TR#1206

Competition and Cooperation in Neuronal Processing

with application to associative memory

Haim Bar[1], Alexander Ambash[2], Willard Miranker[3]
Department of Computer Science, Yale University
New Haven, CT 06520

Abstract

A new type of model neuron is introduced as a building block of an associative memory. The neuron, which has a number of receptor zones, processes both the amplitude and the frequency of input signals, associating a small number of features encoded by those signals. Using this two-parameter input in our model compared to the one-dimensional inputs of conventional model neurons (e.g. the McCulloch-Pitts neuron) offers an increased memory capacity. In our model there is a competition among inputs in each zone with a subsequent cooperation of the winners to specify the output. The associative memory consists of a network of such neurons. A state-space model is used to define the neurodynamics. We explore properties of the neuron and the network and demonstrate its favorable capacity and recall capabilities. Finally, the network is used in an application designed to find trademarks that sound alike.

Key words - associative memory, neural networks, competitive cooperative neuron, CCN, trademarks.

1. INTRODUCTION

Conventional model neurons such as the McCulloch-Pitts neuron use one-dimensional input from each source (the neuronal activity). These neurons themselves are

[1] haim.bar@yale.edu
[2] alexander.ambash@yale.edu
[3] miranker@cs.yale.edu

characterized by a one-dimensional parameter set (the synaptic weights). In this paper we propose a new model neuron which more closely resembles the biological neuron in structure and functionality. The new neuron employs a competition among inputs to each of its input zone and a subsequent cooperation among the winners in each of the zones for specifying the output. For convenience, hereafter we refer to this new model neuron as CCN (Competitive Cooperative Neuron).

There are two critical aspects to the new model neuron. One is that the input signals are characterized by a two-dimensional parameter set (representing the amplitude and the frequency of signals). The other critical aspect of the CCN is that the neuron receives input signals at several distinct and autonomous zones. In each zone there is a competition among the inputs with a winner-takes-all protocol. The winning signal of each zone is passed along to the cell body. There, a threshold protocol is employed for determining neuronal firing and consequent association of inputs at the different zones.

The following diagrams illustrate the conventional and new model neurons, and the way they differ (in Figure 2 we show a CCN with five receptor zones).

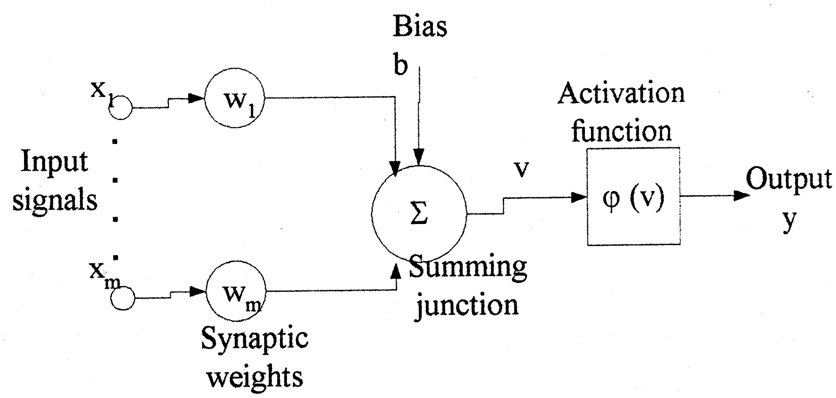

Figure 1: Conventional model neuron

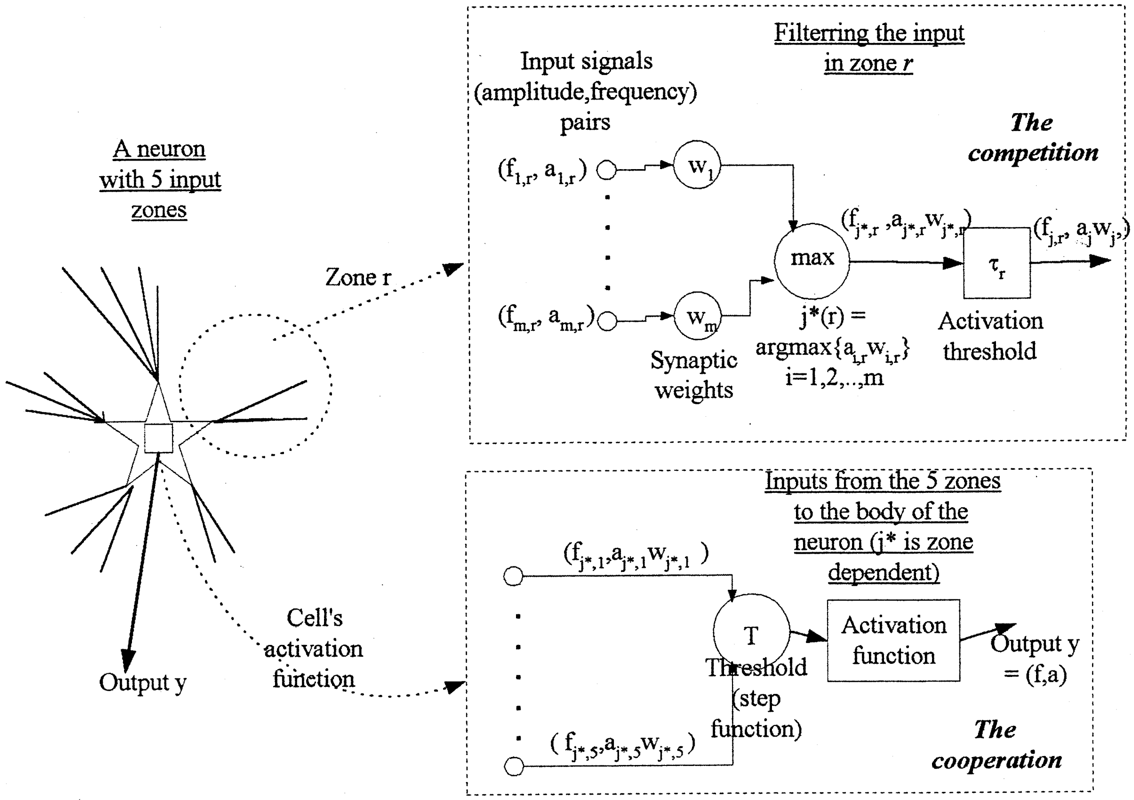

Figure 2: New model neuron (CCN)

The CCN resembles the pyramidal cell found in brain {Arbib, 1972}. It has a physical structure resembling a pyramidal cell, and like such a cell it processes both the frequency and the amplitude of the input signals. We claim that a cognitive neuronal system which uses the CCN has higher memory capacity compared to a McCulloch-Pitts neural network (i.e., the network employing CCN has a higher stored-features/number-of-neurons ratio). We show that the improved capacity is a result of the two new aspects (two-dimensional characterization of signals and autonomous receptor zones), as well, of course, of the structural and operational details of the new model.

In Section 2 we describe CCN in detail. In Section 3 we describe a simple network of such neurons, and in Section 4 we introduce mathematical notation to formalize the description of the network in order to analyze its capabilities. The neurodynamics which affects the neuronal parameters so that learning, association, and processing efficiency are achieved, are presented in Section 5. In Section 6 we describe the training and recall processes of the network. In Section 7 we develop some properties of the neuron and the network including an expression for the memory capacity. We investigate the capacity of the memory and explain how phantom memories can emerge. Finally, in Section 8 we describe a practical application and the results of an experiment in which we train the network to associate between short strings that may sound alike. This application is useful for companies when they register a new trademark. For a trademark to be considered new, it must not sound like any existing one. In particular, when naming a new drug, it is very important to avoid confusion with other drugs.

2. THE NEURON – GENERAL DESCRIPTION

The Competitive Cooperative Neuron consists of a small number of input regions (receptor zones[4]), and one output (axon). Each input region collects input signals from many sources (dendrites). The input signal from each source has two aspects – the frequency (which encodes the information), and the amplitude (the strength of the signal). Each region is especially sensitive to a small range of frequencies (band). The center of this band[5] is chosen randomly at first. After each attempt to learn a specific task (memory), a band that is sufficiently close to the selected (winning) input signal (in a sense made precise in section 4) is preserved when the neuron fires. This resembles natural selection, where the fittest configuration is genetically preserved. If a selected band did not contribute to the improvement in learning the task, a new random value of the band center may be assigned.

The input to a receptor zone is built up by the superposition of the frequencies from the different sources (dendrites) to that zone. A receptor zone of a neuron decomposes the input and detects only the frequencies that are within its band, but only if the amplitude of the input corresponding to this frequency exceeds a certain threshold. These thresholds typically change over time in the process of acquiring memories so that the performance of the network is improved.

Note that the same numerical input frequency value (which encodes the information) in different regions may represent different modalities. For example, if a

[4] We shall use the terms (receptor) zone and (input) region interchangeably.
[5] From now on, when we use the term "band" we mean a small, specified range of frequencies around the center of the band.

neuron associates between names of foods and their taste, then the word "pepper" may be encoded just like (i.e. with the same frequency as) the taste "sweet". Since the inputs arrive from different regions, there is no confusion in the processing. Receiving inputs from distinct zones allows re-use of frequencies, and therefore increases the capacity, as we shall discuss later.

Each input region propagates the winning input signal (amplitude and frequency) to the cell body. The precise nature of the competition determining the winning input is specified in Section 4. The neuron fires if the combined amplitude of the winning input signals exceeds a certain threshold. If the neuron fires the output is arbitrary at first. The neurons that receive this output may or may not detect it. After a few generations, nature (or in the artificial network case, the designer of the network) may create cells which have adapted so that they can detect this output signal. Alternatively the firing neuron can change its output protocol by means of Hebbian learning. In this paper, we shall define the output to be a vector chosen appropriately from training data (exemplars). The components of the output vector are the frequencies of the winning inputs to the neuron. As we shall see, this approach will simplify the recall process and the implementation of the network.

When activated, neurons can decrease the tolerance level in the contributing zones (i.e. zones that have furnished a winning input signal), where the tolerance level is the maximum value of the difference between the band and the winning input frequency that still results in the neuron firing. When a neuron decreases its tolerance levels we say that it specializes. We claim that this feature results in an increased capacity of the network (Section 7). During training, the neuron also decreases the threshold for the amplitude in active zones, and if the neuron fires it also decreases the threshold of the cell body. This allows for a clean, but weaker signal to activate a zone and even cause the neuron to fire.

When a zone is activated by an input signal, it remains active for a short period of time, during which it ignores subsequent inputs. This is one more aspect in which the CCN and the biological neuron resemble one another. This means that the input signals that are required to activate the CCN don't have to be tightly synchronized. From the moment the first receptor zone is activated, the neuron waits for the other signals for a short period of time.

When the neuron is trained and it receives input in some, but not all the regions, it uses that input to recall the previous inputs to the idle regions. For example, if a neuron has three receptor zones, and it fired when the input was the vector that represents the triple (Red, Sweet, Strawberry), then the next time it receives only "Red" and no input from the other regions, it will fire ("Red", "Sweet", "Strawberry"), provided that the amplitude of the input signal ("Red") exceeds the cell's threshold.

In Section 4, we shall introduce mathematical notation to formalize the descriptive representation of our neuronal model.

3. THE NETWORK

The basic implementation uses a feed-forward network with one layer of neurons. The neurons are as described in Section 2. Figure 3 illustrates a simple one-layer feed-forward network with four neurons and three input sources. Each neuron has three receptor zones.

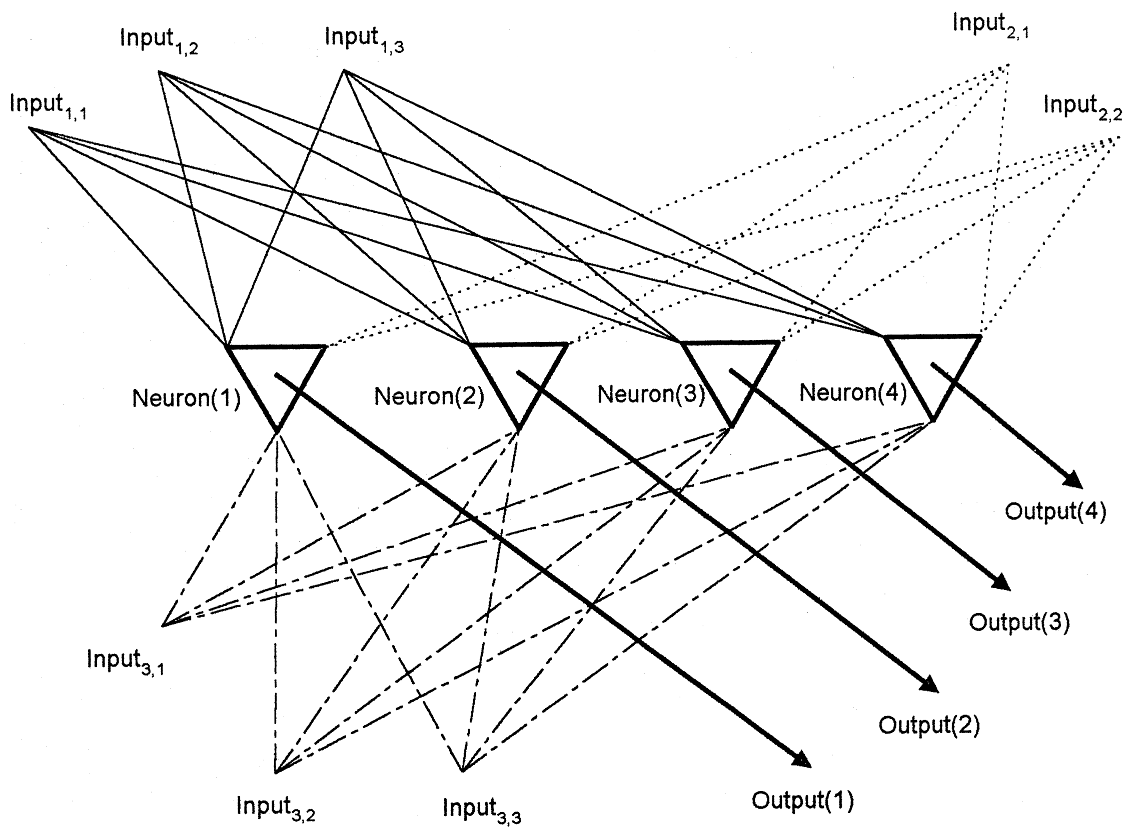

Figure 3: A one-layer feed-forward network with four neurons (CCN)

In future work, we shall use a recurrent network. In that case the vector output of each of the neurons is fed back to that neuron's input lines, where each element of the output vector is fed back only to its corresponding input region. We expect that the recurrent network will be able to recall higher-order memories. For example, if the input "Red" is associated with "Sweet" and "Strawberry", then feeding "Sweet" back to the appropriate input region may result in recalling "White" and "Sugar".

It is also possible to use a multi-layer feed-forward configuration to achieve recall of higher order memories. However this involves a more complicated encoding/decoding scheme for the output of the neurons, which we also defer for later study.

4. THE MATHEMATICAL MODEL

Let N denote the number of neurons. $R(n)$ denotes the number of receptor regions in neuron n. For simplicity we assume that all neurons have the same number of receptor regions, i.e. $R(n) = R$ for $n = 1,2,\ldots,N$.

Receptor region r in *each* neuron receives a finite set of signals, $\mathrm{S}(r) = \{S_1(r), S_2(r), S_3(r),\ldots\}$ (some of which may be null). Note that $\mathrm{S}(r)$ is not a function of the neuron number, n, since in every neuron, the r-th region receives the same set of signals $\mathrm{S}(r)$ (see figure 3). $S_i(r)$, the i-th input to the r-th region of each neuron is a vector: $S_i(r) = (F_i(r), A_i(r))$ where $F_i(r)$ is the frequency and $A_i(r)$ is the amplitude of the input signal $S_i(r)$.

As already noted, R (the number of receptor regions) can take small values (typically 2-5), and $|\mathrm{S}(r)|$ (which denotes the number of input signals) is a larger number, $O(10^3)$-$O(10^5)$ in the human brain.

The totality of possible input frequencies is bounded. Then, without loss of generality, we shall normalize the range of input frequencies $F_i(r)$ so that $F_i(r) \in [0,1]$. The center of the band of input region r of neuron n at time t is denoted by $B(n,r,t)$. The tolerance level of region r in neuron n is denoted by $T(n,r,t)$. $T(n,r,t)$ is a small

We start with a neuron body threshold, $\nu(n,t)$, that is greater than or equal to the sum of the region thresholds, i.e. $\nu(n,t) \geq \sum_r \tau(n,r,t)$, so that in order for the neuron to fire, either all the zones must be active, or some of them must receive a very strong input signal.

5. NEURODYNAMICS

We use a state-space model as defined in Haykin, 1999, p 666. The states are vectors $(B(n,r,t), T(n,r,t), \tau(n,r,t), \rho(n,r,t))$ that describe the state of a receptor zone r in neuron n at time t. We also have to define the dynamics for the CCN body threshold $\nu(n,t)$.

Recall that region r (of every neuron) receives a finite set of signals $\mathrm{S}(r) = \{S_1(r), S_2(r), S_3(r), \ldots\}$ such that $S_i(r) = (F_i(r), A_i(r))$ where $F_i(r)$ is the frequency and $A_i(r)$ is the amplitude of the i-th input signal $S_i(r)$.

If a CCN fires, the active receptor zones reward the winning inputs by increasing their effective amplitude and decreasing the effective amplitude of the other input signals. That is, the change to reduction factors when the neuron fires is:

$$\rho_i(n,r,t+dt) = \begin{cases} \alpha + (1-\alpha)\rho_i(n,r,t), & \textit{if } i \textit{ is the winning input} \\ (1-\alpha)\rho_i(n,r,t), & \textit{otherwise} \end{cases} \tag{2}$$

where $\alpha \in (0,1)$ is the reduction factor change rate. Note that the first alternative in (2) increases ρ while the second decreases it.

When a CCN fires the active receptor zones also decrease the amplitude threshold of active zones:

$$\tau(n,r,t+dt) = \max\{\beta\tau(n,r,t), \tau_{\min}\} \tag{3}$$

where $\beta \in (0,1)$ is the receptor zone threshold change rate, and $\tau_{\min} > 0$ is the minimum threshold that is required to activate a zone.

Similarly, if the CCN fires the tolerance level of active zones is decreased. If a neuron doesn't fire the tolerance level of inactive zones is increased (anti-Hebbian learning):

If neuron n fires and zone r is active:

positive number, typically $T(n,r,t) \leq 0.05$. In Section 5 we describe how the center of the band $B(n,r,t)$ and the tolerance level $T(n,r,t)$ change over time (the neurodynamics).

We also normalize the amplitude $A_i(r)$ so that $A_i(r) \in [0,1]$. Each receptor region can reduce the effective amplitude of the i-th input signal by a factor $\rho_i(n,r,t) \in [0,1]$, where ρ is a function of n (the neuron number), r (the region), and t (the time). We shall call ρ the (amplitude) reduction factor. We assume that there is a threshold $\tau(n,r,t) \in [0,1]$, so that a region can detect an input signal $S_i(r)$ only if $\rho_i(n,r,t)A_i(r) \geq \tau(n,r,t) > 0$. The product $\rho_i(n,r,t)A_i(r)$ will be termed the effective input amplitude.

We say that a region r in neuron n is *active* at time t if for some input $S_i(r)$, we have $F_i(r) \in [B(n,r,t) - T(n,r,t), B(n,r,t) + T(n,r,t)]$ and $\rho_i(n,r,t)A_i(r) \geq \tau(n,r,t)$. In words, the region is active if there is an input signal with effective amplitude that exceeds the positive threshold and with a frequency that is appropriately close to the center of the band of that region.

Let $I(n,r,t)$ denote those values of i such that $\rho_i(n,r,t)A_i(r) \geq \tau(n,r,t)$ and $F_i(r) \in [B(n,r,t) - T(n,r,t), B(n,r,t) + T(n,r,t)]$ at time t. $\{S_i(r)\}_{i \in I(n,r,t)}$ is the set of input signals which activate the r-th region in the n-th neuron at time t. Let $i^* = i^*(n,r,t) = \arg\max_{i \in I(n,r,t)} [\rho_i(n,r,t)A_i(r)]$ (i.e. choose the input signal from the set $\{S_i(r)\}_{i \in I(n,r,t)}$ with the greatest effective amplitude). We say that $W(n,r,t) = (F_{i^*}(r), A_{i^*}(r))$ is the *winning* input of region r of neuron n at time t. In the case that $I = \phi$, there is no winning input, and we take $W(n,r,t) = (-1,0)$.

Neuron n fires if the sum of the effective amplitudes of the winning inputs exceeds the neuron body threshold $\nu(n,t)$. That is, neuron n fires at time t if

$$\sum_{r=1}^{R} \rho_{i^*} A_{i^*}(r) \geq \nu(n,t). \qquad (1)$$

$$T(n,r,t+dt)=\max\{\gamma T(n,r,t),T_{\min}\} \tag{4}$$

and if neuron *n* doesn't fire and zone *r* is inactive:

$$T(n,r,t+dt)=\min\{(1+\gamma)T(n,r,t),T_{\max}\} \tag{5}$$

where $\gamma\in(0,1)$ is the tolerance level change rate and T_{min}, T_{max} are the minimum and maximum values that the tolerance level can take, respectively.

When a CCN fires it decreases the amplitude threshold of the neuron body:

$$v(n,r,t+dt)=\max\{\delta v(n,r,t),v_{\min}\} \tag{6}$$

where $\delta\in(0,1)$ is the neuron body threshold change rate, and $v_{min}>0$ is the minimum threshold that is required to activate the neuron.

Finally, when a neuron fires, the center of the band of an active zone takes the value of the frequency of the winning input:

$$B(n,r,t) = F_{i^*}(r) \tag{7}$$

The neurodynamics is Hebbian since positive correlation between an input region's activity and the neurons firing is rewarded, while anti-correlation results in diminished sensitivity to input signals.

6. TRAINING AND RECALL

To train the network, we stimulate it with the inputs that we want it to memorize. A "*memory*" is an *R*-dimensional feature vector $M=(m(1),m(2),\ldots,m(R))\in[0,1]^R$. Our goal is to create a network that associates each of the features $m(r)$ with the other features, $m(r'),\ for\ r'\neq r$. We denote the set of memories by $\mathbf{M}$. For every memory $M_j\in\mathbf{M}$ where $M_j=(m_j(1),m_j(2),\ldots,m_j(R))\in[0,1]^R$ we define the inputs to each region as follows:

$$S_i(r)=\begin{Bmatrix}(m_j(r),1) & i=1\\ (0,0) & i\neq 1\end{Bmatrix},\ i=1,2,\ldots,|\mathbf{S}(r)|$$

So in the training phase we eliminate noise by defining $S_i(r)=(0,0)$ for $i\neq 1$. The neurons in the network receive only the value $m_j(r)$ in region *r* since it is the only non-null signal. Therefore, there is only one signal that could possibly activate region *r* in each neuron (that signal is $S_1(r)$ according to our definition). We set the threshold of each zone to be $\tau(n,r,t)=\frac{1}{R}$ and the neuron body threshold to 1. Since the non-null amplitudes are unity, we arrange that $\rho(n,r,t)=\frac{1}{R}$ so that $\rho(n,r,t)A_i(r)\geq\tau(n,r,t)$ in the training phase. Given these thresholds the amplitude of each input signal is sufficient to activate an input zone, however a neuron can be activated only if all of its zones are active. Note that this is a form of supervised learning since in this training scheme we eliminated noise from all but one input source of each zone.

A training step consists of a single input $M_j \in \mathbf{M}$ that stimulates the neurons and the adjustment to the bands and the reduction factors associated with that input. We expect that after a number of training steps some neurons will be "attracted" to some memories and fire when such a memory is introduced to the network. That is to say, for some memory $M_j \in \mathbf{M}$, there will be a neuron n such that

$$\max_{r=1,2,\ldots R}\left\{\left|B(n,r,t)-m_j(r)\right|\right\}\leq T(n,r,t) \tag{8}$$

The training stops when the condition in (8) is satisfied for all the memories $M_j \in \mathbf{M}$, or after a specified maximum number of iterations. A trained network can recall a memory $M_j \in \mathbf{M}$ if there exists a neuron n such that $\max_{r=1,2,\ldots R}\left\{\left|B(n,r,t)-m_j(r)\right|\right\}<T(n,r,t)$.

When a neuron fires, i.e. when $\max_{r=1,2,\ldots R}\left\{\left|B(n,r,t)-m_j(r)\right|\right\}<T(n,r,t)$, each active zone in that neuron decreases the tolerance level $T(n,r,t)$. In other words, when a neuron can recall a memory in $\mathbf{M}$, each region in that neuron will become less tolerant to deviations from the band center of that region, $B(n,r,t)$. In Section 7 we will show that a smaller tolerance level in all the neurons results in increased memory capacity of the network.

In the recall phase, the activation of a single region is sufficient to make the neuron fire. That happens when the neuron body threshold becomes sufficiently small so that the amplitude of an input to one of its zones is greater than the neuron body threshold. An input that activates a single region, results in R recalled features ($R-1$, if we exclude the activating input).

7. PROPERTIES OF THE MODEL

Recall our assumption that all neurons have the same number of input regions (R). We define $T = \max_{n,r}\{T(n,r,t)\}$, the largest tolerance level among all the regions of all the neurons. We assume that the r-th band center in neuron n, $B(n,r,t)$ and the frequency $F_W(n,r,t)$ of winning input $W(n,r,t)$ are independent and identically distributed with a uniform distribution over $[0,1]$. In Appendix B we show that

$$prob\left[\left|F_W(n,r,t)-B(n,r,t)\right|<T\right]=2T-T^2 \tag{9}$$

As a result of independence, we get that

$$prob\left[\max_{r=1,2,\dots,R}\left\{\left|F_W(n,r,t)-B(n,r,t)\right|\right\}\leq T\right]=\left(2T-T^2\right)^R \tag{10}$$

So $\left(2T-T^2\right)^R$ is the probability that a neuron will fire, given a specific band configuration, tolerance level, threshold and an input chosen at random.

The memory capacity of the network depends on T and R (equation (10)). Perfect recall cannot occur if two different input signals to a region activate that region. A necessary condition for perfect recall of any of the input frequencies to region r is $\left|F_i(r)-F_j(r)\right|>2T$, $\forall i,j=1,2,\dots,\left|\mathbf{S}(r)\right|$ (recall that $\left|\mathbf{S}(r)\right|$ is the number of all the distinct inputs to region r in all the neurons). To see this note that if there are i,j such that $\left|F_i(r)-F_j(r)\right|<2T$, then there may be a neuron n with a band $B(n,r,t)$ in region r such that $\left|B(n,r,t)-F_i(r)\right|<T$ and $\left|B(n,r,t)-F_j(r)\right|<T$. In this case two *different* input signals ($S_i(r), S_j(r)$) will activate the same region in the n-th neuron.

Since $F_i(r)\in[0,1]$, the condition $\left|F_i(r)-F_j(r)\right|>2T$, $\forall i,j=1,2,\dots,\left|\mathbf{S}(r)\right|$ assures that $1/2T$ input signals (frequencies) can be detected in region r.

The frequency re-use mentioned in Section 2 increases the capacity of each cell. The number of distinct signals that each cell can receive is $R/2T$. Each input signal to a single region will result in $R-1$ recalled features from the other regions, which is more efficient than recalling only one feature from every input (compared with the correlation matrix memory {Haykin, 1999, p 82}).

When the network is trained, it will need only $1/(2T)$ neurons to recall any one of the $R/(2T)$ memories. The correlation matrix memory requires $R/(2T)$ neurons to recall $R/(2T)$ memories. This improved performance of our network memory is a result of the frequency re-use. To get perfect recall of $R/(2T)$ memories using only $1/(2T)$ neurons we have to start with at least $1/(2T)$ neurons (two or more neurons can be activated by the same input, so having $1/(2T)$ neurons may not be sufficient to recall all the memories).

When the neurons specialize ($T(n,r,t)$ are decreased), the capacity grows, allowing more inputs to be associated and then recalled. This follows directly from the fact just observed that the memory capacity is proportional to the inverse of T.

Decreasing the effective amplitude of a winning input and increasing the tolerance level in a region when a neuron doesn't fire may also increase the capacity. The band of that region, $[B(n,r,t)-T(n,r,t), B(n,r,t)+T(n,r,t)]$ gets wider and allows for more input frequencies to be detected (see 5). When the effective amplitude of the winning input frequency becomes small enough, a new input with a different frequency could become the winner. The neuron may then be able to store and recall a different memory.

Note: phantom recalls may result were we to encode the winning input frequency as $F_W(n,r,t) \in [0,1]$ instead of $F_W(n,r,t) = -1$ when there is no input to region r, i.e. recalled memories without any input. This observation is a result of condition (1) in section 4. In the case where $B(n,r,t) \in [0, T(n,r,t))$ the region r will be activated when there is no input.

8. AN APPLICATION: TRADEMARKS

In our experiment we use the network of CCN to find homonyms, words that are spelled differently, but sound alike. The ability to detect such words is very important for companies seeking to register a new trademark. One criterion for a trademark to be considered new is that it doesn't sound like any existing one. Using another company's trademark may result in great liability. Trademarks are not subject to spelling rules. For example, "toysRus" sounds like "toys are us", but the former cannot be found in a dictionary. Handling a phonetic encoding of this sort is different from handling text-to-speech. The latter can benefit from using a database of dictionary words, which makes the pronunciation task somewhat easier. Building a trademark dictionary is a plausible idea, but since hundreds of new words are invented every week, the size of the database will increase, which means that maintaining it is difficult.

During the training phase the network is presented with a list of pairs (rules) of one or two letter strings which sound alike. (The pairs defining rules, all of which are listed in Appendix A, should not be confused with the (frequency,amplitude) pairs introduced earlier.) The objective is to train the network to associate between elements comprising the pairs in the list. For example, "MN" is associated with "N", "GH" is associated with "F", and "PH" is also associated with "F". In most cases we treat all the vowels as if they were the same input. Appendix A contains the complete list of pairs. We shall refer to these pairs as atomic homonyms. We encode every possible input as described in Section 4. For example, the pair of strings (GH,F), an atomic homonym, from Appendix A may be encoded as the two frequency/amplitude pairs (0.313,1), (0.724,1). In the present example there is no competition, so for simplicity we omit the amplitude from now on and denote the encoding of (GH,F) by a frequency pair (0.313, 0.724). Similarly, the encoding of the pair, the atomic homonym (PH, F) is taken as (0.543, 0.724). Assigning frequencies in [0,1] must be done without repetition. We submit all the input pairs to the network, and repeat this step until every input pair has a neuron that fires when it receives this input pair. Then, as we described in Section 2, such a neuron can associate between the two input values, i.e. between strings of each pair from Appendix A.

In the next step, a test word is partitioned into one or two letter sub-strings. For example, the test word "ROCK" will be partitioned in the following ways: R-O-C-K, R-O-CK, R-OC-K, RO-C-K and RO-CK (these are the valid partitions of the word ROCK). Then for every partition of the word, we submit the partitioned components to the network sequentially (see Section 6). If the network associates the input with other strings, it outputs the other strings. Otherwise, it outputs the input string. For example, if the input is R-O-CK, the network will recall (i.e. output) R-A-K (where A is the representation of any vowel). This may be seen from the table in appendix A, because R does not occur in any pair on the left hand side, O is a vowel, and the pair (CK,K) occurs in the table.

At the end of the process we get several outputs which we call "tokens" that are associated with the input (the tokens are strings that are pronounced like the input). Trademarks that sound alike should have a mutual token. For example, it is easy to see that ROCK, ROK, and ROCQ are all associated with the token RAK.

In our experiment we assigned a different random value (frequency encoding) to each component of an atomic homonym that appears in appendix A. We employed a set of N=50 neurons, each having R=2 input regions. We trained the network by submitting 10 input pairs from the list in Appendix A. All the neurons were initialized with the same tolerance level $T(n,r,t)=0.05$ which is smaller than $\min_{i,j}\left|F_i(r)-F_j(r)\right|$ where the minimum is taken over all the input frequencies to region r at t=0. Amplitudes were encoded with a default value A=0.5 and the neuron body activation threshold was set to 1. We trained the network by inputting the input vectors $(F_i(1)+\sigma_1, F_i(2)+\sigma_2)$, $i=1,2,\ldots 10$ corresponding to the rules in Appendix A sequentially. Note that robustness of the processing is enhanced if the input signals have small variations σ_1, σ_2 which are random variables with uniform distribution in the range [-0.005,0.005]. We submitted the first input pair 3 times, then the second input pair 3 times, and so forth. We repeated this process for all 10 input pairs 500 times.

We now list a few observations from our experiment. We only focus on the functionality of the neurons and not on the partitioning of words as described earlier in this section, or any other high-level processing. Recall that in Section 2 we described the recovery time as one of the properties of CCN (as well as of the biological neuron). That is to say, when a receptive zone is activated by an input signal, it remains active for a short period of time to allow other zones to be activated too, and therefore, to permit association between the incoming signals. However, this feature may result in "false recalls". To demonstrate this, consider a network with two CCNs. The first neuron's band centers are 0.1 and 0.9. The second neuron's band centers are 0.3 and 0.7. Suppose also that we train that network by inputting two memories, which are encoded by the frequency pairs M_1=(0.1, 0.2) representing the homonyms 'gh' and 'f', and M_2=(0.5, 0.9) representing the homonyms 'mn' and 'n'. When the neurons receive M_1, zone 1 of neuron 1 becomes active[6]. If the neurons receive M_2 while zone 1 of neuron 1 is still active, neuron 1 will fire, since its two input regions become active at the same time. As a result, neuron 1 will now associate the first feature of M_1 with the second feature of M_2, namely it will associate 'gh' with 'n' (i.e., false association). To train a network of CCNs properly we should repeat each input several times and let it "rest" between different inputs.

The following graphs demonstrate changes to the bands of neurons as a result of the training process. The inputs are displayed as plus signs and the bands by solid lines. Figure 4 shows that the band of zone 1 of neuron 1 was initially [0.118, 0.218] and the band of zone 2 of that neuron was initially [0.858, 0.958]. Some of the inputs were in these ranges and therefore activated the two zones of neuron 1, causing the neuron to start firing. Then, according to the neurodynamics (Section 5) the two bands started to narrow every time the activating input was introduced to the network until the bandwidth was as wide as the variation of the input frequency (0.005).

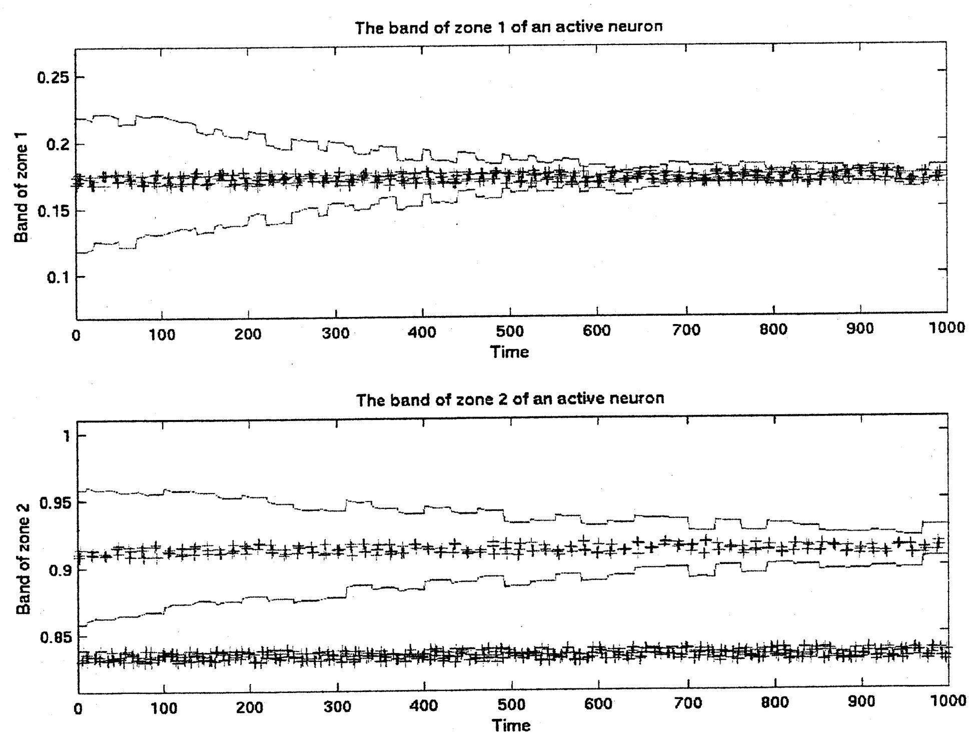

Figure 4: The change in the bands of neuron 1 during training

[6] This zone becomes active because the left frequency 0.1 in M_1 clearly lies in the band of zone 1 of neuron 1 which is centered at 0.1.

Figure 5 shows that the band of zone 1 of neuron 3 was initially [0.127, 0.227] and the band of zone 2 of that neuron was initially [0.222, 0.322]. During the first 300 time cycles zone 1 was activated when the input 0.175 was introduced, but since zone 2 remained inactive, the neuron didn't fire and the bandwidth of zone 2 got wider (see Section 5, equations 4 and 5). At approximately 300 time cycles, the bandwidth was wide enough to detect the input frequency 0.335. At this point, both zones became active every time the input pair (0.175, 0.335) was introduced to the network and the neuron started firing shortly thereafter. When the neuron started to fire, the bands of its two zones narrowed.

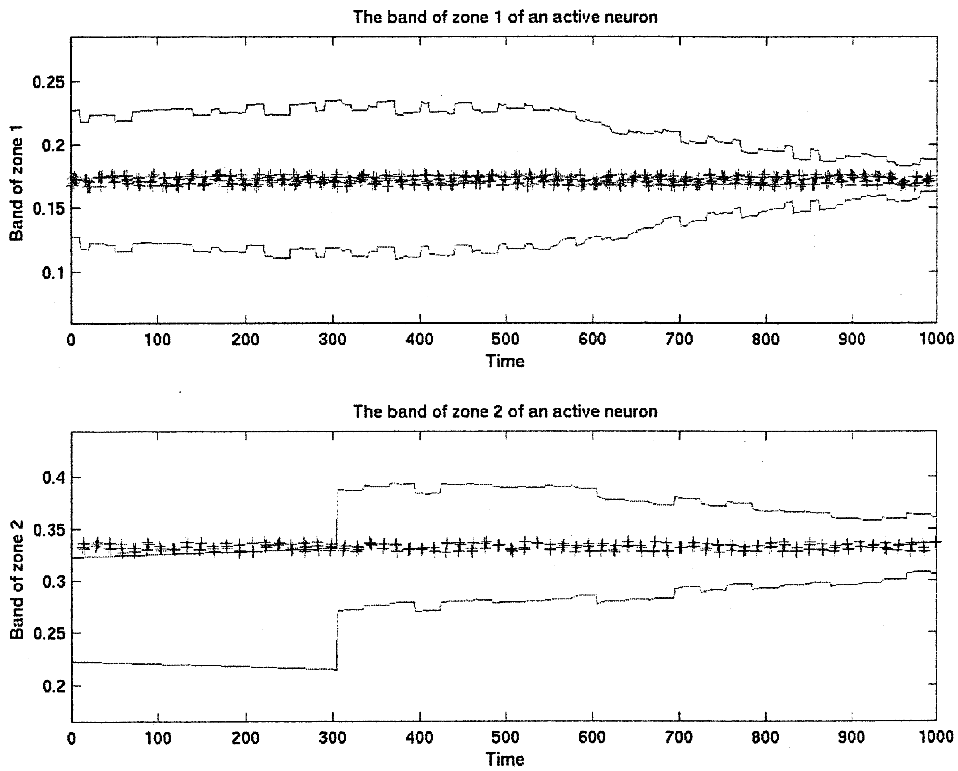

Figure 5: The changes in the bands of neuron 3 during training

Finally, Figure 6 shows that neither zone of neuron 6 became active for more than 650 time cycles, so both bands got wider. After almost 700 time cycles the band of zone 1 was wide enough to detect an input signal. In contrast, the width of the band of zone 2 reached its maximum value, but no input signal to that zone was ever in the band of zone 2.

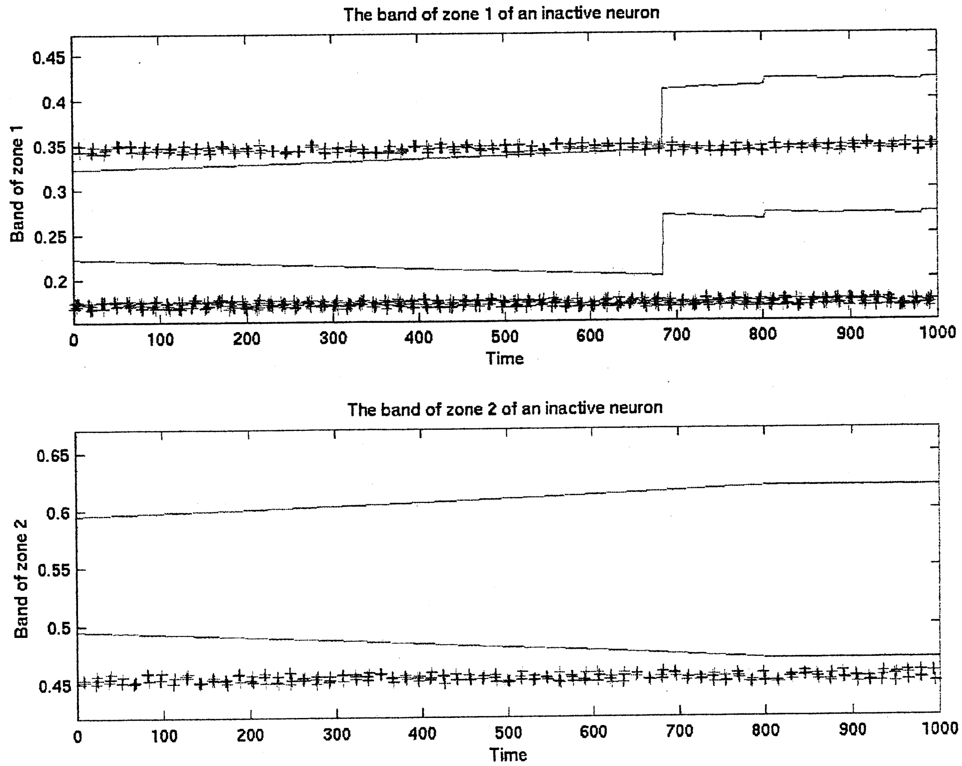

Figure 6: The changes in the bands of a neuron that doesn't fire (neuron 6)

Our application may be interpreted as a process of acquiring a vocabulary of homonyms. The network starts with no prior knowledge. First, the neurons learn to detect the atomic homonyms listed in Appendix A (this happens when the distance between one of the band centers of the neuron and a component of one of the inputs is less than the tolerance level $T(n,r,t)$). Some neurons learn to associate between pairs of homonyms. This happens when both regions in a neuron are activated when one of the memories (input vectors) is presented to the network.

REFERENCES

1. Haykin, Simon (1999). *Neural Networks, a Comprehensive Foundation.* Upper Saddle River: Prentice Hall.
2. Hebb, D.O. (1949). *The Organization of Behavior. A Neuropsychological Theory.* New York: Wiley.
3. Singh, Jagjit (1966). *Great Ideas in Information Theory, Language and Cybernetics.* New York: Dover Publications, Inc.
4. Arbib, Michael A. (1972). *The Metaphorical Brain.* New York: Wiley-Interscience.
5. Anderson, John R. and Bower, Gordon H. (1973). *Human Associative Memory.* Washington D.C.: V.H. Winston & Sons.
6. Kohonen, Teuvo (1977). *Associative Memory. A System-Theoretical Approach.* Berlin; New York: Springer-Verlag.
7. Skapura, David M. (1996). *Building Neural Networks.* New York: ACM Press.

APPENDIX A

The following is a list of associated strings that sound alike. For example "mb" is associated with "m", as in the word "climb". "A" stand for a generic vowel.

bb=b
bt=t
c=ch
c=k
c=s
c=sh
cc=k
cc=ks
ch=ch
ch=k
ch=sh
ci=sh
ck=k
cq=k
cu=k
cz=ch
cz=z
dd=d
dg=j
di=j
dj=j
ed=d
f=v
ff=f
g=j
gg=g
gg=j
gh=f
gh=g
gh=Silent
gi=j
gm=m
gn=n
gu=g
h=Silent
is=i
j=h
ju=w
kn=n
ld=d
lf=f
lk=k
ll=l
lm=m
mb=m
mm=m
mn=m
mn=n
nn=n
pb=b
ph=v
ph=f
pn=n
pp=p
ps=s
pt=t
q=k
qu=k
re=Ar
rh=r
rr=r
s=sh
s=z
sc=s
sc=sh
se=sh
sh=sh
si=ch
si=sh
ss=s
ss=sh
ss=z
st=s
te=ch
th=th
th=t
ti=ch
ti=sh
ts=s
tt=t
vv=v
w=Silent
wh=w
wh=h
x=z
x=ks
y=A
z=s
zz=z
zz=ts

APPENDIX B

Claim: If x, y are independent and identically distributed random variables with a uniform distribution over $[0,1]$ then $prob[|x-y|<T]=2T-T^2$

Proof:

$$
\begin{aligned}
& prob[|x-y|<T]= \\
& 1-prob[|x-y|>T]= \\
& 1-(prob[x-y>T]+\Pr ob[y-x>T])= \\
& 1-(prob[y<x-T]+\Pr ob[x<y-T])= \\
& 1-\left(\int_0^1 \max(0,x-T)dx+\int_0^1 \max(0,y-T)dy\right)= \\
& 1-\left(\int_T^1 (x-T)dx+\int_T^1 (y-T)dy\right)= \\
& 1-\left(\left(\frac{x^2}{2}-xt\right)\Big|_T^1+\left(\frac{y^2}{2}-yt\right)\Big|_T^1\right)= \\
& 1-\left[\left(\frac{1}{2}-T-\left(\frac{T^2}{2}-T^2\right)\right)+\left(\frac{1}{2}-T-\left(\frac{T^2}{2}-T^2\right)\right)\right]= \\
& 1-(1-2T+T^2)= \\
& 2T-T^2
\end{aligned}
$$

The shaded area in the following diagram represents the pairs (x,y) that satisfy $|x-y|<T$.

The area of triangles I and II is $\frac{(1-T)(1-T)}{2}$ each. Therefore, the area of the shaded strip is $1-2\cdot\frac{(1-T)(1-T)}{2}=2T-T^2$