Poker Project - Team #07

Arielyte Tsen Chung Ming, Devarajan Preethi, James Pang Mun Wai, Lee Yi Wei Joel, Yip Seng Yuen
National University of Singapore
chungming.tsen@u.nus.edu, e0203237@u.nus.edu, jamespang@nus.edu.sg, lywjoel@u.nus.edu, yip@u.nus.edu

1 Introduction

The game of poker has long been a field of interest for artificial intelligence (AI) researchers. This is not surprising as developing an AI poker agent capable of competing with humans at the highest level is challenging, given that poker is a game that forces players to make decisions with incomplete and imperfect information. This paper talks about the steps we took in designing an AI Poker Player agent in Heads Up Limit Texas Hold'em. An agent subjected to reinforcement learning must learn to interact with the environment in order to maximize its reward. Reinforcement learning was chosen as a way to get the best possible action based on the current state.

The problem with Limit Texas Hold'em is that there is no input or output training data sets available for the agent to gather feedback from. Hence, the only way an agent can learn is by trial and error based on feedback from its own actions and experiences. A reward function where actions that led to positive outcomes are rewarded, and actions that led to negative outcomes are punished is created, with the goal of maximizing its reward. The design of a reward function good enough for the agent to master poker is the basis for this project. In a game of poker, the agent does not have a complete model of the environment, nor does it know the states that its actions will lead to. Thus, a Q-learning agent, which compares the expected utilities of the remaining choices without the need to know the outcomes nor the model of the environment [Russell and Norvig, 2014], is our agent design of choice as it fits the scenario of a poker game.

2 Pre-Flop Hand Strength

To begin with, we wanted to decide how the agent will determine the strength of its hand, just from the pre-flop round. Due to the limited amount of information that we can manipulate during the pre-flop round, the strength of a hand, HS, at pre-flop is determined by the probability of winning, $Pr(Win)$, based on the prior knowledge of the two hole cards currently in hand.

To estimate HS, we simulated 10,000 games for each possible combination of starting hand. The total combination of starting hands is 13 + $\binom{13}{2}$ = 91, obtained by summing up the combination of pocket pairs and the combination of unsuited

Card 2 \ Card 1	2	3	4	5	6	7	8	9	T	J	Q	K	A
2	0.4986	0.2853	0.2921	0.3138	0.3334	0.3345	0.3702	0.3878	0.4033	0.427	0.47	0.503	0.5371
3		0.531	0.32	0.3365	0.3566	0.3624	0.3728	0.3941	0.4283	0.4583	0.4762	0.5164	0.559
4			0.5686	0.3556	0.3867	0.3834	0.3926	0.4091	0.4433	0.4712	0.504	0.5295	0.5595
5				0.5973	0.3972	0.4139	0.4093	0.4302	0.4456	0.4831	0.5121	0.5386	0.5728
6					0.6327	0.4298	0.4442	0.4552	0.4715	0.4913	0.521	0.5593	0.5883
7						0.6737	0.4608	0.4792	0.4847	0.5081	0.5341	0.5669	0.5929
8							0.6944	0.4863	0.5063	0.5276	0.5403	0.5749	0.6107
9								0.7235	0.5283	0.5423	0.5696	0.5931	0.6254
T									0.7621	0.5734	0.5924	0.6125	0.6406
J										0.7804	0.593	0.6283	0.6407
Q											0.8019	0.6368	0.6612
K												0.8373	0.6697
A													0.8556

Figure 1: Estimated winning probabilities with starting hands

hands. We then calculate $Pr(Win)$ using the formula:

$$Pr(Win) = \frac{\#\, of\ wins}{Total\ games\ played}$$

Where the total games played is fixed at 10,000 in our case. The results of our simulation is as seen in Figure 1.

At every pre-flop street, the agent chooses an action to be performed based on the starting hand's estimated probabilities of winning, which are:

$$\left.\begin{array}{l} raise\colon Pr(Win) > k \\ \\ call\colon Pr(Win) > j \\ \\ fold\colon Pr(Win) \leq j \end{array}\right\} \quad \text{where } 0 < j < k \leq 1$$

3 Post-Flop Rounds

In subsequent post-flop rounds, reinforcement learning is employed. Specifically, the technique of Q-learning, which was described in the Introduction, was used to train our agent.

In Q-learning, we maintain a table of states and actions as inputs for the agent. Each permutation generates an expected reward. An illustration of it can be seen in Table 1.

After each training round, the state of each permutation is updated based on the formula

$$Q\left(s', a'\right) = Q\left(s, a\right) + \sum_{0}^{t} \lambda^{t} R_{t}\left(s, a\right)$$

Where t is the number of turns between the move and the terminal node, R_t is the reward for round t, and λ is the discount factor.

State	Action		
	$fold$	$call$	$raise$
1			
2			
3			
...			
S			

Table 1: Q-learning Training

$$R_t = \begin{cases} +\frac{pot}{2} & \text{if win,} \\ -\frac{pot}{2} & \text{otherwise} \end{cases}$$

λ = discount rate for future reward

4 ε-Greedy Algorithm

At every turn, with a probability of ε, a random action is chosen to be performed, otherwise an action with the best expected reward is chosen.

The value of ε will slowly decrease as the agent acquires more training. A relatively large initial value ensures that all possible paths will be explored before the agent settles into a sub-optimal pattern. The $State$ and $Action$ for the ε-Greedy Algorithm is

$$State = \{HS, ARS, S, P, \#OR, \#SR, OPS\}$$

$$Actions = \{fold, call, raise\}$$

Where

HS refers to the current Hand Strength of a given player. More details can be found in Section 4.1.

ARS refers to the current Average Rank Strength of a given player. More details can be found in Section 4.2.

S refers to the current Street that the game is in. More details can be found in Section 4.3.

P refers to the Pot. More details can be found in Section 4.4.

#OR refers to the Opponent's Raise. More details can be found in Section 4.5.

#SR refers to the agent's Self Raise. More details can be found in Section 4.6.

OPS refers to the Opponent Playing Style. More details can be found in Section 4.7.

4.1 Expected Hand Strength

The Expected Hand Strength, EHS is the probability of the current hand of a given player winning if the game reaches a showdown. It factors in all possible combinations of the opponents' hands, the remaining hidden board cards, and performing a comparison between the agent's hand and the hands in the enumeration to see which is better. The quality of the hand is then measured based on the number of times the hand turns out to be better. For our ε-Greedy Algorithm, the EHS is grouped in increments of 0.01.

Given n number of opponents, the remaining cards, Rem is given by:

$$Rem = [\alpha \backslash \beta]^5$$

Where α is the set of all cards in the deck, and β is the set of all hole cards of a particular player. The formula to calculate the rank of each hand, $Rank(h)$ is:

$$Rank(h) = max(\forall x \in [\beta \cup \Omega]^5 : s(x))$$

Where Ω is the set of community cards. Having Rem and $Rank(h)$, it is now possible to calculate HS through the formulas below:

$$Ahead(h) = \#\{\forall x \in Rem : s(x) > Rank(h)\}$$

$$Tied(h) = \#\{\forall x \in Rem : s(x) = Rank(h)\}$$

$$Behind(h) = \#\{\forall x \in Rem : s(x) < Rank(h)\}$$

Thus, the EHS for player h against n number of opponents is given by:

$$EHS(h) = \left(\frac{Ahead(h) + \frac{Tied(h)}{2}}{Ahead(h) + Tied(h) + Behind(h)} \right)^n$$

The above formulas used to derive the EHS was referenced from Teofilo *et al.* [2013a]. Note that the EHS may be used at any round of the game. However, the number of iterations needed to compute the EHS for a single hand at the pre-flop street is very high. This issue is addressed in the next section using the Average Rank Strength technique.

4.2 Average Rank Strength

A technique developed by Teofilo *et al.* [2013b] called Average Rank Strength (ARS), is used to improve the efficiency of EHS. ARS consists of using the hand score to estimate the future outcome of the match, without having to generate all card combinations. This is simply done by storing the average EHS of a hand for each score in three lookup tables, one for Flop, one for Turn and one for River. Since we are not considering suits, the number of possible scores is reduced. The pre-computation of a given score is as follows:

$$ARS_{n,5}(C_1, C_2) = \left(\frac{\sum_i X_i \in [\alpha \backslash \beta]^5 : Rank(X_i \cup \{C_1, C_2\})}{\#[\alpha \backslash \beta]^5} \right)^n$$

Where n is the number of opponents and X_i is a distinct subset of size 5 of the deck except the pocket cards.

Thus, compared to EHS, ARS had a three orders of magnitude faster response time when querying the lookup table, while also doing so with negligible error [Teofilo *et al.*, 2013b].

To find out the estimated EHS for the three post-flop streets, we simulated 500,000 rounds each. Additionally, for each round of simulation, we ran the Monte Carlo sampling algorithm to obtain the EHS. The results are displayed below: Our reasoning behind such a high number of simulation is. so that we can sample as many combinations as possible. And in doing so, achieve a more accurate estimate.

Street	Number of Combinations
Flop	4569
Turn	9452
River	13987

Table 2: Result of *ARS* Simulation

4.3 Street

Street, *S*, refers to the start of each new round. For our ε-Greedy Algorithm, we only evaluate the three Post-Flop streets, i.e. Flop, Turn and River.

4.4 Pot

Pot, *P*, refers to the total amount in the player's bet. For our ε-Greedy Algorithm, *P* is grouped by the number of big blinds.

4.5 Opponent Raise

Opponent Raise, *OR*, refers to the number of raises by the opponent per street. For our ε-Greedy Algorithm, *OR* is grouped into nine different groups, i.e. from 0 to 8.

4.6 Self Raise

Self Raise, *SR*, refers to the number of raises by the agent per street. For our ε-Greedy Algorithm, *SR* is grouped into nine different groups, i.e. from 0 to 8.

4.7 Opponent Playing Styles

According to Rupeneite [2010], the playing style of an opponent can be classified into four categories. Each style is distinct, in that it describes the opponent's frequency of play and how the player bets. The four categories of playing styles are Loose/Passive, Loose/Aggressive, Tight/Passive and Tight/Aggressive. A brief description of each style is shown in Table 3 below:

Playing Styles	Description
Tight	Plays few hands and often folds.
Loose	Plays multiple and varied hands.
Aggressive	Bets and raises a lot, almost always never checking or call.
Passive	Usually checks and call, unlikely to take the lead.

Table 3: Description of Playing Styles

The Aggressive Factor, AF, is used to classify a player as either aggressive or passive. The formula for AF is as follows:

$$AF = \frac{\#\, raises}{\#\, calls}$$

The Player Tightness, PT, is used to classify a player as either tight or loose. Based on research by Rupeneite [2010], a threshold of 0.28 is used. The formula for PT is as follows:

$$PT = \frac{\#\, folds}{\#\, games}$$

Later, a classification process is conducted to classify the opponent's style of play into four categories, as seen in Table 4.

	$\mathbf{AF \leq 1}$	$\mathbf{AF > 1}$
$\mathbf{PT \geq 0.28}$	Loose Passive	Loose Aggressive
$\mathbf{PT < 0.28}$	Tight Passive	Tight Aggressive

Table 4: Style of Play Classification

4.8 Word Processing Software

As detailed below, IJCAI has prepared and made available a set of LaTeX macros and a Microsoft Word template for use in formatting your paper. If you are using some other word processing software (such as WordPerfect, etc.), please follow the format instructions given below and ensure that your final paper looks as much like this sample as possible.

Note that I did not edit the word document, and it still contains the original IJCAI formatting instructions. Please ignore those!

5 Style and Format

LaTeX and Word style files that implement these instructions can be retrieved electronically. (See Appendix A for instructions on how to obtain these files.)

5.1 Layout

Print manuscripts two columns to a page, in the manner in which these instructions are printed. The exact dimensions for pages are:

- left and right margins: .75″
- column width: 3.375″
- gap between columns: .25″
- top margin—first page: 1.375″
- top margin—other pages: .75″
- bottom margin: 1.25″
- column height—first page: 6.625″
- column height—other pages: 9″

All measurements assume an 8-1/2″ × 11″ page size. For A4-size paper, use the given top and left margins, column width, height, and gap, and modify the bottom and right margins as necessary.

5.2 Format of Electronic Manuscript

For the production of the electronic manuscript, you must use Adobe's *Portable Document Format* (PDF). A PDF file can be generated, for instance, on Unix systems using `ps2pdf` or on Windows systems using Adobe's Distiller. There is also a website with free software and conversion services: `http://www.ps2pdf.com/`. For reasons of uniformity, use of Adobe's *Times Roman* font is strongly suggested. In LaTeX2e, this is accomplished by putting

```
\usepackage{times}
```

in the preamble.[1]

Additionally, it is of utmost importance to specify the American **letter** format (corresponding to 8-1/2$''$ $\times$ 11$''$) when formatting the paper. When working with `dvips`, for instance, one should specify `-t letter`.

5.3 Title and Author Information

Center the title on the entire width of the page in a 14-point bold font. The title should be capitalized using Title Case. Below it, center author name(s) in a 12-point bold font. On the following line(s) place the affiliations, each affiliation on its own line using a 12-point regular font. Matching between authors and affiliations can be done using superindices. Additionally, a comma-separated email addresses list using a 12-point regular font is also allowed. Credit to a sponsoring agency can appear on the first page as a footnote.

5.4 Text

The main body of the text immediately follows the abstract. Use 10-point type in a clear, readable font with 1-point leading (10 on 11).

Indent when starting a new paragraph, except after major headings.

5.5 Headings and Sections

When necessary, headings should be used to separate major sections of your paper. (These instructions use many headings to demonstrate their appearance; your paper should have fewer headings.). All headings should be capitalized using Title Case.

Section Headings

Print section headings in 12-point bold type in the style shown in these instructions. Leave a blank space of approximately 10 points above and 4 points below section headings. Number sections with arabic numerals.

Subsection Headings

Print subsection headings in 11-point bold type. Leave a blank space of approximately 8 points above and 3 points below subsection headings. Number subsections with the section number and the subsection number (in arabic numerals) separated by a period.

Subsubsection Headings

Print subsubsection headings in 10-point bold type. Leave a blank space of approximately 6 points above subsubsection headings. Do not number subsubsections.

Special Sections

You may include an unnumbered acknowledgments section, including acknowledgments of help from colleagues.

Any appendices directly follow the text and look like sections, except that they are numbered with capital letters instead of arabic numerals.

The references section is headed "References," printed in the same style as a section heading but without a number [Russell and Norvig, 2014]. A sample list of references is given at the end of these instructions [Rupeneite, 2010]. Use a consistent format for references, such as that provided by BibTeX. The reference list should not include unpublished work [Teofilo *et al.*, 2013a].

5.6 Citations

Citations within the text should include the author's last name and the year of publication, for example [Teofilo *et al.*, 2013b]. Append lowercase letters to the year in cases of ambiguity. Treat multiple authors as in the following examples: [Abelson *et al.*, 1985] or [Baumgartner *et al.*, 2001] (for more than two authors) and [Brachman and Schmolze, 1985] (for two authors). If the author portion of a citation is obvious, omit it, e.g., Nebel [2000]. Collapse multiple citations as follows: [Gottlob *et al.*, 2002; Levesque, 1984a].

5.7 Footnotes

Place footnotes at the bottom of the page in a 9-point font. Refer to them with superscript numbers.[2] Separate them from the text by a short line.[3] Avoid footnotes as much as possible; they interrupt the flow of the text.

6 Illustrations

Place all illustrations (figures, drawings, tables, and photographs) throughout the paper at the places where they are first discussed, rather than at the end of the paper. If placed at the bottom or top of a page, illustrations may run across both columns.

Illustrations must be rendered electronically or scanned and placed directly in your document. All illustrations should be in black and white, as color illustrations may cause problems. Line weights should be 1/2-point or thicker. Avoid screens and superimposing type on patterns as these effects may not reproduce well.

Number illustrations sequentially. Use references of the following form: Figure 1, Table 2, etc. Place illustration numbers and captions under illustrations. Leave a margin of 1/4-inch around the area covered by the illustration and caption. Use 9-point type for captions, labels, and other text in illustrations.

Acknowledgments

The preparation of this report would not have been possible without the help of Dr. Yair Zick and Arka Maity, National University of Singapore, School of Computing.

A LaTeX and Word Style Files

The LaTeX and Word style files are available on the IJCAI–18 website, `http://www.ijcai-18.org/`. These style files implement the formatting instructions in this document.

The LaTeX files are `ijcai18.sty` and `ijcai18.tex`, and the BibTeX files are `named.bst` and `ijcai18.bib`. The LaTeX style file is for version 2e of LaTeX, and the BibTeX style file is for version 0.99c of BibTeX (*not* version 0.98i).

[1] You may want also to use the package `latexsym`, which defines all symbols known from the old LaTeX version.

[2] This is how your footnotes should appear.

[3] Note the line separating these footnotes from the text.

The `ijcai18.sty` file is the same as the `ijcai07.sty` file used for IJCAI–07.

The Microsoft Word style file consists of a single file, `ijcai18.doc`. This template is the same as the one used for IJCAI–07.

These Microsoft Word and LaTeX files contain the source of the present document and may serve as a formatting sample.

References

[Abelson *et al.*, 1985] Harold Abelson, Gerald Jay Sussman, and Julie Sussman. *Structure and Interpretation of Computer Programs*. MIT Press, Cambridge, Massachusetts, 1985.

[Baumgartner *et al.*, 2001] Robert Baumgartner, Georg Gottlob, and Sergio Flesca. Visual information extraction with Lixto. In *Proceedings of the 27th International Conference on Very Large Databases*, pages 119–128, Rome, Italy, September 2001. Morgan Kaufmann.

[Brachman and Schmolze, 1985] Ronald J. Brachman and James G. Schmolze. An overview of the KL-ONE knowledge representation system. *Cognitive Science*, 9(2):171–216, April–June 1985.

[Gottlob *et al.*, 2002] Georg Gottlob, Nicola Leone, and Francesco Scarcello. Hypertree decompositions and tractable queries. *Journal of Computer and System Sciences*, 64(3):579–627, May 2002.

[Gottlob, 1992] Georg Gottlob. Complexity results for nonmonotonic logics. *Journal of Logic and Computation*, 2(3):397–425, June 1992.

[Levesque, 1984a] Hector J. Levesque. Foundations of a functional approach to knowledge representation. *Artificial Intelligence*, 23(2):155–212, July 1984.

[Levesque, 1984b] Hector J. Levesque. A logic of implicit and explicit belief. In *Proceedings of the Fourth National Conference on Artificial Intelligence*, pages 198–202, Austin, Texas, August 1984. American Association for Artificial Intelligence.

[Nebel, 2000] Bernhard Nebel. On the compilability and expressive power of propositional planning formalisms. *Journal of Artificial Intelligence Research*, 12:271–315, 2000.

[Rupeneite, 2010] Annija Rupeneite. *Building Poker Agent Using Reinforcement Learning with Neural Networks*. SCITEPRESS, 2010.

[Russell and Norvig, 2014] Stuart J. Russell and Peter Norvig. *Artificial Intelligence: A Modern Approach*. Prentice Hall, 2014.

[Teofilo *et al.*, 2013a] Luis Filipe Teofilo, Luis Paulo Reis, and Henrique Lopes Cardoso. Computing card probabilities in texas hold'em. In *Proceedings of the 8th Iberian Conference on Information Systems and Technologies*, pages 988–993, Lisbon, Portugal, June 2013. Canada Institute for Scientific and Technical Information.

[Teofilo *et al.*, 2013b] Luis Filipe Teofilo, Luis Paulo Reis, and Henrique Lopes Cardoso. Speeding-up poker game abstraction computation: Average rank strength. In *Proceedings of the 27th Association for the Advancement of Artificial Intelligence Workshop*, pages 59–64, Bellevue, Washington, USA, July 2013. Association for the Advancement of Artificial Intelligence.