

## SOME THEMES AND PRIMITIVES IN ILL-DEFINED SYSTEMS\*

Oliver G. Selfridge

45 Percy Road Lexington, MA 02173

To say that something is ill-defined is more to describe us than it. That is, for some system we are dealing with, we are more or less ignorant of its working, and that means that we have special problems in trying to control it. Nevertheless, of course, most of the real systems we deal with are ill-defined in that sense -- like other people. But we do find ways to exercise some degree of control.

The main theme here is that -- regardless of exactly what control means -- it is feasible to do better by a succession of quite small improvements in a strategy. Indeed, some of the examples I give suggest to me that evolution has improved the behavior of the organisms in this way just as much as it has improved their somatic organization. Indeed, a long enough sequence of small improvements can lead to enormous changes of effectiveness of control.

Even without a mathematical theory or a complete description of a system, without comprehending how it works, or knowing how to model its interactions, we can examine a system or a situation and choose some forms of control and behavior on our part that will do better than others. As our experience proceeds, we may redefine what "doing better" means, but we will continue to learn from successes and failures.

The word adaptation brings to mind some kind of adjustment to conditions. That it seems to depend on one's point of view may

---

\* Many of the themes in this paper are developed at greater length in O.G. Selfridge: Tracking and Trailing. Bradford Books, Cambridge, MA (1983).

perhaps be illustrated whether the flow of water in a stream can usefully be considered to be adapting to the sinuosities and roughnesses of its banks; or are the banks adapting to the flowing water? For me, terming that adaptation and making it analogous to the governor on a steam engine doesn't seem to be very helpful.

Later I will argue that the apparent presence of adaptation depends on whether our view of behavior lets us usefully describe it in terms of purpose and purposes. For example, over twenty years ago, there was widespread hope that useful complexity of systems could arise merely from "self-organization." The most well know of the efforts concerned a set of machines called perceptrons.\* But the changes in their behavior seemed not to be related to any purposes, and in fact they failed in what was claimed for them.

Here I should like to start by considering some elementary paradigms of adaptation, concerned with adjusting simple controls of movement of a simple organism. The organism is a very elementary one-celled creature called Escherichia Coli. E. Coli is only a couple of microns long, and it moves by twirling a flagellum at one end. The flagellum twirls in one direction like a propellor, and drives the creature in a straight line, more or less, for distances of, say, 10 or 20 microns. Then for some reason it decides to rotate its flagellum in the other direction; but the structure of the flagellum is such that it then separates into separate strands, and E. Coli merely jerks around or "twiddles" in one spot. When the flagellum next reverses, E. Coli takes off again in a straight line in another direction. To some degree, the successive directions are unrelated.

Such a procedure is in mathematics called a random walk, and if that were all that E. Coli did, we should find it duller than it is. The interesting thing is that when things are getting better for E. Coli -- say when the concentration of some edible amino acid in its environment rises -- then E. Coli stops reversing the direction of its flagellum. That is, it keeps swimming much longer in the same direction. Soon after things stop getting better, the flagellum reverses, and twiddling is resumed. The effect is that E. Coli can climb gradients of attractiveness, and find high concentrations of nourishment.

The visual effect is quite startling under a microscope. Take a Petri dish full of E. Coli swimming around uniformly; put a small drop of, say, lysine in the middle of the dish. After it has diffused out slowly across the dish, the germs will be seen clustering closely around it.

---

\* The work was run by the late Dr. Frank Rosenblatt at Cornell Aeronautical Laboratory.

I call such behavior "run and twiddle" or RT for short, and I suggest that it is an adaptive behavioral "primitive," and an elementary fundamental mechanism or strategy of all movement.

It is a very simple mechanism. There is only one sensor, presumably the outside membrane (there is no evidence of spatial discrimination, and the difference in concentration at the two ends must be very small, and beset by Brownian variations.) Such behavior can be moderately well simulated by computer. We can make the behavior probably simpler than it really is by representing it with just two numbers, the probabilities P1 and P2. The former is the probability of twiddling when things are ordinary, the latter when things are getting better. By adjusting P1 and P2 we can produce tracks in a simulation which look quite like real tracks in a microphotograph.

In fact, we can put the principle very simply — if things are getting better, don't change what you're doing.

Let us consider a slightly more complicated expression in a more complicated beast, *Euglena*. *Euglena* is also much bigger than *E. Coli*, and it swims with cilia. It has a certain degree of asymmetry, which results in its moving along a helical path as its own axis rotates. As it moves, then, its axis points in changing directions.

*Euglena* has chloroplasts, which derive energy from light, and so it wants to get to where the light is, usually near the surface of the points it inhabits. It has a light sensitive spot, and also, at the front end, a black spot. Now this black spot casts a shadow; and, because the beast rotates while it swims, this shadow rotates around the eye, just so long as it swims directly towards the light. If *Euglena* deviates from this optimum course, than the shadow covers the light spot when it is pointing far enough off course. This causes suppression of the movement of the inside cilia, just enough to provide a correction.

Again, there is but a single sensor, and a single mode of behavior, which we have called RT. In general, of course, a single sensor cannot determine direction with a single measurement, and the behavior then must depend on the history of the sensor's measurements. With two sensors, much more complex behavior is possible.

An example of two parallel sensors is the ant with its two antennae. One of the behaviors commonly seen in ants is to follow a pheromone trail; a pheromone is a chemical that can be sensitively smelled, and ants lay trails of pheromone that can last for some minutes. Other ants can follow such trails. If one antennae senses more pheromone, then that antenna causes the legs on that side to

take smaller steps, pulling the ant back on to the trail.

Can such trail-following be accomplished with a single sensor? Indeed it can, but not just by the simple hill-climbing RT of *E. Coli*. Let us consider the silk moth. By comparison with *E. Coli* it is a very sophisticated creature, and in fact it takes advantages of several different mechanisms of behavior -- but each is basically a simple RT.

The male silk moth finds its mate at night, when the wind is light but not absent. The female sits on a tree at about 5 feet above the ground and emits a powerful pheromone. This is naturally enormously diffused as it drifts downwind in a plume, spreading out sideways and vertically. Some kilometers away, the male silk moth is waiting, in some tree, with extraordinarily sensitive antennae which can detect a pheromone density on the order of 1 molecule per second. As soon as this happens for a few seconds, he leaves the tree and flies upwind, zigzagging as he goes. The point of the zigzagging is to provide an integration factor for the very small densities involved. He changes direction when he detects that the concentration is falling off on one side of the trail.

A creature as complicated as a moth has, as I said, other forms of behavior. Supposing he loses the trail? Then he flies across the wind, not nearly upwind, in a succession of several attempts to recover the trail.

We can describe this trail following as consisting of several modes. First there is search for an unknown trail, which for the silk moth consists in sitting still and waiting for the trail to come to him. Second, there is following the trail towards the source; third is recovery if the trail is lost; and finally, there must be recognition that the source has been reached.

I hope that the reader will have asked how the silk moth knows which way is upwind, because he can only tell that from the pressure when he is not flying. After a fair amount of discussion in the scientific literature, the following mechanism has been shown: as soon as he smells something, he flies, starting towards the direction the wind is coming. Then he uses his very sensitive eyes (although they fly at night, even the light from stars alone give enough light) to see which way the ground underneath him is going.

And how does the silk moth tell when he has reached his goal, the female? It seems that the trigger for that decision is the mere density of the pheromone, some 10 times large than when he started. At that point, the silk moth merely lands on the nearest tree and walks in circles until he finds the female.

So the individual mechanisms are really pretty simple, but they add up to what certainly looks like purposive behavior of some power and subtlety. The individual mechanisms are the simplest form of adaptation, which I have called RT. Their combining into complex purpose structures, as in the silk moth, is what does the trick.

Now the elementary RT, as a control mechanism, is not easily susceptible to the kinds of analysis we are comfortable with. RT is not linear, it's not inherently continuous or discrete -- and of course it doesn't always work. It requires that the running and twiddling be somehow relevant to the environment.\* For people the easiest way to assure that is to deal with some model of the environment, which they can understand and analyze. But of course the organisms we have discussed are not intellectually rich or powerful enough for that; and yet they survive and adapt and improve. In some sense, when we deal with systems that we do not understand -- which we can say then are ill-defined -- we have to deal with them as those organisms deal with their environments.

I hope it is clear, even from the examples above, that the elementary adaptive mechanisms are usually embedded into hierarchies of them, especially in higher animals. What are the ways that one adaptive mechanism can control another? Let us examine the RT control of *E. Coli* a little more closely. First, there are two probabilities of the model, P1 and P2. Then there is a factor that I did not mention: how long does it take the membrane to react; alternatively, what is its integration time for output?

The elementary unit of classical control theory is usually considered to be the servomechanism. In the control of the servomechanism, there are interesting parallels between it and the RT mechanisms here. The two probabilities are some measure of the gain of the system. That is to say, how much difference in behavior does a small change in control signal make? The integration time of each is analogous, having to do with total loop delays and impulse responses.

It was Wiener who emphasized thirty-five years ago in "Cybernetics" that control of muscles is not effected by direct impulses coming from the motor cortex straight down to the muscle fibers. Rather the cortical impulses are directed at a spinal control loop, changing its gain, as it were.

---

\* This is analogous to the assertion in artificial intelligence that a very important aspect of a problem is to find the right form of representation.

But the most powerful method of control for both mechanisms is one that is in effect their purpose: what is the evaluation function for the RT? -- and what is the set point for their servo, that is, where is the error signal zero?

For example, one form of resetting the evaluation function is to use different sensors. If a membrane is initially sensitive to chemical concentrations in the surrounding fluid, then it might not take a great change in that membrane to make it sensitive to light instead. And in some cases, that would turn out to be profitable for the organism, as with *Euglena*.

My picture is that evolution has built more complex behavior into organisms by building hierarchies and combinations of simple pieces of behavior; control loops controlling other loops in great structures, with little overall understanding in the human sense, and certainly with no overall sets of purposes.

Evolutionary progress has been slow indeed, I think because of that lack of a set of purposes. When we try to find out how to deal with systems that we do not understand, that are ill-defined, we start out with a fair understanding of our own complex purposes and a fair set of intellectual tools for modeling and analyzing parts of the behavior of those systems. We can use these tools to put together some approximation of a strategy for controlling the system.

At that point, however, we must resort to something very like the RT strategies. It is clear, for example, that we have very little profound understanding of how economic systems work. In dealing with them, are we much better than *E. Coli*? Are our models of learning in children so accurate and reliable that we can derive educational strategies?

But as we try simple adaptive techniques, we learn more about the systems; every time we narrow something down, we can apply more sophisticated and powerful analytic tools, because the system we are examining becomes less ill-defined for us.

I have argued that simple adaptive techniques, applied in a responsive environment, and put together in hierarchical control structures, can improve control without necessarily requiring an accurate definition of the system. In that sense, finding out how to do that efficiently, and how to evaluate how well it works, can help to show us how to control systems we do not understand, even ones we are part of ourselves.