STEPS TO A CYBERNETICS OF AUTONOMY

Francisco Varela

In the context of this panel on some fundamental guidelines for cybernetics and systems theory I would like to present today a very specific point of view: that of a biologist. I came upon these issues because I was involved in studying things like nervous systems or immune systems. In cybernetics and systems theory, important notions have been shaped by empirical research dealing with complex biological systems. It is well known to all of you that in the pioneering days - in the 40ies and the 50ies - biology played a fundamental role in asking the questions that then led to the full development of these disciplines.

At the risk of 'dulling' something that is more interesting - let me focus on the contrast between two giants of that time: John von Neumann and Norbert Wiener. Both of them strongly motivated from biology, both of them making enormous contributions to cybernetics and systems theory, but both of them, in their latter days, moving in very different directions (Heims, 1982). Norbert Wiener by emphasizing the quality of independence, autonomy, creativity, the quality of living beings to create their meaning, to create their world. John von Neumann by emphasizing the quality of specifying decision rules, procedures for exact computations, control. During those early days, it was unclear what was going to be the dominant trend of those two sides of the issue - whether control, autonomy, or both. It seems to me that it is quite clear - looking back from the 1980ies - that von Neumann actually prevailed. Cybernetics and systems theory developed most of their effort into the now familiar characterization of a machine (or an automaton) as an input-output device: you have inputs, transformations, and some kind of output, which, of course, can be made very precise with the notion of a Turing machine. The fundamental notion of a Turing machine and all of its abstract and concrete applications is, as far as I can see it, the ultimate evolution of the notion of Cartesian causality.

The point I would like to raise is that it is high time now to develop the Wienerian side quite a lot more, and that it is only in recent years that, out of biology itself, have come the actual grounds to demand the revision of this trend in cybernetics and systems theory. The grounds for that are mostly out of the study of systems like the immune system, the nervous system, or ecological systems. So, as a biologist, my feeling - or my judgement if you want - is that cybernetics and systems theory as it is developed today is <u>incomplete</u>. It is incapable to actually encompass the totality of the relevant biological phenomenology. It is one-sided. What it leaves aside is what I referred to before as the quality of <u>autonomy</u>, the quality of living systems of having an assertion of their internal coherences, their internal determination, as well as the fact that it is this internal determination the one that shapes or imbues a world with meaning.

I am talking then about two basic issues:

(1) self-determination and
(2) the emergence of meaning.

These two qualities are completely outside, cannot be fit into the model or mechanism captured by a Turing machine. I do not mean to say that nobody has ever asked about these qualities before. The incompleteness vis-a-vis this living quality of autonomy has been raised in many different ways. For example it has been said that we must study more the question of self-organization; some people prefer to talk about synergetics; some others speak about co-operative properties. But it seems necessary that we pick up the basic issue out of these various trends in various disciplines, in order to make a more coherent picture of what this line of development of cybernetics ought to be or could be. So let me outline for you, in the next minutes, what I see as the fundamental points that we need to grapple with and deepen in our understanding, if we are to come to grips with the autonomous side of living systems. I said I am speaking as a biologist - in social systems very similar issues arise, but I am very ignorant about them; I'll speak just about those I can put my hands in.

First of all, there is the question of how to <u>characterize</u> a system. In the classical Turing-von Neumann context, a characterization of a system is given by the list of inputs and outputs and their transfer functions. That is to say, one characterizes a system by the way it handles what is given to it as a specified input. This is very familiar to everyone of us. For an autonomous system the characterization is different. One shifts from the emphasis on the inputs - and how they are transformed - to the emphasis on the internal regularities of how the system is constructed. I call this 'operational closure' (Varela, 1979): in an autonomous system we find that its components are so strongly interrelated that it is this internal coherence and interrelatedness what is central, rather than the way inputs are specified. So, instead of inputs and their transformation, one shifts to operational closure, as a characterization of the internal network.

What becomes of inputs then? We have shifted our emphasis: they are not something that is explicitly given, but inputs become simply a background of perturbations which are undefined, 'background noise'. They do not enter into the definition of the machine, system or procedure. What we have learned from the study of systems which have a very clear operational closure, (such as cells, nervous systems or immune systems) can be stated as an empirical conclusion that does not yet have a fundamental theoretical validation: every time there is operational closure, there is also the emergence of <u>internal regularities</u> which arise out of the interconnectedness. Such internal states can be thought of as 'stabilities' or, more appropriately, one can talk about eigen-behaviors, that is, self-determined behaviors.

The study of eigen-behaviors is a big chapter which is of great interest to me, but we have to leave it at that for the time being. Instead, let me backtrack: we have proposed a shift in the characterization from a Turing to an autonomous machine; we have proposed a shift in the characterization from an input and transfer function to one of operational closure and the understanding of how that closure gives rise to eigen-behaviors. These are two fundamentally different modes of approaching and studying a specific situation.

Now, once we have this alternative characterization for an autonomous system, it immediately follows that the mode of relationship of such a system to its environment is completely different. In the first case, (a Turing automaton with inputs) the mode of relationship with its environment is always one of representation, namely items of the environment become instructions that act on the structure of a system - that's why we call them inputs, otherwise we would not call them that. It is fundamentally an instructive mode of relationship or interaction. Instead for an autonomous machine characterized by its closure and its eigen-behavior, what happens is that these eigen-behaviors will specify out of the noise what of that noise is of relevance. So, what you have is a laying down of a world, a laying down of a relevant 'Umwelt'. A world becomes specified or endowed with meaning; out of eigen-behaviors, there arises possibility of generating 'sense'. So what we are talking about here is the contrast between an instructive Turing automaton and an autonomous machine capable of creating (or generating) sense.

So far we have examined what an autonomous system could be by giving a characterization for it, by seeing its mode of relationsship with its environment. A third aspect to consider - to me the most poignant one - is that in this approach we have changed our mode of inference. We have, in fact, an entirely different mode of inference. The mode of inference in the context of a Turing automaton is one where the outside is causal to the inside, and therefore where objectivity is underlined. It is the structure of the environment that has to be well-defined and predetermined. From the point of view of the characterization of an autonomous mechanism, what happens is that the inside is what endows meaning, creates sense - therefore the mode of inference is from the inside to the outside and thereby objectivity is immediately bracketed. That is to say, we suspend what from our point of view looks like a very structured and defined world, and we let the system reveal what is relevant for it. So, objectivity from that point of view is bracketed.

Let me briefly repeat those three points:

(1) Characterization: going from input and output and transfer-functions to operational closure and eigen-behavior;
(2) Mode of relationship: going from an instructive one to laying down a sense;
(3) Mode of inference: going from underlying objectivity to bracketing objectivity.

Please do not take me to be saying that one is better than the other. I am saying they are different. I am saying we need the second, we need the characterization of autonomous mechanisms to actually come to grips with that which is presented in the living world. Of course there are some contexts, in which the Turing characterization is very good and very useful, but I am saying that this is far from being all we need. We need to develop the other tools. We could go here into a long list of the partial tools that have already been developed, the applications in many fields. I see here today some distinguished people who have contributed enormously to do this in different areas - but we don't have time to do that. Instead, let me give you an example which I hope I'll manage to convey the flavor of what I want to say.

Several years ago, a few immunologists in Europe and the United Stated discovered something that was was contrary to what was the belief up until then about the so-called antibodies. These are protein molecules circulating in the blood which normally are said to bind to antigenes; that is, molecules that come from the outside such as viruses and bacteria. Antibodies were taken to be means to guard your body by reacting to an antigene that would come from the outside, bind to the antigene, and then reduce it to nothing. So, it was like a surveillance system. Typically an input - output situation. Now, as it turns out, somebody (almost by accident) discovered that there are antibodies to other antibodies, the so-called anti-idiotypic antibodies. Idiotypes are molecular determinants in the lymphocite cells which make up the immune systems. One can have antibodies against antibodies which of course produce antibodies against those antibodies etc., etc.. If you have ever seen such an infinitely branching structure, you know that it is equivalent to a closed network. So, there is a fundamental immune closure: a lymphocyte talks mainly to lymphocytes; it is not looking outside to antigenes, but is mostly talking to 'his own peers', so to speak. What becomes of an antigene then? An antigene becomes something very different - it is that molecule which resembles enough one of these idiotypes to be able to sneak in into the ongoing closure and produce a change in the network. Antigenes are not determined as a list of what are the relevant bacteria to be kept outside, but rather by the structure of the immune system itself. The rest is simply nonsense. Thus the immune system endows the molecular world with a meaning, in the sense that it is only through its closure that certain molecular items are classified as being relevant. Furthermore, during the development of an organism, the idiotopes are never the same, so that if I were to take the response of my immune system and anyone of yours to the same molecule, we would find that both styles of response are completely different. The family of antibodies against those idiotopes will be completely different, although from the behavioral point of view both you and I would have performed a 'recognition' out of this molecule. The quality of laying down a sense is not given by what you give, but by the structure of the system. It can be arrived at by many possible rules. To say it metaphorically: you can lay down a path by walking in many different ways - the important thing is that you keep walking. That is what the immune system does - it keeps walking. You walk differently than I do, we both walk - that is what matters. There is no 'representation' of the world and its invasive agents.

The examples could be multiplied and discussed in detail, such as for the nervous system. Other people have worked on social applications of similar notions. Let me conclude. My main point today is that I am convinced that we need to understand that autonomous mechanisms are fundamentally different from Turing-Cartesian ones. They address themselves to different issues, they require different tools, they entail a different form of looking at the world. But we need to develop them in order to actually cope with situations which are of a complexity appropriate to the living and the social world. To me, what it comes down to as a fundamental guideline for cybernetics and systems research is the need to actually go back to a Wienerian spirit, and re-take the issues that he raised in the latter part of his life again afresh with new tools and with twenty years of having developed a cybernetics of Turing mechanism. Thank you.

DISCUSSION

Moderator: Any questions to Francis Varela?

From the floor: Yes, one small, and perhaps an important one: I am wondering what the reason was that referring to the founding fathers of our systems theory you have limited yourself to Wiener and von Neumann.

Varela: Just for the sake of making a point; these two gentlemen go so parallel, they are both mathematicians, they are both professors at American universities, they both dealt with war issues during the war, they both exchanged views quite a lot in their lives. It is just a good example, it is a story well told - se non e vero, e ben trovato.

From the floor: ...., if you look at the main part of systems theory you will see of course the type of systems which I would refer to dynamical systems with input and output - I am not sure if the Turing machine is really a very good example for what you mean, because I consider most applications as rather autonomous - you load a program and then you make it run, so at the initial stage ....

Varela: We are not talking about the same thing....

From the floor: I am saying, in engineering of course you are not so much concerned about autonomous systems ....

Varela: I know, that's fine ....

From the floor: The systems theory, or the biologist ... should just come with the problems and then it will be just built out in a way as the contributions ... from engineering.

Varela: That is right, again - please do not misunderstand me. I am not arguing for one type of approach (such as autonomous mechanisms) in favour of another. I am saying: this is fine and good and for engineering it works wonders, but let us not be naive that when we try to transpose that to another level of problems, they just do not work. Basically they do not work, or they work to such a limited extent that we leave out all the real meat of what the issues are. So let us just enlarge the scope. I am not saying that this is all wrong and has to be thrown out the window - no - it is another topic altogether.

From the floor: Also the word 'autonomous' is of course already very much used because we have an autonomous differentiation ....

Varela: I know, but this is in a different sense ...., in a fifteen minutes' talk the words will get fuzzy, but we should pursue that later. Please, Professor Beer ....

Beer: I've realized that the distinction you have drawn is an expository device. But I would like to know to what extent you think that W.Power is a bridge between these two paradigms? You know the work I am referring to?

Varela: Yes, 'Behavior as the Control of Perception'.

Beer: The perception is governed by the process of seeing. It seems to me that that makes a kind of bridge between the two things you have distinguished between.

Varela: No, I do not think so because it seems to me that what he is doing is applying the tool of one field to address the issues of the other. I would classify Powers definitely as a person interested in autonomy. But once you confront a system, you have to make a choice that, as far as I can tell, is always all or none. I either take this perspective, or I take this other one. I am not saying you cannot take both, but at every one moment it is either this or that. Maybe in some future somebody will find a clever way to actually finding some solid complementarity. It is a stance that we take; it is that or that, but not both of them simultaneously. Powers, I believe, is slightly misleading in that he takes one stance and then he talks with the language of the other. But that does not mean that it is a bridge.

Beer: That is what I meant.

From the floor: My question now is: What is your concept of information you have in your framework - is it information as to how to construct a system or how to alter a system?

Varela: The shortest form I have to answer that is saying that information exists on the side of what I am calling here input/output systems - we all know what it is. Here I say that this notion becomes fundamentally different; that is why I do not like to even use the word - but if I were to use it, I would write it something like this: 'in-formation' - that is, something that is formed within.

input/output systems	autonomous
|	|
information	in-formation

In other words, it is the quality of a system to endow the world with a certain regular relationship which then tells you that that item has meaning, e.g. that I am allergic to pollen. You see the difference? These two ideas have nothing to do with each other - that is why I rather drop this word and say that in this side there is <u>no</u> information. There is sense, there is meaning if you want.

From the floor: You said in the beginning, I think, that there are two possible ways in which this subject could go; von Neumann's way or Norbert Wiener's way. One was to do with autonomy and one with control. I was astonished to hear you say that von Neumann was to do with control - I would have thought he was using his logic instrumentally and to do it with autonomous machines and that Norbert Wiener in his book 'The Human Use of Human Beings' at least was trying to get us to trying to control how we think ....

Varela: We have a fundamentally different reading of the same people - I take it exactly the opposite. Von Neumann was the one who advocated the development and use of the atomic bomb and Wiener was the one who always stood against it. Wiener was the one who always advocated the use of technology for social and political problems. Von Neumann was the one who was always trying to make it technological and precise and computable. There is historical evidence for that (Heims, 1982). So to me von Neumann - the old von Neumann - is really leaning towards the control ideology, while Wiener more and more reacted against it and claimed a revival of the revision of values and a non-separation between value and research. I am not a historian of science, my talk is not founded on that, it was meant as an expository device. What I do want to address is the importance of the difference between the two types of mechanism.

REFERENCES

Heims, S., 1982, "John von Neumann and Norbert Wiener", MIT Press, Cambridge, Mass.

Varela, F.J., 1979, "Principles of Biological Autonomy", Elsevier/North-Holland, New York.