1 Storage Capacity in Associative Networks for (Near) Perfect Recall

We consider a Hopfield network with the standard Hebb-like learning rule and ask how many memories we can imbed in a network of N neurons with the constraint that we will accept at most one bit (one neuron's output in only one memory state) of error

$$\begin{aligned} Input \quad &= \sum_{j \neq i}^{N} W_{ij} S_j \\ &= \frac{1}{N} \sum_{\mu=1}^{p} \sum_{j \neq i}^{N} \zeta_i^\mu \zeta_j^\mu S_j \end{aligned} \tag{1.1}$$

where p is the number of stored memories, N is the number of neurons and

$$W_{ij} \equiv \frac{1}{N} \sum_{\mu=1}^{p} \zeta_i^\mu \zeta_j^\mu \tag{1.2}$$

is the synaptic weight matrix given by the Hebb rule.

Now, check stability of stored state. Make $S_j = \zeta_j^1$, one of the stored memory states, so that

$$\begin{aligned} Input \quad &= \frac{1}{N} \sum_{\mu=1}^{p} \sum_{j \neq i}^{N} \zeta_i^\mu \zeta_j^\mu \zeta_j^1 \\ &= \frac{1}{N} \sum_{\mu=1}^{p} \zeta_i^\mu \sum_{j \neq i}^{N} \zeta_j^\mu \zeta_j^1 \\ &= \frac{1}{N}\, \zeta_i^1 \sum_{j \neq i}^{N} \zeta_j^1 \zeta_j^1 + \frac{1}{N} \sum_{\mu \neq 1}^{p} \zeta_i^\mu \sum_{j \neq i}^{N} \zeta_j^\mu \zeta_j^1 \end{aligned} \tag{1.3}$$

On average, the second term is zero, so that

$$\text{Average of Input} = \frac{1}{N}\, \zeta_i^1 (N-1) \; \simeq \; \zeta_i^1 \tag{1.4}$$

What is the variance? The second term, summed over random vectors with zero mean, consists of the sum of $(p-1)$ inner products of vectors with $(N-1)$ terms. Each term is +1 or -1, i.e., binomially distributed, so that

$$\text{Variance of Input} = \frac{1}{N}\left[\frac{1}{N}\cdot(p-1)\cdot(N-1)\right] \simeq \frac{p}{N} \tag{1.5}$$

This results in a fluctuation to the input with a standard deviation, σ, of

$$\sigma = \pm\sqrt{p/N} \tag{1.6}$$

Noise hurts only if the magnitude exceeds 1. The noise becomes Gaussian for large p and N, which is the limit of interest, Thus the probability of an error in the recall of all stored states is

$$\begin{aligned} P_{error} &= \frac{1}{\sqrt{2\pi}\ \sigma}\ \left[\ \int_{\infty}^{-1}\ e^{-x^2/2\sigma^2}\ dx\ + \int_{+1}^{\infty}\ e^{-x^2/2\sigma^2}\ dx\ \right] \\ &= \frac{\sqrt{2}}{\sqrt{\pi}\ \sigma}\ \int_{+1}^{\infty}\ e^{-x^2/2\sigma^2}\ dx \\ &= \frac{2}{\sqrt{\pi}}\ \int_{\sqrt{\frac{N}{2p}}}^{\infty}\ e^{-x^2}\ dx \\ &\equiv \operatorname{erfc}\left(\sqrt{\frac{N}{2p}}\right) \end{aligned} \tag{1.7}$$

where efrc(x) is the complementary error function and we again note that the average of the error term is zero. We recall that for $\frac{N}{2p} >> 1$, the complementary error fucntion may be approximated by an asymptotic form, so that

$$P_{error} \simeq \frac{1}{\sqrt{\pi}}\frac{2p}{N}e^{-N/2p} \tag{1.8}$$

We have a nice and closed expression in a relevant limit!

Now $N \cdot p$ is total number of "bits" in the network. Suppose only less than one bit can be in error. Then we equate probabilities of correct to within a factor of one bit, or $\frac{1}{Np}$. Thus

$$(1 - P_{error})^{Np} \geq 1 - \frac{1}{Np} \tag{1.9}$$

But Np is large and P_{error} will be small by construction, so $1 - Np \times P_{error} \geq 1 - \frac{1}{Np}$ and thus

$$P_{error} < \frac{1}{(Np)^2} \tag{1.10}$$

From the above expansion of the gaussian error:

$$log\,[P_{error}] \simeq -\frac{1}{2}\;log\;\pi - \frac{N}{2p} - log\;\frac{N}{2p} \tag{1.11}$$

From the constraint on the desired error:

$$log\,[P_{error}] < -2\;log(Np) \tag{1.12}$$

Thus

$$-\frac{1}{2}\;log\;\pi - \frac{N}{2p} - log\;\frac{N}{2p}\; < -2\;log\;(Np) \tag{1.13}$$

We now let $N \rightarrow \infty$ with N/p constant. Keeping zero-th and first order terms, we have:

$$\frac{N}{2p} > 2\;log\;(Np) \tag{1.14}$$

or

$$\frac{2p}{N} < \frac{1}{2\;log\;(N) + 2\;log\;(p)} \tag{1.15}$$

so

$$p\;\;<\;\;\frac{1}{4}\;\frac{N}{log\;N} \tag{1.16}$$

Note that p has a similar scaling for the choice of a fixed, nonzero error rate.

Thus we see that an associate memory based on a recurrent Hopfield network stores a number of memories that scales more weakly than the number of neurons if one cannot tolerate any errors upon recall. If a finite number of errors can be tolerated, a statistical mechanical deviation shows that the log (N) term is absent and that $p < 0.14N$.