Thoughts and Thinking – the Mechanism of Intelligence

by

Roger Ellman

Abstract

Most discussions of the mind or brain focus on the "hardware", the neural structure and its biological functioning. But, it is the "software", how the neural components logically interact, that produces the higher level results that we experience in our minds.

The development begins with universals and then proceeds through perception, learning, and the processing of universals to: concepts, thoughts, thinking, purposive behavior, memory and instinct.

Roger Ellman, The-Origin Foundation, Inc.
320 Gemma Circle, Santa Rosa, CA 95404, USA
RogerEllman@The-Origin.org
http://www.The-Origin.org

Thoughts and Thinking – the Mechanism of Intelligence

by

Roger Ellman

1 - INTRODUCTION

This analysis addresses how thoughts and thinking are realized in neural systems whether biological or not.

A universal is a class or category to which a particular specific case or example belongs.

Perception is properly correlating specific examples, with their universals. It includes correlating examples not previously experienced.

A thought is a specific example of a set of a number of universals that are somehow related.

Thinking is a succession of thoughts differing in one or more of the successive set's universals.

2 - UNIVERSALS

Humans perceive a large number of universals. The universal is the common characteristic of all elements of a set: E-ness, moving-ness, shirt-ness in the following three examples:

- the letter E, whether capital or lower case, hand written or mechanically produced, large or small, alone or among other symbols, etc.;
- all moving regardless of speed or direction;
- all shirts regardless of details.

Universals recognition is not always accurate even with a competent recognizer. Everyone is familiar with the problem of another person's handwriting.

Perception involves an input, a data processor, and an output. The processor is a device that operates on the input data to correlate examples with universals. The output is data representing the correlation.

A single input may be a sample of a number of different universals. For example a particular letter E might belong to all of the universals: E, upper case, small, hand written, in ink, red, moving left to right, upside down, getting smaller, etc.

For the present the input is data from an idealized eye. The image appears on the eye's retina, an array of sensors (the "rods and cones"), each registering in an "on"-"off" manner. The image projected onto it is an array of "on" and "off" sensors.

Initially a small array of 16 sensors in a square will be used. To refer to the individual sensors the system of Figure 1-b will be used.

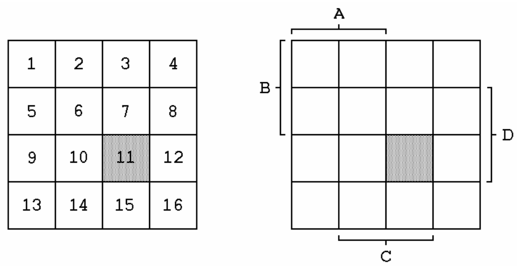

Figure 1

(This procedure is the digitizing of the image into binary elements and the description of the sensors and their states by Boolean Algebra. For those not familiar we continue in simplified terminology.)

The half of the array of Figure 1 that is labeled *A* will be called "A". The other half will be called "not A" and be written $\underline{A}$. We can then refer to element #11 of Figure 1, for example, as being in

(1) $\underline{A}$ and $\underline{B}$ and C and D.

The letters *A*, *B*, etc., are variables. The only allowed values are *1* and *0* ("on" and "off").

Instead of writing "*and*" as in equation *(1)* the notation $\underline{AB}CD$ will be used or, for clarity, $\underline{A} \cdot \underline{B} \cdot C \cdot D$.

To refer to more than one element of the array at a time "*or*" will be used, written as "*+*". To refer to elements #10 and #11 of Figure 1a the reference is

(2) $A\underline{B}CD + \underline{AB}CD = \underline{B}CD$.

Let us consider recognizing a cross, a horizontal line crossing a vertical line. We seek an output of "*1*" for any, and only, such crosses.

First consider some examples as in Figure 2.

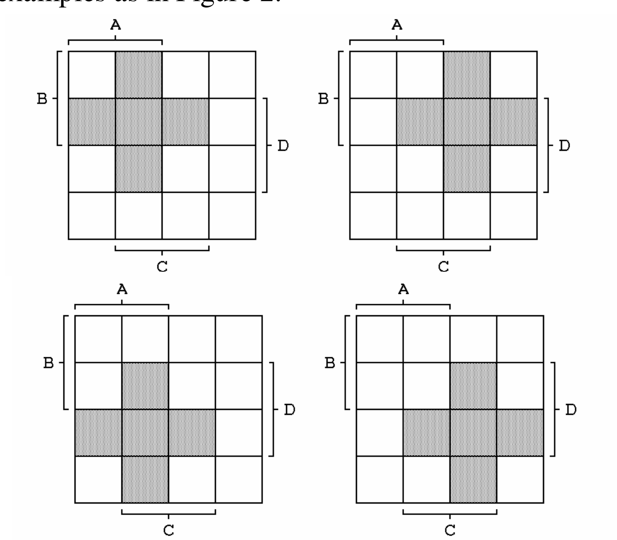

Figure 2

The elements making up each example are:

(3) (a) $ABC\underline{D}+AB\underline{C}D+ABCD+\underline{A}BCD+A\underline{B}CD$

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

(c) $ABCD+A\underline{BC}D+A\underline{B}CD+\underline{AB}CD+A\underline{B}C\underline{D}$

(d) $\underline{A}BCD+A\underline{B}CD+\underline{AB}CD+\underline{ABC}D+\underline{AB}C\underline{D}$.

Using the commutative and associative principals and factoring reduces equation *(3)(a)* to any one of

(4) AB[C + D] + CD[A + B] or

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

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

and similarly for equations *(3) (b, c,* and *d)*

This cross is the "*and*" of any two of the four variables with the "*or*" of the other two, that then "*or*"-ed with the "*and*" of the other two with the "*or*" of the first two, each variable "*not*"-ed or not consistently. That is the logical formulation of the universal, cross-ness from the examples. That is its perception.

So far four crosses have been treated out of *50* possible. All would be identified by the form just derived. The total number of different patterns that can be displayed on this sixteen element array is $2^{16} = 65,536$. The formulation developed perceives the *50* cases having the universal cross-ness in common out of the *65,536* total cases.

The purpose of this example was to illustrate how:

- a sensory input can be analyzed;
- a limited number of samples can be sufficient to establish a formulation for a universal;
- that can be done by digitizing the input into binary representation and Boolean treatment;
- the form of the universal can correspond to contemporary digital logic systems.

3 - Complex Perception Systems

Instead of 16 sensory elements the human eye has about 7,000,000, the retinal rods and cones. The number of patterns in a binary digital array is

(5) Number of possible patterns

$$= 2^{[\text{number of elements in the array}]}$$

The eye can deal with about $2^{7,000,000}$ patterns.

Taking *1,000* as an approximation to $2^{10} = 1,024$

(6) $2^{7,000,000} = (2^{10})^{700,000} \approx (10^3)^{700,000}$

= 1 followed by 2,100,000 zeros.

If the eye saw *10 patterns per second* it would take *30 years* to see *10,000,000,000* patterns (*1* followed by *10 zeros*, not *2,100,000 zeros*) and inconceivable eons to see all of the patterns possible.

And, relationships among patterns are significant, providing information on 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 individually. The available information for the eye is immense.

When an image appears on the retina, a family of signals from the individual sensors goes to the nervous system. The first level of processing identifies basic universals in the input image such as: corners, edges, shape types, motion, etc., universals like the cross of the recent example.

The possible number of such universals is quite large, enough to constitute a complete description of the input, consisting of all its universals with their location and orientation in the image. The input is converted from an array of points in one-to-one correspondence with the original to an array of characteristics of the image.

Referring to the *7,000,000* retina sensors as #*A*, #*B*, ... for all *7,000,000* of them, an image on that retina is the Boolean "*and*" of the *on* sensors "*and*"-ed with the "*and*" of the "*not*" of the *off* sensors.

(7) Sample Image = $AB\underline{C}D\underline{EF}...$ [7,000,000 letters]

A group of images, each in the equation *(7)* form, is the "*or*"-ing together of the expression for each.

(8) Sample Group of Images =

$$= AB\underline{C}D\underline{E}\underline{F}... + \underline{A}BC\underline{D}EF... + A\underline{B}\underline{C}DEF... + ...$$

Such an expression is the universal of that group of images; any image of the group matches part of the expression and any not a member fails to do so.

This kind of expression can be implemented electronically with logic gates that produce the "*and*"-ing and "*or*"-ing, and flip-flops that carry the variables *(A, B*, etc.) and their "*not*"-ing. But, the expressions are cumbersome and require too many logic gates and flip-flops.

More serious, this procedure is only correct for images used in its original set up. It is unable to generalize, learn, and correctly treat images never before experienced.

Referring to equation *(8)*, if every input exhibiting the universal of interest has sensor #*B* = "*on*" and every one that does not exhibit the universal has #*B* = "*off*", both regardless of the other sensors, then sensor #*B*, alone, represents the universal.

The nature of universals is similar simplified expression, generalization, a neglect of specifics in favor of broad commonality. Its expression is simpler than as of equations *(7)* and *(8)*.

But how does a system extract a simplified universal from a group of sample inputs?

4 - Neural-Type Logic Devices

Neural-type logic differs from Boolean logic in two respects: its operator is the "majority" operation and it uses constants in addition to variables. Unlike the Boolean values of 1 or 0, majority variables and constants values are +1, 0, or -1. We use the notation M(1st variable, 2nd variable, ...), where the output is the majority of the variables else zero.

Majority logic can implement Boolean logic. For example (using Boolean values):

(9)

$$M(A,B,C) = AB + AC + BC$$

$$M(A,B,1) = A \cdot B + A \cdot 1 + B \cdot 1 = A + B$$

$$M(A,B,0) = A \cdot B + A \cdot 0 + B \cdot 0 = A \cdot B$$

$$M(\underline{A},\underline{B},C) = \underline{A}\underline{B} + \underline{B}C + \underline{A}C$$

In a neural system, the inputs to a neuron are the outputs of other neurons or sensors. Those inputs are such that some act on the neuron in an excitory fashion (*natural variables*) and some act on it in an inhibitory fashion (*"not"-ed variables*).

There is an internal threshold level in the neuron. If the threshold is *zero* then the logical construct of the neuron is simply the majority of its inputs.

But, if the threshold is *greater than zero*, then for the neuron to have an output of *+1* the number of excitory inputs (*+1's*) must be the threshold amount greater than the number of inhibitory inputs (*-1's*). The effect is as of as many constants equal to *-1* as the level of the threshold. Likewise, a threshold *less than zero* corresponds to there being as many constants equal to *+1* as the level of the threshold.

The neuron "remembers" the value of its threshold, the net value of constants in the logical construct that the neuron effects. The special power of the neuron is that its threshold can be changed, changing its constants, and changing its Boolean logical construct. That adjustment of neural thresholds to achieve a desired result is **learning**.

A network in which the outputs of some neurons are inputs to other neurons creates complex majority logic structures. Networks of neurons use the neuron's majority logic, modified by the individual neuron's thresholds, to represent complex Boolean logical descriptions which are the simplified logical representation of specific universals. Complex neural networks are universals representation systems.

To see how this can be done we now operate a simple neural network with as its input source the earlier `16 element` array. The output of each of its `16 sensors` is interconnected as inputs to a number of neurons and then the outputs of those neurons connected to the inputs of one final neuron. The output of that final, single, neuron is the act of this entire neural network.

But, how should the interconnections be made? Which sensors should be connected to which inputs, excitory or inhibitory, of which neurons? Let us assume, for here, that that has been correctly solved.

We now teach the network to recognize the universal cross-ness, to output `+1` if the input image has `cross-ness` and nothing otherwise.

(A) Project an example image onto the array.

(B) Note whether it has `cross-ness` or not.

(C) Note the network output (is it `+1` or null ?).

(D) Evaluate the network performance:

Input Image	Output	Result
cross	+1	Correct
cross	not	Wrong
not cross	+1	Wrong
not cross	not	Correct

(E) Change each neuron's threshold as follows.

Neuron Output	Result Was	Change Threshold	Makes in Future
+1	Correct	down	likely
not	Correct	up	likely
+1	Wrong	up	unlikely
not	Wrong	down	unlikely

This process inevitably changes the network so that it is increasingly likely to correctly identify crosses versus non-crosses. The network's learning is accomplished by logical adjustments to each neuron's threshold level. Such adjustments change the Boolean logical construct that is effected by the neuron's majority operation in conjunction with the constants represented by its threshold. The Boolean logical operation that the neural network performs on the input gradually changes, approaching the Boolean logical construct of the universal being learned.

5 - Aspects of This System

This simple perception mechanism could not be a prototype for complex rational systems. It requires a **teacher** external to the system, to direct the learning. It is the teacher who decides the threshold changes after each input image is processed. A system without an external teacher is needed.

That is not a difficult problem. We learn things by repetition. Explanation and experience are means to learning, but learning only happens when the lesson is repeated correctly, sufficiently.

Because:

- In real world nervous systems there is no external teacher adjusting thresholds,
- learning does occur in such neural systems,
- learning can only take place by means of threshold changes, and
- things are learned by repetition and emphasis;

there being no alternative it must be that the threshold changes are automatic, that they occur by a simple, inherent process within the neurons.

There being no alternative it must be that when a neuron "fires", that is delivers a “+1” output, its threshold naturally decreases slightly so that a similar "firing" will be more likely at the next similar input. And, when a neuron "does not fire", that is it delivers no output its threshold naturally increases slightly so that a similar "non-firing" will be more likely to occur at the next similar input.

That is learning by repetition. The repeating of success is the repeating of similar inputs and obtaining the same output. That would tend to change the thresholds of the involved neurons in the direction that makes the same outcome more likely.

If we lapse in regular practice or rehearsal of something learned we begin to forget it, to lose the skill, to find it somewhat harder to remember. That corresponds to the gradual decay of thresholds if they are not regularly reinforced by repetition.

The neuron is an electrochemical device. Its operation is the propagation of electrical potentials generated and transmitted within it by chemical actions. The "firing" of a neuron, delivers an amount of electrochemical energy, temporarily depleting the neuron and its electrochemical threshold. Likewise, the absence of "firings" enables the neuron’s on-going metabolism, to accumulate more threshold.

Another issue is the system’s operation’s **synchronization**. It has been implicitly assumed so far that all of the data are available and operated on at each neuron simultaneously.

Such simultaneity is unlikely in a biological neural system. The paths over various different dendrites and axons are of different lengths and the electrochemical signals’ travel times among the various neurons must be different. Also it is too risky for nature to rely on exact synchronized timing.

For this problem digital computers use a synchronizing clock. The clock generates a train of pulses. The input to the *flip-flops* is “*and*”-ed with the *pulses*. Regardless of between pulses, only the conditions at the time of each pulse participate in the next step in the computer’s operation.

For a biological neural system the clock pulses would have to be “*and*”-ed with the neurons’ inputs. The brain is not constructed that way. Neurons connect densely to other neurons and sensors that are physically near to them, less densely to those that are somewhat distant, and rarely or not at all to those that are very distant (excepting sensor neurons that carry signals from distant sensors to the brain and motor neurons that carry signals from the brain to muscles).

How is a biological logic system synchronized? It is not. Much of the time a neuron puts out no signal. Then its firing sends out a *+1* signal. That produces an excitory or inhibitory input in other neurons depending on the connection to them so that they experience in consequence a *+1* or *-1* input.

Instead of synchronized simultaneity the neurons operate on **accumulation**. Immediately after its last firing, which somewhat depletes its electrochemical energy, the neuron accumulates new electrochemical energy from its accumulation of `+1` inputs less the accumulation of `-1` inputs, all in random timing. When that accumulation exceeds its threshold the neuron fires again. That is essentially the complete operation of individual neurons.

The neurons' accumulation mode of operation facilitates another characteristic of living neural systems. Although they are essentially binary, transmitting `+1` pulses that act at inputs excitorially or inhibitorially those pulses can convey information about "how much". Whether a sensor detects touch or temperature at a point on the body or sound in the ear, or light in the eye, the information conveyed to the neural system by sensor outputs is that of both "what" and "how much".

The "what" depends on the type sensor. The "how much" is communicated by the rate of the sensor's neuron firings of `+1` pulses. More frequent pulses are caused by, and therefore signify, greater stimulation of the sensor, e.g.: brighter light, louder sound, or more harsh touch. The greater the rate at which the sensor receives excitation inputs the greater is the rate at which it is able to repeatedly fire.

The combination of neural majority logic and the learning by variable thresholds naturally leads toward the objective: representing the related universal. With synchronization replaced with accumulation the system also responds to "amount of signal". And, it deals with time as well as structure.

The operation of neural systems depends on two aspects: the **interconnections between neurons** and the thresholds. The interconnections are fixed; but, how should they be for optimum performance?

- The output of which neurons should connect to which other neurons ?
- How many ?
- Excitorily or inhibitorially ?

Initially, we have no idea how to interconnect the neurons. Nature had no idea, initially, either. That, is the key. The best choice is randomness.

The interconnections determine the specific Boolean logic implemented by the system. But, we have no idea what the specific logic should be for a particular universal. In any case, we need a system that can deal with any universals, with any input.

A neural system must deal with a great variety of inputs. It is not possible to design in advance for all of the possibilities. Given that, the only way for a neural logic system to maximize its ability to deal with the unknown is to use random interconnections. What is needed is to avoid any pattern that might bias the system in favor of some Boolean logic over others. Random is patternless.

That is also the easiest and most natural system for nature to implement. It requires no plan nor control. It calls for merely allowing what happens to happen.

Then Darwinian variation and natural selection enter. Some "random" interconnection systems turn out to perform better than others. The process tends over time to select optimal interconnection systems.

But "optimal" depends on the situation. Vision systems have existed in nature for hundreds of millions of years. There has been sufficient time and experience for optimal sets of retinal universal processing interconnections to develop.

But, what is optimal brain operation for astronauts, steelworkers, gourmet chefs? Man experiencing so many different geographies, weathers, food supplies, dangers, and so on confronts a thinking need that cannot be predetermined. While his vision system may be well defined, the needs of his thinking system are very broad. The most likely success is one that can adapt to any circumstances.

Thus, at the higher levels of neural systems randomness is still the optimum design even though specific sub-systems, vision, digestion, breathing, can be optimized in special ways.

For designing an artificial intelligence, random interconnections should be used (but employing our brain's greater density of interconnection to near neurons and less to distant ones). That is apparently the case in our cerebral cortex, the "thinking part" of our brains as compared to the retina, the "seeing part", (but lacking understanding of what is seen).

Each neuron receives inputs from tens, hundreds, and in many cases thousands of other neurons. Each neuron's output is input to a number of other neurons. Each universal involves a large number of neurons.

Our brains can express over 10^8 times as many different universals as they have neurons. The number of universals to be perceived is so immense that it is not practical for individual neurons to be dedicated to participating in just one universal. Far too many neurons and far too large a network or brain would be needed.

Each neuron's output inputs to **multiple universals** and their particular Boolean logic expressions, all simultaneously. Each universal is a different logic expression but its participating neurons participate in many other universals. While the same set of neurons might appear in a number of universals (but interconnected differently) most universal's neuron set is at least a little different from the others.

The involvement of individual neurons in a number of universals simultaneously is necessary because otherwise too many neurons would be needed. It is unavoidable given the complex interconnections. And, it has a drawback and a tremendous advantage.

The drawback is that the threshold of an individual neuron is changed by actions involving any of the universals in which it participates. So to speak, having its threshold well trained for cross-ness, it then must learn circle-ness and in the process its threshold is further changed. That most likely degrades its ability to identify crosses. Over a period of inputs randomly varying between circles and crosses the threshold would become a compromise.

We experience this effect on a large scale. Having learned some thing fairly well and then turning to other learning we find that our learning of the former thing degrades somewhat.

But, the tremendous advantage of individual neurons participating in a large number of different universals is that it creates the capability for thinking.

6 - Concepts, Thoughts, Thinking

So far we have dealt with inputs from the external world. Such an input is a specific example and has characteristics peculiar in combination to it, alone. Each of those characteristics is a sample of a universal in which the example participates. The distinction between an input example and a universal is important.

The universal does not exist external to the neural system in the sense that it has no representation there. (It is, of course, the commonality among specific examples with regard to the characteristic that the universal represents.) It exists in the neural system as a configuration of neurons and their thresholds that can discriminate between the presence or absence of that universal in an input example.

The input example exists in the world external to the neural system. It does not exist within the neural system except as a brief representation in terms of the universals in which it participates.

The universal is relatively permanent in the neural system being the "wiring" configuration of interconnections among the neurons plus their threshold settings, which do not change rapidly. In the neural system the input example is a momentary set of appropriate neural firings.

A universal exists whenever a set of examples have something in common. The universal is expressed only by a mechanism, neural logic, able to abstract the universal from a group of examples.

There are material universals, ones dealing with physical objects, and mental universals, ones dealing only with ideas in a brain or neural network.

UNIVERSALS			
Material		Mental	
Blue	Food	Good	Trouble
Round	Heavy	Doing	Anxious
Visually depicted letter "e" – a spatial form		Mentally conceived letter "e" – its language role	

A **concept** is a set of universals that are related (e.g.: material – food, mental – busy). A **thought** is a specific example of a concept (e.g.: material – bread, mental – baking). A concept is the neural logic structure of its universals. A thought is the firing, in that logic structure, of a subset of its neurons for an example that satisfies the universal. Their firing is the thought.

A **thought** can also be of a specific concept itself or a single universal itself (even as in reading this our thoughts are about universals).

Thinking is sequences of thoughts, that is sequences of specific examples of certain sets of universals. Each (momentary) thought is a particular set of (momentary) neuron firings. Thinking is sequences of such firings of particular sets of neurons, the content of the set changing somewhat from firing to firing.

These, form a succession, comprise a trend of thinking, by having in common parts of their universals' logic. For example, greatly simplified, suppose that *thought #1* consists of *universals [a, b, d, f, g]* out of the *26 [a ... z] total universals* in this simple example. The *next thought* consists of the prior *[a, b, d, f, g] plus [k]*. The *third thought* consists of the set comprising the *second thought less [d]*. The three such thoughts in that sequence are "a line of thought", thoughts sharing gradually shifting parts of their Boolean logical expression, thinking.

A thought is the firing of a particular set of neurons. Those neurons as a set represent that thought. But individually, various subsets of that set also represent parts of a number of other somewhat related thoughts. At the moment of the current thought those other thoughts are not active because the current neuron firings are not exactly the required set of those other thoughts.

However, the activation of the current thought could, with only a little help, lead to the activation of one or more of those other thoughts that share a significant portion of their logic with the current thought. That "little help" would be something that has the effect of activating some other related subsets of neurons and / or deactivating some of the currently active ones. And, because of the sharing of neurons, of universals, in common between the two thoughts, the successive thoughts will be related; they will tend to follow properly in terms of thinking.

Participating in multiple universals neurons cannot participate in contradictory universals because a compromise threshold could not form. In general the set of universals in which a neuron participates must be a somewhat related family.

Because the various subsets of the logic of the current thought participate in numerous other potential thoughts, thinking can proceed in numerous possible directions from the current thought.

With the extensive interconnection of neurons, and the recirculation of output firings as inputs elsewhere in the network, and with constant sensor input , items of "little help" to progress to a next thought are constantly present. The existence of a current thought inevitably results in an immediately following next thought *ad infinitum*. It's direction depends on which "help" is stronger, dominant.

The transitions from thought to thought are changes of some of the individual universals that comprise the current thought. In the complex neural thinking structure with thoughts involving inconceivably large numbers of universals the opportunities for a variety of such associations are quite large.

This process is thinking but not yet purposive thinking. It takes place in neural systems over a wide range of size and complexity. Certainly man thinks. But, thinking is also done by dogs, birds, snakes and beetles. In each case as the size and sophistication of the neural system is less the complexity of thinking is reduced. But it does take place.

At the same time each thought modifies the then existing universals. Each neuron's firing reinforces its threshold and each failure to fire de-inforces its threshold. Also, since the universal is an abstraction of a common element from a family of samples, then if the thought is a new sample added to the family, the thought must produce some change in the existing universals.

The result is an iterative process of evolving universals and sequences of specific examples (thoughts) where the examples modify the universals and the universals determine what are the various directions that the sequence may take.

- Most thinking operates with material universals and existing concepts and produces changed concepts.
- Abstract thinking operates purely with mental concepts and produces new and changed concepts.

7 - Purposive Behavior – Consciousness

Purposive behavior involves:

- setting a goal;
- choosing among options to achieve the goal;
- comparing progress against the goal; and
- adjusting behavior by modifying choices.

That is consciousness.

8 - Motivation

How do we "pay attention"; how do our goals come about; what causes our neural system to seek to satisfy our goals not ignore them? A form of pain is the motivation.

A sensory neuron's rate of repeated firings tells "how much": how loud a sound is, how bright a light is, the magnitude of a touch. But if the input is too large it can signal danger. The simple neural networks in early organisms received "how much" data from their sensors. If it was a signal of "too much" the organism might not survive. Some organisms responded to the "too much" sensory inputs by action to avoid the cause.

That could have been common because in simple neural systems the sensors were closely linked to the motor neurons. The most simple such early neural system consisted of a sensor neuron connected directly to a motor one. Such a mechanism could, for example, produce: withdrawal of an exposed body part or waving of flagella (causing swimming away).

That kind of response would have significantly increased the survival rate of the organisms behaving in that manner. They would likely have become the only type organisms surviving into their future and contributing characteristics to their evolved successors. Avoidance of danger or harm must have become an operating principal of simple, early neural networks very early in their existence.

In only slightly more sophisticated neural systems the same response would develop to "too little" sensor input. The Boolean "`not`" operation is an essential of the logic of neural nets and neurons have both excitatory (normal) and inhibitory (not) inputs. The "`not`" of a "too little" input would be a "too much" signal. Some cases of "too little" can be as dangerous as those of too much. We humans react strongly to too little good air to breath.

We retain those early-developed mechanisms. If one puts his finger on an oven at room temperature he can keep it there all day. But if the oven is burning hot, then the moment the finger touches it the finger is quickly withdrawn automatically without awareness of thinking about it. The "too much" neural signals from the finger's sensory neurons directly trigger the arm motion motor neurons by interconnection in the spinal column without the brain as neural logic intermediary.

We also exhibit a moderately more sophisticated neural response to too much. The eye automatically shuts quickly and the head averts when an object is moving rapidly toward it. Our brain is not consciously involved. The reaction occurs before we are aware of the problem. That is not a direct sensor-motor action. Significant neural network processing is needed to convert the raw visual picture into information that, "an object is moving rapidly on a trajectory that endangers the eye" and to generate averting the head.

Eyes developed long after the "early, simple neural networks". But, the long established behavior of those early networks, treating excess as danger and automatically acting to correct the situation, appears developed into a greatly more sophisticated version in the eye's response to a detected danger. A complex set of universals represented by a significant number of neurons as a set producing "too much" signals collectively is involved in that action.

With evolution as neural networks became larger and more complex the operation of their "too much" response became more sophisticated, that is:

- It involved greater numbers of neurons and in more numerous universals which described more complex thoughts.
- It developed the ability to deal with multiple "too much" signals at the same time, and to organize multiple responses, and to relate and prioritize them.
- It included logic to determine whether a response is really needed, to consider alternative responses, and to put together patterns of responses.

Such behavior is the setting of goals and the making of choices among alternative courses of action. It is purposive behavior.

The "too much" signal and the reaction that it triggers ranges from the very simple sensor-motor type cases (the hot oven) through the significant neural processing type cases (the eye shutting and head averting) to increasingly sophisticated motivations and resulting actions. Just as our thoughts are the patterns of which neurons are firing at a particular moment, so our conscious purposive behavior, our performance in life at home, on the job, as parents, in love, and so forth, is our responses to highly sophisticated and complex sets of neural "too much" (and "`not`-ed" "too little") signals.

Those signals involve, are related to, are the equivalent of, <u>are</u> pain and pleasure (pleasure being not-pain). When the signals involve material sensor input the consequent responses normally involve physical action, that is material response to material sensor input. When the signals involve non-sensor input, that is abstract thoughts, the consequent responses normally involve non-material actions, involving intentions and desires.

Early in the evolution of neural systems they evolved to treat extremes of neuron firing rates, low or high, as being: bad, a sign of danger, something to be avoided – triggers of corrective action. At a sophisticated level we now refer to the effect of excessively non-moderate neural firing rates as meaning that the related material objects (described by the universals the neurons of which are so immoderately firing) are painful or unpleasant.

At a high sophisticated level we now refer to the effect of excessively non-moderate neural firing rates as meaning that the related mental objects (described by the universals the neurons of which are so immoderately firing) are unintended or undesired.

The opposite, neuron firing rates that are neither too great nor too small, that is moderate rates, then signify: comfortable, intended, desired.

It could be said that we spend our lives seeking to have our neurons firing at a rate well between the "too much" of a too rapid rate and the "too much" of a too slow rate. One could say that a state of moderate neuron firing rates is what we call happiness, pleasure, contentment, joy.

Or, perhaps, the greatest pleasure or joy, the best sensation, corresponds to neural firing rates that are as near to "too much" as possible without being so strong as to mandate corrective action. Our human experience would tend to indicate that we behave that way, that we crave excitement so long as it does not go over the boundary into the damaging.

Or, perhaps, for different kinds of good or pleasure different neural firing rates apply – contentment corresponding to a moderate rate, great joy to a rate near "too much". Perhaps, the neural network involves a mix of different correspondences between neural firing rates and various subjective feelings of good for the variety of different such feelings. And, perhaps, that mix and the associated firing rates change throughout the individual's life as the neural system has more and more living experience, more learning and adjustment of its thresholds, as it evolves with the person's mental and emotional growth. And perhaps the precise state of the system is a little different for each individual -- each having a unique set of responses.

9 - Response to Motivation

Having treated the input aspect of the evolved "too much" signal, of equal importance is the neural system's action when such signals appear. In early neural systems that was a motor action tending to relieve the cause of the "too much" signal.

A highly evolved system's response to "too much" signals, must also tend to relieve the cause of the "danger" or "pain" signals. No matter how highly evolved the neural system, its response to inputs signifying pain, bad, unintended, or undesired must be a response that tends to, or is intended to, relieve the situation, to remove or reduce the cause of the "too much" signal.

Thus we respond to desire for a cookie by exciting our motor neurons to cause our walking to the cupboard, selecting a cookie, and eating. Even more, in general responses to "too much" signals produce our performance of the routine of living: arising, eating, performing the day's tasks, etc.

But, sophisticated systems are really very complex. They can learn and act not only on that it is bad or painful to fail to experience for example:

- a luscious desert;

- or buying the new sports car one has wanted.

But, even more, they can mandate, for example:

- revenge to relieve the pain of an affront or a loss,

- or declining revenge to relieve the pain that it is contrary to one's standard of moral character.

That can be done by a system as complex as follows.

Our brain has about 100,000,000,000 neurons. Assigning 10% to sensor and motor activities and providing for 100 subsystems, then any one sub-system would have 10^8 neurons and could represent $2^{100,000,000}$ different patterns.

A page can hold about 3,000 zeros. It takes 10,000 pages of zeros just to write out $2^{100,000,000}$ – to write it down not to express its value. (Writing "1000" takes 4 digits, but its value is 1,000.)

10 - MEMORY

A memory, something remembered, is a sequence of thoughts, of neural firings repeating a prior thought sequence. To the extent that the memory is mentally repeated (the thinking through the sequence of thoughts again) to that extent its thresholds become more firmly set; the memory becomes more permanent. To the extent that the memory is not repeated, to that extent other thoughts that use some of the same logic as is used in that particular memory, produce threshold changes that degrade that particular memorization.

Access to the memory, the remembering of it, is via the same kind of associations of thought universals as in any thinking. To access the memory we must think of something associated with it, something that will trigger the sequence of thoughts that are the memory.

Memories reside in a diffused, distributed manner over a large number of neurons. They are not in some separate "library" or "memory file cabinet" of the brain. They are "right out there" intermixed with and inter-operating with the brain's overall activity. The only difference between a remembering and a thinking is whether the pattern of thoughts is new or is the retracing of an earlier one. However, at retrace some of the original thresholds have become changed somewhat blurring the memory in comparison with the original.

What with the vast amount of information that we remember and the complexity of our thinking, one wonders how our neural system can contain it all. Of even more concern could be that, with thresholds being constantly affected by current mental activity how can things learned and things remembered last a long time?

Yes, a certain amount of memory loss occurs because of disuse of some memories or learnings and the consequent blurring of their thresholds. And yes, a brain cell dies here and there regularly and takes its participation in the logic with it when it goes. But those degradations are negligible in the overall system. The number of neurons involved in any single thought or memory is so large that a problem with, or a failure of, a single neuron here or there, now and then, is of no importance.

A saying valid in this context is that "we are what we eat". It is likewise true that our mind (which, after all, is our conscious selves) is what we think. We tend to become, to think as, to behave as, that which we feed into our neural logic networks and threshold settings. That is something to think about.

11 - INSTINCT

What can instinct be except pre-learned behavior somehow embedded in the nervous system? The nervous system consists essentially only of neurons. If instinct is to be embedded in it, the only available way is in the thresholds of various neurons and in the interconnections among them. The passing of instinctive behaviors from generation to generation can only be the passing on of specific neural interconnections and pre-formed thresholds.

It may well be that an aspect of evolution, of variation among individuals of a specie, variation that gradually or suddenly leads to a different type specie, is that of some of the new individual's neurons having different interconnections and threshold settings than the parent had. The new individual would therefore have different pre-learned neural mechanisms, different instincts.