SECTION 3

More Complex Perception Systems

The previous analysis dealt with an extremely simplified system, a four-by-four array of 16 individual sensors. To validly deal with the real problem of *How the Mind Arises from the Brain* it is now necessary to address a more realistic system, the human visual system and brain.

PARAMETERS OF VISUAL PERCEPTION

Instead of sixteen sensory elements as in the preceding example, the human eye has about 7,000,000 such sensory elements, the rods and cones of the retina. The number of different patterns that can be represented by that number on a binary digital array is

(3-1) $$\text{Number of possible patterns} = 2^{[\text{number of elements in the array}]}$$

Thus the human eye can deal with about $2^{7,000,000}$ different patterns, an extremely large number.

Since $2^{10} = 1,024$, then, taking 1,000 as an approximation to 1,024,

(3-2) $$2^{7,000,000} = (2^{10})^{700,000}$$

$$= (1000)^{700,000} \approx (10^3)^{700,000}$$

$$= 10^{2,100,000}$$

$$= 1 \text{ followed by } 2,100,000 \text{ zeros}$$

If the eye saw a different pattern every $^1/_{10}$ of a second it would take 30 years to see 10,000,000,000 patterns (1 followed by 10 zeros, not 2,100,000 zeros) and an essentially inconceivable number of years to see all of the different patterns possible to the eye.

When it is considered further that relationships among different patterns are significant in that they provide information on time sequence, changes, motion, etc., so that different groups of patterns and different orders of occurrence within groups are further input data beyond that of the input patterns taken individually, it is clear that the amount of information available from the human eye, the vision input sensor, is immense.

When an image, an input pattern, is projected onto the retina of the eye, a family of signals from the individual sensory elements of the retina is transmitted to the nervous system for processing. The first level of processing (which actually occurs in the eye, in cell layers of the retina) is to identify all of the *first order* universals in the input image. By *first order* is simply meant any universals identifiable at this first level of processing. These are universals that detect or identify: corners, edges, shape types, motion and so forth, universals similar to the cross of the recent example.

The possible number of such first order universals is quite large, large enough in fact to constitute a complete description of the input image, of any possible input image. Such a description for a particular input image consists of all of the universals identified as present in the input image and their location or orientation in the input image, where they occur. The input is converted from being a "bit map", an array of points in a one-to-one correspondence with the original of the image, to an array of characteristics of the input image, the set of *first order* universals that have been identified as present or absent, located in that array according to their locations in the input image.

This new array, the output of the first level of input processing is the input for the next and all further processing. If we could look at that array as an image on a flat screen it would make little sense to us and would not appear to much resemble the original input. That is because the original input has been re-expressed, encoded, mapped into a new terminology different from the one-to-one correspondence with which we are familiar. But, while meaningless to our conscious selves, that information is quite meaningful to our nervous system. It is the kind of information needed by our nervous system (needed by any rational mechanism) in order to effectively process, to understand and use input information.

However, further processing of the input, the using and understanding of it, must be set aside for the moment in favor of concentrating attention on how the *first order* perception of universals actually takes place. It is convenient here to facilitate discussion and understanding of the processes to deem the images and their realization on the retina as "black and white" not the colors of human vision. Likewise, the process discussed in terms of vision applies to any and all types of input: images, sounds, smells, etc.

If we refer to each of the *7,000,000* sensors in the retina individually as *#A*, *#B*, ... for all *7,000,000* of them, then any single image projected on that retina can be represented as the *and* of the signals from all of the *on* sensors *and-ed* with the *and* of the *not* of the signals from each of the *off* sensors. For example

(3-3) Some image = ABCDEFG … [7,000,000 letters].

A group of input images, each individual one represented in the form of equation *(3-3)*, could be described as a group by the *or-ing* together of the equation *(3-3)* type expression for each of the images of the group. The expression for any single image, image *#1* for example, identifies it as the image having (for example) *Sensor A on and Sensor B off and Sensor C off and* .... The expression for the group of images describes the group as (for example) *Image #1 or Image #2 or Image #3 or* .... It would appear (for example) as

(3-4) Some group of images =

$$= A\underline{BC}DE\underline{F}G... + AB\underline{CD}EF\underline{G}... + \underline{A}BCD\underline{E}FG... + ...$$

where the total number of letters = 7,000,000 letters per image times the number of mages.

Such an expression would be the universal of that group of images. That is, any image belonging to the group matches or fits a part of the expression and any image not a member of the group fails to so satisfy the expression. If an image is tested against the expression then a Boolean output result of *1* or *yes* or *on* or *expression satisfied* means that the image being tested exhibits the universal of the group. If an image is tested and produces a *0* or *no* or *off* or *expression not satisfied* Boolean output result that failure is a signal that the image being tested does not exhibit the universal of the group.

These kinds of Boolean logical expressions are readily implemented electronically with simple devices called *logic gates* that produce the *and-ing* and the *or-ing* and devices called *flip-flops* that represent the Boolean variables (*A*, *B*, etc.) and remember their current value. They also yield the *not* operation where called for.

However, there are several problems with this approach to constructing a mechanism to recognize and implement universals. The first is that the large number of variables makes the Boolean expressions much too large and cumbersome. Implementing those expressions electronically requires far to many logic gates and flip-flops. As a practical procedure it is unworkable.

In addition, however, and far more serious as a problem, is that this procedure can only correctly test input images that were used in the original setting up of the expression. It is unable to generalize, "to get the idea" of what the universal is, and apply that learning to correctly treating new images never before experienced. In the above approach the universal detecting mechanism must be constructed from the beginning using all possible examples of the intended universal plus all possible examples that are not of the universal. Not only would such a device be far too large and expensive; most likely it is impossible to even identify all of the possible input cases called for.

In other words, such a system has no ability to learn, to modify and improve its behavior on the basis of experience. That defect makes the system far too cumbersome to be practical and also leaves the system not corresponding to that which we know about rational systems -- rational systems do learn. Not only do intelligent humans learn; all animals having some form of nervous system exhibit some learning,

learning that varies from the sophistication of chimpanzees to the much simpler, yet still quite complex, worm.

Referring to equation *(3-4)* again, suppose that every input image that exhibits the universal of interest has sensor *#B = on* regardless of the state of any of the other sensors. Likewise suppose that every input image that does not exhibit the universal of interest has sensor *#B = off* regardless of the state of any of the other sensors. Then sensor *#B* alone would represent the universal. The logical expression to represent the universal and test for its presence or absence in input images would be very simple -- a case of examining sensor *#B* and ignoring the rest of the image for this purpose.

In general it is the nature of universals that they exhibit such simplified expressions although not necessarily nor usually as radically simple as the example just used. A universal is a kind of generalization, an omission of non- relevant specifics in favor of a focus on the broad commonality. Its expression tends to be simpler than the expression for the collection of all images exhibiting the universal and all that do not. This simplified representation of commonality among input images is precisely what a universal is.

The problem at this point is, then, how does a rational system operate in a fashion that overcomes the above problems ? How does it extract a simplified universal from a group of sample inputs ? How does it develop the ability to recognize an input never before experienced ? How does a rational system learn ? For, the process of extracting simplified universals from a partial set of input examples is what learning is about.

PARAMETERS OF THE HUMAN BRAIN

As if the foregoing visual complexity were not enough, it is minor, if not minute, compared to the complexity of the human brain, the principle neural mechanism.

The four by four array examined earlier contained only *16* discrete elements -- in effect neurons. Yet that array is capable of representing

$2^{16} = 65,536$ different patterns.

The human brain contains on the order of *one hundred billion* neurons, about 10^{11}. Let us arbitrarily assign *10%* of those to sensory, motor, automatic (for example heart beat, breathing) and intercommunication activities within the body and brain. (That is quite generous. A *Tyrannosaurus Rex* had a total brain size of fewer than *10%* of a human's number of neurons for all purposes yet it did a pretty good job of functioning.)

Let us then recognize that the complex human brain has a number of regions of specialization. One local region interprets vision; another deals with language, another handles emotion, another does abstract reasoning, and so forth. Let us provide for one hundred such sub-systems. Then any one such sub-system would have

(4-1) *total less [10% for body systems] less [100 sub-systems] =*

$$= 10^{(11 - 1 - 2)} = 10^{8} \text{ neurons}$$

and could represent

(4-2) $$[2]^{10^8} = [2^{10}]^{10^7} \approx (1000)^{10^7} \approx (10^3)^{10^7} = 10^{3 \cdot 10^7}$$
$$= 10^{30,000,000}$$
$$= 1,000, \ldots \text{[30 million zeros]} \ldots ,000.$$

different patterns per each such sub-system.

Even our neural system, having that great capacity, is not able to really appreciate what an immense number that is. At the rate of a page being able to contain about 3,000 zeros it would take 10,000 pages of zeros just to write out the number -- to write it down not to express the value of the number. (It takes four digits to write down "1000" but it has the numerical value *1,000* two hundred fifty times greater.)

That vast capability certainly suffices for our neural system, our brain, to readily learn and retain everything that we give it over a lifetime.

> But, how do we deal with, how does our brain deal with, all that size, all that amount, all that complexity ?
>
> It is now time to investigate the neural logic that operates the brain.