The Evolution of Blackjack Strategies

Graham Kendall
University of Nottingham
School of Computer Science & IT
Jubilee Campus,
Nottingham, NG8 1BB, UK
gxk@cs.nott.ac.uk

Craig Smith
University of Nottingham
School of Computer Science & IT
Jubilee Campus,
Nottingham, NG8 1BB, UK
cxs00u@cs.nott.ac.uk

Abstract- In this paper we investigate the evolution of a blackjack player. We utilise three neural networks (one for splitting, one for doubling down and one for standing/hitting) to evolve blackjack strategies. Initially a pool of randomly generated players play 1000 hands of blackjack. An evolutionary strategy is used to mutate the best networks (with the worst networks being killed). We compare the best evolved strategies to other well-known strategies and show that we can beat the play of an average casino player. We also show that we are able to learn parts of Thorpe's Basic Strategy.

1 Introduction

Games can be separated into two distinct types: those of perfect and imperfect information. Games of perfect information offer the players complete knowledge of the game. That is, there is no hidden information and no random events. It is also usual for the players to take turns.

Researchers have been utilising computers to solve perfect information games for over 50 years. Shannon's seminal paper from 1949 [1] first considered using a computer to play chess. He suggested a method of evaluation that looked ahead a certain number of moves, using a game tree, and then selected the most appropriate action

Despite the huge search tree of chess, the first computer chess players were competing in tournaments as early as 1967. This culminated, 30 years later, with Deep Blue beating the reigning world champion Gary Kasparov by 3.5 to 2.5. There has also been notable work done on the game of checkers with Jonathan Schaeffer [2] winning the world championship in 1994 with the Chinook program, taking the title from Marion Tinsley who had held the title for 40 years. David Fogel [3,4] and Kumar Chellapilla and David Fogel [5,6] have also produced checkers programs that can play at an expert level.

The investigation of perfect information games is heavily reported in the literature but imperfect information games have remained relatively untouched, but with some notable work being carried out.

The GAMES (Game playing, Analytical Methods, Minimax search and Empirical Studies) group - headed by Jonathan Schaeffer - at the University of Alberta have stated it wishes to create an automated poker player that is able to compete at world-class level. Its publications date back to 1995 [7]. Its most recent paper ([8], reproduced in [9]) outlines the work their work so far and its aims for the future.

Significant work has been carried out by Gerry Tesauro [10] in investigating backgammon. His approach adopts a neural network that trains itself to learn the game by playing against itself. Tesauro utilised temporal difference to change the weights of the synapses. His program won the first computer Olympiads in 1989 and also went on to gain Master-level status.

2 Related Work

Edward O. Thorpe's seminal work [11] introduced a system (see Appendix) that gave a blackjack player an edge over the casino. Thorpe's strategy presented the player with a set of rules that informed the player what to do in a given situation. For example, if the player has a total of 16 should they draw another card or stand, in the hope that the dealer would bust? In fact, if the dealers upcard is six, or less, the player should stand. If the dealer is showing 7, or greater, the player should draw. The *basic strategy* could easily be learnt and it could be extended by the use of card counting. This keeps track of the proportion of ten cards (10's, Jacks, Queens and Kings) versus low cards that remain in the deck. This can be used to the player's advantage because he can vary the size of his bet, and change his drawing strategy, depending on the card count. *Beat the Dealer*, [11] was so popular that, in 1963, it was on the New York Times bestseller list. Thorpe's strategy was so effective, in that it gave an edge to the player, that casinos were forced to redefine the rules. They increased the number of decks from one to four (or six or eight). They also shuffled the decks when about 75% of the cards had been used. The casino's effectively banned card counters by asking them to leave if they suspected a player of using this skill.

The game of Blackjack has received relatively little interest from the AI community. In the early 1970s, Bernard Widrow [12] used the game to develop the idea of one of the first reinforcement learning algorithms called *"selective bootstrap adaptation."* His work learnt how to play blackjack without knowing the game or the objective of play. In more recent years Olson [13] used neural

networks with temporal difference to learn how to play the game. However his approach only used one player against the dealer and utilised limited blackjack rules. Fogel and Redfield [14] evolved a strategy, over millions of hands, that could improve on a basic counting strategy by about 0.5%. In 1998 Perez-Uribe & Sanchez [15] used a neural network to develop a blackjack player. Their program used improved learning strategies, Q-learning and SARSA[1] algorithms and was able to produce a much higher level of winning. However, again their program used simplified blackjack rules. In this case no doubling down or splitting and only using one deck instead of the more usual six that are used in casinos today.

3 The Game of Blackjack

The first people to analyse blackjack using a mathematical approach were a group headed by Roger Baldwin [16]. In this work they used probability and statistical theory to substantially reduce the house advantage. Although the title of their paper was 'optimum strategy', it was not really optimal because of numerous mistakes in their calculations, as shown by Thorpe in [17].

Since then there have been many different attempts to produce a better blackjack strategy but as the table below shows Thorpe's strategy is still one of the best, if used properly.

Blackjack Strategy	Player's Advantage (in percentage)
Thorpe's Basic Strategy	-0.64 (Using UK Rules)
Mimic the dealer	- 5.7
Never bust	- 6.0 (estimated)
Typical casino player	- 2.0 to -15.0

Figure 1 – Player Advantage for various Blackjack Strategies
(Taken from [11], page 33)

3.1 The rules of the Game

The aim of blackjack is to beat the dealer by obtaining a sum of cards that is as close to 21 as possible, without exceeding that value. If you exceed 21, you lose no matter what total the dealer finally achieves. The card total is made up from the value of the cards you are dealt, with picture cards being counted as 10 and Aces being either a 1 or 11. All other cards take their face value. The value an Ace takes is up to the player and, of course, they will choose the value that gives them the best score. A hand that contains an Ace which can count as 1 or 11 is known

as a *soft* hand. If the Ace must count as 1 then the hand is *hard*.

The rules of blackjack differ from place to place and even from casino to casino and so it is worthwhile describing the rules that are used in this paper (which are the rules used in UK casino's).

The game begins by each player placing their bet in the box in front of them. The dealer then deals two cards to each player and one to himself (in the US, the dealer would deal himself two cards). All cards are dealt face-up.

The players, having seen their two cards, have a decision to make. They can either *hit* (take another card) or *stand* (keep the total they have). If the player hits, the dealer turns up another card and the player now has the same decision. The player keeps hitting until he is satisfied with his hand or he *busts* (exceeds 21). If the player busts the player loses his money irrespective of what total the dealer subsequently makes (even if the dealer busts). Once all the players have completed their play, the dealer plays and keeps taking more cards until his total is 17 or greater. At the end of this play, the winner is defined as follows:

- If the player busted, he loses
- If the dealer busts, and the player did not, the player wins
- If the player has a total which is higher than the dealers, the player wins
- If the player has a total lower than the dealer, the player loses
- If the player and the dealer have the same total, it is draw (also known as a *push*) and the stake is retained

Wins for the player are paid at evens, apart from blackjacks (21 with 2 two cards), which are paid at odds of 3-2.

If the first two cards of the hand total 9, 10, or 11 the player has the option of *doubling down*. This means the player can double their original bet. If a player chooses this option, the dealer only deals one more card to that player.

When a player's first two cards are the same value (for example, a pair of sixes), they have the option to *split*. The player has to place another bet, equal to the original bet. The two cards are then split and played as two separate hands, with another card being dealt to each card before the player decides whether to split again, double down, hit, or stand. A player may only split a maximum of four times. When a player splits Aces, they only receive one more card on each Ace. If they get 21, it is not considered blackjack. Players are not allowed to split 4's 5's or 10's.

Every casino in Britain abides by these same set of rules, which are regulated by the Gaming Board. The rules, although still maintaining a house edge for the casino, do protect the player. For example, it is very bad to split a pair of fives, as not splitting them gives the player a score of ten, which is a good total on which to draw. In addition, splitting fives gives two hands of fives, with five being a bad card on which to draw. Therefore, the gaming board do not allow players to split 5's.

---

[1] The name SARSA actually comes from the fact that the updates are done using the quintuple Q(s, a, r, s', a'). Where: s, a are the original state and action, r is the reward observed in the following state and s', a' are the new state-action pair. The procedural form of SARSA algorithm is comparable to that of Q-Learning.: from http://www.cse.unsw.edu.au/~s2229705/rl/algorithms.html

2475

Sometimes this causes arguments in casino's when a player wants to split, yet the casino does not allow them (as the regulations forbid it). The players do not realise they are being protected and, if the truth were known, the casino would like you to have the opportunity to split. For similar reasons, players are not allowed to split 4's or 10's.

4 Learning the Game

In this paper we propose to utilise artificial neural networks to evolve good strategies for blackjack, in a similar way that Fogel used neural networks to evolve strategies for checkers [3].

Due to the distinct situations that arise in blackjack (splitting, doubling down and drawing) it was felt that a single neural network would not be appropriate and that three separate networks should be employed. One of the main reasons for this is that the different phases of the game require a different number of inputs (splitting and doubling down requires 3 inputs whereas the decision as to whether to hit or stand requires 16 inputs). In addition, having, for example, a separate split network means that we did not have to learn when we had to split. If the player's card values totaled 6, then in certain situations these cards could be split e.g. when the cards are both 3's, but in other situations this would not be allowed e.g. when the cards were 4 and 2. To train a network to recognise these different situations would be difficult, or, at least, it was felt more sensible to have a series of networks which were trained for a distinct task, and the networks were only used when we knew the input was valid (e.g. the split network would only ever be passed two player's cards of the same value so it only had to decide whether *to* split or not and not *if* it was legal to split. Finally, if we just used a single network, the output would be more complicated as there would have to be 4 possibilities, these being to tell the player whether to double down, split, draw, or stand.

Therefore, we investigate the use of three neural networks to learn the game of blackjack. One network deals with the possibility of splitting, the second deals with doubling down, and the third deals with normal game playing. In this way, we will be able to isolate certain parts of the game to see if the networks are learning the different aspects.

Each player during a game of blackjack has a number of decisions to make that are summarised in the flow chart below (figure 2):
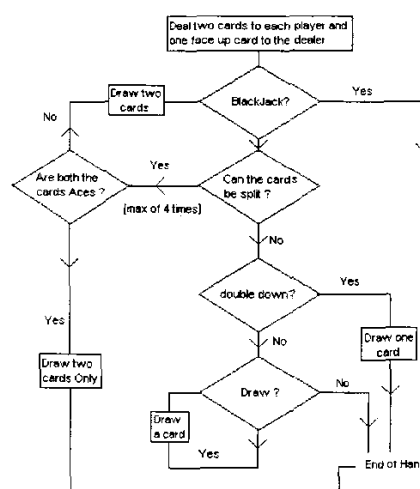

Figure 2 - Flow chart of players desisions during a hand of blackjack

4.1 Initialising the population

The population initially consisted of 60 players spread across 10 tables (in a casino a maximum of six blackjack hands can be played at a single table). Each table held six players and we dealt 1000 hands of blackjack for every generation. The players were assigned three neural networks, intialised with random (-0.5..0.5) synapse values. The first network, dealing with splitting, had three inputs, the values of the players' two cards and the value of the dealer's up card. This network had one hidden layer, comprising one node. The network outputs either 1 or 0, which indicated whether to split the cards or play them as a single hand.

The cards were only passed into the network if they could be split; therefore the network did not have to learn to deal with illegal situations.

If the cards equaled 9,10, or 11, they were passed into the second neural network, which handled doubling down. Again this network had an input of the players' two cards and the dealer's up card and produced an output of either 1 or 0, which again indicated if the player should take a card or not. The diagram below (figure 3) shows the basic structure of these two networks.
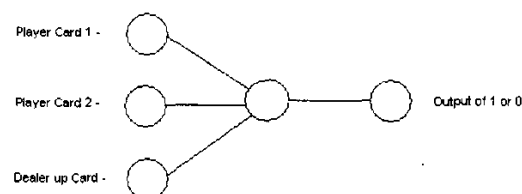

Figure 3 - Architecture for doubling down and splitting networks

The final network dealt with the majority of the hands. It had 16 inputs. This number was required to cope with the largest number of cards a player could be dealt, this being 2,2,2,2,2,2,A,A,A,A,A,A,A,A,A^2, fifteen cards for the player and then a single card for the dealer. In most cases (probably all) the players hand would not comprise 16 cards and the unused nodes were assigned a value of zero. The hit/stand network (figure 4) had 5 hidden nodes and one output node that will tell the player to either take another card or to stand on the cards that they already have.
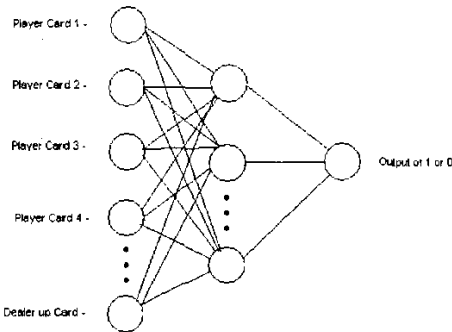

Player Card 1 -
Player Card 2 -
Player Card 3 -
Player Card 4 -
Dealer up Card -
Output of 1 or 0

Figure 4 - Architecture for the hit/stand neural network

4.2 Evolving the players

Once each player had played its 1000 games the population was evolved using an evolutionary strategy. The top 30 performing individuals were cloned to make the new population of 60 players (the lower 30 performing players were killed). The same mutation strategy was used as suggested by Fogel and Chellapilla [4].

1) $\sigma'(j) = \sigma_i(j)\exp(\tau N_j(0,1))$ $j = 1,..,N_w$

2) $w'(j) = w_i(j) + \sigma'(j)N_j(0,1)$ $j = 1,..,N_w$

where τ is: $\tau = (2(N_w)^{0.5})^{0.5}$

and where N_w is the number of weights in the specific neural network. This mutation is applied to all three of the neural networks assigned to each player.

5 Results

5.1 Doubling Down

Doubling down is a smaller game inside the large game of blackjack and because we used a separate neural network, experiments could be carried out in order to demonstrate that the player was learning this aspect of the game. Due to the rules of blackjack, once a player has doubled down only one more card can be taken and so we

could conduct the double down experiments in complete isolation.

In this initial experiment we ignored any player's card values that could not be doubled down i.e. two vard hands that do not equal 9, 10, or 11. Therefore, the only neural network that we use in this initial experiment was the doubling down network, all other hands being ignored. If the player did not have a doubling down hand then no money changed hands and the cards were discarded.

The following discussion is based on 1000 hands being played at each table (6 players = 6000 hands). 1000 generations, were run, making a total of 6,000,000 hands. About 15% of the hands were suitable for doubling down (i.e., 9, 10, or 11), so the doubling down network learning was based on around 900,000 hands.
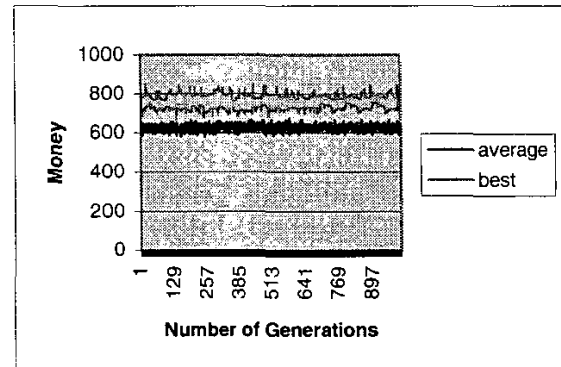

Figure 4 - Doubling down with 1 hidden node

As figure 4[3] shows there was no real learning, as the amount of money did not increase over successive generations. That is, the players did not learn when to double down and when not to.

We next used a neural network with no hidden nodes (a simple perceptron). In fact the two networks (one with a single hidden node and a simple perceptron), in theory, have the same capability. We did try using a network with two hidden nodes, with similar results as those shown in figure 4; although this does require further investigation. The simple perceptron produced much better results, as figure 5 shows.

---

[2] The cards will have to be dealt in this order as it forces all the Ace cards to be counted as 1 instead of 11.

[3] As with all the graphs the average player is the sum of all the player's values divided by the total number of players.
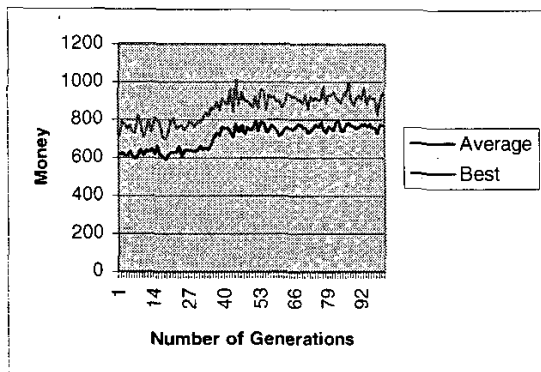
Figure 5 - Doubling down as a Perceptron

After around 50 generations the players began to learn which cards to double down on and which cards not to double down on as the average and best money values for the population begin to increase.

5.2 Normal Play Only

For normal play, the architecture has 16 input nodes. 5 hidden nodes and 1 output node. We ignored any splitting or doubling down situations, so all the hands were passed through the normal play network. We ran the program for 1000 generations and figure 6 shows the results.
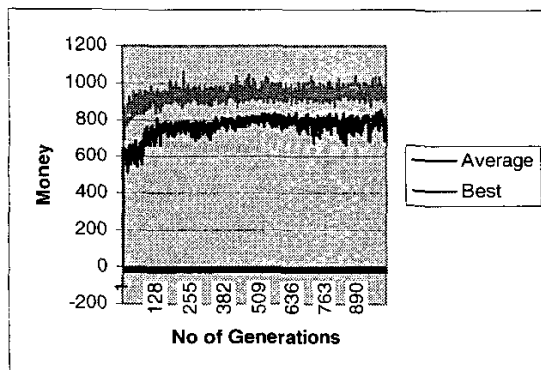

Figure 6 - Normal Play Only

The figure shows that learning clearly occurred, as at around generation 100 the best players begin to learn when to hit or stand.

5.3 Splitting The Cards

Experimentation with the splitting network is not quite as easy to carry out as, unlike the doubling down and normal play networks, there is no direct gain from splitting the cards. When the cards are split, they are then passed into the normal playing network or, in some cases, the double down network and then the results of these networks determine if the hand wins or not. So in order to test the splitting network all three networks were run together. We hoped that when the splitting network was included in the experiments that the winnings would increase.
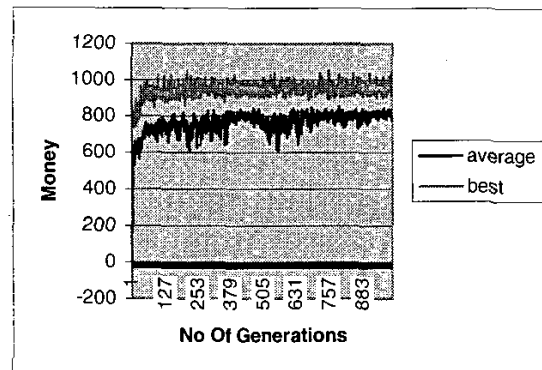
Figure 7 - All three networks running

Figure 7, showing the results from using all three networks, is no different from figure 6. This indicates that the splitting network is not working as we would expect. Therefore, we simplified the architecture for the splitting network as we had used for the doubling down network (i.e., a simple perceptron).
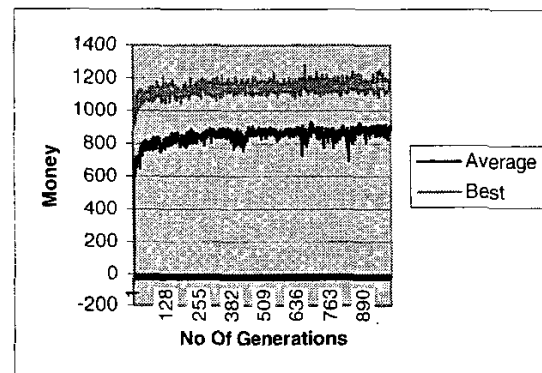
Figure 8 - All three Networks running

Figure 8 shows these results with the simple perceptron and demonstrates that there is an increase in the money over figure 6. This suggests that all three networks were working and the players were learning how to play the game of blackjack to some degree.

5.4 Comparing Strategies

We compared the population's best strategy against four other well-known blackjack strategies (figure 1). We took the average of the last 1000 generations from the data shown in figure 8 and plotted these (figure 9).

2478
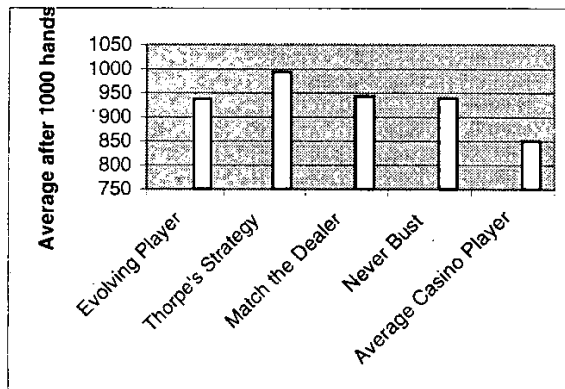
Figure 9 - Comparing Blackjack strategies

Our evolved strategy is superior to an average casino player and on a par with the 'Match the Dealer' and 'Never Bust' strategies. However the figure does show that the players are still around 5% down on Thorpe's strategy. We have compared the best players' output values with the outputs that Thorpe suggests in his strategy (See appendix for the Basic Strategy).

Player's Hand	Dealer's Upcard									
	2	3	4	5	6	7	8	9	10	A
9	D	D	D	D	D	*D*	*D*	*D*	N	N
10	D	D	D	D	D	*N*	D	*N*	N	N
11	D	D	D	D	D	*N*	*N*	*N*	*N*	N

Figure 10 - Doubling Down Strategy

The values in italics are the values that disagree with Thorpe's strategy. The table shows that the players learn a general trend that if the dealer has a low card, there is more chance of him going bust, so the player should double down. As the dealer up cards get higher, the evolved player begins to disagree with Thorpe's strategy. We believe that this is because the players will occasionally win when the dealer has a high card and so there will always be players in the population that will be slightly off Thorpe's strategy and still win.

Player's Hand	Dealer's Upcard									
	2	3	4	5	6	7	8	9	10	A
2	H	H	H	H	H	H	H	H	H	H
3	H	H	H	H	H	H	H	H	H	H
4	H	H	H	H	H	H	H	H	H	H
5	H	H	H	H	H	H	H	H	H	H
6	H	H	H	H	H	H	H	H	H	H
7	H	H	H	H	H	H	H	H	H	H
8	H	H	H	H	H	H	H	H	H	H
9	H	H	H	H	H	H	H	H	H	H

10	H	H	H	H	S	S	H	H	H	H
11	H	S	S	S	S	S	S	H	H	H
12	S	H	*H*	*H*	*H*	H	H	H	H	H
13	S	S	*H*	*H*	*H*	H	H	H	H	H
14	S	*H*	*H*	*H*	*H*	H	H	H	H	H
15	S	S	*H*	*H*	*H*	H	H	H	H	H
16	S	S	S	S	S	*H*	*H*	*H*	*H*	*H*
17	S	S	S	S	*H*	*H*	*H*	*H*	*H*	*H*
18	S	S	S	S	S	S	S	S	S	S
19	S	S	S	S	S	S	S	S	S	S
20	S	S	S	S	S	S	S	S	S	S
21	S	S	S	S	S	S	S	S	S	S

Figure 11 - Normal Play Desisions

The table in figure 11 shows that when the player has a hand that has a value of less than 9 the player learnt that it is best to take a card as you cannot lose by taking one more card and it will always improve your total, without busting. The player also learnt that if you have a hand with a value higher than 17 then there is no point taking another card (when using basic strategy) because there is a very high chance you will go bust and so lose the hand. We believe that the relatively quick rise in the amount of money won by the players is due to this fact. However we found that for hands with the values 12, 13, 14, 15,16, and 17 the neural networks struggle to learn what the best strategy is and so disagrees with Thorpe's strategies

Player's Hand	Dealer's Upcard									
	2	3	4	5	6	7	8	9	10	A
AA	Sp	Sp	Sp	Sp	Sp	Sp	Sp	*N*	*N*	*N*
22	Sp	Sp	Sp	Sp	Sp	*N*	N	N	N	N
33	Sp	Sp	Sp	Sp	Sp	*N*	N	N	N	N
66	Sp	Sp	*N*	*N*	*N*	N	N	N	N	N
77	Sp	Sp	*N*	*N*	*N*	*N*	Sp	*Sp*	*Sp*	*Sp*
88	Sp	Sp	*N*	*N*	*N*	*N*	*N*	*N*	*N*	*N*
99	Sp	Sp	Sp	Sp	Sp	Sp	*N*	N	N	N

Figure 12 – Splitting Decisions

In figure 12 "Sp" indicates that the player will split the cards and "N" means that the player will not split. Again the table show that for the 'clear cut' cases e.g. the high cards 9,10,A and the low cards 2,3 the players learn what actions to take and what actions not to take but it is the middle cards where the players are not learning he correct actions. We feel this is due to the same problems as the previous two networks, that sometimes the statistical difference between the right action and the wrong action are too marginal for the networks to learn.

2479

6 Conclusions

Blackjack has been shown to be a game of not just luck but skill as well. Mathematicians have derived strategies and this paper attempted to evolve strategies by allowing a "knowledge-less" population to develop its own strategies by evolution. The results show the players learnt the basic actions fairly quickly and then struggled to learn the appropriate actions in more complex situations where the values cards were not as clear cut. The fact that there is a large element of luck involved in blackjack does not help the population to learn these complex hands. Sometimes the players were actually taking the wrong actions and still winning the hand. This was good in the short term but the incorrect decision in the longer term. However we were hoping that with each player playing 1000 hands per generation these lucky hands would be counter balanced with unlucky hands. Looking back this is not a large number of hands by any means, in fact when new strategies are being tested, sometimes the players will play up to 800 million hands.

7 Acknowledgements and Further Comments

We would like to thank the reviewers for their very helpful comments on this paper. All three reviewers offered very constructive and detailed comments. Their remarks raised a number of questions many of which we have addressed in the revision of this paper and some which will motivate our future research. One common theme was to ask why we had used neural networks rather than directly evolving a series of look up tables. This is a reasonable question and we must admit it may be a more fruitful research direction, However, one of the main motivations for this paper was to take the work of Fogel and the evolution of checkers strategies and to apply this idea to a (relatively simple) game of imperfect information.

The architecture of the networks is also worthy of further investigation, as is the number of generations and statistical analysis of the results.

We believe that it should be possible to learn Thorpe's Basic Strategy *perfectly* (or improve upon it) and then we can investigate other aspects of the game such as card counting and betting strategy.

Although we have almost exactly implemented the rules as used in UK casinos there are some rules that we still need to incorporate. For example, we do not include any aspects of *insurance* bets. For example, in the UK, if you have blackjack and the dealer is showing an Ace you can elect to take even money (rather than the 3:2 normally paid for blackjack) before the dealer plays. If they turn up a 10, you still get paid but not so highly.

This paper was originally written as a brief investigation of the game but the reviewers comments have motivated us to carry out a much more in depth study of this game with respect to evolving strategies.

Appendix – Thorpe's Strategy

Player's Hand	Dealer's Upcard									
	2	3	4	5	6	7	8	9	10	A
12	H	H	S	S	S	H	H	H	H	H
13	S	S	S	S	S	H	H	H	H	H
14	S	S	S	S	S	H	H	H	H	H
15	S	S	S	S	S	H	H	H	H	H
16	S	S	S	S	S	H	H	H	H	H
A2	H	H	H	D	D	H	H	H	H	H
A3	H	H	H	D	D	H	H	H	H	H
A4	H	H	D	D	D	H	H	H	H	H
A5	H	H	D	D	D	H	H	H	H	H
A6	H	D	D	D	D	H	H	H	H	H
A7	S	D	D	D	D	S	S	H	H	H
A8	S	S	S	S	S	S	S	S	S	S

Table 1 - Normal Game Playing

NOTE : any hard hands of greater than 16 should always stick and any hard hands of less than 9 should always take a card

Player's Hand	Dealer's Upcard									
	2	3	4	5	6	7	8	9	10	A
9	H	D	D	D	D	H	H	H	H	H
10	D	D	D	D	D	D	D	D	H	H
11	D	D	D	D	D	D	D	D	D	H

Table 2 - Doubling Down

Player's Hand	Dealer's Upcard									
	2	3	4	5	6	7	8	9	10	A
AA	Sp	Sp	Sp	Sp	Sp	Sp	Sp	Sp	Sp	Sp
22	Sp	Sp	Sp	Sp	Sp	Sp	H	H	H	H
33	Sp	Sp	Sp	Sp	Sp	Sp	H	H	H	H
44										
55										
66	Sp	Sp	Sp	Sp	Sp	H	H	H	H	H
77	Sp	Sp	Sp	Sp	Sp	Sp	H	H	H	H
88	Sp	Sp	Sp	Sp	Sp	Sp	Sp	Sp	Sp	Sp
99	Sp	Sp	Sp	Sp	Sp	S	Sp	Sp	S	S
1010										

Table 3 - Splitting Rules

Where *Sp* stands for Splitting, *D* stands for Doubling Down, *H* stands for hit and finally *S* stands for stand.

2480

References

[1] Shannon, C. E., 1950, *"Programming a computer to play chess"*. Philosophical Magazine [Series 7] 41: pp 256-275

[2] Schaeffer J., 1997, *"One Jump Ahead: Challenging Human Supremacy in Checkers"*. ISBN: 0387949305

[3] Fogel D.B., 2002, *"Blondie24: Playing at the Edge of AI"*. Morgan Kaufmann, SF, CA

[4] Fogel D.B. 1999, *"Anaconda Defeats Hoyle 6-0: A Case Study Competing an Evolved Checkers Player Against Commercially Available Software"*, Proceedings of the 2000 Congress on Evolutionary Computation, IEEE Press, Piscataway, NJ, 857-863

[5] Chellapilla K and Fogel D.B., 2001, *"Evolving an Expert Checkers Playing Program without Using Human Expertise"*, IEEE Transactions on Evolutionary Computation, Vol. 5:4, pp. 422-428

[6] Chellapilla K. and Fogel D.B., 1999, *"Evolving Neural Networks to Play Checkers without Expert Knowledge"*, IEEE Trans. Neural Networks, Vol. 10:6, pp. 1382-1391

[7] Billings D., 1995, *"Computer Poker"*, MSc Thesis, University of Alberta

[8] Billings D., Davidson A., Schaeffer J. and Szafron D., 2002, *"The challenge of Poker"*. Artificial Intelligence 134 (1-2) : pp 201 –240.

[9] Billings D., Davidson A., Schaeffer J. and Szafron D., 2002, *"The challenge of Poker"*. Chips Challenging Champions: Games, Computers and Artificial Intelligence, Elsevier, pp. 243-282.

[10] Tesauro G., 1995, *"Temporal difference learning and TD-Gammon*, Communications of the ACM"*, 38(3): pp 58-68.

[11] Thorpe E.O., 1963, *"Beat the Dealer. A winning strategy for the game of twenty one"*, ISBN 0-394-70310-3, Vintage Books, NY.

[12] Widrow B., Gupta N. and Maitra S., 1973, *"Punish/Reward: Learning with a Critic in Adaptive Threshold Systems"* IEEE Transactions on Systems, Man and Cybernetics, vol. 3, no.5, pp. 455-465

[13] Olson K., 1993, "Learning to play games from experience: An application of artificial neural networks and temporal difference learning" M.S. thesis, Pacific Lutheran University, Washington

[14] Fogel D.B. and Redfield R., 1993, *"The Evolution of Blackjack"*. TIMS/ORSA Conference, Invited by S. Yakowitz, Phoenix, AZ, Nov 3.

[15] Perez-Uribe A. and Sanchez E., 1998, *"Blackjack as a Test bed for Learning Strategies in Neural Networks"*. Proceedings of Logic Systems Laboratory, Department of Computer Science, Swiss Federal Institute of Technology, Switzerland

[16] Baldwin R., 1953,*"The Optimum Strategy in Blackjack"*, Journal of the American Statistical Association Vol. 51, pp. 429-439

[17] Thorpe E.O., 1961,*"A Favorable Strategy for Twenty-One"*, Proceedings of the National Academy of Sciences, vol. 47, No.1, pp. 110-112