SESSION 1

PAPER 5

CONDITIONAL PROBABILITY COMPUTING IN A NERVOUS SYSTEM

---

by

DR. A. M. UTTLEY

BIOGRAPHICAL NOTE

Dr. Uttley took an Honours degree in Mathematics at King's College, London where he also took a degree in Psychology and did post-graduate research in Visual Perception. At the Royal Radar Establishment he designed and built analogue and digital computers. For the last five years Dr. Uttley has been working on theories of computing in the nervous system.

CONDITIONAL PROBABILITY COMPUTING IN A NERVOUS SYSTEM

by

DR. A. M. UTTLEY

ABSTRACT

IN two previous papers it has been suggested that two particular mathematical principles may underlie the organization of nervous systems; the first is that of classification (Uttley, 1954, *ref.* *13*) and the second is that of conditional probability (Uttley, 1956, *ref.* *14*). The suggestion is based on the similarity of behaviour of these formal systems and of animals. The design of classification computers is discussed in the first paper; the design of conditional probability computers is discussed in a third paper (Uttley, 1958, *ref.* *15*); in both papers working models are described. Further reference to these papers will be by date only. It is the aim of the present paper to consider whether the two principles might operate in nervous systems.

There are four requirements for the principle of classification to operate in an area of a nervous system. Firstly, in that area, signalling must be binary; this would be the case if, for example, the impulse frequency were at either a very low rate or at a maximal rate, or if signalling were in terms of standard volleys; in general, if the fibre activity were in one of only two states.

The second requirement is that the fibres which form the input to the area be connected to neurons in as many different ways as possible; there are many areas in which this condition is met.

The third requirement is that more than one synapse of a neuron must become active for it to fire; this appears to be met.

The fourth requirement is that there shall be some way of delaying signals for periods of the order of seconds. A block of isolated cortex does remain active for such periods when stimulated briefly so in this way the requirement might be met.

If these conditions are all met each neuron will indicate, by firing, the occurrence of a particular spatio-temporal pattern of activity in the input to the system.

If the system is to have the additional property of computing conditional probabilities and so of making inferences it must meet four additional requirements three of which have been suggested by Burns (1955, *ref.* 4) to explain the spontaneous firing of a single neuron. The first two demand that some physical quantity in a neuron shall vary (e.g. a chemical concentration). The quantity must change when the neuron fires and recover at a slow rate when the neuron is not being activated.

The third requirement is that the effect of one firing neuron upon another to which it is connected must depend on their relative state. This hypothesis has been made before by Burns, (loc. cit) who has recently found evidence to support it.

The last requirement is met if there are short axon neurons, embedded in the population of indicator neurons already described, whose axonal systems have an inhibitory effect.

From the three papers, two important points emerge. Firstly, a theory designed to explain only conditioning and extinction was found to explain, without extension, a number of other forms of plastic behaviour. Secondly, from the theory, the rules for the design of a model were deduced which can be translated into reasonably biological requirements.

There are at least two lines of work which may lead to tests of the theory. Firstly, a study of the mathematics of random connexions, using anatomical data of cell and fibre density, will lead to laws describing the limited powers of classification of randomly connected neurons. These powers can be compared with known psychological limits to discrimination. Secondly, in terms of the theory, learning laws in experiments on very simple animals may be translatable into laws of reversible chemical reactions; such work may contribute to an understanding of synapse chemistry.

CLASSIFICATION SYSTEMS

The input to a classification system is a number of separate channels each in a changing state of activity; the system can distinguish complex spatio-temporal patterns of activity in these channels. The output from a classification system consists of a number of channels too and the system can synthesise complex patterns of activity in them.

The conditions for the existence of a classification system have been stated formally (1954, *ref.* 13 and 1958, *ref.* 15) and they are repeated here.

1. Each input channel must be always in one of two states active and inactive.
2. The inputs must be combined in as many ways as possible - ideally in all possible ways.

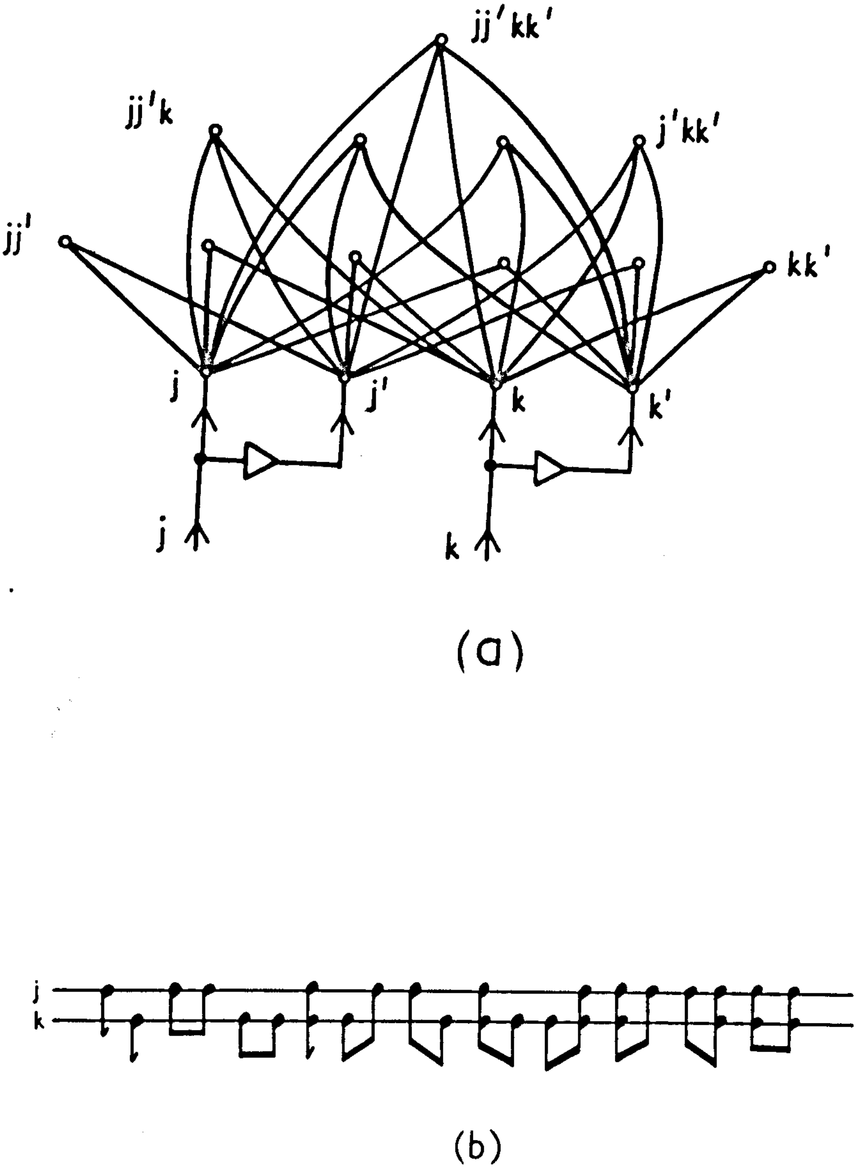

Fig.1. The Classification of canalo-temporal patterns of activity in two binary inputs (a). A system of coincidence units and delays, (shown as triangles) which distinguishes only 'before', 'simultaneous' and 'after'. (b) The patterns which can be distinguished by the system.

3. There must be a unit for every combination of inputs, which *indicates* if every input of the combination is active. A combination or set of inputs is said to define a *pattern* of activity. The connexions between inputs and units are called *counting connexions.*
4. If temporal patterns are to be distinguished each input must pass through a series of delays; the output of each delay must provide a separate input to the system of indicating units.

A very simple classification system is shown in *figure 1a*; there are only two inputs *j* and *k* and because each input suffers only one delay the system can distinguish temporally only 'before', 'simultaneous' and 'after'; nevertheless it can distinguish all the patterns of *figure 1b*. Without any extension of principle or of rules of design the number of inputs and of delays can be extended indefinitely so that the system can distinguish very complex patterns; there is always one indicating unit for each pattern.

Binary Inputs

Nervous transmission is in terms of standard impulses which meet the requirements of binary classification. However, at low levels in nervous systems, intensity is signalled in terms of impulse frequency. It would not be possible for a train of impulses to signal both an intensity and the number of times that a pattern had occurred. If, at higher levels, patterns are distinguished by classification then intensity must not be signalled in terms of frequency but in terms of 'place'. There are several possible mechanisms by which this coding could occur.

Firstly, by means of a threshold mechanism a system can distinguish signals which are either above or below a certain energy level; all neurons appear to possess such thresholds. The second principle is that of adaptation whereby changes are emphasised; for a signal which consists of pulses of varying frequency if the pulse rate suffers such differentiation many times the signal will be changed into a series of standard volleys - standard because the firing rate of a neuron can lie only between two definite limits. Such a form of activity meets the first requirement.

In a number of areas there is a tendency for firing rates to be standardised; in the auditory nerve, for example, signalling is in terms of pulse frequency (Galambos and Davis 1943, *ref. 8*) but Hilali and Whitfield (1953, *ref. 9*) have shown that, in efferent fibres of the cochlear nucleus and for steady stimulus tones, this frequency is, broadly speaking, either at about 50 pulses per second or it is at the slow background rate. A similar effect appears to occur in ventral horn ganglia so that the force in a muscle is determined more by the number of muscle fibres contracting maximally than by the impulse rate in individual fibres.

SCIENTIFIC LIBRARY
APR 22 1960
U. S. PATENT OFFICE

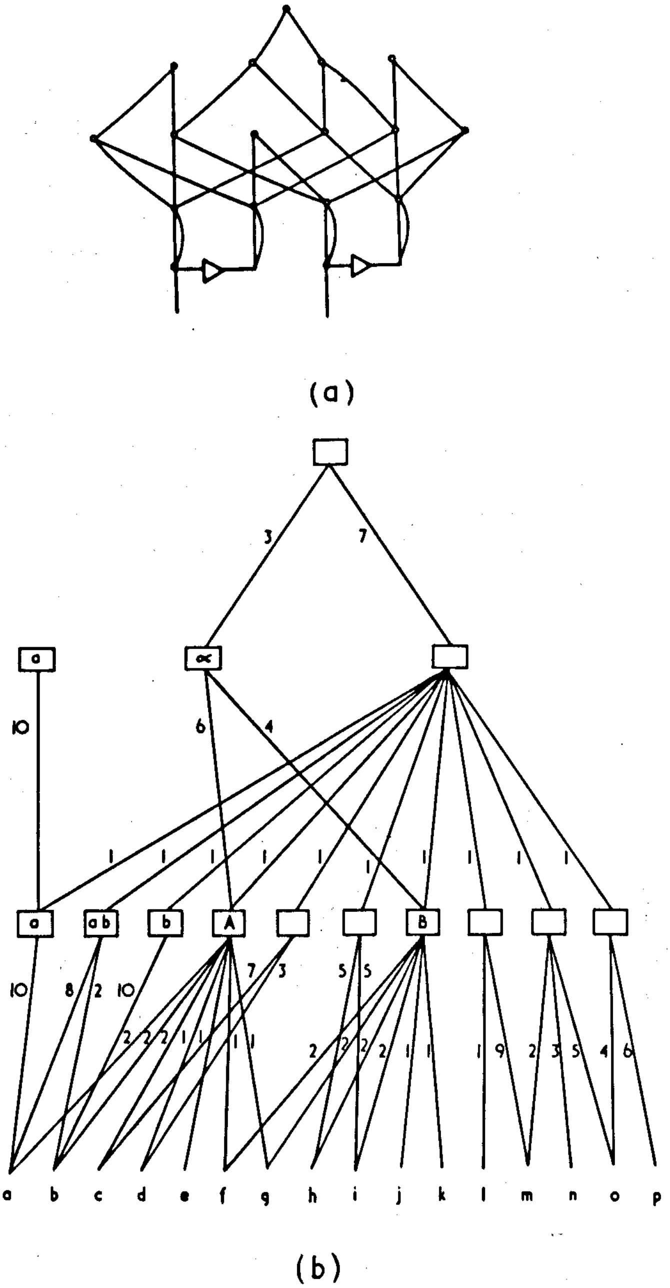

Fig.2. Classification by means of identical units each of which indicates if a fixed number of inputs to it become active, (a) when the number is two, (b) when the number is 10.

The Indicating Units

It has been pointed out (1958, *ref. 15 p. 14*) that a classification system can be constructed from units of a common design if each unit indicates when more than a critical number of the inputs to it are active; the critical number must be two or more. It is easiest to see this if the critical number is taken as two: the system of *figure 1a* then takes the form of *figure 2a*. There are two inputs to each unit; if they are active simultaneously an output of the unit is active and can serve as an input to further units. A classification system using units for which the critical number is ten takes the form of *figure 2b*; it has been discussed in an earlier paper. (Uttley, 1954, *ref. 13*). If the critical number of impulses is different for different neurons classification will still arise but its form will vary in the way shown in *figures 2a and 2b*.

It can be seen that in *figures 2a and 2b* a principle of coincidence is employed so that an *ab* unit, for example, could not fire more often than an *a* or a *b* unit. It follows that the higher the level in such a system the lower the rate at which its units indicate; it would appear inevitable, if more than one synapse must become active for a neuron to fire, that the same statement must be true of a nervous system.

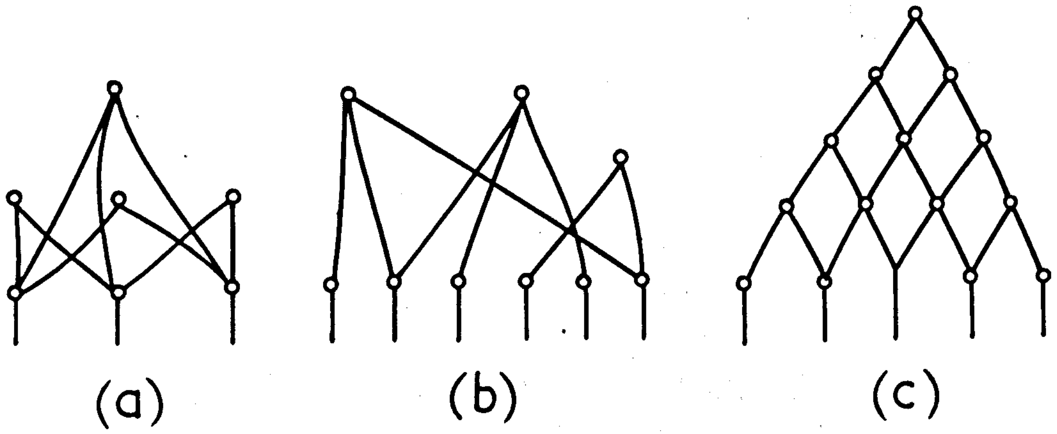

Fig.3. (a) Complete classification of three inputs.
(b) Incomplete classification of specially chosen patterns.
(c) Incomplete classification arising because connexions are limited in length.

The Connexions

A classification system can be complete, as in *figure 3a*, and able to distinguish all possible patterns - this is a mathematical ideal; or it can be incomplete and designed to distinguish special patterns, as in *figure 3b*; or it can be incomplete and yet have no special structure to enable it to distinguish special patterns - such an arrangement is shown in *figure 3c*. It has been suggested (1954, 1956, *refs. 13,14*) that in the second form the system can imitate releaser mechanisms and pointed out (1954, *ref. 13*) that the third form arises if the connexions between input channels and units are entirely random; this possibility will now be considered further.

Consider a set of afferent fibres arborising in a population of identical neurons each possessing a random dendritic system; the axonal systems of such neurons will not be considered at this stage. Suppose that synapses form by chance proximity and that a neuron fires if it receives two simultaneous impulses. Such a system is shown diagrammatically in *figure 4a* and more realistically in *figure 4b*; it possesses three features which do not occur in the formal classification system. Firstly there is more than one neuron with, for example, one connexion to an afferent fibre *a* and another to a fibre *b*; so, corresponding to the *ab* unit of the formal system, there is a diffuse cluster of *ab* neurons, and there is such a cluster for every unit of the system. The different clusters overlap but there is no confusion because each neuron is permanently labelled in terms of its connexions; it does not matter where a neuron is but what it is connected to. Because the units are duplicated such a neural system can sustain much damage before it is unable to distinguish patterns. Secondly, there are neurons with less than two connexions; so they will never fire. In the limit, if fibres *a* and *b* are sufficiently far apart there will be no *ab* neurons, so classification will be incomplete; the actual number of neurons which distinguish the pattern *ab* will therefore depend upon the separation of the afferent fibres. Thirdly, if a neuron has connexions to more than two fibres it will indicate ambiguously, that is, when any two of the connexions are active.

If the neurons possess axonal systems these will make similar connective arrangements with the dendritic systems of further neurons; this is shown in *figures 4a and 4b*. Because these neurons are indirectly connected to the original fibres via one intermediate neuron they will be said to effect *first order indirect classification*. There are two consequences; firstly, there will be more duplication of indicating units; secondly, there will be units which distinguish more complex patterns. The connective arrangements of *figure 2a*. can therefore arise, and the second of the above conditions is met.

The system can be extended to higher orders indefinitely, but it will be stable only under special circumstances. Consider only the *a* neurons of *figure 4*; the number which are indirectly connected to the *a* afferent fibre through one intermediate neuron might exceed the number of directly connected *a* neurons; similarly the second order neurons might exceed the first order neurons, and so on indefinitely; if the *a* fibre became active there would then be an uncontrollable increase in the number of active neurons. The magnification ratio per order depends on the total length of the axonal system of a neuron (1955, *ref. 16*); if the ratio exceeds unity any input can fire the whole population. In terms of classification each unit is then an *a* unit OR a *c* unit OR ... and there is complete ambiguity of classification.

If, on the other hand, the magnification ratio is kept below unity so that the system is stable, then the actual number of neurons of a given order will decrease with the order. In consequence, as the complexity of a pattern increases the number of neurons which distinguish that pattern will decrease; but the nearer to the instability point that a system can be held, the greater the complexity of the patterns that it can distinguish.

The stability of a population of neurons has been studied by Beurle (1955, *ref. 3*), Allanson (1955, *ref. 1*) and the writer (1955, *ref. 17*); this work must be related to experiments on isolated cortex (e.g. Burns, 1951, *ref. 5*) and to measurements of fibre density (e.g. Sholl, 1953, *ref. 12* and Eayrs, 1955, *ref. 6*).

The Number of Units

It has been suggested above that each neuron distinguishes a particular pattern of activity by virtue of its connexions. The hypothesis will be made that these connexions do not change, once having formed; although their function may. Consequently it is most important to consider whether the number of neurons in a nervous system are adequate to account for all the patterns of activity that an animal can distinguish.

The first point to be made is that one is here considering the patterns of activity which are initiated within the system - the tunes which are being played upon the receptors - not the limitless patterns of the external environment which can arouse the internal patterns. If there were n fibres which signalled in a binary way to a neural classification system then, for complete classification, that system would have to contain 2^n neurons. From the point of view of this paper the evolution of nervous systems is seen a struggle to solve that intractable problem; the failure to solve it can be seen in a number of known psychological limits to pattern discrimination. Incidently, if each input fibre possessed not two but m discriminable levels of activity, the required number of units would be, not 2^n but m^n; this is the fundamental reason why the input fibres should use a binary system.

At least three methods appear to be used in nervous systems in order to reduce the number of discriminating units. Firstly, the number of fibres entering the brain is reduced to much less than the number of

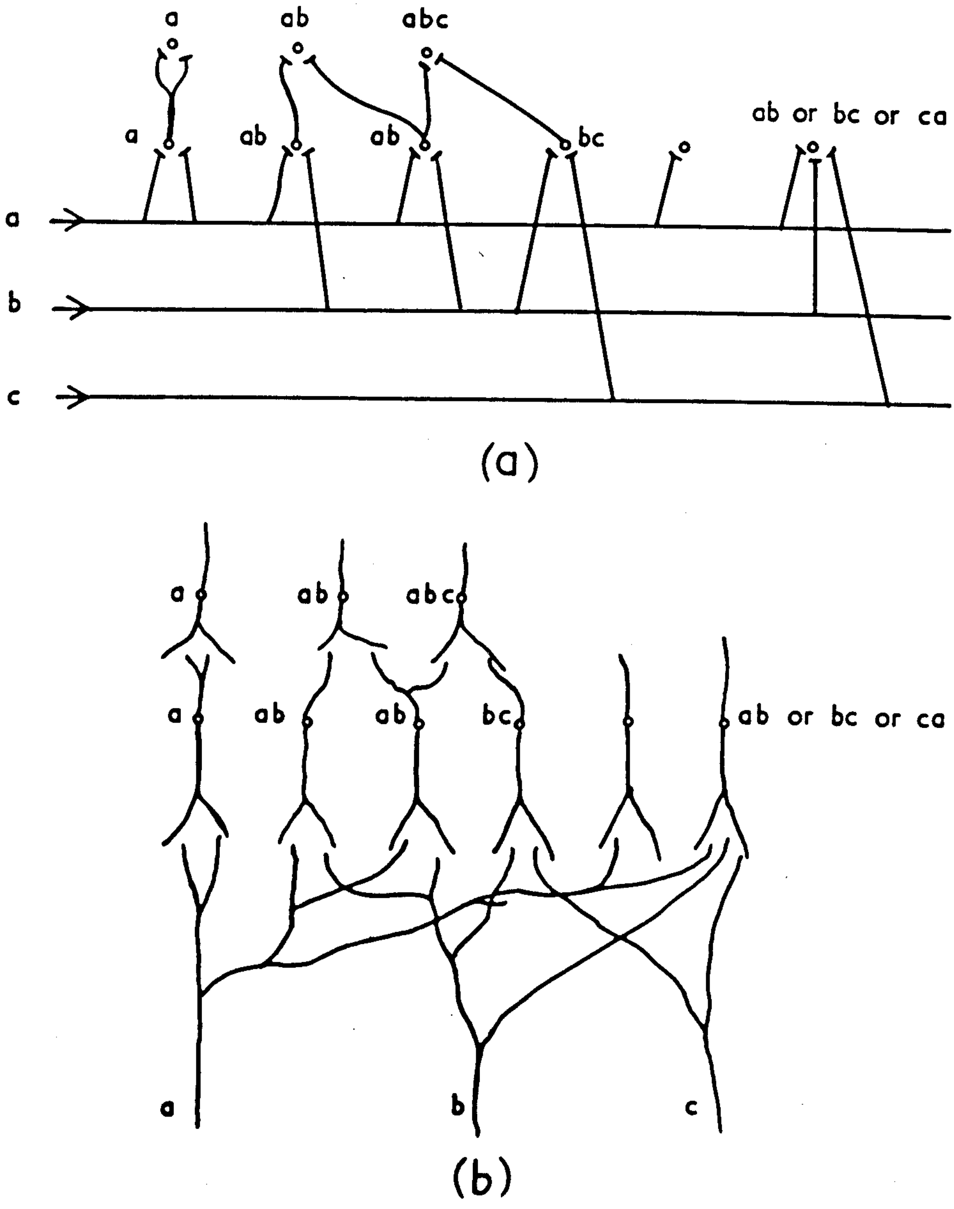

Fig.4. Imperfections in classification due to random connexions; some units never indicate, some do so ambiguously. (a) Formal connexions to units which indicate if two inputs are active (b) Neural realization of these connexions.

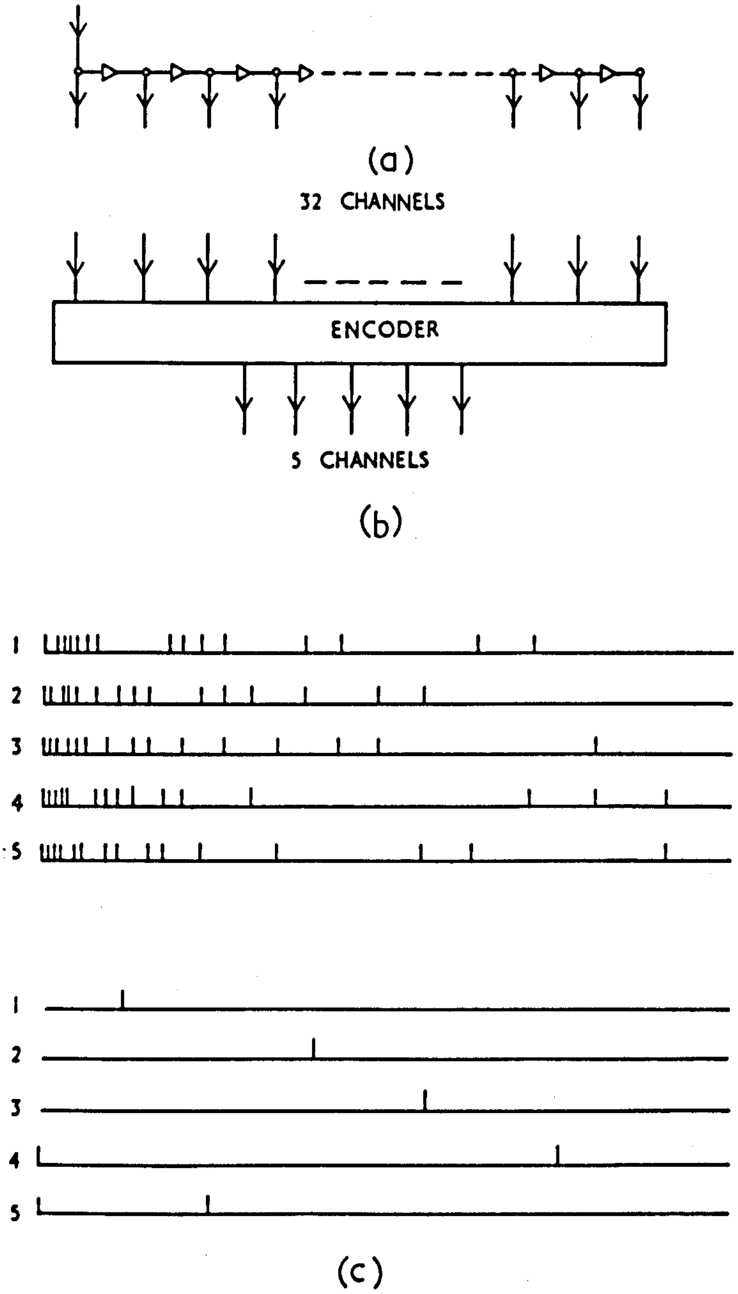

Fig.5. Discrimination of temporal intervals by means of (a) a series of delays followed by (b) an encoder. The output of the combined system, when a single pulse enters, is shown at (c).

receptors. It is suggested that this is one of the primary functions of the neural bodies which are interposed between receptors and brain; in other words they may be said to abstract key features and reduce redundancy. For example, even at the early stage of the retina, there is a hundredfold convergence from receptors to optic nerve. There is a corresponding loss of discrimination in that the information passed on is concerned primarily with contours. In consequence, for man, n is of the order of 10^6, so that the information reaching the brain from all senses is about as much as is contained in black-white television screen.

Secondly, the nervous system does not achieve complete classification of the inputs. If it were so there would be a unit for the Eroica and if it fired one would perceive the whole symphony in that magic instant. But, through the phenomenon of short term memory, one can perceive only the last few seconds' worth. So if r is the number of fibres which fire during the span of short term memory N is of the order of nCr.

Thirdly, there is the phenomenon of span of discrimination whereby a set of unrelated objects can be distinguished, at first glance, only if the number in the set is less than about seven.

Opposing these factors is that of duplication. It has been pointed out earlier that, by random connexions, duplication is likely to occur and that this will enable the system to function even if damaged; but such a gain will be balanced by a loss of discrimination if the total number of neurons is not increased.

To conclude, at the present time there appears to be no evidence to show that the nervous system of any particular animal does not contain enough neurons to match its powers of discriminating patterns of receptor activity.

Delays

It has been shown (1954, *ref.* 13) that for temporal discrimination each binary signal must pass through a series of delays, as in *figure 5a*, and that the output of each delay must provide an input to a classification system.

The human can distinguish temporal intervals between brief stimuli in accordance with Weber's law for intervals up to about three seconds; it will be assumed, for the convenience of binary arithmetic, that one can distinguish 32 such intervals. A complete classification system to distinguish temporal patterns to this extent would require, not 2^n but 2^{32n} units. However, if at any instant only one of the 32 outputs of *figure 5a* is active it is possible to design a system which recodes the information from the 32 channels into 5 channels, each active or inactive: the output from such a coder, shown in *figure 5b*, has 32 possible states. The coding system could be that of binary arithmetic which is shown in column two of the table below.

Number of Delays	Normal Code	Rearranged Code	Total Lapsed Time assuming Weber's Law
0	00000	11111	a
1	00001	01111	ar
2	00010	11101	ar^2
3	00011	10111	ar^3
4	00100	11011	ar^4
5	00101	11110	ar^5
30	11110	00100	ar^{30}
31	11111	00000	ar^{31}

Any fixed arrangement of this code would suffice; let it be according to the number of units in the code number. Such a code is shown in column three; the undelayed signal is represented by five units; the next five discriminable intervals are represented by numbers containing four units and one zero and placed in random order; the last interval produces zero activity in the output of the coder. The fourth column shows the time at which each number should appear at the output of the coder for Weber's law to hold. The data of columns three and four are shown graphically in *figure 5c*; this is the output of the combined delay and coding system when a single pulse occurs at the input. It can be seen that the system behaves exactly like the nearly unstable population of neurons which has been postulated already. Consider all the a neurons suggested by *figure 4e*; when the a afferent becomes active the directly connected a neurons will fire after one synaptic delay; this activity will be passed to a neurons of higher and higher order, their number steadily decreasing, until the reverberation finally ceases. It is suggested therefore that a nearly unstable population of neurons, each of which fires if a critical number of impulses falls on it, not only classifies but also incorporates the necessary system of delays. However this system is not exactly of the form originally proposed and in which delays occur only at the input level. Not only will the clusters of a neurons behave as in *figure 5c*; to every pattern of input activity there will correspond a cluster of neurons which will behave as a reverberating delay system.

The original formal arrangement of delays is shown in *figure 6a* and the suggested neural realisation of it is shown diagrammatically in *figure 6b*. In the latter there is a redundance of delays but it will have a certain advantage; the number of neurons which indicate the pattern *'A one second before B'* will be independent of the complexity of the patterns *A* and *B*; in other words temporal discrimination of the stimuli will be independent of their complexity.

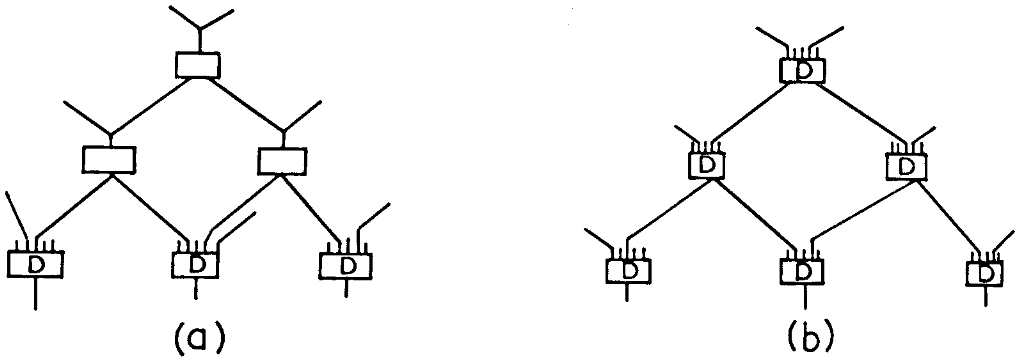

Fig.6. Delays in a classification system may be (a) all at the input or (b) distributed throughout the system.

In making an estimate of the number of neurons required to distinguish spatio-temporal patterns one concludes that it is not necessary to allow one factor for delays and another for duplication; the duplication provides the delays.

CONDITIONAL CERTAINTY SYSTEMS

A conditional probability system is a classification system with an additional function; it has been defined (1956, *ref.* 14) as one in which, when any set of inputs becomes active, the conditional probability is computed of every other possible set of inputs; the conditional probability of each set is contained in the unit which distinguishes that set. The design of such a system has been discussed in a previous paper (1958, *ref.* 15) from which there were the following conclusions.

The system is much easier to construct if each unit indicates only whether the conditional probability of the corresponding set of inputs exceeds some arbitrary threshold. For example, suppose that this threshold is taken as one half and that, in the past, patterns A and B have occurred as follows:-

A 1111111
B 1111111111

If A occurs $p(B/A)$ is 4/7 so B is indicated (inferred); on the other hand if B occurs $p(A/B)$ is 4/10 so A is not inferred. Each unit therefore indicates in two different ways; if the corresponding set of inputs becomes active it *indicates occurrence;* if this set is only inferred it *indicates conditionally.* Such a system is called a *conditional certainty system.*

The first new requirement of the system *is* that each unit shall count, on a common scale, the number of times that it has indicated. The design of the whole system is eased if counting is on an approximately logarithmic scale, and this can be achieved in the following way.

1. When any unit is not indicating occurrence, some physical quantity associated with it grows - preferably in an approximately exponential manner.
2. When a unit indicates occurrence this physical quantity is decreased, preferably by a fairly constant fraction of its present value.
3. To imitate spontaneous recovery in conditioning there must be storage in depth.

The stored quantity is called the rarity of the set of inputs; it is approximately log K/N, where K is a constant and N is the number of occurrences.

If a set of inputs B includes a set A, then the unit which distinguishes set B is called a superunit of the A unit; conversely the A unit is called a *subunit* of the B unit.

There are two new systems of connexions in a conditional certainty system in addition to the counting connexions. There is a connexion from each unit to all its superunits; its function, which is called *supercontrol* is as follows:-

4. A unit which is indicating occurrence causes any superunit of it to indicate conditionally if the rarity stored in the latter does not exceed that stored in the former by more than some critical amount.

There is also a connexion from each unit to all its subunits; its function, which is called *subcontrol* is as follows:-

5. A unit which has been supercontrolled to indicate conditionally causes all of its subunits to indicate conditionally. If the system has more than two inputs there is one further rule.
6. A unit which is counting inhibits supercontrol by its subunits.

The possibility that these six requirements may be met in nervous tissue will now be considered.

Counting in a Neuron

In a classification system the units must possess two states, active and inactive, in order to indicate occurrence. It has been suggested that in the nervous system there correspond neurons which are either firing or not firing and which possess a rapid recovery process. The additional requirements of counting (1 and 2 above) are met by a reversible process in the neuron whereby its state is changed when it fires. An example of such a process is the transfer of sodium ions across neural membrane; here the recovery process (the sodium pump mechanism) seems too rapid but it has been measured so far only in fibres of large diameter which are not associated with learning; for fibres of smaller diameter the recovery may be slower. However, because of the long retention of learned inferences, the slow recovery process is more likely to be a chemical than a physical one.

The fourth and fifth requirements imply that a unit can indicate in quite a different way, conditionally. In doing so it must not count; but, by sub-control, it must affect other units to which it is connected. For the present the indication of occurrence and conditional indication will be both identified with the firing of a neuron; the problem of distinguishing them and the consequence of not doing so will be discussed later.

Connexions for Counting Control, Supercontrol and Subcontrol

For a classification system with three inputs the counting connexions are shown in *figure 7a.* For supercontrol (requirement 4) there are connexions from each unit to all its superunits and for subcontrol (requirement 5) there are connexions from each unit to all its subunits; the two systems of connexions are identical and are shown in *figure 7b,* but the two forms of control are in opposite directions and have different functions.

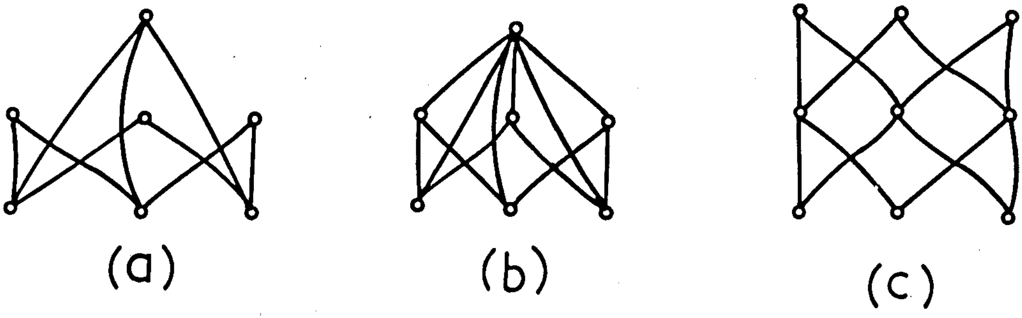

Fig.7. Connexions for (a) counting control, (b) both super-control and subcontrol, and (c) for all three forms of control.

The connexions for supercontrol and subcontrol include some which are required for counting control and it is an important question whether two or even all three forms of control could be effected by the same physical connexions; this has not been achieved in the electrical computer described in the previous paper and which uses three separate connective systems.

However, at the beginning of this paper, it has been shown that if a unit indicates when more than a fixed number of impulses reach it the counting connexions take a different form. If the critical number is two and there are only three inputs the system of counting connexions is as in *figure 7c.* It is interesting that this system of counting via subunits also meets exactly the requirements of supercontrol and subcontrol. Not only are no additional connexions needed but incorrect connexions cannot arise; for example, there can be no connexion between a *c* unit and an *ab* unit. This is true by definition; if a unit has been labelled a *ab* unit, for example, this can be only because either it possesses one connexion to an *a* unit and one to a *b* unit, or it possesses two connexions to another *ab* unit. These conclusions are also true if the critical number of impulses is more than two, for example, ten as in *figure 2b*; but then there will be further features of design which have been discussed elsewhere (1954, *ref. 13*). In the rest of this section the critical number of impulses will be taken as two.

The Function of Connexions

An attempt will now be made to discover the conditions under which a neuron must fire in order to meet all the requirements of a conditional certainty system while at the same time using a single connective system for all three forms of control; only the three input system of *figure 7c* will be considered.

For counting control there is the following rule.

A *Whatever the state of a neuron it fires when it receives two or more simultaneous impulses from other neurons whatever their state*

For supercontrol a firing neuron must be able to fire a neuron to which it has one connexion but only if the two are in approximately the same state. This difference in state could be effective physically only in an area affected by both neurons but unaffected by others; the synapse meets these conditions.

For a neuron which distinguishes the pattern *a*, suppose that *R(a)*, the rarity of *a*, is stored on the *a* side of all its synapses; similarly for any other neuron *b*. If there is a synapse between neurons *a* and *b*, the independently variable quantities *R(a)* and *R(b)* will be stored on either side of it. A possible mechanism which would meet the fourth requirement is this; suppose that the spike and threshold potentials of a neuron depend on its postulated variable state so that they both decrease after the neuron has fired, and then slowly recover. If neuron *a* fires it is then less able to activate neuron *b*; and if neuron *b* fires, neuron *a* is more able to activate neuron *b*. For supercontrol there is therefore the following rule.

B. A neuron fires when it receives one impulse from a firing neuron if the rarity of the former exceeds that of the latter by less than a threshold amount.

A very simple neural model of a conditional probability system is sketched in *figure 8a.* If neurons *a* and *b* fire then the *ab* neuron will also fire, by rule A; this is counting control. If the *a* neuron fires alone, the *ab* neuron will fire only if *R(ab)* exceeds *R(a)* by less than the critical amount; this is supercontrol. Now the conjunction of *a* and *b* cannot occur more often than that of either *a* or *b* alone so *R(ab)*, the state of the *ab* neuron, cannot be less than *R(a)* or *R(b)*; if the *ab* neuron now does fire conditionally then it must fire the *b* neuron by rule B because *R(b)* is less than *R(ab)*; in consequence, the fifth requirement, of subcontrol, has also been met but with one further physiological hypothesis - that *synapses function in both directions.* However, this hypothesis can be avoided if recurrent collateral axons are introduced in neuron *ab* to mediate subcontrol; in this arrangement, shown in *figure 8b,* there is one system of connexions for counting and supercontrol and another for subcontrol.

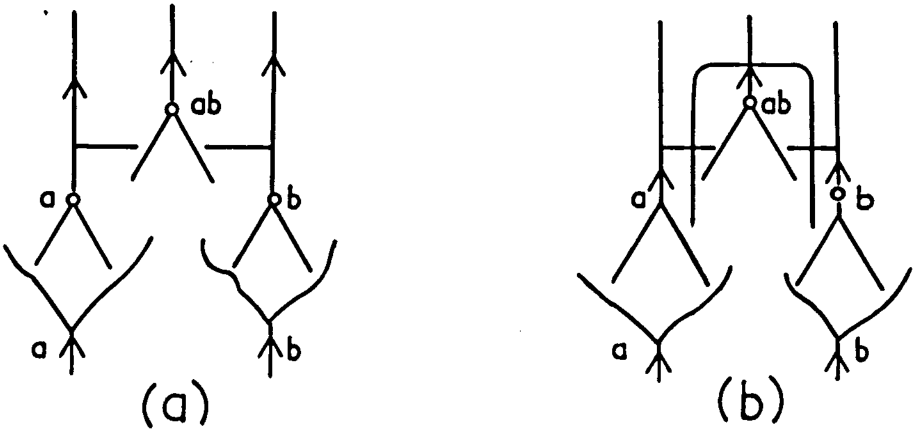

Fig.8. Neural connective system in which (a) each synapse exerts all three forms of control, (b) subcontrol is effected by recurrent collateral axons.

Rule B may be compared with the hypothesis of varying synaptic resistance; according to this theory synaptic resistance is not the same as electrical resistance but it involves the notion of impulses sometimes passing and sometimes not passing over a synapse and of this passage being facilitated by a previous passage of impulses. According to the conditional probability hypothesis the passage of impulses is determined by the difference between two independent variable quantities, the states of the two neurons; but the hypothesis of synaptic resistance refers to only one variable. In consequence, experimental extinction is more easily explained by the former theory as an active process (1956 *ref. 14* p. 5).

Rules A and B and the connective systems discussed do not seem biologically impossible; indeed they have been suggested by Burns (1955, *ref. 4*) together with the requirements 1, 2 and 3, for an entirely different reason. As will be seen later, the sixth requirement is not unreasonable. The major difficulty must now be discussed; this is that a neuron can give a binary indication in two quite different ways in only one of which its state is changed. In *figure 8b* suppose that some event external to the system causes the *a* afferent fibre to become active; this fibre, by means of two connexions, will cause the *a* neuron to fire by rule A. Suppose that the *a* neuron fires the *ab* neuron by rule B so that pattern *ab* is inferred from pattern *a*. Now supercontrol of further neurons should, by requirement four, be exerted only by the *a* neuron and not by the *ab* neuron. There is no way of making this distinction in the arrangement shown.

This difficulty occurs in an acute form when considering connexions from the system to motor units. Consider a classical alimentary conditioned reflex in which there are the three patterns *Sc*, a conditioned stimulus, and *Su*, an unconditioned stimulus, which always produces a response *R*. Suppose that the three patterns are reported to a conditional certainty system as in *figure 9*; the primitive deterministic control path is shown in full lines; broken lines show the channels by which the activity of the lower level system and of an unrelated receptor system are reported to the higher level system. Whether this system is designed according to *figure 8a or 8b* the input points will also be output points and the broken lines must indicate to two-way signalling. Now suppose that conditioning occurs as in the first sequence of the following table, where a unit in brackets indicates an inferred pattern.

S_c	1111	1	1	1	1
S_u	1111	(1)	(1)	(1)	(1)
R	1111	(1)	(1)	(1)	(1)

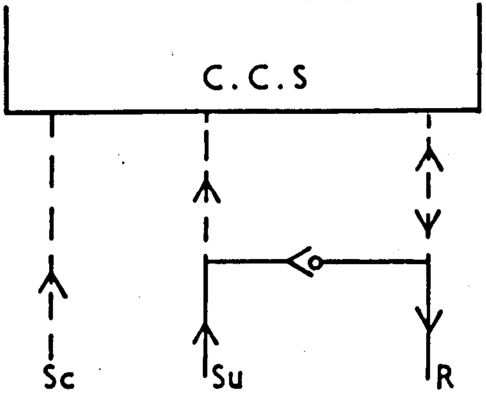

Fig.9. Connexions between receptors and effectors and a Conditional Certainty System which will lead to Classical Conditioning Behaviour.

If S_c occurs alone, as in the second sequence, then S_u and R will be inferred and impulses will pass down the R channel from the conditional probability system to the effector organ in accordance with the facts. But how is the R impulse from the R unit of the conditional probability system to the effector organ to be treated as inferred and not actually counted? If the inferred events, which are shown in brackets in the above table, are counted the probability of R given S_c will remain at unity even though S_c is not reinforced; and this situation will be perpetuated.

The same difficulty occurs in the goal-seeking situation of an instrumental conditioned reflex; this has been discussed in a previous paper (1956, *ref. 14*). In the simplest example there are two situations which arouse the stimulus patterns S_1 and S_2; the animal makes responses which give rise to the proprioceptor pattern P_1 and P_2 in random order. Food which arouses a stimulus pattern G is given for the conjunctions S_1P_1 and S_2P_2 but not for S_1P_2 or S_2P_1. These facts are shown in the following table.

S_1	1		1	
S_2		1		1
G	1	1		
P_1	1			1
P_2		1	1	

Then S_1 will not arouse P_1 rather than P_2 because $p(P_1/S_1)$ and $p(P_2/S_1)$ are both equal to one half. If, however, the pattern G is aroused internally in the system, to represent goal choosing, and S_1 occurs, the P_1 and P_2 units will contain $p(P_1/S_1G)$ and $p(P_2/S_1G)$ which are 1 and 0 respectively; so the pattern will be evoked which has the highest probability. The goal G may now actually occur. How then is the single unit G to be used to indicate at one moment that the pattern G is inferred, imagined, chosen, and at another moment that it is actually occurring? A further important point is that P_1 and P_2 are proprioceptive patterns signals reporting what has happened to joints, tendons, muscles spindles and so on; they are quite different from the efferent patterns of signals sent to muscles. A conditional certainty system which received signals from exteroceptors and proprioceptors but with no input signals reporting the motor patterns emitted could learn what to do but now how to do it.

The problem of connexions from a conditional probability system to effector units cannot be discussed until the problem of the two forms of indication in a unit have been clarified; for this reason the theory developed in the three present papers relates only to the analysis of patterns of afferent signals. The synthesis of efferent signals is the subject of further work.

If the two forms of indication were not distinguished in the neurons of a conditional certainty system there would be important consequences which have been discussed in the previous paper (1958, *ref. 15*). There would, of course, be a failure to distinguish actual and inferred events; but this would occur only in the conditional certainty system; if in other parts of a nervous system neurons were fired only by actual occurrence, then the system as a whole could distinguish actual from inferred events. A further consequence, within the conditional certainty system, would be that a single actual event could set off chains of inferences; if such chains formed closed cycles there would be the possibility of regenerative storage. This would occur if, for example, the past had been as follows:-

A 1 1 1 1 1 0 0 0 0 0

B 0 0 1 1 1 1 1 0 0 0

C 0 0 0 0 1 1 1 1 1 0

D 1 0 0 0 0 0 1 1 1 1

E 1 1 1 0 0 0 0 0 1 1

Then, if *A* occurs, *A* implies *B* implies *C* implies *D* implies *E* implies *A* implies.... An isolated brain does exhibit such continuous activity.

Inhibitory Connexions

The sixth requirement arises in the following way. If a pair of inputs *a* and *c* becomes active then the conditional probability of any input *b* is obtained by subtracting the rarity stored in the *ac* unit, which is counting, from that of its superunit *abc*. But the *(a)* unit is counting too and it must not exert supercontrol, hence the rule. The system of units is shown diagrammatically in *figure 10a*. If *a* and *c* are occurring there will be false supercontrol of the *ab* unit by the *a* unit; this will be prevented if the *ac* unit, in counting, can exert an inhibitory effect on the *ab* unit; similarly for the *bc* unit. In neural terms, the sixth requirement is met, therefore, by the following rule.

C. *A counting neuron inhibits supercontrol by other neurons at the same level.*

The necessary inhibitory links have been indicated in *figure 10a* and shown more realistically as short axon neurons in *figure 10b*. Such neurons occur in large numbers in cerebral cortex and the hypothesis is here made that their axonal systems form inhibitory synapses. This hypothesis has been made to explain inhibitory phenomena in a number of areas, in the cochlear nucleus (Allanson and Whitfield, 1955, *ref. 2*) in the retina and in spinal ganglia (Renshaw, 1946 *ref. 10*).

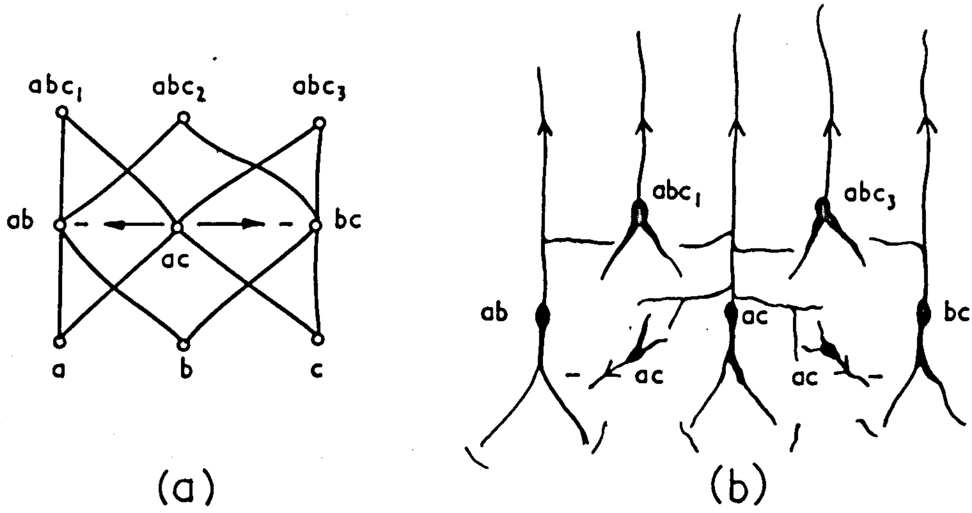

Fig.10. Inhibiting connexions needed in a Conditional Probability Computer.

Storage in Depth

The third requirement has been left until last because it is not essential to the main argument; it is concerned with an additional property of counters which has been discussed in a previous paper (1958, *ref.15*). In the counting unit of *figure 11a* the first two requirements of this section are met and past events are given an approximately exponential weighting; in consequence the weighting of two temporally separated events is scaled down in exactly the same way as they sink into the past. But in the spontaneous recovery of conditioned reflexes recent events are given enhanced weight in comparison with earlier ones although this effect disappears in time; this property is possessed by the counter of *figure 11b* which gives even greater weight to recent events than does an exponential weighing function. Spontaneous recovery has been demonstrated in a conditional certainty system which uses such counters. The circuit contains two time constants and it is necessary that C_3R_3 exceed C_2R_2; to count an event the switch S is moved to the right and then returned; the voltage R drops immediately and then recovers with the time constant C_2R_2 to the voltage on the condenser C_3. This condenser is hardly affected by a single count - it stores the average rarity over the long period C_3R_3. In a conditional certainty system which uses such a unit inferences (conditional probabilities) can be modified rapidly but there will be a reversion to those inferences which are based on long maintained consistencies in external events.

Spontaneous recovery can also be demonstrated with the circuit of *figure 11c*. A secondary battery B is charged to count and discharged slowly to recover; short time storage occurs in the surface of the electrodes, long term storage occurs in deeper layers. It is suggested that this storage in depth might also occur in synaptic material lying between the membranes of two neurons; the two surfaces so formed could store the rarities of the patterns distinguished by the corresponding neurons, each backed up by storage in depth.

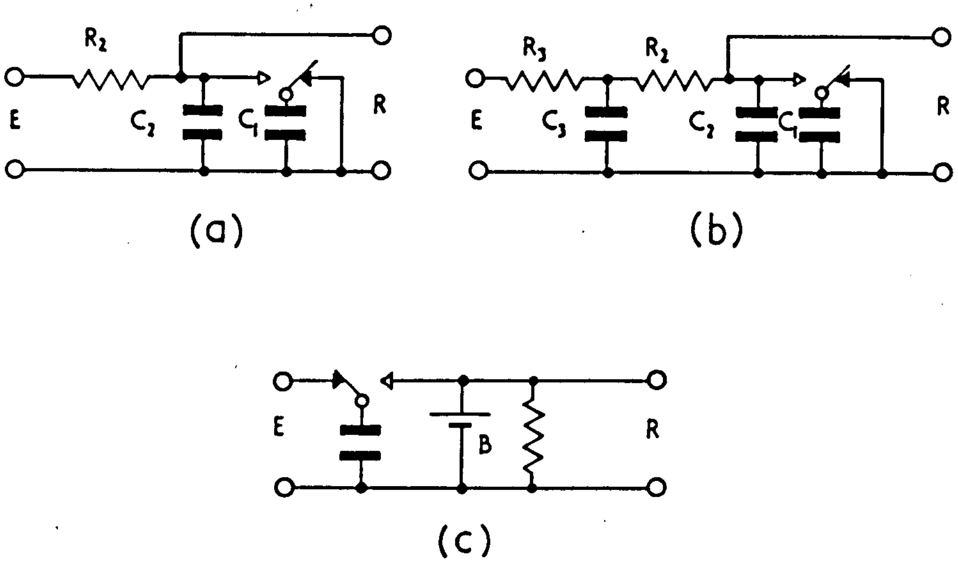

Fig.11. Counters whose recovery processes possess
(a) One time constant
(b) Two time constants
(c) A distributed time constant
The last two imitate the spontaneous recovery of conditioned reflexes.

VARIATIONS OF DESIGN

COUNTING

If there were neurons which met the second requirement of counting but not the first - recovery process, then the system would be capable of learning to make inferences but not of modifying or forgetting them; such a property would be akin to imprinting.

Regenerative storage systems used in digital computers, exhibit no slow recovery processes; however information stored by such methods can be destroyed suddenly and completely if the regenerative loop is broken. If regenerative storage methods were employed in a nervous system, for example in the form of reverberating chains of neurons, then inferences would be made slowly and then either remain fully active, as in imprinting, or be suddenly destroyed.

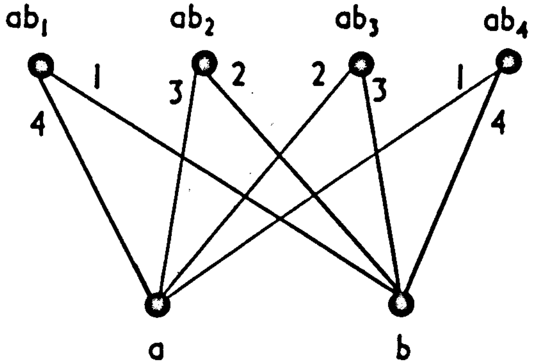

Fig.12. Conditional probability computed in terms of the number of active indicating units.

Conditional Probability Systems

It has been stated at the beginning of this paper that a conditional probability system is more difficult to design than a conditional certainty system; no such computer has yet been built. Yet a salivary reflex, for example, does not suddenly appear; the quantity of saliva secreted increases continuously as the conditional probability of food increases. This suggests that the nervous system does succeed in computing the probability of inferences in a continuous rather than a binary way; two possible mechanisms will be suggested.

First suppose that the critical number of impulses to fire a neuron is more than two, say five; then any neuron, the sum of whose connexions to an *a* and a *b* neuron is five, will be an *ab* neuron; this is shown in *figure 12* in which the different numbers of connexions are indicated. Rule A, suitably modified, implies that, if there have been no conjunctions of *a* and *b*, none of the *ab* neurons will fire if the *a* neuron fires alone; for example, the ab_1 neuron will not fire because it receives only four impulses. But after a sufficient number of conjunctions, $R(ab)$ exceeds $R(a)$ by less than the threshold value so, by Rule B, the *a* neuron fires the ab_4 neuron to which it has one connexion; even more easily it fires the other three *ab* neurons since it possesses more than one connexion to each of them. This suggests that at intermediate stages of relative rarity the *a* neuron might fire neurons to which it has two, three or four connexions. If this can occur, the number of impulses (the excitation) arriving at *B* by subcontrol will vary according to the number of *ab* neurons fired by *a*, and hence according to the value of *R(ab) - R(a)*: this is shown in the following table.

R(ab) - R(a) on an arbitrary scale	*ab* neurons fired by *a* neuron	Number of impulses arriving at *b* neuron
5	none	none
4	ab_1 only	1
3	ab_1, ab_2	3
2	ab_1, ab_2, ab_3	6
1	ab_1, ab_2, ab_3 ab_4	10

With the above mechanism the system succeeds, by digital methods, in transferring to the *b* neuron, a *measure* of conditional probability and it solves a problem found difficult by the computer designer. Rules A and B, modified as below, are the criteria for this principle to occur in the nervous system.

(A') *Whatever the state of a neuron, it fires if it receives a critical excitation from firing neurons whatever their state.*

(B') *The excitation contributed by a firing neuron increases as the rarities of the firing and excited neurons approach one another.*

A second possibility has been suggested by Russell (1955). In an electrical conditional probability computer the practical difficulty is that of transferring the difference between two voltages to many parts of the system. It is much easier to transfer a delay. A conditional probability computer has been constructed in which probabilities are computed as delays; the rule for supercontrol (Rule 4 of this section) is as follows:-

4a. *A unit which is indicating occurrence causes any superunit of it to indicate conditionally with a delay proportional to the difference in rarity of the two units.*

For the neural model rule B is modified similarly. Such a delay can be transferred easily from one part of the system to another; it will be approximately proportional to the negative logarithm of conditional probability so that a delay *t* will refer to a probability *exp(-t)*. It is suggestive of this idea that in an alimentary conditioned reflex the delay in response decreases as the probability of food increases. The design of computers using this principle has been described by Russell *(ref.11)*.

Equivalence

An instrumental conditioned reflex is extinguished if the correct response is not rewarded. It has been shown (1956, *ref. 14*) that this property arises only if inference is redefined symmetrically as follows:-

A implies B if p(A/B) and p(B/A) exceed a threshold value

It has been shown (1958, *ref. 15*) that for a system to infer in this way the fifth requirement, of subcontrol, must be re-stated as follows:-

A unit which has been supercontrolled to indicate conditionally causes any subunit of it to indicate conditionally only if the rarity stored in the former does not exceed that stored in the latter by more than some critical amount.

In consequence rule B takes the following form.

B A neuron fires when it receives one impulse from a firing neuron if the rarities of the two neurons differ by less than a threshold amount.

CRITICISMS OF THE THEORY

A number of important criticisms have been made of the theory.

Barlow (private communication) has referred to the property demanded by the theory that the threshold pontential of a neuron should decrease after firing (p.16 *l.* 22) and so have an increased sensitivity. He has pointed out that this is the opposite of adaptation. Burns (private communication) has shown that it is, in fact, adaptation which occurs in cortical neurons in the isolated brain of cat.

If this reversal in the property of the unit is made the system computes, not probabilities, but rarities; it produces increased outputs as probabilities decrease and should be called a *Conditional Rarity System.* (C.R.S). In trial-and-error learning the system must therefore have an inhibitory function. The all-important Rule B' now becomes

The excitation contributed by firing neuron decreases as the rarities (states) of the firing and excited neurons approach one another.

This rule is met in physiological terms by two neural properties.

1. After a neuron has fired its sensitivity to incoming excitation is reduced. This is adaptation.
2. After a neuron has fired its potential excitation of connected neurons is increased. This is in agreement with the findings of Eccles and others. This is postactivation potentiation (*Eccles, ref. 7*).

Such a change of design is in line with Barlow's demand that there should be a decreased output from units which distinguish patterns of high probability and, conversely, increased outputs for unexpected patterns. In its changed form the system reduces the redundancy in the output channel compared with that of the input channel.

The combination of a C.R.S. with over-connectivity between units appears to resolve the dilemma of the number of units required.

Consider the extreme right-hand unit of *figure 4a* which possesses more input connexions than the two which are necessary for classification.

Suppose that, in the past *a*, *b* and *c* occurred as follows:-

a	111
b	1111
c	1

Then the *a*, *b* and *c* units (which are not shown) will contain 3, 4 and 1 while the ambiguous unit will contain 4. Now, being a C.R.S., the output from the ambiguous unit will be zero when *b* occurs alone. This is reasonable since *b* provides no information as to whether *ab* or *bc* are occurring. The link from *b* to the unit is therefore broken and ambiguity disappears. If now *ab* occurs the unit signals -log(3/4); if the rarer event *bc* occurs it signals the larger quantity -log(1/4).

One is led to the concept of an over-connected system with a reasonable number of units and in which confusion occurs initially. Learning consists in removing the connexions which carry low information; this resolves ambiguities so that the units eventually distinguish uniquely those events which are rare and carry high information.

Watson (private communication) has criticised the demand for symmetrical inference in trial-and-error learning. He considers that there is experimental evidence to show that for a goal *G* an animal chooses a reaction *R* for which the inverse probability $p(G/R)$ is a maximum regardless of the value of $p(R/G)$. This, in fact, effects a simplification in design.

The neural properties demanded by the system in its changed form will now be summarized. Referring to *figure 8a* and *b*, conditions 1, 2 and 3, (that rarities shall be stored in the units) remain unchanged. Because of the change in Rule B the excitation of the *ab* unit by the active *a* unit is to be proportional to the difference in their states. The output from the *ab* unit and, in general, from all affected superunits thus form a channel of reduced redundancy. Also, the excitation of the *b* unit by the *ab* unit depends only on the difference in their rarity. In consequence, the *b* unit and, in general, all affected subunits form a channel which signals the inverse rarities demanded by Watson.

ACKNOWLEDGEMENTS

Thanks are due to J. A. V. Bates, H. B. Barlow, G. Russell, W. H. A. Rushton and D. A. Sholl for many discussions and suggestions.

This paper is published by permission of the Director of the National Physical Laboratory, Teddington, Middlesex.

REFERENCES

1. ALLANSON, J. T. Some properties of a randomly connected network. Third London Symposium on Information Theory, 1955. London: *Methuen* (1956).
2. ALLANSON, J. T. & WHITFIELD, I. C. The cochlear nucleus and its relation to theories of hearing. Third London Symposium on Information Theory, 1955. London: *Methuen* (1956).
3. BEURLE, R. L. Properties of a mass of cells capable of re-generating pulses. *Roy. Soc. Phil. Trans. B.*; 1956, **240**, No. 699, 55.
4. BURNS, B. D. The mechanism of after-bursts in cerebral cortex. *J. Physiol.*; 1955, **127**, 168.
5. BURNS, B. D. Some properties of isolated cerebral cortex. *J. Physiol.*; 1951, **112**, 156.
6. EAYRS, J. T. The cerebral cortex of normal and hypothyroid rats Acta. *Anat.*, 1955, **25**, 160.
7. ECCLES, J. C. The Physiology of nerve cells. *Oxford University Press*, (1957).
8. GALAMBOS, R. & DAVIS, H. The response of single auditory nerve fibres to acoustic stimulation. *J. Neurophysiol.*; 1943, **6**, 39.
9. HILALI, S. & WHITFIELD, I. C. Responses of the trapezoid body to acoustic stimulation with pure tones. *J. Physiol.*, 1953, **122**, 158.
10. RENSHAW, B. Central effects of centripetal impulses in axons of spinal ventral roots. *J. Neurophysiol.*; 1946, **9**, 191.
11. RUSSELL, G. Learning-machines and adaptive control mechanisms 1957. R.R.E. Memo. No. 1369.
12. SHOLL, D. A. Dendritic organisation in the neurons of the visual and motor cortices of the cat. *J. Anat.*; 1953, **87**, 387.
13. UTTLEY, A. M. The classification of signals in the nervous system. *E.E.G. Clin. Neurophysiol.*; 1954, **6**, 479.
14. UTTLEY, A. M. Conditional probability as a principle of learning. International Congress on Cybernetics Namur, 1956. *Publisher Gauthier-Villars, Paris.*
15. UTTLEY, A. M. The design of conditional probability computers. *Information and Control*. U.S.A., 1958. In the Press.
16. UTTLEY, A. M. The probability of neural connexions. *Proc. Roy. Soc.* (B), 1955, **144**, 229.
17. UTTLEY, A. M. The stability of a population of neurons. 1955. Unpublished work.

DISCUSSION ON THE PAPER BY DR. A. M. UTTLEY

DR. H. B. BARLOW: I am not sure that I have fully absorbed the implications of the modified scheme outlined at the end of Dr. Uttley's paper, but I think this modification brings his ideas and mine much closer together, and it may well answer some of the points I am going to raise.

One thing that worried me in the original conditional certainty scheme indicated in *fig. 8* was this: if it infers that b occurs when a occurs because $P(ab)$ is nearly equal to $P(a)$, then there is no way of distinguishing the rare event a from the common event ab. This does not seem biologically desirable, for rare events are often supremely important. The cat surely keeps a watchful eye for the rare mouse which fails to scuttle to its hole. Does the modified machine detect such rare events?

There are two other points I would like to know more about. Biologists tend to be intimidated by "requirements" proposed by mathematicians, physicists, or engineers, and therefore tend to accept them without argument. I do not, for instance, question the physicist's "requirement" that a light-sensitive receptor cell cannot absorb a fraction of a quantum. I don't feel that you want your "requirements" to be understood as absolute in that sense, but it might help us to understand them correctly if you were to indicate how much they depend upon your particular model, or how far they follow from general properties of machines using conditional probabilities.

The last point is on what Dr. Uttley describes (p.128) as the "intractable problem" of classifying 2^n possible inputs when n is of the order 10^6. There are only about 2^{30} neurones available, so the ratio of cells you have to cells you need is $2^{-999,970}$; in other words, you have virtually none of the cells you need. Of course our capacity to discriminate patterns may be very bad compared with a complete classification system, but I'm sure it's not as bad as that figure suggests.

The conclusion I would like to draw from this absurd ratio is that the nervous system cannot afford to leave to chance the pattern of effective connections a cell makes, and probably cannot allow it to be rigidly predetermined genetically because the environment is not rigidly predetermined. On the contrary, neurones must be allotted to input patterns according to a scheme which achieves the utmost economy. Does Dr. Uttley's machine, either in its original or in its modified form, achieve this economy?

DR. F. ROSENBLATT: Dr. Barlow has really stolen half of my thunder here with his last comment. However, I would like to congratulate Dr. Uttley for being the first speaker on the programme to propose a theory rigorous and concrete enough in its references to actual physical or biological systems to permit me to take exception to it.

The main point of exception I would like to take is the quantitative one which I believe Dr. Uttley himself has recognised. He has proposed that it might be possible to replace the two to the Nth units which seem to be required for complete coverage of the sensory field by the number of combinations of N things taken M at a time where M is the number of units actually activated or illuminated by a particular stimulus. Or if we use some of the techniques which Dr. Taylor has recommended (*ref. 1*), we might even reduce M to the number of elements necessary to represent a contour which is still smaller. I still think it is perfectly clear that such techniques alone are really not sufficient to adequately reduce the number of units to anything like a biologically plausible system.

On the other hand, this does not mean we are required to reject Dr. Uttley's proposition that the brain is essentially a classifying system in some sense. I think this is a point which does require some clarification, however. There is an error in his assertion that in order to have a classifying system it is necessary to have a single unique element to represent each of the sensory events or patterns that his system may classify. In other words, in Dr. Uttley's system, if we have a hundred billion possible images of a cat on a retina this would require a hundred billion neurons representing these cats. However, a system of only a million neurons could easily represent one hundred billion events by assuming a hundred billion different states; that is the number of possible states is very much greater than the number of elements in the system. It is indeed possible to devise systems capable of classifying many more events than the number of elements in the system. Admittedly the classification is not perfect in these cases, but (as shown by the systems considered in my own paper) it can be made surprisingly reliable on a statistical basis.

I think it might be profitable for Dr. Uttley to reconsider his model

REFERENCE

1. TAYLOR, W. K. Electrical simulation of some Nervous System Functional Activities in Information Theory: Third London Symposium, Ed. by Colin Cherry. *Butterworths Scientific Publications, London* (1956).

in terms of representation of individual events by states of the system rather than by individual elements.

The suggestion that it is necessary to have a unit's delay of each of the impulses coming into the system in order to get recognition of time-patterns, I think is subject to the same sort of error in reasoning. It is not necessary to have a single unit to represent each possible sequence of events. That is, it is not necessary to have a unique unit to represent each entire sequence of configurations of impulses at times 1, 2, 3 and so on. It is really sufficient that the state of the system at time t should be some contingent function of the state of the system at time $t-1$, which is some contingent function of the state at $t-2$, and so on, and we can show that this accomplishes fantastic economies in the number of elements required. I think such systems will indeed prove to be economical.

I think it is out of place in the programme now to go into this in any detail. These comments really apply not only to Dr. Uttley but much more generally to a great many of the people working in this field. Dr. Uttley has fortunately presented us with a rigorous enough theory so that we can be specific. On the other hand, I think this kind of problem (the problem of "neuroeconomy") rears its head over and over again and must really be dealt with by the theorist if the theory is to prove more useful than a mere conjecture. I would like to suggest a methodological postulate (this really comes very close to what Dr. McCarthy would call commonsense) that wherever methods of gaining economy become essential in order to demonstrate that a proposed perceiving mechanism remains practical in a complex environment (not merely the environment of 2 or 3 impulses, but the environment of cats, dogs, and the things it must ultimately recognise) wherever this is the case, I think that a theory of the mechanism can be considered complete only if some specific means of economy are provided and worked out mathematically. Really I am developing a plea for something beyond heuristics here. I think we are all rather enraptured with heuristics at this point. On the other hand I think there is still some hope that rigorous mathematics may have some value. There are still one or two things that can be accomplished with it. I think its main place in this field is the study of the economy problem. I think many of the qualitative behaviours of our models can be discovered by heuristics, but I think these problems of the neuron economy almost always require a very careful mathematical evaluation before a theory can be acceptable.

DR. A. M. UTTLEY (in reply): Regarding Dr. Barlow's first point, the event a cannot be rarer than the event a and b. On the other hand the event a and not b can be rarer than a and b. This distinction cannot be made in a system based on unitary classification, i.e. which counts only the presence of properties. For this to be done there must be binary classification with the operator *not*. All this is unrelated to the scale upon which probabilities are stored, i.e. upon whether the output at any unit is large for high probabilities, as in my original scheme, or large for low probabilities as in my modified proposal. The latter is very close to Dr. Barlow's proposal and both have the biologically useful property of giving large signals for rare events i.e. when there is a high information content in the incoming message.

On the second point, my "requirements" are necessary deductions from the original specification that conditional probabilities are to be computed.

On Dr. Barlow's last point, which is also Dr. Rosenblatt's first I consider that my modified proposal of a conditional rarity computer together with vast over-connection does achieve the necessary economy of units - with chance connections. Because of the over-connection there is, initially, ambiguity of classification, i.e. failure to distinguish, to a high degree. As information is accumulated those connections which carry little information become less effective until they are disconnected. Ambiguity is thus eliminated and the system learns to discriminate. I suggest that such a system meets the physiological and psychological facts very well.

Dr. Rosenblatt's second criticism is invalid. True, there can be 10^9 different images on a retina of 10^6receptors; the retina alone can *represent* them all. But it cannot *distinguish* them all; for this we should definitely need 10^9 output channels, the first to indicate Cat. No.1, the second Cat. No.2, and so on. To distinguish all the patterns one would need a complete Boolean lattice with 2^{10^6} units. The vital point is that we know we are not seeking to discover such a system in the brain. To sum up, classification, with computation of conditional rarity, with over-connection, does seem to meet the requirements.