

# Laboratory work in AI: First steps in Poker Playing Agents and Opponent Modeling

Avram Golbert  
015724669  
agolbert@gmail.com

## **Abstract:**

While Artificial Intelligence research has shown great success in deterministic full-knowledge games such as chess and checkers, and in full-knowledge non-deterministic games such as back-gammon, the best artificial players in poker are still far behind the top human players. Poker, unlike the games mentioned above involves hidden information, and each player plays deceptively to intentionally hide information about their cards, causing even best known search algorithms such as alpha-beta to be incapable of finding a good strategy in reasonable time.

In this work I implement an agent along the general lines described by Aaron Davidson<sup>1</sup>. Instead of searching the tree looking for minimax, the agent searches the game tree looking to maximize the expected outcome according to the opponents' expected behavior.

## **Introduction:**

### **Rules of Limit Texas Hold'em Poker<sup>2</sup>:**

#### **The Blinds**

In Texas Hold'em, a disc called "the button" indicates which player is the nominal dealer for the current game. Before the game begins the player immediately clockwise the button posts the "small blind", typically half a small bet (for example, \$0.5 in a \$1/\$2 game). The player immediately clockwise from the small blind posts the "big blind", which is always a full small bet (\$1 in a \$1/\$2 game). Now each player receives his or her two hole cards.

#### **Pre-Flop**

After seeing his or her pocket cards, each player now has the option to play his or her hand by calling or raising the big blind. As mentioned before, the big blind is the size of a full small bet. Thus in a \$1/\$2 Texas Hold'em game, it would cost \$1 to call in this initial round of betting (known as the "pre-flop").

In Texas Hold'em the available actions are fold, bet, call or raise. These options are available depending on the action taken by the previous player. Each poker player always has the option to

fold. The first player to act has the option to check or bet<sup>i</sup>. Subsequent players have the option of calling or raising. To call is to match the same amount as the previous player has bet. To raise is to match the previous bet and increase the bet.

Now, the "flop" is dealt face-up on the board. The flop consists of the first three community poker cards available to all active players. Betting begins with the active player immediately clockwise the button. All bets and raises occur in small bets (increments of \$1 in a \$1/\$2 game). The same rules apply from above to complete this round of poker betting.

### **The Flop**

Now three cards are dealt face-up on the board - this is known as the flop. In Texas Hold'em, the three cards on the flop are community cards available to all players still in the hand. Betting begins with the active player immediately clockwise of the button. All bets and raises occur in small bets (increments of \$1 in a \$1/\$2 game).

When betting action is completed for the flop round, the "turn" is dealt face-up on the board. The turn is the fourth community card in a Texas Hold'em poker game. Play begins with the active online player immediately clockwise the button. On this round, same rules apply as above except that poker betting doubles from the small bet to the big bet. In a \$1/\$2 game, betting on the turn is done in \$2 increments.

### **The Turn**

When the betting action is completed for the flop round, the "turn" is dealt face-up on the board. The turn is the fourth community card in Texas Hold'em. Play begins with the active player immediately clockwise of the button. On this round, the betting doubles from the small bet to the big bet. Thus, in a \$1/\$2 game, betting on the turn is upped to \$2 bets.

### **The River**

When betting action is completed for the turn round, the "river" is dealt face-up on the board. The river is the fifth and final community card in a Texas Hold'em game. Betting begins with the active player immediately clockwise the button and the same poker rules apply as they do in the fourth card explained above.

### **The Showdown**

At the end of the final betting round, if there is more than one remaining poker player, the last bettor or raiser shows his or her cards first. If there was no bet on the final round, the player immediately clockwise the button shows his or her cards first. The player with the best five-card hand wins the pot. In the event of identical hands, the pot will be equally divided between the players with the best hands. Texas Hold'em rules state that all suits are equal, so split pots are more common than in other variations of poker.

---

<sup>i</sup> "Check" and "bet" are a specific instance of call and raise respectively, when the amount to call equals zero.

After the pot is awarded, a new game of Texas Hold'em is ready to be played. The button now moves clockwise to the next player and new hands are dealt to each player.

### **Analysis of Poker Characteristics:**

Limit Texas hold'em Poker is a game played sequentially. Each hand is played with a new deck and is independent of the previous games. While each hand is not a zero value game, a sequence of hands is since position of players, and therefore the order of actions, is rotated from one hand to the next.

Each hand is played in 4 stages; each stage starts with the dealer revealing some information followed by a round of betting<sup>ii</sup>. In the first round the information revealed is hidden, i.e. revealed only to each individual player in the form of 2 private cards (hole cards) in the following rounds the information is communal; 3 cards in the second round (the flop) and one card in each of the last 2 rounds (the turn and the river).

In each round of betting each player in his turn can choose from 3 options; fold, call or raise. The round of betting is over when all players have the same amount invested. If at any point of the game all players but one have folded, the remaining player is the winner and is awarded all money invested (the pot). After all 4 rounds are played the winner is chosen according to each remaining players' best 5 card hand chosen from the 5 community cards and his 2 hole cards. Although many variations exist the standard rules are mostly set and vary only slightly, example of such variations regard putting a limit on how many raises are allowed during each betting round and order of showing hands in showdown. I assume 3 raises per round and a player is never forced to reveal his hand if he's not contesting the pot (more on this later).

Notice there are two distinct ways to win a hand; show the best hand at Showdown, or get everyone to fold. Another factor you want to consider is the size of the pot when you win and the cost you pay when you lose. Obviously our goal is to win more then we lose, this can be achieved by winning a lot of small pots (typical of pots that don't go to Showdown) or winning a few large pots (typical of multi contested pots won at Showdown). In the former scenario, the winner's hand is irrelevant; in the later the winner's hand is typically very strong.

Strategies to achieve these goals are contradictory; to achieve high win rate when you go to Showdown you need to carefully choose your battles, while "taking stabs" at small pots can be done with higher frequencies. To simplify a point, a player's strategy is a balance of aggressiveness and conservatism - when aggressiveness is successfully applied it muscles out opponents and wins pots uncontested, but expensive when fails. A proper balance is hard to come by and should take into account the opponent's playing habits and predictability. Previous work has showed for simplified versions of poker that game-theory Nash equilibrium includes bluffing and slow playing [1].

This work is based on the assumption that the past is a good indication of the future. If our program (a.k.a agent, bot) successfully studies past behaviors of his opponent and can successfully

---

ii There are 2 exception; 2 players are forced to enter money into the pot before the hole cards are dealt. This is done to "jump start" the betting and force money into the pot. This slightly alters the order of betting in the first round.

The other is in the case were a player runs out of money. In this work I assume all players have enough money to finish the hand, avoiding "all-in" exceptions.

predict their actions and hole cards, then a best move can be found that will maximize the agent's payoff.

I make two basic (and real-world unrealistic) assumptions; an unlimited amount of hands to study and the opponent's strategy is constant, i.e. he's not a learning opponent. In addition, I only tested the program on games with 2 players although the code is written to be able to run with more.

## **Game representation:**

The game is represented by its game tree. In naïve form each node has as many children as possible moves by the player whose turn it is to act. Each node representing a state when its player p's turn to act, will have 3 children corresponding to fold, call or raise (2 in cases when raising has been capped). Regarding chance as a player (called the dealer) then when it's the dealer's turn to act there will be as many children nodes as possible deals;  $(2^n \text{ choose } 52) * 2^n / 2^n$  for the hole cards  $3 \text{ choose } (52 - 2^n)$  on the flop... etc.

It's easy to see that implementing this game tree and trying to run a game search is infeasible even using clever search methods such as alpha beta. Therefore the following reductions were made on a depth first search:

- Dealer buckets – instead of regarding each possible outcome for the dealer we can bucket different boards. A simple example is symmetry regarding suits. It's possible to characterize boards according to parity (and trinity) of cards; number of same suited cards and distribution of rank are good candidates since they indicate likelihood of flushes, full houses and straits. In this work only one single bucket was used; meaning only the agent's relative hand strength was calculated but no adjustments were made to the expected hand rank of opponents.
- Simulations – even with buckets, the possible outcomes are too great to enumerate. Instead of searching through all sub trees, a predefined number of outcomes are simulated and the expectancy on those is calculated. Since the interesting events (hitting a set on the flop, a straight or flush draw) have likelihood greater than 1:10, simulating 100 flops are enough to get a good expectation.
- Cutoff search with heuristic functions - a depth first search is used with Cutoff on the probability of the hand following the current branch. The Cutoff number is in direct relation to allowed time per move and in opposite relation to the speed of the system; number of nodes calculated per second. If a branch is cut off before reaching an edge a heuristic guess is used to calculate the expected value of the node.
- Heuristic function estimates the expected value for nodes with learned history as current pot size times the probability all opponents will fold plus the estimated pot size at Showdown times probability agent beats all opponents. The probabilities are calculated by the Opponent Profiler.

## **Opponent modeling:**

For each opponent a profiler is created. This profiler holds a simplified tree correlating to possible game states when it's the opponent's turn to act. Each node hold information regarding number of times reached, number of times an action was observed, and information of events in the sub-tree such as chance of eventually folding and hand strength.

The hand strength is calculated as relative hand strength between 0 and 1 correlating to its rank against all other hole cards with the given table. This normalization favors hole cards that “fit” with the board. A strong starting hand might not fare well if a highly suit-coordinated board comes out while low suited connectors can connect strongly on the flop.

When a Showdown occurs, the relative hand strength is calculated and added to a list held in the showdown node. If in the Showdown the opponent folds his cards, we can’t know what his hand strength was but we use the winning hand as an upper bound. This causes our observation to be a biased statistical sample. This situation is called a “right censored observations” in biology. The mean is calculated using the area under the Kaplan–Meier curve<sup>3</sup>.

The hand distribution was assumed to be normal for this work, and the standard deviation was calculated also from the Kaplan-Meier curve although this is a bad estimator and will always underestimate the true variance.

Once the mean and variance are calculated for the edge nodes in the profile tree the mean and variance are calculated recursively using the simple equations:

$$E[x]_{father} = \frac{\sum_{s \in Son} (E[x]_s * n_s)}{\sum_{s \in Son} n_s}$$

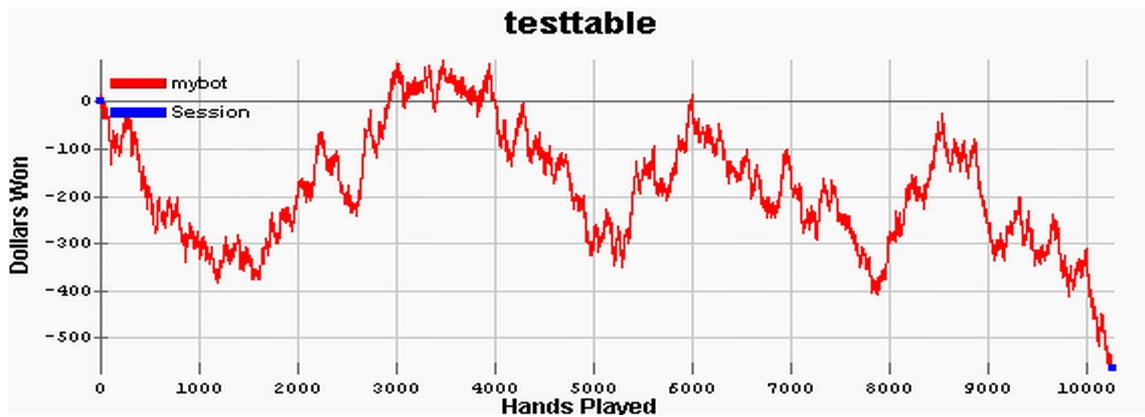
$$E[x^2]_{father} = \frac{\sum_{s \in Son} (E[x^2]_s * n_s)}{\sum_{s \in Son} n_s}$$

$$var(x) = E[x^2] - E[x]^2$$

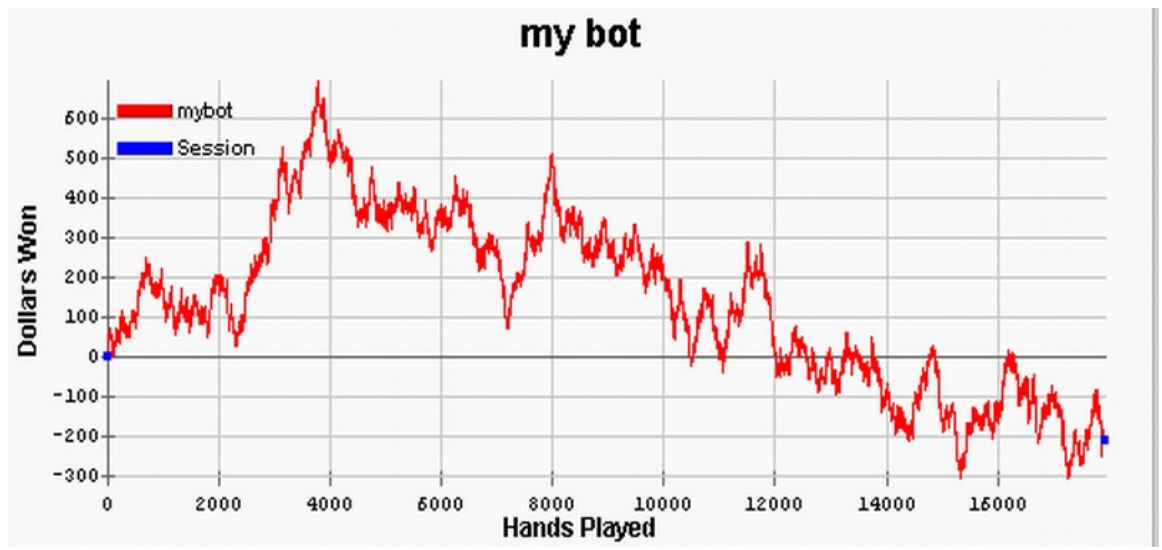
Each node holds a number of children corresponding to possible events that might occur between the opponent’s turns to act. Once again it’s not possible to hold a child for every possible path. Buckets are used here as well; each node holds up to n\*4 children that code information of how many players are left in the hand and how many bets the opponent must call. This bucketing smoothes over possibly vital information such as which opponents raised, how many have called the raise and table position.

## **Results:**

The agent played \$1/\$2 against a static aggressive opponent. Over 10,255 hands the agent lost an average of 0.06 cents/hand.



Then the agent observed the opponent for another 10,000 hands against an exploratory player. This was expected to help the agent learn strategies that he himself is not programmed to explore. After this learning period the agent lost a 30 cents/hand over 5,000 hands then evened out over 18,000 hands losing only 1 cent/hand:



Each hand averaged close to \$5. It's hard to quantify the inherent variance of the game, but using a model of random walks on a graph, over N hands two equal players should statistically end up anywhere in the range of  $\pm\sqrt{N}$  hands. The agent is just barely out of that range in the first set of hands but is well within that range for the second round, not including the initial 5000 hands.

There is too little data to try to determine trends in the agent's learning graph, although it's we can see a learning improvement that tends to "get stuck". As the agents profiler gather's more and more data it gets harder to unlearn bad habits and close to impossible to find new directions of play.

The agent seems to find local optimums that are highly dependant on the agent's initial assumptions. In this work, when an unknown state is reached, the agent searches up through the tree until it reaches a learned state. Clearly this is over simplistic and can cause entire sub trees to be avoided because of one

bad-beat. This was partially overcome by the 10K hands observed while the opponent was playing against an exploratory agent.

**Future work:** This agent makes two unrealistic assumptions; static opponents and unlimited learning curve. More work should be done to overcome these 2 points by finding better correlations between game states to faster learn the entire tree, starting with default opponent profiles and ways to quickly adapt them to form a satisfying opponent profile. More work is also needed to find ways to quickly change a given profile to find trends in opponent behavior to quickly recognize and adapt to opponent learning and “gear changing”.

Future work should be focused on finding “related” game states. Given a game state, either end node or mid game we would like to find other states that have the same, or close expected outcomes. This would greatly speed up the learning process, and bring better results on the heuristic function and predicting opponents.

#### **Poker Terms<sup>4</sup>:**

- **Blinds:** Short for "blind bets," these are the forced bets made before the cards are dealt. In Hold'em, blinds take the place of the classic "ante."
- **Button:** Nickname for the player acting as the dealer in current hand.
- **Check:** Similar to a call, but no money is bet. If there is no raise preflop, the big blind may check.
- **Flop:** The first three community cards dealt.
- **Preflop:** Anything that occurs before the *flop* is dealt is preflop.
- **River:** The final (5th) community card dealt; also known as *fifth street*.
- **Showdown:** When players reveal their hands to discover the pot's winner.
- **Turn:** fourth community card dealt; also known as *fourth street*.
- **Fold:** The player pay nothing to the pot, and throw away their hand, waiting for the next deal to play again.
- **Call:** The player matches the amount of the biggest bet.
- **Raise:** The player raise the bet by adding, on top of the biggest bet, the amount of the current betting round.

<sup>1</sup>References:

Opponent Modeling in Poker: Learning and Action in a Hostile and Uncertain Environment. Aaron Davidson

<sup>2</sup> [www.pokerstars.com](http://www.pokerstars.com) (copied and edited without permission)

<sup>3</sup> Kaplan, E.L. & Meier, P. "Nonparametric estimation from incomplete observations," Journal of the American Statistical Association, 53, 457-481 (1958).

<sup>4</sup> [www.pokerlistings.com](http://www.pokerlistings.com) (copied and edited without permission)