IT3105 Project 1

Texas Hold 'EM

Alexandre Balon-Perin

Norvegian University of Science and Technology - NTNU

Academic Year 2011-2012

Table of Contents

Introduction

In this project of artificial intelligence programming, we have been asked to implement a simulator for the famous poker game Texas Hold 'Em. The difficulty of Poker remains in the imperfect information that a player has about the current game and the stochasticity due to the fact that the cards are chosen randomly for each game.

The goal of the project is to develop artificial agents capable of making decision regarding the different poker actions (call, bet and fold). The project will be divided in three phases. In each phase the artificial agent will acquire new tools that will allow it to analyse the game and its opponents strategy in order to choose the best possible action.

The complexity of the game makes it difficult for a computer to compete with the best poker players in the world. For example, the concept of bluffing is not so evident to implement when you want it to be smart. Obviously, the computer could just raise on weak hand but against good opponent this strategy will not last forever. Only the computational power and the huge amount of memory give the intelligent agent an advantage against low to average level players. Experienced opponent are harder to trick and quite a bit of cleverness will be necessary before winning against those experts.

The report will go through the code structure of each phase. An overview of the betting decision's logic will follow. Afterwards, the opponent model will be analysed. Eventually, the results will be shown and a discussion on the effectiveness of the strategies used by the agents will take place.

Code Structure

General

The programming language used in this project is Python. The code is separated in different packages that group classes working togheter to achieve the same task. The different packages are: game, main, modeler, simulation, strength and utilities. The concepts of heritage and polymorphism are used to model the different agent types as shown on the following figure.

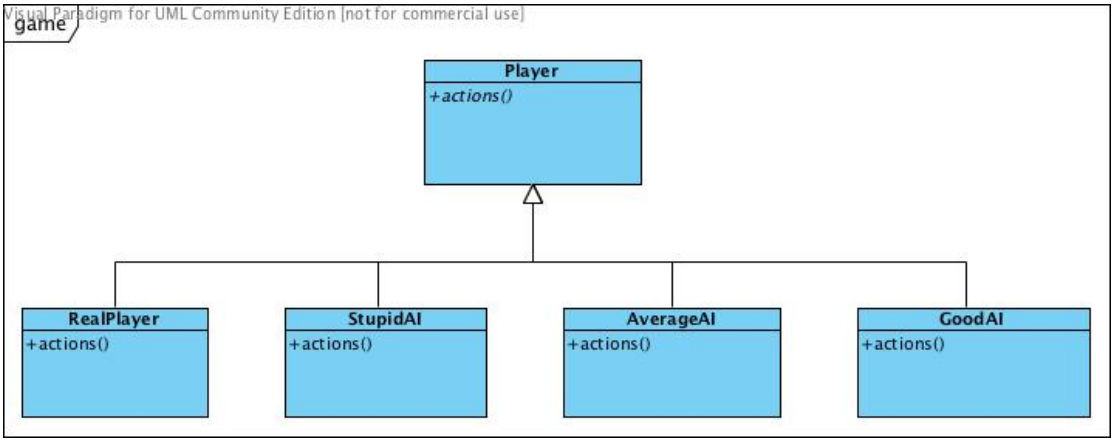


Figure 1: Agent types

The package "utilities " contains various tools such as a FileHandler and a Decorator. The use of the Decorator facilitates the transition between a run with a simulation and a normal run. A class Displayer is in charge of the messages display and can be turned off easily when running a simulation thanks to the use of decorators.

Phase I

In this phase, the focus will be on the game itself and the classes RealPlayer and StupidAI. The main class is Poker_Manager that coordinates the entire game. The Player class is in charge of all the actions of the players and has two subclasses to model the real players and the computers. The CardDeck class is used to deal cards and to convert the powers to rankings. Finally, the Evaluator is used to calculate the powers of the hand to determine the winner(s).

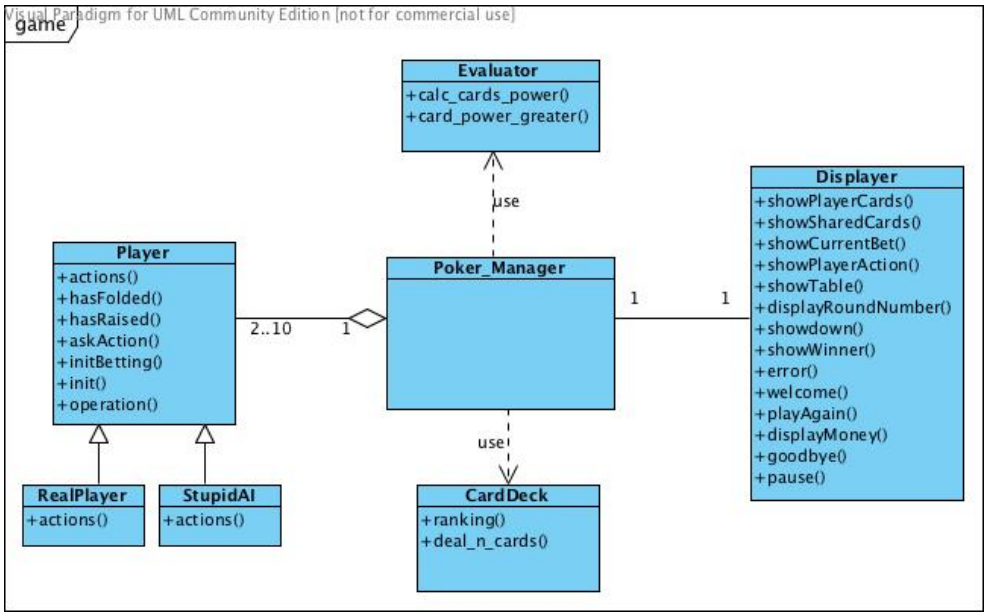
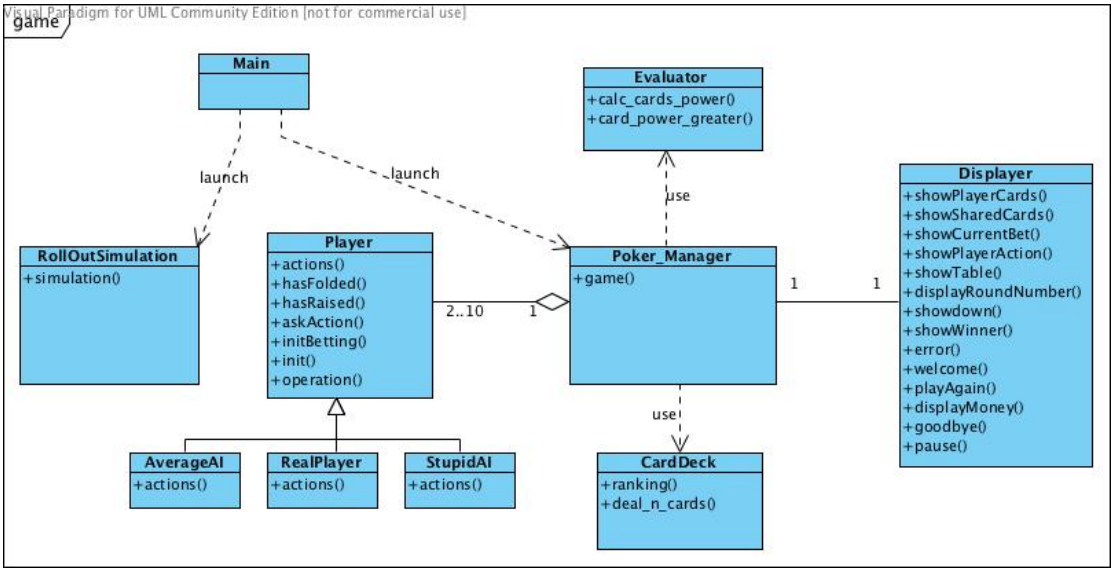


Figure 2: Dependencies of Poker_Manager

NB:
The list of methods in each class is not exhaustive and the parameters of the functions have been left out for simplicity.

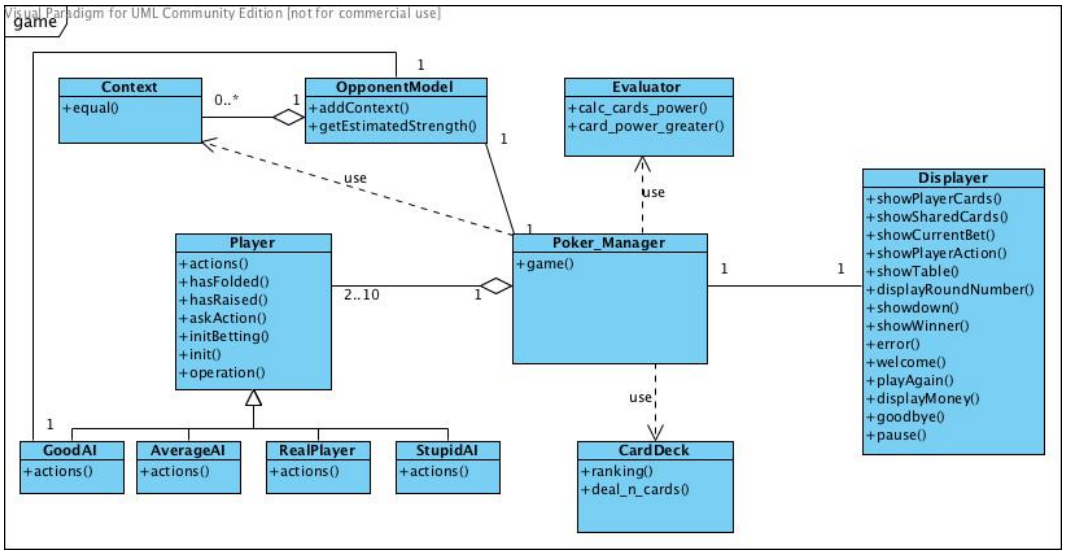
Phase II

The important classes are AverageAI and the one from the package "simulation". AverageAI is a new subclass added to the subclasses of Player to model a new type of computer a little bit more intelligent than the computers from phase I. This new type of computer use rollout probabilities in the preflop and the hand strength during the rest of the game. The class RollOutSimulation is responsible for the creation of the table associating equivalence classes and probability to win. The main (which is not exactly a class but is represented as one in the following diagram) launches the simulation() first and then the game(). The simulation can be removed thanks to the decorator.



Phase III

Eventually, the third phase adds a new subclass GoodAI to the subclasses of Player. The rollout simulation was already done in the second phase and must not be computed again. A class OpponentModel is in charge of the developpment of the opponents' model. It contains a list of objects Context. The list is built online at showdown and is then used during the game by GoodAI.



Betting Decision's Logic

General

The difference between aggressive players and conservative ones is the same for all phases. The thresholds used to trigger the possible actions are lower for the aggressive players than for the conservative ones.

Phase I

The betting decision is somewhat trivial in this phase. Before the flop, the agent has 40% of chance to raise, 40% of chance to call and 20% of chance to fold. After the flop, the agent assesses the power rating of its hand and decides which action to take according to thresholds. To keep it simple, only the first number of the power rating is taken into account. This is not very accurate but, even with a more complex logic, the result would be quite close. The power rating does not consider the number of players left in the game.

Phase II

In the preflop round, the agent uses the probabilities gathered from the rollout simulation to decide if its hand is good enough to continue. The agent's hand is compared to the equivalence classes present in the table of probabilities obtained from the rollout simulation. Once a matching equivalence class is found the probability value is extracted and compared to a threshold that varies with the betting round. After the flop, the decision is made according to the hand strength. The agent makes a simulation on the fly to estimate its chances to win against all possible other hands left in the deck using the same shared cards for itself and its opponent. The result of the hand strength is also compared to thresholds in order to decide which action to take. The rollout probability and the hand strength give a good estimation of the winning chances but they take into account neither the game context nor the opponent behavior. Besides, the potential of a hand is disregarded. A hand with high probability to improve is estimated weak if at the moment its strength is low.

Phase III

Last but not least, agents from phase III use the most complex betting decision's logic from the three phases. The logic is similar as in phase II but information about opponents is added in order to counter bluffing or conservative behavior. Concerning the aggressiveness of the player, two factors come into play. First, the thresholds are lower for aggressive players as said before. Second, conservative computers will raise only if their hand strength is higher than the estimated hand strength of all other players remaining in the game. Similarly, they will call if their hand strength is equal to the highest strength. On the other hand, aggressive computers will raise only if their hand strength is higher than any other hand and will call if it is equal to any other hand.

Opponent Model

Context

The variables chosen for the context are the following:

- Betting round's number
- Player's number
- Number of raises
- Pot odds

To have equal contexts, the value of the three first variables must have the same value and the absolute value of the difference between the two pot odds must be below 0.05. The pot odds was then removed because it was misleading as the other type of agents do not use this information to choose their actions. Obviously this information is relevant when playing against more complex agent.

The context is an object containing all the variables cited above plus the number of the player and its action as well as the player's cards if available and the shared cards.

General

The class OpponentModel has an attribute called context_table, which is a list of all the contexts gathered so far. This list is updated at showdown with all the contexts from the previous game except the contexts from players who folded. If the context is already in the table, the hand strength of the opponent is calculated and averaged with the existing value in the table.

Before acting, the agent looks in the table for contexts equals to the previous contexts of all the players who played before. If a matching context is found, the estimated strength of the opponent is extracted. The estimated strength of all remaining opponents is then used to decide the next action.

Results

Phase I

1000 runs with the following players:
All players use strategy from phase I
Players 0-3 are aggressive players
Players 4-7 are conservative players

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 46952
Player 1 's money: 32387
Player 2 's money: 29944
Player 3 's money: 39719
Player 4 's money: -28655
Player 5 's money: -34370
Player 6 's money: -22684
Player 7 's money: -23293
************************************************************

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 45799
Player 1 's money: 41242
Player 2 's money: 50955
Player 3 's money: 36159
Player 4 's money: -33352
Player 5 's money: -20608
Player 6 's money: -38163
Player 7 's money: -42032
************************************************************

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 42111
Player 1 's money: 41443
Player 2 's money: 46989
Player 3 's money: 48401
Player 4 's money: -32330
Player 5 's money: -31375
Player 6 's money: -34713
Player 7 's money: -40526
************************************************************
```

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 50934
Player 1 's money: 57523
Player 2 's money: 32931
Player 3 's money: 33778
Player 4 's money: -33455
Player 5 's money: -34015
Player 6 's money: -37141
Player 7 's money: -30555
************************************************************

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 39804
Player 1 's money: 55758
Player 2 's money: 44815
Player 3 's money: 26622
Player 4 's money: -21758
Player 5 's money: -37324
Player 6 's money: -35184
Player 7 's money: -32733
************************************************************
```

Discussion

The difference between the two strategies is quite evident. The aggressive strategy seems to be the best. At each additional round, the thresholds increase and the possibility to call or raise for the consevative player becomes too limited. On the other hand, the aggressive players call even on weak hands and win the most of the time without reaching the showdown or at least there is no more players with strategy I when reaching showdown.

Phase II

1000 runs with the following players:

```
Players 0-3 use strategy from phase II
      Player 0 and 1 are aggressive
      Player 2 and 3 are conservative
Players 4-7 use strategy from phase I
      Player 4 and 5 are aggressive
      Player 6 and 7 are conservative

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 63074
Player 1 's money: 95059
Player 2 's money: 6057
Player 3 's money: 18917
Player 4 's money: -25677
Player 5 's money: -17441
Player 6 's money: -48060
Player 7 's money: -51929
************************************************************

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 91955
Player 1 's money: 70643
Player 2 's money: 3681
Player 3 's money: 12676
Player 4 's money: -18770
Player 5 's money: -10907
Player 6 's money: -50770
Player 7 's money: -58508
************************************************************

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 124142
Player 1 's money: 134398
Player 2 's money: 25291
Player 3 's money: 18018
Player 4 's money: -59023
Player 5 's money: -52873
Player 6 's money: -76390
Player 7 's money: -73563
************************************************************
```

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 87244
Player 1 's money: 128537
Player 2 's money: 10433
Player 3 's money: 40777
Player 4 's money: -37248
Player 5 's money: -30971
Player 6 's money: -89255
Player 7 's money: -69517
************************************************************

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 106627
Player 1 's money: 119025
Player 2 's money: 27443
Player 3 's money: 16343
Player 4 's money: -30923
Player 5 's money: -43700
Player 6 's money: -87280
Player 7 's money: -67535
************************************************************
```

Discussion

In this phase, the difference in strategy is obvious. As expected, all the computers with strategy I have negative amount of money whereas all computers playing with the strategy II have positive money. This result is normal for several reasons. First, in the preflop, the players using strategy II use the probabilities from the rollout table to decide if their hand is good enough to continue whereas the players using strategy I decide randomly. A player using strategy I might fold on the preflop with a pair of aces as holecards. Second, after the preflop, strategy II assess its hand strength for each betting round whereas strategy I calculates the power rating of its hand. The power rating contains too limited information to be relevant. Knowing that you have 2 pairs is not enough. Players using strategy I might raise with two pairs in the river phase even when there are still ten players remaining in the game. Another interesting fact is that the aggressive players earn more money than the conservative one. This is probably because they force all other players to fold even when they have good hands. The last players standing at showdown are the aggressive one which have weak hands and share the pot between themself. However, the conservative players from strategy II are playing better than the aggressive players from strategy I. This is because they use more relevant information than the players from strategy I. As a consequence, they are not really affected by the aggressive game of those players.

Phase III

1000 runs with the following players:

```
Players 0-3 use strategy from phase III
      Player 0 and 1 are aggressive
      Player 2 and 3 are conservative
Players 4-5 use strategy from phase II
      Player 4 is aggressive
      Player 5 is conservative

************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: -4895
Player 1 's money: -4257
Player 2 's money: 19264
Player 3 's money: 22583
Player 4 's money: -11024
Player 5 's money: 8329
************************************************************
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 4259
Player 1 's money: -17775
Player 2 's money: 10684
Player 3 's money: 28356
Player 4 's money: -1842
Player 5 's money: 6318
************************************************************
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: -3049
Player 1 's money: 10997
Player 2 's money: 18025
Player 3 's money: 21370
Player 4 's money: -16039
Player 5 's money: -1304
************************************************************
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: -11949
Player 1 's money: -4343
Player 2 's money: 18944
Player 3 's money: 11749
Player 4 's money: 5857
Player 5 's money: 9742
************************************************************
```

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: -10378
Player 1 's money: -9145
Player 2 's money: 18323
Player 3 's money: 9444
Player 4 's money: 3929
Player 5 's money: 17827
************************************************************
```

Discussion

The ressource tables 1-4 display a clear result. The agents using an aggressive strategy lose more often than the conservative ones. In phase II, those agents were the most successful ones. This result shows that the opponent modeler works as expected. Aggressive players are ruined with both strategy II and III. Besides, even when playing conservative strategy III wins against strategy II. The difference between the conservative players is not as big as with the aggresive ones but this only means that conservative players are more difficult to model than aggressive ones. With more runs, the players using strategy III would probably take over all players using strategy II.

The last table is the exception. The aggressive players are still losing but the conservative player using strategy II obtains better result than one of the conservative players using strategy III. As said before, the conservative players give probably a harder time to the opponents' modeler. A big problem with the opponent modeler is that if the game does not go until showdown no information is gathered. As a consequence, if only conservative players are in the game, there will not be enough showdown.

One player of each phase

Player 0 uses strategy from phase III and is aggressive
Player 1 uses strategy from phase II and is conservative
Player 2 uses strategy from phase I and is conservative

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 8618
Player 1 's money: 1977
Player 2 's money: 4405
************************************************************
Player 0 must be aggressive in order to push the others to showdown.
If not, too little information is gathered for the opponents' model.
```

Player 0 uses strategy from phase III and is aggressive
Player 1 uses strategy from phase II and is conservative
Player 2 uses strategy from phase I and is aggressive

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 9578
Player 1 's money: 178
Player 2 's money: 5244
************************************************************
```

Playground

After the completion of phase III, a new type of agent has been added called ExpertAI. The purpose of this class is to try to improve the type III agents with new methods from the paper [1].

The agent is now able to calculate its hand potential. The hand potential gives an idea of the possible improvement of the hand as the turn and river are played. This calculation is made only when the flop is dealt. The use of the hand potential ensures that the agent does not fold with a hand that has a high probability of improvement. For example, (3♣, 5♥) with a flop (2♣, 4♣, 10♥) is a pretty good hand but the agent will fold if the hand potential is not computed.

A second addition is the possibility to bluff. Bluffing diversifies the strategy and increases the possibilities of winning as long as it is done only on rare occasions. The agent calculates the pot odds to make sure that the bluff is worth the winnings.

1000 runs with the following players:
Player 0 is an aggressive expert
Player 1 is a conservative expert
Player 2 is a aggresive good player (phase3)

```
************************************************************
MONEY
------------------------------------------------------------
Player 0 's money: 5042
Player 1 's money: 5703
Player 2 's money: 4255
************************************************************
```

[1] Darse Billings, Denis Papp, Jonathan Schaeffer, and Duane Szafron. 1998.

Possible Improvements

- Modify the agent strategy over time might give hard time to an opponent modeler
- Modify the threshold dynamically to adapt to the opponent's strategy
- General improvement of the opponent model to obtain more accurate and faster results
- Record the number of time each opponent has folded, raised or called in the preflop and flop
- Use different amount of money for the betting

Conclusion

In conclusion, the opponent modeler gives good results as long as the player behavior is strongly aggressive or conservative. When playing against an opponent which strategy lies in between those two behaviors, the opponent modeler is not as good as expected. In this case the players using strategy II and III obtain similar results. If all players are conservative, the lack of showdown leads to a worse phase III player.

How to get started

- Launch the Main.py
- Select the number of players
- Choose the player type (Phase I, II, III or a real player)
- Choose their aggressiveness (except for real players)
- To turn off the display, go to the class Displayer and assign True to the variable "simulation"