which, sooner or later, is registered as the consequence of the activity. The consequence is in fact the reason why particular activities are linked to particular perturbations.

On the evolutionary level, natural selection tends to eliminate individuals that have non-adaptive reactions to perturbations from the environment, whereas those that happen to have adaptive reactions survive. Phylogenesis, thus produces results which, considered retrospectively, look as though they were the result of induction: what survives are only those mutants that happen to weather the perturbations of the environment.

On the ontogenetic level, the pattern is similar. The Law of Effect, "Other things being equal, connections grow stronger if they issue in satisfying states of affairs" (Thorndike, 1931), is essentially equivalent to the paradigm

> The living system, due to its circular organization, is an inductive system and functions always in a predictive manner what occurred once will occur again. Its organization (both genetic and otherwise) is conservative and repeats only that which works. (Maturana, 1970; p.39).

For Maturana, speaking as a biologist, the expression "it works" means that, what the system does, successfully eliminates a life-threatening perturbation.

However, the same inductive principle is inherent also in Piaget's concept of *scheme*, but there it is a principle of cognition. Schemes serve not only biological survival but also organisms' cognitive goals whose non-attainment is not fatal. They are part of a theory of learning and incorporate the processes of *assimilation* and *accommodation*.

In order to be activated, a scheme requires the perception of a particular pattern of sensory signals. In actual experience, however, no two situations are quite the same. The sensory pattern that triggers a particular scheme must, therefore, be isolated by the organism in a perceptual field that usually provides vastly more signals than those needed for the particular pattern, and at other times it does not provide all the necessary ones. In other words, differences must be disregarded and this disregarding of differences, so that the pattern can be obtained in spite of them, is called assimilation.

The acting system or organism, does not notice specific differences because it is looking for the signals required to complete a pattern that might trigger a scheme. In contrast, an observer who does register extraneous signals could say that the organism is assimilating (cf. Ch.III).

Sophisticated cognitive organisms, however, have the capability deliberately to disregard such differences and, for them, assimilation becomes a crucial instrument in the construction of regularities and rules, as well as for the practical extension of their schemes. To give an example, if Mr. Smith urgently needs a screwdriver to repair the light switch in the kitchen, but does not want to go and look for one in his basement, he may "assimilate" a butter knife in the context of that particular repair scheme, although he is quite aware of the fact that the butter knife is perceptually and functionally different from a screw driver.

Whenever a scheme is activated, but the triggered activity does not yield the expected result, the discrepancy from the accustomed sequence of events creates a perturbation in the system. As this perturbation springs from the mismatch of an actually sensed situation and an expected one that served as reference, it is equivalent to negative feedback in a cybernetical control loop. It is a novel kind of perturbation. It is not associated with a specific sensory pattern, nor an activity that might eliminate it. However, because it arises as the result of an enacted scheme, it may direct the agent's attention to the sensory material that was present when the scheme was activated (cf. Piaget, 1974a, p.264) and this may then lead to an *accommodation* of the scheme or the formation of a new one (see Ch.III).

As in the case of assimilation, such an accommodation may take place without the agent's awareness, or it may be deliberate. Every time we sit down on an unfamiliar chair, the physical movements that constitute the motor part of our sitting-down scheme may have to be slightly adjusted to fit the particular circumstances, but we usually remain quite unaware of that accommodation. When, on the other hand, we drive a new car, we also have to make certain adjustments: we deliberately accommodate our motor acts and sometimes even construct (by trial and error) novel sub-schemes to fit into, or partially replace, the ones we had.

Such sensory-motor schemes constitute the lowest but nevertheless essential level of cognitive development; and the concepts of scheme, assimilation. and accommodation are no less applicable to the higher levels of cognition.

From the systems point of view, the conception of the scheme with its inherent processes of assimilation and accommodation and the conception of the learning feedback mechanism are analogous and wholly compatible. In both cases, all vital knowledge is constituted by rules that indicate which particular actions are successful in eliminating particular perturbations. No knowledge of an independent external reality is gained, nor is any such knowledge needed.

Analogously to a learning cybernetic system, a living organism must be able to experiment and to construct, by inductive learning from experimental outcomes, a repertoire of schemes that enable it to maintain its sensory perceptions within an acceptable range of the reference values.

The situation is similar to that of biological organisms in the theory of evolution, in that only the viable structures survive, because natural selection does away with those that cannot in some way meet the environmental conditions (Bateson's "explanation").

On the cognitive level, of course, the perturbations are not, as a rule, immediately fatal. Ontogeny provides opportunities for learning, phylogeny proceeds by pruning. In both dimensions the organism meets reality only in its failures. As Warren McCulloch said "To have proved a hypothesis false is indeed the peak of knowledge" (McCulloch, 1965, p. 154). That statement, from the perspective of traditional epistemology, declares the bankruptcy of the discipline. Ever since the pre-Socratics, knowledge was supposed to correspond to a real world. If it did, it was true, if it did not, it was worthless. But already at the time of the pre-Socratics there were doubting voices, and Pyrrho, a little later, formulated the sceptics arguments that have troubled epistemologists until now. How can we ever tell whether or not the pictures our senses convey are accurate and true, if the only way they could be checked is again through our senses? Kant compounded the dilemma by casting doubt on the "thinghood" of objects in our experience, for if space and time are inherent characteristics of the way in which we humans structure our experience, it seems clear there is no possibility at all for us to imagine the structure of a world before we have experienced it. The traditional epistemologist is thus left with no more than Descartes pious hope that God could not have been so mischievous as to equip us with deceptive senses. The radical constructivist theory of knowledge I have suggested (von Glasersfeld, 1976, 197B, 1979, 1981) breaks with tradition and relinquishes the iconic conception of knowledge. If one takes seriously the proposition that the only cognitive contact organisms make with ontological reality is when their schemes to eliminate perturbations break down, one can adopt a more positive albeit less metaphysical view of knowledge. On the level of schemes that involve action, the value of schemes has always been assessed on the basis of whether or not they achieve what they are expected to achieve. In other words, it is a question of know-how that has functional value and, as with all functional values, criteria of economy with regard to effort, speed, cost, elegance, etc., can be added. This, however, is not the only level. With the construction of schemes for the construction of schemes, the first step is made into a virtually infinite hierarchy of levels of reflection and abstraction; and the assessment of cognitive structures and schemes, though it never loses the connection to the functional level of action at the bottom of the ladder, comes to adopt criteria of homogeneity, compatibility, and consistency as it moves up the rungs of abstraction.

The crucial aspect of this theory of knowledge is that the idea of correspondence with reality is replaced by the idea of fit. Knowledge is good knowledge if it fits within the constraints of reality and does not collide with them. That fit manifests itself whenever a cognitive structure, a scheme, a theory, remains viable in the face of new experience or experiments. The epistemology is parallel to Poppers (1962) refutations but it puts the stress on the rather than on their refutation.

The concept of viability, understood as a function of fitting into a scheme, an environment, or an experiential context, is a concept with which we are quite familiar in the realm of linguistic communication. The receiver of a piece of language, be it a word, a sentence, or a text, faces a task of interpretation. As members of a linguistic community, users of the language have formed, in the course of experience, semantic connections between the experiential items that constitute language and other items that they have isolated in the flow of experience. A piece of language, then, may enable the receiver to build up a conceptual structure whose building blocks are wholly subjective. They cannot be anything but subjective, because elements of his or her own experience are the only material a cognitive interpreter has acC8SS to. Insofar as the receiver succeeds in completing the conceptual structure, he or she will consider that the piece of language has been understood. Interpreting a communication, therefore, is the process of weaving a conceptual web such that it satisfies the constraints that are indicated by the received signs or signals. Neither signs, signals, or words can supply the conceptual material to build that web, but they do delimit what is eligible. In that sense they have a selectional function, much as Nature or the environment selects living organisms by eliminating others. In English, for instance, almost every lexical item in a sentence or text has more than one meaning if it is taken as an isolated item. The communication context, however, ordinarily eliminates all but one of the potential meanings (instances of unresolvable ambiguity are remarkably rare). In evolution, what survives, does so, because it has the wherewithal to cope with, and thus to fit into, the environmental constraints. In communication, the result of an interpretation survives and is taken as the meaning, if it makes sense in the conceptual environment which the interpreter derives from the situational context and from his prior experience. The constraints that are inherent in conceptual environments are, of course, far less tangible and definite than degrees of temperature and humidity, speed of locomotion, rate of reproduction, etc., which are the factors that delimit an organisms potential for survival. Nevertheless, opaque though the conceptual conditions may often be, it if they that determine whether or not a word or a sentence can be fitted meaningfully into the web of an interpretation and whether or not that interpretation can be fitted into the context of the interpreters general experience. The point to be emphasized in the present discussion is that neither in the realm of evolution nor in that of interpretation do the constraints specify the actual properties of the items that do or could fit into the allowed space. The constraints merely eliminate what does not fit. Norbert Wiener’s definition of cybernetics hinges on the concepts of control and communication. While he viewed control mechanisms mainly from the perspective of the engineer who uses feedback devices as proxies for himself, he did not stress the epistemological implications that arise if one considers these devices as independent, self-regulating systems. There is, however, no contradiction between the epistemological position I have outlined for a learning feedback system and the use of such a system by its designer, they merely live in very different experiential worlds. With regard to communication. the analysis of interpretation I have given is, in fact, no more than an amplification of Wieners (and Shannon’s) theory. There is, however, a more explicit formulation of the purpose and character of cybernetics. The study of control and communication can be explicated as the endeavor actually or hypothetically to construct possible and plausible contents for black boxes. The formalistic branch of that discipline aims at the development of mathematical models, i.e., networks of functions that mathematically account for and predict observable output from observable input. The more concrete, heuristic branch of the discipline aims at the development of conceptual or physical models that are operationally equivalent to the unobservable mechanisms inside a black box. In both these branches of cybernetics one works towards a fit and not towards an iconic replication. Hence, a model is a good model whenever the results of its functioning show no discrepancy relative to the functioning of the black box. That relation, I contend, is the very same as the relation between our knowledge and our experience and because our experience is the only contact we have with what philosophers call ontological reality, I have suggested that that absolute reality has for us the status of a black box.

References

Craik, K.J.W. (1966). The Nature of Psvchology. Cambridge University Press, Cambridge.

Hume, D. (1963). An Enquiry Concerning Human Understanding (First published 1748). Washington Square Press, New York.

Maturana, H.R. (1970). Neurophysiology of cognition. In P.L.Garvin (Ed.) Cognition: A Multiple View. Spartan, New York.

Mayr, O. (1970). The Origin of Feedback Control. M.I.T. Press, Cambridge, Massachusetts.

McCulloch, W.S. (1965). Embodiments of Mind. M.I.T. Press, Cambridge, Massachusetts.

Piaget, J. (1967). Biologie et Connaissance. Bailimard, Paris.

Piaget, J. (1974). La Prise de Conscience. Presses Universitaires de France, Paris.

Popper, K.R. (1962). Conjectures and Refutations. Basic Books, New York.

Powers, W.T. (1978). Quantitative analysis of purposive systems: Some spadework at the foundations of scientific psychology. Psychological Review, 85: 417–435.

Thorndike, E.L. (1931) Human learning. Century, New York.

von Glasersfeld, E. (1976). Am. Soc. Cybernetics Forum, 3, 115–120.

von Glasersfeld, E. (1978). In Murray (Ed.), Impact of Piagetian Theory. University Park Press, Baltimore.

von Glasersfeld, E. (1979). In M.N.Ozer (Ed.), A Cybernetic Approach to the Assessment of Children: Toward a more Humane Use of Human Beings. Westview Press, Boulder, Colorado.

von Glasersfeld, E. (1981) In P. Watzlawick (Ed.), Die Erfindung der Wirklichkeit. Piper, München.

Preprint version of 25 Feb 2006