Neural Network with Memory and Cognitive Functions

Janusz A. Starzyk, and Yue Li

School of Electrical Engineering and Computer Science
Ohio University, Athens, OH 45701, U.S.A.
starzyk@bobcat.ent.ohiou.edu
David D. Vogel

Ross University School of Medicine, Commonwealth of Dominica
Dvogel@rossmed.edu.dm

Abstract. This paper provides an analysis of a new class of distributed memories known as R-nets. These networks are similar to Hebbian networks, but are relatively sparsly connected. R-nets use simple binary neurons and trained links between excitatory and inhibitory neurons. They use inhibition to prevent neurons not associated with a recalled pattern from firing. They are shown to implement associative learning and have the ability to store sequential patterns, used in networks with higher cognitive functions. This work explores the statistical properties of such networks in terms of storage capacity as a function of R-net topology and employed learning and recall mechanisms.

I. R-nets Neural Network Organization

Main Concept of R-nets

R-nets have been used as components in the modular construction of larger networks capable of computations analogous to serial memory, classical and operant conditioning, secondary reinforcement, refabrication of memory, and fabrication of possible future events. R-net components of these larger networks appear to be appropriate objects of more detailed analysis than has been performed [Vogel, 2005].

R-nets stress biological plausibility and have demonstrated large storage capacities with the sparse connectivity of mammalian cortex. The number of synapses of principal cells on interneurons is at least 320 [Sik, Tamamaki, & Freund 1993]; the number of synapses of interneurons on principal cells is 1000 to 3000 [Freund & Buzsáki, 1996] and the ratio of interneurons to principal cells is roughly 0.2. The R-net modeled for this paper has 40% of excitatory neuron pairs linked though at least 1 inhibitory neuron [Vogel, 2005]. The detailed network structure is described in previous studies [Vogel, 2005]. The biological plausibility of this arrangement is discussed by Vogel [2001, 2005] and also by Fujii, Aihara, and Tsuda [2004].

Mathematically, R-nets are defined as randomly connected artificial neural networks with primary and secondary neurons. The network structure and connections between primary and secondary neurons are discussed by Vogel [2001, 2005].

R-nets implement distributed memories able to recall input patterns. During training, an input pattern is presented to the R-net by activating a selected cluster C of primary neurons. All links between active primary neurons are trained. During recall a subset of one of the stored patterns is presented to the input, activating corresponding primary neurons (initial recall set). The initial recall set is expected to activate all neurons of one of the stored patterns that include the activated neurons as a subset. Each primary neuron is activated if it is not inhibited.

During recall, inhibitory neurons linearly sum the weighted projections.

$$a_{i,x} = \sum W_{i,e} a_{e,x} \qquad (1)$$

where $a_{i,x}$ represents the activity of the ith inhibitory neuron on the xth cycle, $a_{e,x}$ is the current activity of the eth excitatory neuron with possible values 0 or 1, and $W_{i,e}$ is the weight of the projection of the eth excitatory neuron onto the ith inhibitory neuron with possible values of 1(untrained) or 10(trained). The excitatory neurons are then synchronously updated according to the rules [Vogel, 2005].

II. Statistical Model of R-nets

A series of papers [Vogel and Boos, 1997; Vogel, 1998; Vogel 2001] demonstrated the substantial storage capacities of networks progressively approximating the R-nets. An R-net with 10^6 excitatory neurons and brain-like connectivity will store at least $2x10^8$ bits of information [Vogel, 2005]. In this section, a statistical model of R-nets is presented and is compared with simulated R-nets.

Let us assume that an R-net is characterized by the set of primary neurons ***P***, the set of secondary neurons ***S***, the primary neurons' outgoing sets, $\boldsymbol{k_p}$, and the secondary neurons' outgoing sets, $\boldsymbol{k_s}$. These numbers are related through $P * k_p = S * k_s$.

Let us define P_{ci} as a set of primary neurons reachable from a primary neuron C_i through the secondary neurons, and α as probability that $c_j \in P_{ci}$ for a selected primary neuron c_j and a given P_{ci}, so that $\alpha = \frac{P_{ci}}{P}$. The expected value of the number of different primary and secondary neurons reaching to (or reached from) a secondary and a primary neuron, are respectively

$$k_{sp} = P * \left(1 - \left(1 - \frac{1}{P}\right)^{k_s}\right) < k_s, k_{ps} = S * \left(1 - \left(1 - \frac{1}{S}\right)^{k_p}\right) < k_p \qquad (2)$$

If the links through other secondary neurons reached from a primary neuron are considered, the number of primary neurons linked to a given primary neuron is

$$P_{ci} = P * \left(1 - \left(1 - \frac{1}{P}\right)^{k_s k_p}\right) < k_s k_p \qquad (3)$$

Eliminating Spurious Neurons

Spurious neurons are defined as neurons that are not a part of the original pattern and that are activated during the recall process. The probability that a potential spurious neuron, $\boldsymbol{c_j}$, will be inhibited depends on the probability of the existence of an inhibitory link from an activated primary neuron.

The probability that a projection out of a primary neuron in a trained set is trained with T patterns stored in the R-net is estimated as:

$$P_{t1} = 1 - \left(1 - \frac{C}{P}\left(1 - \left(1 - \frac{\alpha}{k_{ps}}\right)^{C}\right)\right)^{T} \quad \textbf{(4)}$$

P_{t1} also equals $P_{t2,}$ the probability of a projection out of a secondary neuron being trained with T patterns stored.

As shown in Fig. 1, a neuron is spurious if it meets both of the following conditions: a) It has no projection from S_{wa}; and b) All its projections from S_{sa} are trained, where S_{sa} is the strongly activated set of secondary neurons and S_{wa} is the weakly activated set of secondary neurons.

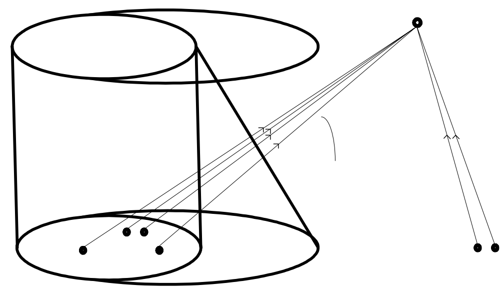

Fig. 1. Spurious neurons and activated secondary

The probability that a secondary node, y, belongs to S_a is the same as the probability that a selected secondary neuron is active during recall.

$$p_{asr} = 1 - \left(1 - \frac{C_r}{P}\right)^{k_{sp}} \quad \textbf{(5)}$$

Thus the probability that a node in S_a is strongly activated is approximately

$$P_{ssa} = 1 - (1 - P_{t1})^{(k_{sp}-1)P_{sr}+1} \quad \textbf{(6)}$$

and consequently the probability that a node in S_a is weakly activated is $1 - P_{ssa}$.

The probability that a potentially spurious node, z, is not linked to any node in S_{wa} is

$$P_{nwa} = (1 - P_{asr}(1 - P_{ssa}))^{k_{ps}} \quad \textbf{(7)}$$

Assume that a node z is not linked to S_{wa}. The probability that such primary node z is connected to a node in S_{sa} is

$$P_{pssa} = \frac{P_{asr}P_{ssa}}{1 - P_{asr}(1 - P_{ssa})} \quad \textbf{(8)}$$

Using this result, we can obtain the probability that a primary node z has k projections to S_{sa} and no projections to S_{wa}, and all of the links to S_{sa} are trained, thus obtaining the probability that z is a spurious node as

$$P_z = P_{nwa} \sum_{k=0}^{k_{ps}} \binom{k_{ps}}{k} P_{pssa}^k (1-P_{pssa})^{k_{ps}-k} P_{t1}^k \qquad \textbf{(9)}$$

The increasing probability of spurious neurons with increasing numbers of patterns stored limits the maximum number of patterns that can be stored in the R-net memory. Since, in the recall process, we can tolerate no more than S_{max} spurious neurons, and each neuron in the P-C set has probability of being spurious equal to P_z, then the recall set has no more than S_{max} spurious neurons with the probability

$$P_{NO\_spurious} = \sum_{i=0}^{S_{max}} \binom{P-C}{i} P_z^i (1-P_z)^{P-C-i} \qquad \textbf{(10)}$$

The analysis is in reasonable agreement with actual simulations of modestly large networks, and anticipate that increasing the size of both the networks and the subsets used for recall will only make the network stochastically smoother and the analysis more accurate. This anticipation does not replicate the error found in Marr [1971] (discussed in [Vogel, 2005]).

Eliminating Missing Neurons

A missing neuron is a neuron from C-C_r which is suppressed by an inhibitory projection to an activated primary neuron. The following lemma can be established.

Lemma: Each missing neuron is suppressed by an inhibitory link to a spurious neuron connected through a secondary node **w**, where **w** is different from all nodes in S_a, as shown in the Fig. 2.

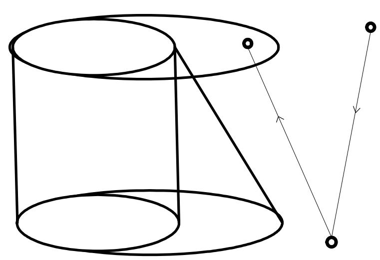

Fig. 2. Creation of missing neurons

Proof: If **m** connects to S_a, then there is a node **x** in C_r such that **x** and **m** are linked. Since m and **x** are a part of the same pattern, both parts of their link (projections from the primary node **x** to the secondary node in S_a and from the secondary node in S_a to the primary node **m**) are trained. Obviously the inhibition cannot result from such link. Therefore, the inhibitory link to **m** must pass through a secondary node outside of S_a.

To prove the argument that inhibition must come from a spurious neuron we may notice that no neuron in C_r can be connected to **w** and that if a node **x** in C is connected to **w** then m and x are connected through a trained link, since they are in the same pattern C. Therefore no other node in C can inhibit m. This leaves, as the only option that an inhibitory link to m comes from a spurious node outside of C.

The probability that a given primary node will be missing due to a single spurious node can be estimated to be less than

$$P_{missing\_one} = \frac{1}{P}(1-P_{t1}P_{t2})k_{ps}k_{sp}(1-P_{asr})P_z \quad \textbf{(11)}$$

By connecting all possible locations of missing neurons, the probability that a single primary neuron is missing is

$$P_{missing} = 1-(1-P_{missing\_one})^{P-C} \quad \textbf{(12)}$$

So the probability that the recall set has no more than S_{max} missing neurons is

$$P_{NO\_missing} = \sum_{i=0}^{S_{max}} \binom{C}{i} P_{missing}^{i}(1-P_{missing})^{C-i} \quad \textbf{(13)}$$

Results of the Statistical Model

The statistical model is in a good agreement with simulated R-nets and can be applied to estimate the computational performance of very large R-nets. We simulated the storage capacity for 20 to 100 neuron patterns of networks up to 10^7 primary neurons with 1000 projections from each primary neuron to $2x10^6$ secondary neurons with 5000 projections from each back to primary neurons. Result is shown in Fig. 3.

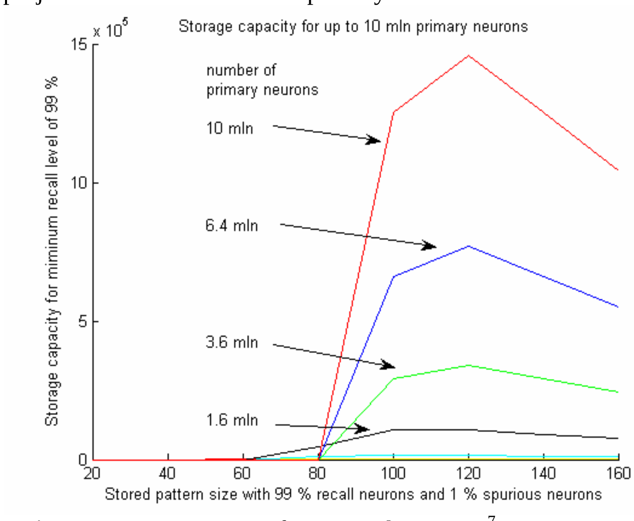

Fig. 3. Storage capacity of R-net with up to 10^7 primary neurons

In addition, from the conducted analysis of R-net properties based on the presented model, we can conclude that their storage capacity grows faster than the number of primary neurons and that a slope of growth is close to 10/7 on the logarithmic scale which agrees with experimental results reported in [Vogel, 2005]. When the network size reaches 10^9 primary neurons (with average number of projections 10^4 that is similar to interconnection density of human brain), the network can store over 10^9 patterns and the optimum storage for these very large memories is achieved with a pattern size of about 150 neurons.

Conclusion

In this paper a statistical model of R-nets was presented and results were compared to results observed in simulated R-nets. This model has already demonstrated that work on the role of disinhibition in paired-pulse induced long-term potentiation may be of fundamental importance to understanding memory and higher cognitive functions. It suggests an entirely new understanding of the role of massive projections of excitatory neurons onto neurons of distant regions [Vogel, 2005]. These models have produced computations analogous to serial memory, context dependent classical and operant conditioning, secondary reinforcement, refabrication of memory, and planning. They distinguish between perceived and recalled events, and predicate responses on the absence as well as presence of particular stimuli. Analysis suggests that the models may be expected to scale up to brain-sized networks efficiently.

References:

Freund, T.F., and Buzsáki, G. (1996). Interneurons of the Hippocampus. *Hippocampus.* 6, 347-470.

Marr, D. (1971). Simple memory: a theory for archicortex. *Philosophical Transaction of the Royal Society of London, B.* 262, 23-81.

Sik, A., Tamamaki, N., and Freund, T.F. (1993). Complete axon arborization of a single CA3 pyramidal cell in the rat hippocampus, and its relationship with postsynaptic parvalbumin-containing interneurons. European Journal of Neuroscience. 5: 1719-1728.

Vogel, D., and Boos, W. (1997). Sparsely connected, Hebbian networks with strikingly large storage capacities. *Neural Networks* 10, 671-682.

Vogel, D. (1998) Auto-associative memory produced by disinhibition in a sparsely connected network. *Neural Networks.* 11: 897-908.

Vogel, D. (2001). A biologically plausible model of associative memory which uses disinhibition rather than long term potentiation, Brain cogn. 45, 212-228.

Vogel, D. (2005). A neural network model of memory and higher cognitive functions. International Journal of Psychophysiology. 55 (1), 3-21.