International Joint Conference on Neural Networks, IV, 1992.

Correlation Matrices and the Construction of Successful Network Recodings

Dawn M. Adelsberger-Mangan [1,2] and William B Levy [2]
[1] Department of Biomedical Engineering
[2] Department of Neurosurgery
University of Virginia Health Sciences Center Box 420
Charlottesville, Virginia 22908

Abstract

Previous work in our laboratory has demonstrated the importance of the amount of synaptic connectivity and output firing thresholds, and the relative unimportance of an activity based weight modification rule, in the construction of neural networks which perform successful recodings. We define successful recodings as those which minimize information loss and maximize the decrease in statistical dependence. Here our previous results on connectivity are extended and strengthened. But in contrast to our earlier result synaptic modification is important. When degree of synaptic connectivity is high, recoding networks incur increased information loss, in spite of efforts at synaptic weight and cell firing threshold optimization. Knowledge which can be obtained from the afferent correlation matrix can be used to set synaptic weight strengths and output cell firing threshold. Networks constructed in this fashion perform near optimal recodings, losing very little of the input information while reducing the statistical dependence of the input representations. Knowledge of the optimal firing threshold eliminates the need for dynamic threshold adjusting mechanisms. Additionally, it is shown that networks constructed using an approximation in the calculation of optimal synaptic weights and output firing thresholds, perform almost as well as those constructed using the exact formulations.

Introduction

As early as 1959, Barlow proposed that a fundamental function of the successive recoding found in the sensory systems of the brain is the extraction of representations with reduced redundancy from very highly redundant input representations. As the redundancy in the neuronal layer firing patterns decreases, fewer neurons are needed to completely represent the input, so called "tight packing" (Levy, 1989).

Reduced redundancy is also desirable from a computational perspective. When redundancy is completely removed, statistical independence obtains. Probabilities of joint events A and B, P(A,B), are easily calculated when the neurons specifying events A and B are firing independently. Therefore, a recoding mechanism that produces an output with reduced redundancy is highly desirable. Statistical dependence measures the redundancy. A desirable recoding mechanism results in output representations with reduced statistical dependence. However, any recoding methodology risks information loss (Richards & Levy, 90). When the value of the lost information is unknown, any information loss must be considered worrisome. Therefore, we propose two measures for assessing the quality of a recoding mechanism: the information loss incurred during the recoding and the change in statistical dependence. During a successful recoding, information loss is minimized while the decrease in statistical dependence is maximized.

Previous work in our laboratory has illustrated the paramount importance of appropriate output firing levels in the construction of successful recodings (Adelsberger-Mangan & Levy, submitted). Using feed-forward networks composed of only summate and fire excitatory neurons, the importance of output firing thresholds and synaptic connectivity levels in the regulation of firing levels was documented. When firing thresholds are dynamically adjusted so that the neuron maintains a firing rate of approximately 50%, the optimal percentage for maximum representational capacity, there is a marked increase in the ability of the simple networks to perform a successful recoding. When firing threshold is optimized, a sparser synaptic connectivity results in a better recoding network than its more fully connected counterpart. However, an activity based mechanism of synaptic weight modification (that is, synapse weights are modified only when the postsynaptic neuron overcomes its firing threshold) was shown to confer only a very small advantage in the construction of a successful recoding network.

In the present work, we attempt to construct near optimal recoding networks incorporating and building upon the results of our previous work. In the first series of simulations, we generalize our previous results on optimal network connectivity levels. In the second series of simulations, we illustrate the desirability of a mechanism that allows individual setting of the afferent connection weights and postsynaptic neuron firing threshold. Additionally, the success of an approximation technique in the determination of optimal synaptic weights and firing thresholds is demonstrated.

Simulations

Input Layer

The are 250 binary input layer neurons. The simulation begins with the construction of the firing patterns of the input layer neurons over a 64 time step period. From these input firing patterns, an input layer correlation matrix is generated.

Two variants were employed in the generation of input firing patterns. When constructing firing patterns for the connectivity study, binary firing patterns were constructed such that nearby neurons had similar firing patterns. The firing rate of all input neurons is approximately 0.30. The mean entry in the input layer correlation matrices is 0.094 and the variance of the entries averages 0.0017. The input representations created in this fashion are highly redundant, with an average representational information content of 6.0 bits and an average statistical dependence of 211.5 bits.

The input firing patterns for the experiment series studying the desirability of individual firing thresholds are created in a slightly different fashion. Here the firing patterns of the input neurons are created independently of one another with a small variance introduced in the firing rates of the input cells. Average firing is 0.30, but the firing of individual neurons varies from 0.20 to 0.40. The mean input layer correlation matrix entry is 0.0906 and the average variance of the entries is 0.0027. The input patterns are highly redundant with an information content of 6.0 bits and an average statistical dependency of 205.5 bits.

Output Neurons

All input neurons innervate the output neurons in an excitatory fashion. The net excitation of each output neuron at each time step is determined from the simple summation of the binary activities of the input layer multiplied by the corresponding synaptic weights. If the excitation to the output neuron is greater than its firing threshold the cell fires (with an output of 1); if the excitation is less than the firing threshold, the cell does not fire (output 0).

Three distinct paradigms for determining synaptic weight and firing threshold are investigated:

Paradigm 1 is the simplest. In this paradigm all synaptic weights are set at 0.50. All output layer firing thresholds are equal and are determined by:

Synaptic Weight = 0.50 * Number of afferent to the output cells * 0.30

In paradigms 2 and 3, the correlation matrix of the afferent to a given output cell determine the output firing threshold and synaptic weights of that neuron.

In **Paradigm 2**, the vector of afferent synaptic weights is calculated from:

Synaptic weight = $(\lambda_1 e_1) / (E[X] \cdot e_1)$

where e_1 is the dominant eigenvector of the afferent correlation matrix, and λ_1 is the corresponding dominant eigenvalue. This determination of synaptic weight provides the asymptotic weight of an associatively modifiable synapse (Levy & Geman, 1982). A power method is used in the determination of e_1; λ_1 is calculated using the Rayleigh quotient.

Given this form for synaptic weight determination, the expected net excitation of each output cell is equal to λ_1, the dominant eigenvalue of the afferent correlation matrix. Therefore, by setting the output firing threshold equal to λ_1, we set the output cell to a firing rate of approximately 50%, the optimal firing rate for representational capacity.

In **Paradigm 3**, approximations of the dominant eigenvector and corresponding eigenvalue are employed. If we consider the afferent correlations matrices as approximately stochastic, we can approximate the dominant eigenvector as aligned with the vector composed entirely of ones. The corresponding eigenvalue, is approximated by the average of the sum of the elements in each row, the average row sum, of the correlation matrix. The sum $(E[X] \cdot e_1)$ is approximated by the average firing of the afferents, Avefir, multiplied by the number of afferents. Therefore, all synaptic weights on a given output cell and are equal and determined as:

Synaptic Weight = (Average row sum) / (Avefir * Number of afferents)

The firing threshold of the output neuron equals the average row sum of the afferent correlation matrix.

Afferent Selection

In the experiments studying the influence of synaptic connectivity on network performance, there were no constraints on afferent selection. Afferent sets were chosen randomly and all sets were accepted as suitable innervation to an output neuron. However, when studying the interaction between increasing afferent correlations and synaptic weight and firing threshold paradigm, afferent sets with increasing correlation were chosen by reslecting afferents which did not meet some specified level of average correlation with the other afferents to a output cell.

Network Performance Measures

The representational information, or entropy of binary layer X is:

$$H(X) = E[-\log P(X=x)] = -\sum P(X=x) \cdot \log P(X=x)$$

The transformations performed by the networks are noiseless, memoryless and deterministic. Therefore, the joint entropy H(X,Y), is equal to H(X), the entropy of the input patterns. Conditional entropy measures the information lost in recoding and is calculated:

$$H(X/Y) = H(X,Y) - H(Y) = H(X) - H(Y)$$

The statistical dependence (Stat. Depend.) of the firing patterns of layer X is determined as:

$$\text{Stat. Depend. of } X = \sum P(X=x) \cdot \log \frac{P(X=x)}{\prod P(X_i)} = \sum H(X_i) - H(X)$$

All input environments for a given experiment are created in an equivalent fashion. Therefore, they have on average equal input representational information and statistical dependence. For ease of illustration, in figures 1 and 2, output representational information (rather than information loss) and statistical dependence (rather than change in statistical dependence) are graphed.

Results

Figure 1 illustrates the deleterious effect of increasing synaptic connectivity on the recodings regardless of synaptic weight and firing threshold paradigm. For this series of experiments, afferent sets are chosen with no restriction, therefore the average row sum of the afferent correlation matrix is small, equaling 1.33. Paradigm 2, which explicitly calculates the dominant eigenvector and eigenvalue of the afferent correlation matrix, has the highest levels of output representational information and the smallest output statistical dependence, at all connectivity levels. However, at the highest connectivity levels there is a relatively small increase in performance of this methodology when compared to that of of the most trivial paradigm, Paradigm 1 which sets the synaptic weights and firing threshold uniformly over the network. Optimum selection of synaptic weight strength and firing threshold cannot compensate for connectivity levels which are too high.

At low levels of connectivity the average performance of the three paradigms is nearly identical. The best network recodings are achieved with synaptic connectivity levels of 4.8%, which corresponds to 12 afferents innervating each output cell. Networks constructed with this level of connectivity lost less than 2 % of the input layer representational information, and construct representations with average output statistical statistical dependence of only 3.9 bits.

Figure 2 illustrates the dramatic increase in performance in those paradigms which set synaptic weight and firing threshold in a neuron by neuron fashion, when compared to a paradigm in which synaptic weight and firing thresholds are set uniformly throughout the network. When more highly correlated afferents are selected, as indicated by a higher average row sum of the afferent correlation matrix, the paradigms which set synaptic weights and firing thresholds using this knowledge of this higher degree of correlation maintain a successful recoding performance. Paradigm 3, which sets a uniform synaptic weight and firing threshold based on the average statistics of the input layer, fails when afferents are more highly correlated than would on average be expected or when afferent firing rates do not equal the average firing rate of the input layer.

Networks constructed using approximations of the afferent correlation matrix dominant eigenvalue and eigenvector (Paradigm 3) very nearly match the recoding performance seen in networks constructed with the exact values (Paradigm 2). Both paradigms suffer extremely little information loss over a wide range of afferent correlations. However, networks constructed using the eigenvalue and eigenvector approximations fashion suffer from a small increase in output layer statistical dependence.

Discussion

This research is an extension of our earlier work on the construction of simple feed forward neural networks which are able to maintain the information content of the input layer while reducing the redundancy of the input representations.

Previous work (Adelsberger-Mangan & Levy, submitted) had demonstrated that when threshold is optimal, networks with sparse connectivity are better at constructing recodings than their more fully connected counterparts. This observation is generalized with the results of this report; judicious setting of synaptic weight and firing threshold cannot overcome the deleterious effects of an over-connected network.

When networks are constructed with proper connectivity levels the use of afferent correlation matrix knowledge provides appropriate synaptic weight strengths and output neuron firing thresholds for constructing networks which perform successful recodings. This knowledge allows the construction of very simple networks composed of excitatory interactions only and with no need for dynamic firing threshold regulation to maintain output firing rates at optimal levels. By using afferent correlation matrix knowledge, successful recodings can be performed over a wide range of afferent coactivity.

The use of and approximation method for determining afferent correlation matrix statistics was extremely effective in constructing recoding networks. Over a greater than two fold range in afferent correlation matrix average row sum, the approximation methodology maintained an average information loss of only 1.5%.

Acknowlegments

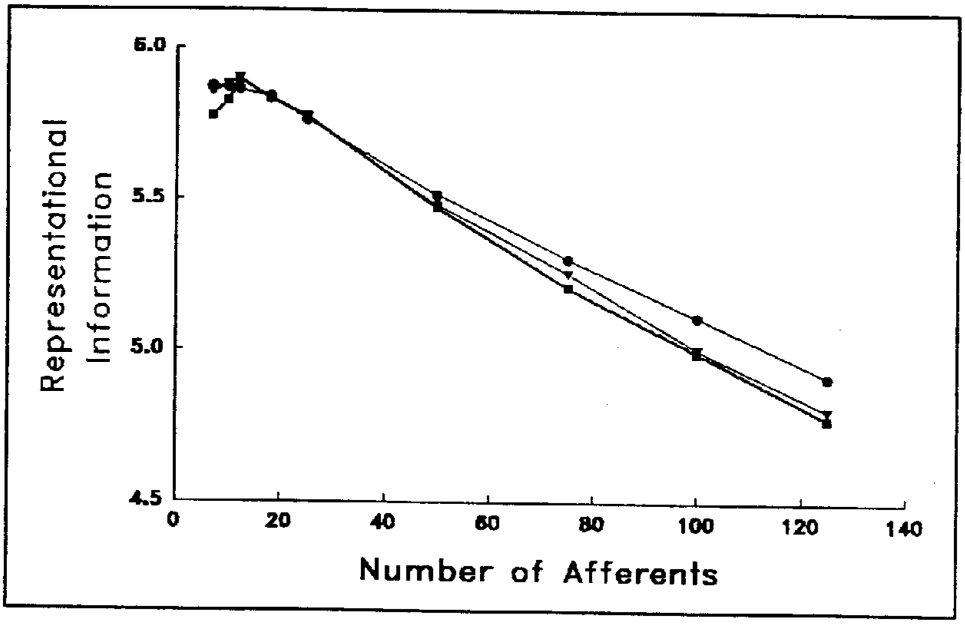
This research was supported by NIH NS15488 and NIMH MH48161 to WBL, and by NIMH MH09967 to DMA, and by the Department of Neurological Surgery, Dr. John A. Jane, Chairman.

References

Barlow, H. B. Sensory mechanisms, the reduction of redundancy and intelligence, In: Mechanisation of Thought Processes, Vol. II, National Physical Laboratory, Symposium No. 10, 1959, 535-560.

Levy, W. B A computational approach to hippocampal function. In: Computational Models of Learning in Simple Neural Systems. (R. D. Hawkins and G. H. Bower, Eds.), New York: Academic Press, pp. 243-305, 1989.

Richards, D. St.P. and Levy, W.B Optimal preprocessing networks and a data processing theorem. International Joint Conference on Neural Networks, 1990, I, 19-22.



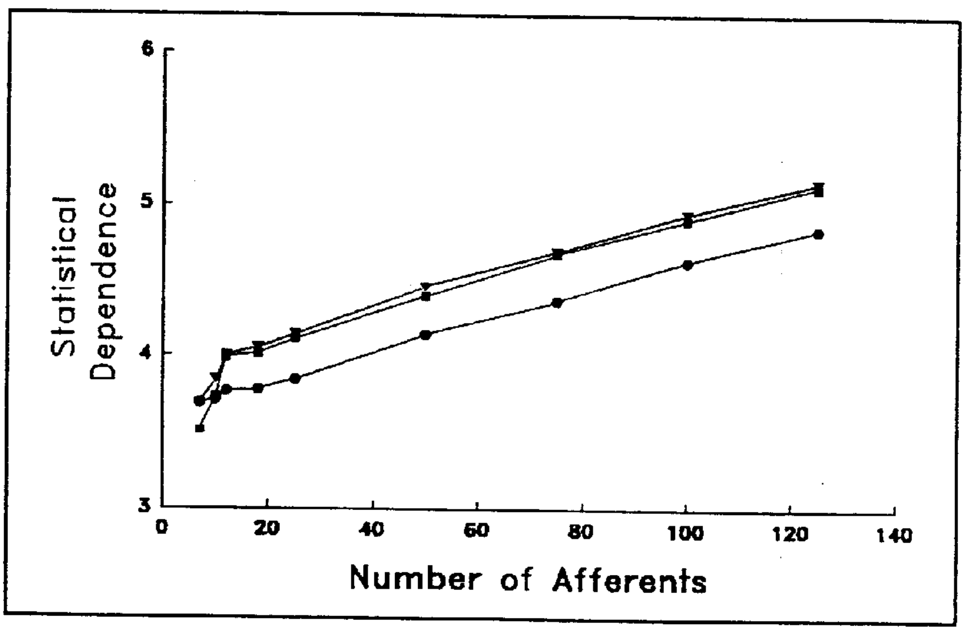
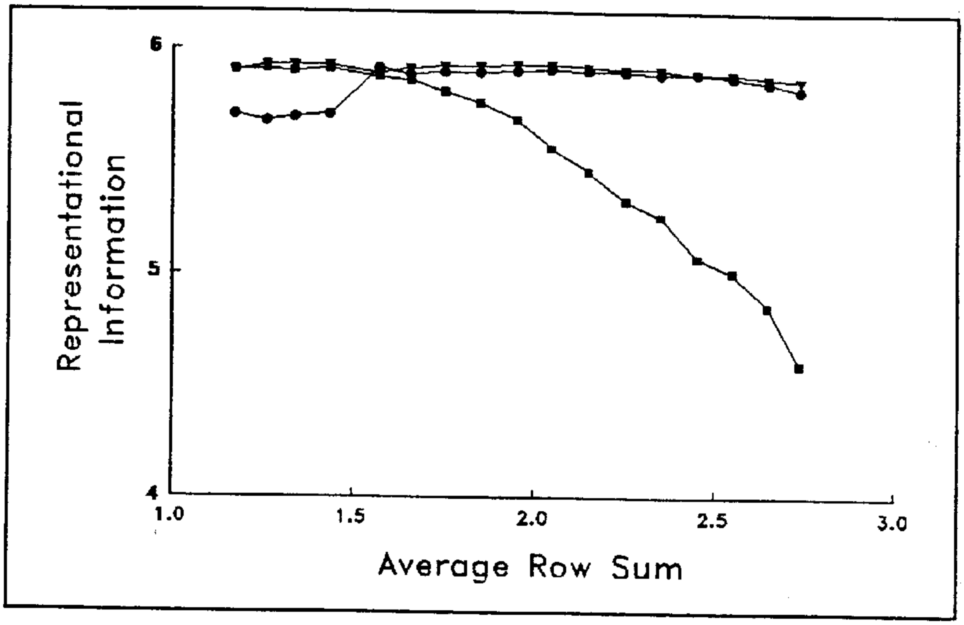


Figure 1
The effect of increasing number of afferents on the information measures. Paradigm 1 is represented as ■, paradigm 2, as ●, and Paradigm 3 as ▼.



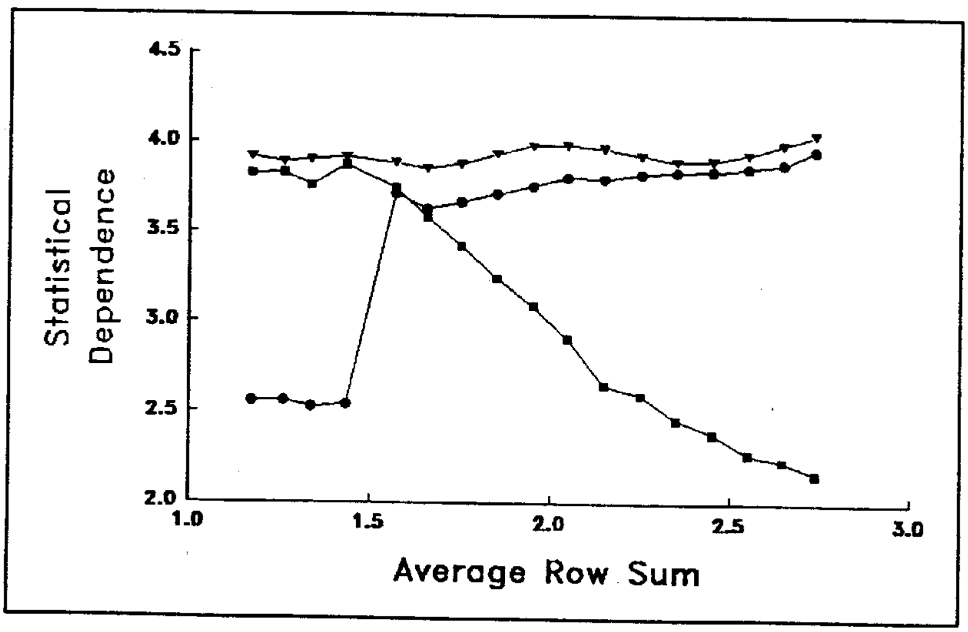


Figure 2
The effect of increasing average afferent correlation matrix row sum on the information measures. Paradigm 1 is represented as ■, paradigm 2 as ●, and Paradigm 3 as ▼.