

at the same rate as its α analogue⁸. Such loss could be caused by tissue enzymes distinct from those identified in plasma or, as suggested by Biron, to storage at receptor sites. The destruction of β aspartyl-angiotensin II reported by Biron *et al.*⁸ had been previously investigated with conflicting results¹⁰. Edetic acid and dimercaprol are potent inhibitors of angiotensin destroying enzymes¹¹. Addition of these compounds to the perfusing medium resulted in significant reduction in the loss of angiotensin II suggesting a dominant if not exclusive role for tissue peptidases.

The techniques that we have used provide a simple and reproducible means of examining organ destruction of angiotensin in a variety of situations. Preliminary work¹² has indicated that the rate of angiotensin disappearance can be altered profoundly by changes in sodium balance or blood pressure in albino rats.

We thank Sir G. Pickering and Sir H. Krebs for making facilities available and Drs H. F. Woods,

A. J. Honour and D. Hems for technical instruction and advice. W. P. L. is a Nuffield Dominions Fellow and is also grateful to the CSIR, Pretoria, for their support.

Received April 11, 1969.

¹ Hems, R., Ross, B. D., Berry, M. N., and Krebs, H. A., *Biochem. J.*, **101**, 284 (1966).

² Nishitsutsuji-Uwo, J. M., Ross, B. D., and Krebs, H. A., *Biochem. J.*, **103**, 852 (1967).

³ Peart, W. S., *Biochem. J.*, **59**, 300 (1955).

⁴ Lee, M. R., thesis, Oxford University (1965).

⁵ Boucher, R., Biron, P., and Genest, J., *Canad. J. Biochem. Physiol.*, **39**, 581 (1961).

⁶ Regoli, D., and Vane, J. R., *J. Physiol.*, **183**, 513 (1966).

⁷ Goffinet, J. A., and Mulrow, P. J., *Clin. Res.*, **11**, 408 (1965).

⁸ Biron, P., Meyer, P., and Panisset, J.-C., *Canad. J. Physiol. Pharmacol.*, **46**, 175 (1968).

⁹ Hodge, R. L., Ng, K. K. F., and Vane, J. R., *Nature*, **215**, 138 (1967).

¹⁰ Wernze, H., and Fujii, J., *Zeit. Ges. Exp. Med.*, **140**, 128 (1966).

¹¹ Ryan, J., thesis, Oxford University (1967).

¹² Leary, W. P., Ledingham, J. G., and Woods, H. F., *Abstract, Proceedings of the Fourth International Congress of Nephrology, Stockholm* (in the press).

Non-Holographic Associative Memory

by

D. J. WILLSHAW
O. P. BUNEMAN
H. C. LONGUET-HIGGINS

Department of Machine Intelligence
and Perception,
University of Edinburgh

The features of a hologram that commend it as a model of associative memory can be improved on by other devices.

THE remarkable properties of the hologram as an information store have led some people^{1,2} to wonder whether the memory may not work on holographic principles. There are, however, certain difficulties with this hypothesis if the holographic analogy is pressed too far; how could the brain Fourier-analyse the incoming signals with sufficient accuracy, and how could it improve on the rather feeble signal-to-noise ratio³ of the reconstructed signals? Our purpose here is to show that the most desirable features of holography are manifested by another type of associative memory, which might well have been evolved by the brain. A mathematical investigation of this non-holographic memory shows that in optimal conditions it has a capacity which is not far from the maximum permitted by information theory.

Our point of departure is Gabor's observation^{4,5} that any physical system which can correlate (or for that matter convolve) pairs of patterns can mimic the performance of a Fourier holograph. Such a system, which could be set up in any school physics laboratory, is shown in Fig. 1. The apparatus is designed for making "correlograms" between pairs of pinhole patterns, and then using the correlogram and one of the patterns for reconstructing

its partner. One of the pinhole patterns is mounted at A , and the other at B . The distance between them equals f , the focal length of the lens L . A viewing screen is placed at C , at a distance f from the lens, and a diffuse light source is mounted behind A . The pattern of bright dots appearing at C is the correlogram between the pattern at A and the pattern at B . Formally, $C = \bar{A} * B$, where the asterisk stands for convolution and \bar{A} is the result of rotating the pattern A through half a turn round the optical axis. If A and B were interchanged, the pattern at C would be $\bar{B} * A = A * \bar{B} = \bar{C}$, so that the correlogram would be inverted. This is clear enough if B is a pinhole, and shows that the order of the patterns is important.

To recover pattern A from pattern B we convert the correlogram into a pattern of pinholes in a black card and place the light source behind it, so that the light shines through C and B on to a viewing screen at A (Fig. 2). A pattern of spots now appears on the viewing screen. All the spots of the original pattern A are present, but a number of spurious spots as well. If the pinholes were infinitesimal and there were no diffraction effects the reconstructed pattern would be $\bar{C} * B = A * \bar{B} * B$, just as in Fourier holography. If B were a random pattern, one could argue, $\bar{B} * B$ would approximate to a delta function at the origin, so that the reconstructed pattern would look like a slightly bespattered version of the original pattern A . How can we pick out the genuine spots from the others?

To solve this problem let us simplify the set-up by removing the lens (Fig. 3). Suppose, for example, that A has two holes and B has three. Then the pattern C will consist of six bright spots (barring coincidences). When these spots are converted into pinholes and illuminated from the right, a total of 18 ($= 6 \times 3$) rays will emerge from B and impinge on the screen at A . But we shall not see eighteen spots on this screen, because six of the rays will converge, in sets of three, on to the two points

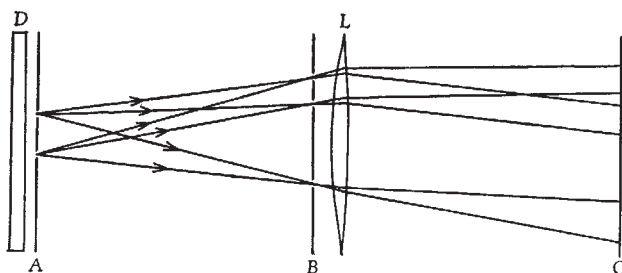


Fig. 1. Constructing a correlogram. D is a diffuse light source, L a lens and C the plane of the correlogram of A with B .

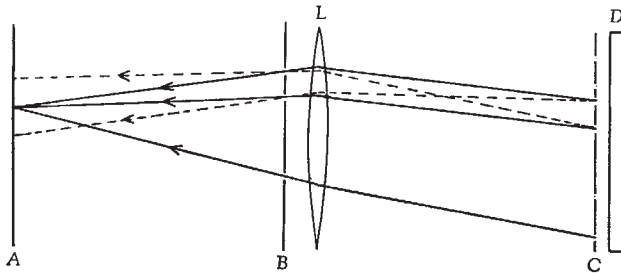


Fig. 2. Reconstructing a pattern. —, Paths traversed in Fig. 1; - - -, paths not traversed in Fig. 1.

of the original pattern. The other twelve rays will give rise to spurious spots, but (again barring coincidences) these spots will be fainter than the genuine ones. We can therefore expect to be able to pick out the wheat from the chaff with a detector with a threshold slightly less than three units of brightness.

This reasoning applies equally to the "correlograph", with lens, illustrated in Figs. 1 and 2. So, having found how to get rid of the unwanted background in reconstructing *A* from *B* and *C*, we can now envisage the possibility of constructing multiple correlograms, comprising all the spots present in $C_1 = \bar{A}_1 * B_1$ or in $C_2 = \bar{A}_2 * B_2$, and so on. The presentation of *B*₁ should evoke *A*₁, presentation of *B*₂ should evoke *A*₂, and so on, up to the limit set by the information capacity of the system. But what is this limit?

To answer this question let us evade the complicated (and basically irrelevant) issues raised by the finite wavelength of light, edge effects and so on, and pose the question in terms of a discrete, and slightly more abstract, model. We suppose *A*, *B* and *C* to be discrete spaces, each containing *N* points, *a*₁ to *a*_{*N*}, *b*₁ to *b*_{*N*}, and *c*₁ to *c*_{*N*}. The point-pair (*a*_{*i*}, *b*_{*j*}) is mapped on to the point *c*_{*k*} if *i* - *j* = *k* or *k* - *N*. Conversely, the point-pair (*c*_{*k*}, *b*_{*j*}) is mapped on to *a*_{*i*} if the same condition is met. Imagine now that we have *R* pairs of patterns which we wish to associate together, each pair consisting of *M* points selected from *A* and another *M* selected from *B*. The total number of point-pairs determined by all the pairs of patterns will be *RN*², and we may think of this number of "rays" striking *C*. If they impinge at random, the probability of any point *c*_{*k*} not being struck will be

$$\exp(-RM^2/N) = 1 - p, \text{ say}$$

The correlogram for the whole set of *R* pairs will then consist of the remaining *pN* points of *C*.

Now consider the reconstruction process. One of the *B*-patterns, comprising *M* of the points *b*₁ to *b*_{*N*}, is selected, and combined with the correlogram to produce *pNM* "rays" impinging on *A*. Each point of the original *A*-pattern will receive exactly *M* rays, so that we should set the threshold of our detector at *M* if we want to pick up all the original points. Now consider any one of the *N* - *M* other points in *A*. It may receive a ray through any one of the *M* "holes" in *B*; the probability that it

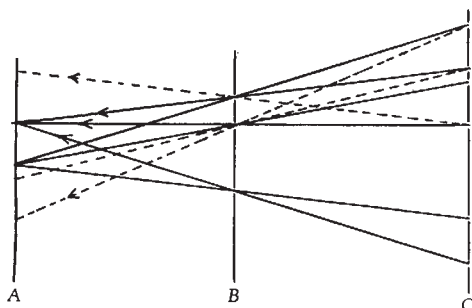


Fig. 3. Showing that original spots are generally brighter.

receives a ray through a given hole is just *p*, for this is the chance that the point on *C* "behind" the hole belongs to the correlogram. The chance of an unwanted point reaching the threshold is thus *p*^{*M*}, and the probable number of spurious points of brightness *M* is consequently (*N* - *M*)*p*^{*M*}. If *M* is a fairly large number, this will be a sensitive function of *p*, and for given *N* and *M* the critical value of *p* above which spurious points begin to appear may be found from the relation

$$(N - M)p^M = 1$$

Alternatively, this may be viewed as a relation which sets a lower limit to the value of *M* for given values of *N* and *p*. A slightly safer estimate is given by

$$Np^M = 1, \text{ or } M = -\log N / \log p$$

If *M* falls below this value, the reconstruction will be marred by spurious points.

Next we enquire about the amount of information stored in the memory when *R* pairs have been memorized and *M* satisfies the aforementioned condition for accurate retrieval. We can evoke any one of *R* *A*-patterns by presenting the appropriate *B*-pattern. There are $\binom{N}{M}$ possible *A*-patterns altogether, so the amount of information needed to store any one of them is $\log \binom{N}{M}$, which is roughly *M* log *N* natural units of information. The total amount of information stored is, therefore, approximately

$$I = RM \log N \text{ natural units}$$

But according to our original calculation of *p*

$$RM^2 = -N \log(1 - p)$$

and if we are working at the limit of accurate retrieval

$$M = -\log N / \log p \simeq \log_2 N \text{ (see below)}$$

It follows immediately that

$$I = N \log p \log(1 - p)$$

As one might have anticipated, this expression has its maximum value when *p* is 0.5—when the correlogram occupies about half of *C*.

What is remarkable is the size of *I*_{max}.

*I*_{max} = *N*(log 2)² natural units = *N* log 2 bits. The maximum amount of information that could possibly be stored in *C* is *N* bits. So the correlogram, in this discrete realization, stores its information nearly (log_e2 = 69 per cent) as densely as a random access store with no associative capability.

As described, the discrete correlograph, like the holograph, will "recognize" displaced patterns. If an *A*-pattern {*a*_{*i*}} and a *B*-pattern {*b*_{*j*}} have been associated, then presentation of the displaced *B*-pattern {*b*_{*j*+*a*}} will evoke the displaced *A*-pattern {*a*_{*i*+*a*}}.

But the resemblance does not cease there. Just as in holography, the information to be stored is laid down (i) in parallel, (ii) non-locally and (iii) in such a way that it can survive local damage. In parallel, because each mapping (*a*_{*i*}, *b*_{*j*}) → *c*_{*k*} can be effected without reference to any other; the same applies to the reconstructive mappings (*c*_{*k*}, *b*_{*j*}) → *a*_{*i*}. Non-locally, because the presence of *a*_{*i*} in an *A*-pattern is registered at *M* separate points on the correlogram, one for each point of the *B*-pattern. And robustly, because if the system is not stretched to its theoretical limit it can (as we shall show elsewhere) be used for the accurate reconstruction of *A*-patterns even when some of the correlogram is "ablated" and/or the *B*-patterns are inaccurately presented. But it can only be made secure against such contingencies by sacrificing storage capacity—as one would expect.

In our discussion of the process of reconstruction we had occasion to note that a point *c*_{*k*} might owe its presence on the correlogram to the joint occurrence of (*a*_{*i*}, *b*_{*j*}); but that if a pattern were presented containing the point *b*_{*j*+*a*}, the "ray" (*c*_{*k*}, *b*_{*j*+*a*}) would light up the point *a*_{*i*+*a*}, which might never have occurred in any *A*-pattern. It was

this feature which underlay the ability of the system to recognize displaced patterns; but the same feature is a slight embarrassment when one comes to consider how a discrete correlograph, with the reconstructive facility, could be realized in neural tissue. We will not dwell on this point, except to acknowledge that it was drawn to our attention by Dr F. H. C. Crick, to whom H. C. L.-H. is indebted for provocative comments. But it led us on to a further refinement of our model, in which a given point c_k is admitted to the correlogram only if the particular pair (a_i, b_j) occurs in one of the pairs of patterns, and not otherwise. On this assumption there might be as many as N^2 separate point-pairs to take into account, and a correspondingly large number of points in the space C .

In this form our associative memory model ceases to be a correlograph, having lost the ability to recognize displaced patterns, but its information capacity is now potentially far greater than before. To show this, we will adopt a rather different type of representation, in which the points of A become N_A parallel lines, and those of B become N_B parallel lines. The points of C are the $N_A N_B$ intersections between the lines a_i and the lines b_j .

In this network model, as before, a particular point of C is included in the active set if the pair of lines (a_i, b_j) which pass through it have been called into play in at least one association of an A -pattern with a B -pattern. Let us suppose that R pairs of patterns have been associated in this way, each pair comprising a selection of M_A lines from A and M_B lines from B . Then the chance that a given point of C has not been activated by the recording is

$$\exp(-RM_A M_B / N_C) = 1 - p, \text{ say}$$

where we have written N_C for $N_A N_B$. If B -patterns are being used to recall A -patterns, then there will be a minimum value of M_B such that if the threshold on the A -lines is set at M_B (so as to detect all the genuine lines) spurious lines will begin to be detected as well. (The argument is just the same as that applied to the correlograph earlier on.) This minimum value of M_B is given by

$$N_A p^{M_B} = 1$$

or $M_B = -\log N_A / \log p \approx \log_2 N_A$

Now the amount of information stored in the memory when R pairs of A -patterns have been memorized is roughly

$$I_A = R M_A \log N_A$$

But from our equation for $1 - p$

$$R M_A M_B = -N_C \log(1 - p)$$

therefore

$$I_A = N_C \log p \log(1 - p)$$

showing that, as in the correlograph, the density with which the associative net stores information is 69 per cent of the theoretical maximum value. We may note, in passing, that I_B , defined as $R M_B \log N_B$, is also equal to $N_C \log p \log(1 - p)$.

An associative network of this kind also operates (i) in parallel (ii) non-locally and (iii) in such a way that local damage or inaccuracy is not necessarily disastrous. We intend to go into the details of (iii) elsewhere. We now succumb to the temptation of indicating how such an associative memory might be realized in neural tissue though, as Brindley has pointed out⁶, function need not determine structure uniquely.

The system we have in mind is represented diagrammatically in Fig. 4. The horizontal lines are axons of the N_B input neurones b_1, b_2, \dots , while the vertical lines are dendrites of the N_A output neurones a_1, a_2, \dots . At the intersection of b_j with a_i is a modifiable synapse c_{ij} . This synapse is initially inactive, but becomes active after a coincidence in which a_i and b_j are made to fire at the same time by some external stimulus. Such a coincidence is

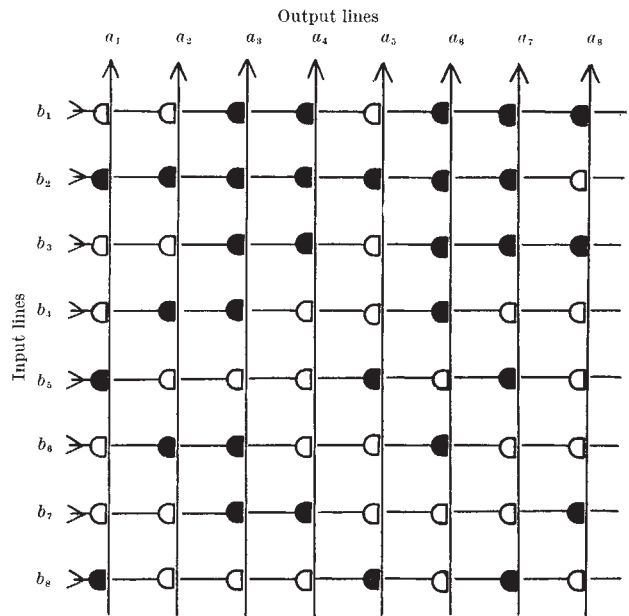


Fig. 4. An associative net.

supposed to occur if an A -pattern containing a_i is presented in association with a B -pattern containing b_j . After the activation of c_{ij} (which we regard as a permanent effect) the firing of b_j will locally depolarize the membrane of a_i . The output neurone a_i is then supposed to fire if M_B or more input cells depolarize it simultaneously.

In Fig. 4 we indicate what the state of the network would be after it had learned to associate the following pairs of patterns:

B-pattern	A-pattern
1,2,3	4,6,7
2,5,8	1,5,7
2,4,6	2,3,6
1,3,7	3,4,8

The synapses indicated by solid semicircles would be active, those indicated by open semicircles being still inactive. In this particular example, N_A and N_B are both 8, and $M_A (\approx \log_2 N_B)$ and $M_B (\approx \log_2 N_A)$ are both 3. R , the number of pairs of patterns associated, has been chosen so as to make p , the proportion of synapses active, close to 0.5; in fact p equals 0.5 exactly. These various numbers illustrate the system working near its maximum capacity. The reader may verify that every B -pattern except the first evokes the correct A -pattern at a threshold of 3; the only mistake the system makes is that when supplied with the B -pattern 1,2,3 it responds with an A -pattern 3,4,6,7 containing four elements.

To summarize, we have attempted to distil from holography the features which commend it as a model of associative memory, and have found that the performance of a holograph can be mimicked and actually improved on by discrete non-linear models, namely the correlograph and the associative net just described. Quite possibly there is no system in the brain which corresponds exactly to our hypothetical neural network; but we do attach importance to the principle on which it works and the quantitative relations which we have shown must hold if such a system is to perform, as it can, with high efficiency.

Received March 17, 1969.

¹ Van Heeden, P. J., *App. Optics*, **2**, 393 (1963).
² Pribram, K. H., *Sci. Amer.*, **220**, 73 (1969).
³ Willshaw, D., and Longuet-Higgins, H. C., *Machine Intelligence 4* (edit. by Michie, D.) (Edinburgh University Press, 1969).
⁴ Gabor, D., *Nature*, **217**, 1288 (1968).
⁵ Gabor, D., *Nature*, **217**, 584 (1968).
⁶ Brindley, G. S., *Proc. Roy. Soc., B*, **168**, 361 (1967).