CASE STUDY

Using a simple poker game to introduce mixed strategies in game theory

Salim Neil Khan, School of Computing, Electronics and Mathematics, Coventry University, Coventry, UK. Email: aa9803@coventry.ac.uk

Abstract

This article presents a classroom activity that introduces the idea of mixed strategies in game theory. In the activity, the students investigate a simple two-player poker game with the aim of determining the best strategies for both players. It comprises a mixture of hands-on playing, analysis using game theory, and simulation using computers.

Keywords: Practical Activity, Game Theory, Poker, Simulation.

1. Introduction

At Coventry University, the mathematics undergraduate students, as part of a Problem Solving first year module, were taught some introductory game theory over two sessions. The first session involved the students engaging in a Prisoners' Dilemma activity. This is a scenario first formulated by Tucker (1950), and it is briefly outlined in Section 2 of this article. In that activity, the students met concepts such as maximin, saddle point, and pure strategy; ideas in game theory originated by Von Neumann and Morgenstern (1944).

This article concentrates mainly on the students' second session. This built on their knowledge from the first session, by introducing a situation where there was no saddle point, and a mixed strategy was required. Section 3 describes a simple poker game that the students played, and measures empirically their resulting profits. Section 4 analyses the game theoretically and produces recommended playing strategies. In Section 5, the recommended strategies are implemented by the students and a comparison of profits is made. In Section 6, the theory is generalised so that recommended strategies can be obtained for different versions of the game. Then, in Section 7, a computer simulation is used to test the effectiveness of the theory.

2. Prisoners' Dilemma

This is an often-used example as an introduction to game theory. There are many different versions, but they all roughly follow along similar lines as detailed below:

- Arry and Butch are gangsters. The police have arrested them, but do not possess enough information for a conviction. Following the separation of the two men, the police offer both a similar deal;
- If one testifies against his partner, and the other remains silent, the former receives a 1 year sentence and the latter receives 10 years;
- If both remain silent, both are sentenced to only 3 years in jail for a minor charge;
- If each testifies against the other, each receives a 6 year sentence;
- Each prisoner must choose either to betray or remain silent. Whilst choosing, they do not know the other's decision. What should they do?

The actions and outcomes can be displayed in a 'payoff matrix' (Table 1):

Table 1: Payoff matrix for Prisoners' Dilemma

Prison sentences (Arry, Butch)		Butch	
		Betray	Silent
Arry	Betray	(-6,-6)	(-1,-10)
	Silent	(-10,-1)	(-3,-3)

- Arry's line of reasoning could be as follows. If Butch betrays, then Arry should also betray to maximise his own outcome. However, if Butch is silent, again Arry should betray to maximise his own outcome. Therefore Arry should betray whatever Butch does. This is called a maximin strategy as in essence he maximises his minimum gain;
- Similarly, Butch could follow the same line of reasoning, and conclude that he himself should betray;
- The result would then be that both would betray and consequently both end up with a 6 year prison sentence. The cell (-6,-6) is called the saddle point as neither prisoner can improve their outcome by changing their mind if the other doesn't change;
- However, if they both agree to cooperate and both keep silent, they would both be better off with only a 3 year sentence. But, this does rely on trust as one of them could renege on the deal to personally benefit - hence the dilemma.

3. Simple Poker Game

The students were split into groups of three with each group given a standard pack of playing cards and 40 casino style chips. Two of the group played the game (Player 1 and Player 2), and the third dealt the cards and recorded the number of hands played.

Player 1 (P1) and Player 2 (P2) were given 20 chips each and the rules were as follows:

- P1 and P2 both put 1 chip into the pot.

- P1 is given a card. He can either:
  - Fold – in which case P2 wins the pot; or
  - Bet 1 (i.e. put another 1 chip into the pot).

- If P1 bets, P2 can then either:
  - Fold – in which case Player 1 wins the pot; or
  - Call (i.e. put another 1 chip into the pot).

- If P2 calls, P1 shows his card:
  - If P1 has 9 or higher, P1 wins the pot;
  - If P1 has 8 or lower, P2 wins the pot.

- Shuffle the cards and play again several times.

Figure 1: Cards and chips for simple poker game

Before they started playing, the students were asked which player they thought was most likely to be in profit after several hands. The majority who responded suggested that it would be P2, as there were only 6 winning cards (9 to ace) for P1 but 7 winning cards (8 to 2) for P2.

The game was played for three or four minutes, after which the profit/loss was recorded for P1 along with the number of hands played. The process was repeated twice more with the students swapping roles within each group so that each had a turn at each role.

There were 33 students in the class. Some of the results are shown in Table 2.

Table 2: Students' P1 Profits

Student No.	P1 Profit	Number of Hands
1	4	20
2	-8	16
3	-3	22
etc.	…	…
Total	-8	455

The total number of hands played was 455 and the P1 total profit was -8 chips. Thus the mean profit per hand was -8/455 = 0.018 i.e. P1 lost on average 0.02 chips per hand.

The students were asked which strategies they employed, and it was evident that those P1's who ended up in profit were mainly those who employed 'bluffing' i.e. Betting on a losing card (an 8 or lower) in the hope that P2 would Fold.

4. Game Theory Analysis

Let Strategy 1 (S1) for P1 be ‘Fold if he has a losing card i.e. 8 or lower, Bet if he has a winning card i.e. 9 or Higher’ (F if L, B if W).

Let Strategy 2 (S2) for P1 be ‘Bet if he has a losing card i.e. 8 or lower, Bet if he has a winning card i.e. 9 or Higher’ (B if L, B if W).

Let S1 for P2 be Fold, and S2 for P2 be Call.

Thus for each hand played there are 4 possible strategy pairs for the two players – (S1, S1), (S1, S2), (S1, S1), (S2, S1), and (S2, S2). (For each pair, the first number represents the strategy that P1 employs and the second number represents the strategy that P2 employs).

Consider (S1, S1). This means that P1 will only Bet if he has a winning card and P2 will always Fold if P1 has Bet. Indeed this was a common pair of strategies amongst the students, particularly at the beginning. Cautious P1’s didn’t Bet if they had a losing card, and cautious P2’s assumed that if P1 had Bet then he must have a winning card. Here the expected profit for P1 can be calculated as follows:

P1 Folds if he has 8 or lower, so

$$P(P1\ Folds) = \frac{7}{13}, P(P1\ Bets) = \frac{6}{13}.$$

If P1 Folds, his profit is -1. If P1 Bets, his profit is 1 (because P2 will Fold). Hence

$$E(P1\ Profit\ if\ (S1, S1)) = -1 \times \frac{7}{13} + 1 \times \frac{6}{13} = -0.077.$$

Now consider (S2, S1). This means that P1 Bets when he has a winning card, and also Bets when he has a losing card. In the classroom session, as the playing progressed many P1’s who were employing S1 tended to realise that they were Folding a lot of hands, and thus started to bluff some hands – this was either through reasoning that they might have more chance of obtaining a profit, or through getting a bit bored with sticking with the same approach.

$$E(P1\ Profit\ if\ (S2, S1)) = 1$$

because here we are assuming that P1 is always Betting and P2 is always Folding.

For (S2, S2), P1 Bets whatever his card, and P2 Calls. In the activity, as the playing progressed it was found that P2 Called more often, as he became more aware that P1 was bluffing.

$$E(P1\ Profit\ if\ (S2, S2)) = -2 \times \frac{7}{13} + 2 \times \frac{6}{13} = -0.154$$

because P1 wins 6/13 of the time.

Finally, consider (S1, S2), i.e. P1 only Bets if he has a winning card, and P2 Calls. Here,

$$P(P1\ Folds) = \frac{7}{13}, P(P1\ Bets) = \frac{6}{13}.$$

If P1 Folds, his profit is -1. If P1 Bets, his profit is 2 (because P2 has Called). Hence

$$E(P1\ Profit\ if\ (S1, S2)) = -1 \times \frac{7}{13} + 2 \times \frac{6}{13} = 0.385.$$

These profits are displayed in a payoff matrix in Table 3:

Table 3: Payoff matrix for Poker Game

P1 Expected Profit			P2	
			S1	S2
			F	C
P1	S1	F if L, B if W	-0.077	0.385
	S2	B if L, B if W	1	-0.154

Key:
S1 – Strategy 1,
S2 – Strategy 2,
F – Fold, B - Bet, C – Call,
L – has losing card,
W – has winning card.

There are some differences between this payoff matrix and the one from the Prisoners' Dilemma (Table 1). Firstly, it can be seen that there is only one value in each cell here, as opposed to two in Table 1. This is because the poker game is a 'zero-sum game'. That means that a gain for one player corresponds to a loss of equal magnitude for the other player. Hence for (S1, S1) the payoff for player 2 is 0.077, for (S1, S2) it is -0.385, and so on. As these values are assumed, it is not necessary to display them in the matrix. Also, as there is some uncertainty in the outcomes in terms of the probabilities of the card being dealt, the entries are expected values. These can appropriately be interpreted as the expected profit in the long run, because the game is repeated. The Prisoners' Dilemma however would often be perceived as a game that would only be played once if in a realistic context.

Being a zero-sum game, there is no possibility of cooperation helping both players, as a change in strategy could not increase the pay-out for both players. Moreover, there is no saddle point here i.e. there is no cell in the matrix with a value representing the optimal strategy pairing for the two players. Hence we would say that there is no 'pure' strategy solution.

Nevertheless we can obtain a 'mixed' strategy (Maschler, Solan and Zamir 2013) solution i.e. to maximise his expected payoff, a player can use S1 some of the time and S2 some of the time. P1's optimal mixed strategy solution is to use S2 a proportion, x, of the time so that his expected payoff is the same whatever P2 does, and P2's optimal mixed strategy solution is to use S2 a proportion, y, of the time so that his expected payoff is the same whatever P1 does. This will result in a mixed strategy equilibrium.

First, consider P1's mixed strategy. Using the values from Table 3;

$$E(P1\ Profit\ if\ P2\ Folds) = -0.077(1-x) + 1x,$$

$$E(P1\ Profit\ if\ P2\ Calls) = 0.385(1-x) - 0.154x.$$

Equating these to obtain the optimal mixed strategy for P1:

$$-0.077(1-x) + x = 0.385(1-x) - 0.154x$$

$$x = 0.29.$$

This means that P1 should employ S2 29% of the time and S1 71% of the time.

For P2's optimal mixed strategy; again using the values from Table 3;

$$E(P2\ Profit\ if\ P1\ uses\ S1) = -0.077(1-y) + 0.385y,$$

$$E(P2\ Profit\ if\ P1\ uses\ S2) = 1(1-y) - 0.154y$$

Equating these gives:

$$-0.077(1-y) + 0.385y = 1 - y - 0.154y$$

$$y = 0.67.$$

Thus, P2 should use S2, 67% of the time, and S1, 33% of the time.

In Table 4 these proportions are added to the payoff matrix.

Table 4: Optimal Mixed Strategies and Payoffs in Poker Game

<table>
<tr><td colspan="4" rowspan="4">P1 Expected Profit</td><td colspan="2">P2</td></tr>
<tr><td>0.33</td><td>0.67</td></tr>
<tr><td>S1</td><td>S2</td></tr>
<tr><td>F</td><td>C</td></tr>
<tr><td rowspan="2">P1</td><td>0.71</td><td>S1</td><td>F if L, B if W</td><td>-0.077</td><td>0.385</td></tr>
<tr><td>0.29</td><td>S2</td><td>B if L, B if W</td><td>1</td><td>-0.154</td></tr>
</table>

If these mixed strategies are used, then

$$E(P1\ Profit) = -0.077(0.71)0.33 + 0.385(0.71)0.67 + 1(0.29)0.33 - 0.154(0.29)0.67 = 0.23$$

So, the optimal mixed strategies are that P1 should bluff 29% of the time, and P2 should Call 67% of the time. This would result in a 23% profit for P1. If one of the players deviates from this, his expected profit would decrease.

5. Playing the Game again

The students were then asked to play the game again several times, utilising the optimal strategies that they had just learnt. In order to randomly bluff approximately 29% of the time, P1 would glance at their watch or phone and bluff if the last digit of the seconds of the time was a 3, 6 or 9. Similarly P2 would Call if the first digit was 0-3. (Alternatively, appropriate sided dice could be used but these would not be as easy to conceal from the opponent and may inhibit the flow of the game.)

The result for the class was a mean gain of 0.36 chips for P1, a lot better than the 0.02 average loss that P1 had the first time they played the game, highlighted in Section 3. The result was also considerably higher than the 0.23 predicted by the theory in Section 4. This is likely attributable to the fact that, when questioned afterwards, many of the P2's admitted that they hadn't stuck to the optimal 67% for Calling. They had taken it upon themselves to try and 'spot' when P1 was bluffing, resulting in a higher Call rate.

6. Generalising the Theory

The students were then asked to find optimal strategies in the general case where, in the poker game, the probability that P1 gets dealt a winning card is p. Using the same reasoning as in section 4,

$$E(P1\ Profit\ if\ (S1,S1)) = -1(1-p) + 1p = 2p - 1;$$

$$E(P1\ Profit\ if\ (S2,S1)) = 1;$$

$$E(P1\ Profit\ if\ (S2,S2)) = -2(1-p) + 2p = 4p - 2;$$

$$E(P1\ Profit\ if\ (S1,S2)) = -1(1-p) + 2p = 3p - 1.$$

These expected profits are displayed in a payoff matrix in Table 5:

Table 5: Payoff Matrix where P(Winning Card) = p

P1 Expected Profit				P2	
				1-y	y
				S1	S2
				F	C
P1	1-x	S1	F if L, B if W	2p-1	3p-1
	x	S2	B if L, B if W	1	4p-2

Finding the equilibrium point, P1 should bluff a proportion, x, of the time, where

$$(2p-1)(1-x) + x = (3p-1)(1-x) - (4p-2)x$$

$$x = \frac{1}{3}\left(\frac{p}{1-p}\right) \tag{1}$$

P1 should Call a proportion, y, of the time, where

$$(2p-1)(1-y) + (3p-1)y = 1 - y + (4p-2)y$$

$$y = \frac{2}{3}.$$

Interestingly, P2 should Call 2/3 of the time, whatever the probability of P1 receiving a winning card. This results in

$$E(P1\ Profit) = (2p-1)\frac{1}{3}\left[1-\frac{1}{3}\left(\frac{p}{1-p}\right)\right] + x + (3p-1)\frac{2}{3}\left[1-\frac{1}{3}\left(\frac{p}{1-p}\right)\right] + (1)\frac{1}{3}\left[\frac{1}{3}\left(\frac{p}{1-p}\right)\right] + (4p-2)\frac{2}{3}\left[\frac{1}{3}\left(\frac{p}{1-p}\right)\right]$$

$$= \frac{8p-3}{3}$$

Using the value of $p = 6/13$ from the poker game, from Eq. (1) the proportion of time that P1 should bluff,

$$x = \frac{\left(\frac{1}{3}\right)\left(\frac{6}{13}\right)}{1-\frac{6}{13}} = \frac{2}{7} = 0.29,$$

and from Eq. (2),

$$E(P1\ Profit) = \frac{8\left(\frac{6}{13}\right) - 3}{3} = 0.23$$

confirming the values found in Section 4.

The students then found the optimal strategy for P1 if the rules of the game were changed slightly, so that P1's winning cards were reduced to 10 or higher:

Here $p = 5/13$, so

$$x = \frac{1}{3}\left(\frac{5/_{13}}{1 - 5/_{13}}\right) = \frac{5}{24} = 0.21$$

and

$$E(P1\ Profit) = \frac{8 \times \frac{6}{13} - 3}{3} = 0.03.$$

Hence P1 should bluff 21% of the time and will end up on average with a profit of 3%.

Many students were surprised or sceptical of this finding, in that P1 only has five winning cards to eight losing cards, but will still end up in profit through strategic game play.

7. Computer Simulation

To convince the sceptics, the students were asked to test the optimal strategy for the '10 or higher' rules. This time, rather than deal cards, to save time they played the game on a computer simulation created in Excel. An example of a screenshot is shown in Figure 2.

The spreadsheet operated as follows: The student clicked on the Deal tab and a card was randomly dealt. In the example in Figure 2, this was an 8. The program then recommended to Bet or Fold according to the optimal strategy i.e. it randomly recommended 21% of the time to Bet on a card 9 or lower. The student then clicked on the Bet tab or Fold tab. If the student selected Bet, they then clicked on the P2 Decision tab. The program would then Call 67% of the time and Fold 33% of the time. The number of chips won or lost for that hand were then displayed. The students then repeated the process, playing as many times as they wished, and the program displayed their mean profit and the percentage of the time that the student had bluffed.

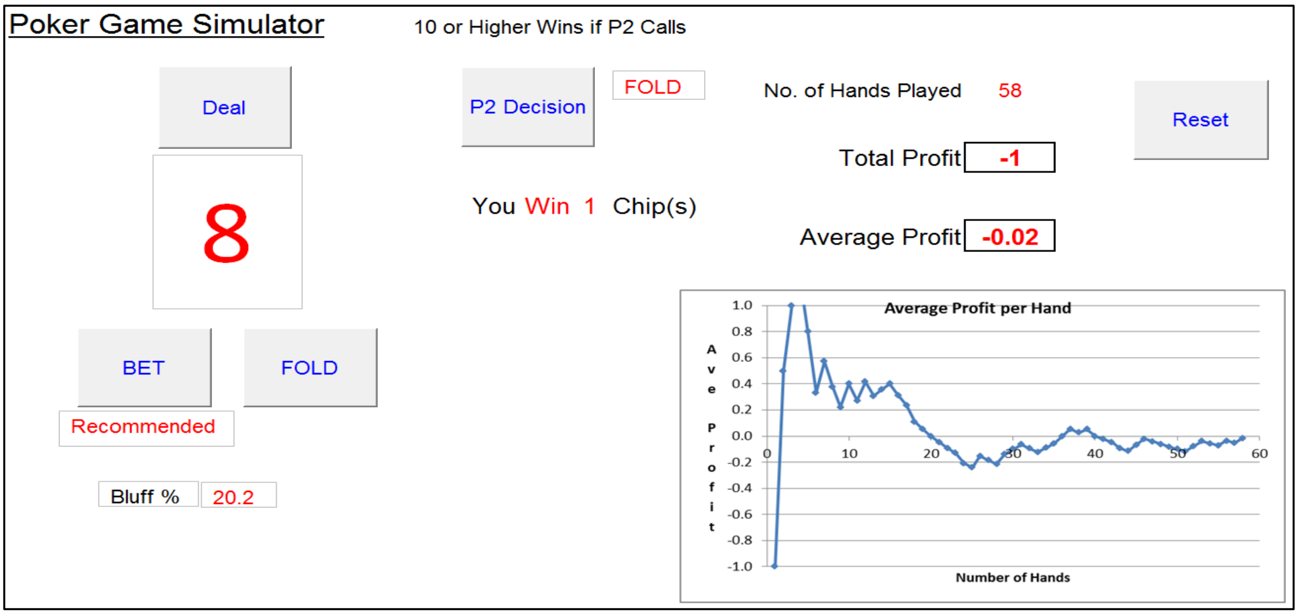

Figure 2: Simulation of ’10 or Higher’ Poker game

In Figure 2, the student’s bluff percentage was 20.2 which was close to the recommended 21%. This suggested that the student had likely followed the automatically generated recommendations by the program. Consequently the mean profit here of -0.02 was close to the expected profit of 0.03 calculated in Section 6. Obtaining results from the other students in the class, those who had deviated from the recommended action generally ended up with a lower average profit. This helped to convince the sceptical students of the validity of the theory.

8. Conclusions

Through participating in a hands-on activity, the students investigated and applied mixed strategies in game theory to see them work in practice. There is much research that activity led learning is a successful learning mechanism, for example see Freeman et al. (2014). Hence the activity outlined in this article could be considered a useful tool when teaching an introduction to mixed strategies in game theory.

9. References

Freeman, S., Eddy, S.L., McDonough, M., Smith, M.K., Okoroafor, N., Jordt, H. and Wenderoth, M.P., 2014. Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, 111(23), pp. 8410-8415. Available at http://www.pnas.org/content/111/23/8410.short [Accessed 19 June 2017].

Maschler, M., Solan, E. and Zamir, S., 2013. *Game Theory*. Translated from Hebrew by Z. Hellman and edited by M. Borns. New York: Cambridge University Press. pp. 144-218

Tucker, A.W., 1950. A two-person dilemma (unpublished notes): In: E. Rasmusen, ed. 2001. *Readings in Games and Information*. Oxford: Blackwell. pp. 7–8.

Von Neumann, J. and Morgenstern, O., 1944. *Theory of Games and Economic Behaviour*. Princeton, NJ: Princeton University Press.