Recall and Recognition in an Attractor Neural Network Model of Memory Retrieval.

E. Ruppin and Y. Yeshurun
Department of Computer Science
School of Mathematical Sciences
Sackler Faculty of Exact Sciences
Tel Aviv University
69978, Tel Aviv, Israel

October 23, 1996

Abstract

This paper presents an Attractor Neural Network (ANN) model of Recall and Recognition. It is shown that an ANN Hopfield-based network can qualitatively account for a wide range of experimental psychological data pertaining to these two main aspects of memory retrieval. After providing simple, straight-forward definitions of Recall and Recognition in the model, a wide variety of 'high-level' psychological phenomena are shown to emerge from the 'low-level' neural-like properties of the network. It is shown that modeling the effect of memory load on the network's retrieval properties requires the incorporation of noise into the network's dynamics. External projections may account for phenomena related with the stored items' associative links, but are not sufficient for representing context. With low memory load, the network generates retrieval response times which have the same distribution form as that observed experimentally. Finally, estimations of the probabilities of successful Recall and Recognition are obtained, possibly enabling further quantitative examination of the model.

1 Introduction

In recent years, considerable progress has been made in investigating the properties of Attractor Neural Network (ANN) models as content addressable memory devices. The goal of this paper is to demonstrate that a Hopfield-based [Hop82] ANN model can qualitatively account for a wide range of experimental psychological data pertaining to the two main aspects of memory access, Recall and Recognition. In the context of psychological experimental conditions, *Recall* is defined as the ability to retrieve an item from a list of items (words) originally presented during a previous learning phase, given an appropriate cue (*cued Recall*), or spontaneously (*free Recall*). *Recognition* is defined as the ability to successfully acknowledge that a certain item has or has not appeared in the tutorial list learned before.

Attractor neural networks are part of the more general connectionist framework that has been developed in recent years in parallel to the classic artificial intelligence mainstream of symbolic processing. While in the latter information is stored at specific memory addresses and processed by explicit rules, in the connectionist framework no such distinction between the information and the processing algorithm exists. The modeled entities are represented not locally (at specific nodes of the network), but their corresponding representations are distributed, involving many nodes of the network, activated in parallel. Parallel Distributed Processing modeling of a broad spectrum of experimentally established phenomena concerning human memory have begun only recently [Sch89]. Previous memory modeling efforts do include some models which are computationally related to connectionist models, as the holographic distributed model [Eic82], and a distributed convolution-correlation model [Mur82]. However, these models lack the biological flavor accompanying connectionist modeling.

An ANN is a network of formal neurons (the network's nodes) connected by synapses (the network's links). After the network's state is set in accordance with a given input pattern, the dynamics of the network are characterized by the following iterative process: Every neuron receives inputs from all other neurons to which it is connected, and fires only if the sum of the inputs is above a certain threshold. When a neuron fires, its output (weighted by the synaptic strengths) is communicated to other neurons, and as a consequence the network's state evolves. By using specific learning rules, governing the way the strength of the synapses in the network are established, a specific set of input patterns can be learned. I.e., these memorized patterns are made to be attractors of the network, such that the network will converge to a memory state if a closely related pattern is presented as an input to the system. Since ANNs have the ability for performing error correction, they can perform content addressable memory retrieval. Indeed, ANN models of psychological data concerning specific aspects of memory retrieval have been presented; e.g., of high speed scanning experiments of the Sternberg type [ASU90], and of semantic memory queries [RU90].

Previous classical 'mathematical' models of memory retrieval (reviewed in [GS84]) have shown a remarkable ability to fit experimental data. However, these models entail the existence of numerous parameters bearing arbitrarily assigned high-level cognitive 'interpretation'. Moreover, the broad spectrum of values possibly assigned to these parameters results in the existence of a large scope of 'model manipulation' possibilities [And91], which contributes significantly to their capability of obtaining a close fit with the data. The ANN model presented in this paper contains very few such high-level parameters explicitly. Various cognitive attributes will be shown to be implicitly manifested as 'bottom-up' emerging properties of the network. Another important motivation for ANN modeling of memory retrieval is that 'reaction times' are naturally represented in such models by the time required for the network computation to be completed [And91]. Indeed, it will be shown that with low memory load, the network generates retrieval response times which have the same distribution form as that observed experimentally.

The model of memory retrieval presented here is based on a particular ANN model, the Hopfield model [Hop82]. The Hopfield framework has been selected because of two main reasons: the first is that this model has been subject to considerable research efforts, which have yielded several results enabling the analysis of Hopfield-based models. The second reason is that in recent years an extensive family of Hopfield-based models have been derived [Wei86], rendering possible the construction of more 'real' biologically-oriented networks, and making it plausible that the model presented can be further extended in the future.

The Hopfield model's dynamics are composed of a non-linear, iterative, asynchronous transformation of the network state [Hop82]. The process may include a stochastic noise which is analogous to the 'temperature' T in statistical mechanics. Formally, the Hopfield model is described as follows:

Let neuron's i state be a binary variable S_i, taking the values ± 1 denoting a firing or a resting state, correspondingly. The network's state is denoted by a state vector S specifying the binary values of all its N neurons. Let J_{ij} be the synaptic strength between neurons i and j. Then, h_i, the input 'field' of neuron i is given by

$$h_i = \sum_{j \neq i}^{N} J_{ij} S_j \tag{1}$$

and the neuron's dynamic behavior is described by

$$S_i(t+1) = \begin{cases} 1, & \text{with probability } \frac{1}{2}(1 + tgh(\frac{h_i}{T})) \\ -1, & \text{with probability } \frac{1}{2}(1 - tgh(\frac{h_i}{T})) \end{cases} \tag{2}$$

Storing a new memory pattern ξ^μ in the network is performed by modifying every ij element of the synaptic connection matrix according to

$$J_{ij}^{new} = \frac{m-1}{m} J_{ij}^{old} + \frac{1}{m} {\xi^\mu}_i {\xi^\mu}_j \tag{3}$$

where m denotes the number of currently stored memory patterns. A few properties of the Hopfield model as a memory device, bearing relevance to the model presented, should be noted:

- The maximal number m of (randomly generated) memory patterns which can be stored with good retrieval quality is $m = \alpha_c \cdot n$, $\alpha_c \approx 0.14$. If more memories are presented in a storage attempt, the network becomes *overloaded* and the retrieval of all stored memories deteriorates abruptly [AGS85], a phenomena termed the *blackout catastrophe* [Ami89]. When noise is incorporated into the neuron's dynamics, the memory capacity $\alpha_c(T)$ monotonically decreases.

- Every stored memory is an *attractor* having an area surrounding it termed its *basin of attraction.* This 'area' denotes a subgroup of the n-dimensional state space. When the network is initiated by an input vector belonging to the basin of attraction of some memory ξ^μ, the networks' state will gradually evolve to the vicinity of ξ^μ, if the noise (temperature) is not too high.

- In addition to the stored memories, there exist also other attractors, denoted as *spurious* states, which do not have a high level of similarity with any single stored memory vector [AGS85]. The dynamic behavior of a Hopfield network can be pictured as following a descending trajectory in the 'energy' plane, where all the attractors are minima of the 'energy' function $E = -1/2 \sum_{i \neq j} \sum S_i S_j J_{ij}$. In a noiseless deterministic Hopfield network the network's attractors are stable states of the dynamics and the network is ensured to converge to such a state [Hop82].

Section 2 provides a description of the model. In section 3, it is demonstrated that a low level ANN model can indeed display the global dynamics that characterize high level, memory related phenomena. We relate some basic principles of human memory performance to the behavior of the model, by showing a qualitative analogy between them, and assuming that the encoding reflects the level of similarity existing between the various items. It should be stressed, however, that a more accurate comparison heavily depends on the nature of the encoding of the stored memories. This encoding is obviously currently unknown, expressing our present ignorance of the nature of the 'representation' used in the brain, thus inherently limiting our ability to provide a detailed quantitative account of memory retrieval. In section 4, the probabilities of successful Recall and Recognition are estimated, providing for a further quantitative examination of the model in the future, when more is known about the encoding. Finally, in section 5., we briefly discuss the modeling of correlated patterns and of free recall.

2 The Model.

The model consists of a Hopfield ANN, in which distributed patterns representing the learned items are stored during the learning phase, and are later presented as inputs during the test phase. In this framework, successful Recall and Recognition are defined. Some additional components are added to the basic Hopfield model for the modeling of the relevant psychological phenomena.

The distributed representation incorporated is well suited to account for Tulving's encoding specifity principle, which asserts that "... remembering of events always depends on the interaction between encoding and retrieval conditions, or the compatability between the engram and the cue *as encoded* ... " [Tul83]. Indeed, by assuming that successful cues are represented by patterns closely related to the stored memory patterns, Tulving's encoding specifity principle is 'naturally' conserved. The only assumption we make on the representation of the items is that the Hamming distance [1] between the test pattern and the various memories reflects their similarity.

The psychological data accounted for by the model is composed of experiments where memory is assessed with recognition and recall tests that make explicit reference to a specific previous experience, i.e., the tutorial phase. This deliberate recollection of recent experiences has been referred to as *explicit* memory, as opposed to another kind of episodic long-term memory, referred to as *implicit* memory. There exists a considerable amount of evidence supporting the existence of fundamental performance differences between implicit and explicit memory [Sch89]. In this work, we have adopted the assumption that different memory systems underlie the various differences in memory performance [Tul85]. Following [GM84], it is further assumed that explicit memory storage depends on the formation of new episodic representations. It is assumed that these representations are formed in a distinctly 'allocated' network, shared by all items learned within the same learning episode (i.e., context).

Recall is considered successful when upon starting from an initial cue the network's state is transformed to the vicinity of the stored memory nearest to the input pattern. Since every stored memory pattern is an attractor, the network's state is guaranteed to remain in this vicinity for a considerable amount of time. If the network converges to a spurious stable state, its output will stand for a 'failure of recall' response. The question of "how do such non-memory states bear the meaning of 'recall failure'?" is out of the scope of this work. However, a possible explanation is that during the learning phase 'meaning' is assigned to the stored patterns via connections formed with external patterns, and since non-memory states lack such associations with external patterns, they are 'meaningless', yielding the 'recall failure' response. Another possible mechanism is that every output

[1] The Hamming distance between a state vector S and a memory pattern $\xi\mu$ is defined as $H(x) = \frac{1}{2n}\sum_{i=1}^{n} |S_i - \xi^{\mu}{}_i|$

pattern generated in the recall process passes also a recognition phase so that non-memory states are rejected, (see the following paragraph describing recognition in our model). For an interesting interpretation of spurious states as pathological phenomena see [Hof87].

Let the rate of change of activity (RCA) denote the number of neurons flipping their state during a time unit. *Recognition* is considered successful when the network's RCA becomes lower than some threshold γ during a time interval Δ, beginning from input presentation. When a pattern precisely identical to a stored memory is presented as an input to the network, successful Recognition is obviously instantaneous; the state of the network will remain for some time in the vicinity of the stored memory, with a very small amount of change in its activity. However, allowing for partially corrupted inputs, the model incorporates a certain degree of 'tolerance' towards slightly distorted inputs, expressed in the length of the interval Δ. Simulation results show that convergence is monotone [AM88] and this is analytically proven for the synchronous case [KP88]. Therefore, the shorter the distance between an input and its nearest memory, the faster is its convergence. The length of the interval Δ hence determines the probability of successful Recognition. Since non-memory (non-learned) stable states have higher energy levels and shallower basins of attraction than memorized stable states [AGS85, LN89], convergence to such states takes significantly longer timer. This observation has also been strongly supported by simulations we have performed, in which convergence of random input vectors to non-memory states takes a much larger number of asynchronous iterations than their convergence to memory states. Therefore, there exists a range of possible values of Δ that enable a successful distinction between stored to non-learned inputs.

The definition of successful Recognition presented is close in spirit to Hopfield's original observation [Hop82] that familiarity can be recognized in a Hopfield network, by monitoring the initial flip rate, which is slower for familiar (memory) states). Our definition also reflects the notion been suggested before in the psychological literature, that the check of novelty of an input pattern is one of the first steps in information processing [Pos78].

The context of the psychological experiments is represented as a substring of the input's encoding. However, since the storage of strongly correlated memory patterns may lead to ambiguity in memory retrieval due to interference, the size of the context encoding relative to the total size of the memory encoding is kept small, preserving the low level of inter-pattern correlation.

The external influence exerted upon the behavior of the allocated network by connections from other networks, expressing the total associational linkage of a learned item, is modeled as an external field vector E. When a learned memory pattern ξ^μ is presented to the network, the value of the external field vector generated is $E = h \cdot \xi^\mu$, where h is a 'projection' coefficient, expressing the association strength supporting ξ^μ.

The weakening of the networks inter-connections as a function of the elapsed time between the tutorial and testing session is modeled by a decay parameter λ. Denote the connection matrix at the end of the learning phase as ${J_{ij}}^0$, then the connection matrix after t time has elapsed is $J_{ij}(t) = e^{-\lambda(t-t_0)}{J_{ij}}^0$. Such a decaying factor has no effect on any of the other test phase phenomena analyzed since it operates on a much longer time scale than the length of the test phase being modeled.

The model may shed some light upon the historical dispute whether Recall and Recognition involve a basically similar mechanism (one-process theories), or whether there exist essential differences between them (two-process theories) [Kin70]. Indeed, in the model presented, both Recall and Recognition are performed in the same network, sharing the same dynamics; the dynamic behavior of an ANN can be viewed as performing a parallel, mutual exclusive search in the phase space [RU90]. Yet, during a Recognition assignment, the 'read-out' of this similar process is different than in Recall, since familiarity is then examined. For a similar observation regarding feed-forward networks see [YRD79].

3 The modeling of experimental data.

In this section we describe how various psychological phenomena are accounted for. Memory retrieval research has yielded an accumulation of a broad spectrum of experimental data, giving rise to elaborate mathematical models where the data is modeled in excruciating detail in a highly sophisticated manner. We have used as reference the list of experimental phenomena described by Gillund and Shiffrin in their comprehensive SAM model [GS84]. The interested reader may find there a detailed review of the literature, providing extensive support to the psychological phenomena described. When appropriate, more recently acquired data has been considered. Regarding every phenomenon discussed, a brief description of the psychological findings is followed by an account of its modeling. We rely on the known results pertaining to Hopfield models to show that the psychological phenomena reviewed are emergent properties of the model. When such analytical evidence is lacking, simulations were performed in order to account for the experimental data. Several phenomena can be accounted for by considering the memory load of the network. When associations with items not presently studied have an important role, the influence of the external 'projection field' is shown to account for the observed data.

3.1 Phenomena accounted for by the effect of memory load

The List-Length Effect: It is known that the probability of successful Recall or Recognition of a particular item decreases as the length of list of learned items increases [GS84, AJ73, RM76b].

List length is expressed directly as the memory load. It has been shown (in networks with several hundred neurons) that the width of the memories basins of attraction monotonically decreases following an approximately inverse parabolic curve [Wei85]. Hence, Recall performance should decrease as memory load is increased. Simulations previously performed with networks having noiseless dynamics have shown that retrieval times of successful trials remains about the same, with various memory loads [Ami87], as long as the memory load is sub-critical. However, we have found that this is not true when noise is incorporated into the dynamics of the network: We have examined the convergence time of the same set of input patterns at different values of memory load. Figure 1 shows that as the memory load is increased, successful convergence has occurred only after an increasingly growing number of asynchronous iterations. Hence, convergence takes more time and can result in Recognition failure, although memories' stability is maintained till the critical capacity $\alpha_c(T)$ is reached.

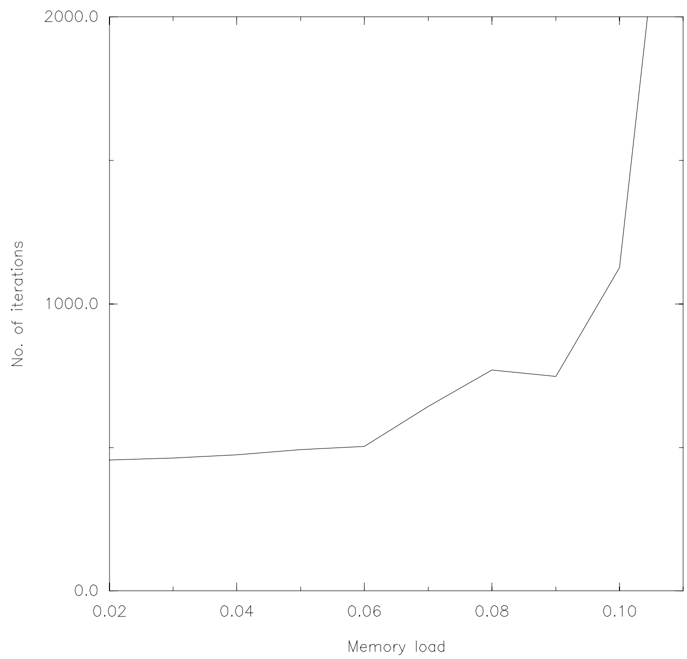

Figure 1: Recognition speed (No. of asynchronous iterations) as a function of the number of stored memories (determining the memory load). The network has $n = 500$ neurons, and $T = 0.2$.

Presentation Time: Increasing the presentation time of learned words is known to improve both their Recall and Recognition [RM76a].

This is explained by the phenomenon of maintenance rehearsal; items presented for a longer time are considered as if presented repeatedly for a number of times, and thus their memories' basins of attraction get deeper: Assume that rehearsing a word for k times is equivalent to presenting it as a learned memory for $f(k)$ times, (f monotonically increasing), then its corresponding energy level will get deepened by a factor of $f(k)$. Deeper basins of attraction are also wider [HFP83, KPKP90]. Therefore, the probability of successful Recall of rehearsed items is increased.

The effect of a uniform rehearsal of all learned items (without renormalizing the synaptic weights) is equal to a global uniform energy decrease by a factor of k, i.e., a decrease in the temperature T by a factor of $1/k$. The temperature decrease is known to increase the memory capacity of the network [AGS85], and thus to decrease the actual memory load, leading to a higher probability of successful Recognition. The model predicts that increasing presentation time will attenuate and delay the List length phenomenon. It can be seen that there is a certain limit beyond which the increase of the presentation time will not have any significant effect since the temperature is already approaching zero. This is in accordance with the findings [PA85] that beyond a certain amount of daily repetitions, there is little impact on fact retrieval. Moreover, the temperature decrease will lead to a lesser variation of retrieval times with memory load and therefore the model predicts that increasing presentation time will decrease the magnitude of the list-length phenomenon.

Test Delays: When a list of items is learned, it is known that as the time delay between the tutorial and test phases get longer, Recall and Recognition performance deteriorates [She67].

The Test Delay phenomena is accounted for by entering the decay parameter λ into the model, simulating the weakening of neuronal synapses as a function of the elapsed time between the tutorial and testing session. Relative to the length of the test delay, the duration of the test phase itself is short, and this justifies discarding the synaptic decay occurring during the latter phase, and treating the synaptic matrix as fixed through the test period. Contrary to the case of an increase in presentation time, now the energy level of stored memory patterns is increased, and their basins of attraction get shallower and smaller, resulting in a deterioration of Recall performance. The actual memory load is increased (since decreasing the neurons' input field h is equivalent to a temperature raise), leading to a parallel decrease in Recognition performance.

Age differences in Recall and Recognition: It was found that older people perform more poorly on Recall tasks than they do on Recognition tasks [CM87]. An analysis of covariance, with recognition performance as the covariate, showed a reliable age decrement in recall.

These findings can be accounted for by assuming that synapses are being deleted in the progress of life. The data supporting this assumption is inconclusive. However, some evidence supports an age-related decrease of synaptic membrane surface densities in the hyppocampus [BFFMRU89], and a progressive decrease in hippocampal dendritic spine density [MR88].

It has been previously shown that most of the retrieval failures occurred, even when the network's synapses were quite radically diluted, only with input states which were far apart from stored memories [HD89]. The quality of the retrieved pattern is only mildly affected by synaptic dilution; up to 80% dilution, the final converged states have final overlaps above 0.9 with the nearest memory states [Som86].

We have investigated the retrieval performance as a function of the input's initial similarity (overlap) with its nearest memory pattern, for various levels of synaptic dilution and memory load: At low levels of synaptic dilution, memory retrieval remains intact. For every level of initial overlap of the input patterns, there exists a 'critical' level of dilution whereupon a rapid decrease in the retrieval of such patterns occurs. As demonstrated in Figure 2, when the memory load is increased, this 'critical' level begins at lower levels of synaptic dilution. Accordingly, it is predicted that during the early phases of synaptic dilution, the deterioration of Recall performance might be noticeable only when the tutorial phase involves long lists of items. Only a mild decrease (up to 12%) in recognition speed was found. taken together, these findings could account for the relative stronger decrease in Recall vs. Recognition Performance. One may further speculate that during aging the the network's noise level may decrease, and hence attenuate the effect of the actual increase of memory load on recognition speed.

3.2 Considering the effect of external projections

The total sum of external, currently active associations, projecting on the allocated network from all other networks (i.e., 'the rest of the brain') is represented in the model by an external field vector E. It has been shown that the probability of retrieval of a memory pattern associated with an external field vector increases monotonically with the amplitude of such a field [AGS87].

The word-frequency effect: The more frequent a word is in language, the probability of recalling it increases, while the probability of recognizing it decreases. [She67, Rao89].

A word's frequency in the language is assumed to effect its retrieval through the stored word's semantic relations and associations [Kat85, NCBK87]. Thus, activation in various other networks (i.e., 'cortical areas'), and not just exclusively in the specifically allocated network has to be considered.

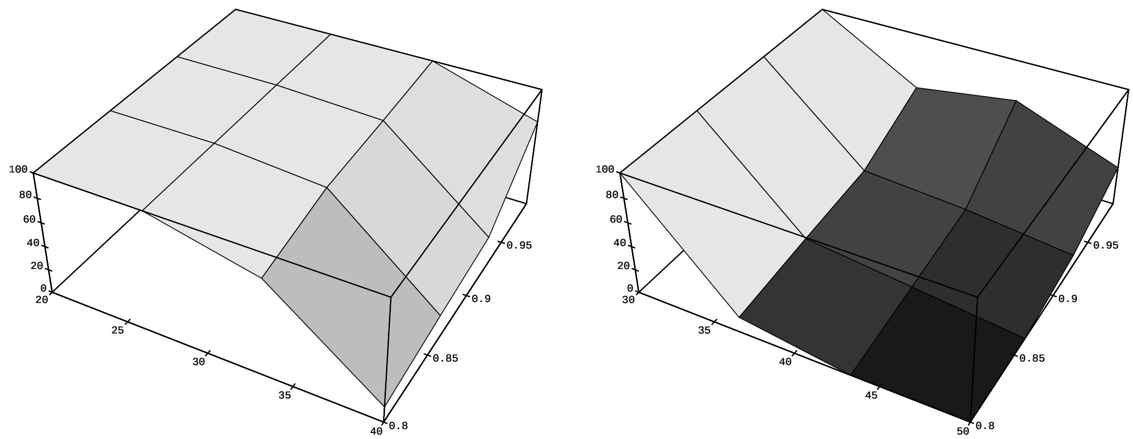

Figure 2: The probability of successful Recall as a function of the number of stored memories and the input's initial overlap (similarity) with its nearest memory pattern. The right figure pertains to 50% dilution and the left figure to 55% dilution. The network has $n = 500$ neurons, and $T = 0.2$.

When, during a Recognition test phase, a stored memory pattern ξ^μ is presented to the network, it is assumed that the external field vector generated has a component projecting at ξ^μ, such that $E = h^\mu \cdot \xi^\mu$, where h^μ is a 'projection coefficient' representing the strength of associational linkage supporting ξ^μ.

It is assumed, that relative to low frequency words, high frequency words have more semantic relations and therefore more connections between the patterns representing them and other patterns stored in the memory (i.e., in other networks). This one-to-many relationship is assumed to be reciprocal, i.e., each of the externally stored patterns has also connections projected to several of the stored patterns in the allocated network. Let the overlap H^μ, denoting the level of similarity between the current network's state S and a corresponding memory pattern ξ^μ, be defined as $H^\mu = \frac{1}{n} S \cdot \xi^\mu = \frac{1}{n} \sum_i S_i {\xi^\mu}_i$. The process leading to the formation of the external field E (acting upon the allocated network), generated by an input pattern nearest to some stored memory pattern ξ^μ is assumed to be characterized as follows:

1. There is a threshold degree of overlap θ_{min}, such that $E > 0$ only when the allocated network's state overlap H^μ is higher than θ_{min}. This reflects the initiation of activation of associated items in the external networks. As the network's state evolves towards the nearest memory pattern ξ^μ, (the overlap H^μ is increased) more patterns corresponding to items associated with this memory are activated.
2. Patterns which are activated already at low overlap (H^μ) values are strongly associated with ξ^μ. They are assumed, in turn, to strongly support ξ^μ, reflected

in an increase in the value of h^μ. Patterns which are activated only at high overlap values (and thus are only loosely connected to the memory represented by ξ^μ) generate a more diffuse projection on the allocated network, which results in a decrease in the value of h^μ. Consequently, as H^μ rises, h^μ rises monotonically till some maximal point beyond which it monotonically decreases.

3. High-frequency words have lower θ_{min} values than low-frequency words. This reflects the widely held assumption that high-frequency words have stronger associational linkage, which enables the earlier excitation of their related items' patterns of activity.

In Recognition tests, since the input patterns represent stored memories, a high value of overlap H^μ to some memory ξ^μ is obtained. Such high values of H^μ are already much larger than the relatively low θ_{min} values of high-frequency words, and therefore the value of h^μ and E generated is already post-optimal and therefore small (smaller than in the case of low-frequency words which have higher θ_{min} values). Thus, high-frequency words generate a more diffuse external field E, accounting for their relatively lower probability of successful Recognition. Since the effect of the external field on memory retrieval is increased at conditions of high memory overload [AGS87], the model predicts that the relative inferiority of Recognition of high-frequency words will be chiefly manifested with long lists of learned words.

In Recall tests, however, only a cue pattern is presented to the network. The initial situation is therefore characterized by low values of overlap H^μ to some nearest memory ξ^μ. At such conditions, only the overlap value of high-frequency words may suffice for activating associated items, i.e. $H^\mu > \theta_{min}$, and thus an external field strongly oriented at ξ^μ is generated. Hence, the relatively higher probability for successful Recall of high-frequency words is accounted for.

We have presented here the analysis of a 'classical' case of Recall testing. Generally, however, a wide spectrum of Recall cues, differing upon their overlap value to their nearest memory, may be presented. This may lead to various scenarios, depending on the relation between the initial H^μ and θ_{min}. It is interesting to note that this is in fact in accordance with some newer inconclusive psychological experiment results regarding the influence of word frequency upon Recall [GMC80], while the findings regarding Recognition remain robust. In addition, the model predicts the creation of a specific type of Recall error; namely, successful recall of a learned-list word, but not the one nearest to the cue. This can happen since a Recall cue nearest to a low-frequency stored word may generate an external field oriented toward a near (but not nearest) high-frequency word, which may modify the convergence toward it.

Word Fragment Completion Tests: In this version of Recall testing, graphemic word fragments are presented as Recall cues. In contradiction to the generally accepted positive correlation found between successful Recall and Recognition (i.e., the probability of successful Recall is greater for items that can be recognized than for those that cannot), no such positive correlation is found when using graphemic fragment completion for Recall [TSS82]. The relation actually found was one of stochastic independence; a word appearance in the study list does enhance the subject's ability to generate the word to its fragment cue, but such enhancement was identical for the recognized words and for those not recognized.

The results reported in [NCBK87] indicate that while cued recall has also a semantic search component in addition to lexical search, fragment completion does not. In accordance, it is assumed that while in regular Recall tests the semantic associations have an important role and exert considerable influence on the final form of the energy plane, in word fragment completion the situation is entirely different: The external field vector is inactive since the cue is presented in a fragmented form, incapable of generating the semantic re-enforcing associations. The energy plane is therefore totally different during such tasks than in Recognition testing (where, like in Recall testing, semantic associations are activated), resulting in the lack of correlation described.

3.3 Modeling other phenomena

Context Shift: The term Context Shift refers to the change in context from the tutorial period to the test period. Studies examining the effect of context shift have shown a decrement in Recall performance with context shift, but little change in Recognition performance [SGB78].

The context is represented as a substring of the patterns' encoding. Let β denote the size of the common context sub-string relative to the size of the stored pattern, and let α denote the memory load. In simulations we had performed it has been demonstrated that there exists a certain range of values of α and β where Recall and Recognition are successful in the majority of trials. However, as shown in Figure 3, when context shift is simulated by entering random input patterns, but with half of the context substring inverted, Recall performance severely deteriorates. We have found only a small increase (of 5%) in the number of asynchronous iterations required for convergence, in accordance with the experimental findings.

It is interesting to note that representing the context of memorized words as part of the external connections projecting upon the network cannot successfully account for the context shift phenomena. In such an approach (inspired by the modeling of context used in Semantic networks models), the tutorial phase's context would be represented as an external field C projecting equally on *all* stored memories. However, such a field, averaged on all stored memories, ($E = h \cdot \sum_{\mu=1}^{m} \xi^{\mu}$) has been demonstrated to cause only a marginal increase in the retrieval properties of the network [AGS87], and thus is incapable of accounting for the decrease in Recall accompanying context shift.

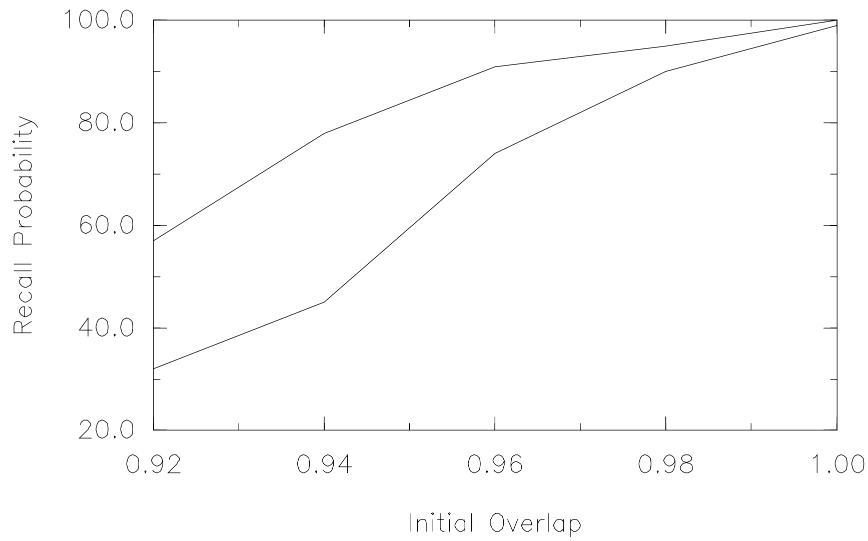

Figure 3: The probability of successful Recall as a function of the initial overlap between the input vector and its nearest memory pattern, after a context shift. The network has $n = 500$ neurons, $T = 0.2$, and $m = 30, 40$. The corresponding pre-shift probabilities have been 100% success.

Distractor Similarity effect: Distractors are unlearned words that are intermixed with the learned items while performing a Recognition test. It turns out that data concerning distractor similarity effects on Recognition are mixed; some studies indicate that increased similarity of the distractors to certain learned words decreases Recognition performance [GS84, AK68] while others do not report any significant change [UF68, CE72].

By definition, the model predicts that increased distractor similarity will cause an increase in false positive recognition (i.e., of wrongly 'recognizing' non-stored patterns). However, this depends on the degree the encoding preserves this similarity. It is interesting to note in this respect, that physical similarity produces greater amount of false alarms than semantic similarity [GS84], possibly indicating that the encoding conserves better the first kind of similarity than the latter. Moreover, one should note that the retrieval of an input pattern in a Hopfield network is not determined by the initial Hamming distance solely [AM88]. The precise form of the memories' basins of attraction in the network's n-dimensional space is not known [Ami89]. These observations therefore make the model not necessarily inconsistent with the the mixed results obtained concerning the similarity effect of distractors.

Serial position effect It has been found that the serial position of the learned items in the tutorial phase has a considerable influence on the probability of their successful Recognition, and on the mean response times (MRT) obtained: The items presented earlier in the learning phase had been recognized with more accuracy and during a shorter MRT [RM76b, Rat76].

Serial position is modeled by ascribing higher synaptic weights to the patterns stored earlier in the tutorial phase. Following the original experimental data, we have divided the stored patterns into four groups of 'serial position', in accordance with the period in which the item was learned during the tutorial phase. The original synaptic learning rule is slightly modified so that the strength ϕ_k by which a memory pattern is embedded in the synaptic matrix is a function of the period k in which the item was learned, i.e., $J_{ij} = \sum_{\mu=1}^{m} \phi_k {\xi^\mu}_i {\xi^\mu}_j$, where m denotes the number of stored memories. Selecting $\phi_1 = 1.15$, $\phi_2 = 1.1$, $\phi_3 = 1.05$ and $\phi_4 = 1$ leads to the results depicted in Figure 4. It can be seen that the MRTs obtained have a monotonically increasing trend similar to that obtained experimentally. The analysis of all other phenomena remains valid when the latter modified learning rule is considered.

We have selected a temporal threshold and examined the Recognition performance as a function of the serial position. Figure 4 displays the percentage of correct hits obtained in the network, simulated by entering random input vectors generated with a high degree of similarity to the stored memories. Indeed, the results obtained have a decreasing trend, similar in general to the experimental data. However, in the experimental data, the two groups memorized in the second half of the learning phase (3 and 4) have shown essentially the same MRT and recognition accuracy. Obviously, a closer fit with the data could have been obtained simply by selecting $\phi_3 = \phi_4$. Such a close fit with the data can be achieved in a less arbitrary manner by modifying the synaptic learning rule so that it will generate a relative 'weighting' of the stored memories in accordance with the time t of their storage, for example the rule $\phi(t) = \phi_0(1 + e^{-(t-\tau)})$, where τ is some time constant of the learning phase.

3.4 Response Times

In addition to the set of data shown to be accounted for qualitatively, the model presented can account for the data concerning retrieval response times. Figure 5 displays the distribution of retrieval times of correct hits at a low memory load. The response times distribution

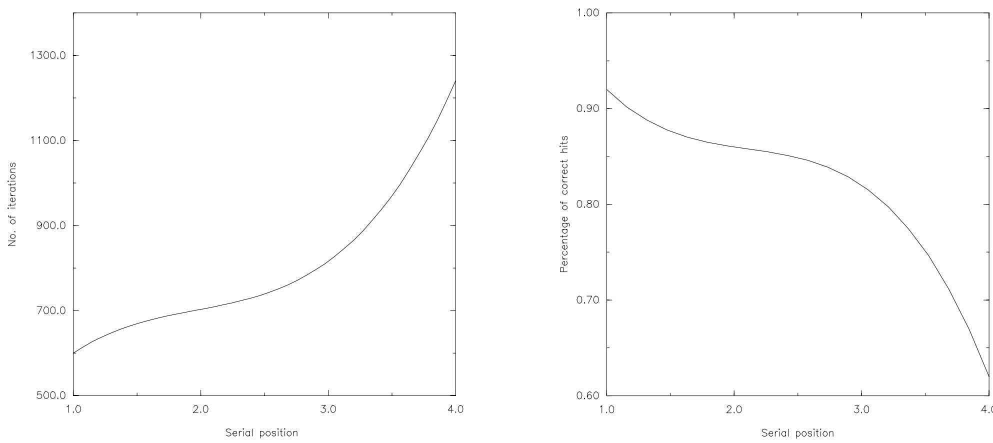

Figure 4: Left: The MRT (No. of asynchronous iterations) as a function of the serial position. The network has $n = 500$ neurons, $m = 30$ memories, and $T = 0.2$. Right: Percentage of correct hits as a function of the serial position. The network has $n = 500$ neurons, $m = 30$ memories, and $T = 0.2$. The temporal threshold was set to 950 asynchronous iterations

obtained looks like the experimentally observed distributions [RM76b], which are asymmetric and typically have a long, low density, right-sided tail. Such experimentally-matching response times distributions have been obtained previously by Anderson in an analog ANN [And91]. However, it should be noted that as the memory load is increased towards maximal capacity, the distribution tends to loose its typical asymmetric form and becomes essentially symmetric. It would be interesting to examine experimentally, if considerably extending the list length may lead to a similar change in the response times distribution.

The experimentally observed response times distribution of correct rejections has a form similar to the response times distribution of correct hits [RM76b]. However, using a single, fixed Δ time interval seems to lead to a narrow distribution of response times of correct rejections. Moreover, a naive 'implementation' of the interval Δ requires one to assume the existence of an internal 'clock' in the network. However, a possible neural realization of Δ may be achieved by assuming the existence of positive feedback between the network's RCA and the thermal noise level. I.e., the level of thermal noise is monitored by the level of the network's RCA [LN89]: When the network converges into an attractor its RCA diminishes, resulting in a temperature decrease. If convergence does not occur during some time interval, the noise will gradually increase leading to a chaotic state which denotes recognition failure. Such a process can account for an actual dynamically varying temporal recognition interval, and will 'spread-out' the response times distribution of correct rejections. The previous results concerning positive hits would still hold since in the case of correct recognition a

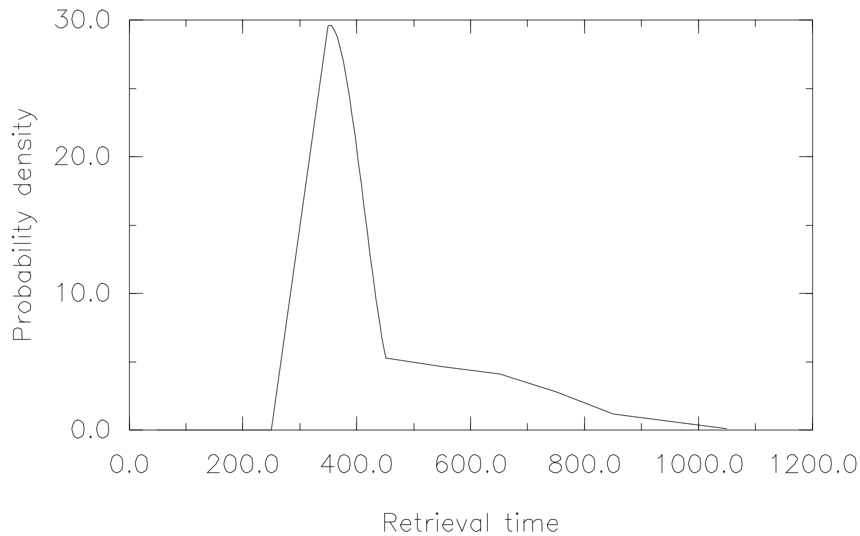

Figure 5: Response times distribution of correct hits. The network has $n = 500$ neurons, $m = 30$ memories, and $T = 0.2$. Time is denoted by the number of asynchronous iterations.

noise surge will not occur.

4 Quantitative Estimations of Recall and Recognition.

In this section we present some estimations of the expected probabilities of successful Recall and Recognition, in the framework of the model presented. Since the precise form of the memories basins of attraction is yet unknown, our calculation is based on the assumption that a memory's basin of attraction can be approximated as a 'sphere' of radius r in the n-dimensional space surrounding it. We recall that any attempt to quantitavely analyze 'high order' experimental data as the psychological manifestation of 'low order' neural activity requires prior knowledge of the encoding. Yet, one hopes that perhaps by trying to evaluate some predictions of such ANN models, some insight about the encoding itself may be gained.

4.1 Estimating Recall performance

In a given network, with n neurons and m memories, the radius r of the basins of attraction of the memories decreases as the memory load parameter ($\alpha = m/n$) is increased. McEliece et. al. [MPRV87] have related n, m, and r by the expression $m = \frac{(1-2 \cdot r)^2}{4} \cdot \frac{n}{\log n}$. Using this result in the framework of our model, we can estimate the probability of successful recall. The concept of the basins of attraction implies a non-linear probability function with low probability when input vectors are further than the radius of attraction and high probability otherwise. The slope of this non-linearity increases as the noise level T is decreased. Even without 'thermal' noise (T = 0) this probability is not unity, in light the presence of non-thermal noise created by the attraction of the other memories. Similarly, the probability of successful Recall (to the nearest memory) when the input is further than the radius of attraction is small but non-zero [Cot88].

In Figure 6, the calculation of r as a function of the memory load α, is shown. r denotes a vector being $r \cdot n$ bits (neurons) apart from its closest memory. The calculated r's denote the maximal allowed distance between the cue and its closest memory so that Recall will still be successful, with high probability.

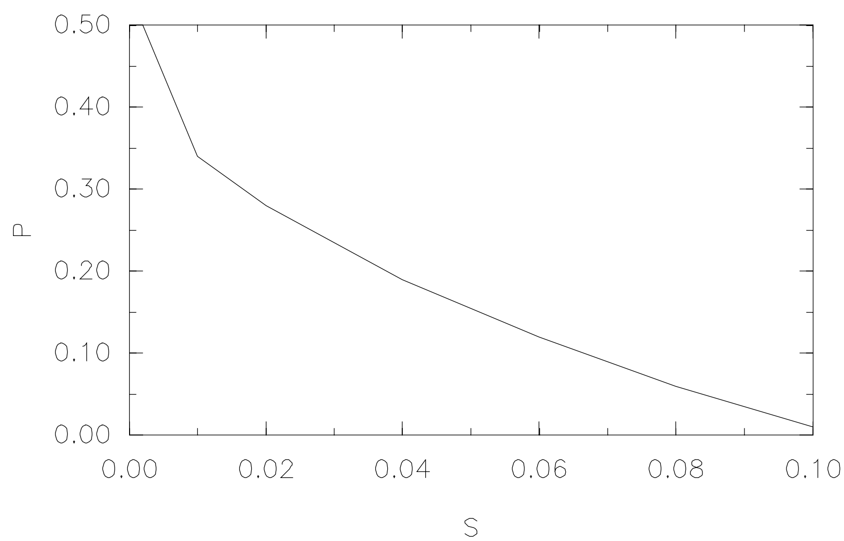

Figure 6: The radius of attraction as a function of memory load. The network has $n = 500$ neurons, and $T = 0.2$

The probability P_c that a *random* input vector will converge to one of the stored memories is $\approx \frac{\sum_{d=1}^{rn} \binom{n}{d}}{2^n}$. It converges very fast toward zero as r is decreased. Thus, at moderate memory load conditions we already obtain a relatively small cumulative attraction size, since the memories' basins of attractions defined by these radia cover only a small part of the total state space.

Figure 7 presents the results of a simulation performed in order to observe the probability P_c as a function of the memory load. These results are in correspondence with the estimation of P_c presented, i.e, P_c, decreases very fast (faster than r) with the increase of memory load. Hence, Recall tests beginning from randomly generated cues would yield a very low rate of successful Recall (P_c). Yet, if one examines Recall by picking a stored memory, flipping some of its encoding bits, and presenting it as an input to the network (determining r), reasonable levels of successful Recall can still be obtained even when a considerable number of encoding bits are flipped.

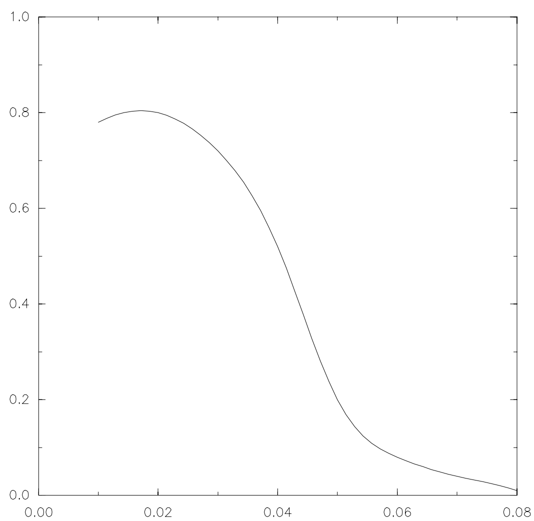

Figure 7: The probability of successful Recall of randomly generated input patterns as a function of memory load. The network has $n = 500$ neurons, and $T = 0.2$

It is interesting to note that the allocated networks' size n can be estimated: The number of learned items (stored memories) m in the psychological experiments is known (of the order of a few dozens), and since the capacity ratio $\alpha = m/n$ is bounded by $\alpha_c = 0.14$, then n is between several hundred to several thousand neurons (i.e., of the order of the size of cortical columns). When considering the context representation, P_c can be estimated as $\approx \frac{\sum_{d=1}^{r \cdot n} \binom{n-c}{d}}{2^{n-c}} \cdot m$. Finding a value of c providing a good fit with experimental data may provide an estimation of the relative size of the context encoding. The size of the context encoding could also be estimated by examining the effect of total context shift: Accordingly, the relative expected deterioration in the probability of successful random Recall due to a change of context involving all c bits of the encoding ($RD(c)$), is estimated by $RD(c) \approx \frac{\sum_{d=1}^{r \cdot n - c} \binom{n-c}{d}}{\sum_{d=1}^{r \cdot n} \binom{n-c}{d}}$.

4.2 Estimating Recognition performance

The probability of correct Recognition depends mainly on the the length of the interval Δ. Assume that after an input pattern is presented to the network, during the time interval Δ

k iterations steps of a Monte Carlo simulation are performed: In each such step, a neuron is randomly selected, and then it examines whether or not it should flip its state, according to its input. Since successful Recognition involves input patterns with a high level of overlap to one of the stored memories, it is plausible to assume that a selected neuron will indeed update its state correctly; i.e., in accordance with its corresponding state in the input pattern's nearest memory.

Let n denote the number of neurons in the network, and let d denote the Hamming distance between an input pattern presented to the network and its nearest stored memory pattern. Obviously, the set of neurons which are selected must include all d 'faulty' neurons so that successful convergence could be achieved. Let us now estimate the probability P_g that the input pattern will be successfully recognized, i.e., that the network will converge to its nearest memory pattern during (at most) k iterations: Let $\bar{i_k}$ denote the event 'neuron i was not selected (and if necessary, updated) during k iterations'. Its probability, $Pr\{\bar{i_k}\}$, is given by $(1-\frac{1}{n})^k$. The event 'at least one of the d faulty neurons was not selected' is equivalent to the union of all the events $\bigcup_{i=1}^{d} \bar{i_k}$. Therefore, its probability $\bar{P_g}\{d\}$ is bounded by

$$\bar{P_g}\{d\} \leq \sum_{i=1}^{d} pr\{\bar{i_k}\} = d(1-\frac{1}{n})^k \leq d \cdot e^{\frac{-k}{n}} \tag{4}$$

and the probability of successful recognition of an input at distance $H(d)$ $P_g\{d\} = 1 - \bar{P_g}\{d\}$, is bounded by $P_g\{d\} \geq 1 - d \cdot e^{\frac{-k}{n}}$. It can be seen from equation 4 that the dependence of the successful Recognition on the initial Hamming distance is multiplicative, which reflects our intuitive notion that Recognition's success depends strongly on the initial input proximity to a stored memory. However, successful Recognition is even more critically dependent on the number of allowed asynchronous iterations k, determined by the length of Δ. For a selection of $k = n(ln(d)+c)$, one obtains $P_g \geq 1 - e^{-c}$.

The expected number of iterations $Exp(X)$ till successful convergence is achieved is calculated as follows: Let the random variable X_i denote the number of iterations performed when i faulty neurons are left, till one of them is selected and corrected. Therefore, $Exp(X) = \sum_{j=1}^{d} X_j$. Since there are i faulty neurons left in the network, the probability of selecting a faulty neuron is i/n, the expected number of iterations X_i performed is $E(X_i) = n/i$, and

$$E(X) = \sum_{i=1}^{d} E(X_i) = n \cdot \sum_{i=1}^{d} \frac{1}{i} \approx n \cdot ln(d) \tag{5}$$

The calculations presented above pertain to the case of 'complete' retrieval, where the network converges to a stable state which is precisely identical to the input's nearest stored memory. However, it is known that in most times, when presented with a successful input cue, a Hopfield network converges to a stable state which is very similar, but not identical, to its nearest stored memory [AGS85]. Let o denote the Hamming distance below which retrieval is considered successful, then a similar calculation yields $P_g \geq 1 - \binom{d}{o} \cdot e^{-\frac{o \cdot k}{n}}$ and $E(X) \approx n \cdot ln(\frac{d}{o})$. In simulations we have performed, ($n = 500$, $m = 10$, $d = 20$, $o = 10$), the average number of iterations until successful convergence was in the range of 300 - 400, in excellent correspondence with the predicted expectation, $E(X) = 500 \cdot ln(2)$.

Two kinds of Recognition errors exist: False negative (FN) recognition – in which we do not recognize a learned item that should be recognized, and false positive (FP) recognition – in which we do recognize an item that should not be recognized, (i.e., was not learned). In order to estimate the probabilities of FP and FN recognition errors, one must adopt a criterion for determining which input patterns should be recognized; let us assume that there exists a constant j such that any input pattern X with Hamming distance $H(X) < j$ should be recognized, and otherwise rejected. Since all *correctly* recognized input patterns have a Hamming distance smaller than j, than, using equation 4, $P\{j\}$ is a lower bound on the probability of correct recognition, and an upper bound on the probability of FP errors. Similarly, $(1 - P\{j\})$ is a lower bound on the probability of correct rejection and an upper bound on the probability of FN errors. The measurement of these probabilities experimentally can set constraints on the possible value of j.

5 Discussion

A prime motivation behind neural networks research is the claim that such networks' behavior may be viewed as a simplified model of the brain. True, the Hopfield model incorporates extensive simplifications in comparison with the known biological data pertaining to neural activity. Yet, some of these simplifications have been challenged in a family of Hopfield-based models been recently derived. These models preserve the basic properties of the original Hopfield model, and therefore it is plausible that the model presented could be modified in accordance, thus better approximating real biological data. The interested reader may find an account of such models in [Wei86], and a detailed discussion of ANN modeling from a neurophysiological point of view in [Ami89, Abe91].

The work presented is limited to the storage of patterns with low correlation, since the storage of strongly correlated patterns will lead to erroneous memory retrieval. Indeed, it is known that a list composed of highly similar items is considerably more difficult to learn than one with dissimilar items. Efficient storage and retrieval of patterns with considerable inter-item correlation is possible, using Hopfield-based Hierarchical neural networks (HNN) [FI87, Gut86]. Such networks have already been utilized for modeling Semantic memory [RU90], and they can potentially be used for modeling of Recall of categorized items. HNN networks can be used as a starting point for an alternative approach to that presented in this work; instead of assuming the existence of specifically allocated networks, an HNN could be used to store different 'assemblies' of stored memories, each characterized by common a substring representing its learning context. The context substring could be then viewed as a 'key' for obtaining access to the appropriate assembly of memories, stored during the same learning phase.

In the model presented, only cued Recall is accounted for. The modeling of free Recall seems a more complicated task, since it requires the understanding of an autonomous, self generated cognitive processes. In principle, a model of an autonomous cognitive process must include two components; the first being a mechanism for generating a continuous movement in the model's phase-space between the various stored memories. The second component being, that once the network has reached a memory state, it should remain in this state for a certain amount of time. Free recall is therefore envisioned as a continuous flow of the system between some of its states, whereas upon arriving at a stored memory state, it remains there for a certain duration. This period is essential for rendering these states as 'cognitively significant' and thus as representing successfully recalled items. Both requirements can be fulfilled using Hopfield-like ANNs: The motion between the various stored memories may be either deterministic [CS87, HU89], or stochastic (for a review of such mechanisms see [Bel88]). Once the network's state approaches the vicinity of a stored memory it remains there for some time before moving to the vicinity of another attractor. The context of the tutorial list is used as a cue guiding the recall process to the stored patterns of learned items.

We have shown in this paper that modeling the effect of memory load on the network's retrieval properties requires the incorporation of noise into the network's dynamics, and that external projections are not sufficient for representing context. In addition to numerous qualitative accounts of the experimental data, the network is shown to generate retrieval response times which have the same distribution form as that observed experimentally, at low memory load. As Schacter claims, "... no current theory gracefully accommodates all or most of the important empirical facts" [Sch89]. However, we feel that the ANN model presented has the intuitive appeal of being considered as an approximation (as coarse as it may be) of some principles of neural correlates of cognitive processes related to memory retrieval.

Acknowledgment: We are indebted to Professor J. Anderson for his many illuminating comments, and for bringing to our attention R.D. Luce's book 'Retrieval Times'.

References

[Abe91] M. Abeles. *Corticonics: Neural Circuits of the Cerebral Cortex.* Cambridge University Press, 1991.

[AGS85] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Storing infinite numbers of patterns in a spin-glass model of neural networks. *Phys. Rev. Lett.*, 55:1530, 1985.

[AGS87] D. J. Amit, H. Gutfreund, and H. Sompolinsky. Statistical mechanics of neural networks near saturation. *Annals of Physics*, 173:30, 1987.

[AJ73] R.C. Atkinson and J.F. Juola. Factors influencing speed and accuracy of word recognition. In S. Kornblum, editor, *Attention and Pefromance (Vol. 4)*. 1973.

[AK68] M. Anisfeld and M. Knapp. Association, synonymity and directionality in false recognition. *Journal of Experimental Psychology*, 77:171, 1968.

[AM88] S. I. Amari and K. Maginu. Statistical neurodynamics of associative memory. *Neural Networks*, 1:63, 1988.

[Ami87] D.J. Amit. Neural networks - achievements, prospects, difficulties. In W. Guttinger, editor, *The Physics of Structure Formation*. Springer-Verlag, 1987.

[Ami89] D.J. Amit. *Modeling brain function: the world of attractor neural networks*. Cambridge University Press, 1989.

[And91] J. Anderson. Why, having so many neurons do we have so few thoughts? In W.E Hockley and S. Lewandowsky, editors, *Relating Theory and Data: Essays on Human Memory in Honor of Bennet B. Murdock*. Erlabaum, 1991.

[ASU90] D.J. Amit, D. Sagi, and M. Usher. Architucture of attractor neural networks performing cognitive fast scanning. *Network*, 1(2):???, 1990.

[Bel88] T. Bell. Sequental processing using attractor transitions. In Carnegie Mellon, editor, *Proceedings of the Connectionist Models Summer School*. 1988.

[BFFMRU89] C. Bertoni-Freddari, P. Fattoretti, W. Meier-Ruge, and J. Ulrich. Computer-assisted morphometry of synaptic plasticity during aging and dementia. *Pathol. Res. Pract.*, 185(5):799, 1989.

[CE72] P. Cramer and M. Eagle. Relationship between conditions of crs presentation and the category of false recognition errors. *Journal of Experimental Psychology*, 94:1, 1972.

[CM87] F.I.M. Craik and J.M. McDowd. Age differences in recall and recognition. *Journal of Experimental Psychology; Learning, Memory, and Cognition*, 13(3):474, 1987.

[Cot88] M. Cottrell. Stability and attractivity in associative memory networks. *Biol. Cybern.*, 58:129, 1988.

[CS87] A. Cristiani and H. Sompolinsky. Dynamics of spin systems with randomly assymetric bonds: Langevian dynamics and a spherical model. *Phys. Rev. A.*, 36:4922, 1987.

[Eic82] J.M. Eich. A composite hollographic associative recall model. *Psychological Review*, 89(2):627, 1982.

[FI87] M.V. Feigelman and L.B. Ioffe. Models of associative memory asymmetric interaction and hierarchy of patterns. *Int. Jour. of Mod. Phys. B*, 1:51, 1987.

[GM84] P. Graf and G. Mandler. Activation makes words more accessible, but not neccessarily more retrievable. *Journal of Verbal learning and Verbal Behavior*, 23:553, 1984.

[GMC80] V. H. Gregg, D.C. Montgomery, and D. Castano. Recall of common and uncommon words from pure and mixed lists. *Journal of Verbal Learning and Verbal Behavior*, 19:240, 1980.

[GS84] G. Gillund and M. Shiffrin. A retrieval model for both recognition and recall. *Psychological Review*, 91:1, 1984.

[Gut86] H. Gutfreund. Neural networks with hierarchically correlated patterns. *Tech. Report. NSF-ITP-86-151, Univesity of California*, 1986.

[HD89] R. Hoffman and S. Dobscha. Cortical pruning and the development of schizophrenia: A computer model. *Schizophrenia Bulletin*, 15(3):477, 1989.

[HFP83] J.J. Hopfield, D. I. Fienstien, and R. G. Palmer. Unlearning' has a stabilizing effect in collective memories. *Nature*, 304:158, 1983.

[Hof87] R.E. Hoffman. Computer simulations of neural information processing and the schizophrenia-mania dichotomy. *Arch. Gen. Psychiatry*, 44:178, 1987.

[Hop82] J.J. Hopfield. Neural networks and physical systems with emergent collective abilities. *Proc. Nat. Acad. Sci. USA*, 79:2554, 1982.

[HU89] D. Horn and M. Usher. Neural networks with dynamical thresholds. *Physical Review A*, 40:1036, 1989.

[Kat85] T. Kato. Semantic-memory sources of episodic retrieval failure. *Memory & Cognition*, 13(5):442, 1985.

[Kin70] W. Kintsch. Models for free recall and recognition. In D. A. Norman, editor, *Models of human memory*. Academic Press, 1970.

[KP88] J. Komĺos and R. Paturi. Convergence results in an associative memory model. *Neural Networks*, 1:239, 1988.

[KPKP90] B. Kagmar-Parsi and B. Kagmar-Parsi. On problem solving with hopfield neural networks. *Biol. Cybern.*, 62:415, 1990.

[LN89] M. Lewenstein and A. Nowak. Fully connected neural networks with self-control of noise levels. *Phys. Rev. Lett.*, 62(2):225, 1989.

[MPRV87] R.J. McEliece, E.C. Posner, E.R. Rodemich, and S.S. Venkatesh. The capacity of the hopfield associative memory. *IEEE Transactions on Information theory*, IT-33(4):461, 1987.

[MR88] N.A. Muma and P.P. Rowell. Effects of chronic choline and lecithin on mouse hippocampal dendritic spine density. *Exp. Aging Res.*, 14(2-3):137, 1988.

[Mur82] B. B. Murdock. A theory for the storage and retrieval of item and assocciative information. *Psychological Review*, 89(6):609, 1982.

[NCBK87] D.L Nelson, J.J. Canas, M.T. Bajo, and P.D. Keelan. Comparing word fragment completion and cued recall with letter cues. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 13(4):542, 1987.

[PA85] P.L. Pirolli and J.A. Anderson. The role of practice in fact retrieval. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 11(1):136, 1985.

[Pos78] M. I. Posner. *Chronometric Exploration of Mind.* Erlbaum, 1978.

[Rao89] K.V. Rao. Word frequency effect in situational frequency estimation. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 9:73, 1989.

[Rat76] R. Ratcliff. A theory of memory retrieval. *Psychological Review*, 85:59, 1976.

[RM76a] R. Ratcliff and B. B. Murdock. Retrieval processes in recognition memory. *Psychological Review*, 83:190, 1976.

[RM76b] R. Ratcliff and B.B. Murdock. Retrieval processes in recognition memory. *Psychological Review*, 83:190, 1976.

[RU90] E. Ruppin and M. Usher. An attractor neural network model of semantic fact retrieval. *Network*, 1(3):325, 1990.

[Sch89] D.L. Schacter. Memory. In M.I. Posner, editor, *Foundations of Cognitive Science*, page 683. 1989.

[SGB78] S.M. Smith, A. Glenberg, and R.A. Bjork. Environmental context and human memory. *Memory and Cognition*, 6:342, 1978.

[She67] R.N. Shepard. Recognition memory for words, sentences and pictures. *Journal of Verbal Learning and Verbal Behavior*, 6:156, 1967.

[Som86] H. Sompolinsky. Neural networks with nonlinear synapses. *Phys. Rev. A.*, 34:2571, 1986.

[TSS82] E. Tulving, D.L. Schacter, and H. Stark. Priming effects in word fragment completion are independent of recognition memory. *Journal of experimental Psychology: Human Learning and Memory*, 8(4):336, 1982.

[Tul83] E. Tulving. *Elements of Episodic Memory.* The Clarendon Press, 1983.

[Tul85] E. Tulving. How many memory systems are there? *American Psychologist*, 40:385, 1985.

[UF68] B.J. Underwood and J.S. Freund. Errors in recognition learning and retention. *Journal of Experimental Psychology*, 78:55, 1968.

[Wei85] G. Weisbuch. Scaling laws for the attractors of hopfield networks. *J. Physique Lett.*, 46:L–623, 1985.

[Wei86] G. Weisbuch. Neworks of automata and biological organization. *J. Theor. Biol.*, 121:255, 1986.

[YRD79] Y. Yeshurun and N. Richter-Dyn. Analysis of recall and recognition in a certain class of adaptive networks. *Biol. Cybern.*, 32:35, 1979.