

# Thoughts and Notes on Cognition

Heinz von Foerster

University of Illinois

I am deeply Indebted to Humberto Maturana, Gotthard Gunther,(1) and Ross Ashby for their untiring efforts to enlighten me in matters of life, logic, and large systems, and to Lebbeus Woods for supplying me with drawings that illustrate my points better than I could do with words alone. However, should there remain any errors in exposition or presentation, it is I who am to blame and not these friends who have so generously contributed their time.

## THOUGHTS

Projecting the image of ourselves into things or functions of things in the outside world is quite a common practice. I shall call this projection “anthropomorphization.” Since each of us has direct knowledge of himself, the most direct path of comprehending *X* is to find a mapping by which we can see ourselves represented by *X*. This is beautifully demonstrated by taking the names of parts of one’s body and giving these names to things which have structural or functional similarities with these parts: the “head” of a screw, the “jaws” of a vise, the “teeth” of a gear, the “lips” of the cutting tool, the “sex” of electric connectors, the “legs” of a chair, a “chest” of drawers, etc.

Surrealists who were always keen to observe ambivalences in our cognitive processes bring them to our attention by pitching these ambivalences against semantic consistencies: the legs of a chair (Fig. 1(2)), a chest of drawers (Fig. 2(3)), etc.

At the turn of the century, animal psychologists had a difficult time in overcoming functional anthropomorphisms in a zoology populated with animals romanticized with human characteristics: the “faithful” dog, the “valiant” horse, the “proud” lion, the “sly” fox, etc. Konrad Lorenz, the great ornithologist, was chased from Vienna when he unwisely suggested controlling the population of the overbreeding, underfed, and tuberculosis-carrying pigeons of the city by importing falcons which would raid the pigeons’ nests for eggs. The golden heart of the Viennese could not stand the thought of “pigeon infanticide.” Rather, they fed the pigeons twice as much. When Lorenz pointed out that the result of this would be twice as many underfed and tuberculosis-carrying pigeons, he had to go, and fast!

Of course, in principle there is nothing wrong with

anthropomorphizations; in most cases they serve as useful algorithms for determining behavior. In trying to cope with a fox it is an advantage to know he is “sly,” that is, he is a challenge to the brain rather than to the muscles.

Today, with most of us having moved to the big cities, we have lost direct contact with the animal world, and pieces of steel furniture with some functional properties, the computers, are becoming the objects of our endearments and, consequently, are bestowed now with romanticizing epithets. Since we live today, however, in an era of science and technology rather than in one of emotion and sentimentality, the endearing epithets for our machines are not those of character but of intellect. Although it is quite possible, and perhaps even appropriate to talk about a “proud IBM 360–50 system,” the “valiant 1800,” or the “sly PDP 8,” I have never observed anyone using this style of language. Instead, we romanticize what appears to be the intellectual functions of the machines. We talk about their “memories,” we say that these machines store and retrieve “information,” they “solve problems,” “prove theorems,” etc. Apparently, one is dealing here with quite intelligent chaps, and there are even some attempts made to design an A.I.Q., an “artificial intelligence quotient” to carry over into this new field of “artificial intelligence” with efficacy and authority the misconceptions that are still today quite popular among some prominent behaviorists.

While our intellectual relationship with these machines awaits clarification, in the emotional sphere we seem to do all right. I wish to make this comment as a footnote to Madeleine Mathiot’s delightful observations in this volume about various degrees of “awesomeness” associated with the referential genders “it,” “he,” and “she.” She develops a three-valued logical place-value system in which the nonhuman “it” carries no reference

to awesomeness either in the negative (absence) or else in the affirmative (presence), while the human “he” and “she” indeed carry reference to awesomeness, the masculine “he” referring to its absence, the feminine “she,” of course, to its presence.

When in the early fifties at the University of Illinois ILLIAC II was built, “it” was the referential gender used by all of us. The computer group that now works on ILLIAC III promises that “he” will be operative soon. But ILLIAC IV reaches into quite different dimensions. The planners say that when “she” will be switched on, the world’s computing power will be doubled.

Again, these anthropomorphisms are perfectly all right inasmuch as they help us establish good working relations with these tools. Since most of the people I know in our computer department are heterosexual males, it is clear that they prefer the days and nights of their work spent with a “she,” rather than with an “it.”

However, in the last decade or so something odd and distressing developed, namely, that not only the engineers who work with these systems gradually began to believe that those mental functions whose names were first metaphorically applied to some machine operations are indeed residing in these machines, but also some biologists — tempted by the absence of a comprehensive theory of mentation — began to believe that certain machine operations which unfortunately carried the names of some mental processes are indeed functional isomorphs of these operations. For example, in the search for a physiological basis of memory, they began to look for neural mechanisms which are analogues of electromagnetic or electrodynamic mechanisms that “freeze” temporal configurations (magnetic tapes, drums, or cores) or spatial configurations (holograms) of the electromagnetic field so that they may be inspected at a later time.

The delusion, which takes for granted a functional isomorphism between various arid distinct processes that happen to be called by the same name, is so well established in these two professions that he who follows Lorenz’s example and attempts now to “de-anthropomorphize” machines and to “de-mechanize” man is prone to encounter antagonisms similar to those Lorenz encountered when he began to “animalize” animals.

On the other hand, this reluctance to adopt a conceptual framework in which apparently separable higher mental faculties as, for example, “to learn,” “to remember,” “to perceive,” “to recall,” “to predict,” etc., are seen as various manifestations of a single, more inclusive phenomenon, namely, “cognition,” is quite understandable. It would mean abandoning the comfortable position

in which these faculties can be treated in isolation and thus can be reduced to rather trivial mechanisms. Memory, for instance, contemplated in isolation is reduced to “recording,” learning to “change,” perception to “input,” etc. In other words, by separating these functions from the totality of cognitive processes one has abandoned the original problem and now searches for mechanisms that implement entirely different functions that may or may not have any semblance with some processes that are, as Maturana<sup>1</sup> pointed out, subservient to the maintenance of the integrity of the organism as a functioning unit.

Perhaps the following three examples will make this point more explicit.

I shall begin with “memory.” When engineers talk about a computer’s “memory” they really don’t mean a computer’s memory, they refer to devices, or systems of devices, for recording electric signals which when needed for further manipulations can be played back again. Hence, these devices are stores, or storage systems, with the characteristic of all stores, namely, the conservation of quality of that which is stored at one time, and then is retrieved at a later time. The content of these stores is a record, and in the pre-semantic-confusion times this was also the name properly given to those thin black disks which play back the music recorded on them. I can see the big eyes of the clerk in a music shop who is asked for the “memory” of Beethoven’s Fifth Symphony. She may refer the customer to the bookstore next door. And rightly so, for memories of past experiences do not reproduce the causes for these experiences, but — by changing the domains of quality — transform these experiences by a set of complex processes into utterances or into other forms of symbolic or purposeful behavior. When asked about the contents of my breakfast, I shall not produce scrambled eggs, I just say, “scrambled eggs.” It is clear that a computer’s “memory” has nothing to do with such transformations, it was never intended to have. This does not mean, however, that I do not believe that these machines may eventually write their own memoirs. But in order to get them there we still have to solve some unsolved epistemological problems before we can turn to the problem of designing the appropriate software and hardware.

If “memory” is a misleading metaphor for recording devices, so is the epithet “problem solver” for our computing machines. Of course, they are no problem solvers, because they do not have any problems in the first place. It is our problems they help us solve like any other useful tool, say, a hammer which may be dubbed a “problem solver” for driving nails into a board. The danger in this subtle semantic twist by which the responsibility for ac-

<sup>1</sup> See Chapter 1, pages 3–23.

tion is shifted from man to a machine lies in making us lose sight of the problem of cognition. By making us believe that the issue is how to find solutions to some well-defined problems, we may forget to ask first what constitutes a “problem,” what is its “solution,” and — when a problem is identified — what makes us want to solve it.

Another case of pathological semantics — and the last example in my polemics — is the widespread abuse of the term “information.” This poor thing is nowadays “processed,” “stored,” “retrieved,” “compressed,” “chopped,” etc., as if it were hamburger meat. Since the case history of this modern disease may easily fill an entire volume, I only shall pick on the so-called “information storage and retrieval systems” which in the form of some advanced library search and retrieval systems, computer based data processing systems, the nationwide Educational Resources Information Center (ERIC), etc., have been seriously suggested to serve as analogies for the workings of the brain.

Of course, these systems do not store information, they store books, tapes, microfiche or other sorts of documents, and it is again these books, tapes, microfiche or other documents that are retrieved which only if looked upon by a human mind may yield the desired information. Calling these collections of documents “information storage and retrieval systems” is tantamount to calling a garage a “transportation storage and retrieval system.” By confusing vehicles for potential information with information, one puts again the problem of cognition nicely into one’s blind spot of intellectual vision, and the problem conveniently disappears. If indeed the brain were seriously compared with one of these storage and retrieval systems, distinct from these only by its quantity of storage rather than by quality of process, such a theory

would require a demon with cognitive powers to zoom through this huge system in order to extract from its contents the information that is vital to the owner of this brain. *Difficile est saturam non scribere*<sup>2</sup>. Obviously, I have failed to overcome this difficulty, and I am afraid that I also fail in overcoming the other difficulty, namely, to say now what cognition really is. At this moment, I even have difficulties in relating my feelings on the profoundness of our problem, if one cares to approach it in its full extension. In a group like ours, there are probably as many ways to look at it as there are pairs of eyes. I am still baffled by the mystery that when Jim, a friend of Joe, hears the noises that are associated with reading aloud from the black marks that follow

#### ANN IS THE SISTER OF JOE

— or just sees these marks — knows that indeed Ann is the sister of Joe, and, *de facto*, changes his whole attitude toward the world, commensurate with his new insight into a relational structure of elements in this world.

To my knowledge, we do not yet understand the “cognitive processes” which establish this insight from certain sensations. I shall not worry at this moment whether these sensations are caused by an interaction of the organism with objects in the world or with their symbolic representations. For, if I understood Dr. Maturana correctly, these two problems, when properly formulated, will boil down to the same problem, namely, that of cognition *per se*.

In order to clarify this issue for myself, I gathered the following notes which are presented as six propositions labeled  $n = 1-6$ . Propositions numbered  $n.1$ ,  $n.2$ ,  $n.3$ , etc., are comments on proposition numbered  $n$ . Propositions numbered  $n.m1$ ,  $n.m2$ , etc., are comments on proposition  $n.m$ , and so on.

Here they are.

<sup>2</sup>It is difficult not to write satire.