J. WUNDERLICH PhD

MACHINE INTELLIGENCE

APPENDIX A) Historical Perspective

FROM J. WUNDERLICH MASTERS THESIS
"DESIGN OF A NEURAL NETWORK MICROPROCESSOR"
1991/92

PART OF FIRST CHAPTER
OF J. WUNDERLICH BOOK
ON ROBOTICS & MACHINE
INTELLIGENCE

Plus attached 2003 study of the biological
structure and function of the human brain

This appendix includes many historical advances in neurocomputing. The back-propagation neural network model was selected after researching and comparing the many models and concepts summarized here.

The selection of a neural network model for implementation should not be solely a function of what is most recently popular. It should be done by analyzing the historical developments in neurocomputing and brain research. The most appropriate model can then be selected for implementation as a digital microprocessor chip. This approach not only aids in the selection of an already developed model, but also provides a foundation of knowledge for the possible development of a new neural network model.

The following list by no means includes all of the significant advances in neurocomputing; it does, however, contain many models which have had a significant impact on the evolution of artificial neural networks. This list also includes some advances in the neurosciences which give insight into how the brain may actually work:

1890 William James

(1)

He formulated a general elementary principal of association:

> "When two brain processes are active together or in immediate succession, one of them, on reoccurring, tends to propagate its excitement into the other."

(2)

James also noted the general tendency of the brain to be inherently application specific:

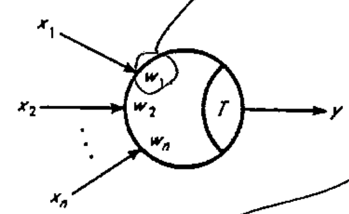
> "The brain is not constructed to think abstractly...it is constructed to ensure survival in the world. It has many of the characteristics of a good engineering solution applied to mental operation: do as good a job as you can, cheaply, and with what you can obtain easily. If this means using ad hoc solutions, with less generality than one might like, so be it." [1]

ϕ OR 1

1943 Warren S. McCulloch and Walter Pitts

They modeled a neuron as a binary device (threshold logic unit) which is on only if the weighted sum of the input signals to it exceeds a given threshold. This is threshold logic. (See Fig.5) [2]

WEIGHTS (MULTIPLIERS)



Y=1 if and only if $\sum w_i x_i \geq T$
Y=0 if and only if $\sum w_i x_i < T$

Figure 5. The Threshold Logic Unit (Figure from [2])

1945 John von Neumann "PRINCETON" VS. "HARVARD" ARCH.

He formulated idea of digital computer which has program stored along with data in computers memory. [1]

1949 Donald O. Hebb

He presented first explicit statement of the physiological learning rule for synaptic modification:

"When an axon of a cell A is near enough to excite a cell B, and repeatedly or persistently takes part in firing it, some growth process or metabolic change takes place in one or both cells such that A's efficiency, as one of the cells firing B, is increased." [1]

AB CONNECTION STRENGTH MEMORY

Note the similarity to the elementary princple of association of William James (1890).

1950 K.S. Lashley

He observed that rats demonstrated learning behavior even though large areas of their cerebral cortex was removed. This demonstrated the "robustness" of the brain. It also showed that the representation of information in the brain is distributed rather than localized. [1]

1954 B. Farley and W.A. Clark

They developed first computer software simulation of nervous system. [1]

1958 O.G. Selfridge

He developed the "pandemonium model: a paradigm for learning" in which mulitiple independent systems simultaneously look at the input and respond according to their own bias. Selfridge also used the mathematical concept of "hill-climbing" to modify the connection strengths (weights) between the units of his model. This consisted of modifying the weights by a small amount in all directions and choosing the direction that gives the largest increase in effectiveness. The process stops at the peak of the first hill it encounters. This is a local maximum. [1]

MORE ON THIS WHEN WE DERIVE "BACKPROPAGATION" LEARNING

1958, 1960 F. Rosenblatt

He developed the PERCEPTRON; the first precisely specified, computationally-oriented neural network; a learning machine potentially capable of complex adaptive behavior. [1] In 1958 Rosenblatt developed the basic structure of the PERCEPTRON and in 1960 he developed the perceptron learning rule. Fig. 6 shows the operation of the PERCEPTRON. The input layer consists of threshold logic units of McCulloch and Pitts (1943). The adaptive threshold element collects a weighted sum from the input layer and applies the hard-limiting signum function to them. The inputs can be either 0 or 1. The output is +1 if the weighted sum exceeds a fixed threshold, else the output is -1. The perceptron learning rule adjusts the weights by the following rule [3]:

$$\overline{W(k+1)} = \overline{Wk} + ETA(ek/2)\overline{Xk} \qquad \text{(EQ. \#1)}$$

Where: $\overline{Wk}$ = the present weight vector; each element is a connection weight

$\overline{W(k+1)}$ = the next weight vector

SCALERS

ek = (dk - yk); this is the error signal.

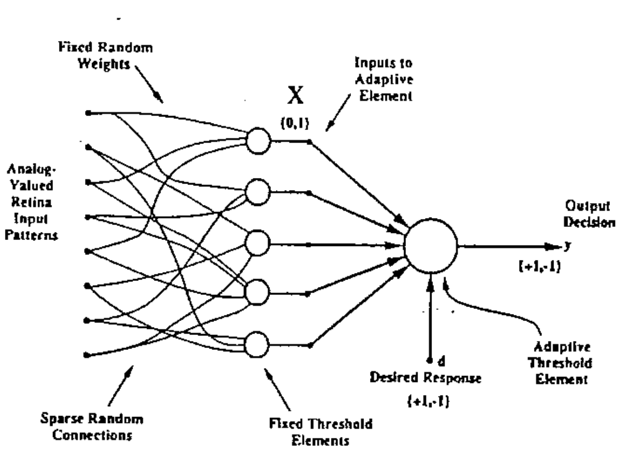
dk = desired response (i.e., supervised learning)

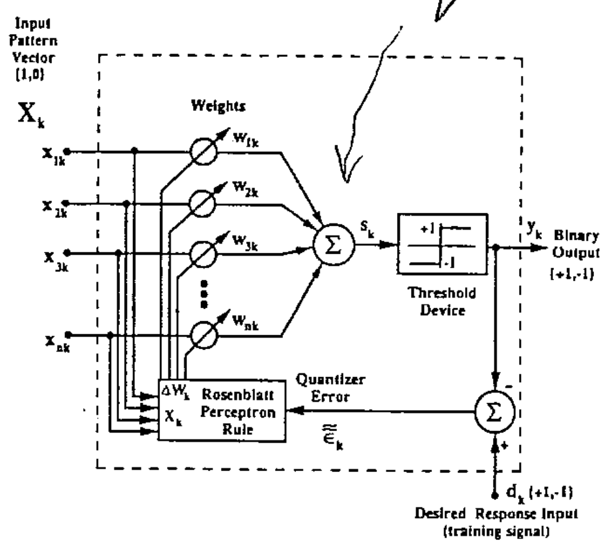
yk = binary output

$\overline{Xk}$ = input vector ← BINARY INPUTS!

ETA = the learning rate (Rosenblatt normally set this to 1)

Σwx





The adaptive threshold element of the Perceptron.

Figure 6. The PERCEPTRON (Figure from Widrow [3])

They developed the ADALINE (Adaptive linear element) which has a structure similar to the adaptive threshold of the PERCEPTRON of Rosenblatt (1958) except that the input signals are not limited to 0 and 1; any analog value can be input. Also, the ADALINE threshold is variable:

$$TH = w0 * (\text{Bias}) \qquad \text{(EQ. \#8)}$$

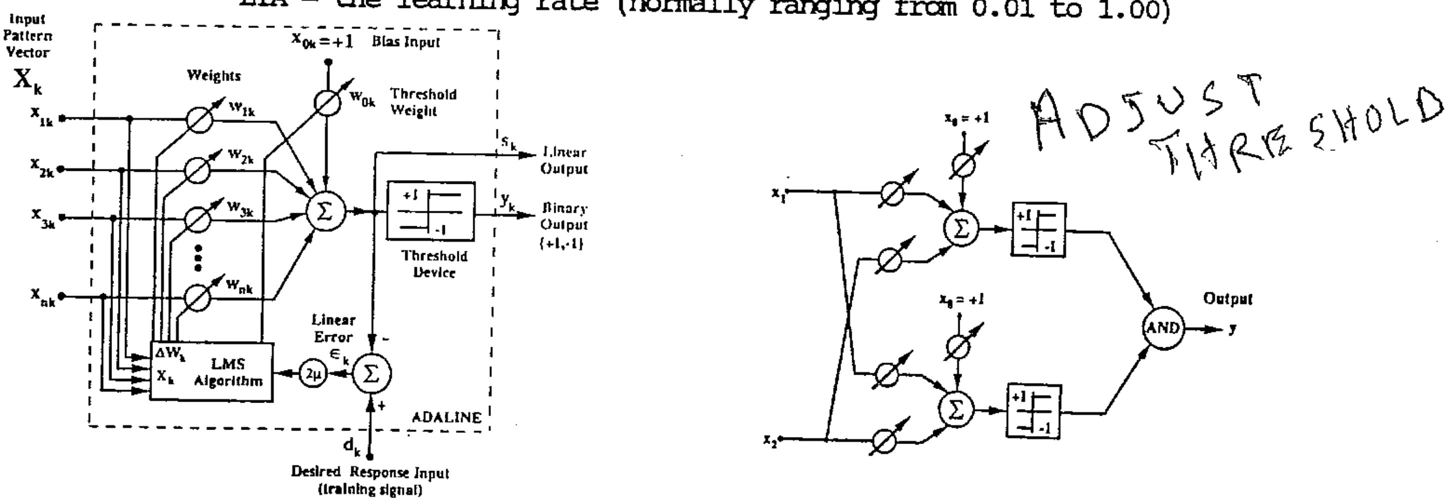
WE WILL EXPLORE THIS LATER

This is a weighted sum from a bias connection. Widrow and Hoff used an error correcting algorithm for adjusting weights known as the least mean square (LMS) algorithm which minimizes the sum of squares of the linear errors over the training set. The error signal used by the algorithm is the difference between the desired output of the network and the linear output of an ADALINE before the threshold is applied. (See Fig. 7). These errors create an "error surface in the weight space." This error surface represents a topography in which the valleys contain minimum errors and the lowest "elevation" is the global minimum representing the optimal answer. The error surface is searched by using an instantaneous gradient to follow the path of steepest decent and minimize the mean square error. This is similar to "hill-climbing" but in an inverted reference frame (i.e., find the bottoms of the valleys rather than the peaks of hills). The ADALINEs and MADALINE (multiple ADALINEs) of the 1960's only had adaptive elements in one layer. Fig. 7 shows the operation of an ADALINE and a MADALINE. The learning rule for each ADALINE can be stated as follows [1] [3]:

$$\overline{W(k+1)} = \overline{Wk} + ETA(ek(\overline{Xk})/(\|\overline{Xk}\|^2)) \qquad \text{(EQ. \#9)}$$

Where: $\overline{Wk}$ = the present weight vector; each element is a connection weight
- W(k+1) = the next weight vector
- ek = (dk - sk); this is the error signal.
- dk = desired response (i.e., supervised learning)
- sk = ($\sum$wk(xk)) this is the linear output before the threshold device! (i.e., the weighted sum)
- $\overline{Xk}$ = input vector ($\|\overline{Xk}\|$ is the length of the vector)
- ETA = the learning rate (normally ranging from 0.01 to 1.00)

SCALER!



ADJUST THRESHOLD

Figure 7. The ADALINE and MADALINE (Figure from Widrow [3])

1958 to 1969

The state of neurocomputing research during this period of time is depicted in a statement by James A. Anderson, a neurophysiologist and well-published neurocomputing researcher:

"In the popular history of neural networks, first came the classical period of the PERCEPTRON, when it seemed as if a neural network could do anything. A hundred algorithms bloomed, a hundred schools of learning machines contended. Then came the onset of the dark ages, where, suddenly, research on neural networks was unloved, unwanted, and most important, unfunded. A participating factor in this sharp decline was the publication of the book "PERCEPTRONS" by Minsky and Papert." (1969)

Anderson does however point out that this was a brilliant book, but unfortunately in the 1960's there was a general public suspicion of neurocomputing so that this book merely amplified the already growing public discontent with neurocomputing. [1]

1969 Minsky and Papert

They observed several shortcomings of the PERCEPTRON. The shortcoming which received the most public attention was that the PERCEPTRON could not classify all patterns presented to it. (e.g.: the output classification for certain sets of inputs, such as the 2-input exclusive-or, could not be done). This is the problem of linear seperability. [1]

GO to NOTES

1972 Teuvo Kohonen

He developed theories on associative memory. [4]

EX) CHARAC. RECOGNITION

2 → Z

1973 Chr. von der Marlsburg

He observed direct topographical mapping of visual field onto the surface of the brain's visual cortex. [1]

1974 P. Werbos

He developed the theoretical concept of back-propagation learning for neural networks [1].

1975 W.A. Little and Gordon L. Shaw

They observed that brain memory, for short time periods, is "bound-up" with the development of stable states in the nervous system. [1]

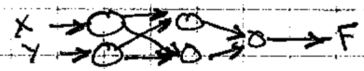
1976 S. Grossberg

He developed Adaptive Resonance Theory (ART); a number of mathematical hypotheses about the underlying principles governing biological neural systems. [3] (e.g., cortex tunes itself to pick-up the most useful features of the environment and ignores or suppresses other information). [1]

SELECTIVE AWARENESS!

LINEAR SEPERABILITY

characteristic of certain problems that was criticized by influential people in the 1960s that said neural networks at that time could not solve non-linearly separable problems. There would be however a later solution to this by layering neurons and having a hidden there in the middle.



EX) TWO INPUT BINARY PROBLEM

BINARY INPUTS X, Y → [ ] → F BINARY OUTPUT

X	Y	F (FUNCTION)
0	0	?
0	1	?
1	0	?
1	1	?

$2^n = 4$ PERMUTATIONS WHEN n = # OF INPUTS

$2^{(2^n)} = 16$ POSSIBLE OUTPUTS

X	Y	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
0	0	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1
0	1	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	0	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	1	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1

#6 #8

TWO EXAMPLES:

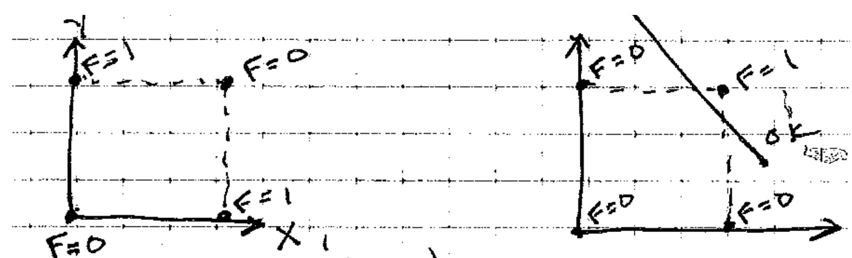
CALLED "XOR" (EXCLUSIVE-OR) FUNCTION

X	Y	F
0	0	0
0	1	1
1	0	1
1	1	0

"AND" FUNCTION

X	Y	F
0	0	0
0	1	0
1	0	0
1	1	1

First plot all of the F values on a graph of X vs Y and show next to each point the value of F, then see if you can draw a line to separate the F=0 and F=1 points



XOR CAN'T DRAW LINE!

AND

1980 Stephen Grossberg

He proposed that a neural network should do error correction by itself, (i.e., unsupervised learning). [1]

1982 J.J. Hopfield

He published a paper containing a coherent, new neural network model. This paper reestablished the favorable public perception of neural networks which was lost during the 1970's. The Hopfield model relies on an exhaustive connecting of all neurons in the network. In the original Hopfield model, each neuron is a binary threshold element performing the hard-limiting, signum function (i.e., output equal to +1 if the threshold exceeded, else -1). The exhaustive connecting of neurons allows lateral inhibition between neurons (also known to exist in brain functioning). The learning algorithm is not a weight changing function of error, common in many other models, but rather an algorithm which determines the degree of connectiveness between all neurons (i.e., active, inhibited, or partial connectivity). This is done by minimizing a global "energy" for the system. This "energy" is a function of the present state of each neuron (+1 or -1) and the fixed weights on each connection. [1] [5]

1982 Teuvo Kohonen

He developed artificial system of self-organizing feature maps that can show the same behavior as the brain's visual cortex for direct topographical mapping of visual images. The structure of the model is a "slab" of processing units in which nearby "neighboring" units respond similarly. When an input is provided to the model, the unit which behaves in the most desirable way has its synaptic weights (from the input relay network) strengthened. The synaptic weights of nearby neighbors are changed so that they behave more like the "well-behaved" unit. [1]

1983 Kunihiko Fukushima, Sei Miyake and Takayuki Ito

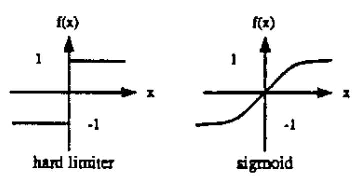
They developed the NEOCOGNITRON; a multilayered neural network dedicated to the recognition of handwritten characters. This meant that the system was considerably constrained because of the nature of its inputs (i.e., only lines of varying degrees of complexity). The NEOCOGNITRON was modeled after the actual anatomy and physiology of a biological visual system. Although this network is multilayered, it learns by a sequentially-directed learning procedure where the expected behavior of each layer is well-known and therefore there is no interdependence between layers during the learning process. [1]

1983 S. Kirkpatrick, C.D. Gelatt and M.P. Vecchi

They developed the concept of SIMULATED ANNEALING to deal with the problem of local maxima (or inversely, local minima) when using "hill-climbing" (or a variation of it such as gradient decent). The problem is that if a topography is being searched for the highest peak (or deepest valley), how do you know if you have merely found a local peak (or valley) which is not the optimum? These undesired outcomes are a result of the step-wise processes normally used which effectively "walk" across the terrain one step at a time and come to a complete stop when the next step in any direction would no longer be uphill (for "hill-climbing"). The solution proposed by SIMULATED ANNEALING is to essentially get a "birds-eye-view" of the terrain before settling to a local hill. The algorithm used is taken from thermodynamics and works with the probability of finding a system in a particular configuration with a given energy. For a neural network error surface (or energy surface, for Hopfield and BOLTZMANN models), the given energy is a function of the elevation of the topography. [1]

1984 J.J. Hopfield

Hopfield substitutes the non-linear sigmoid for the hard-limiting signum function previously used as a transfer function in his network's neurons. (see Fig.8) [1] [6]



CONTINUOUSLY DIFFERENTIABLE

Figure 8. The Signum and Sigmoid Transfer Functions (Figure from [6])

1985 David H. Ackley, Geoffrey E. Hinton and Terrence J. Sejnowski

They developed the BOLTZMANN MACHINE; a neural network with the same basic units as Hopfield's original 1982 model (signum transfer function) and with learning based on a similar energy function. However, for the BOLTZMANN MACHINE, the state of the neuron is made probabilistic (hence a BOLTZMANN probability distribution). The BOLTZMANN MACHINE also uses SIMULATED ANNEALING (Kirkpatrick, et al. 1983) to avoid getting stuck at local minimum energies. Another significant difference is that the BOLTZMANN MACHINE was made to work for a multilayer network. [1]

1986 D.E. Rumelhart, G.E. Hinton and R.J. Williams

They developed the BACK-PROPAGATION multilayered neural network. The neurons of this network use a non-linear sigmoid transfer function applied to a weighted sum of input signals (including a bias) to give an analog output ranging from 0 to 1. However, since the sigmoid function is asymptotic, output values of 0.1 and 0.9 are considered as 0 and 1 for classifications requiring binary transformations. The most significant aspect of this network is the learning process via backward propagation of error signals from one layer of neurons backwards (toward the input) to a previous layer. This kind of learning is called the "generalized delta rule". The error surface created is searched by the method of "gradient decent" somewhat similar to that used by the LMS algorithm of the MADALINE (Widrow, 1960). The network learns as follows: [1] [3] [4] [7]

1) Feed input layer an input vector.

2) Propagate the signals forward via non-linear sigmoidal transformation at each layer until an output vector is created at the last layer.

3) Create an error signal from the difference of the actual output signals and desired outputs provided to the network.

4) Use the error signal to change the synaptic connection weights between the output layer and the previous layer.

5) Create an "effective" error signal from the previous error signal and the weights that were just changed.

6) Use this new error signal to change the synaptic connection weights located one level closer to the inputs.

7) Repeat steps 5 and 6 until weights have been changed at all levels (i.e., this procedure works for any number of levels).

8) Repeat steps 1 to 7 for each input/desired-output exemplar pair in the training set.

9) Repeat steps 1 to 8 until the desired outputs have been obtained for every input/desired-output exemplar pair.

For the back-propagation network shown in Fig. 9 the mathematical representation of the learning algorithm is as follows:

$$w(k+1,j+1) = w(kj) + ETA(ek)(xj) \qquad \text{(EQ. \#10)}$$

Where:

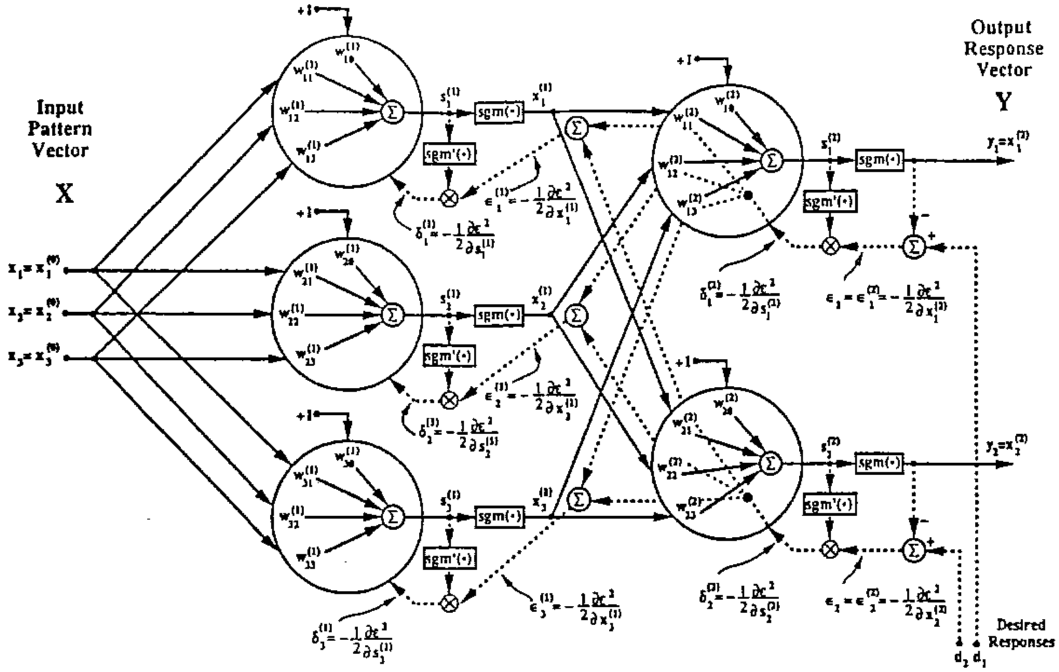
w(kj) = one of the elements of the weight matrix [Wkj] for the connections between the hidden layer (level j) and the output layer (level k).
w(k+1,j+1) = the next value of element w(kj)
ek = the error signal for one of the neurons of the output layer
= (dk - yk)(yk(1-yk)) where yk(1-yk) is the derivative of the sigmoid transfer function with respect to yk
dk = desired response (i.e., supervised learning)
yk = the analog output from one of the neurons of the output layer
xj = one of the outputs from the hidden layer
ETA = the learning rate (usually ranging from 0.01 to 1)

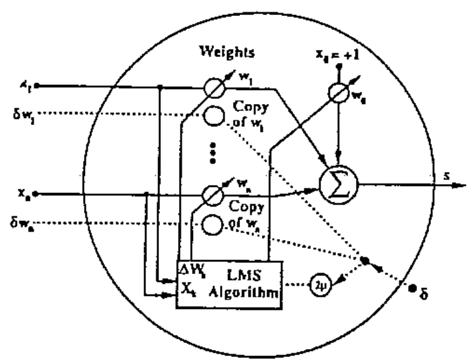
And for the connection weights between the input layer and the hidden layer:

$$w(j+1,i+1) = w(ji) + ETA(ej)(xi) \qquad \text{(EQ. \#11)}$$

Where:

w(ji) = One of the elements of the weight matrix [Wji] for the connections between the input layer (level i) and the hidden layer (level j).
w(j+1,i+1) = the next value of element w(ji)
ej = The "effective" error signal back-propagating through one of the neurons of the hidden layer.
= (back_error)(xj(1-xj))
back_error = The back-propagating weighted sum of error signals (ek's) coming backward from each of the neurons in the output layer (i.e., (ek)(Wkj) for a neuron in level j)
xi = One of the inputs to the network





Detail of linear combiner and associated circuitry

Figure 9. The back-propagation network (Figure from Widrow [3])

1987 B. Widrow, R.G. Winter and R. Baxter

They developed the MADALINE II by expanding the MADALINE to a multilayer network. The learning rule is a least mean square algorithm which uses hamming errors and corrects weights in a direction co-linear to the forward feed of the input vector (i.e., does not propagate backwards). The transfer function is still a hard-limiting signum function (See Fig. 10) [3]

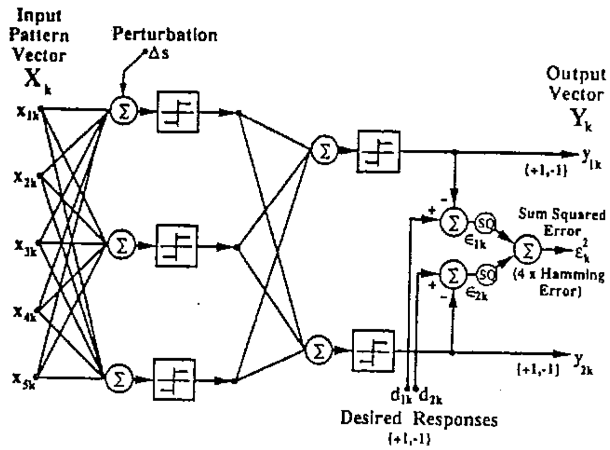


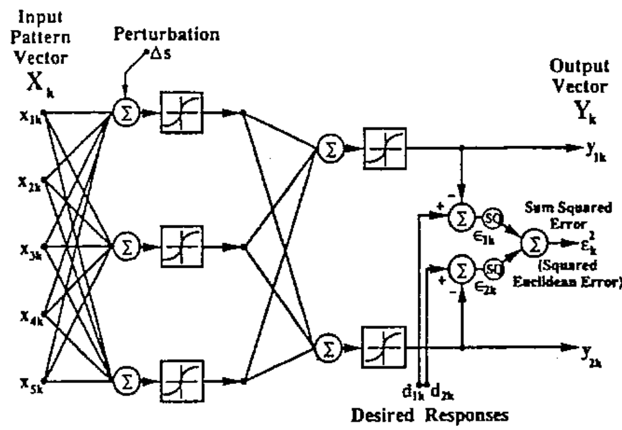
Figure 10. MADALINE II (Figure from Widrow [3])

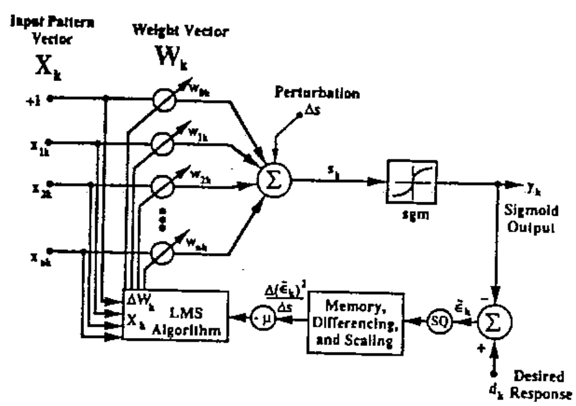
1987 B. Kosko

He developed bidirectional associative memory (BAM), a two layered neural network used to determine if two patterns match according to a predetermined association set in the network. [5]

1988 David Andes

He developed the MADALINE III by modifying the MADALINE II. The hard-limiting signum transfer function was replaced with the non-linear sigmoid function. [30] Widrow observed that the resulting learning rule was mathematically equivalent to back-propagation for small perturbations of the linear outputs. (See Fig. 11) [3]

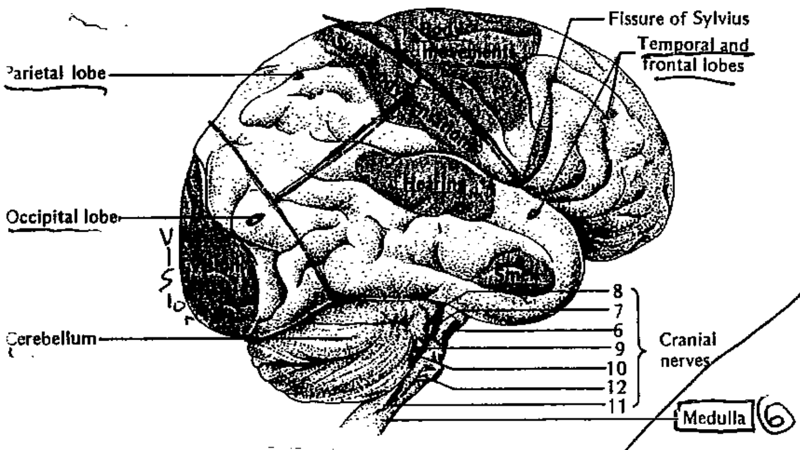


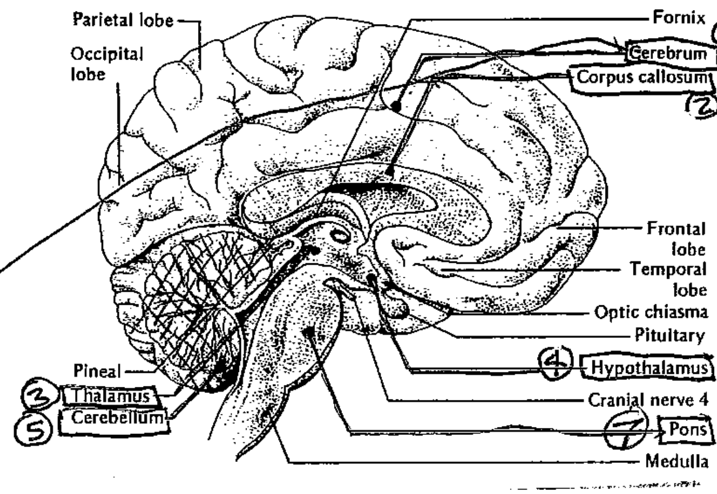


Implementation of the MRIII algorithm for the sigmoid Adaline element.

Figure 11. MADALINE III (Figure from Widrow [3])

HUMAN BRAIN 2003





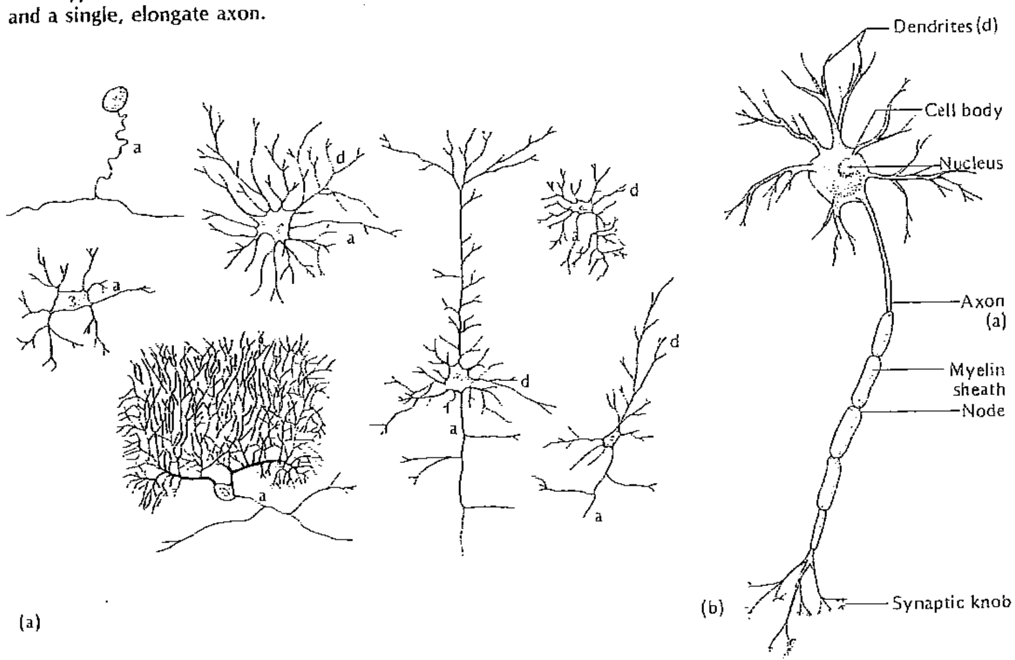
The structures of the human brain

Structure	Description	Function
1 Cerebrum		
Cortex (outer surface)	Thin (3–6 mm), highly folded sheet with large surface area; most recent evolutionary development; high degree of development	Center of the mind; seat of conscious thought, memory, speech, intelligence, personality, and judgment; center of sensory-motor coordination
White matter (inner cerebrum)	Composed of nerve tracts connecting parts of cortex to each other and to rest of brain and spinal cord	Center of sensation, including sight, hearing, taste, etc.
2 Corpus callosum	Consists of 300 million separate neuronal lines	Connects the halves (hemispheres) of the cortex
3 Thalamus	Groups of cell bodies arranged in shape of a football	Relays sensory information to cortex (vision, hearing, touch, taste); regulates sleep and wakefulness in conjunction with limbic system
NOT SHOWN Limbic system Hippocampus Fornix Cingulate gyrus Amygdala NEGATIVE EMOTION • ANGER • FEAR • SADNESS	Internal region of forebrain, divided into several regions	Involved in emotion, motivation, and reinforcement
4 Hypothalamus	Linked to sympathetic nervous system and pituitary	Controls basic drives (eating, drinking, sleeping, sexual behavior); regulates temperature, blood pressure, heart rate
5 Cerebellum ("little brain")	Baseball-sized lump of gray-white tissue lying on either side of the brain stem	Involved in sensory-motor coordination; if damaged, loss of equilibrium and motor coordination results
6 Brain stem (medulla)	Extension of spinal cord within the skull	Involved in control of respiration, heart rate, and gastrointestinal function
NOT SHOWN Reticular formation	Consists of fibers and cell bodies running from spinal cord to cerebral hemispheres	Involved in control of arousal and alertness
7 Pons	Consists mainly of white matter lying anterior to cerebellum and between midbrain and medulla	Acts as bridge linking the various parts of the brain and as relay station from medulla to higher centers

FROM "BIOLOGY: A HUMAN APPROACH" I.W. SHERMAN AND V.G. SHERMAN

NEURONS

FIGURE 17–2 (a) A variety of neurons from different parts of the human nervous system. The axon of each neuron is designated by the letter a, and the dendrites are indicated by the letter d. Dendrites receive excitation from other cells and conduct impulses toward the cell body, whereas axons conduct nerve impulses away from the cell body. (b) The "typical" neuron shows many short dendrites and a single, elongate axon.



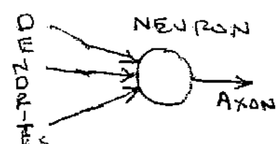
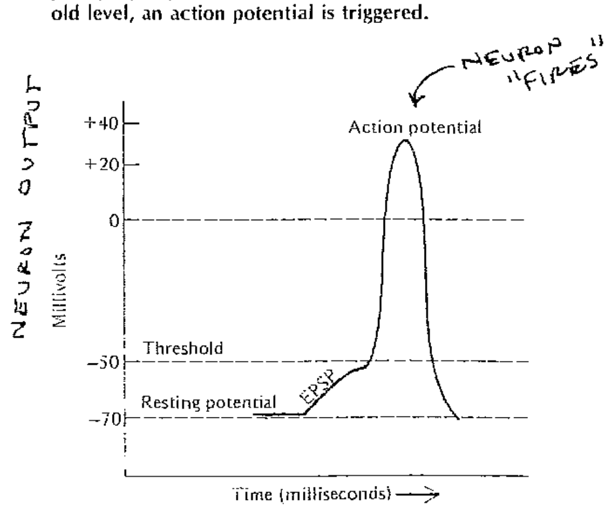


FIGURE 17–12 Initiation of an action potential (nerve impulse) by transmitter substance. Transmitter substance reaches the postsynaptic neuron, reducing its potential and creating an EPSP (excitatory postsynaptic potential). If the EPSP reaches a threshold level, an action potential is triggered.



17–6 INTEGRATED CIRCUITS OF NEURONS: THE NERVOUS SYSTEM

Basic design

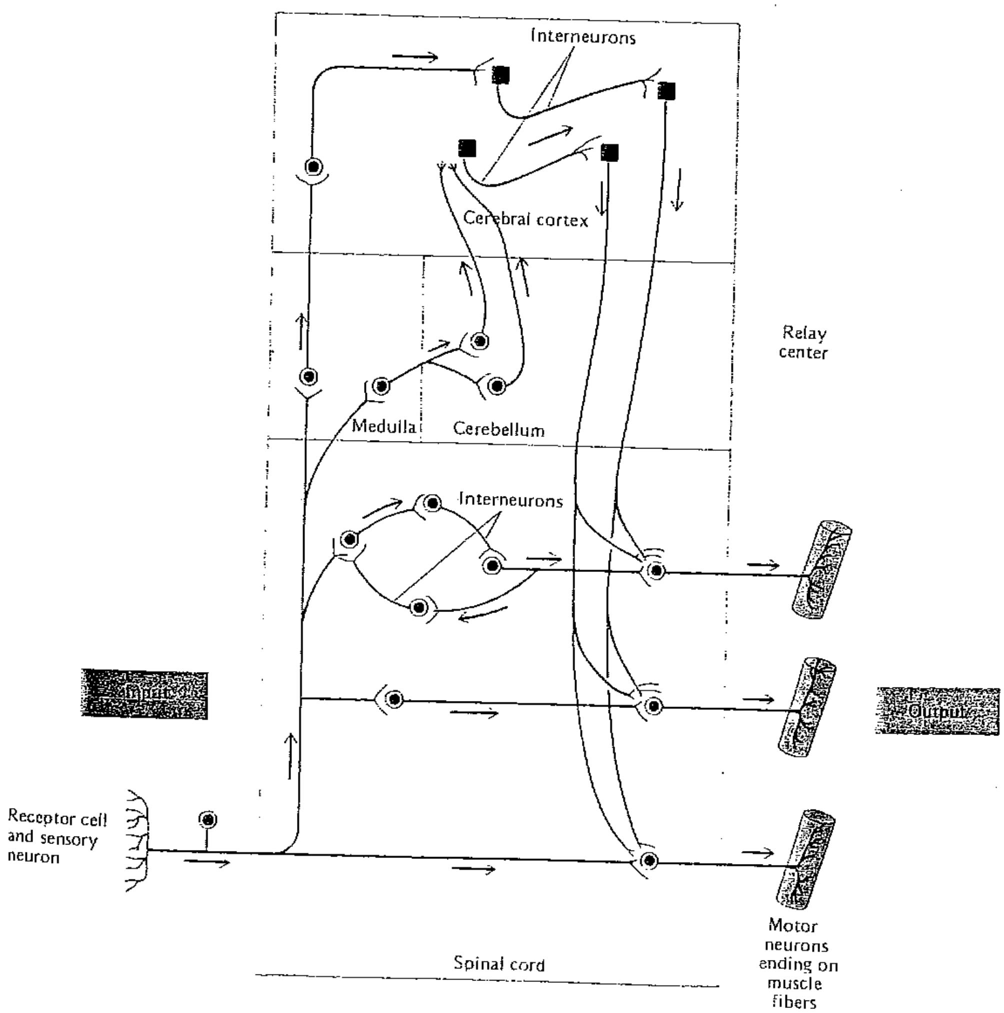
Open the back of a television set or a radio and you will see a maze of wires and printed circuits. Look inside a computer or a telephone switchboard and you will find a tangle of interconnecting wires and other electronic gadgetry. The wires and circuits of these communication devices are the fundamental units of operation, but the sequence and arrangement of the components determine how the machine operates. The fundamental unit of structure in the nervous system is the neuron, and similarly the way neurons are hooked together determines the manner in which the nervous system functions (Figure 17–13).

There are three different functional classes of neurons: (1) **sensory neurons**, which receive stimuli from the environment and transmit information to the central nervous system (brain and spinal cord); (2) **motor neurons**, which conduct messages from the brain and spinal cord to the glands and muscles; and (3) **interneurons**, which act in an integrative capacity and shuttle signals back and forth between the neurons of various parts of the brain and spinal cord. Over 99% of the neurons of the body and brain are interneurons.

a simplified way Figure 17–13 illustrates ematic wiring diagram of the nervous system. Note that an impulse entering the spinal cord via a sensory neuron has many possible pathways. Rarely does the signal that traverses a sensory neuron directly activate a motor neuron leading to an effector (for example, muscle); typically the signal travels upward via the spinal cord interneurons and through a number of relay centers in the brain before reaching the higher centers. From there the command signals travel down the spinal cord, again via relay centers, and via a motor neuron the effector is triggered to activity. Note that the more neuronal cells in the circuitry, the more flexible can be the response.

The nervous system's circuitry is composed of two basic subdivisions: the central nervous system (CNS), comprising the structures encased within the skull and the vertebral column, and the peripheral nervous system (PNS), which lies outside the skull and vertebral column but connects up with the CNS via spinal and cranial nerves.

FIGURE 17–13 Simplified circuitry of the nervous system. Neurons are arranged into cables consisting of many axons and dendrites. Axons, bundled together to form a multistranded cable, form the nerve fiber or nerves we commonly see in a dissected specimen. The collections of axons and dendrites in the brain and spinal cord—the information centers—are called tracts.



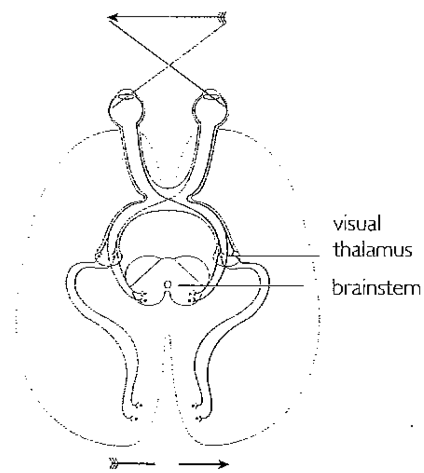
2003
J. WUNDERLICH HANDOUT

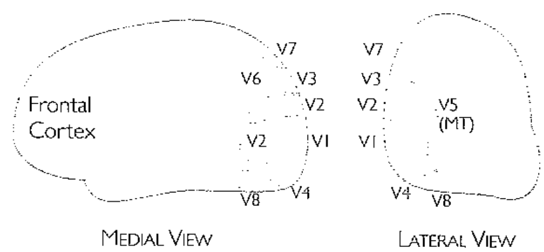
FROM
"MAPPING THE MIND"
BY RITA CARTER 1998

SENSING THE WORLD

Vision

Light from a visual stimulus is inverted as it passes through the lens. It then hits the retina at the back of the eye, where light-sensitive cells turn it into a message of electrical pulses. These are carried along the optic nerve from each eye and cross over at the optic chiasma – a major anatomical landmark. The optic track then carries the information to the lateral geniculate body, part of the thalamus. This shunts it on to V1 at the back of the brain. The visual cortex is split into many areas, each processing an aspect of sight, such as colour, shape, size and so on.





Above: *Awareness of sighted objects is conveyed to the limbic system but is not consciously visual.*
Below: *Each visual element is processed by a separate brain area.*

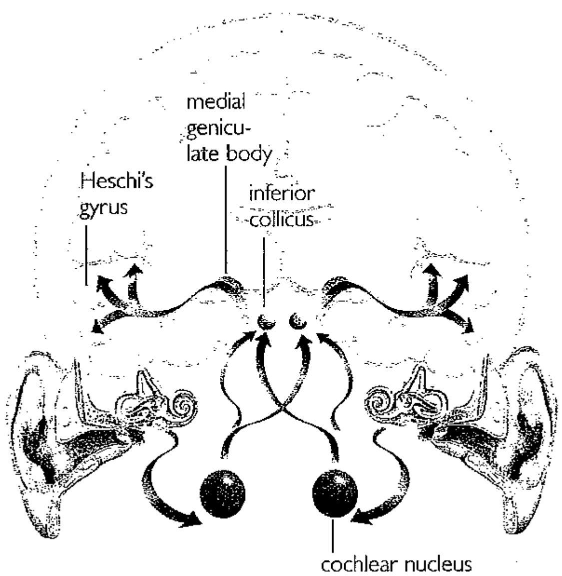
Layout of visual cortex:
V1 – general scanning
V2 – stereo vision
V3 – depth and distance
V4 – colour
V5 – motion
V6 – determines objective (rather than relative) position of object
'Where?' path: V1-V2-V3-V5-V6
'What?' path: V1-V2-V4

V1 mirrors the world outside in which each point in the external visual field matches a corresponding point on the V1 cortex. When a person stares at a simple pattern like a grating the image is reflected by a matching pattern of neuronal activity on the surface of the brain.

The 'map' is distorted, as the neurons responding to the central area of the visual field take up a much greater cortical area – so the 'picture' painted on V1 is a little like that seen through a fish-eye camera lens.

The centre of the retina, the fovea, is much more densely packed with neurons and sees far more detail. The eyes therefore dart around, in a series of leaps called saccades, in order to scan the visual field in detail. Saccades are triggered by the attention system of the brain and are not generally under conscious control.

You do not need eyes to see. Blind patients have been fitted with a device that turned



The neural pathways that convey sound information to different parts of the brain.

low level video pictures into vibrating pulses that could be 'read' tactilely, rather like Braille. A camera mounted next to the subjects' eyes, spread the pulses – which felt like a grid of tingles – over their backs, so they had continuous sensory input from the visual world. The patients soon started to behave as though they were 'really' seeing. They ceased to be aware of the tingles and their 'point of view' shifted to the camera. One of the devices had a zoom lens, and when an experimenter – without warning – operated the zoom, causing the image on the subject's back to expand suddenly as though the world was looming in, the subject ducked and raised his arms to protect his head.

However, there seemed to be a limit to the impact of visual information presented in this way. After the (male) subjects became practiced 'viewers', an erotic picture was projected – the subjects were able to describe it accurately but were unmoved by it.

Hearing

The neural pathways carrying sound information from each ear divide into unequal parts once they leave the ear.

On each side the broader path goes off to the brain hemisphere opposite the ear from which it came, so sound from each ear reaches both hemispheres – but most of the left ear's signals go to the right hemisphere and vice versa.

Both hemispheres have a distinct role in sound processing, and this means that sounds are dealt with (and therefore experienced) slightly differently according to which ear they enter. For example a person deaf in the right ear will receive most sound signals in the left auditory cortex (the side of the brain opposite the 'good' ear). This is the side that deals mainly with the identification and naming of sounds rather than their musical quality, so rhythm and melody perception may be blunted.

Conversely, a person deaf in the left ear may find that words are more difficult to distinguish than music, irrespective of loudness.

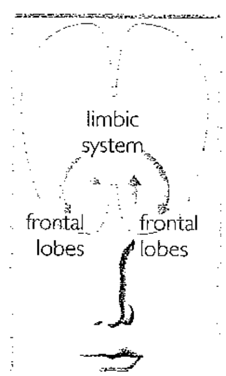
Smell

Flavour perception seems to be processed separately from either smell or taste. In one study students who learnt new words while sniffing an unusual smell and then sniffed the smell again when they had to recall the words showed a 20 per cent boost in memory power.

Whether we find a smell nice or nasty depends crucially on what memories we associated with it. For one person the smell of a bonfire may bring back happy memories of fireworks and winter barbecues, for another it may bring melancholy recollections of summer's end. The first person is likely to find the smell pleasant and the second to dislike it.

Scanning studies suggest that pleasant odours mainly light up the frontal lobes' smell area, particularly on the right-hand side. Unpleasant odours activate the amygdala and the cortex in the temporal lobe (insula). Unlike other senses smell passes directly to the limbic system. This fast route to the brain's emotional centre gives smell its power to elicit strong emotional memories.

Smell is different from other senses because it goes straight to the limbic system – a fast route to the brain's emotional centre. Unlike other senses it does not cross from nostril to opposite hemisphere.



Taste

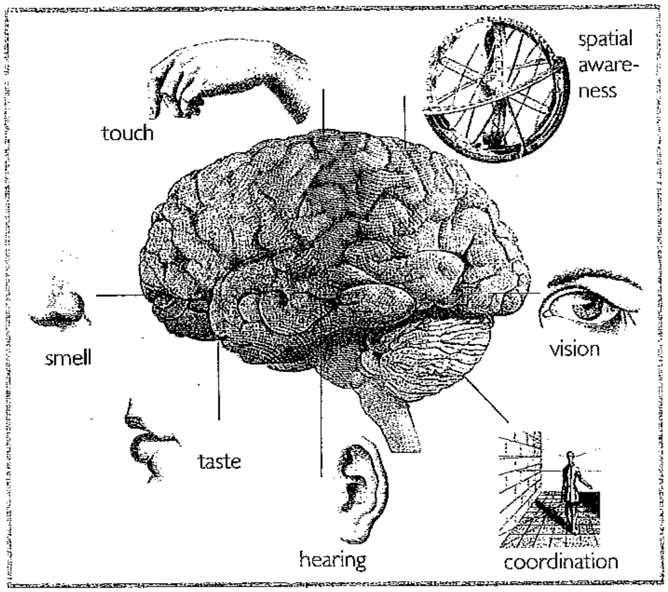
Damage to the frontal lobe of the right hemisphere may turn ordinarily hungry people into fanatic seekers of fine food. Gourmand syndrome has been identified by Swiss researchers who first suspected it when two of their patients developed foodie obsessions after sustaining brain injuries. The researchers subsequently scanned thirty-six gourmands: thirty-four of them had lesions in the right frontal lobes. The mechanism causing the new interest in food has yet to be revealed – serotonin levels in the frontal lobe may play a part.

Sensation

Sensation travels along several different types of nerves to the brain. Pain is carried by two types of nerve – fast, which carries sharp pain; and slow, which carries deep, burning pain. Stimulation of one type blocks messages from the other by closing a 'gate' in the spine. That is why 'rubbing it better' is effective.

The anterior cingulate cortex – an area primarily associated with emotion and attention – is essential for conscious pain. Opioid-type analgesics (including morphine and codeine) are the most effective type of painkillers. They block the receptors in brain neurons normally be filled by enkephalins – the brain's own painkilling chemicals, which are released by acute pain stimuli. Opioids also damp down activity in the anterior cingulate cortex.

The importance of the anterior cingulate cortex in pain perception is demonstrated by brain scans showing that people with cardiovascular disease appear to get angina – the chest pains associated with lack of oxygen to the heart – only when the anterior cingulate cortex is active. In some people, it seems, the anterior cingulate cortex lights up as soon as the heart is short of oxygen. This creates conscious pain, warning them to stop doing whatever is straining the heart. In others the heart can be severely short of oxygen before the anterior cingulate cortex is activated. These people can develop potentially dangerous heart disease without angina, making them vulnerable to surprise heart attacks.

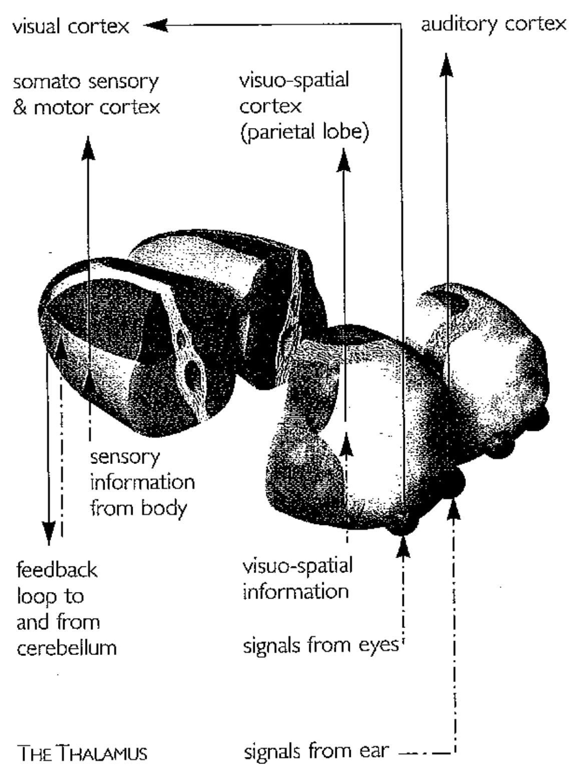


Left: *The vast majority of the cortex is given over to sensory processing – only the frontal lobes are dedicated to non-sensual tasks.*

Below: *All incoming sensory information (except smell) goes first to the thalamus. This limbic nucleus acts like a relay station, shunting the data onto appropriate cortical areas for processing.*

The sixth sense

Proprioception is the sense of body awareness telling us the position of our limbs, our posture and equilibrium. It involves the integration of several sensory inputs: touch and pressure sensations from skin, muscles and tendons; visual and motor information from the brain; and data about our balance from the inner ear. Together they amount almost to a sixth sense. Proprioception uses so many different brain areas that it is very rare for it to be lost altogether. Occasionally, though, people suffer brain injuries that so disturb proprioception that they lose all sense of having a body. Certain meditative states involve dissociating the conscious brain from proprioceptive input, inducing a feeling of disembodiment and maybe give the impression of floating or levitation. Out-of-body experiences, in which people report becoming detached from their bodies and floating around in mid-air, may be due to temporary loss of proprioception.



THE THALAMUS