Hariharan Abishek

USC ID: 2743190572

abishekh@usc.edu

Theme 3: Cognition, Planning & Schemas

A neural network approach for complex cognition and planning in adversarial environments.

Abstract:

The model proposed in this report aims to build on the strengths of previous goal directed models in the domain of complex cognition and planing behavior while addressing some of the deficiencies present among the models proposed by: Dehaene and Changeux [DC97] [1] [CD2000] [2]along with the one proposed by Polk et al . [PSLF02][3] . A realistic neurally inspired model has been proposed to simulate progressive degeneration of processing units and pathological effects typically observed in subjects affected by Alzheimer's‘ Disease. The model has been adapted to the adversarial game of English Draughts similar to the work laid down by Arthur Samuel[5] to simultaneously normal and impaired behavior.

Introduction:

The game of English Draughts is (1) played by two opponents, alternating moves on opposite sides of the game board. The pieces are traditionally black, red, or white. Enemy pieces are captured by jumping over them. (2)Each player starts with twelve pieces on the dark squares of the three rows closest to that player's side (see diagram 1). The row closest to each player is called the crown head or kings row. The player with the darker coloured pieces moves first.(3) A player can make two types of moves in English Draughts: Simple moves consists of sliding a piece one square diagonally to an adjacent unoccupied dark square. Uncrowned pieces may move only diagonally forward; kings may move in any diagonal direction.And jump moves from a square diagonally adjacent to an opponent's piece to an empty square immediately beyond it, in the same line;Jumping is always mandatory: if a player has the option to jump, he must take it, even if doing so results in disadvantage for the jumping player.(4)If a player's piece moves into the kings row on the opposing player's side of the board, that piece is said to be crowned , becoming a king and gaining the ability to move both forward and backward. (5)A player wins by capturing all of the opponent's pieces or by leaving the opponent with no legal move.

Apoptosis is the process of programmed cell death (PCD) that may occur in multicellular organisms that is brought about through biochemical changes.The rationale for adapting the model to the game of english draughts is to provide a two player adversarial environment for planning and cognition in order to simulate the effects of apoptosis for a control subject (Player A) and a test subject (Player B) and a over processing cycles and gather the relevant empirical findings.This is in contrast with the approaches presented in [DC97][CD2000] as well as [PSLF02]. These models focused on the Tower of London task to elicit responses from their respective frameworks that were in line with complex problem solving in non adversarial environments.Since the the work presented in [PSLF02]provides argument for the advantages of neurally based model architecture that arises from their ease of mapping onto biological brains the current model falls in line with this approach.Further the [PSLF02] introduces a problem solving architecture that is not only task independent but also method independent.Conversely, the [PSLF02] model suffers from the lack of a proper look ahead mechanism or explicit state evaluation.This is overcome in the proposed model by implementing a game tree approach of problem formulation.

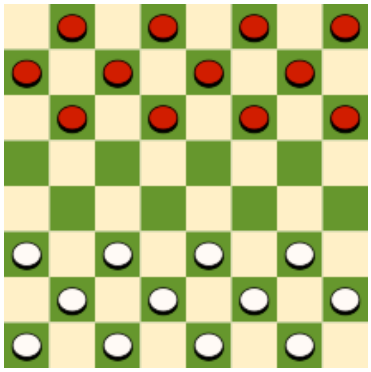

Figure 1:Illustrates the setup of the game of English draughts

The game tree is essentially a search problem over the state space formed by plausible moves from the current state.(Diagram 2) Similar to the functioning of the plan level in [DC97] finding a solution via the game tree is non trivial and involves the selection of a sequence of legal moves that must be simulated and evaluated and then committed to the board or withdrawn depending on the heuristic and distance to the goal.

The key features that the [DC97] model put forth were, (1)the hierarchal nature of processing levels in the form of plan,operation, and gesture levels.(2)A simple greedy algorithm that chose a move that increases the propensity of reaching the goal as well as a mechanism to explore random moves.The model not only ensured that the system stayed focused on task, but also provided randomization. In the game of checkers , this can be used to arbitrarily break ties that occur when distinct search paths lead to the same evaluation. This random exploration prevented the[DC97] system from getting stuck at a sub optimal pre-goal state.Another mechanism that prevented the system from diverging from the goal path employed the use of memory to remember the last desirable state.This remembered state could be returned to if the system ended up in an undesirable state.This is analogous to the notion of backtracking.In the current model, the game tree is constructed by considering plausible moves from the immediate state space , while enumerating the various states from the initial state, if a move generates a sub tree that is lower valued than a previously generated subtree choice , it may be abandoned and the hypothesis can return to a previous state in the search tree and proceed to explore a different move if better move options are present.This approach is consistent with the human notion of 'train of thought'.Exploring one train of thought during strategy formulation and abandoning it if it leads to undesirable consequences in favor of another line of exploration.(Diagram 2)

Furthermore the model incorporated task-independent but method-specific architecture for looking ahead. This notion of forecasting is an important feature that has been conserved for the purpose of constructing a better model , in the present model the game tree aids in forecasting future moves in terms of opponent response and available moves.

For the purpose of constructing the present model , the [CD2000] framework is used as an anchor while adding relevant modifications to improve and adapt it to the current domain.In the following section , the framework of the current model is postulated and the necessary modifications and adaptations from previous models is specified in order to tailor the model for the game of checkers.

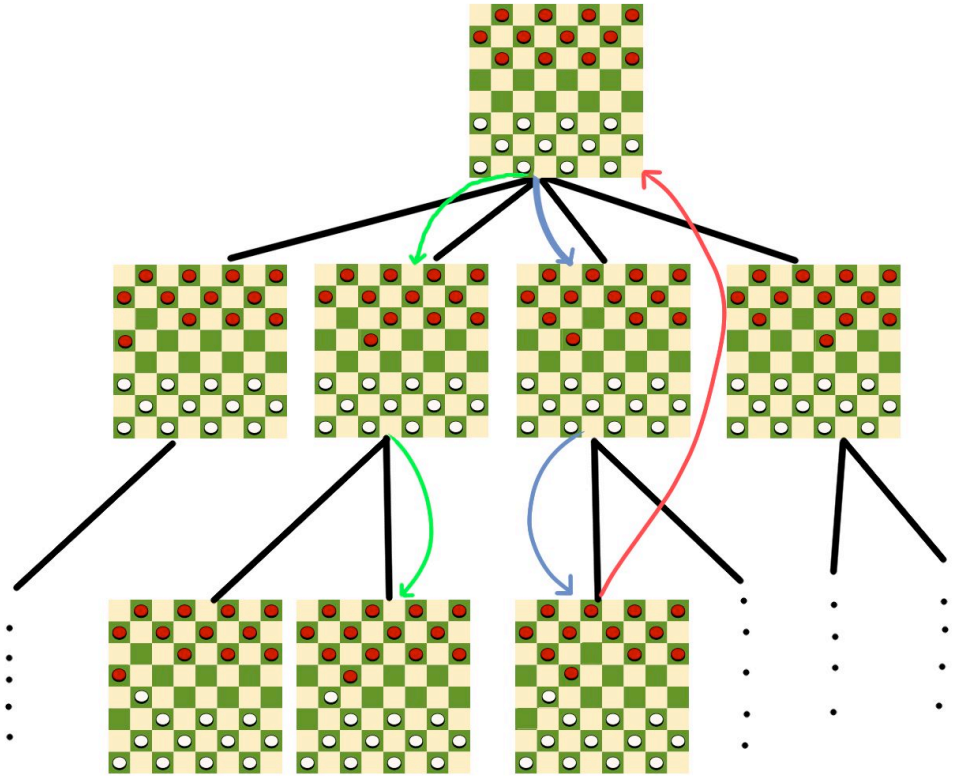

Figure 2: The construction of the game tree is shown with the initial state at the root. The initial line of exploration for the 1st player is along the blue connectives where the player receives reward for making the first move but needs to retreat to the previous state to avoid the consequence of losing a red piece to the opponent , but once the player realizes that it leads to an undesirable outcome , it is abandoned - this is shown by the red connective .The player then proceeds to explore the path along the green connective.

Training:

The individual processing units for player A and player B are initialized with the same values and subject to the same training set consisting of 1000 chess games and 100 tests sets .The training given is similar to Samuel's model [6]. Wherein with progressive games the system begins to adjust the 16 parameters using hyperplane approximations similar to the one shown in (Figure 3),Samuel's model also addresses the issue of temporal credit assignment by adjusting the evaluation function over successive move cycles by comparing the current evaluation with past evaluations.

$$z = w_1x_1 + \ldots + w_{16}x_{16} - \theta$$

Later during simulations Player A will act as the control subject while specific impairments are introduced to the test-subject B.It must be noted that two similar models are created for the purpose of separation.The only difference between the model used for player A is that it does not contain degeneration units.This ensures that A remains healthy throughout processing cycles.

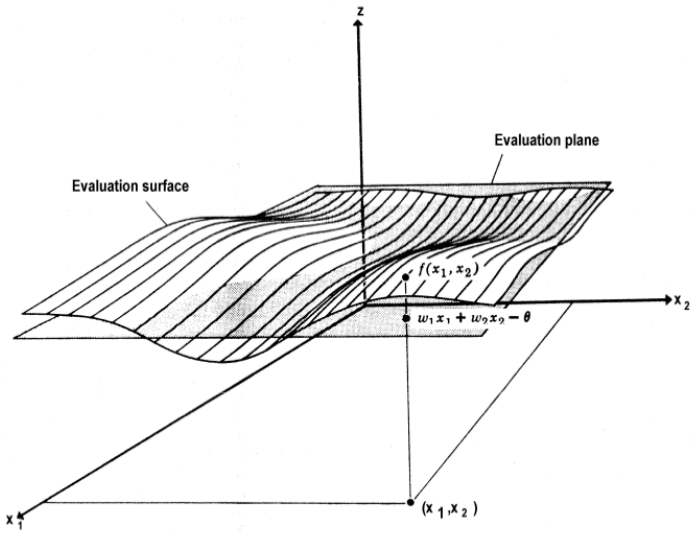

Figure 3 : Evaluation of weights using hyperplane approximations to roughly fit the trend exhibited by the function.

The Model:

The model discussed here abandons classical physiology in favor of a projective style, according to which the system constantly anticipates and tests the outside world as its own private world by producing a diversity of 'pre-representations' in the form of the game tree.The elementary building blocks for this model will be similar to the single cell networks adapted from [CD2000]. This is in order to preserve the realistic nature for simulations of progressive degeneration.An addition to features extracted from [CD2000] the model adds DEGENERATION parameters to each of the processing units and memory units. At every move cycle the DEGENERATION units dilute the efficiency of the processing units by a factor proportional to the synaptic weights and activation frequency of the unit.These units weaken the processing units by reducing the synaptic weights and firing potential to a point where the activation of a processing unit is no longer possible and hence it does not contribute to the computation. They also act upon working memory by randomly flushing stored values in memory,this is analogous to humans being forgetful and not retaining recently provided information.The memory storage mechanisms will have to frequently refresh themselves to keep up with the memory loss.The frequency of refreshing will increase over move cycles as more units are affected. The rationale behind the introduction of these units is to gain a more realistic network response during impairment simulation by introducing the concept of Apoptosis.These units do not come into play during training and are used specifically for simulation purposes.

Similar to [CD2000] this model also deals at the level of the cell and synapse , but the neural network is adapted to checker playing through an expedited temporal difference learning process[5][6].The process is hastened through the addition of the distance to goal criteria *'d'*[DC97]. This attribute of the game state at a particular level in the search indicates the number of relevant moves that are between the current state and the intended goal.The motive is to make moves that increase the value of the heuristic for a particular player while reducing the heuristic for the opponent.The player must also make moves that reduce the distance to a particular goal .

The plan (Prefrontal Cortex) units interact with the WORKING MEMORY which is included to incorporate the hippocampal function of short term memory(the issue of cross-temporal contingency[BOPL01]

(Botvinick,Plaut 2001), where choosing the correct action depends on retaining information about earlier actions is applicable only to the evaluation values as the system needs to compare its past evaluation with the current one to decide if it is on track to the goal). For this model the working memory need only represent the current and intended board configurations along with the storage of plan variables and the past evaluation value.

This is done by representing each board position with 'null' indicating that the system has no knowledge yet of that particular board position ; 'color' to indicate the color of the square and 'piece' to indicate the presence of a particular piece at that position. If there is no piece at a particular position the system represents it with a 'blank' value.A 'blank' value differs from 'null' and should not be confused for each other.

The idea of using 'null' values for a particular position is to encourage the system to collect information of the entire game board before contemplating moves.Acquiring game board information is done by scanning the board either row wise or column wise and updating the corresponding memory position this is referred to as a memory refresh cycle and models realistic human interactions.Humans without photographic memory , need to refresh their knowledge about the environment by diverting their attention to objects of interest.In order to form an action plan , the working memory needs to be filled for all board positions ,so as to consider all the available pieces for play.

The working memory is made transient through the DEGENERATION UNITs that are associated with flushing random memory cells.The frequency of decay and number of attributes decayed depend on the progression of game cycles(the more number of game cycles , the more the cells get utilized and hence the faster they age) and random selection of memory cells which enables flushing or blanking of arbitrary board positions.

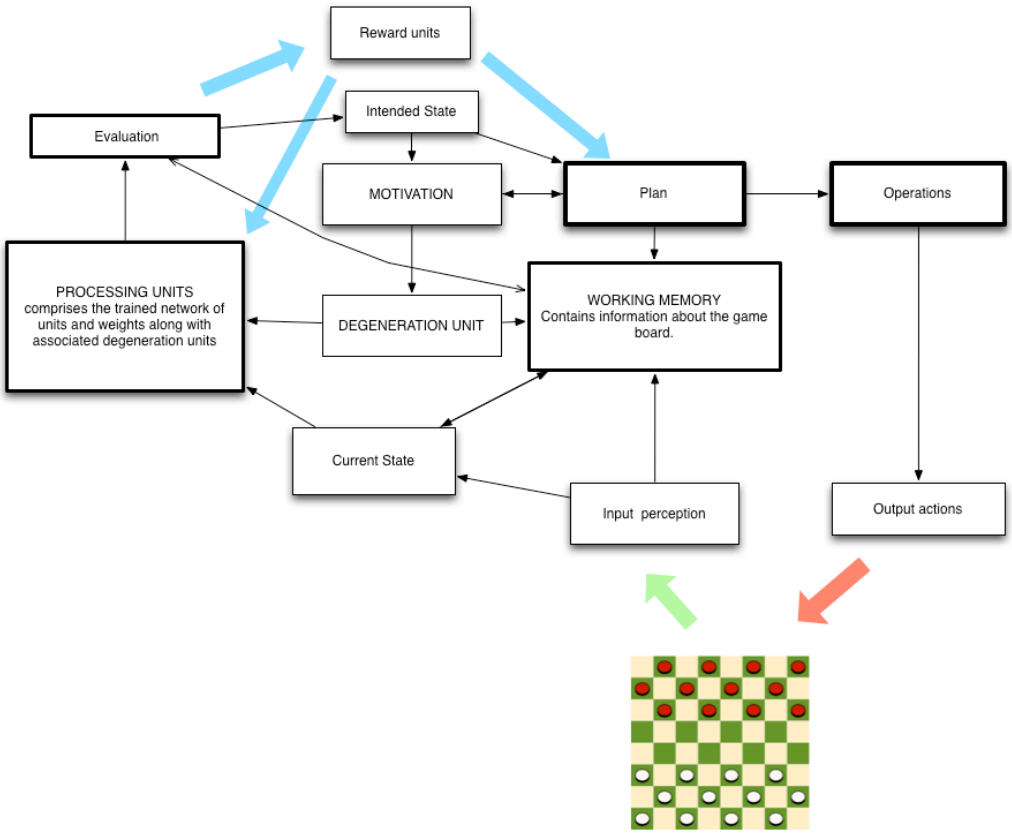

Figure 4:The above illustration captures the various components of the model.The DEGENERATION UNIT is only present for Player B

Once a memory location is flushed it is set to ‘null’ and this motivates the perception modules to update it.But once random noise alters the color and piece values , there is no way for the system to realize their alteration until the next refresh cycle.This is analogous to forgetting or not accounting for certain cues on the game board.Humans do not realize that they have forgotten something until confronted with a requirement for the specific knowledge.Thus the model proposed not only provides more realistic memory organization but also goes on to improve and provide a much more realistic cognitive organization model based on [CD2000]. PLAN units that cause activation patterns among OPERATION units. REWARD UNITS analogous to the Substantia Nigra activate in response to the evaluation function , indicating a “good ” move. The STORE plan unit is activated to save the details of the new plan.

[BOPL01] criticizes [CS2000] for not explaining the mechanism of goal monitoring and inhibition. For this purpose, MOTIVATION UNITS are introduced in the current model which activate while a sub problem remains unsolved (Example: during a multiple jump move the player shouldn't cease operation after the first jump.) , this maintains the generation of new plans until the problem is solved(the new plan here would be to jump again). This also addresses the point brought up in [BOPL01] of reflex inhibition which relates to Estes’ refractory period. We avoid this problem in the current model by the keeping the MOTIVATION UNITS active throughout a move cycle.

Prediction of Empirical data obtained from the model:

Normal Behavior:

Under normal apoptosis which matches the level seen in healthy individuals . Player A and Player B play under no degeneration.

1.An overall task tends to be composed of distinguishable subtasks, which are themselves made up of distinguishable actions ; the game of checkers involves perceiving the board , committing it to memory , passing it to the processing units for evaluation in order to planning the next move .Once a plan is formed it is passed to the operation level which executes the necessary action.

2.Individual actions may figure in more than one subtask, and a given subtask (example: simply moving a piece out of the way or moving a piece with intent to trap an opponents piece.) may feature in a number of different tasks.

3. In some cases, it may be permissible to execute the elements of a sequence in variable order if they lead to a similar outcome. For example, there may be situations where only a single piece of he opponent remains on the board , the player can then chose between several different move sequences to advance on the opponents piece , resulting in the same final result of its capture as long as the distance to goal criteria has not been violated.

4.In some cases, actions or subroutines may be substituted for one another. For example, the player can chose arbitrarily between moves that have the same evaluation.

5.The details of certain action sequences may depend on the state of the board in which they are performed. For example moving a piece to the opponents king row requires that a spot in the king row be available to move to.

6.Both players are observed to be equally matched. Each winning about the same number of games as the other. Figure 5a.

Slips of Action:

Patients with Mild cognitive impairment (MCI) have memory problems and other symptoms of cognitive decline that are worse than might be expected for their age.The degeneration unit for Player B is activated to simulate mild cognitive impairment.

1. As noted earlier, slips tend to occur under conditions of distraction or preoccupation (Reason, 1990). The degradation of the systems processing units , will cause poor evaluation of the board in successive cycles .Player B will end up making suboptimal moves and may not stay on path towards the goal or may take longer more convoluted routes to achieve victory even though simple ones existed.(The system will tend to be at the mercy of the remaining healthy processing units that only enable some actions but not consider others that may hold more pertinence in decreasing the distance to the goal) . This deviation of the system qualifies as distracted behavior. Moreover , the fewer surviving plans units , means that fewer effective plans will be generated.
2. Slips tend to occur at "branch points" or "decision points," junctures where the immediately preceding actions and/or the environmental context bear associations with different subsequent actions.As noted in construction of this model , multiple branch points exist at every stage of the game tree.This is trivially obvious, as a player can make multiple move choices from a particular game state.
3.As observed from the Figure 5b , until about 30 games the mild cognitive impairment cause Player B to exhibit below average scoring.Player A on the other hand takes advantage of B's impairment and begins to win more number of games.
4.Furthermore under MCI , mild loss of memory function is also noted (Figure 5e);while Player A requires less memory refreshes per move cycle (since there are less pieces to keep track of) while Player B experiences mild losses of memory , and requires on average more memory refresh cycles per move cycle.

Moderate to severe loss of cognitive and memory function

Moderate to severe cognitive impairment is characterized by forgetfulness and wandering behavior,impaired ability to perform challenging mental arithmetic,greater difficulty performing complex tasks such as planning and loss of ability to respond to the environment.

1.In the current system severe loss of cognitive function is characterized by repeated inability to produce correct moves,leading to very poor performance as illustrated by (Figure 5c).
2.Since the processing units yield bad evaluations , Player B ends up making bad move choices,which only increase as the number of game cycles progress due to further bad evaluations
3.The move choices seem sporadic and exhibit a kind of guess behavior in the game tree, this can be associated with wandering. The player may also exhibit an abandonment of intended plans , due to insufficient motivation or severe memory losses.
4.(Figure 5d) , illustrates an interesting case wherein the simulation has to be aborted due to sudden non participation of Player B.The player is incapable of performing due to catastrophic degradation of the processing structure and working memory.Such abandonment is comparable to the inability to respond to ones environment.
5.Severe memory loss is is indicated by (Figure 5f) wherein Player B is consuming an average of about 42 refresh cycles per move cycle.

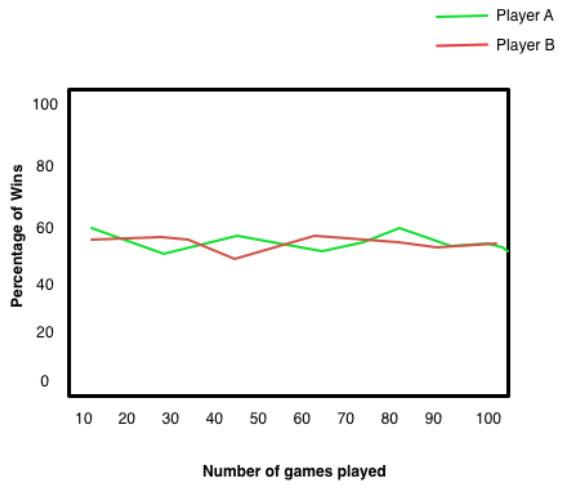

Figure 5a:Normal response

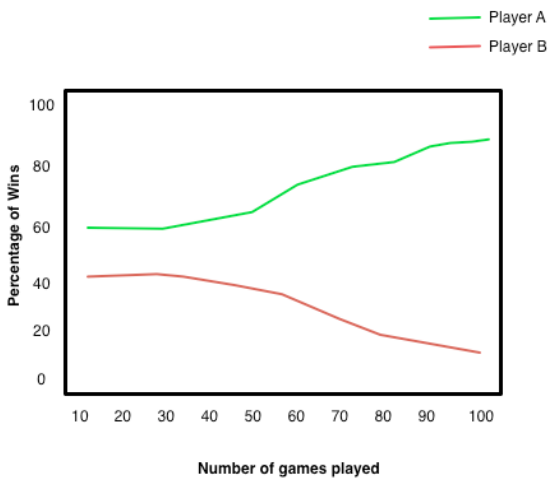

Figure 5b: MCI

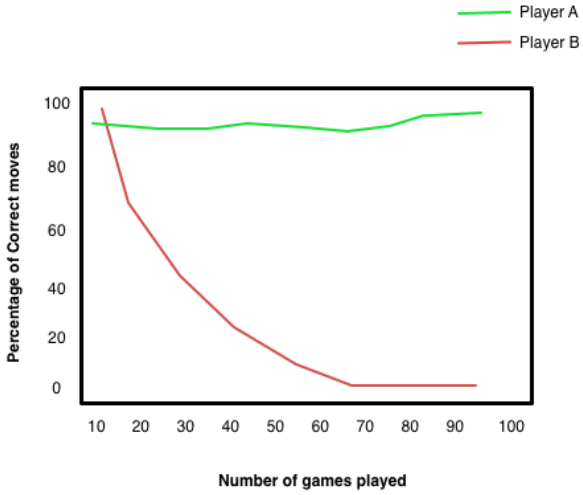

Figure 5c: Player B is progressively incapable of producing correct moves

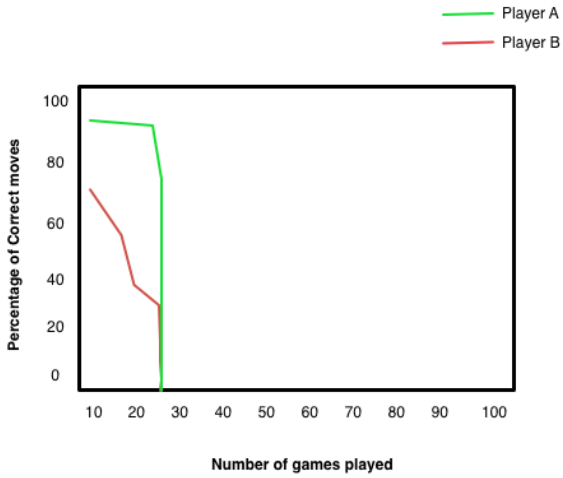

Figure 5d: Simulation prematurely aborted

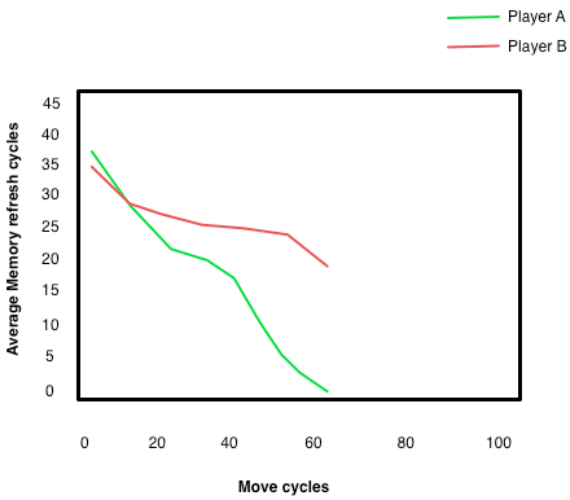

Figure 5e :Average Memory refresh cycles for MCI

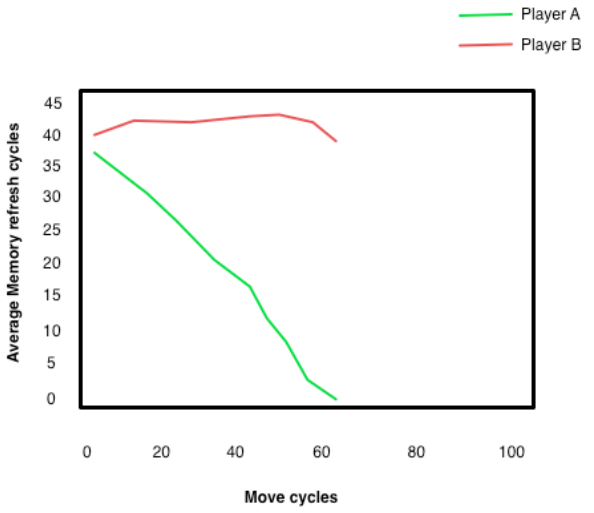

Figure 5f:Average memory cycles for moderate to severe loss

Treatment of Alzheimer's Disease:

There is still no cure for Alzheimer's. People with the disease progressively lose memory and the ability to function.As the disease advances, symptoms can include confusion, irritability, aggression, mood swings, trouble with language, and long-term memory loss. Reduction in the activity of the cholinergic neurons is a well-known feature of Alzheimer's disease.But treatment options exist to cope with the progressive degeneration and allow patients to manage their symptoms to an extent.These medications won't stop the disease, but they can slow down the progression of symptoms for a few months or even years.

Memantine is used to treat moderate to severe confusion (dementia) related to Alzeheimer's disease. It does not cure Alzheimer's disease, but it may improve memory, awareness, and the ability to perform daily functions. This medication works by blocking the action of a certain natural substance in the brain (glutamate) that is believed to be linked to symptoms of Alzeheimer's disease.[10] Glutamate (glutamic acid) is especially prominent in the human brain where it is the body's most prominent neurotransmitter, the brain's main excitatory neurotransmitter, and also the precursor for GABA, the brain's main inhibitory neurotransmitter.

Predicted Summary of Simulation Results:

ACh:

The pathogenesis of Alzheimer's disease (AD) has been linked to a deficiency in the brain neurotransmitter acetylcholine.Acetylcholinesterase inhibitors are employed to reduce the rate at which acetylcholine (ACh which is important for learning, memory, and attention. [9]) is broken down, thereby increasing the concentration of ACh in the brain and combating the loss of ACh caused by the death of cholinergic neurons.

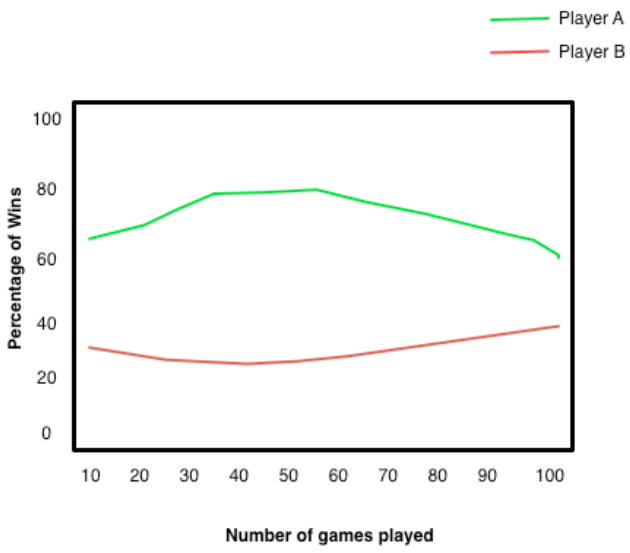

Figure 6:Effects of slowing down the degeneration.

To simulate this effect of treatment we lower the frequency of the DEGENERATION UNIT updating for each processing unit and we do not apply degeneration to WORKING MEMORY. We set the DEGENERATION UNIT to update its projections every 10 cycles , instead of every cycle.This ensures that the processing units degenerate slower.

As observed with mild symptoms of AD , patients are able to overcome some of the effects of the disease while exhibiting small improvements in cognitive and global assessments.Similar conditions are observed (Figure 6)

where Player B started with 30% wins , but prolonging the life of processing units enabled B to reorganize their synaptic weights over time and achieve a slight improvement of 5% .

Glutamates:

NMDA (N-methyl-D-aspartate) receptor antagonist, works by regulating the activity of glutamate, a chemical messenger involved in learning and memory.Attachment of glutamate to cell surface "docking sites" called NMDA receptors permits calcium to flow freely into the cell. Over time, this leads to chronic overexposure to calcium, which can speed up cell damage.Preventing this destructive chain of events is possible by partially blocking the NMDA receptors.

As put by Walton and Dodd [7], “Glutamate plays major roles in normal brain function and in brain development. Perturbations of glutamate neurotransmission have severe consequences. Prolonged elevation of the extracellular glutamate level tonically activates glutamate receptors, causing sustained local depolarization of neurons, which in turn triggers a sequence of intracellular events that culminate in Na+ and Ca2+ influx and further exocytosis of glutamate. Ca2+ influx leads to delayed necrosis of the neuron, and to a lesser extent activation of apoptotic pathways. Continuing release of glutamate leads to a spreading of the process (Choi, 1992), which is termed excito-toxicity. The capacity of glutamate to be highly toxic, yet necessary for neurotransmission, sets a fine balance between plasticity and pathology. Glutamate excito-toxicity has been implicated in a number of brain disorders, including epilepsy, amyotropic lateral sclerosis, Huntington’s disease, Alzheimer’s disease.”

We can simulate the effects of excito-toxicity in the current model , by increasing the value of the reward unit to very high levels , as well as triggering very high motivation levels which in turn trigger degeneration to occur. By scaling up the reward units to very high values , we saturate the processing circuitry. Which causes it to excite with high activity after receiving input stimuli.Initial mild improvement in performance is observed , followed by rapid apoptosis of large number of processing units which have burnt through their firing thresholds.

Relation to other work :

While to the best of the author’s knowledge , no computational model has been found to exist to replicate , observe and study the effects of apoptosis on planning and cognition. The behavior of the model closely relates to previous biological and psychological studies like that of (Bennett et al., 2005) [9].

Persons with mild cognitive impairment had intermediate levels of AD pathology from those without cognitive impairment and those with dementia .(Figure 7)The study goes on to indicate that that loss of memory is not an inevitable consequence of aging but rather is usually the consequence of age-related diseases.This is similar to the effects displayed by the current model when modeling normal ,MCI and moderate to severe AD.

Additional studies have provided data on the trials of various treatments to address the effects of AD.The model presented here is in line with the general notions presented in those studies.

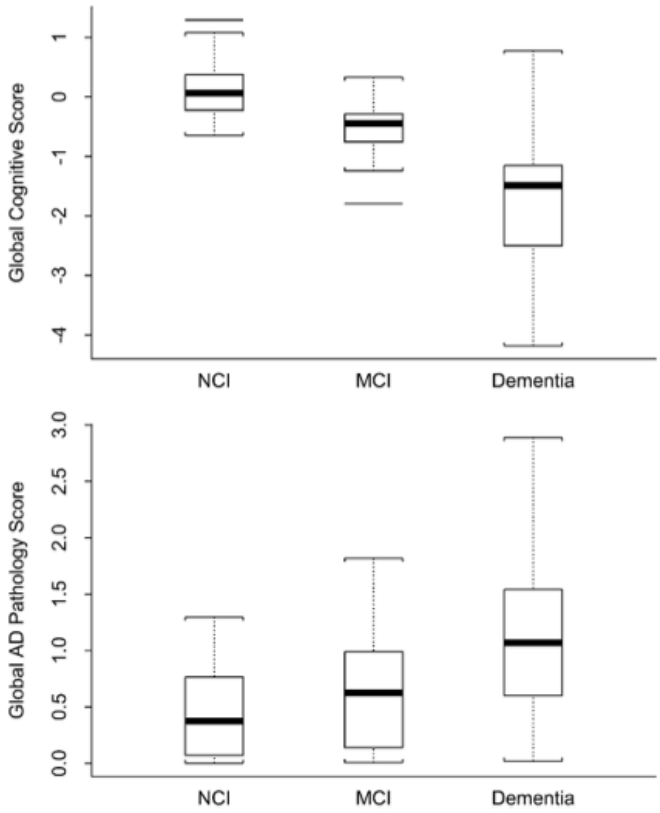

Figure 7: The global cognitive score and global Alzheimer Disease (AD) pathology score by diagnostic group in persons with no cognitive impairment (NCI), mild cognitive impairment (MCI), and dementia.[9]

Conclusion:

The current model that has been proposed will not only provide a good platform for the understanding the early learning mechanisms of neural networks but also provide a basis for replicating and studying the effects of apoptosis on cognitive systems.The work presented here also reports on the pertinent empirical findings of simulated neurological degeneration when playing the game of checkers.We observe that the effects of mild degeneration can be managed to an extent while preserving participation in the task. But untreated MCI can rapidly progress to unmanageable levels, impeding cognitive and memory functions.Moderate to severe AD cannot be managed effectively for a prolonged period.The degeneration has become so prominent that recovery is very unlikely and often terminal in patients.Thus an early detection and diagnosis is key in prolonging cognitive function and possibly postponing the onset of dementia.

References:

1. S. Dehaene, J.P. Changeux, A hierarchical neuronal network for planning behavior, *Proc. Natl. Acad. Sci*. USA 94 (1997) 13293– 13298.

2. Changeux JP, Dehaene S. Hierarchical neuronal modeling of cognitive functions: from synaptic transmission to the Tower of London. *Int J Psychophysiol. 2000 Mar*;35(2-3):179-87.

3. Thad A. Polk, Patrick Simen, Richard L. Lewis, Eric Freedman, A computational approach to control in complex cognition, *Cognitive Brain Research, Volume 15, Issue 1, December 2002, Pages 71-83,* ISSN 0926-6410, 10.1016/S0926-6410(02)00217-3.(http://www.sciencedirect.com/science/article/pii/S0926641002002173)

4. Doing Without Schema Hierarchies: A Recurrent Connectionist Approach to Normal and Impaired Routine Sequential Action.Botvinick, Matthew; Plaut, David C.Psychological Review, Vol 111(2), Apr 2004, 395-429. doi: 10.1037/0033-295X.111.2.395

5. *Cooper R, Shallice T*: Contention scheduling and the control of routine activities. *Cogn Neuropsychol. 2000 Jun 1;17(4):297-338*. DOI: 10.1080/026432900380427.(http://www.ncbi.nlm.nih.gov/pubmed/20945185)

6. Richard Sutton (May 30, 1990). "Samuel's Checkers Player". Reinforcement Learning: An Introduction. MIT Press. Retrieved April 29, 2011.

7. Samuel, A. L. Some studies in machine learning using the game of checkers. IBM *Journal of Research and Development 3(3)*

8. Glutamate–glutamine cycling in Alzheimer's disease Heather Scott Walton *, Peter R. Dodd*;School of Molecular and Microbial Science, University of Queensland, Australia 2006*

9. Neurodegeneration in Excitotoxicity, Global Cerebral Ischemia, and Target Deprivation: A Perspective on the Contributions of Apoptosis and Necrosis ,Lee J. MartinA, B, Nael A. Al-AbdullaA, Ansgar M. BrambrinkD, Jeffrey R. KirschC, Frederick E. SieberC, Carlos Portera-CailliauB

10. Mild cognitive impairment is related to Alzheimer disease pathology and cerebral infarctions D. A. Bennett, J. A. Schneider, J. L. Bienias, et al. Neurology 2005;64;834 DOI 10.1212/01.WNL.0000152982.47274.9E