6
Responsibilities of Competence*

Heinz von Foerster
Biological Computer Laboratory, University of Illinois

At our last Annual Symposium I submitted to you a theorem to which Stafford Beer referred on another occasion as "Heinz Von Foerster's Theorem Number One". As some of you may remember, it went as follows:

"The more profound the problem that is ignored, the greater are the chances for fame and success."

Building on a tradition of a single instance, I shall again submit a theorem which, in all modesty, I shall call "Heinz Von Foerster's Theorem Number Two". It goes as follows:

"The hard sciences are successful because they deal with the soft problems; the soft sciences are struggling because they deal with the hard problems."

* Adapted from the keynote address at the Fall Conference of the American Society for Cybernetics, Dec. 9, 1971, in Washington, D.C. Published in the *Journal of Cybernetics*, *2* (2), pp. 1–6, (1972).

Should you care to look closer, you may discover that Theorem 2 could serve as a corollary to Theorem 1. This will become obvious when we contemplate for a moment the method of inquiry employed by the hard sciences. If a system is too complex to be understood it is broken up into smaller pieces. If they, in turn, are still too complex, they are broken up into even smaller pieces, and so on, until the pieces are so small that at least one piece can be understood. The delightful feature of this process, the method of reduction, "reductionism", is that it inevitably leads to success.

Unfortunately, the soft sciences are not blessed with such favorable conditions. Consider, for instance, the sociologist, psychologist, anthropologist, linguist, etc. If they would reduce the complexity of the system of their interest, i.e., society, psyche, culture, language, etc., by breaking it up into smaller parts for further inspection they would soon no longer be able to claim that they are dealing with the original system of their choice. This is so, because these scientists are dealing with essentially nonlinear systems whose salient features are represented by the *interactions* between whatever one may call their "parts" whose properties in isolation add little, if anything, to the understanding of the workings of these systems when each is taken as a whole. Consequently, if he wishes to remain in the field of his choice, the scientist who works in the soft sciences is faced with a formidable problem: he cannot afford to loose sight of the full complexity of his system, on the other hand it becomes more and more urgent that his problems be solved. This is not just to please him. By now it has become quite clear that his problems concern us all. "Corruption of our society", "psychological disturbances", "cultural erosion", the "breakdown of communication", and all the other of these "crises" of today are our problems as well as his. How can we contribute to their solution?

My suggestion is that we apply the *competences* gained in the hard sciences—and not the method of reduction—to the solution of the hard problems in the soft sciences. I hasten to add that this suggestion is not new at all. In fact, I submit that it is precisely *Cybernetics* that interfaces hard competence with the hard problems of the soft sciences. Those of us who witnessed the early development of cybernetics may well remember that before Norbert Wiener created that name for our science it was referred to as the study of "Circular-Causal and Feedback Mechanisms in Biological and Social Systems", a description it carried even years after he wrote his famous book. Of course, in his definition of Cybernetics as the science of "communication and control in the animal and the machine" Norbert Wiener went one step further in the generalization of these concepts, and today "Cybernetics" has ultimately come to stand for the science of *regulation* in the most general sense.

Since our science embraces indeed this general and all-pervasive notion, why then, unlike most of our sister sciences, do we not have a patron saint

or a diety to bestow favors on us in our search for new insights, and who protects our society from evils from without as well as from within? Astronomers and physicists are looked after by Urania; Demeter patronizes agriculture; and various Muses help the various arts and sciences. But who helps Cybernetics?

One night when I was pondering this cosmic question I suddenly had an apparition. Alas, it was not one of the charming goddesses who bless the other arts and sciences. Clearly, that funny little creature sitting on my desk must be a demon. After a while he started to talk. I was right. "I am Maxwell's Demon", he said. And then he disappeared.

When I regained my composure it was immediately clear to me that nobody else but this respectable demon could be our patron, for Maxwell's Demon is *the paradigm for regulation*.

As you remember, Maxwell's Demon regulates the flow of molecules between two containers in a most *unnatural* way, namely, so that heat flows from the cold container to the hotter, as opposed to the natural course of events where without the demon's interference heat always flows from the hot container to the colder.

I am sure you also remember how he proceeds: He guards a small aperture between the two containers which he opens to let a molecule pass whenever a fast one comes from the cool side or a slow one comes from the hot side. Otherwise he keeps the aperture closed. Obviously, by this maneuver he gets the cool container becoming cooler, and the hot container getting hotter, thus apparently upsetting the Second Law of Thermodynamics. Of course, we know by now that while he succeeds in obtaining this perverse flow of heat, the Second Law remains untouched. This is because of his need for a flashlight to determine the velocity of the upcoming molecules. Were he at thermal equilibrium with one of the containers he couldn't see a thing: he is part of a black body. Since he can do his antics only as long as the battery of his flashlight lasts, we must include into the system with an active demon not only the energy of the two containers, but also that of the battery. The entropy gained by the battery's decay is not completely compensated by the negentropy gained from the increased disparity of the two containers.

The moral of this story is simply that while our demon cannot beat the Second Law, he can, by his regulatory activity, retard the degradation of the available energy, i.e., the growth of entropy, to an arbitrary slow rate.

This is indeed a very significant observation because it demonstrates the paramount importance of regulatory mechanisms in living organisms. In this context they can be seen as manifestations of Maxwell's Demon, retarding continuously the degradation of the flow of energy, that is, retarding the increase of entropy. In other words, as regulators living organisms are "entropy retarders".

Moreover, as I will show in a moment, Maxwell's Demon is not only an entropy retarder and a paradigm for regulation, but he is also a func-

tional isomorph of a Universal Turing Machine. Thus, the three concepts of regulation, entropy retardation, and computation constitute an interlaced conceptual network which, for me, is indeed the essence of Cybernetics.

I shall now briefly justify my claim that Maxwell's Demon is not only the paradigm for regulation but also for computation.

When I use the term "computation" I am not restricting my self to specific operations as, for instance, addition, multiplication, etc. I wish to interpret "computation" in the most general sense as a mechanism, or "algorithm", for *ordering*. The ideal, or should I say the most general, representation of such mechanism is, of course, a Turing Machine, and I shall use this machine to illuminate some of the points I wish to make.

There are two levels on which we can think of "ordering". The one is when we wish to make a description of a given arrangement of things. The other one when we wish to re-arrange things according to certain descriptions. It will be obvious at once that these two operations constitute indeed the foundations for all that which we call "computation".

Let A be a particular arrangement. Then this arrangement can be computed by a universal Turing machine with a suitable initial tape expression which we whall call a "description" of A: D(A). The length L(A) of this description will depend on the alphabet (language) used. Hence, we may say that a language α_1 reveals more order in the arrangement A than another language α_2, if and only if the length $L_1(A)$ of the suitable initial tape description for computing A is shorter than $L_2(A)$, *or mutatis mutandis*.

This covers the first level of above, and leads us immediately to the second level.

Among all suitable initial tape descriptions for an arrangement A_1 there is a shortest one: $L^*(A_1)$. If A_1 is re-arranged to give A_2, call A_2 to be of a higher order than A_1 if and only if the shortest initial tape description $L^*(A_2)$ is shorter than $L^*(A_1)$, or *mutatis mutandis*.

This covers the second level of above, and leads us to a final statement of perfect ordering (computation).

Among all arrangements A_i there is one, A*, for which the suitable initial tape description is the shortest L*(A*).

I hope that with these examples it has become clear that living organisms (replacing now the Turing machine) interacting with their environment (arrangements) have several options at their disposal: (i) they may develop "languages" (sensors, neural codes, motor organs, etc.) which "fit" their given environment better (reveal more order); (ii) they may change their surroundings until it "fits" their constitution; and (iii), they may do both. However, it should be noted that whatever option they take, it will be done by computation. That these computations are indeed functional isomorphs of our demon's activity is now for me to show.

The essential function of a Turing machine can be specified by five operations:

(i) *Read* the input symbol x.
(ii) *Compare* x with z, the intenal state of the machine.
(iii) *Write* the appropriate output symbol y.
(iv) *Change* the internal state z to the new state z′.
(v) *Repeat* the above sequence with a new input state x′.

Similarly, the essential function of Maxwell's Demon can be specified by five operations equivalent to those above:

(i) *Read* the velocity v of the upcoming molecule M.
(ii) *Compare* $(mv^2/2)$ with the mean energy $\langle mv^2/2 \rangle$ (temperature T) of, say, the cooler container (internal state T).
(iii) *Open* the aperture if $(mv^2/2)$ is greater than $\langle mv^2/2 \rangle$; otherwise keep it closed.
(iv) *Change* the internal state T to the new (cooler) state T′.
(v) *Repeat* the above sequence with a new uncoming molecule M′.

Since the translation of the terms occurring in the correspondingly labeled points is obvious, with the presentation of these two lists I have completed my proof.

How can we make use of our insight that Cybernetics is the science of regulation, computation, ordering, and entropy retardation? We may, of course, apply our insight to the system that is generally understood to be the *cause célèbre* for regulation, computation, ordering, and entropy retardation, namely, the human brain.

Rather than following the physicists who order their problems according to the number of *objects* involved ("The one-body problem", "The two-body problem", "The three-body problem", etc.), I shall order our problems according to the number of *brains* involved by discussing now "The one-brain problem", "The two-brain problem", "The many-brain problem", and "The all-brain problem".

1. The Single-Brain Problem: The Brain Sciences

It is clear that if the brain sciences do not want to degenerate into a physics or chemistry of living—or having once lived—tissue they must develop a theory of the brain: T(B). But, of course, this theory must be written by a brain: B(T). This means that this theory must be constructed so as to write itself T(B(T)).

Such a theory will be distinct in a fundamental sense from, say, physics which addresses itself to a (not quite) successful description of a "subject-less world" in which even the observer is not supposed to have a place. This leads me now to pronounce my Theorem Number Three:

"The Laws of Nature are written by man. The laws of biology must write themselves."

In order to refute this theorem it is tempting to invoke Gödel's Proof of the limits of the Entscheidungsproblem in systems that attempt to speak of themselves. But Lars Löfgren and Gotthard Günther have shown that self-explanation and self-reference are concepts that are untouched by Gödel's arguments. In other words, a science of the brain in the above sense is, I claim, indeed a legitimate science with a legitimate problem.

2. The Two-Brain Problem: Education

It is clear that the majority of our established educational efforts is directed toward the trivialization of our children. I use the term "trivialization" exactly as used in automata theory, where a trivial machine is characterized by its fixed input-output relation, while in a non-trivial machine (Turing machine) the output is determined by the input *and* its internal state. Since our educational system is geared to generate predictable citizens, its aim is to amputate the bothersome internal states which generate unpredictability and novelty. This is most clearly demonstrated by our method of examination in which only questions are asked for which the answers are known (or defined), and are to be memorized by the student. I shall call these questions "illegitimate questions".

Would it not be fascinating to think of an educational system that detrivializes its students by teaching them to ask "legitimate questions", that is, questions for which the answers are unknown?

3. The Many-Brain Problem: Society

It is clear that our entire society suffers from a severe dysfunction. On the level of the individual this is painfully felt by apathy, distrust, violence, disconnectedness, powerlessness, alienation, and so on. I call this the "participatory crisis", for it excludes the individual from participating in the social process. The society becomes the "system", the "establishment" or what have you, a depersonalized Kafkanesque ogre of its own ill will.

It is not difficult to see that the essential cause for this dysfunction is the absence of an adequate input for the individual to interact with society. The so-called "communication channels", the "mass media" are only one-way: they talk, but nobody can talk back. The feedback loop is missing and, hence, the system is out of control. What cybernetics could supply is, of course, a universally accessible social input device.

4. The All-Brain Problem: Humanity

It is clear that the single most distressing characteristic of the global system "mankind" is its demonstrated instability, and a fast approaching singular-

ity. As long as humanity treats itself as an open system by ignoring the signals of its sensors that report about its own state of affairs, we shall approach this singularity with no breaks whatsoever. (Lately I began to wonder whether the information of its own state can reach all elements in time to act should they decide to listen rather than fight.)

The goal is clear: we have to close the system to reach a stable population, a stable economy, and stable resources. While the problem of constructing a "population servo" and an "economic servo" can be solved with the mental resources on this planet, for the stability of our material resources we are forced by the Second Law of Thermodynamics to turn to extra-planetary sources. About $2 \cdot 10^{14}$ kilowatts solar radiation are at our disposal. Wisely used, this could leave our earthy, highly structured, invaluable organic resources, fossilized or living, intact for the use and enjoyment of uncounted generations to come.

If we are after fame and success we may ignore the profundity of these problems in computation, ordering, regulation, and entropy retardation. However, since we as cyberneticians supposedly have the competence to attack them, we may set our goal above fame and success by quietly going about their solution. If we wish to maintain our scientific credibility, the first step to take is to apply our competence to ourselves by forming a global society which is not so much *for* Cybernetics as it *functions* cybernetically. This is how I understand Dennis Gabor's exhortation in an earlier issue: "Cyberneticians of the world, unite!" Without communication there is no regulation; without regulation there is no goal; and without a goal the concept of "society" or "system" becomes void.

Competence implies responsibilities. A doctor must act at the scene of the accident. We can no longer afford to be the knowing spectators at a global disaster. We must share what competence we have through communication and cooperation in working together through the problems of our time. This is the only way in which we can fulfill our social and individual responsibilities as cyberneticians who should practice what they preach.