PART IV

ON THE MECHANISM OF INTELLIGENCE AND ITS ORIGIN

SECTION 22

The Problem of Intelligence

Introduction

The problem is to explain the phenomenon of human intelligence. Scientific knowledge has developed to the point where there is generally a sound understanding of most phenomena in the world, and for those phenomena not yet thoroughly understood there is confidence that development of the knowledge is only a matter of a little more time. But, for the phenomenon of human intelligence there is no well developed scientific explanation corresponding to that for evolution, physics, or biology.

To obtain a purely material explanation of the universe in all of its aspects, part of the objective set forth in the first section of this work, it is necessary, then, to demonstrate that the origin and mechanism of human intelligence are purely material results of the material universe thus eliminating any reliance on the hypothesis of god. This is done in the next several sections of this work by presenting a description of the mechanism by which intelligence operates. It is shown that intelligence and all of the higher order intelligent activities (e.g. those of consciousness, creativity, emotion, conscience, abstract thought, etc.) are the natural behavior of any brain-like structure and can, in fact, be realized in a man-made such structure.

The human brain and nervous system is a very complicated and sophisticated system. It not only performs the human functions of thought, intelligence, self-awareness, and so forth, but the lesser functions found in most animals such as purposive behavior and control of voluntary actions of the body. Furthermore, it is also an involuntary control system that monitors and controls all of the bodily functions so as to make the total biological system of the person (or animal) function in its best overall biological interest.

For example, the brain and nervous system control:

- fuel and materials input (food)
- oxidant input (breathing)
- processing and distribution of these (digestion, blood circulation, waste elimination)
- temperature control
- growth and repair
- reproduction, and so forth.

Involved in these processes are systems of nervous and chemical (endocrine) signals and controls and semi- and fully-automatic sub-systems (heart beat, reflexes, etc.)

This system operates on an evolved design. Humans have sub-systems quite like those of lesser animals, These are apparently retained as evolutionary "carry-overs". They can also be viewed as the retaining of well-developed and well-proven systems of evolutionary precursors upon the base of which, as sub-systems, the more sophisticated human systems are built. A partially true biological paradigm is that "ontogeny recapitulates phylogeny" or, in other words, that the development of the fetus from conception to birth recapitulates the evolution of the specie. An analogous, and related partially true paradigm is that the evolved human nervous system recapitulates and has as functioning sub-systems the evolutionary history of the earlier developed stages of nervous system.

Whether, if one were to design a human "from scratch", one would include all of these mechanisms is a hypothetical question to be perhaps answered in the future. Certainly most of the functions would appear to be needed. However, for the present purposes the issue is intelligence, explanation of that high order human function. Digestion is well understood by science and humans have no monopoly on the process. The same is true for reflexes, temperature control, and so forth. Consequently there is no attempt here to go into the detail of the brain's control of all those type activities of the human brain.

The objective is intelligence. How do we see, think, remember, know ourselves, learn, plan, create ?

In setting out to describe and explain these sophisticated functions, probably the most complex and sophisticated in the universe to our knowledge, it is necessary to start with simple first steps, building blocks, and gradually erect the total structure.

That procedure is followed in the next several parts. The reader is urged to be patient with the review of fundamentals in the earlier portion, which lays the basis for the development.

OVERVIEW

Until assigned a name, things are identified by their description. For example the letter "t" in the last word of the prior sentence can be described fairly definitively as: roman letter "t", in the last word of the prior sentence, black, on a white background, Times New Roman font, size eleven point, lower case, non-italic. Each of the components in that description can apply to a variety of other things, but together they specify the particular instance. A number of other things have some, but not all, of the characteristics of that "t" and have other characteristics that the "t" does not.

The specific individual momentary concepts in our heads are likewise describable in terms of a set of characteristics -- ones that collectively are the particular concept of that instant, ones that are partially shared with a variety of different other concepts.

The process that goes on in our minds is a progression of such specific momentary concepts, thoughts. Successive thoughts are linked by having most of their characteristics in common but one or more changed. A chain of such successive thoughts is thinking. In the following analysis and development the characteristics are referred to as universals.

Our minds have thoughts by supporting representations of universals and by detecting various universals amid a mass of other data. We think by chains of successive specific momentary sets of universals progressing from set to slightly different set in a systematic (logical, rational) fashion. But, ... how ?

SECTION 23

Universals and Perception - (1)

By the word *universal* is meant a class or category to which a particular specific example belongs. *Perception* is the process of properly correlating individual specific cases, examples, with universals. Perception particularly includes the proper identifying or recognizing of examples that have not previously been specifically experienced.

Humans perceive, recognize, a very large number of universals, of course. Some examples, in order to clarify the concept, are:

- recognition of the letter *E*, whether capital or lower case, hand written or mechanically produced, large or small, alone or among other symbols, even though the particular *E* being recognized may be different from any ever before seen;

- recognition of all beings that are human as human beings;

- recognition of all shirts.

The universal is the common characteristic of all elements of the group, that is *E*-ness, human-ness, shirt-ness in the above three examples.

Not only humans recognize universals; most animals do also, but the ability in non-humans is apparently more limited. Nevertheless, for example, a dog can recognize another dog as a dog even though the dog recognized was never before seen and is of a significantly different breed or appearance.

Recognition of universals is not always accurate even though the recognizer is competent. The sample may be a marginal case. For example, everyone is familiar with the problem of reading another person's handwriting, which involves properly recognizing various sample letters as samples of particular letters of the alphabet, which is a set of letter universals.

The process of perception involves an input, a data processor, and an output. For the present case the input is data from a sensory organ: eyes, ears, nose, etc. The processor is some mechanism that operates on the input data so as to correlate examples with universals. The output is data representing that correlation or identification. Of course a given input sample may be a sample of a number of different and perhaps unrelated universals. For example a particular letter *E* might belong to all of the following classes simultaneously: *E*, upper case, small, hand written, in ink, red, moving left to right across the field of vision, upside down, appearing progressively smaller, etc.

The process of recognizing universals is most easily understood by using the case of the sense of sight as the input. The procedures and conclusions apply equally to the other senses or to any coherent or systematic input system. For the purposes here, the sample is projected onto a screen (or the retina of the eye). The screen is not continuous, however. Rather, it is divided into an array of more or less uniformly spaced essentially identical sensors (the "rods and cones" of the retina). Each individual sensor can only register in an on-off manner (for the present); that is, if the part of the image projected onto the screen and falling on a particular sensor is light then the sensor is in the *on* state, if dark the sensor is *off*.

Thus the image projected onto the screen is represented on the screen as an array of black and white dots (*off* and *on* sensors) similar to a photograph in a newspaper as viewed with a magnifying glass.

To initially discuss the process an example using a relatively small array of sixteen sensors arranged in a square of four rows of four sensors each will be used as in Figure 23-1(a), below. It is necessary to be able to refer to each of the individual sensors (elements) of the array. This could be done by sequentially numbering them as in Figure 23-1(a); however, it will be more useful to use the system of Figure 23-1(b), in which the array is divided in half four different ways.

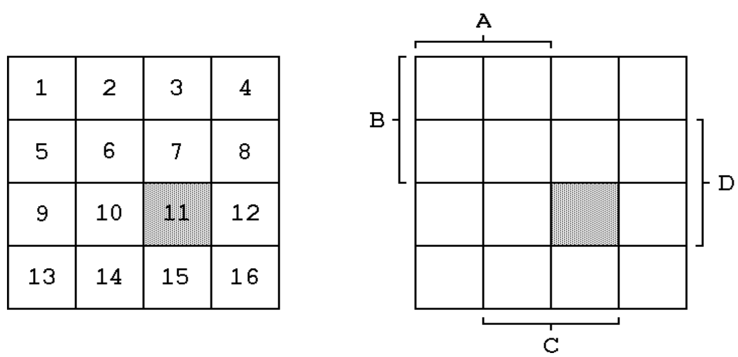


Figure 23-1

(The procedure being used is, of course, the digitizing of the image into binary elements and the description of the sensors and their binary states by means of Boolean algebraic variables and functions. In fact the eye, also, essentially digitizes the image on its retina and supplies signals that are essentially binary to the brain; however, the human processor is not quite Boolean. Boolean discussion will be used for the moment and the conversion to the biological mode of processing will then be presented. For those who are not familiar with these techniques the explanation is continued in simplified terminology.)

The half of the array of Figure 23-1(b) that is labeled *A* will be called *A*. The other half will be called *not A* and be written $\overline{A}$. We can then identify element number #11 of Figure 23-1(a), for example, in Figure 23-1(b) as being in

(23-1) $\overline{A}$ and $\overline{B}$ and C and D.

a description that fits no other element of the array.

This procedure makes use of Boolean Algebra, a mathematics of logic originally developed by the Englishman, George Boole, for the purpose of testing and interpreting the logical construct of verbal statements. Although it was developed well before even the notion of digital computers had occurred or could have occurred, Boolean logic is the underlying principle on which digital computers operate.

The letters *A, B, etc.*, are called *variables* meaning that they may vary in value. The allowed values in the present case are *1 ≡ on* or *"true"* or *"yes"* and *0 ≡ "not on"* (i.e. *"off"*) or *"not true"* (*"false"*) or *"not yes"* (*"no"*).

Instead of writing "and" over and over as in equation 23-1 the notation

(23-2) $\bar{A}\bar{B}CD$ or, when needed for clarity $\bar{A} \cdot \bar{B} \cdot C \cdot D$

will be used and understood to mean the same as equation 23-1. It is read as *"not A and not B and C and D"* and means the state in which *A is not true*, *B is not true*, *C is true* and *D is true*. It also means that portion of the array of Figure 23-1 which is *not in A*, *not in B* and is *in C* and is *in D*.

To refer to more than one element of the array at a time the connective *or* will be used, written as *+*. Thus to refer to the combination of the elements #10 and #11 of the four by four array of sixteen elements (per Figure 23-2, below) the reference is

(23-3) $A\bar{B}CD + \bar{A}\bar{B}CD$

which is read as *"A and not B and C and D or not A and not B and C and D"*.

This reference (equation 23-3) can be stated more simply as

(23-4) $\bar{B}CD$

because if it is true for *A* or for $\bar{A}$ then it is independent of the value of *A*, whether it is *0* or *1*. That is, for the two elements, #10 and #11, as an area of the array, designation in terms of *A* is to no point. As Figure 23-2 shows, that area is correctly described as the area simultaneously in *not B and C and D* as equation 23-4 presents.

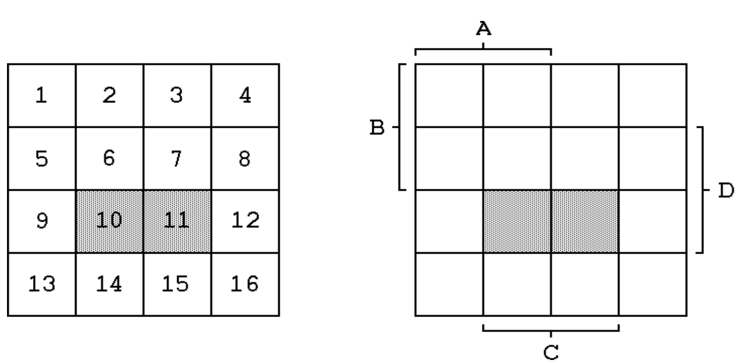
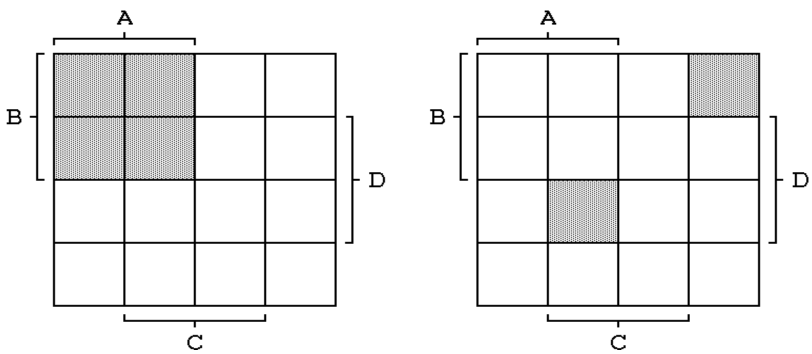


Figure 23-2

Other combinations of elements may yield similar such simplifications of the reference and still others may not. For example, the indicated set of four elements in Figure 23-3(a), below, is

$$AB\overline{C}\overline{D} + ABC\overline{D} + AB\overline{C}D + ABCD \equiv AB \qquad (23\text{-}5)$$

On the other hand, the two elements indicated in Figure 23-3(b), below,



(a) *(b)*

Figure 23-3

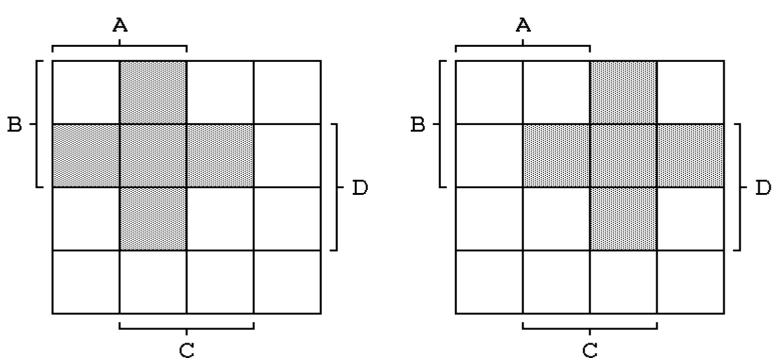
are given by

$$\overline{A}B\overline{C}\overline{D} + A\overline{B}CD \qquad (23\text{-}6)$$

which cannot be further simplified or reduced.

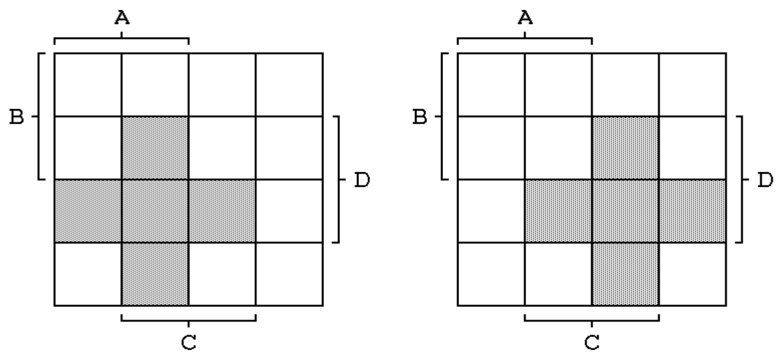
Now let us consider the problem of recognizing (that is identifying the universal of) a simple cross, a horizontal line crossing a vertical line, in this system. More specifically, we wish to obtain a method for recognizing any such cross and only such crosses. Within the special case of the sixteen element array we wish to be able to properly assign any input, as it is projected onto the array, as a member or a non-member of the universal *cross*.

First we consider some examples of the specified input as in Figure 23-4, below and continued on the next page.



(a) *(b)*

Figure 23-4



(c) *(d)*

Figure 23-4 (continued)

The elements making up each sample are as follows.

(23-7) (a) $ABC\overline{D} + AB\overline{C}D + ABCD + \overline{A}BCD + A\overline{B}CD$

(b) $\overline{A}BC\overline{D} + ABCD + \overline{A}BCD + \overline{A}B\overline{C}D + \overline{A}\overline{B}CD$

(c) $ABCD + A\overline{B}\overline{C}D + A\overline{B}CD + \overline{A}\overline{B}CD + A\overline{B}C\overline{D}$

(d) $\overline{A}BCD + A\overline{B}CD + \overline{A}\overline{B}CD + \overline{A}\overline{B}\overline{C}D + \overline{A}\overline{B}C\overline{D}$

These can be simplified by expression in terms of two-element areas as follows

(23-8) (a) $ABC + ABD + ACD + BCD$

(b) $\overline{A}BC + \overline{A}BD + \overline{A}CD + BCD$

(c) $A\overline{B}C + A\overline{B}D + ACD + \overline{B}CD$

(d) $\overline{A}\overline{B}C + \overline{A}\overline{B}D + \overline{A}CD + \overline{B}CD$

(That these expressions include the central element of the cross four times instead of once is mere redundancy and does not affect the accuracy or effect of the expression.)

The commutative principal of mathematics applies to this mathematics; that is, the order of stating variables has no effect on the result. For example

$$\overline{A}BC + A\overline{B}C = A\overline{B}C + \overline{A}BC = B\overline{A}C + CA\overline{B}$$

Likewise, the associative mathematical principal also applies; that is, factoring and the related grouping of variables has no effect on the result. For example

$$A\overline{B}C + AB\overline{C} = A\cdot[\overline{B}C + B\overline{C}]$$ [A is "and-ed" with the bracketed expression.]

Making use of those principals equation 23-8(a) can be expressed as

(23-9) AB[C + D] + CD[A + B]

or as

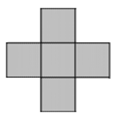
AC[B + D] + BD[A + D]

or as

AD[B + C] + BC[A + D]

Each of equation 23-8(b), (c) and (d) can be similarly expressed. All of the resulting formulations have the same general form:

(23-10)



$$= V1 \cdot V2 \cdot [V3 + V4] + V3 \cdot V4 \cdot [V1 + V2]$$

where Vi, i = 1, 2, 3, or 4, is a Boolean variable like A, B, etc.

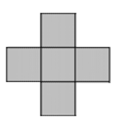
That is, the cross in the four by four element array being treated, is apparently characterized by an identification in terms of the four Boolean variables as the and-ing of any two variables with the or of the other two, that whole then or-ed with the and-ing of the other two variables with the or of the first two, any or all of the variables being natural or not-ed consistently.

Examination of the four by four arrays being used demonstrates the validity of the following identity

(23-11) $A + B = [\,\overline{\bar{A} \cdot \bar{B}}\,]$

with the use of which equation 23-10 can be rewritten as

(23-12)



$$= V1 \cdot V2 \cdot (\overline{\overline{V3} \cdot \overline{V4}}) + V3 \cdot V4 \cdot (\overline{\overline{V1} \cdot \overline{V2}})$$

Either of the two logically equivalent formulations, equation 23-10 or 23-12, is the extraction of the indicated universal, cross-ness from the samples with which the analysis began. Either formulation is, therefore, the means to the perception of that universal in the sample array being studied.

So far in this analysis four sample crosses have been examined. In the simple four by four array being studied there are a total of ten possible symmetrical crosses. The other six are displayed in Figure 23-5 on the following page. All of the ten would be correctly identified by the formulation just derived.

Some asymmetrical crosses would also be identified by the formulation, for example as in Figure 23-6. For each symmetrical cross consisting of five elements there are three ways that it can be asymmetrical: horizontally, vertically or both. Thus each such cross can appear in four forms. The total number of possible crosses, symmetrical and asymmetrical is that four times the eight possible symmetrical five-element crosses, equals 32. That plus the two larger crosses equals a total of 34.

The total number of different patterns that can be displayed on this sample sixteen element array is

$$2^{16} = 65{,}536$$

The logical formulation just developed detects the *34* cases having the universal *cross-ness* in common out of the *65,536* total cases. (If one wished the universal to be so defined as to reject the asymmetrical crosses, it can be done with only a little more complexity.)

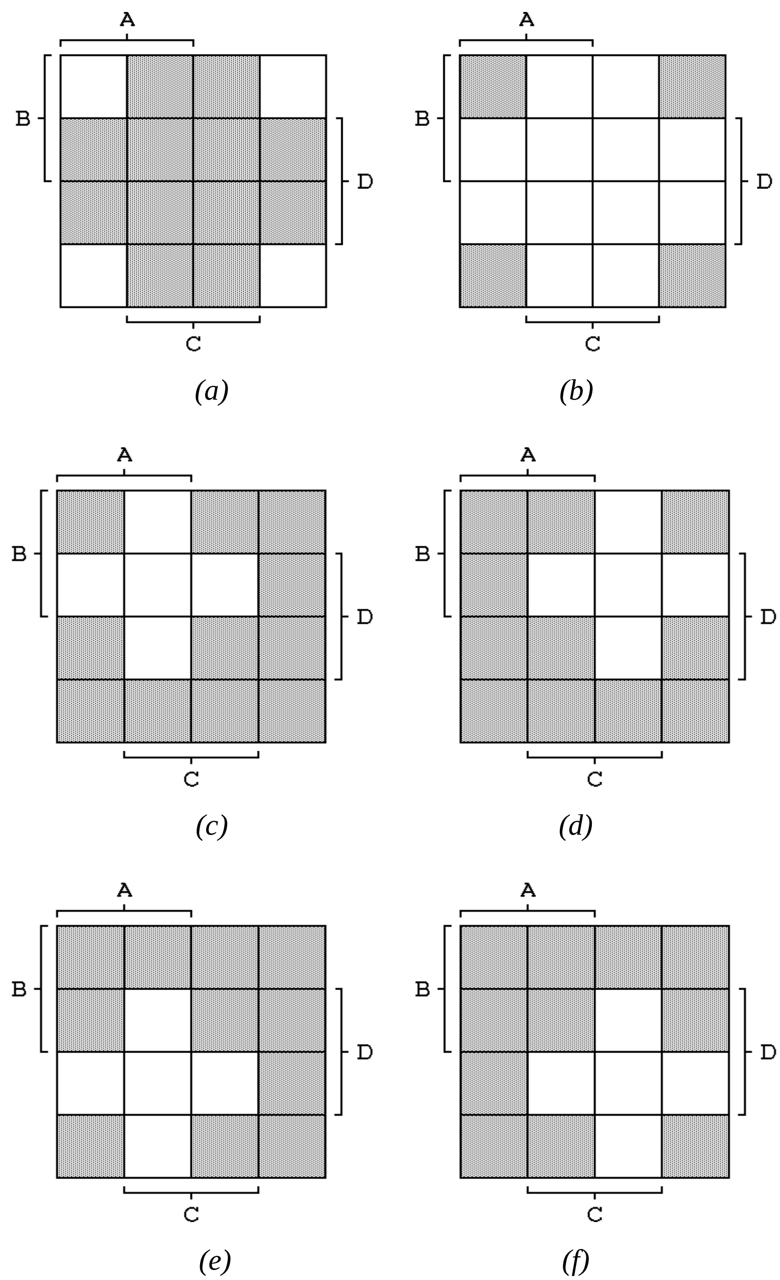


Figure 23-5

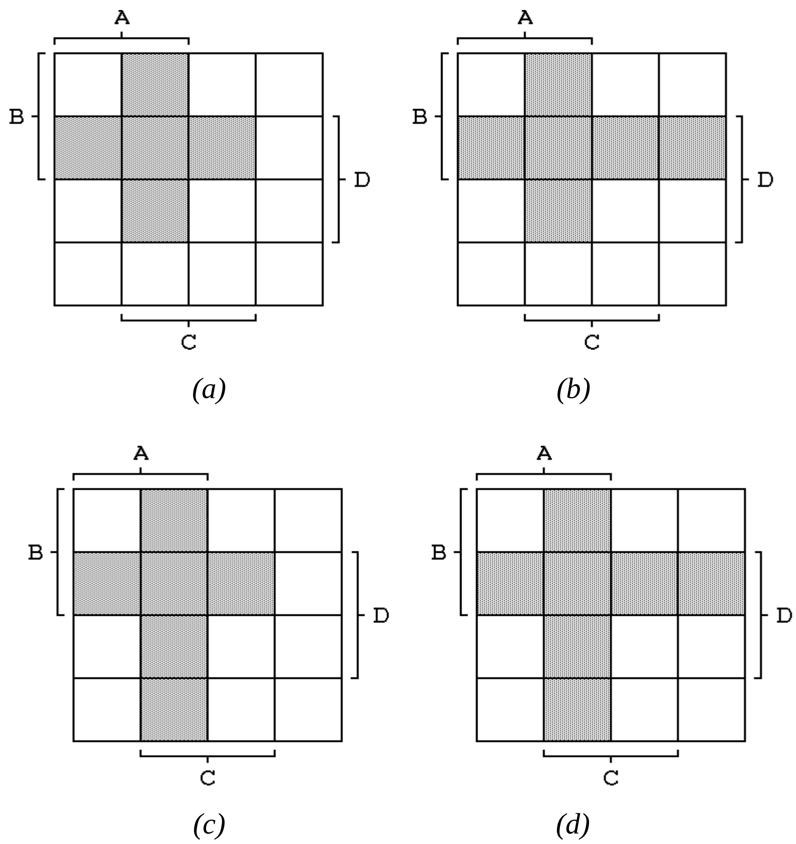


Figure 23-6

At this point the example of the four by four array of sixteen elements can be abandoned in favor of the general case of a practical, human perception system. The purpose of the example was to illustrate in a general sense that:

· a sensory input can be analyzed,

· a limited number of input samples can be sufficient to reasonably well establish a formulation for a universal,

· that can be done by digitizing the data into binary representation,

· and the formulation of the universal can be of a kind corresponding to well known digital logic arrangements as used in digital computers and some automatic control systems.

(However, intelligence functions quite differently from the functioning of a digital computer.)

Complex Perception Systems

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 on a binary digital array is

(23-13) 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. This is an extremely large number.

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

(23-14) $$2^{7,000,000} = (2^{10})^{700,000}$$

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

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

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

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 an array of points in a one-to-one correspondence with the original of the image (each point being light or dark, *on* or *off* as its corresponding point in the original) 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 location in the input image.

This new array, the output of the first level of input processing is the input for 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.

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

(23-15) Some image = $A\overline{B}\overline{C}DE\overline{F}G\ldots$ [7,000,000 letters].

A group of input images, each individual one represented in the form of equation 23-15, could be described as a group by the *or-ing* together of the equation 23-15 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

(23-16) Some group of images =

= $A\overline{B}\overline{C}DE\overline{F}G\ldots + AB\overline{C}\overline{D}EF\overline{G}\ldots + \overline{A}BCD\overline{E}FG\ldots + \ldots$

[Total number of letters = 7,000,000 letters per image times the number of images.]

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 *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 to 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 23-16 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.

NEURAL-TYPE LOGIC DEVICES

The *neuron* is a special type of biological cell which is the operating component in the nervous system of all life on Earth that has a nervous system, whether human, animal, insect or whatever. By *neural-type logic* is meant systems in which the principal operating component is the *neuron* or systems in which the principal operating component is a device, a man-made device, that operates *logically in the same way as a neuron*.

The logic technique used in such neural-type rational systems, including the human brain, is slightly different from the *and / or* logic examined so far. The basic logic function (procedure) used in biological systems is *majority logic*. Using the notation `M(...)`, where the `M` stands for `majority of` and the variables involved (e.g. array or retina element signals) are listed in the parentheses, then a translation between *majority* and *and / or* logic is, for example

(23-17) M(A,B,C) = AB + AC + BC

That is, a majority logic operator has an output of *on* if a majority of its inputs are *on* and otherwise an output of *off*. In the example of equation 23-17 any two of the three variables is a majority of them.

For convenience of notation, and because Boolean algebra employs binary logic (a logic based on the binary number system having base *2* instead of *10* and digits *0* through *1* instead of *0* through *9*), the binary digit *1* will be used to represent *on* or *yes* or *satisfied* hereafter with regard to Boolean algebra expressions and the digit *0* to represent the opposite. In those terms equation 23-17 states that the output is *1* if any two or all three of the inputs are *1*. Otherwise the output is *0*.

In addition to variables such as the *A*, *B*, etc. already used, majority logic can also use *logic constants*. Here a constant is like a variable in all respects except that it always has the same, fixed value. Since the system is binary there are only two values that a constant can have, *1* or *0*.

In *and / or* logic, constants are essentially meaningless as the following examples illustrate.

(23-18) A + B + 1 = 1 [In spite of the variables the result is always "1". The variables are meaningless because of the constant.)

A + B + 0 = A + B [The constant has no effect.]

A·B·1 = A·B [The constant has no effect.]

A·B·0 = 0 [In spite of the variables the result is always "0". The variables are meaningless because of the constant.)

However, in *majority logic*, *constants* play a useful and important role; they enable *majority logic* to represent *Boolean logic*. For example:

(23-19) M(A,B,1) = A·B + A·1 + B·1 = A + B

M(A,B,C,1,1) = ... = A + B + C

M(A,B,0) = A·B + A·0 + B·0 = A·B

M(A,B,C,0,0) = ... = A·B·C

The *not* operation still applies in *majority logic*; that is, the majority operation may operate on *natural* or *not-ed* variables. For example

(23-20) $M(A,\overline{B},\overline{C},1,1) = A + \overline{B} + \overline{C}$

$M(\overline{A},B,0) = \overline{A}\cdot B$

Thus majority logic with both constants and variables can produce all of the fundamental type logical constructs that Boolean logic uses.

Likewise, a majority operation's output can be an input variable in another majority operation just as in Boolean logic. For example

```
           ┌-------┐
(23-21) M(A,M(B,C,0),1)
           └-------┘
```

where the bracket indicates the "*The Majority of B*, *C* and *0*" as one of the variables in the overall expression, which reads as "*The Majority of A*, *The Majority of B*, *C* and *0*, and *1*. Such complex majority operations, which can have many more levels than the two-level case illustrated in equation 23-21, enable majority logic to implement any Boolean logic whatsoever.

In fact majority logic can do more than that. The very same physical structure, that is the same connection of inputs to a given majority processor, can yield controllably different logical constructs, logical results, depending on the value of the constants applied to that majority processor. Majority logic makes possible fixed "pre-wired" interconnections in a configuration where the logical effect of the physically fixed structure can be controlled and varied by varying the values of the constants involved.

That is precisely the process that goes on in a rational system based on neurons, whether that system is in a human, a cow, an ant or whatever. The inputs to a neuron are the outputs of other neurons or of sensors (e.g. the retina of the eye). Those inputs are such that some act on the neuron in an *excitatory* fashion and some act on it in an *inhibitory* fashion. That is, *excitatory* inputs are analogous to *natural* variables (as opposed to *not-ed* ones) and have the logical effect of an input of *1* if activated and *0* if not. *Inhibitory* inputs are analogous to *not-ed* variables and have the logical effect of an input of 0 *if activated and* 1 if not.

In a neuron the presence or absence of a majority is not determined by counting the total possible inputs and comparing the number of them that are *1* to that count. Rather the effect is as if the *1* inputs are each *+1* (excitatory) and the *0* inputs are each *-1* (inhibitory). If the algebraic sum, the excitatory plus the inhibitory (the number of excitatories less the number of inhibitories), is greater than zero then a majority is present.

There is still another component of a neuron's operation, however. That *algebraic sum* of the excitatory *+1* and the inhibitory *-1* inputs is not compared to *zero* as such. Rather it is compared to a *threshold* level present in that neuron. If the *threshold* happens to be *zero* then the logical construct of the neuron is simply the majority of its inputs.

But, if the threshold is greater than *zero*, meaning that for the neuron to have an output of *1* the number of excitatory inputs must be that much (the *threshold* amount) greater in number than the number of inhibitory inputs, then the effect is the same as if there were as many constants equal to *0* present and acting as the level of the *threshold*. Likewise, a *threshold* less than *zero* corresponds to there being that many constants equal to *1* present and acting. Thus the value of the *threshold* represents the net value of constants in the input and variation of the *threshold* produces variation of the net value of the constants which produces variation in the Boolean logic that the majority operator is equivalent to.

For example, if the inputs to the neuron are *A*, *B*, *C*, *...* and all of them are excitatory (simply for this example), then:

(23-22)	With Threshold	The Neuron Performs
	0	M(A,B,C,...)
	+1	M(A,B,C,...,0)
	+2	M(A,B,C,...,0,0)
	-1	M(A,B,C,...,1)
	-2	M(A,B,C,...,1,1)

The threshold is equivalent to the net number of constants involved, constants of *-1* for positive threshold and of *+1* for negative threshold. The output is *1* if the majority of the input variables and those constants is greater than *zero*.

But, the special power of the neuron is that its threshold can be changed. That means that its constants can be changed and that means that the logical effect, the Boolean logic that the neuron is implementing, can be changed. The neuron "remembers" the value of the threshold so that the threshold is, in that sense, some set number of majority logic constants operating as such in the logical construct that the neuron effects. However, that set value or level of the threshold can be changed, adjusted so that the logical construct that the neuron effects is slightly, gradually changed. It is that process that enables learning. Learning is, in effect, the directed adjustment of neural thresholds to achieve the desired result.

The input to the neuron from other neurons or from sensors is received by the neuron as various excitatory and inhibitory, *+1* and *-1*, inputs. The neuron emits an output that is *1* or *0* depending on the internal operation of the neuron. That output acting as an excitatory input to another neuron is a *+1* input to it if the output was *1*. That output acting as an inhibitory input to another neuron is a *-1* input to it if the output was *1*. The internal operation of the neuron simply determines whether the majority of the inputs plus the threshold is greater than zero (neuron output is *1*) or not (neuron output is *0*). (How the threshold changes occur will be treated shortly, in the next section of this work.)

Actual biological neurons operate in this manner. A single biological neuron consists of a central cell body, a number of input lines (filaments or fibers of cell material) called dendrites, and an output line (also a filament or fiber of cell material) called an axon. Output signals of neurons travel to the end of the axon where they then communicate, as inputs, with the dendrites of other neurons. The junction where the signal transmission from neuron to neuron takes place is called a synapse. Within a neuron some of the dendrites (inputs) are excitatory and some are inhibitory. The threshold, at the main cell body, determines whether the net effective input signal causes or fails to cause an output signal on the axon. The processes within the neurons and at the synapses are electrochemical in nature.

When neurons, whether biological or man made neural-type electronic devices, are interconnected so that the outputs of some neurons are inputs to other neurons then a multilevel neural network exists. Such a network makes possible neuron-implemented complex majority logic structures that can effect logic such as illustrated in equation 23-21. Multilevel networks of neurons use the neuron's majority logic, modified by the individual neuron's thresholds, to represent the equivalent of complex Boolean logical descriptions. Such descriptions are the logical representation of universals. Complex neural networks can thus represent specific universals if the individual neural thresholds are correctly set to make them do so.

Let us now operate a simple such neural network using as its input the sample four-by-four, *16* element, array used in the first part of this section. That array was there used to illustrate the universal *cross-ness* among the various possible images that could appear as input on the array.

An individual neuron or neural-type device will be symbolized as in Figure 23-7, below.

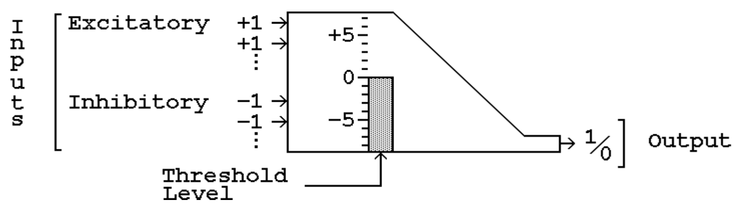


Figure 23-7

The outputs of the four-by-four array will be interconnected to the inputs of a number of such neurons and then the outputs of those *first level* neurons will be interconnected to the inputs of one more neuron. The output of that final, single, neuron will be deemed the representation of the action of the entire neural network. (See Figure 23-8 on the following page.)

But, how should the interconnections be made; that is, which sensors should be connected to which inputs of which neurons ? This question is quite fundamental to neural networks as is the matter of how threshold changes occur. As with the control of threshold changes, the subject will be treated fully in the following section. For the moment let us assume that those aspects of the problem have been correctly implemented in the sample neural network being used.

Let us now teach the neural network to recognize the universal *cross-ness*; that is, let us cause it to learn how to discriminate between input images exhibiting *cross-ness* and those lacking it. Our objective is that the neural network should give an output of *1* if the input image has *cross-ness* and *0* otherwise.

We use the following procedure.

(1) Show the input array an input image (project an image onto the four-by-four, *16* element array). That is, cause various of the *16* elements in the array to be *on* and others *off* so that the desired pattern is represented on the array

(2) (Being the teacher in this case, the authority, we) note whether the image exhibits the universal *cross-ness* or not. (The problem of where, in general, the teacher comes from is also addressed in the next section.)

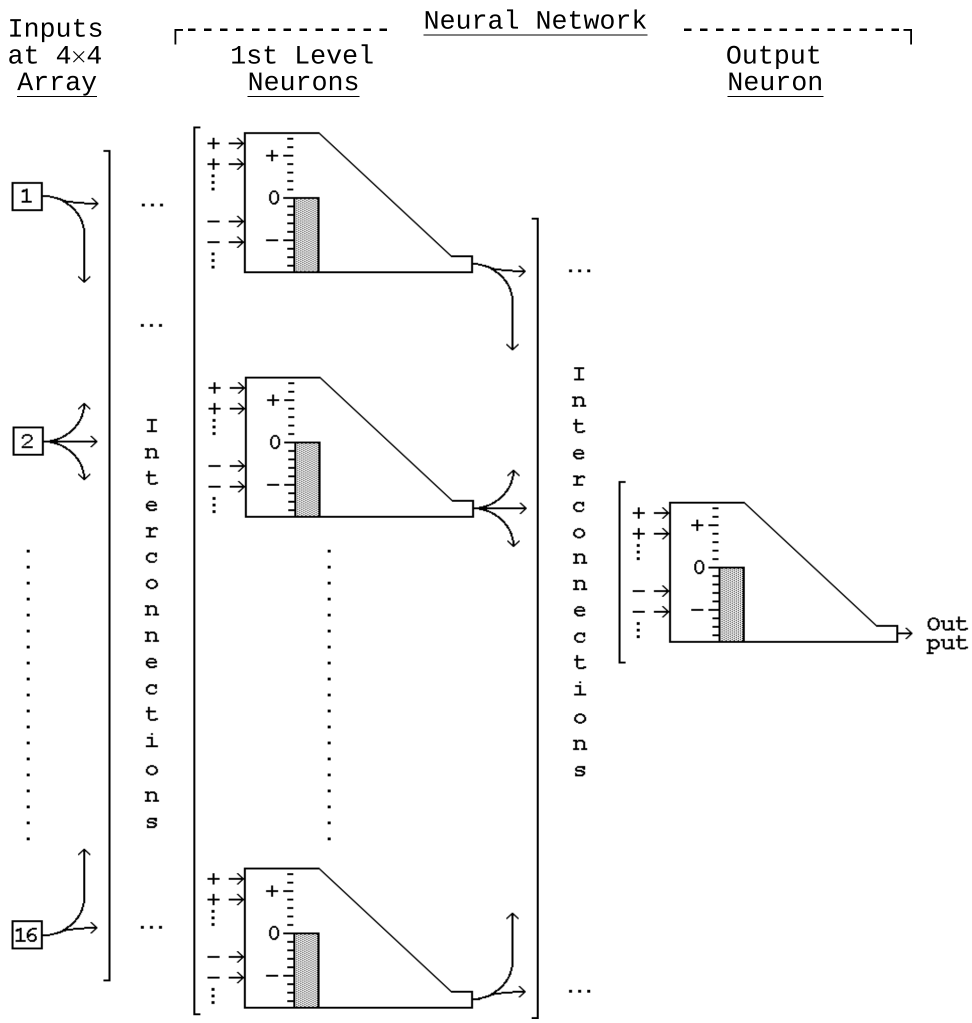


Figure 23-8

(procedure continued)

(3) Observe the output of the neural network (whether it is *1* or *0*).

(4) Evaluate the performance of the network, which could be any of the following four possible cases.

Input Image	Output	Result
cross	1	correct
cross	0	wrong
not cross	1	wrong
not cross	0	correct

(5) Change the threshold of each neuron of the neural network as follows:

- If the neural network output was correct reinforce that behavior by adjusting each neuron's threshold in the direction that makes that result more likely.

  · If its output was `1` lower its threshold by `1` unit (making even more likely a `1` output for another input like this one).

  · If its output was `0` raise its threshold by `1` unit (making even more likely a `0` output for another input like this one).

- If the neural network output was wrong discourage that behavior by adjusting each neuron's threshold in the direction that makes that result less likely.

  · If the output was `1` raise its threshold by `1` unit (making less likely a `1` output for another input like this one).

  · If the output was `0` lower its threshold by `1` unit (making less likely a `0` output for another input like this one).

(6) Repeat the above five steps using a different input image each time until the neural network's performance is sufficiently consistently correct.

This has the appearance of a reward-and-punishment type procedure but that is not the case here. The neurons do not understand anything, certainly not reward and punishment. The procedure simply changes the thresholds in a direction tending to increase the chances that for input images similar to the one just processed the neural network's operation on the input variables, with its now changed thresholds, will yield the desired correct output.

But, whether the neurons "understand" this or not is irrelevant. The end result of the process is that the neural network actually becomes able to discriminate *cross-ness* even though at the start of the process it could not do so. The neural network has learned, been taught by the teacher, to discriminate. It effectively *perceives* the *universal* taught, *cross-ness* in this example, having *learned* to do so.

That learning was accomplished by directed, logical adjustments to each neuron's threshold level. Such adjustments have already been shown to change the Boolean logical construct that is effected by each neuron's majority operation in conjunction with the constants represented by its threshold.

In other words, the above described learning process causes the Boolean logical construct or operation that the neural network performs on the input variables to gradually change until it is identical to, or it sufficiently resembles, the Boolean logical construct that corresponds to the universal being taught.

The accomplishment of that is the learning to perceive that universal. The subsequent using of that to make correct outputs in response to input images is the perceiving of that universal.

[This concept and laboratory research with regard to it were first developed and pursued at the Cornell Aeronautical Laboratory in the latter 1950's. The research was reported in the Proceedings of the Electronic Computers Group of the (then) Institute of Radio Engineers, IRE, (now the Institute of Electrical and Electronic Engineers, IEEE) circa 1960. The neuron simulator device, operating as herein described, was called the "perceptron". Laboratory development demonstrated that the type device does learn and operate as here described.

[The first generation of commercially produced machines using these principles were in the mid 1990's appearing on the market and being used. The machines employ neural networks similar to those described above. The machines are used to perceive patterns in data in situations where humans may be too slow or unable to perceive the pattern.]

In general summary so far:

· Perception is the correlating of an experienced example with a universal, a class to which it belongs.

· Learning is the developing of the ability to so perceive.

· The perception is accomplished by having -- the learning is the process of constructing -- a logical mechanism that operates on the experienced example in a fashion that detects the presence or absence of the universal.

· That "logical mechanism" is a physical implementation that is, in effect, a Boolean logical expression that conforms to the universal.

· The "logical mechanism" is "constructed", exists and operates, by means of majority logic with constants as implemented by neurons or neural-type devices having majority logic and adjustable thresholds.

While this process has been discussed in terms of our sense of vision the same process operates with regard to all of the senses: hearing, smell, touch, etc. Hearing involves the universals in sounds and hearing and understanding language involves universals just as numerous and complex in their effect as in the case of vision. The blind read by their sense of touch and process a similarly numerous and complex set of universals through their fingertips. And some of the animals, unlike we humans, derive quite extensive information from their sense of smell.