

A Balanced Hebbian Algorithm for Associative Learning in ACT-R

Robert Thomson (thomsonr@andrew.cmu.edu),

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA, 15213, USA

Christian Lebiere (cl@cmu.edu)

Department of Psychology, Carnegie Mellon University
5000 Forbes Avenue, Pittsburgh, PA, 15213, USA

Abstract

Associative learning is a mechanism ubiquitous throughout human and animal cognition, but which is absent in ACT-R 6. Previously, ACT-R 4 had implemented a Bayesian learning algorithm which derived the strength of association between two items based on the likelihood that one item was recalled in the context of the other (versus being recalled outside of this context). This algorithm suffered from asymmetries which tended to lead all associations to become strongly inhibitory the longer a model ran. Instead, we present a Hebbian learning algorithm inspired by spiking neurons and the Rescorla-Wagner model of classical conditioning, and show how this mechanism addresses asymmetries in the prior Bayesian implementation. In addition, we demonstrate that balanced learning of both positive and negative associations is not only neurally- and behaviorally-plausible, but has benefits in both learning and in constraining representational complexity. This is demonstrated using a simple model of list learning derived from Anderson et al. (1998).

Keywords: cognitive architectures; Hebbian learning; associative learning; representation; list learning

Introduction

Associative learning occurs when a relationship is formed between two possibly-unrelated items that are presented in a close temporal proximity. This kind of learning is flexible and experience-driven, and is implicated in many major cognitive phenomena such as classical and operant conditioning (Rescorla & Wagner, 1972), context sensitivity (Wickelgren, 1969), sequence learning (Gheysen & Fias, 2012), and expectation-driven learning (Lukes, Thompson, & Werbos, 1990). It is describable as both a micro (e.g., neural specialization) and macro (e.g., behavioral conditioning) function of both human and animal cognition.

Despite being an essential element of many basic cognitive functions, associative learning is not currently implemented in recent versions of cognitive architectures such as Soar (Laird, 2012) and ACT-R 6 (Anderson et al., 2004). Previously, ACT-R 4 (Anderson & Lebiere, 1998) had implemented Bayesian-inspired associative learning which derived the strength of association between two items based on the likelihood that one item was recalled in the context of the other (versus being recalled outside of this context). This implementation, however, suffered from some asymmetries which led to issues with stability.

We argue that difficulties in balancing the Bayesian-inspired learning were due to the asymmetry in log-likelihood calculations and by deriving associative strength from a global variable such as the number of times an item was recalled. Instead, we present an associative learning algorithm derived from both neural evidence from spiking neurons (Caporale & Dan, 2008) and behavioral evidence

from the Rescorla-Wagner (1972) model of classical conditioning.

For the remainder of this paper, we will provide an overview of the neural and behavioral evidence used to derive our associative learning algorithm, argue how this algorithm avoids the instabilities of ACT-R 4's Bayesian algorithm, and present a sample ACT-R model using associative learning in a list learning paradigm.

Neural Evidence (AKA: the Micro Level)

At the micro level, associative learning is the process by which neuronal firing patterns are specialized, such that *neurons that fire together, wire together* (Hebb, 1949). In its simplest form, hebbian learning can be described by:

$$(1) \quad \Delta S_{ji} = x_i x_j$$

where S_{ij} is the synaptic strength of the connection between an input neuron i and output neuron j , and x_i and x_j are the inputs to i and j (Hebb, 1949). When both i and j are active together in a given temporal window, S_{ij} is strengthened.

While this simple hebbian rule is unstable due to a lack of mechanisms to control for weakening of connections or to set a maximum state of activation (e.g., a softmax function; Sutton & Barto, 1998), several variants such as Oja's rule (Oja, 1982) and the generalized hebbian algorithm (Sanger, 1989) have addressed these issues to provide a stable learning rule across multiple output neurons:

$$(2) \quad \Delta S_{ji} = \alpha (x_i x_j - x_j \sum_{k=1}^j w_{ik} x_k)$$

where S_{ij} is the synaptic strength of the connection between neurons i and j , and x_i and x_j are the vectors of i and j , and α is the learning rate.

More recently, improvements in measuring techniques have found evidence that the timing of the pre-synaptic input neuron i and post-synaptic output neuron j firings influence whether the likelihood of j firing in the presence of i increases (hebbian learning) or decreases (anti-hebbian learning). More specifically, when the pre-synaptic neuron fires just before (5 - 40ms) the post-synaptic neuron then their association is strengthened, whereas if the post-synaptic neuron fires just before the pre-synaptic neuron then their association is weakened (Bi & Poo, 1998; Dan & Caporale, 2008). This timing creates a neurally-driven temporal coding (Gerstner et al., 1999), and has been experimentally reproduced in human hippocampal cells (Lamsa et al., 2007; Wittenberg & Wang, 2006). While an understanding of the mechanisms of spike-timing is incomplete, the flow of calcium at voltage-specific calcium channels and NMDA receptors at the post-synaptic neuron are a consistent finding amongst researchers, with the same receptors triggering both hebbian and anti-hebbian learning.

Behavioral Evidence (AKA: the Macro Level)

At the macro level, associative learning occurs when a stimulus is reliably paired with another stimulus or behavior. In terms of classical conditioning, an association is formed when a neutral stimulus is reliably paired with an unconditioned stimulus which produces a response. Eventually the neutral stimulus alone is able to elicit the response, at which point it is considered a conditioned (i.e., learned) stimulus. In the canonical example, when Pavlov rang a bell (a neutral stimulus) just before feeding his dogs (an unconditioned stimulus) then the dogs would salivate (a response). He discovered that the dogs began to salivate before receiving the food, and determined that the dogs had learned to associate the previously-neutral act of ringing the bell (now a conditioned stimulus) with the imminent presentation of food. Further research determined that the conditioned stimulus *predicting* the unconditioned stimulus drove learning, rather than simply their co-occurrence.

Much like the shift in thinking from Hebbian learning to timing-dependent spiking neurons, Chang, Stout, and Miller (2004) varied whether a conditioned stimulus was presented either before (forward conditioning) or after (backward conditioning) an unconditioned stimulus. Similar to the spiking neuron account, in the forward conditioning trials, the conditioned stimulus tended to be positively associated with the unconditioned stimulus, whereas in the backward conditioning trials the conditioned stimulus tended to be negatively associated with the unconditioned stimulus.

Computational models of classical conditioning are a common macro-level application of associative learning. Perhaps the most well studied implementation is the Rescorla-Wagner model (1972), which is based on the following equations:

$$(3) \quad \Delta S_x^{t+1} = \alpha_x \beta (\gamma - S_{total})$$

$$(4) \quad S_x^{t+1} = S_x^t + \Delta S_x^{t+1}$$

where ΔS_x is the change in strength of the association, α_x is the salience of the conditioned stimulus, β is the salience of the unconditioned stimulus, γ is the maximum amount of learning that can occur, and S_{total} is the total associative strength of all conditioned stimuli. Learning is proportional to the degree that the conditioned stimulus predicts the unconditioned stimulus, which is set by the salience parameters. Thus, when there is no prediction error (i.e., no violation of expectation) the salience of the unconditioned stimulus is low, and no significant learning occurs. However, there is no *a priori* justification to initially set saliences other than to tune them to fit data patterns. Thus, the model is effectively explanatory rather than predictive.

In summary, at the behavioral level associative learning allows humans to predict outcomes based on prior experience, with learning mediated by the saliency of the stimuli and the mismatch between the predicted outcome and the actual result. There is also a strong analogy between forward and backward conditioning with the balanced hebbian and anti-hebbian learning seen in spiking neurons. We argue that an associative learning algorithm should be consistent with these balanced timing-based approaches.

The ACT-R Architecture

ACT-R is a computational implementation of a unified theory of cognition which accounts for information processing in the mind via task-invariant mechanisms constrained by the biological limitations of the brain. ACT-R 6 includes declarative memory and perceptual-motor modules connected through limited-capacity buffers. Each module exposes a buffer, which contains a single chunk, to the rest of the system. A chunk is a member of a specific chunk type, and consists of a set of type-defined slots containing specific values.

The flow of information is controlled by a procedural module which is implemented by a production system. Modules process information in parallel with one another. For instance, the visual and motor modules may both operate at the same time. There are two serial bottlenecks: 1) only one production may execute during a cycle; and 2) each module is limited to placing a single chunk in a buffer.

Each production consists of if-then condition-action pairs. Conditions are typically criteria for buffer matches, while the actions are typically changes to the contents of buffers that might trigger operations in the associated modules. The production with the highest utility is selected to fire from among the eligible matching productions. Production utilities are learned using a reinforcement learning scheme.

When a retrieval request is made to declarative memory (DM), the most active (highest A_i) matching chunk is returned according to the following function:

$$(5) \quad A_i = B_i + S_i + P_i + \epsilon_i$$

where activation A_i is computed as the sum of base-level activation (B_i), spreading activation (S_i), partial matching (P_i) and stochastic noise (ϵ_i). Spreading activation propagates activation from the contents of buffers to declarative memory proportionally to the strength of association between buffer contents and memory chunks. Partial matching allows for chunks in memory that do not perfectly match a retrieval request to be recalled if their activation overcomes a similarity-based mismatch penalty.

Bayesian Associative Learning in ACT-R 4

Associative learning was deprecated in ACT-R due to a lack of scalability in spreading activation as the number of chunks in a model increased and as new productions fired (i.e., new contexts generated). The reason for this was that the Bayesian associative learning algorithm would render larger and longer-running models unstable. In ACT-R 4, the strength of association (S_{ji}) represented the log-likelihood that chunk N_i was relevant given context C_j :

$$(6) \quad S_{ji} = \ln \left(\frac{P(N_i|C_j)}{P(\bar{N}_i|C_j)} \right) = \frac{P(N_i)}{P(\bar{N}_i)} \prod_j \frac{P(C_j|N_i)}{P(C_j|\bar{N}_i)}$$

When C_j is usually not in the context when N_i is needed, $P(N_i|C_j)$ will be much smaller than $P(\bar{N}_i|C_j)$ and the S_{ji} will be very negative because the log-likelihood ratio will approach 0 (see Figure 1). In a long-running model with many chunks, these chunks may have been recalled many times without being in context together, leading to strongly inhibitory S_{ji} .

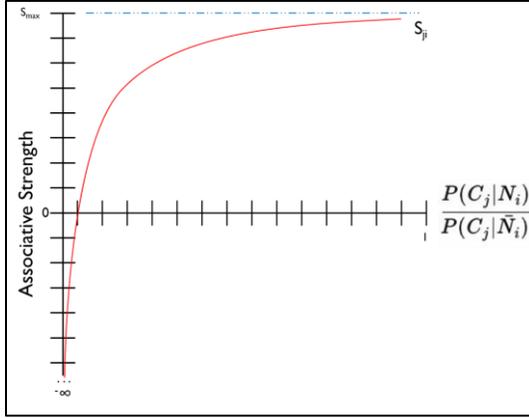


Figure 1. The Influence of Retrieval Context on S_{ji}

Once a connection between chunks j and i was made, the initial S_{ji} was set by the following equation:

$$(7) \quad S_{ji} = \ln(m/n)$$

where m is the total number of chunks in memory and n is the number of chunks which contain the source chunk j . This ratio is an estimation of the likelihood of retrieving chunk i when j is a source of activation. Unconnected chunks were set at 50% likelihood because, before C_j appears in a slot of N_i , the total probability of retrieving a chunk unconnected to C_j is 0 ($S_{ji} \rightarrow -\infty$).

As can be seen from Equations 6 and 7, given sufficient experience or sufficient numbers of chunks in the model, S_{ji} will become increasingly and unboundedly negative as more chunks are learned and more unique contexts experienced. This is a direct result of S_{ji} reflecting the statistics of retrieval of chunk j given that source i is in the context. The issue is with the ratio-driven global term (C_j) which alters S_{ji} values for a chunk whenever a new chunk is added and/or production fires, which means that learning an unrelated fact causes all existing associations to weaken. This effect is further magnified by the log-likelihood calculation which penalizes the inevitable low context ratio.

Spreading Activation Rule in ACT-R 6

Due to the abovementioned issues with stability, associative learning was deprecated in ACT-R 6 and a spreading activation function was implemented that does not learn activation, but instead spreads a fixed amount of activation:

$$(8) \quad S_{ji} = \text{smax} - \ln(\text{fan}_{ji})$$

where smax is a parameterized fixed spread of association (replacing the m term from the previous equation), and fan_{ji} is the number of chunks associated with chunk j (the n term). More specifically, fan_{ji} is the number of times chunk j is a slot value in all chunks in DM and represents the effect of interference. It is important to note that under this definition, activation is only spread symbolically by overlapping slot contents, and not due to experience.

With a default smax usually between 1.5 and 2 (Lebiere, 1999), a given chunk can appear as a slot value in 6 chunks before its strength becomes inhibitory (see Figure 2). In the context of modeling a single session psychology experiment this may be reasonable, but if ACT-R is used to model long-

term knowledge effects, then the S_{ji} term will become inhibitory for most commonly-used chunks. One solution (applied after a version of this argument was presented at the ACT-R 2012 Workshop) was to institute a floor of zero activation to the strength of association.

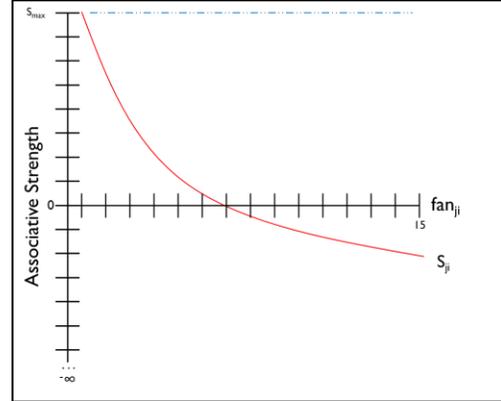


Figure 2. The Influence of Fan_{ji} on Strength of Association

Another difficulty is that spreading activation has no mechanism to chain together sequences of events due to the fact that a source chunk in a buffer only spreads activation to target chunks in DM which match slot values in source chunk's contents (i.e., the chunk *tree* with slot values *branches* and *green* only spreads activation to the chunks *branches* and *green* in DM). Due to this limitation, it is very difficult to chain together sequences of retrievals without including additional slots with contextual and/or indexical information. Furthermore, since there is no learning of associations, it is up to the modeler (as opposed to the architecture) to determine the nature of the indexical/contextual information. Finally, fan_{ji} is a global term (i.e., it changes based on adding new chunks to the model) thus changing the S_{ji} between two chunks without either chunk ever being retrieved. For these reasons, we argue that such contextual information should be implicit, sub-symbolic, and driven by experience; rather than be explicit and symbolic as is currently implemented in ACT-R 6.

A Balanced Associative Learning Rule

As previously discussed, associative learning is ubiquitous in both human and animal cognition, and serves as a kind of statistical accumulator which is applicable at both the micro (neural) and macro (behavioral) level. In abstracting away this essential learning mechanism, we are losing out on the exact kind of human-model comparisons that might lead to more general models of learning (as opposed to task-specific models). Perhaps, it is in part for this reason that ACT-R (and other architectures) have had their generative power limited due to a lack of newer, more complex models being built from successful extant models (ACT-R Workshop, 2012). To reconcile the difficulties in prior implementation of associative learning, we will now present an associative learning algorithm inspired by spiking neurons and models of classical conditioning.

The major issues with ACT-R 4's Bayesian associative learning rule were the reliance on asymmetric log-

likelihoods and the fact that context (C_j) was a global term which altered S_{ji} whenever a new chunk was created and whenever a production fired. Low log-likelihoods become strongly inhibitory, and the generation of context-based ratios necessitates low likelihoods in a long-running model. An alternative framework is to eliminate the ratio function and remove the global nature of the context term, while also moving away from of probability-based algorithm.

By switching to a frequency-based algorithm, it is possible to reshape the range of S_{ji} values and make S_{ji} independent of changing global context. Learning, rather than being a global property of the system, is now a local property based on co-occurrence and sequential presentation. As previously discussed, our algorithm is influenced by time-dependent spiking neurons and timing-dependent forward and backward conditioning. Unlike traditional Hebbian implementations which simply give a bump to association so long as the pre-synaptic and post-synaptic neurons both fire within a given temporal window, with spiking neurons, if the pre-synaptic neuron fires before the post-synaptic then the association is strengthened (long-term potentiation; LTP). Conversely, if the post-synaptic neuron fires before the pre-synaptic then the association is inhibited (long-term depression; LTD).

This theory of neural plasticity was adapted to our modeling approach by assuming that the set of chunks (and their contents) in all ACT-R buffers at the time a request is initiated (such as a declarative retrieval) act analogously to pre-synaptic firings, and the set of chunks in the buffers at the time the request is completed is analogous to post-synaptic firings. In behavioral terms, the requested chunk is analogous to the conditioned stimulus, the set of chunks at the time the request is initiated are analogous to the unconditioned stimuli, and the set of chunks at the time the request is completed are analogous to the response.

The Associative Learning Algorithm

The associative learning algorithm is implemented as an ACT-R module which detects when a request is made to any module (by the request-detected event), stores the current context (i.e., state of chunks in all the buffers), and when the request is fulfilled (by a scheduled set-buffer-chunk event) learns a positive association between the chunks at the time the request is initiated and a negative association between the chunks at the time the request is completed. Note this learning may not be on the next production fired (as requests may span the time-course of several productions), thus the context may be markedly different. Furthermore, the chunks in the buffer must be a match (i.e., equal-chunks) to a chunk already existing in DM in order to be a source.

Associations are by default buffer-specific, thus an object only experienced by the visual buffer would only spread activation when in the visual buffer, and not if the same object were moved to the imaginal buffer. This buffer-specificity is important for neural plausibility and to distinguish chunks as being cues from extrinsic (e.g., sensory) and intrinsic (e.g., memory) sources. Furthermore, associations are learned asymmetrically ($S_{ji} \neq S_{ij}$).

Formally, the strengths of association are learned according the following algorithm:

$$(9) \quad \Delta S_{ji} = F(N_i|C_j^{pre}) - F(N_i|C_j^{post})$$

$$(10) \quad S_{ji} = \alpha S_{ji}^{prior} + \Delta S_{ji}$$

where S_{ji} are the strengths of association between contextual chunks j and the successfully requested chunk i , α is a learning rate parameter, $F(N_i|C_j^{pre})$ is the context of chunks C_j^{pre} at the time of the retrieval request for chunk N_i , and $F(N_i|C_j^{post})$ is the context of chunks C_j^{post} after chunk N_i has been retrieved. Context is defined as the set of valid chunks in each of ACT-R's buffers.

When a request is initiated the chunk in the requested buffer is cleared. An important modeling decision was to include the successfully requested chunk in its post-request context. Thus a chunk learns to be self-inhibitory. We argue that this self-inhibition, while seemingly unintuitive, is actually a relevant and necessary mechanism to account for local refractory periods in neural firings and for driving expectation-driven behavior. Self-inhibition and forward-chaining positive association have the effect of inhibiting the current context and spreading activation to the next experienced context, thus driving a form of temporal sequencing between requests. This feature is essential for expectation-driven learning (such as that needed in sequence learning paradigms) and provides an implicit measure for traversing a list in serial recall paradigms without the need for extraneous contextual/ indexical markers.

The learning rate parameter controls the maximum strength of association by introducing an interference-based squashing mechanism, as opposed to a decay-based mechanism such that seen as in base-level activation. At each learning step (i.e., at each completed request), the equation is updated by squashing the prior S_{ji} by the learning rate before adding a positive and/or negative increment. At a value of .85, the maximum S_{ji} is approximately 6.6 per source, which is similar to the practical maximum of the base-level activation equation. In practice, the S_{ji} values should not be so high.

Much like the base-level activation equation, S_{ji} are by default incremented or decremented by a value of 1; however, when j is a slot value of a chunk in a buffer, then the increment/decrement is divided by the number of valid slots in the buffer's chunk (i.e., slots whose values match chunks in DM). Thus, by having more slots in a target chunk, ΔS_{ji} is reduced. It is possible, by having a chunk with many slots, to reduce the S_{ji} between a strongly associated slot and the retrieved chunk even during a positive association (due to the squashing influence of the learning rate). Thus, it is important to mediate the number of extraneous slots to not dilute the effect of learning.

Note that only changes in context will have a net ΔS_{ji} due to the balanced positive and negative learning phase. That is, the only significant learning that occurs at a successful recall is where the contents of the buffers have changed between the initiation and completion of the request. In the module code, the negative learning is applied serially before the positive learning, which leads to a very small positive

association between chunks that occur in both the pre and post context. This is due to the serial application of the squashing mechanism (α), and leads to only a very small association (asymptotes at .54 after 30 – 50 presentations).

While the balanced hebbian/anti-hebbian phase at each request is analogous to the positive and negative associations seen at both the neural and behavioral level, the learning rule is also similar to the Rescorla-Wagner rule (compare Equations 3 and 4 to Equations 9 and 10). Specifically, the strength of association in the Rescorla-Wagner rule is derived by the salience of the conditioned and unconditioned stimulus, the learning rate, and the difference between the current and maximum level of activation. Our associative learning algorithm utilizes the learning rate to both determine the maximum strength of association and control the degree of learning (lower rates = slower learning and learning with a lower theoretical maximum). The salience parameters of Rescorla-Wagner are implicitly captured by the base-level activation of the source and requested chunks, in addition to any mismatch penalty driven by the difference between the requests chunk specification and the chunk successfully requested.

Variants of Learning and Spreading Activation

There are several variants of associative learning and spreading activation which may be enabled. It is possible to learn associations between elements in context and a successfully requested chunk: 1) from a chunk in a buffer (i.e., a context chunk) to the retrieved chunk (C-C); 2) from a chunk in a buffer (i.e., a context chunk) to the slots of the requested chunk (C-S); 3) from the slot of a context chunk to the retrieved chunk (S-C); and 4) from the slots of a context chunk to the slots of the retrieved chunk (S-S). In the case of S-C and C-S variants, the learning increment for each association is mediated by the number of valid slots in either the source chunk or the retrieved chunk's slots, respectively (see Figure 3). In the S-S variant, the increment is mediated by both the number of slots in the source chunk and the number of slots in the retrieved chunk (thus 3 valid slots in source chunk j and 4 valid slots in retrieved chunk i will result in a learning increment of $1/12$ per S_{ji}).

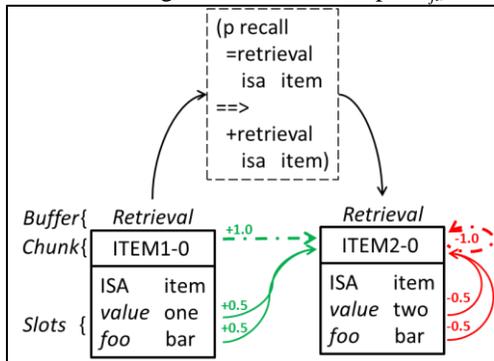


Figure 3. Hebbian (positive) and anti-Hebbian (negative) associative learning after the successful retrieval of ITEM2-0 when ITEM1-0 was in the retrieval buffer at the time the retrieval was initiated. Solid lines represent S-C learning and dotted lines represent C-C learning.

In the ACT-R 4 associative learning algorithm, only slot contents in the goal buffer were associated with the chunk in the retrieval buffer (i.e., S-C learning). Further testing of the mechanism will determine the best learning variants to be enabled; however, ideally all four kinds of learning will be enabled by default in the final release.

Similarly, spreading activation (by default) is handled similarly to the original mechanism in ACT-R 6 where the slots of chunks in buffers spread activation to chunks in DM (consistent with S-C and S-S spread). This decision was made to directly compare the predictions of the original and updated implementations. Additionally, it is also possible to enable the full chunk in each buffer to spread activation as well (see C-C and C-S), which has benefits in tasks where one is recalling an ordered list of chunks, as there is a more direct associations between sequences of chunks passing through a buffer (generally the retrieval buffer).

Testing the Model: Serial Order Memory

There is evidence that humans do not explicitly encode positional information when sequentially recalling a list of items (Henson, 1998), yet a previous implementation of serial memory in ACT-R utilized mismatches over explicit positional information to drive recall (Anderson et al., 1998). While very successful at fitting human data, this model (much like models of classical conditioning) was not predictive, only explanatory. The goal of the present model was to demonstrate a proof-of-concept that serial order effects can be explained as an implicit process of learning inter-item associations, rather than applying a mismatch penalty to explicit positional indices.

The data to be modeled include the single-digit stimuli (from 1-9) and methodology used in Anderson et al. (1998), including the same presentation time used in their human trials (1 second per stimulus, presented serially on a display). Our model was serially presented single-digits in the center of a 300x300 pixel window. The model perceived the stimuli, stored them to memory, and upon perceiving the recall cue (“forward” presented on the display) was given 10 seconds to recall the list. The model was run on lists of length 1 to 9 with no repeating digits, and was run for 100 trials at each list length. The results of both aggregate model and human performance are presented in Figure 4 and describe a close fit to human data ($r^2 = .977$).

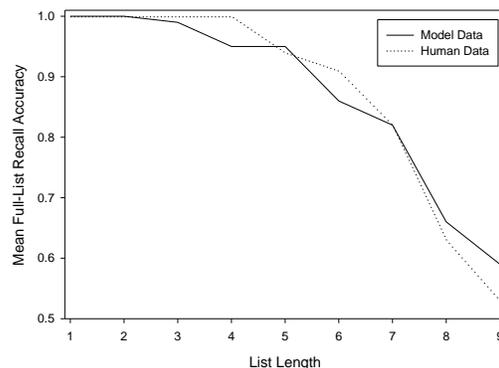


Figure 4. Comparing aggregated full-list model and human results on a serial recall task. Model fit was $r^2 = .977$.

The current model utilizes the same kind of associative learning and spreading activation as in the 1998 ACT-R 4 model, however, the current model required no mismatch penalty over explicit positional indices to recall the complete list with similar accuracy as humans. One modeling assumption was to explicitly mark the beginning and end of the list. The beginning-of-list marker spreads activation to the first list element to recall, which in turn inhibits itself and spreads activation to the next element in the list. The final list element spreads activation to the end-of-list marker. Recall is a matter of retrieving the beginning-of-list marker, which then drives an implicit temporal sequencing of list elements. Finally, the choice of rehearsal strategy drives the strengths of associations formed (implicitly through the ordering of requests), which is left up to the modeler to determine via productions.

Discussion

There are several beneficial effects seen in the current associative learning implementation. The first is that the algorithm is more balanced and geared towards specializing associative activations rather than just increasing all activations. Thus, the mechanism is more stable as it grows (i.e., it will not tend towards all associations becoming either strongly excitatory or inhibitory; S_{ji} doesn't vary with number of chunks in memory). Second, since the retrieved chunk self-inhibits, this reduces the likelihood that it will be the most active chunk in the following retrieval, causing a natural refractory period which avoids self-feedback loops. Third, by self-inhibiting and spreading activation to the next context, there is a forward momentum for the serial recall of chunks. Combined with recency and frequency of base level activation, this provides a mechanism for automatic serial recall of lists without the need for coding of explicit positional information and marking of previously retrieved chunks through FINST-like mechanisms. The uniqueness of the subsequent context drives order effects.

This associative learning mechanism is an attempt to increase constraints within the ACT-R architecture and promote a broader explanatory power to numerous cognitive phenomena. A major contribution of this mechanism is its balance between hebbian and anti-hebbian learning, which provides stability and benefits over prior implementations.

Future Work

The current associative learning algorithm is still under development, with several points to focus on. Currently, slots are equally weighted and a set increment of 1 unit of association is spread by default. It may be possible to determine the strength of association increment based in part on retrieval time, which would be more consistent with both neural and behavioral data. However, in ACT-R retrieval time is totally determined by the base-level activation of the retrieved chunk. Thus, this would reduce associative learning for highly activated chunks, which is not necessarily a desirable outcome. Furthermore, while not applicable to all requests, the mismatch penalty between the request specification and the actual chunk recalled may be

used to mediate the degree of learning. Finally, existing models of list memory and sequence learning need to be re-envisioned in terms of the new associative learning algorithm to determine the best parameters to set by default.

Acknowledgments

This work was conducted through collaboration in the Robotics Consortium sponsored by the U.S Army Research Laboratory under the Collaborative Technology Alliance Program, W911NF-10-2-0016; and by Intelligence Advanced Research Projects Activity, via DOI contract D10PC20021. The views and conclusions contained herein are those of the authors and should not be interpreted as representing the official policies or endorsements, either expressed or implied, of IARPA, ARL, DOI, or the U.S. Government.

References

- Anderson, J. R., Bothell, D., Byrne, M. D., Douglass, S., Lebiere, C., Qin, Y. (2004). An integrated theory of Mind. *Psychological Review*, 111, 1036-1060.
- Anderson, J. R., Bothell, D., Lebiere, C. & Matessa, M. (1998). An integrated theory of list memory. *Journal of Memory and Language*, 38, 341-380.
- Anderson, J. R., and Lebiere, C. (1998). *The atomic components of thought*, Erlbaum: Mahwah, NJ.
- Caporale, N., & Dan, Y. (2008). Spike Timing-Dependent Plasticity: A Hebbian Learning Rule. *Annual Review of Neuroscience*, 31, 25-46.
- Chang, R. C., Stout, S., & Miller, R. R. (2004). Comparing excitatory backward and forward conditioning. *Quarterly Journal of Experimental Psychology*, 57 (1), 1-23.
- Dayan, P., Kakade, S., & Montague, P. R. (2000). Learning and selective attention, *Nature Neuroscience*, 3, 1218-1223.
- Deneve, S. (2008). Bayesian spiking neurons II: Learning. *Neural Computation*, 20, 118-145.
- Gheysen, F., & Fias, W. (2012). Dissociable neural systems of sequence learning. *Advances in Cognitive Psychology*, 8 (2), 73-82.
- Hebb, D. O. (1949). *The Organization of Behavior: A neuropsychological theory*. New York: Wiley.
- Henson, R. N. A. (1998). Short-Term Memory for SerialOrder: The Start-End Model. *Cognitive Psychology*, 36,73-137.
- Lamsa, K. P., Heeroma, J. H., Somogyi, P., Rusakov, D. A., & Kullmna, D. M. (2007). Anti-Hebbian Long-Term Potentiation in the Hippocampal Feedback Inhibitory Circuit. *Science*, 315, 1262-1266.
- Lebiere, C. (1999). The dynamics of cognition: An ACT-R model of cognitive arithmetic. *Kognitionswissenschaft*, 8 (1), 5-19.
- Lukes, G., Thompson, G., & Werbos, P. (1990). Expectation driven learning with an associaitve memory. *In Proceedings of the IEEE International Joint Conference on Neural Networks*, 521-524.
- Pearce J. M. and Hall G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. *Psychological Review*, 87, 532-552.
- Rescorla, R. A. & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A.H. Black & W.F. Prokasy (Eds.), *Classical Conditioning II*, Appleton-Century-Crofts.
- Sanger, T. D. (1989). Optimal unsupervised learning in a single-layer linear feedforward neural network, *Neural Networks*, 2 (6), 459-473.
- Sutton, R. S., & Barto, A. G. (1998). *Reinforcement Learning: An Introduction*. MIT Press, Cambridge, MA.
- Wickelgren, W. A. (1969). Context-sensitive coding, associative memory, and serial order in (speech) behavior. *Psychological Review*, 76 (1), 1-15.
- Wittenberg, G. M., & Wang, S. S. H. (2006). Malleability of spike-timing-dependent plasticity at the CA3-CA1 synapse. *Journal of Neuroscience*, 26, 6610-6617.