A Simple Oscillatory Short-Term Memory

James Reggia[2], Jared Sylvester[2], Scott Weems[1], and Michael Bunting[1]

University of Maryland

[1]Center for the Advanced Study of Language, University of Maryland, College Park, MD 20742
[2]Dept. of Computer Science & UMIACS, A. V. Williams Bldg., University of Maryland, College Park, MD 20742
reggia@cs.umd.edu, jsylvest@umd.edu, sweems@casl.umd.edu, mbunting@casl.umd.edu

Abstract

Oscillatory neural network models have been an increasing focus of study over the last several years. These models consist of recurrent neural networks whose dynamics are characterized by persistent learned/designed rhythmic activity. Here we consider *simple oscillatory memories* for short-term retention of items occurring as temporal sequences. By incorporating decay as well as interference, we find that it is easy to match behavioral data from human subjects recalling temporal sequences under different situations by adjusting a single parameter in the model. These results suggest that simple oscillatory memories capture at least some key properties of human short-term memory, and might be used effectively in future biologically-inspired cognitive architectures.

Modeling Short-Term Memory Using Oscillatory Networks

While neural networks using fixed-point attractor states can be effective as memory models and have generated substantial theoretical and experimental analysis, they are typically limited to maintaining only a single pattern at a time in short-term memory. Further, it is difficult to relate activity in these models to neurobiological systems where rhythmic activity, rather than fixed-point attractor states, is the rule (Buzsaki, 2006). In response to these and other concerns, a number of oscillatory memory models have been created and studied during the last several years. In these models, stored/recalled memory patterns are typically represented as rhythmic network activity in which multiple memory patterns are essentially simultaneously active over the same neural substrate. This is possible because the network's activity oscillates between activity states representing different stored patterns.

A remarkably diverse set of oscillatory memory models exists today. Some models are based on theories about the mechanisms underlying theta/gamma activity in specific brain regions such as the hippocampus or neocortex (Hasselmo et al, 2002; Ingber, 1995; Koene and Hasselmo 2007, Lisman & Idiart, 1995). Other models that use individual spiking neurons are based on more abstract architectures (Raffone & Wolters, 2001), while still others have adopted a higher-level approach such as Wilson-Cowan oscillators (Chakravarthy & Ghosh, 1996; Hayashi, 1994; Wang, 1995).

Here we are specifically concerned with modeling *short-term memory,* the human memory system that retains information over brief time intervals (on the order of seconds) and that has substantial capacity limitations, in contrast to the relatively limitless capacity of more permanent long-term memory (Baddeley, 2000; Cowan, 2001). For example, current evidence suggests that human short-term memory capacity is approximately four items (Cowan et al, 2005).

A particularly simple and elegant approach to creating oscillatory short-term memory models is based on minimally modifying Hebbian associative memories having fixed-point attractor states so that they become oscillatory. In the following, we will refer to all such models as *simple oscillatory memories*. For example, Horn and Usher (1991) produced a simple oscillatory memory by introducing "dynamic thresholds" into Hopfield networks (Hopfield, 1982; Amit, 1989). With this approach, whenever a node has a particular activity level ±1, the threshold of that node gradually changes so that eventually the node switches its activity level to the complementary value. When such a network is presented with an input that is a superposition of multiple stored memories, it is found to oscillate between activity states that represent these individual memories, thereby indicating its recognition/recall of the memories in parallel. Similar behaviors have been produced based upon Hopfield networks modified to use "dynamic synapses" (Pantic et al, 2002) or negative feedback with asymmetric connection weights (Brown & Collins, 2000).

Simple oscillatory memories derived from Hopfield networks are intriguing in their simplicity as models of short-term memory. In this paper, we present a simple oscillatory memory based on dynamic thresholds as used by Horn and Usher (1991), except that with our approach

the model is extended to include rapid decay of connection weights. This weight decay allows the network to have a dependency upon the order in which it sees input stimuli, something that is not the case with classical Hopfield networks. Further, it lets us examine the relative roles of interference and decay as mechanisms underlying forgetting. To evaluate the model, we used data that we collected from human subjects performing a running memory span task (Winder et al, 2009). This task involves rapidly presenting a sequence of stimuli that are to be recalled subsequently by the subject. We found that our model can demonstrate recall performance similar to the behavioral data that we obtained: a capacity limit of approximately three items, and a prominent recency effect. To our knowledge, no previous work has examined how the performance of simple oscillatory memories compares to real-world behavioral data collected from human subjects on a short-term memory task. We summarize our results in the followings and discuss their implications for biologically-inspired cognitive architectures.

Model Formulation

Short-term memory is modeled using a fully connected network of N linear threshold nodes similar to many past neural network models. There are two possible values for the activity state a_i of each node i, -1 and 1 (simplified from biology for computational convenience), and these values change over time as described below. Memory patterns to be stored are essentially arbitrary. We label each individual memory state $\vec{a}$ that is being stored with a specific letter (A – Z).

Connection strengths for this N node network are kept in an N x N weight matrix W, where each weight w_{ij} is a real-valued number. Connection strengths between two nodes are the same in both directions (w_{ij} = w_{ji} everywhere), so W is symmetric. Node activation levels a_i are updated according to a stochastic linear threshold activation rule with the input h_i to node i being a linear weighted sum of the activities of the other nodes minus θ_i, where θ_i is the threshold associated with node i. Note that h_i, a_i, w_{ij}, and θ_i are all functions of time.

Unlike in fixed-point attractor networks where after learning a node's threshold is fixed, the threshold values here are "dynamic", changing with each time step t. For example, when the activation level a_i = +1, this causes the threshold for node i to rise slowly, making it more likely that the node will become negative during the next time step. Similarly, when the value of the node is -1, the threshold drops, making it more likely that the node will become positive. This way, when the network is run for any length of time, the network state can oscillate and explore different patterns stored in its weight matrix.

To simulate the presentation of a sequence of stimuli that are being consecutively stored in a subject's short-term

memory during a running memory span task, the N x N weight matrix W is initialized with $w_{ij} = 0$ for all i and j, and then a sequence of memory states corresponding to that sequence of stimuli is imposed on the network. As each presented memory state is transiently present, the connection strengths w_{ij} are all concurrently updated according to the weight change rule

$$w_{ij} = (1 - k_d)w_{ij}^{old} + \frac{1}{N} a_i a_j (1 - \delta_{ij})$$

where N is the number of nodes in the network, k_d is a decay rate capturing how weights diminish over time ($0 \le k_d < 1$), and δ_{ij} is Kronecker's delta (the latter insures $w_{ii} = 0$ for all i, so weights on self-connections are fixed at zero). For the computational experiments that follow, we used $N = 35$, while k_d values varied between different simulations. The second term on the right side of this weight change rule implements Hebbian weight changes as in many past neural network models of memory. However, our weight change rule is unusual in explicitly incorporating a weight decay factor $-k_d w_{ij}$ that gradually reduces the influences of old memory patterns. It still produces a symmetric weight matrix W with zero entries on the main diagonal. If relatively few memory patterns are in memory at any point in time and the decay rate is small, with more traditional Hopfield networks having constant thresholds the stored memories would typically tend to be fixed point attractor states (energy minima). Thus, when our model starts in an arbitrary initial state $\vec{a}$, its activity state would be expected to change until it reaches one of these stored memory states, but would then not remain fixed in that state due to the dynamic thresholds as explained above. This leads to oscillatory behavior.

Assessment Methods

To assess how well the network's current activity state $\vec{a}$ matches the activity state $\vec{a}^\lambda$ corresponding to one of 26 specific stimuli λ, we first compute a measure of the distance d_λ between $\vec{a}$ and $\vec{a}^\lambda$. This measure is essentially the same as the Hamming distance between two binary numbers. The similarity s_λ of current state $\vec{a}$ to stimulus pattern $\vec{a}^\lambda$ is then computed based on this distance measure, such that the larger the distance the less the similarity (Winder et al, 2009). Measure s_λ lies between 0.0 and 1.0. A value $s_\lambda = 1.0$ at any time step indicates a perfect match between the current state $\vec{a}$ and the stimulus pattern $\vec{a}^\lambda$, while progressively lower values of s_λ indicate progressively worse matches.

The process of testing for the retention of specific stimuli that have just been presented sequentially to the model is done as follows. The network is started in a random initial activity state $\vec{a}$, and then the network's state is allowed to evolve according to the dynamics described above for a 200 time step test period. Because of

the changing nature of the thresholds the network does not reach a fixed state during the test period, but instead typically oscillates between states that are at or close to some of the activity patterns $\vec{a}^{\lambda}$ that were shown to it during the simulated running memory span task. During this testing period, the similarity measure s_{λ} for each of the stimuli λ that were in the just-seen test sequence is recorded at each time step.

In the simulations that follow, we label a specific stimulus λ that was presented to the network as being actively present in memory, and thus recalled by the model, only if s_{λ} reaches a value of 1.0 during the 200 time step testing interval. This means the stimulus must be perfectly recalled by the network during this testing period at least once. We use this strict criterion because the similarity measure of some stimuli that are similar to each other can tend to rise and fall synchronously.

Finally, we used behavioral data that we collected previously on a running memory span task for comparison with the model's performance, roughly following the designs of Pollack et al. (1959) and Bunting et al. (2006). Our human experimental data was obtained from 38 adult subjects (13 females, 25 males, mean age 25) who were shown a rapidly presented, two per second sequence of 12 to 20 randomly ordered stimuli under computer control, and were asked to remember the most recent six items in the order of their presentation[3]. Subjects indicated the stimuli that they recalled by clicking on a subsequent graphical display of all possible stimuli. Recall was measured by assessing accuracy of recall as a function of stimulus position. A stimulus was counted as accurately recalled only if: 1. it was presented in the retention window (i.e., the last six items, depending on instructions), 2. it was correctly recalled by the participant; and 3. it was recalled in the same position as it was presented. Details of the behavioral data collection methodology are given in (Weems et al, 2009).

Model Performance

Following presentation of a sequence of stimuli, network activity measured using s_{λ} oscillates between memory states representing some of the presented stimuli, indicating the retention of those patterns in short-term memory. As an example, suppose that a decay rate of $k_d = 0.2$ is used and a randomly-selected sequence of stimuli M, L, X, N, E, F, H, and B, listed by their labels from first to last, serves as inputs. Typically the oscillations associated with earlier stimuli have relatively small amplitudes (only partial matches), while those of the more recently presented stimuli are more prominent. Figure 1 shows when the oscillations in s_{λ} values peaked at 0.8 or above for the eight stimuli of this example during just the middle of the testing period. Using our criteria that a stimulus λ is retained in short-term memory (recalled) if and only if its activity pattern $\vec{a}^{\lambda}$ occurs exactly ($s_{\lambda} = 1.0$) during the test

period, the input patterns E, H and B would be labeled as recalled. Peaks in the oscillations associated with the different recalled stimuli (E, H and B) alternate with each other, allowing the three remembered stimuli to be retained in short-term memory "simultaneously", unlike with fixed point neural associative memories. Note that the oscillations are irregularly spaced and not periodic.
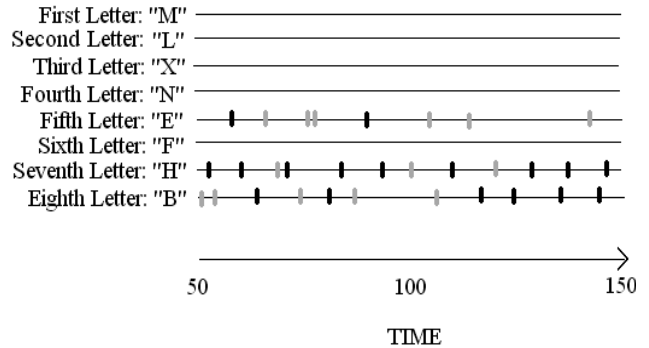

Figure 1: Plot over time of when the values of s_{λ} reached their peaks for the eight stimuli during an example run (see text). Solid black marks indicate when s_{λ} reached the maximum possible value of 1.0, while lighter gray marks indicate when s_{λ} exceeded 0.8 but did not reach 1.0. One can see that the oscillatory states rotate between the three recalled memory patterns for the 5th, 7th and 8th stimuli (E, H and B). Only the middle 100 time steps of the testing period are shown here, expanded horizontally.

Figure 2 displays the fraction of stimuli recalled (maintained as oscillatory states in short-term memory) for stimuli in each position averaged over 1000 random input sequences of 8 different stimuli. When there is no decay ($k_d = 0$), the fraction of input patterns recalled is largely independent of a stimulus' position in the input sequence, resulting in a flat curve. This result can be related to fixed-point attractor networks where the final weights, and hence network performance as an associative memory, is independent of the order in which the input patterns are stored. Further, recall of *any* observed stimulus in this case is quite poor as would be expected; for a network of the size used here the number of patterns used far exceeds the expected memory capacity of an equal size fixed-point attractor associative memory (Hopfield, 1982; Amit, 1989). With no decay, the interference between the stored memory patterns is excessive, preventing almost all stimuli from being retained effectively in short-term memory. In contrast, when the decay rate is very large ($k_d = 0.5$), a very steep curve is seen (Figure 2), with the single most recently presented stimulus always being retained. This occurs because the weight changes from previously stored stimulus patterns quickly dissipate, and even recently presented earlier stored patterns no longer interfere significantly with the final stimulus's retention. In this case, the fraction of the first six presented patterns that are recalled is almost zero, reflecting that they have been

erased from memory. This can be contrasted with the roughly 20% recall rate of presented stimuli when there is no decay at all. Intermediate behaviors are seen for intermediate decay rates, as is shown in Figure 2.
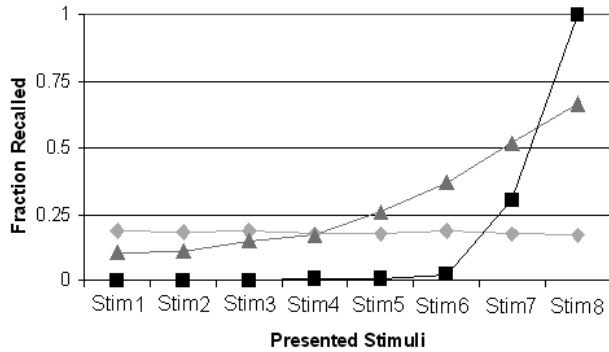

Figure 2: Fraction of stimuli recalled versus stimulus position using different decay rates k_d ($\blacklozenge$ = 0.0, $\blacktriangle$ = 0.1, $\blacksquare$ = 0.5). Each curve plots the fraction of presented stimuli in each position of eight-stimulus sequences that were recalled correctly during a following test period averaged over 1000 trials using the same decay rate. Stimulus 1 was the first stimulus presented, and stimulus 8 the last.
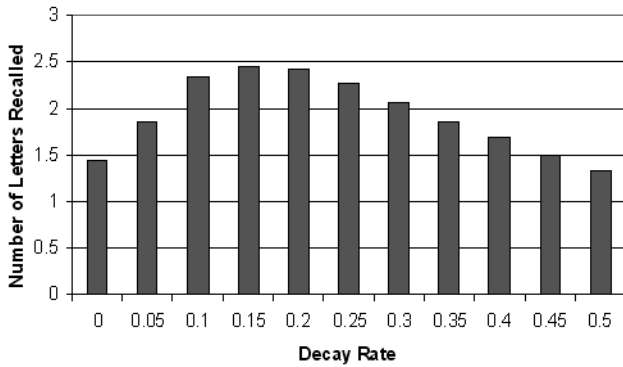

Figure 3: Mean number of stimuli recalled for different decay rates. This peaks at a memory capacity of roughly 2.5 for intermediate decay rate values (around 0.15).

Figure 3 shows, for each decay rate value used, the mean total number of stimuli recalled (memory capacity), averaged over the 1000 stimuli sequences tested. These results show that memory capacity is highest with moderate, intermediate decay rates. If decay is very low, older input stimuli are substantially retained and the attempt to store an excessive number of memory patterns results in too much interference. If decay is too high, even recently observed stimuli are lost. Maximum recall of observed stimuli occurs when there is a balance between information lost due to decay and interference.

How does position-specific recall by the model compare with that exhibited by human subjects? We found that the model is able to approximate closely the position-specific recall rates found with the human subjects when it is run with a 0.1 decay rate with input sequences of length six (Winder et al, 2009). The mean total memory capacity for

recall of six stimuli was 2.73 items for human subjects versus 2.69 items for the model. Thus, both the model's total memory capacity, and its position-specific stimulus retention patterns, were in close agreement with those seen with the human subjects.

Our human subject data was collected under conditions where the subjects were presented with stimulus sequences of 12 to 20 items, and were told to recall only the last 6, regardless of the actual sequence length. To explore how the model was affected by seeing more stimuli than were to actually be recalled, we repeated the simulations as above but now using a total of 20 stimuli in each sequence, even though recall of only the last 6 stimuli was of interest. The model's performance in position-specific recall can, with an adjusted decay rate, still reasonably match the human subject results. The model results are again averaged over 1000 different random sequences of stimuli. In this case we found that a modestly higher decay rate (0.185) provides an approximate match to the six-back human data. The model's mean total memory capacity for recall of six stimuli was now 2.28 items (vs. 2.73 for human subjects).

Discussion

In this work we introduced a simple oscillatory memory model of short-term memory, examined some of its properties, and compared its behavior to that of human subjects on a running memory span task. Our model's dynamics are intrinsically oscillatory due to the use of rapidly varying threshold values, and recall of an item is dependent upon the time elapsed since it was observed due to the use of rapidly decaying weights. Unlike with many past neural models of memory, we assessed recall by initializing the model's activity to a random state, rather than by initializing it to a noisy/partial stored memory pattern, or by biasing the network's dynamics by applying an external input pattern that represents a noisy/partial stored pattern. When moderate decay rates were used, this approach resulted in a short-term memory capacity of between two to three items, a value that is comparable to what has been observed in past experimental studies by others (Baddeley, 2000; Cowan, 2001; Cowan et al, 2005), and that matches the memory capacity that we observed in a group of human subjects performing a similar running memory task. The model also showed a prominent recency effect as would be expected given the use of weight decay, and as is also seen in human subjects.

Our model is intended to simulate short-term memory processing only. It is not intended to address any processes by which semantic or other long-term memory information is accessed to aid storage or recall. It is well established that short-term memory capacity is higher for familiar items compared to novel stimuli (see Cowan, 2001 for review). This is likely due to the fact that for novel stimuli, representations must be created before retention can successfully occur. While the results of our model are

promising in suggesting that oscillatory neural models can show similar capacity limitations as with humans, they do not allow us to make predictions regarding frequency specific contributions to EEG, especially as the model oscillations recorded are in terms of the extent to which a specific distributed memory pattern is present (quantity s_λ) and not in terms of amount of network activity.

Our study adds to a rapidly growing literature on computational models of short-term memory by examining the role of weight decay on simple oscillatory memories. Many past models of short-term memory have employed lateral inhibition between representational units to establish competition between activated entities, and thus capacity limitations (e.g., Haarmann and Usher, 2001). Our approach differs in not explicitly building in such lateral inhibition (although inhibitory weights do occur during pattern storage), with competition between memory patterns arising in the dynamics due to the interference occurring between the non-orthogonal memory patterns. Other recent models of short-term memory, inspired by specific neuroanatomical structures, have used separate modules for memory representation, maintenance, and selective gating. For example, (Frank et al, 2001; O'Reilly & Frank, 2006) incorporate modules representing prefrontal cortex and basal ganglia. Our approach does not use a complex architecture or gating mechanisms, and thus shows that some basic behavioral properties of human short-term memory (limited memory capacity, recency effect, and shifts in position-specific stimulus recall) can be captured by a surprisingly simple neurocomputational mechanism. Still other recent short-term memory models have been based on modulation of persistent neuronal firing by rhythmic changes to membrane potential at theta frequencies (Koene and Hasselmo, 2008). Our approach is quite different in that storage is based primarily upon synaptic connectivity, and memory capacity limitations arise mainly due to synaptic decay and pattern interference.

Perhaps the most interesting finding with the model is that, by adjusting just the weight decay rate, one can produce shifts in the model's memory capacity and position-specific recall rates. This represents a prediction of the model that by adjusting the decay rate, one could reasonably match the shifts exhibited by human subjects who were instructed to recall different length stimuli sequences. This is remarkable, given the simplicity of our model and that it requires adjustment of only a single parameter. This prediction relates to long-standing issues in the cognitive science literature concerning the nature of forgetting. For example, one view of forgetting is that short-term memory is subject to decay (Brown, 1958), while an alternative view is that forgetting is due to interference between competing elements that are simultaneously vying for attention (Waugh & Norman, 1965). Our model incorporates both interference and decay as mechanisms for forgetting, and shows that the latter can partially mitigate effects from interference, consistent with past behavioral studies (Altmann & Gray, 2002).

The observation that adjustments to decay rate control not only the total short-term memory capacity (Figure 5), but also position-specific stimulus recall rates (Figure 4), raises the issue of whether altering decay rate might be a useful mechanism permitting a cognitive system to control short-term memory characteristics. Specifically, our model is consistent with the hypothesis that dynamic adjustments to activity decay rate may be an important aspect of the human attention mechanisms that control forgetting (Altmann & Gray, 2002).

It is already well established that attention is a cognitive property that can be manipulated based on the needs of the task at hand (Broadbent, 1982; Downing & Pinker, 1985; Eriksen & St. James, 1986), and that attentional scope can be adjusted during visual search and memory recall, between being more focused or more diffuse (Engle, 2002; Kane & Engle, 2002). Based on our modeling results, we hypothesize that altering the decay rate could serve as a means via which attentional mechanisms could act to manipulate attentional scope. More focused attention is simulated in the model by a higher decay rate, so that attention is directed more intently on a smaller number of items. In this way, decay is used as a means for combating proactive interference, with higher decay rates leading to more effective retention of recent information, but also at the expense of that which was presented before it.

For the running memory span task used here involving rapid presentation of stimuli, human subjects attempt to hold presented stimuli in a limited capacity memory without the use of rehearsal (Bunting et al, 2006). Assuming that maintaining such stimuli depends on attentional resources, then changing instructions requiring subjects to retain varying numbers of stimuli (i.e., not just six as we did in our behavioral experiments) would be expected to have a great effect. Specifically, if attention is drawn sufficiently thin so that activation maintenance is small across all retained stimuli (a low decay rate in our model), then with longer stimulus sequences (e.g., a task requiring human subjects to recall 12 stimuli) few/none of the stimuli would be expected to retain activation levels above some cognitive threshold required for successful recall due to interference, although no doubt some attenuated recency effect will still be present. This is both a surprising and informative prediction from the model, and it suggests that overloading subjects' attentional resources, i.e., drawing attention sufficiently thin, has a detrimental effect on retention. Future behavioral testing with varying length recall task could therefore either refute or strongly support the model we have presented here. Future computational studies will explore this issue as well as the ordered recall of stored stimuli.

References

Altmann, E., and Gray, W. Forgetting to remember: the functional relationship of decay and interference. *Psychological Science.* 13 (1), 27-33, 2002.

Amit D. *Modeling Brain Function*. Cambridge University Press, 1989.

Baddeley A, Short-Term and Working Memory, in *The Oxford Handbook of Memory*, E. Tulving and F. Craik (editors), Oxford University Press, 77-92, 2000.

Broadbent D. Task combination and selective intake of information, *Acta Psychologica*, 50, 253-290, 1982.

Brown A, Collins S. An oscillatory associative memory analogue architecture, in *Oscillations in Neural Systems*, D. Levine et al (eds.), Lawrence Erlbaum, 327-341, 2000.

Brown J. Some tests of the decay theory of immediate memory. *Quarterly J. Exper. Psych..* 10, 12-21, 1958.

Bunting, M., Cowan, N., and Saults, J. How does running span work? *Journal of Experimental Psychology: Human Perception and Performance.* 59 (10), 1691-1700, 2006.

Buzsaki G. *Rhythms of the Brain*, Oxford Univ., 2006.

Chakravarthy S, Ghosh J. A complex-valued associative memory for storing patterns as oscillatory states, *Biological Cybernetics*, 75, 229-238, 1996.

Chen K, Wang D. A dynamically coupled neural oscillator network for image segmentation, *Neural Networks*, 15, 423-439, 2002.

Cowan, N. The magical number 4 in short-term memory. *Behavioral and Brain Sciences.* 24, 87-185, 2001.

Cowan, N., Elliot, E., Saults, J., et al. On the capacity of attention. *Cognitive Psychology.* 51, 2005, 42-100.

Deb K. Multi-Objective Optimization, in *Multi-Objective Optimization Using Evolutionary Algorithms*, K. Deb, John Wiley & Sons, 2001, pp. 13-46.

Downing C, Pinker S. The spatial structure of visual attention. In: M. Posner & O. Marin (eds.), *Attention and Performance*, XI, Erlbaum, 171-188, 1985.

Engle R. Working memory capacity as executive attention, *Current Directions in Psychological Sci.*, 11, 19-23, 2002.

Eriksen C, St. James J. Visual attention within and around the field of focal attention, *Perception and Psychophysics*, 40, 225-240, 1986.

Frank, M., Loughry, B., and O'Reilly, R. Interactions between frontal cortex and basal ganglia in working memory: a computational model. *Cognitive, Affective, and Behavioral Neuroscience.* 1, 137-160, 2001.

Haarmann, H., and Usher, M. Maintenance of semantic information in capacity-limited short-term memory. *Psychonomic Bulletin.* 8 (3), 568-578, 2001.

Hasselmo M, Bodelon C, Wyble B. Proposed function for hippocampal theta rhythm, *Neural Comp*, 14, 793-817, 2002.

Hayashi Y. Oscillatory neural network and learning of continuously transformed patterns, *Neural Networks*, 7, 219-231, 1994.

Haykin, S. *Neural Networks*. Prentice Hall, 1999.

Hopfield, J. Neural networks and physical systems with collective computational abilities. *Proceedings of the National Academy of Science.* 79, 2554-2558, 1982.

Horn D, Usher M. Parallel activation of memories in an oscillatory neural network. *Neural Comp.* 3, 31-43, 1991.

Ijspeert A. A connectionist central pattern generator for the aquatic and terrestrial gaits of a simulated salamander, *Biological Cybernetics*, 84, 331-348, 2001.

Ingber L. Statistical mechanics of neocortical interactions: Constraints on 40-hz models of short term memory, *Physical Review E*, 52, 4561-4563, 1995.

Kane M, Engle R. The role of prefrontal cortex in working memory capacity, executive attention, and general fluid intelligence, *Psychonomic Bulletin Rev.*, 9, 637-671, 2002.

Kimura H, Fukuoka Y, et al. Three-dimensional adaptive dynamic walking of a quadruped, *Proc. 2002 IEEE Intl. Conf. Robotics and Automation*, 2228-2233, 2002.

Koene, R., and Hasselmo, M. First-in-first-out replacement in a model of short-term memory based on persistent spiking, *Cerebral Cortex*, 17, 1766-1781, 2007.

Koene, R., and Hasselmo, M. Consequences of parameter differences in a model of short-term persistent spiking buffers provided by pyramidal cells in entorhinal cortex, *Brain Research*, 54-67, 2008.

Lisman J, Idiart M. Storage of 7 ± 2 short-term memories in oscillatory subcycles, *Science*, 267, 1512-1516, 1995.

Nunez P, Srinivasan R. *Electric Fields of the Brain*, Oxford University Press, 2006.

O'Reilly R., Frank M. Making working memory work: a computational model of learning in the prefrontal cortex and basal ganglia. *Neural Computation.* 18, 283-328, 2006.

Pantic, L., Torres, J., Kappen, H., and Gielen, S. Associative memory with dynamic synapses. *Neural Computation.* 14, 2903-2923, 2002.

Pollack, I., Johnson, I., Knaff, P. Running Memory Span. *Journal of Experimental Psychology.* 57, 137-146, 1959.

Raffone A, Wolters G. A cortical mechanism for binding in visual working memory, *Journal of Cognitive Neuroscience*, 13, 766-785, 2001.

Wang D. Emergent Synchrony in locally coupled neural oscillators, *IEEE Trans. Neural Netw.*, 6, 941-947, 1995.

Waugh N, Norman D. Primary memory, *Psychological Review*, 72, 89-104, 1965.

Weems S, Winder R, Bunting M, Reggia J. Running Memory Span, *Cognitive Systems Res.*, 10, 2009, 161-171.

Winder R, Reggia J, Weems S, Bunting M. An Oscillatory hebbian network Model of Short-Term Memory, *Neural Computation*, 21, 2009, 741-761.