

Neural networks that use three-state neurons

This article has been downloaded from IOPscience. Please scroll down to see the full text article.

1989 J. Phys. A: Math. Gen. 22 2265

(<http://iopscience.iop.org/0305-4470/22/12/023>)

View [the table of contents for this issue](#), or go to the [journal homepage](#) for more

Download details:

IP Address: 129.252.86.83

The article was downloaded on 31/05/2010 at 11:41

Please note that [terms and conditions apply](#).

Neural networks that use three-state neurons

Jonathan S Yedidia

Department of Physics, Jadwin Hall, Princeton University, Princeton, NJ 08544, USA
and

Laboratoire de Physique Statistique, Ecole Normale Supérieure, 24 rue Lhomond,
75231 Paris Cedex 05, France

Received 17 January 1989

Abstract. The dynamics of a neural network which uses three-state neurons (1, 0 and -1) is solved exactly in the limit of non-symmetric and highly dilute synapses. Recursion relations for the 'activity' (the fraction of non-zero neurons) and overlap of the network with a given pattern are derived which have three generic kinds of fixed points: a retrieval fixed point, a chaotic fixed point which corresponds to non-zero activity but no overlap, and a 'zero' fixed point where all the neurons go to the 0 state. As the non-retrieval fixed points both have activities different from the retrieval fixed point, one can easily tell whether a pattern has been recovered. An analysis of which fixed points occur as a function of the thresholds and the storage ratio of the system yields remarkably rich phase diagrams. Optimising the threshold level can be very important, especially when low-activity patterns are stored. A similar analysis can be applied to 'biased' networks using two-state (1, 0) neurons. Finally, we find that mixture states which have an overlap with two patterns can be stabilised by a threshold in the networks using three-state neurons. This property allows 'larger' (higher activity) memories to be naturally constructed out of smaller ones.

In this paper, we introduce a model of neural networks which uses three-state (1, 0, -1) neurons, and solve it in the dilute non-symmetric limit using the methods introduced for two-state (1, -1) neurons by Derrida, Gardner and Zippelius [1]. The motivation for considering such a model is twofold. First, one can consider it as a first step towards the analysis of analogue neurons, which can be modelled in the limit $S \rightarrow \infty$ by S -state neurons where the S states correspond to 'voltage levels' between -1 and 1. The move from two states to three already highlights many of the qualitative changes in moving from digital to analogue neurons. Second, one can think of the 0 ('off') neurons as a kind of background on top of which one can combine patterns of $+1$ and -1 ('on') neurons which contain information. For real-world applications of neural networks such a distinction between background and pattern is often very natural. The model turns out to have many interesting features, especially a kind of 'threshold control', which can best be understood by turning directly to its solution.

The model consists of a system of N neurons S_i which can take on the values 0, ± 1 , and whose interactions J_{ij} depend on p stored patterns. The J_{ij} are given by

$$J_{ij} = \frac{C_{ij}}{C} \sum_{\mu=1}^p \xi_i^{\mu} \xi_j^{\mu} \quad (1)$$

where $\xi_i^{\mu} (=0, \pm 1)$ is the value at neuron i of pattern μ and the C_{ij} are random independent parameters (independent of C_{ji}) which take the value 1 with probability C/N and 0 with probability $1 - C/N$. C represents the mean number of synapses per

neuron. Each pattern has an 'activity' given by $a_\mu = (1/N) \sum_i |\xi_i^\mu|$. Aside from this constraint on the total activity of each pattern, we will assume that $\xi_i^\mu = +1$ or -1 with equal probability.

One can use parallel dynamics, when all spins are updated simultaneously at discrete time steps, or sequential dynamics, when spins are updated one at a time in a random order. The spins are updated according to the rule

$$S_i(t+1) = \begin{cases} +1 & \text{with probability } \exp[(h_i - \theta)/T]/Z \\ 0 & \text{with probability } 1/Z \\ -1 & \text{with probability } \exp[(-h_i - \theta)/T]/Z \end{cases} \quad (2)$$

where $Z \equiv \exp[(h_i - \theta)/T] + 1 + \exp[(-h_i - \theta)/T]$; $h_i \equiv \sum_j J_{ij} S_j$ is the 'local field'; θ is a 'threshold'; and T is the temperature. The zero-temperature response of a neuron is shown in figure 1: a neuron will point in the direction of its local field if the local field exceeds a threshold, otherwise it will have value 0. We have chosen the positive and negative thresholds to be equal for simplicity, but it is already clear that a simple generalisation to an S -state neuron with $S - 1$ possibly non-symmetric thresholds will give a discrete approximation to any analogue deterministic response function. In this paper, we restrict ourselves to three-state neurons, in order to understand fully what happens when we take this first step.

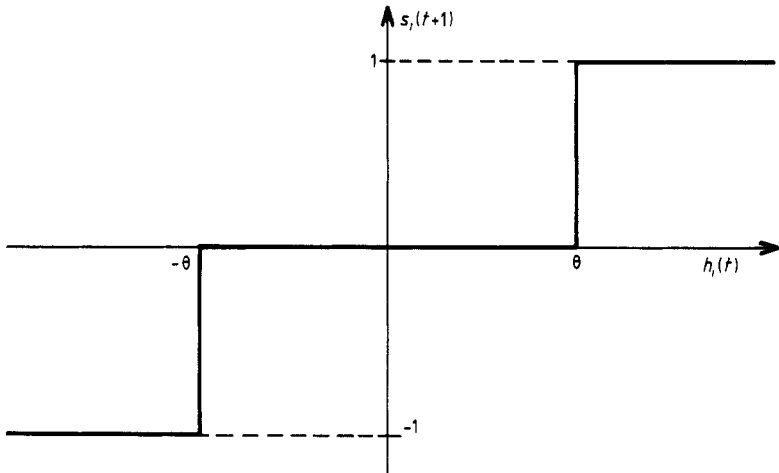


Figure 1. The zero-temperature 'response function' of a three-state neuron.

We consider first the case when the network has some macroscopic overlap with one of the patterns (say pattern 1) and no macroscopic overlap with any of the other patterns. We solve the model by deriving recursion relations for two 'order parameters': the 'dynamic activity' $a_D(t)$ of the network defined by $a_D(t) \equiv (1/N) \sum_i |S_i(t)|$ and the scaled overlap $m_1(t)$ of the network with pattern 1 defined by $m_1(t) \equiv (1/Na_1) \sum_i \xi_i^1 S_i(t)$. We use this definition of the overlap, rather than other possible definitions which, for instance, might not be indifferent to the state of those neurons which should be 'off' in pattern 1, because this definition arises naturally in the solution of the model.

Using the methods of [1] (see [2-6] for other applications of this method) for parallel dynamics, we find that in the limit $1 \ll C \ll \log N$, these two order parameters

satisfy the recursion relations

$$m_1(t+1) = \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dz e^{-z^2/2} \left(\frac{e^{(A-\theta)} - e^{-(A+\theta)}}{e^{(A-\theta)} + 1 + e^{-(A+\theta)}} \right) \tag{3}$$

and

$$a_D(t+1) = \frac{a_1}{\sqrt{\pi}} \int_{-\infty}^{\infty} dz e^{-z^2/2} \left(\frac{e^{(A-\theta)} + e^{-(A+\theta)}}{e^{(A-\theta)} + 1 + e^{-(A+\theta)}} \right) + \frac{(1-a_1)}{\sqrt{\pi}} \int_{-\infty}^{\infty} dz e^{-z^2/2} \left(\frac{e^{(B-\theta)} + e^{-(B+\theta)}}{e^{(B-\theta)} + 1 + e^{-(B+\theta)}} \right) \tag{4}$$

where

$$A \equiv [a_1 m_1(t) - z(2\alpha a_p^2 a_D(t))^{1/2}] / T$$

$$B \equiv z(2\alpha a_p^2 a_D(t))^{1/2} / T$$

$$a_p^2 \equiv \sum_{\mu \geq 2} a_\mu^2 / (p-1)$$

and $\alpha = p/C$. A slightly different set of recursion relations exists for sequential dynamics, which lead, in any case, to the same fixed points [1]. The equation for the overlap is derived by considering the neurons which overcome the threshold to point in the ‘right’ direction for pattern 1, and subtracting the neurons which actually overcome the threshold to point in the ‘wrong’ direction. The equation for the dynamic activity naturally separates into a part for the fraction a_1 of the neurons that should be ‘on’ in pattern 1, and are therefore affected by the overlap with that pattern, and the fraction $1 - a_1$ of the neurons that should be ‘off’ in pattern 1. For both parts of this equation, we add the neurons which overcome the thresholds to point in either the positive or the negative direction. These equations give us a complete solution for the dynamics of this model when there is a macroscopic overlap with one pattern—that is, a trajectory in (a_D, m_1) space parametrised by $\theta, a_1, \alpha a_p^2$ and T . It should also be mentioned that identical equations arise for a fully connected system if one makes the approximation of simply iterating the one-time-step parallel dynamics [7-9]. At zero temperature and when the activity of all patterns is equal ($a_1 = a_p = a$) the equations reduce to

$$m_1(t+1) = \frac{1}{2} \left[\operatorname{erf} \left(\frac{m_1(t) - \theta/a}{\sqrt{2\alpha a_D(t)}} \right) + \operatorname{erf} \left(\frac{m_1(t) + \theta/a}{\sqrt{2\alpha a_D(t)}} \right) \right] \tag{5}$$

and

$$a_D(t+1) = \frac{a}{2} \left[2 + \operatorname{erf} \left(\frac{m_1(t) - \theta/a}{\sqrt{2\alpha a_D(t)}} \right) - \operatorname{erf} \left(\frac{m_1(t) + \theta/a}{\sqrt{2\alpha a_D(t)}} \right) \right] + (1-a) \left[1 - \operatorname{erf} \left(\frac{\theta/a}{\sqrt{2\alpha a_D(t)}} \right) \right] \tag{6}$$

where

$$\operatorname{erf} \left(\frac{X}{\sqrt{2}} \right) \equiv \sqrt{\frac{2}{\pi}} \int_0^X e^{-z^2/2} dz. \tag{7}$$

These equations are a generalisation of the equations derived for two-state (1, -1) neurons [1] which can be recovered by taking the limits of threshold $\theta = 0$ and pattern

activity $a = 1$. The main difference is the existence of the activity order parameter $a_D(t)$ (which one should note enters into the 'noise' term $\sqrt{2\alpha a_D(t)}$).

These exact recursion relations have been studied by numerical iteration. It should be noted that the dynamic activity $a_D(t)$ is generally not a constant during this iteration. We concentrate in the following on the fixed points of the iteration. At any given pattern activity a , threshold θ and storage ratio α , we will find one of three generic kinds of fixed points. Which fixed point the system actually reaches will generally depend also on the initial values of the scaled overlap $m_1(t)$ and dynamic activity $a_D(t)$. The different kinds of fixed points are: a 'retrieval' fixed point where the final overlap and activity are greater than zero ($m_1^* > 0, a_D^* > 0$ —the star superscript indicates the value of the quantity at the fixed point); a 'chaotic' (or 'paramagnetic') fixed point where there is no overlap, but the final activity sustains itself anyway ($m_1^* = 0, a_D^* > 0$); and a 'zero' fixed point where all the neurons turn off ($m_1^* = 0, a_D^* = 0$). Of course, the 'retrieval' and 'chaotic' fixed points are only fixed points in the sense that $m_1(t)$ and $a_D(t)$ are fixed—which spins are up, down or off will continually change. The overall situation is best summarised in the form of a phase diagram which shows which fixed points coexist at a given threshold θ , storage ratio α and pattern activity a . In figures 2(a) and 2(b), we draw such phase diagrams in the (θ, α) plane for pattern activities $a = 1$ and $a = 0.01$.

Consider first the case where $a = 1$, which means that the stored patterns consist only of +1s and -1s, even though the neurons can in the course of the dynamics also take on the value 0. To understand this phase diagram better, imagine that we fix the storage ratio α at 0.5 and increase the threshold θ (figure 3). At $\theta = 0$, all the neurons will be 'on' and we have a situation equivalent to a network that uses two-state (1, -1)

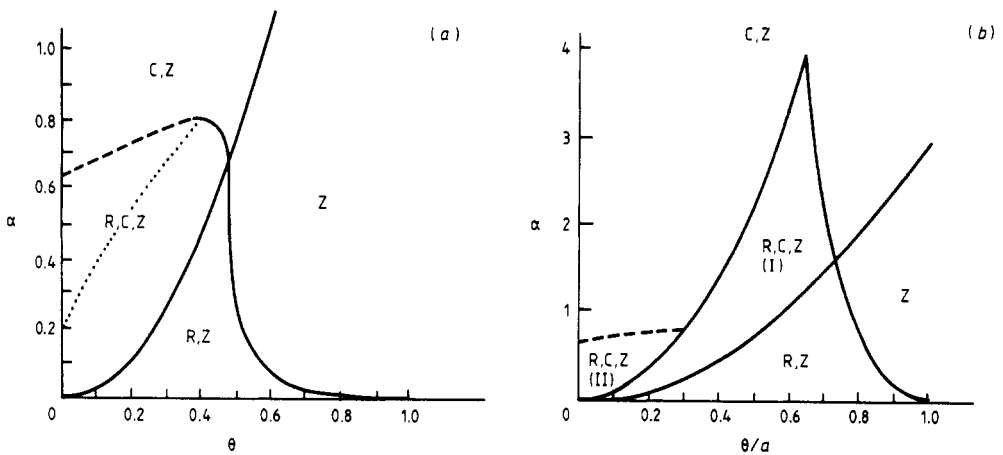


Figure 2. (a) The zero-temperature pattern activity $a = 1$ phase diagram in the (threshold θ , storage ratio α) plane. 'R' stands for the retrieval fixed point (fixed point overlap $m_1^* > 0$, fixed-point dynamic activity $a_D^* > 0$), 'C' for the chaotic fixed point ($m_1^* = 0, a_D^* > 0$) and 'Z' for the zero fixed point ($m_1^* = 0, a_D^* = 0$). When more than one fixed point coexists in a region of the phase diagram, which one is actually reached depends on the initial conditions. The full curves represent first-order phase transitions. The broken curve represents a second-order phase transition. The dotted curve gives the θ which optimises the fixed-point overlap m_1^* at a given α . (b) The $T = 0, a = 0.01$ phase diagram in the (θ, α) plane. The abbreviations are as in (a). In region (I), the retrieval fixed point dynamic activity a_D^* is approximately equal to the pattern activity a . In region (II), a_D^* is nearly equal to 1.

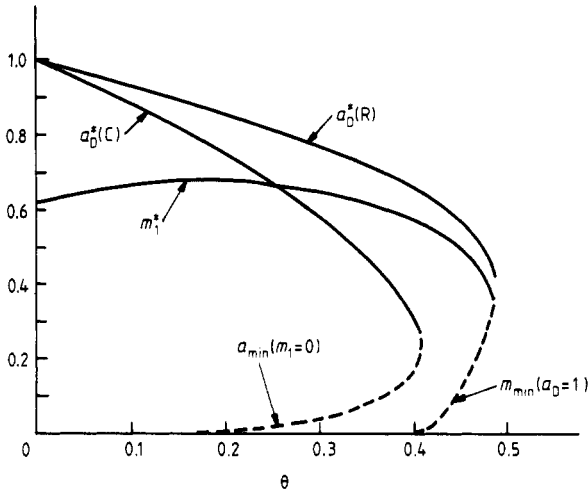


Figure 3. Data for the case temperature $T = 0$, pattern activity $a = 1$, storage ratio $\alpha = 0.5$. Plotted are the retrieval fixed-point overlap m_1^* , the retrieval fixed point activity $a_D^*(R)$, the chaotic fixed point activity $a_D^*(C)$, the minimal initial activity $a_{\min}(m_1 = 0)$ to go to the chaotic fixed point when the overlap $m_1 = 0$, and the minimal initial overlap $m_{\min}(a_D = 1)$ to go to the retrieval fixed point when the initial activity $a_D(\text{init}) = 1$.

neurons. Because α is so near to the critical capacity $\alpha_c = 2/\pi$, where there is a transition to the chaotic phase, the fixed point overlap m_1^* is quite low. As we increase the threshold θ , the fixed point retrieval is actually improved, as the threshold tends to turn off those neurons which have the lowest local field, which are the most important ones in causing errors. We find that the fixed point overlap is optimised at $\theta \approx 0.17$. Note that we can calculate explicitly the 'basin of attraction' (or equivalently the minimal initial overlap given some initial activity). We see in figure 3 that the minimal initial overlap is extremely low in the region around $\theta = 0.17$ which means that the basin of attraction is very large. If, however, the initial overlap is equal to zero, then the network goes instead to the chaotic fixed point (if the initial activity is large) or the zero fixed point (if the initial activity is small). Note that the fixed-point activity is always different for the chaotic fixed point compared with the retrieval fixed point. This is because in the chaotic fixed point there is no coherent piece from pattern 1 sustaining the activity. For $\alpha = 0.5$, beyond the point $\theta \approx 0.41$, the chaotic fixed point cannot sustain itself and disappears through a first-order transition. In this region of the phase diagram, it is particularly easy to tell whether or not a pattern has been recovered—if no pattern is recovered, the network simply turns off. The retrieval fixed point itself vanishes through a first-order transition at $\theta \approx 0.48$. All the full curves in figure 2(a) indicate first-order phase transitions and the broken curve indicates a second-order transition. The dotted curve indicates the optimal threshold at a given storage ratio α .

Next we consider the phase diagram when the pattern activity $a = 0.01$, shown in figure 2(b). We see immediately from this figure that for small patterns, it is very important to optimise the threshold (which, to a first approximation, scales with the size of the stored patterns). At the optimal threshold of $\theta/a \approx 0.65$, the network has a capacity of $\alpha_c \approx 4.0$, considerably higher than it did for $a = 1$. Note also that the transition at the optimal threshold is now first order, which means that the fixed-point

overlap is quite high as long as one is in the retrieval phase. For example, for $\theta = 0.65$, the fixed point overlap $m_1^* \approx 0.89$ even at a storage ratio of $\alpha = 3.9$.

Note that the region of the phase diagram where the three fixed points coexist is now divided into two subregions which are separated by a first-order phase transition (figure 2(b)). In subregion I, which includes the vicinity of the optimal threshold, the fixed-point activity a_D^* of the retrieval fixed point is of the order of the pattern activity a . In subregion II, which includes the area where the threshold is zero, a_D^* is $\mathcal{O}(1)$. The second-order transition above subregion II is the remnant of the second-order transition found for networks of two-state (1, -1) neurons [1]. These technical points concerning subregion II are not really very important to the functioning of the system as an associative memory.

Very similar phase diagrams exist for 'biased' dilute non-symmetric neural network models which use two-state (1, 0) neurons[†]. Tsodyks [6] has solved such a model which uses the altered Hebbian prescription

$$J_{ij} = \frac{C_{ij}}{C} \sum_{\mu} (\xi_i^{\mu} - a)(\xi_j^{\mu} - a)$$

and found zero-temperature recursion relations which can be rewritten as

$$m_1(t+1) = \frac{1}{2} \left[\operatorname{erf} \left(\frac{(1-a)m_1(t) - \theta/a}{\sqrt{2\alpha a_D(t)}} \right) + \operatorname{erf} \left(\frac{am_1(t) + \theta/a}{\sqrt{2\alpha a_D(t)}} \right) \right] \tag{8}$$

$$a_D(t+1) = \frac{a}{2} \left[1 + \operatorname{erf} \left(\frac{(1-a)m_1(t) - \theta/a}{\sqrt{2\alpha a_D(t)}} \right) \right] + \frac{(1-a)}{2} \left[1 - \operatorname{erf} \left(\frac{am_1(t) + \theta/a}{\sqrt{2\alpha a_D(t)}} \right) \right] \tag{9}$$

where the 'dynamic activity' $a_D(t)$ is now defined by $a_D(t) = (1/N) \sum_i S_i(t)$, the activity of the patterns is defined by $a_{\mu} = (1/N) \sum_i \xi_i^{\mu}$ and all $a_{\mu} = a$, and the scaled overlap is defined by $m_1(t) = (1/Na) \sum_i S_i(t)(\xi_i^1 - a)$. Tsodyks concentrated his analysis on the limit $a \rightarrow 0$. In that limit, and if we restrict ourselves to the region around the optimal threshold where we can assume that

$$\operatorname{erf} \left(\frac{\theta/a}{\sqrt{2\alpha a_D(t)}} \right) \approx 1 \tag{10}$$

the sets of equations (5), (6) and (8), (9) become identical. Thus, the models have identical behaviour in the limit of small pattern activity and nearly optimal thresholds. The intuitive reason for the similarity of the two models in this limit and region is that nearly all the 'errors' for the three-state model will consist of neurons having value 0 when they should be +1 or -1, rather than the more serious error of a neuron having value +1 when it should be -1 or vice versa. Thus the three-state model behaves very much like a two-state (1, 0) model, albeit with a considerably more convenient and efficient storage prescription.

To be more specific, the three-state model has the following advantages over the two-state model. First, it stores about twice as much information because the critical capacity α_C is nearly the same, but an 'on' neuron can now be 1 or -1, instead of always 1. Secondly, the storage prescription is simpler, and it is much easier to implement with digital synapses. Finally, as we will see below, it is easy to combine small patterns into bigger ones using the three-state neurons. Using two-state neurons, bigger patterns would ultimately consist entirely of 1 neurons, with no information stored at all.

[†] I thank Werner Krauth for pointing out to me that this might be true.

Nevertheless, as a consequence of the mathematical equivalence of the two models, we can read off some interesting results from reference [6]. In particular, we find that the optimal threshold and the optimal capacity at that threshold have the following asymptotic dependence on the pattern activity a as $a \rightarrow 0$:

$$\theta_{\text{OPT}} \approx a(1 - |\ln a|^{-1/2}) \quad \alpha_C^{\text{OPT}} \approx \frac{-1}{2a \ln a}. \quad (11)$$

Comparing the two models when the pattern activity $a = 1$ for the three-state model and $a = \frac{1}{2}$ for the two-state model, so that both models store one bit per neuron, we find by solving the recursion relations given above that the three-state model again can store about twice as much information as the two-state model; the difference now being that α_C is half as large for the two-state model. In fact we find that the phase diagrams are nearly identical if we make the replacements $\alpha^{(3)} \rightarrow 2\alpha^{(2)}$ and $\theta^{(3)} \rightarrow 2\theta^{(2)}$, where the superscripts denote the model using three-state or two-state neurons respectively. Finally, returning to the limit $a \ll 1$, we find that the full phase diagram for the two-state model (with for example $a = 0.01$) is very similar to the phase diagram for the three-state model shown in figure 2(b), the only difference being in the lower left-hand corner, where the second-order transition between region II and the chaotic phase occurs at a storage ratio α of around $1/\pi$ for the two-state model, and $2/\pi$ for the three-state model. These phase diagrams put in more complete perspective the observation made by Tsodyks [6] that the transition for $a \ll 1$ in the two-state model is first order, which is strictly speaking only true in the region around the optimal threshold.

Finally, we consider the case of a configuration which has macroscopic overlaps $m_1(t)$ and $m_2(t)$ on two patterns (1 and 2) and zero overlap with the other $p - 2$ patterns. We assume that the two patterns have an overlap Q and an 'activity overlap' R :

$$Q \equiv \frac{1}{N} \sum_i \xi_i^1 \xi_i^2 \quad R \equiv \frac{1}{N} \sum_i |\xi_i^1| |\xi_i^2|. \quad (12)$$

We find, in the limit $1 \ll C \ll \log N$ (and specialising immediately to the case $T = 0$), that

$$a_1 m_1(t+1) = \sum_{\tau=\pm 1} \left[\frac{R+Q}{4} \operatorname{erf} \left(\frac{m_1(t) + m_2(t) + \tau\theta/a_p}{D} \right) + \frac{R-Q}{4} \operatorname{erf} \left(\frac{m_1(t) - m_2(t) + \tau\theta/a_p}{D} \right) + \frac{a_1 - R}{2} \operatorname{erf} \left(\frac{m_1(t) + \tau\theta/a_p}{D} \right) \right] \quad (13)$$

($m_2(t+1)$ is obtained by making the substitutions $m_1 \leftrightarrow m_2$, $a_1 \leftrightarrow a_2$) and

$$a_D(t+1) = \sum_{\tau=\pm 1} \left\{ \frac{R+Q}{4} \left[1 + \tau \operatorname{erf} \left(\frac{m_1(t) + m_2(t) - \tau\theta/a_p}{D} \right) \right] + \frac{R-Q}{4} \left[1 + \tau \operatorname{erf} \left(\frac{m_1(t) - m_2(t) - \tau\theta/a_p}{D} \right) \right] + \frac{a_1 - R}{2} \left[1 + \tau \operatorname{erf} \left(\frac{m_1(t) - \tau\theta/a_p}{D} \right) \right] + \frac{a_2 - R}{2} \left[1 + \tau \operatorname{erf} \left(\frac{m_2(t) - \tau\theta/a_p}{D} \right) \right] + \frac{(1+R-a_1-a_2)}{2} \left[1 + \tau \operatorname{erf} \left(\frac{-\tau\theta/a_p}{D} \right) \right] \right\} \quad (14)$$

where $D = \sqrt{2\alpha a_D(t)}$ and $a_p^2 = \sum_{\mu \geq 3} a_\mu^2 / (p-2)$. These formulae, although they are rather long, are not too difficult to derive. One needs to consider a neuron at site i and note the following probabilities:

$\xi_i^1 = \xi_i^2 \neq 0$	with probability	$(R+Q)/2$
$\xi_i^1 = -\xi_i^2 \neq 0$	with probability	$(R-Q)/2$
$\xi_i^1 \neq 0; \xi_i^2 = 0$	with probability	$a_1 - R$
$\xi_i^1 = 0; \xi_i^2 \neq 0$	with probability	$a_2 - R$
$\xi_i^1 = \xi_i^2 = 0$	with probability	$1 + R - a_1 - a_2$.

The corresponding terms in the equations then follow by calculations which are very similar to those necessary when one considers an overlap with a single pattern. These equations give us trajectories in $(a_D(t), m_1(t), m_2(t))$ space parametrised by $\theta, a_p, a_1, a_2, Q, \alpha$ and R .

Obviously, this is too large a parameter space to explore thoroughly, but we can make some general observations. In particular, it should be noted that using a threshold $\theta > 0$, we can stabilise states which are mixtures of two patterns. In previous work with two-state $(1, -1)$ models, it was found that for fully connected networks [10, 11], feed-forward networks [12], and non-symmetric diluted networks [1], mixture states of two uncorrelated patterns ($Q = 0$) were always unstable. Here we find that if, for example, $a_1 = a_2 = a_p = R = 1$, and $Q = 0$, then the mixture state $m_1^* \approx 0.5, m_2^* \approx 0.5, a_D^* \approx 0.5$ is stable for low α and $\theta > 0$. The reason is that the threshold simply turns off all neurons which contribute to the two patterns in opposite directions, and leaves on all the other neurons. This general behaviour is also true for mixture states of two smaller patterns. In that case, one generally finds that the few neurons that contribute to the two patterns in opposite directions ($\xi_i^1 = -\xi_i^2 \neq 0$) are turned off by the threshold, leaving on those neurons which contribute to only one of the patterns or neurons which contribute to the two patterns in the same direction. Thus, we generally find that mixture states of two small patterns are stable with fixed-point-scaled overlaps for the two patterns both nearly equal to 1. A simple extension of these results shows that this is also true for mixture states of more than two patterns.

The usual philosophy about mixture states is that they are something of a nuisance. This three-state model leads us to consider instead the idea that mixture states might in fact be quite valuable, in that we can combine smaller memories into larger ones, thus seeing 'words' where before we had only seen 'letters'. One might imagine that if two small patterns were consistently recovered concurrently, synapses might eventually be strengthened between the neurons forming the patterns, thus creating a larger pattern. If the mixture state had not been stable in the first place, this could not have happened.

In this paper, we have introduced a neural network that uses three-state $(1, 0, -1)$ neurons and have solved for its dynamics. The main results of this paper are the phase diagrams of figure 2. Similar phase diagrams exist for models using two-state $(1, 0)$ neurons but have not, perhaps, been fully appreciated. The three-state model stores twice as much information as the two-state model and uses a storage rule which is much more easily implemented using digital synapses. These phase diagrams make obvious the significance of the threshold, which can be used to optimise retrieval, 'zero out' the network, or search for different-sized patterns (the optimal threshold scales with the pattern size). Another important result is that for the three-state model,

mixture states of two patterns are inherently stable, allowing the network to ultimately form larger patterns out of smaller ones. Some extensions of the analysis are evident. First, one can repeat the analysis using the replica [11] or cavity method [13] to solve for the case of a fully connected network (see also [14]). The main difference in this case will be the existence of order parameters which correspond to the freezing of the system. We can expect that the 'chaotic' phase will be replaced in this case by a spin-glass phase. Secondly, one can consider S -state neurons where S becomes very large. These neurons are an excellent discrete approximation to deterministic analogue neurons. Using this formalism, some long-standing questions about the advantages of analogue neurons can be addressed.

Acknowledgments

I thank P W Anderson, Ido Kanter, Werner Krauth and Marc Mézard for helpful discussions. I am grateful to the physicists at the Laboratoire de Physique de l'Ecole Normale Supérieure for their warm hospitality. I acknowledge the financial support of AT&T Bell Laboratories.

Note Added. C Meunier, D Hansel and A Verga [14] independently considered neural networks using three-state neurons and performed numerical simulations on fully connected networks. Their results are in qualitative agreement with the analytical results for highly dilute networks reported here. I thank them for communicating their results to me.

References

- [1] Derrida B, Gardner E and Zippelius A 1987 *Europhys. Lett.* **4** 167
- [2] Kree R and Zippelius A 1987 *Phys. Rev. A* **36** 4421
- [3] Derrida B and Nadal J-P 1987 *J. Stat. Phys.* **49** 993
- [4] Kanter I 1988 *Phys. Rev. Lett.* **60** 1891
- [5] Gutfreund H and Mézard M 1988 *Phys. Rev. Lett.* **61** 235
- [6] Tsodyks M V 1988 *Europhys. Lett.* **7** 203
- [7] Kinzel W 1985 *Z. Phys. B* **60** 205
- [8] Krauth W, Nadal J-P and Mézard M 1988 *J. Phys A: Math. Gen.* **21** 2995
- [9] Krauth W, Nadal J-P and Mézard M 1988 *preprint*
- [10] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. A* **32** 1007
- [11] Amit D J, Gutfreund H and Sompolinsky H 1985 *Phys. Rev. Lett.* **55** 1530; 1987 *Ann. Phys., NY* **173** 30
- [12] Meir R and Domany E 1988 *Phys. Rev. A* **37** 608
- [13] Mézard M, Parisi G and Virasoro M A 1987 *Spin Glass Theory and Beyond* (Singapore: World Scientific)
- [14] Meunier C, Hansel D and Verga A 1988 *preprint*