A Historical Population in a Coevolutionary System

Phillipa M. Avery	Zbigniew Michalewicz	Martin Schmidt
Dept. of Computer Science	Dept. of Computer Science	SolveIT Software
University of Adelaide	University of Adelaide	Adelaide, Australia
pippa@cs.adelaide.edu.au	zbyszek@cs.adelaide.edu.au	martin.schmidt@solveitsoftware.com

Abstract—The use of memory in coevolutionary systems is considered an important mechanism to counter the Red Queen effect. Our research involves incorporating a memory population that the coevolving populations compete against to obtain a fitness that is influenced by past generations. This long term fitness then allows the population to have continuous learning that awards individuals that do well against the current populations, as well as previous winning individuals. By allowing continued learning, the individuals in the populations increase their overall ability to play the game of TEMPO, not just to play a single round with the current opposition.

Keywords: Computational Intelligence, Adaptive memory, TEMPO

I. Introduction

This paper discusses applications of memory to address the 'forgetfulness' of the coevolutionary processes [1]. The issue with the system forgetting previous good solutions is well documented in coevolutionary research, and in the case of this research it lead to the creation of strategies that were not highly competitive. By introducing memory we have attempted to increase the aptitude of the TEMPO players to play against all strategies, not just the current opposition. We also discuss other improvements made to the system, particularly the change from a Gaussian membership function for the fuzzy rules, to a triangular one.

Using coevolution, the evolutionary process can 'lose' the best found solution, as the goal of each individual is to beat the opposition for the current generation only. As discussed by Cliff and Miller [2], a later generation individual should be able to beat an early generation opponent but this is not always the case. Ancestors of the opposition could have strategies that the later generations have forgotten about and have no way of beating. One solution to this is to incorporate some form of memory into the coevolutionary process, one that helps the populations to remember previous solutions[3][1].

The use of memory in coevolution to retain previous winning (best of generation) solutions has been researched using different memory mechanisms. Some involve retaining the best of generations for insertion or replacement into populations in latter generations [4], while others use the historical memory to compete against and influence the fitness [3]. Whichever mechanism used, there remain the same questions regarding the historical population's representation and selection. These questions can be summarized as [4]: how should individuals be selected for insertion into the history, what size should the history be, and how should

an individual from the history be selected for use in the evolution.

Most systems choose to use best of generation individuals, however this is not the only option as discussed by Bader-natal and Pollack [5]. However, the purpose of using history for this research is to remind current generations about the successful strategies of past generations, so only the best will be stored. Having decided *which* individuals to be inserted, the next step is to decide *when* the individuals should be selected and stored. The selection of individuals directly affects the question relating to the size of the population, and the selection of an individual from the historical population for use in the evolution. For example, storing the best individual from both populations at every generation can cause the historical population to grow rapidly, and the selection of individuals will be influenced by the larger scale.

We examined a variety of methods for the representing the historical population. One particular choice was whether to periodically remove older (and perhaps less efficient individuals) or constantly grow the population until program termination. We chose the latter for the experiments, due to possible relevance of older individuals. A number of experiments on the effectiveness of this method is discussed in section IV. We also performed experimentation to demonstrate that forgetting was happening, and to experiment with the most beneficial memory mechanism.

For the case of this research, we use the term "memory" and "history" interchangeably, as our metaphor for the memory was that it was a recording of history for the individuals to learn from.

The paper is organized as follows. Section II gives a background of the main topics of this paper, including a description of the TEMPO game and the previous research performed by Johnson et al., and a discussion on the way the human memory system works. Section III discusses the different methods we have used to investigate and test the addition of memory to the system. Section IV outlines the results of the experiments performed, and a discussion of observations from the results. The final section concludes with a discussion on the overall findings.

II. Background

For a background to the research described in this paper, we discuss the game used; the TEMPO military planning game. The discussion includes the game play involved in TEMPO, and the previous computationally intelligent system developed for the game.

Following this, a background is given into the psychology of human memory. This gives a very brief introduction into how the human memory system works, and in particular focuses on the way humans use short and long term memory.

A. The TEMPO military planning game

TEMPO is a zero sum game played between two opposing parties by allocating resources in a cold war style simulation. The goal of the game is to acquire more utilities than the opposition before war breaks out. The decision making process requires allocating a yearly budget resource between offensive and defensive weapons of various strength and cost. The purchase of intelligence is also provided to give insight into opponent's tactics. Lastly, investment in research and development is available to provide for future weaponry.

The resource allocation involved in the game is in concept relatively simple; determine who needs what and then allocate accordingly. The reality is however very different, as the number of different combinations of allocation plans can be high due to the amount of areas to allocate to. This complexity is then magnified by the changing environment that occurs yearly, such as the increase in the chance of war breaking out, and the addition of new weaponry. This situation is representative of a number of real world situations in the corporate and defense world alike, where resource allocation can be a very complex and difficult task to manage. As a result the TEMPO game provides a training facility for staff to practice and refine their skills, as well as test bed for the development of computer systems to tackle the allocation issue.

The TEMPO game was used in a coevolutionary system by Johnson et al. [6][7] to test the use of evolving a self-learning artificial player for the game. The system followed similar work done by Chellapilla and Fogel [8][9] where a computer player developed its own method of play through coevolution. Instead of using the neural network approach done by Chellapilla and Fogel however, the player was developed by coevolving a set of fuzzy rules.

The system developed in [7] used a Mamdani fuzzy logic system with Gaussian membership functions. The individuals consisted of a rule base that were randomly initialized and then coevolved to create a strategy of play. The coevolution occurred by using two populations that competed against each other, with a fitness calculated from the average net offensive utilities and average wins over all the games played. A linear penalty deduction was then applied to minimize the rule set.

The chromosomes of the coevolutionary process are rule bases with each individual comprising a number of concatenated rules up to a maximum of m rules where m is made up of w weapon rules and q intelligence rules. Each of these rules are built from data relating to: whether the rule is to be used, whether each of the input values for the rule are to be used, the corresponding fuzzy membership function for each input, and the output membership function of the rule. These are then used to decide if a particular intelligence category or weapon should be bought when the yearly budget allocation is performed, using the production operation rule for fuzzy-AND. After each generation the populations are evolved with a chanced crossover and mutation operation (for more details see [7]).

The research presented in this paper used the system from [7] as a baseline for the research, and conducted further experiments to refine and improve the former computer player. There were a number of issues that were observed by the research from [7] relating to the ability of the player to compete against humans, and this research is intended to further this work. One particular observation was that the rules being developed by the coevolution process were difficult for a human to beat initially, but easy to overcome with practice. We felt that a large reason for this situation was due to a lack of memory on behalf of the computer player. Without being able to remember previous games, the player was only focusing on the strategy of the moment, instead of developing generic rules of play.

B. A discussion on the psychology of human memory

In our experimentation it was decided to investigate the use of two types of memory for selection, one as a long term memory (LTM) and one as a short term memory (STM). The approach was to simplistically mimic the way the human brain stores and uses memory. The concept of applying STM and LTM to heuristics has been widely used over the years, and the use with tabu search has been thoroughly documented by Glover and Laguna [10]. In this research however, we seek to mimic the human mechanism of STM and LTM in a coevolutionary process.

Early theories of human memory consisted of one large memory system. However, in the 1960's researchers recognized that the memory consisted of two parts: a short term memory that acted as a temporary storage mechanism, and a long term memory that was a permanent memory storage [11]. This was re-examined in the 70's and the two memory systems broken down into further levels of processing [12]. The STM was re-termed the working memory, and broken down into articulatory and visuospatial components. The LTM was broken into explicit and implicit memory. Implicit memory involves learning that does not involve active recollection of information, but rather retrieval through indirect performance. Implicit memory includes information associated with: skills and habits, priming, conditioning and non-associative learning [12]. Explicit memory on the other hand is the active recollection of past incidents, and the semantic memory representing general knowledge of the world. Explicit memory includes storage and retrieval of facts and events.

The actual linking between short and long term memory was originally thought to happen serially; that the information being taken in by our sensors was processed in the STM and then if it stayed in the STM long enough, it would be transferred into LTM [13] [12]. This was then challenged by Craik and Lockhart [14] who introduced the concept of a *levels-of-processing* framework. This framework had a sequence of analytic stages that show how memory is set.

The idea is that memory is not a separate faculty, but instead reflects the outcome of attempts to perceive and comprehend information. Thus, the ability to affect long term memory is directly related to the comprehension or ability to relate the knowledge with the meaning (semantic understanding). In this way, the retrieval of information is not done in a 'brute force' search mechanism, but instead the memories are encoded into the whole cognitive system, and there is an increased potential for the same pattern to be repeated on a subsequent occasion [13].

The understanding of how the STM and the LTM work, both independently and together, is still a wide area of investigation. However, the knowledge that the memory is divided into two forms: a short term memory aimed at recalling recent and relevant information, and a long term memory to store necessary information in the long term, are fundamental theories for human memory.

III. METHODOLOGY

We have addressed several of the issues mentioned by Johnson et al. [7] regarding system performance, and also investigated memory mechanisms used in addition to the original system. The following section is broken into the improvements we made to the system to increase the processing speed, the method we used to seed the populations with human strategies, and what methodologies were used in the experiments we performed.

A. Processing improvements

To address the speed of the system, the process of coevolving with fuzzy logic rules was evaluated. It was decided that while the general concept of the system was well founded, changes to the fuzzy logic system and the representation of the genes could possibly improve the performance of the system.

The original TEMPO fuzzy logic system was based on the Mamdani fuzzy logic system with Gaussian membership functions (see [7] for further details). The use of a Gaussian membership function with floating point precision resulted in the evolutionary process fluctuating over small variations that did not improve very much. To counter this, we decided to change the Gaussian membership function to a triangular membership function with integer precision. It was thought this would stop the minor fluctuations observed with the floating point precision, and the system would spend the time looking at a larger area of the search space. The change was implemented at the start of the experimentation, and the results showed that the processing time was cut down to a third of the previous speed, and the results against the static expert were also marginally improved.

The new representation of the fuzzy logic system meant that the chromosome structure used in [7] needed to be updated. The new structure can be seen in figure 1. As per the old system, there are $m = w + q$ rules (where w is the maximum number of weapon rules, and q is the maximum number of intelligence rules). Each rule is built from the following (Figure 1 expands rule 3): U_3 is a Boolean
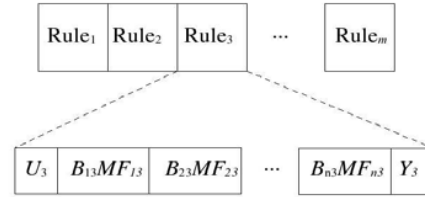

Fig. 1. New structure of a chromosome

defining if the rule is used, B_{i3} are Booleans defining if input i is used, MF_{i3} is the membership function used for the input i, and Y_3 is the output in range [0,1] for Rule$_3$. The membership functions were created for each input value and were a variable number of functions for each input. For further information on how the rules are used, refer to [7].

B. Seeding the populations

One suggestion made by [7] was to seed the population with hand crafted individuals to determine how more human-like individuals would fare in the co-evolutionary process. We decided to investigate this area, and provided a simple experiment to determine the reaction of the co-evolutionary process. The existing base-line measurement of the system consisted of a static expert that used static rules to represent the strategy of; buy weapons based on their utilities per operation cost. The higher the ratio, the more of the weapons would be bought. The term "expert" is used loosely here, as the strategy is a simple concept. However, the strategy has proved capable of winning against novice human players of the tempo game, and can be considered a good base-line player for measurement of the system performance.

When using the expert for performance measurement, it was noted that the best individuals from the generations performed relatively poorly on average against the expert. This was to be expected to some degree, in that the populations were focused on beating each other, and they had no incentive to find strategies to beat the expert. We decided to insert the static expert as an individual into each of the population and record how it fared in the evolutionary process. We already had the measurement abilities in place, as the best individuals from each generation would play the static expert at the end of the generation, and the alien expert would be seen in the results from the static game play. To differentiate between the two functions of the same expert, we named the base-line measurement use of the expert as the "static expert", and the seeded expert the "alien expert".

The alien expert mechanism worked by placing the static expert as another individual in each of the populations. The individual would be subject to the same variation operators as all the other individuals. The base-line static expert measurement system remained the same. At the end of each generation, the best individual from each population would play 100 games against the static expert and the amount of times the individual won would create the won ratio against the individual for that population.

C. Use of a history

An effect of coevolution has been described as the Red Queen effect [15][2], where two or more populations are constantly changing their traits to compete with the other players. This can lead to cyclic evolution, where previous strategies are constantly forgotten and rediscovered. After experimenting with the alien expert, we realized that the coevolutionary strategy did just this. The populations were focused only on outdoing each other, and even after expert knowledge was seeded, as soon as the populations had beaten the expert in the other population, they then developed new strategies, which the other population focused on beating, hence ignoring the expert strategy. Occasionally they might rediscover bits of strategies that would beat the expert, but as they were not rewarded for these discoveries, they were promptly forgotten.

We wanted to provide a player that would be competitive against all strategies, not just the strategy of the moment. This meant developing players that could beat the current opposition, as well as previous opposition; an ancestral tree where the progeny learn from the ancestors. Following this line of thought, we decided to incorporate a memory structure that would act as a record of historical strategies. We decided to store the best individuals of each population into a separate population that we called the History Bank. We then used the History Bank in the evolution of the individuals of the population, where at the end of the generation, the fitness of each individual would be created by playing against the opposition and the individuals from the History Bank. We wanted this to act as a trigger so that old winning strategies would not be lost in the time line, but would help in developing the future generations.

As discussed before there are a number of different ways we could implement and use this History Bank. We started by implementing a separate History Bank for each population, and each generation the best of the opposition was stored in the other population's associated History Bank. This was intended to remind Population A of Population B's winning strategies, and vice versa. After further experimentation and analysis however, we came to the conclusion that to create a truly competitive individual it should be able to beat the opposition, as well as any previous strategies of it's own. We decided to experiment with having a single combined (and decidedly larger) History Bank that both populations would add to and play against.

The mechanism used to remind the current players of the history, was to use the History Bank as opposition against the current populations. At the end of each generation, each individual would play a defined number of games against the opposition population, followed by a defined number of games against the individuals in the History Bank. The fitness we use is calculated as the won ratio of the individual (the number of games won divided by the number of games played) plus the total net utilities divided by the times played, with a penalty applied to minimize rules and inputs used for the fuzzy rule base. Because the fitness is calculated against totals of all games played, it is directly influenced by each game it plays. The strategy uses this by playing each individual against random individuals from the opposition r_1 times. By using random selection, each individual plays a different sample from the opposition population and may play more games than other individuals (if chosen to play by the opposition's random game selection). This is slightly countered in the fitness however, as the win ratio is calculated as an average of all games played. The individual is then played an additional r_2 times against individuals from the history, where the selection mechanism changes as explained in section IV and it's subsections. So the total games played to assign the fitness for the individual is $r_1 + r_2$.

We decided to include every generation's best individual from each population into the History Bank, as we wanted to keep a full history of the past winning strategies used. Initially we implemented the History Bank as only containing unique entries, and later experimented with the inclusion of all individuals regardless of uniqueness.

The History Bank is a continually growing population, with an increase of two individuals every population (when uniqueness is not applied). This means that whatever selection mechanism we apply, it will always be effectively a time based probability distribution. The more generations that pass, the larger the population will become and the less probability an individual will have of being selected for game play. We have included different mechanisms to bias this selection probability, but it can be said that all are still a probability of time. The three probability distributions experimented on are: random uniform selection, selection with a linear distribution and selection with a Laplace distribution probability curve.

After experimenting with the History Bank, we decided to investigate the use of a system mimicking the way humans use memory. As discussed before, humans have a current short term memory, and a larger long term memory. We wanted to replicate this ability, as it is the short term memory that humans use to determine the current situation, and the long term memory to bring past knowledge on how to act given the situation. This is precisely what we wanted the players to do. The short term memory function was done by identifying the top ten individuals in the History Bank and playing an additional amount of games against them. The long term memory was then applied by playing another set of games against the entire history. So where previously the fitness was created by playing $r_1 + r_2$ games, now the games against the history is split into r_{2S} for the games played against the short term memory, and r_{2L} for the games against the long term memory. The History Bank individuals selected for the long term memory game play are selected with a linear time based probability, so the more recent individuals have a higher chance of being selected. The short term memory is also selected from with the same linear time based probability, but as the size is a static 10 individuals, as the entire population grows, the probability becomes more uniform.

The last experiment performed was to further mimic the way the long term memory worked in the human brain. Instead of every single bit of information ever learnt being stored in long term memory, the human brain would sift the information and only store relevant information. We attempted to mimic this by using a ranked long term history. When adding a new individual to the History Bank, the individual would play the top individual from the ranking and if it won it would be ranked above, otherwise it would continue down the ranking until it found it's appropriate rank. This was only performed for the top 1000 ranked individuals. If the individual was not in this top 1000, it would not be used in the long term memory. We nicknamed the 1000 top ranked individuals the "gladiators".

This mechanism provided a very loose ranking, and it was likely that the ranking would be incorrect in some cases. To address this issue, we included a mechanism where at each generation a random individual would be chosen from the gladiators and played against the surrounding 20 neighbours, and adjusted in ranking where required.

The fitness calculation process was then modified as follows. After playing r_1 games against the opposition, the individual would be required to play against the short term memory as before (r_{2S}). They would then be required to play an additional number of games against the gladiators (r_{2G}).

IV. EXPERIMENTS AND RESULTS

This section gives a record of the experiments performed to investigate the different memory mechanisms we used. All the graphs of results in this section depict how well the best individual from each generation performed against the static expert player, which was the baseline measurement. The graphs show the won ratio against the expert on the y axis, and the generations on the x axis. The results depicted show the average of ten runs, all with the same environmental and evolutionary configuration. The system was run each time for 50,000 generations with a population size of 100 for both populations.

Each experiment had the following evolutionary parameters. The ratio of individuals to be replaced in each generation was 90%. The variation operators applied were mutation and crossover. Mutation was applied with a 70% probability of occurrence, where if applied each gene had 50% chance of mutation, with a 10% chance of a large mutation, or otherwise a small mutation occurred. The crossover applied was two point crossover with a 30% probability of occurrence. For further details please refer to [7].

Following the discussion of the experiments and their results is a table summarizing all the experiments and the different mechanisms used in each one.

A. Experiment 1

As a baseline measurement, we ran the system without any of the memory or seeding mechanisms with the fitness calculated solely from game play against the opposition ($r_1 = 20$). The results can be seen in Figure 2. The results show that the players were not doing overly well against the

expert, they were achieving an average of 36% win ratio. The results also showed that there were no trends to beating the expert, just occasional jumps in performance followed by decreases (this is not visible in the results, as the results show the average of 10 runs). This was to be expected, as the coevolutionary process was not training the individuals against the expert, only against the other population.

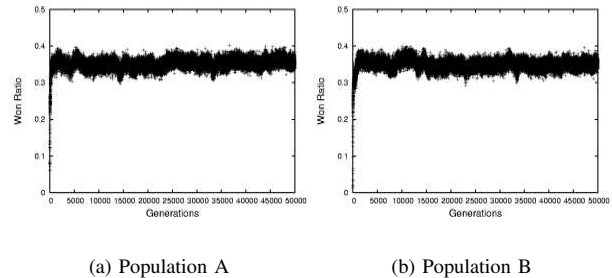(a) Population A   (b) Population B

Fig. 2.   Success ratio against the expert for the Original System

To try and encourage the system to include the expert in the process, we included the 'alien expert' discussed previously. The expert was inserted as one of the individuals for each population A and B. The results from the insertion looked much the same as the original results. After the addition of the alien expert, there was a very brief increase in the won ratio followed by a sharp drop back to the previous average won ratio. Once again, this was mostly expected, as the nature of coevolutionary systems is to focus on beating the opposition in it's current form, and promptly forgets any previous opponents once they have been beaten. The system was only concentrating on beating the opposition, and so the next experiment aimed to include the expert and other past strategies in the whole evolutionary process.

B. Experiment 2

For the next experiment, we decided to implement a mechanism to allow the individuals to remember past winning opponents, by including another population of historical individuals. The logic behind this, was that to begin with the expert is the best individual, so it will be placed in the history. When the history is then used in the evolutionary process, there is a chance that there will be repeated promptings to continue to beat the expert.

The initial memory experiments consisted of randomly selecting individuals from the History Bank, and playing a set number of games against them. The fitness was calculated by playing a number of games against the opposition ($r_1 = 20$) and re-adjusting after each time. To include this strategy of evolving against the History Bank, we then provided a mechanism for playing an additional number of games ($r_2 = 20$) against the History Bank, and adjusted the fitness of the current individuals in the same manner as that for the opposition games.

The graph of the won ratio against the expert can be seen in Figure 3. The use of a History Bank gave us the better

results we were looking for. We were now achieving an average of around 50% won ratio, and were achieving an over 90% top win ratio against the expert for single runs as well as longer trends of winning.
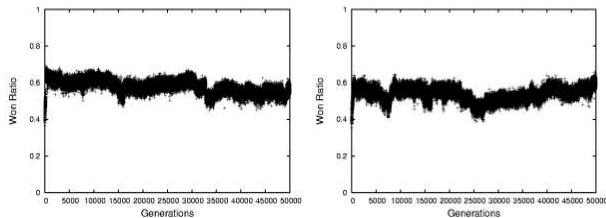
(a) Population A   (b) Population B

Fig. 3. Success ratio against the expert using random selection from History Bank for population A

These results were promising, but we noted that the populations were still not tending overall towards beating the static expert and were instead regularly staying around the average to below average won ratio. The populations were still 'forgetting' previous solutions and focusing more on beating the current opposition. We reasoned that the cause of this was the random nature of the selection from the History Bank. By selecting at random, we were increasing the potential for larger amounts of the earlier, and presumably simpler, individuals from the history to be chosen. What we actually wanted, was for the populations to mainly grow against the harder (more recent) individuals, and then be occasionally prompted not to forget older strategies and so address the Red Queen effect. To create this scenario, we decided to include a probability distribution for game play selection from the History Bank.

C. Experiment 3

The initial probability distribution was a simple linear time distribution, where the latest historical individuals had a higher probability of getting chosen then the older individuals (see Figure 4).
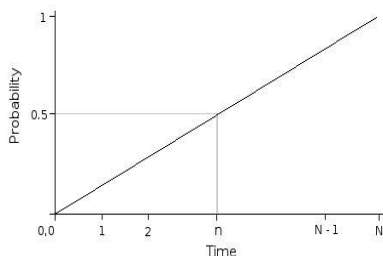

Fig. 4. Probability Distribution by Function of Time P(t) (with the time function represented with N as the maximum generation number)

This distribution would force a higher playing ratio against the more recent individuals, but would still encourage play

against the older strategies. The results of this experiment can be seen in Figure 5, and showed a slight increase in the win ratio average against the expert, but was still not to a satisfactory standard.
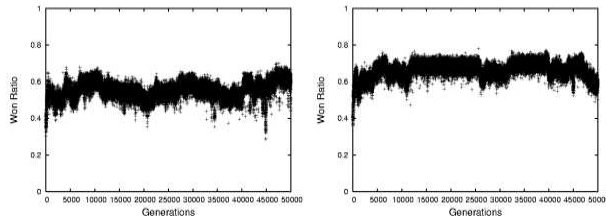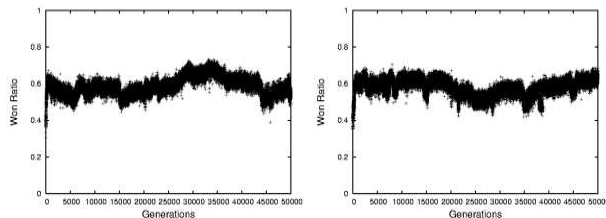

(a) Population A   (b) Population B

Fig. 5. Success ratio against the expert using Single History Bank with Linear Distribution

D. Experiment 4

The next question following on from the previous results, was whether the linear probability distribution was the best distribution, and if perhaps a curved distribution might provide better results. To test this theory, we replaced the linear distribution with a Laplace (otherwise known as Double Exponential) probability distribution function. We were expecting that by introducing a curved distribution function with a sharp regression, we would force the populations to play more against the most recent players in the history. The results of this experiment can be seen in Figure 6.

(a) Population A   (b) Population B

Fig. 6. Success ratio against the expert using Laplace Probability Distribution

The individual results showed a slight improvement with more trends towards winning against the expert. However, there was still a lot of fluctuation and the results were still lacking. What we really wanted was to see an overall increase over time in the statistics against the expert.

E. Experiment 5

The inclusion of a history to evolve against had produced some nice peaks, however even with the probability selection distribution it still seemed to average around a 60% win ratio with a lot of variation. The time function probability selection we were utilizing allowed a biased selection mechanism,

but it was still allowing a higher play rate against the older solutions than we wanted. Even though the time-based probability selection was biased towards the top of the scale, as the History Bank grew, the chances of playing against the top end decreased. To address this we decided to introduce a specified top end window to the probability selection - the short term memory window. This window would be evolved against a set number of times prior to evolving against the whole history - the long term memory.

The experiments included a window of the top ten individuals and played against individuals from the short term memory ($r_{2S} = 10$) followed by games against the long term memory - the entire History Bank ($r_{2L} = 10$). The games played against the opposition remained the same ($r_1 = 20$). The results were very promising, with a higher ratio for both populations against the expert as depicted in Figure 7. We were now seeing trends towards beating the expert, which was exactly what we were after.
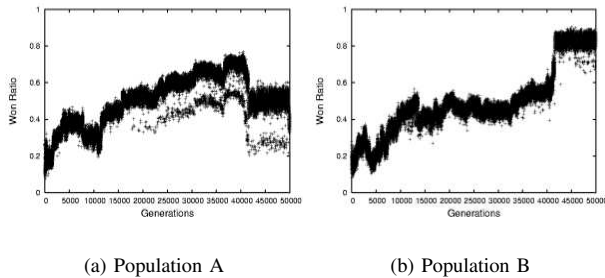

(a) Population A (b) Population B

Fig. 7.   Success ratio against the expert using Short and Long Term Memory

F. Experiment 6

One of the decisions made early on in the process was to make the History Bank unique and then increase their probability of getting chosen if the same individual was repeatedly added to the population. We decided to test this decision and performed a run with no uniqueness checking, and instead all the best individuals for all the generations were automatically added to the history.

Although there was not much of a difference, this was actually the first time the populations had achieved an outright 100% win ratio against the 'expert'. We were not expecting this, as we thought the unique population would have fared better due to the forced diversity of the populations. It occurred to us that this higher achievement might actually be caused as the short term memory would be filled with the best strategy, and therefore we are forcing the populations to find a solution to that single strategy.

G. Experiment 7

The gladiator long term memory system was used in this final experiment, and followed along the same lines as the previous history experiments. The fitness was calculated by playing first against the opposition ($r_1 = 20$), then against the short term memory ($r_{2S} = 10$) and finally against the

gladiator long term memory ($r_{2G} = 10$). The results can be seen in Figure 8.

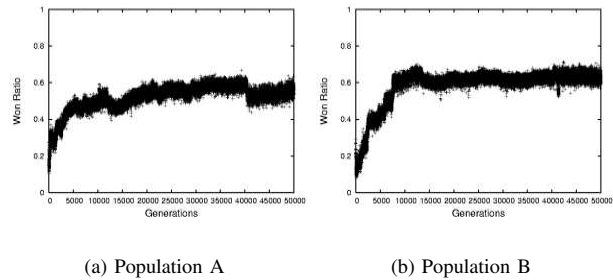(a) Population A (b) Population B

Fig. 8.   Success ratio against the expert using Gladiator System

The interesting thing about this system is the individual results for the system, as almost every run had one population performing better that the other. An example of this can be seen in Figure 9 where a distinct trend of high performance is recorded. The other population however averaged around 40-50% win ratio. This observation is still being analyzed, however it is thought that this is caused when one population fills its gladiator ranking with similar strong individuals and the result is similar to the unique and not unique phenomena described in section IV-F.
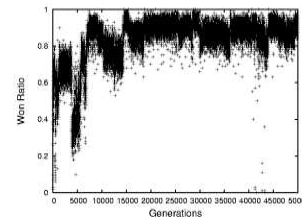

Fig. 9.   Success ratio against the expert for single run of Population B

H. Comparisons

A summary of the different experiments and the difference in parameters can be found in Table 1. The information depicted is as follows. The type of History Bank mechanism used is single for the initial case, then split into short and long, and lastly gladiator. The Probability Selection Distribution depicts the type of distribution used to select an individual for game play from the History Bank. The Game number shows the number of games each individual from each population must play to calculate the fitness: r_1 is the games played against the opposition, r_2 is games against the History Bank, r_{2S} is games against the short term memory, r_{2L} is games against the long term memory and r_{2G} is games against the gladiator long term memory. The History Bank lengths show what size each of the history bank mechanisms are, where L is the single History Bank length, L_s is the short term memory length, L_l is the long term memory length and L_g is the gladiator memory length. The unique field is a Boolean field representing whether the History Bank applied

uniqueness of individuals or not. Finally, the Alien Expert field depicts if the populations were seeded with the Alien Expert for the experiment.

TABLE I
SUMMARY OF THE EXPERIMENTS

Exp.	History Bank	Prob. Select Dist.	Game No.	HB Length	Unique	Alien Expt.
1	$\emptyset$	N.A.	$r_1 = 20$	N.A.	N.A.	no
2	Single	Uniform	$r_1 = 20$, $r_2 = 20$	$L = \infty$	yes	yes
3	Single	Linear	$r_1 = 20$, $r_2 = 20$	$L = \infty$	yes	yes
4	Single	Laplace	$r_1 = 20$, $r_2 = 20$	$L = \infty$	yes	yes
5	Short, Long	Linear	$r_1 = 20$, $r_{2S} = 10$, $r_{2L} = 10$	$L_s = 10$, $L_l = \infty$	yes	yes
6	Short, Long	Linear	$r_1 = 20$, $r_{2S} = 10$, $r_{2L} = 10$	$L_s = 10$, $L_l = \infty$	no	yes
7	Short, Gladiator	Ranked	$r_1 = 20$, $r_{2S} = 10$, $r_{2G} = 10$	$L_s = 10$, $L_g = 1000$	yes	yes

V. CONCLUSIONS AND FURTHER WORK

The experimentation of memory in a coevolutionary system has shown that the short and long term memory approach was beneficial. We demonstrated that the sorting and selection mechanism for the history population affects the usefulness of the memory, and that by including a specialized memory selection mechanism represented by the short term memory, the system could make better use of its memory.

The short term memory mechanism works like the human memory mechanism for learning; it allows a person to overcome information overload and use the most useful and recent memory rather than their large long term knowledge base. The long term memory is still there for prompting historical memories and reminding the person of past successes and mistakes, and allows the person to grow overall. We have attempted to replicate this in the TEMPO system, allowing the system to have reminders of past strategies, while still giving the best information for the current situation.

The inclusion of more than one selection mechanism from a history based on long and short term memory has many different areas that could still be investigated. The idea of an even more biased history, where the current situation could affect which strategies from the history to play against based on environment and ranking could be one area to investigate. There are also areas involving different ways to rank the individuals in the gladiator system, which is currently a very simple stochastic mechanism. There are many ways that the evolutionary parameters could be experimented with to determine what the optimum values for the games played against each memory should be.

One area we have begun to investigate is to have an adaptive number of games played against each memory set dependent on the environment and the current stage of evolution. At the moment there are user defined parameters representing each set of games to be played to get the fitness,

including games against the opposition and the history. It might be that a better alternative would be to let the system itself decide what number of games to play. To this end, investigation has begun into the use of a fitness threshold, where the number of games played adjusts itself as the size of the fitness variation increases or diminishes.

We will also be investigating further into the ways the long term memory is stored and retrieved. As mentioned in section II-B, the long term memory in a human is divided into explicit and implicit areas. It might be beneficial to try and find a way to mimic this. In particular, this involves looking at the linking of relevant information for retrieval from the long term memory. One possibility is to categorize the rules in the long term memory into the areas they effect, for example, if a rule set is focusing on a low budget then that might be relevant to other current rules that are also focusing on a low budget. Another possible mechanism could follow the case-based memory technique used by Sushil and Johnson [16]. Lastly, we shall investigate possibilities to incorporate some generic semantics into the system to represent 'social' learning for the whole population.

REFERENCES

[1] S. G. Ficici and J. B. Pollack, "A game-theoretic memory mechanism for coevolution," in *GECCO 2003, Lecture Notes in Computer Science*, vol. 2723/2003. Springer Berlin/Heidelberg, 2003, pp. 286–297.

[2] D. Cliff and G. F. Miller, "Tracking the Red Queen: Measurements of adaptive progress in co-evolutionary simulation," in *European Conference on Artificial Life*, 1995, pp. 200–218.

[3] C. D. Rosin and R. K. Belew, "New methods for competitive coevolution." *Evolutionary Computation*, vol. 5, no. 1, pp. 1–29, 1997.

[4] J. Branke, "Memory enhanced evolutionary algorithms for changing optimization problems," *Congress on Evolutionary Computation*, pp. 1875–1882, 1999.

[5] A. Bader-Natal and J. B. Pollack, "A population-differential method of monitoring success and failure in coevolution," Tech. Rep., 2004.

[6] R. W. Johnson, Z. Michalewicz, M. E. Melich, and M. Schmidt, "Coevolutionary tempo game," *Congress on Evolutionary Computation*, vol. 2, pp. 1610–1617, 2004.

[7] R. W. Johnson, M. E. Melich, Z. Michalewicz, and M. Schmidt, "Coevolutionary optimization of fuzzy logic intelligence for strategic decision support," in *IEEE Transactions on Evolutionary Computation*, vol. 9, 2005.

[8] D. B. Fogel, *Blondie24: Playing at the Edge of AI*. San Francisco, CA: Morgan Kaufmann, 2002.

[9] K. Chellapilla and D. B. Fogel, "Evolving an expert checkers playing program without using human expertise," in *IEEE Transactions on Evolutionary Computation*, vol. 5, no. 4, 2001, pp. 422–428.

[10] F. Glover and M. Laguna, *Tabu Seach*. Kluwer Academic Publishers, 1997.

[11] R. Atkinson and R. Shiffrin, "Human memory: a proposed system and its control processes." in *The Psychology of Learning and Motivation: Advances in Research and Theory*, vol. 2. New York: Academic Press, 1968, pp. 89–195.

[12] A. D. Baddeley, *The Psychology of Memory*. John Wiley and Sons, 2005.

[13] F. Craik, A. Routh, and D. Broadbent, "On the transfer of information from temporary to permanent memory," in *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, vol. 302, no. 1110, 1983, pp. 341–359.

[14] F. Craik and R. Lockhart, "Levels of processing: a framework for memory research," in *Journal of Verbal Learning and Verbal Behavior*, vol. 11, 1972, pp. 671–684.

[15] L. van Valen, "A new evolutionary law," *Evolutionary Theory*, vol. 1, pp. 1–30, 1973.

[16] S. J. Louis and C. Miles, "Combining case-based memory with genetic algorithm search for competent game AI," *ICCBR Workshops*, pp. 193–205, 2005.