Artificial Poker Player
IT3105

Nordmoen, Jørgen H.
Østensen, Trond

September 17, 2012

Abstract

In this text we will explain our implementation of an A.I. Poker player in the course IT3105. We will explain each phase of the project, what choices we made and our decisions that have effected the outcome in each phase. We will include some results of the different phases and wrap up with our thoughts on the implementation and possible future work.

Contents

1 Basic Structure

Since we chose to go with Java instead of Python as our programming language in this course our structure might not be as simple and straight forward because of limitations or design off Java.

We chose to utilize Java because we both recognized the advantages of having the possibility of threading our code. Python does not support true multithreading[1] and since we both have good experience with Java we chose to go with that. Of course multithreading is not everything and the speed afforded to us by Java is quite apparent compared to Python. This choice however did mean that we could not use the code that was given to us in the course which meant we had some catching up to do in the starting phases.

1.1 Poker playing

We started out by implementing the basic of any card game, the cards, the deck and power rating. The later was much inspired by the code that we had gotten from the course, but implemented as a class which could be compared to others with most of the Java interfaces[2]. We spent quite some time getting this class right and have had some strange problems from time to time, but we ended up with something that can really do its job.

Most of the work of this class is just to determine the best rank that a player can get from either an array of cards or a poker hand and some community cards and be able to compare it self with other power ratings. Looking through this code there might be some odd bits and pieces that stick out, the lazy evaluation is probably a bit of a surprise. The reason for this choice is that we quite quickly realized that when we compared power ratings in the roll-out simulation most of the time we only need the rank it self because most evaluations of *compareTo* will

[1] `http://wiki.python.org/moin/GlobalInterpreterLock`
[2] `http://docs.oracle.com/javase/7/docs/api/java/lang/Comparable.html`

end then and there and doesn't need to do the costly evaluation of determining which cards should be kept and which kickers to use.

To implement the game it self we used a class called *PokerMaster* which deals with all the poker playing in our code. To enable it to support multiple phases of poker players we designed an interface which allows the *PokerMaster* to interact with all of the phases in a uniform manner. This interface, which can be found in *PokerPlayer*, supports all the methods that we needed in the later phases, but have seen several revisions before it got to this stage. This interface is then backed up by an abstract class which implements some of the tasks which every phase needed.

AbstractPokerPlayer does most of the work regarding chip count and make sures every phase pays what they need in order to participate further in a game. It also makes sure that blinds are payed when that is needed. When it comes to paying to the table we decided, for simplicity, that we would allow each phase to go negative. This has some ramifications for the game, because it means that there could potentially be much more chips involved than we intended at the start, but that have not been a huge problem as we force some of the phases to bet less when they are out of chips.

To make it easier for us on determining which player goes when we created the *PokerTable* class which contains some methods for keeping track of the blinds at the table and also which player is big and small blind. The *PokerMaster* uses this class to retrieve the small and big blind and also who to deal cards from.

To give each player cards we designed the *PokerHand* class which is just a helper class to enable us to compare hands, easily create power ratings and have a simple way of passing two cards around. It can compare it self to other hands, but this is not used much since we mostly deal in power ratings.

Since we share the *PokerMaster* code between the different phases there is not much else to say about this code in regard to the phases. We have some opponent modeling code which is only used by phase 3, but we will come back to that in section 4.

2 Roll-out simulation

As the code progressed we started looking at roll-out simulation in phase 2 since that was one of the big challenges in this project.

We started by looking at what we needed to simulate a complete poker game and came to the conclusion that we should make a separate poker simulator from *PokerMaster*. Since the roll-out simulation is not dependant on any players playing the game or any notion of betting and calling we decided that the *PreFlopMaster* would only deal cards and compare them at the end of the game to speed up the calculation. The *PreFlopMaster* gets the hand we want to gather statistics about and the number of players to play against. It then deals out cards to the "other" players and begins to deal the flop, turn and river. Once this is done it declares a winner and updates its statistics for that hand. It then goes on to simulate X amount of simulations, decidable on the command-line, with the given cards and a deck to play with. Once all the simulations are done it returns a *TestResult* instance with the given wins, ties and loses and a ratio.

As we mentioned before we decided to go with Java because we realized that

this roll-out simulation could easily be threaded to increase performance and so we did. We start by creating all the possible poker hands according to the hole-card equivalence classes[3]. Then we split these poker hands into lists of equal lengths with the number of lists equal to the number of processors reported by Java[4]. We then create *RolloutSimulators* which get a list each which then for each card create a new *PreFlopMaster* and simulates from 2 to 10 players and writes the results to a file. From what we can see this threading makes it possible to do quite a lot more simulations than without the threading and the factor that dominates the time is the shuffling of the deck which is done 7 times for each simulation.

From our limited experiments this allows us to do 100 000 simulations per hand with 2 to 10 players in about 15 minutes on a 4 core Intel i7 with Hyper-Threading.

3 Hand strength

As the roll-out simulation had started to work at this point we started looking at the phase 2 player and started to implement *Hand strength* calculation. Hand strength is used after the flop in order for the phase 2 player to determine if the current hand can be beaten by other hands. This is done by comparing the current hand and the current community cards to all other combination of hands that is possible to achieve without the cards in the hand or on the table.

This to is a simulation that can be done in parallel and we did this to.

We have implemented two different calculations the one described in our handed out materials which does not take into account the possible turn and river cards and one which does. The first one is simply implemented by creating a fresh deck, remove the cards on the players hand and the given community cards and then create all permutations of poker hands that is possible to get and then compare each with the players hand to see how many times the player wins, ties and loses. To run this in parallel we just split the poker hand permutations into lists as before and calculated all the wins, ties and loses in parallel before we collected all the statistics together and return a hand strength ratio.

For the second implementation we deal out all possible turn and river cards, again except for the players cards and the community cards, and create all possible poker hand permutations and do the same calculation as above. The result of this calculation is the lowest possible score that a pair of cards can give, meaning it is much more accurate than the above calculation.

A flaw in the first hand strength calculation above is that it will most likely over estimate the strength of a hand after the flop. For instance lets say the player we are looking at are sitting with two aces on h*s hand and one ace is on the table together with a 7 and a 10 of some suits. Since there can not be a straight present and lets say for this example that only the 7 and 10 are of the same suit, while the ace on the table is of a different suit there can not be a flush either. So the simple hand strength calculation reports that the strength is 1.0. This might be a basis for a player to bet because he/she thinks that they have the best cards and nothing can beat them. But the flaw is that the turn

[3] `www.idi.ntnu.no/emner/it3105/lectures/ai-poker-players.pdf`

[4] `http://docs.oracle.com/javase/7/docs/api/java/lang/Runtime.html#availableProcessors%28%29`

could be a card with the same suit as the 7 and 10 and then someone might have a flush or even a straight might begin to appear. The second implementation would account for this and give the lowest ratio that could possibly appear which would be an underestimate, but it could be much more accurate because it doesn't overestimate.

4 Opponent modelling

When it comes to opponent modelling most of the interesting features is in the contexts that we use to define a situation in the poker game that we think might come up again.

Each context gets added by the *PokerMaster* after each player make a move. When a game reaches showdown the *PokerMaster* informs the opponent modeller that a showdown has been reached and gives the modeller the players that have reached the showdown. The modeller then looks through all the contexts that have been added this round and selects those that have the players in the showdown in them. It then calculates their hand strength at each context and add these contexts to the pool of all seen contexts. When a player later gives the opponent modeller a context it checks the list to see if that context has been seen before and if it has it calculates an average hand strength and a standard deviation[5] and returns this to the player. The player can then use this information to assess the potential hand of an opponent and select an action based on this.

4.1 Context model

In our setup we have chosen to take the below factors into account when creating a context.

- The player
- The action the player took. I.e. Bet or call
- The current round. E.g. Pre-flop, post-flop etc.
- The amount of players still playing. We chose to split this into three categories, "FEW", "NORMAL" and "MANY". Where few players are anything below 1/3, normal is below 3/4 and many is above that.
- The pot odds size which is split into three again "SMALL", "MEDIUM" and "LARGE". Where small is pot odds above 0.45, medium is above 0.25 and large is below that.

From testing this seems to be a good fit, because we get a few different context, but we also seem them again quite often. For the most part this is a good fit, we could extend the number of different possibilities for the amount of players and pot odds, but that has not been a big problem up until now.

[5] `http://mathworld.wolfram.com/StandardDeviation.html`

5 Player behavior

In this section we will describe each player phase in a bit more detail than what we have seen up until now. As we said in section 1.1 some of the code is shared between all phases, but for the betting decision everything is up to the implementation.

5.1 Phase 1

The first phase player is the simplest player of any of the phases. It has no information about the start of the game or opponent modelling and bases most of its decisions on chance and power rating. Before the flop this player has nothing to base its decision on so it just uses chance to decide whether to bet, call or fold. Most of the time the player should just try to call to get to the flop, but sometimes it will bet and sometimes it will fold with the folding as the least possible outcome.

The intuition here is much the same as for a novice poker player which has little grasp of the game. Before the flop it is hard to estimate what sort of chances one has and thus we just leave it up to chance, but with a preference for calling. This is done because none of the other player will likely bet a lot at this stage and a flop will give us so much more information to continue so seeing it is a smart choice to make.

If this phase 1 player get to see the flop be do things a bit different. Now we can use our power rating as a basis for telling us if the hand is worth something. We still utilize random chance, but this time we lay more emphasis on the value of the hand. What we do is to see what kind of hand we have, if it is a good hand our value will be higher than random chance and we do not fold, but we retain this randomness to sometime bluff and throwaway somewhat good cards in order to not play static all the time. Then we see if we have a very good hand, if so we bet and if not we just call to see if we can improve or have the other players fold.

TODO: Add results

5.2 Phase 2

The second phase player is quite a bit more sophisticated than the first one. This is due to the new information available in this phase. In phase 2 the player has access to both roll-out simulation(2) and hand strength(3) calculation. With all this information the player can take a much more qualified decision about what to do.

With the roll-out simulation the player has some statistics to tell whether or not the hole-cards is worth continuing with before the flop. This information enables the player to know whether or not he/she should continue to see the flop or fold the cards. This enables strategies based on folding before the flop because the player has an idea of how strong the cards are or have the potential to become later in the game.

The hand strength calculation is used after the flop. Because there are now cards on the table we need to know if we have gotten anything better than just the two hole-cards. The hand strength calculation helps with this because it calculates, as mentioned above, the strength of the current hole-cards and

community cards compared to all the other possible hole-cards. This helps the player decide if he/she should continue to play or, if there is a high chance anyone else has better cards, fold.

In our code, before the flop, we consult the roll-out simulation statistics to see if our cards are any good. If the statistics is over a certain threshold we decide to bet. Should the statistics not be good enough to bet we check to see if the hole-cards have the potential to become something after the flop and call. If this is not the case we fold as this means the hole-cards have little potential for improvement.

After the flop we move over to combining hand strength, a bit of chance and a configurable aggressiveness. We have again used chance to incorporate some element of bluffing. If the random dice throw is large enough and the hand strength is weak we fold, but if the dice throw had been smaller we could potentially call with a weak hand just to try our luck. We also calculate whether or not the winnings are good enough that we should play, if we might lose more than we could win, we fold. For the decision the hand strength is what we rely on most because it tells us the most about our hand[6]. Although with a more aggressive player this can be changed.

For the betting decision we look at what is on the table and what we can potentially win. We calculate an expected amount and bet according to this. If the player has a negative amount of chips we only put in the amount needed to call so we don't end up with chip counts in the billions.

TODO: Add results

5.3 Phase 3

In this phase we finally have all the bells and whistles to play poker. Combining roll-out simulation, hand strength calculation and opponent modelling. Phase 3 does not have the big advantage that phase 2 had over phase 1, but it has the potential to become better than phase 1 and 2.

The big change is of course opponent modelling(4) which enables this player to take a more informed decision about what the other players might be doing.

As with phase 2 we just check the statistics when it comes to the pre flop betting round. Since very little information is available to us and since using the opponent modelling would be quite uncertain at this stage there is not much else to do. After the flop however we can start using our arsenal to our advantage.

What we then do is to look at our hand strength compared to the what the opponent modeller thinks is the best other player at the table. We then combine the average strength from the opponent modeller with the standard deviation[7]. To incorporate some aggressiveness we can apply different levels of confidence in the weight of the standard deviation, meaning a conservative player might like 95% confidence, but a more aggressive player might be good with only the average or maybe even less.

TODO: Add results

[6] At the time of writing we are currently using the simple calculation because we had some problems with the more advanced. The advanced should be working now, but we have to little time to test and so we have used the simple one

[7] `http://mathworld.wolfram.com/StandardDeviation.html`

6 Results

Below is the results from the different phases of play. All results are 10 000 simulations and for phase 3 there are and additional 1000 to see if the opponent modeller could learn anything. Each player start with a 1000 chips and blinds are at 10 and 20 and increase.

6.1 Phase 1

Phase 1, Player 0, Aggressiveness: 0.0	chip count: 63960280
Phase 1, Player 1, Aggressiveness: 1.0	chip count: -21112167
Phase 1, Player 2, Aggressiveness: 2.0	chip count: -11393325
Phase 1, Player 3, Aggressiveness: -1.0	chip count: -31450803

Phase 1, Player 0, Aggressiveness: 0.0	chip count: 62938731
Phase 1, Player 1, Aggressiveness: 1.0	chip count: -20831844
Phase 1, Player 2, Aggressiveness: 2.0	chip count: -11176039
Phase 1, Player 3, Aggressiveness: -1.0	chip count: -30926860

Phase 1, Player 0, Aggressiveness: 0.0	chip count: 63160394
Phase 1, Player 1, Aggressiveness: 1.0	chip count: -20883887
Phase 1, Player 2, Aggressiveness: 2.0	chip count: -11252784
Phase 1, Player 3, Aggressiveness: -1.0	chip count: -31019735

In the above tables we have some a