Summary

Buying a house can be a time consuming process. After the buyer places its first bid the seller may offer a counter bid, which might be followed by a counter bid from the buyer. After a certain time period this can result in an agreement when the bid is accepted.

These multiple rounds of bids can be described mathematically, and the scientific discipline dealing with them is called game theory. In this discipline the focus lies on the analysis of strategic choices of the players involved.

When we go back to the example of the buyer and the seller we see that the action that both players can play is to place a bid. If the bid is not accepted by the other player the game continues and a counterbid has to be placed. However, if the bid is accepted the game ends and the payoff is paid from the buyer to the seller. When an agreement is made the payoff is equal to the final bid, otherwise, if the players do not come to a settlement the game may go on forever, in that case the payoff equals 0.

A strategy for a player is a decision rule that prescribes a mixed action, that is, a probability distribution on the possible actions. The combination of these strategies yields some payoff for all players. We assume that all players seek a strategy that maximizes their payoff.

In reality the bids of the buyer and seller tend to come closer to each other and to the value of the house as the game proceeds. In the end a price agreement is made someplace in the midst. In this paper we assume however that the strategies of the players are stationary and are played simultaneous. That is, they do not depend on the history of the game and in each round all players play their action at the same time.

We define the value of the game as v if, for every error-term $\varepsilon > 0$, the buyer of the house has a strategy such that he can guarantee never to pay more than v plus ε, and if in the same time the seller has a strategy that guarantees him never to get a payoff less than v minus ε. These strategies are called $\varepsilon-$optimal.

If a strategy guarantees the exact value of the game, we call it 0-optimal. In these games, it turns out that at least one player always has a 0-optimal strategy. The second player however might not have a strategy that exactly

guarantees v, but he still has an ε-optimal strategy for every $\varepsilon > 0$.

In this case we see that the buyer can guarantee never to pay more than the value of the game by rejecting bids higher than the true value of the house. The seller on the other hand wants that the buyer accepts the bid otherwise nothing will happen and he will not receive any payoff. He can not guarantee the true value of the house because the buyer would have no incentive to agree with a bid equal to v. Yet he can get arbitrarily close to the value. In this case by bidding v - ε, which the buyer will accept. Therefore we see that in this case the seller only has an $\varepsilon-$optimal strategy.

This paper elaborates on repeated games, such as the multiple bidding in the example. It shows how to find the value and (ε-)optimal strategies by using theoretical and practical techniques. Also we will introduce an algorithm that approximates the value for games with a finite, yet long horizon and prove that in the limit the algorithm always converges to the value of the game.

2.1 The game

We have a two-player zero-sum game G in which both players have two or more different actions. The game is given by a matrix in which the rows represent different actions for player 1 and the columns for player 2. The cells in the matrix represent the payoff of the game and can take two different forms: *termination** or *continue*.

   The game is played as follows: Both players play their actions simultaneously. If we hit a cell with a * the game terminates and the corresponding payment is paid from player 2 to player 1. If the game hits a cell with continue, in the matrices displayed as an empty cell, we play the game once again. This might go on an infinite amount of times, in that case the payoff for both players will be 0.

   For the moment we will assume that terminating payoffs are non-negative. Important to notice is that due to this the incentives for both players differ. Player 1 would like to terminate the game at a terminating cell with payoff > 0 which gives him a positive payoff. Player 2 however is indifferent between ending at a cell 0* or continuing the game until infinity.

   This game belongs to a special class of stochastic games, namely the recursive games. A recursive game has the property that all continuing cells have a payoff of 0 [Everett (1957)]. This assumption is essential in our game, hereafter we will show an example with non-zero payoff in continuing cells to illustrate its importance.

   We take the horizon of the game to be infinity since this is a very good approximation for the game with a finite, yet long horizon. We will devote a subsection to eleborate on this.

2.1.1 Strategies

A strategy for player $i \in \{1, 2\}$ is a decision rule that prescribes a mixed action, that is a probability distribution on the actions of player i depending on the history of the game. A strategy is called stationairy if it prescribes the same mixed action regardless past plays. A strategy is called pure if one of the actions is played with probability 1.

   It is known that if a player uses a stationary strategy then the other player has a pure best response. This has the convenient implication that we do not need to investigate all possible mixed response strategies.

2.1.2 Value and optimality

Before we can dig deeper into optimal strategies we have to talk about another important concept for our game: the value v of the game. We will

use the following definition for the value in the further course of this paper:

v is the value of the game if for every $\varepsilon > 0$ player 1 has a strategy σ_1 that ensures him an expected payoff of $v - \varepsilon$ and player 2 has a strategy σ_2 that ensures him not to pay more than the expected $v + \varepsilon$.

If the value exists, then these strategies are called ε-optimal. If a strategy guarantees the exact value of the game we call it 0-optimal. However, it is not always the case that the players have a 0-optimal strategy.

[Everett (1957)] and [Thuisman and Vrieze (1992)] proved that for any two-player zero-sum recursive games the value exists, and there are stationary ε-optimal strategies for both players. That is why we only need to look at stationary strategies. In addition it follows from this result that if $v \leq 0$ or $v \geq 0$ respectively player 1 or 2 has 0-optimal strategies.

This is a very robust result. Since we look at nonnegative payoffs we see that player 1 always has ε-optimal strategies and player 2 0-optimal strategies. This has the quite intuitive reason that the task of player 1 is harder than that of player 2. As we mentioned before, player 2 would be equally happy with termination at a cell 0* as going through the game an infinite amount of times, whereas player 1 only wants to hit a terminating cell with positive payoff.

We can encounter these ε-optimal strategies because the payoffs behave as a discontinuous function. In this example we see that the payoff function for player 1 is discontinuous: Let player 1 play top with probability p and bottom with probability $1 - p$.

We see that if player 1 plays $p = 1$ the payoff is 0, but if he puts some positive probability on bottom $0 \leq p < 1$ the payoff is 1 since eventually the game will hit 1*.

Typically in the cases where we cannot guarantee an exact payoff, but can get arbitrarily close to the value we refuge to ε-optimal strategies. This makes sense in several different situations [Flesch et al. (2012)]. The main point is that if we can assure stability by giving up a small part of the profit this is worth it. Especially if the profits are just approximations the small extra loss is remote compared to the measurement errors that might occur.

2.1.3   Examples

Now let us introduce some examples of the game to get a feeling how it might look like and how we can determine the value and optimal strategies.

Example one: Here we see that the value of the game is 0 since both player 1 and 2 can assure themselves a payoff of 0. The optimal strategy for player 2 is to play right with probability 1 for the entire course of the game. All strategies for player 1 are optimal since they all guarantee the value 0. So the sets of optimal strategies are $\sigma^1 = \{(p, 1-p)|p \in [0,1]\}$ and $\sigma^2 = \{(0,1)\}$.

Example two: If player 1 uses both actions top and bottom with a positive probability the game will eventually hit a cell 1*. Hence we see that the value of the game is 1. We have that all strategies for player 2 are optimal since they all guarantee him to pay 1. $\sigma^1 = \{(p, 1-p)|p \in (0,1)\}$ and $\sigma^2 = \{(q, 1-q)|q \in [0,1]\}$.

Example three: This example is slightly more involved and allows us to show that 0-optimal strategies do not always exist. Now, let us assume that player 1 plays top with probability p and bottom with probability $1-p$. Recall that since the strategy $(p, 1-p)$ is stationary, player 2 has a best response in $(q, 1-q)$ with $q = 0$ or $q = 1$.

For convenience we will seperate this example into two parts: $p = 1$ and $p < 1$.

- $p = 1$: We see that by playing this player 1 can guarantee himself a value of at least 0.

- $p < 1$: We have two possibilities: Player 2 can play either left or right. If he were to play left, $q = 1$, we will eventually hit 1* with positive probability $1-p$ and the payoff would be 1. So let us assume player 2 plays right, $q = 0$. Then the payoff for player 1 would be $p * 1 + (1-p) * 0 = p$. We see that player 1 can assure himself of a payoff of p (arbitrary close, but not equal to 1).

Therefore the value of this game is 1. We see that in this case player 1 only has an ε-optimal strategy. Every strategy of player 2 guarantees him the exact value and is 0-optimal.

2.1.4   The Big Match

A famous example of a non-recursive game in the literature is "The Big Match". In this game we disregard the assumption that continuing cells have a 0 payoff. The overall payoff will be determined by taking the average payoff of all rounds.

The Big Match was first stated by Everett in 1957, the game looks as follows[Everett (1957)].

0	1
1*	*

We have 2 continuing cells in the upper row, with payoff 0 and 1 respectively. The lower row contains two terminating cells also with payoff 1 and 0. [Blackwell and Ferguson (1968)] Blackwell and Ferguson proved that the value of the game is $1/2$ and the ε-optimal strategy for player 1 in this specific case is not stationary since it has to base his choice of action on the previous actions of player 2. We see that if we also include continuing cells with payoff other than 0 we can get non-stationary $varepsilon$-optimal strategies for the players.

2.2   Subclasses

Now let us look beyond the 2 by 2 zero-sum games that we used in our previous discussions and assume payoffs can take any number, not neccessary non-negative. Continuing cells contain a payoff of 0. The incentive for both players is still to maximize their payoffs, yet from now on player 2 could also make a profit. We see that there are some different feasible forms of the game. The overall idea is that the game consists of a combination of a normal matrix game with only terminating cells and value w and continuing cells.

What is very important in to think about is that the players should always play their optimal strategy for the subgames. If for example $w_1 > w_2 > 0$ it is not the case player 1 only has to play his optimal strategy for game 2. Without playing the optimal strategy one cannot guarantee a payoff equal to the value of the (sub)matrix. Therefore the optimal strategy for the entire game has to have the same proportion as the optimal strategies in the subgames.

In the following examples the terminating matrix game is represented as $**$. This matrix game might take any size, so the representation of $**$ in one cell does not imply that it consists of only 1 row or/and column. Now let us check the different possibilities where we omit the pure matrix game

since we know how to deal with that.

Case 1: We have some rows with only terminating cells and one or more rows with only continuing cells.

Player has the possibiltiy to choose between playing with positive probability for the matrix game with value w or playing only continuing rows and receive a payoff of 0.

Now we consider two possibilities:

1. $w \leq 0$

2. $w > 0$

1. In the first case player 1 would avoid playing the matrix game, since the value of the game is the highest payoff he can guarantee himself and it is < 0. By playing a row with only continuing cells he can always assure himself of 0 payoff, which in this case is his optimal strategy. We see that for player 2 it does not matter what strategy he chooses since he will always receive 0 payoff. The value of the game $v = 0$ and optimal strategies for the players are $\sigma_1 = continue$, $\sigma_2 = \{(q, 1 - q)|q \in [0, 1]\}$

2. In the second case player 1 can guarantee himself a positive payoff of w when playing the matrix game. He will therefore play his optimal strategy for the matrix game with positive probability. The best thing for player 2 to do is to play his optimal strategy for the matrix game to guarantee that he does not need to pay more than w. We see that the value $v = w$ and optimal strategies for the players are: $\sigma_1 = \alpha * continue + (1 - \alpha) * \sigma_{w1}$ for $0 \leq \alpha < 1$ and $\sigma_2 = \sigma_{w2}$

Ofcourse we can easily convert this game such that player 2 gets the option to choose between continue and the matrix game, this works intuitively the same way.

Case 2: We have one or more rows and columns with only continuing cells.

We see that both player 1 and 2 can assure themself a value of at least 0, by playing the continuing strategy. It follows that the value of the game $v = 0$ and optimal strategies for both players are to play continue: $\sigma_1 = \sigma_2 = continue$.

**		
	**	
		**

Case 3: Both player 1 and 2 have multiple matrix games to choose from, they don't have an option that only gives continue. Let us use the example where we have 3 seperate matrices with only terminating cells and no overlapping rows or columns to choose from with values w_1, w_2 and w_3.

We can distinguish 3 different possibilities:

1. $w_1 \geq w_2 \geq w_3 \geq 0$

2. $w_1 \leq w_2 \leq w_3 \leq 0$

3. $w_1 \leq w_2 \leq 0 \leq w_3$

1. All values of the individual matrix games are larger or equal to 0. Player2 cannot guarantee himself a payoff of 0 so he has to minimize his costs. As we see the value of w_3 is smallest so the best strategy for player 2 is to play his optimal strategy for w_3 with probability 1: $\sigma_v^2 = \sigma_{w3}^2$. Player 1 can guarantee himself of a payoff of at least w_3, but should not ignore his other possibilities. If he would play the third matrix with probability 1 it would be possible for player 2 to deviate and the game will hit continuing cells. If player 1 puts positive probability on each of his optimal strategies for the subgames the best response for player 2 is σ^2 and player 1 is guaranteed a payoff of w_3. $\sigma_v^1 = \alpha_1 * \sigma_{w1}^1 + \alpha_2 * \sigma_{w2}^1 + \alpha_3 * \sigma_{w3}^1$ where $\alpha_1, \alpha_2, \alpha_3 > 0$ and $\alpha_1 + \alpha_2 + \alpha_3 = 1$. The value of the game is equal to the value of matrix 3; $v = w_3$.

2. This case is symmetric to the first one, but now player 2 has the advantage of positive payoffs, the value of the game is again $v = w_3$.

3. In this case the matrix games may have positive payoff for both players. Let them both play the optimal strategy, or a combination of optimal strategies, of the matrix game with for them favorable payoff. Now check the example $w_3 \geq 0$, player 1 plays these rows, now we either hit the matrix game or continuing cells. The same happens for the columns that player 2 chooses, in this case they would hit w_2, w_3 or continuing cells. Since we have that none of the matrix games have similar rows and columns, we see that both players can guarantee themselves a payoff of 0.

The value of the game $v = 0$, and optimal strategies for the players are $\sigma_1 = \alpha * w_1 + (1 - \alpha) * w_2$ and $\sigma_2 = w_3$.

2.3 Value iteration

2.3.1 Introduction

As we have seen we can find the values and optimal strategies for the above three cases. These cases cover almost all possibilities we could encounter except for the following case that might lead to a problem considering the existence of 0-optimal strategies.

	**
**	**

Value iteration can be used to compute an estimation of the value and optimal strategies. It has been proven that for some types of stochastic games this gives a good approximation of the true value. We will explain the concept of value iteration by using some examples we have implemented in Mathematica and then prove that also for recursive games the method of value iteration can be used.

2.3.2 Examples

2x2 matrix

	1*
1*	0*

Let us revisit Example 3 we used in the beginning. We know the value is 1 and player 1 only has ε-optimal strategies.

The iterating process: We introduce the matrix $A(v_0)$ where we replaced the continuing cells of the original game G with an initial guess of the value, say $v_0 = 0$. We now have a matrix with only terminating cells that is displayed below.

0*	1*
1*	0*

Obviously the value of this game $v_1 = \frac{1}{2}$ and optimal strategies for player 1 and 2 are respectively $\sigma_1 = (\frac{1}{2}, \frac{1}{2})$ and $\tau_1 = (\frac{1}{2}, \frac{1}{2})$. The next step is to define $A(v_1)$ where we fill the previous value v_1 in the continuing cell. The game now looks as follows:

½*	1*
1*	0*

Again we compute the value and optimal strategies of this

matrix game: $v_2 = \frac{2}{3}$, $\sigma_2 = (\frac{2}{3}, \frac{1}{3})$ and $\tau_2 = (\frac{2}{3}, \frac{1}{3})$. The next step would be to insert $\frac{2}{3}$ etcetera. If we continue this iteration process n times, where $n \in \mathbb{N}$, we get a value of $v_n = \frac{n}{n+1}$ and $\sigma_n = \{\frac{n}{n+1}, \frac{1}{n+1}\}$ and $\tau_n = \{\frac{n}{n+1}, \frac{1}{n+1}\}$. We see that v converges to 1 so the value iteration works and we get ε-optimal strategies for large n.

3x3 matrix

Now let us consider an interesting example of a 3x3 matrix.

|		1*
	1*	0*
1*	0*	0*

This game has some resemblence with our 2x2 example. It has been proven that the value $v = 1$ and player 1 has an ε-optimal strategy of $(1 - \varepsilon - \varepsilon^2, \varepsilon, \varepsilon^2)$. We would like to check with value iteration if we get the same results, and how quickly outcomes converge to their true values, the initial value $v_0 = 0$ again. Since we have three continuing cells we fill in v_0 in all three of them.

⅓*	⅓*	1*
⅓*	1*	0*
1*	0*	0*

In the first step we have a value of $v_1 = \frac{1}{3}$ and optimal strategies of $\sigma_1 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ **and** $\tau_1 = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$**.** If we continue the iterations we get that after 58 steps the value of the game reaches the $0,9$ and continues to increase slowly to 1. It is also interesting to check is if the ratio between p_3 and p_2 goes indeed, as in the theory, to ε and ε^2. $\frac{p_3}{p_2} - p_2$ should $\to 0$ for $n \to \infty$. Since $p_2 \to 0$ we check $\frac{p_3}{p_2}$ which is represented in the right graph. We see that this ratio indeed goes to 0 as n increases.
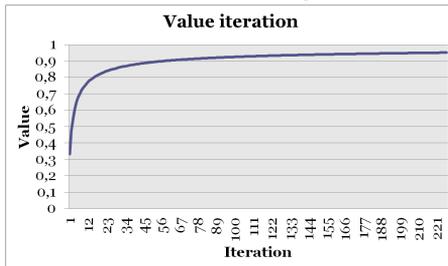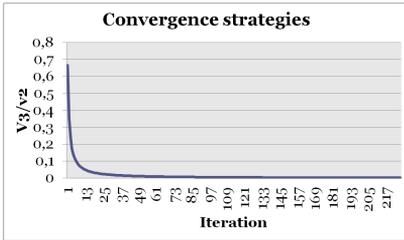
2.3.3 Proof of correctness

Now let us prove that we can use value iteration for our type of game:

We assume that all terminating payoffs ≥ 0. We set $v_1 = 0$ and v_{n+1} as the value of the matrix game $A(v_n)$. Clearly the value $A(v)$ is equal to the value v of G. We would like to prove that if $n \to \infty$ $v_n \to v$, because if this is true we know our value iteration method is invariably valid. The proof is constructed in 5 different parts:

1. v_n is increasing
   Proof by induction: We introduce for $n \in \mathbb{N}$ the statement $P(n)$: $v_n \geq v_{n-1}$. First we show that $P(1)$ is true: clearly $v_2 \geq v_1$, since all payoffs ≥ 0 and $v_1 = 0$.
   Now let $k \in \mathbb{N}$ and let us assume that $P(k)$ is true. So $v_k \geq v_{k-1}$. We know that $v_{k+1} =$ value $A(v_k)$ and since $P(k)$ is true it is more favourable to play $A(v_k)$ than $A(v_{k-1})$ for player 1. We see that $v_{k+1} \geq v_k$. According to the principle of induction, the statement $P(n)$ is true and v_n is an increasing function.

2. $v_n \leq v$
   Proof by induction: We introduce for $n \in \mathbb{N}$ the statement $P(n)$: $v_n \leq v$. First we show that $P(1)$ is true: $v_1 = 0$ and $v \geq 0$, so $v_1 \leq v$.
   Now let us assume that $P(k)$ is true so $v_k \leq v$. Then we know that value $A(v) = v \geq v_k$. $v_{k+1} =$ value $A(v_k)$, since $v_k \leq v$ and $A(v_k)$ is more favourable than $A(v_{k-1})$ for player 1 $v_{k+1} =$ value $A(v_k) \leq$ value $A(v) = v$. By the principle of induction we have that $v_n \leq v$ for all $n \in \mathbb{N}$.

3. $v_n \to \bar{v} \leq v$
   By 1, 2 we have that v_n is an increasing sequence which is bounded above by v. The Monotone Sequence Property implies that the sequence converges to some $\bar{v} \leq v$.

4. $\bar{v} = $ Value $A(\bar{v})$

We know that the sequence $v_n = $ value $A(v_{n-1}) \to \bar{v}$ for $n \to \infty$. By taking limits, we immediately get that $\bar{v} = $ value $A(\bar{v})$.

5. $v = \bar{v}$

Let x be an optimal strategy in G and let b be a best response against x in $A(\bar{v})$. Let us define $u_{A(v_i)}(x,b)$ as the expected utility in game A at time period i for the strategy combination (x,b). We can distinguish 2 cases:

- The strategy combination (x,b) only hits terminating cells.

  We know that these terminating cells are the same in $A(\bar{v})$ as in G. Since x is optimal and b is a best reply, player 1 can guarantee himself at least v and player 2 can guarantee himself at most $\bar{v}$. So value $A(\bar{v}) \geq v$ and $A(\bar{v}) \leq \bar{v}$. We see that indeed $\bar{v} = v$.

- The strategy combination (x,b) can hit continuing cells.

  We know that in G the expected payoff of playing an optimal strategy x for player 1 is at least v: $u_G(x,b) \geq v$. Now in the matrix game $A(\bar{v})$, since b is a best response for x for player 2 we know that he can guarantee himself an expected payoff of at most the value $\bar{v}$ of $A(\bar{v})$: $u_{A(\bar{v})}(x,b) \leq $ value $\bar{v}$. Proof by contradiction: let us assume that $v > \bar{v}$. Then value $A(\bar{v}) = (1-\alpha)*v + \alpha*\bar{v} > \bar{v}$, where $0 < \alpha < 1$ is the probability of immediate termination for the strategies (x,b). This contradicts 4, therefore $\bar{v} = v$ has to be true.

2.4   GAUSS

We would like to try to investigate how many different cases we encounter of 2 by 2 games where player 1 does not have an optimal strategy. The only combination that gives an "error" is the matrix form of example 3. The remaining combinations of 0^*, 1^* and continue are good in the sense that both players have a 0-optimal strategy.

We use a program that is written in GAUSS. We first define a matrix with randomly assigned cells $0^*, 1^*$ and continue (all with probability $\frac{1}{3}$). Then we check all different combinations of 2 rows and 2 columns and we verify how many of these give us a game with the form of example 3. We run this simulation a lot of times for different matrix sizes and came with the following outcomes:

We investigated the following different simulations: For a matrixgame of 5 by 5 we have done 1000 simulations. For a matrixgame of 6 by 6 we

also did 1000 simulations, but since this didn't prove sufficient we also ran 10000 simulations. The graphs belonging to our findings are displayed on the next page.

On the x-axis it is represented how many errors occur in a single matrix, and the y-axis shows in how many simulations this amount of errors has occured.

2.5 Final remarks

2.5.1 Horizon

For our theoretic models we took the horizon of the game equal to ∞. [Mertens and Neyman (1981)] showed that the value can be approached by solving for a finite horizon of n.

As we shortly stated in the beginning ∞ is a good approximation for the game with a finite but long horizon. What [Mertens and Neyman (1981)] proved is that $v_n \to v$ for $n \to \infty$.

We see that this is a good approximation, but there is another advantage namely that if we consider a finite game in the ultimate stages strange outcomes might occur. If the players are certain the game will stop at a certain point backwards induction might lead to other strategies than normally would be played. Since we assume that the horizon for our games is large enough so that we do not have to consider this it is better to calculate for an infinite amount of stages.

2.5.2 Discounted value

In our model we only considered undiscounted payoffs, this means that the utility for the players of receiving payoff now is equal to receiving it in the future. It is also possible to take a discountfactor δ into account. Some of the characteristics that we saw in this paper such as existence of v_δ and (ε-)optimal strategies also hold for discounted models. [Bewley and Kohlberg (1978)] proved that for $n \to \infty v_\beta \to v$. It is quite interesting to investigate the discounted game in further research since this can have some striking similarities with reality.
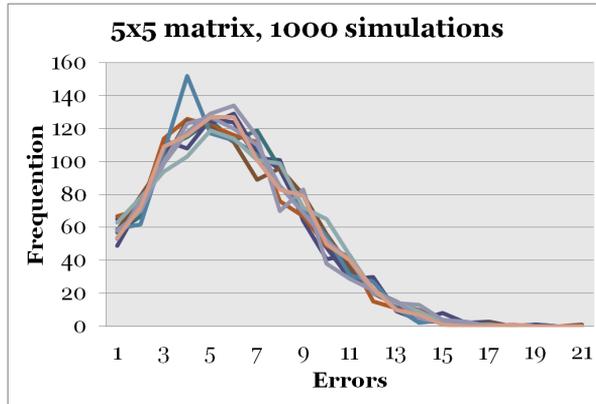
Figure 2.1: 1. 1000 simulations of 5x5
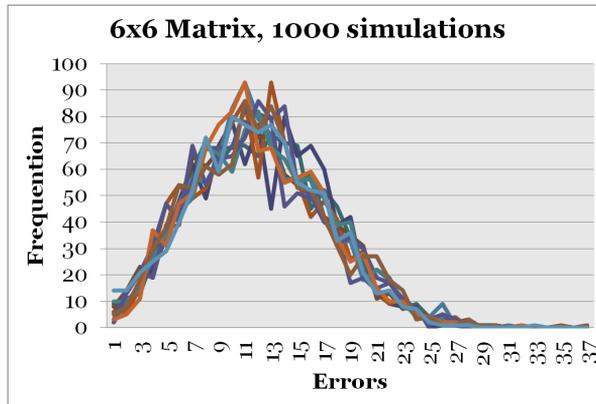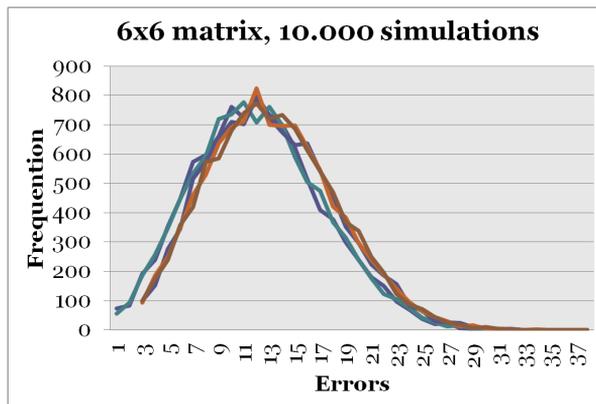

Figure 2.2: 2. 1000 simulations of 6x6

Figure 2.3: 3. 10000 simulations of 6x6

2.6 Conclusion

We have examined some different subclasses of repeated recursive games and came with theoretical solutions for them. With mathematica we examined the concept of value iterations on some examples and wrote an algorithm to show that value iteration is applicable to our model. Finally we used GAUSS to check the quantity of 2x2 games where player 1 does not have a 0-optimal strategy in matrices of different sizes.

Bibliography

[Bewley and Kohlberg (1978)] T. Bewley and E. Kohlberg. On Stochastic Games with Stationary Optimal Strategies. *Mathematics of Operations Research*, 3:104–125, 1978.

[Blackwell and Ferguson (1968)] D. Blackwell and T. S. Ferguson. The Big Match. *The Annals of Mathematical Statistics*, 39:159–163, 1968.

[Everett (1957)] H. Everett. Recursive games. *Contributions to the Theory of Games*, 3:67–78, 1957.

[Flesch et al. (2012)] J. Flesch, J. Kuipers, A. Mashiah-Yaakovi, G. Schoenmakers, E. Shmaya, E. Solan, and K. Vrieze. Subgame Perfection in Games with Infinite Horizon. 2012.

[Mertens and Neyman (1981)] J. F. Mertens and A. Neyman. Stochastic games. *International Journal of Game Theory*, 10:53–66, 1981.

[Thuisman and Vrieze (1992)] F. Thuijsman and O. J. Vrieze. Note on recursive games. *Game Theory and Economic Applications*, 389:133–145, 1992.