Multi Agent Systems

Homework Assignment 2

Dirk Hoekstra, Luisa Ebner, Philipp Lintl

September 26, 2018

2 Nash Equilibrium

2.1 Let's watch some TV shows

BBC's Split or Steal - Game *Golden Balls*

Game description:

- finite strategic game
- 2 agents = {Player 1 (P1), Player 2 (P2)}
- a finite number of possible action profiles $A_i, i = 1, 2$
- A_1 = {split (SP), steal (ST)}
- A_2 = {split (SP), steal (ST)}
- Idea/Goal: decision on how to share price money (pm)

The strategy profiles:

- $s_1 : \{SP, SP\} \rightarrow u_1(s_1) = \frac{pm}{2}$, $u_2(s_1) = \frac{pm}{2}$
- $s_2 : \{SP, ST\} \rightarrow u_1(s_2) = 0$, $u_2(s_2) = pm$

- $s_3 : \{ST, SP\} \rightarrow u_1(s_3) = pm,\ u_2(s_3) = 0$
- $s_4 : \{ST, ST\} \rightarrow u_1(s_4) = 0,\ u_2(s_4) = 0$

The payoff matrix:

P2

		SP	ST
P1	SP	$(\frac{pm}{2}, \frac{pm}{2})$	(0, pm)
	ST	(pm, 0)	(0,0)

The two players are playing a

- simulatneous (moves/choices are made at the same time)
- finite strategic (finite number of agents and action profiles)
- non-cooperative (both agents act independently and self-interested to maximize own payoff)
- symmetric
- complete (there is no private information)

game.

What are the values of R,S,T and P?

P = $\frac{pm}{2}$ S = 0 T = pm P = 0

Golden Balls has two pure, weak Nash Equilibria. The weak NE's are the strategy profiles $s_2 = s^*$ and $s_3 = s^{**}$.

$$\forall agent\, i, i = \{1, 2\}, \forall s_i \neq s_i^* : u_i(s_i^*, s_{-i}^*) \geq u_i(s_i, s_{-i}^*)$$

$$\forall agent\, i, i = \{1, 2\}, \forall s_i \neq s_i^{**} : u_i(s_i^{**}, s_{-i}^{**}) \geq u_i(s_i, s_{-i}^{**})$$

Golden Balls has no mixed Nash Eqilibria.

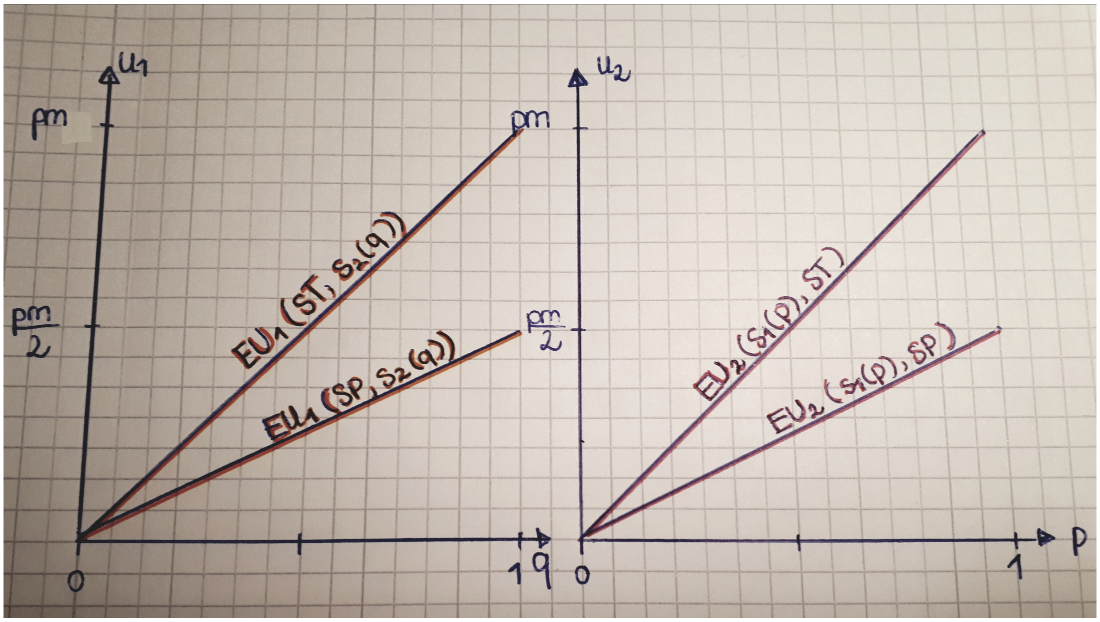
		P2	
		SP (q)	ST (1-q)
P1	SP (p)	$(\frac{pm}{2}, \frac{pm}{2})$	(0, pm)
	ST (1-p)	(pm, 0)	(0,0)

$$P1: \quad EU_1(SP, s_2(q)) = \frac{pm}{2} * q$$
$$EU_1(ST, s_2(q)) = pm * q$$

$$\frac{pm}{2} * q = pm * q$$
$$\leftrightarrow q = 0$$

$$P2: \quad EU_2(s_1(p), SP) = \frac{pm}{2} * p$$
$$EU_2(s_1(p), ST) = pm * p$$

$$\frac{pm}{2} * p = pm * p$$
$$\leftrightarrow p = 0$$



Does communication help?

As seen in the figure above, there is no mixed-strategy NE. There is no p: $0 < p < 1$ and q: $0 < q < 1$ that both players would stick to their choice if they knew about the other's choice.

As the two players communicate, they could find out more about the opponent's probabilities p or q to split or steal. However, the game is constructed in a way, that communication does NOT help. If P1 assures to play split and P2 believes him, this would even more encourage P2 to play steal and get all the money (pm) with a very high probability. If P2 does not believe P1, he would also play steal. Therewith he at least keeps a chance to win all the money. If he played split he would get nothing very likely and half the money elsewise. If P1 assures to Steal and P2 believes him, P2 would end up getting nothing no matter his choice. Nevertheless, he would play steal to keep a tiny chance of winning everything. If he does not believe P1, he would play steal to get everything in case P1 plays split. As it is a symmetric game, the same holds for what P1 assures to do in their communication. To conclude, the game is constructed in a way that communication does not help. No matter what both players say and whether they believe each other, both their best, rational choices would always be steal.

What happens in clip 1?

The man tries to use communication to agree upon a guaranteed half of the prize money.He assures/guarantees the girl to play *split* and asks her to do the same. She promises to play *split* either. However, she is lying. As a rational, self-interested and non-cooperative player, she uses the communication to make the man play *split* and to assess his probability of playing *split* close to 1 or even 1. If the man plays *split* with probability 1, she plays *steal* as her best response. It would be irrational to play *split* either.

What is the man's strategy in clip 2?

The man (P1)'s strategy is to guarantee himself the half price money by changing the rules of the game. He assures P2 to play *steal* no matter what and promises, that if P2 would play *split*, he would equally share the total payoff he gets with him. As such, he changes the game's payoff matrix to the following:

		P2: SP	P2: ST
P1	SP	$(\frac{pm}{2}, \frac{pm}{2})$	$(\frac{pm}{2}, \frac{pm}{2})$
	ST	(pm, 0)	(0,0)

Now, P2 has the following options, given that P1 plays ST with p = 1.

$$EU_2(ST, ST) = 0$$
$$EU_2(ST, SP) = \frac{pm}{2}$$

P2's best response to P1's affirmation is to play split. No matter, whether he trusts P1 or not, he can now guarantee himself a payoff of $\frac{pm}{2}$ by playing *split*. P1 changed the rules of the game to end up at the strategy profile $s_1 = (SP, SP)$. Thereby, he renounced his chance of getting the whole price, but clearly raised his chance of getting the half. P1 confused P2 to end up at s_1, which is rationally unlikely in the game elsewise.

2.2 Hawk or Dove

1. Explore how the Nash equilibria for this game depend on the parameters v (value of resource) and c (cost of aggression). If possible, explain the result in layman terms.

Given is the following payoff matrix:

		Player 2	
		$Hawk\ (q)$	$Dove\ (1-q)$
Player 1	$Hawk\ (p)$	$v/2-c, v/2-c$	$v, 0$
	$Dove\ (1-p)$	$0, v$	$v/2, v/2$

We get the mixed Nash equilibrium by finding the values for p and q. We start with the expected utility for player 1 for its 2 moves: playing Hawk and playing Dove:

$eu_1(Hawk, q) = (\frac{1}{2}v - c)q + (1-q)v = \frac{qv}{2} - qc + v - qv$

$eu_1(Dove, q) = (1-q)\frac{1}{2}v = \frac{v - qv}{2}$

Where eu_1 is the expected utility of player 1. To find the value of q we set the expected utilities equal to each other:

$$\begin{aligned} eu_1(Hawk, q) &= eu_1(Dove, q) \\ \frac{qv}{2} - qc + v - qv &= \frac{v - qv}{2} \\ qv - 2qc + 2v - 2qv &= v - qv \\ -2qc &= -v \\ q &= \frac{\frac{1}{2}v}{c} \end{aligned}$$

Since the payoff matrix is symmetric we know that $p = q$

We note that if $\frac{1}{2}v \geqslant c$ then $q = p \geqslant 1$ meaning that the best strategy will be to always play Hawk. This makes sense as half the value is greater than the cost. So if a Hawk meets another Hawk they will both have a positive utility. And if a Hawk meets a Dove the Hawk will also have a positive utility.

When the cost c reaches 0 then q and p become very big. Meaning that it is once again good to play the Hawk strategy all the time. This makes sense as there is no negative effect on fighting for the resource anymore.

The interesting cases are when $0 \leqslant \frac{0.5v}{c} \leqslant 1$.

In these cases the probability that the players play Hawk is simply $\frac{0.5v}{c}$. This makes sense, as the cost c increases relative to the value v, fighting over the value becomes more costly. Thus choosing Hawk less times yields a better expected utility in this scenario.

Likewise, if cost c decreases relative to the value v, fighting over the value becomes less costly. Thus choosing Hawk more times yields a better expected utility.

2. Suppose that reproductive success depends on how successful you are in these encounters. If your trait gives you above-average pay-off, then a higher rate of reproduction will mean that the proportion of your trait in the population grows. Now imagine a 100% D-population in which a single H-mutant is introduced. What will happen? What about the opposite scenario (i.e. a homogeneous H-population in which a D-mutant appears)?

What will happen depends on the value of v and c. Lets define r as

$$r = \frac{0.5v}{c}$$

If $r \leqslant 0$ then there is no extra cost when fighting over the value. In this case 100% of the population will become a Hawk.

If $r \geqslant 1$ then the cost of fighting over the food outweighs the extra value gained. Thus the entire population will stay a Dove.

If $0 \leqslant r \leqslant 1$ then, as we have seen in the previous question, the Hawk will give an above-average payoff if the percentage of Hawks in the population is smaller than r. Thus the population will stabilize with an r percentage of Hawks.

If the entire population is a Hawk and a single Dove mutant appears then it will die, horribly. As the Dove cannot meet another Dove it will always have an expected utility of 0. However, there is the special case where the cost c is greater than $0.5v$. In this case the Hawks will have a negative utility when fighting over the value. And thus the average utility of the Dove is better, increasing the amount of Doves until there are only Doves left.

2.3 Correlated equilibrium

The traffic game:

		Driver 2	
		Go	Wait
Driver 1	Go	(-10, -10)	(1, -1)
	Wait	(-1,1)	(-1,-1)

1. What are the NE's (mixed & pure)?

The strategy profile $s^* = (Go, Wait)$ and the strategy profile $s^{**} = (Wait, Go)$ are the two pure, strict Nash Equliibria of this game. Those two are stable strategy profiles, in which no driver would want to change his strategy knowing that of the other. They both play their (only/one) best response.

Is there a mixed NE?

		Driver 2	
		Go (q)	Wait (1-q)
Driver 1	Go (p)	(-10, -10)	(1, -1)
	Wait (1-p)	(-1,1)	(-1,-1)

Driver 1: $EU_1(Go, s_2(q)) = -10 * q + 1 * (1 - q)$
$= -11q + 1$

$EU_1(Wait, s_2(q)) = -1 * q - 1 * (1 - q)$
= -1

$-11q + 1$ = -1 $\leftrightarrow q = \frac{2}{11}$

Driver 2: $EU_2(s_1(p), Go) = -10 * p + 1 * (1 - p)$
$= -11p + 1$

$EU_2(s_1(p), Wait) = -1 * p - 1 * (1 - p)$
= -1

$$-11p + 1 = \text{-1} \leftrightarrow p = \frac{2}{11}$$

$\rightarrow$ There is one mixed NE for $p = q = \frac{2}{11}$.

2. What is the EU for each player in the MNE?

Applying $p = q = \frac{2}{11}$, the expected utility for each player in this game is -1.
As it is a symmetric game, both players have the same expected utility of -1.

3. What is the expected utility if we mix the two PNE's as described above?

An equal mixing of the two PNE's with equal probabilities could be done by a third party. Naturally, this is done by a traffic light. It mixes the strategies $(Wait, Go)$ and $(Go, Wait)$ with equal probabilities. As such, it is a fair, randomizing device, that alterningly tells one of the drivers to *Go* and the other to *Wait*. The unwanted strategy profiles *(Wait, Wait)* and *(Go, Go)* are assigned probability 0 / can be avoided completely.

The traffic light's action recommendation would follow the subsequent distribution:

		Driver 2	
		Go	Wait
Driver 1	Go	0	$\frac{1}{2}$
	Wait	$\frac{1}{2}$	0

$\rightarrow$ If Driver 1 is assigned *Go*, he knows for sure that Driver 2 will simultaneously be assigned *Wait*. His expected utility of *Going* is $EU_1(Go) = 1 * 1 + 0 * (-1) = 1$.
If Driver 1 is assigned *Wait*, he knows for sure that Driver 2 will simultaneously be assigned *Go*. His expected utility of *Waiting* is $EU_1(Wait) = 1 * (-1) + 0 * 1 = -1$. As this situation is symmetric, the same holds for Driver 2 vice versa.

4. Why is there no mixed strategy profile, that can achieve this utility?

In any mixed strategy, the strategy profiles *(Wait, Wait)* and *(Go, Go)* are not avoided completely. They occur with a probability greater than 0. Both profiles mean negative payoffs for both players. As such, they lower their expected utilities. For this reason, no mixed strategy can give the driver's the utility from (3.).

The game of chicken

		P2	
		Chicken (q)	Dare (1-q)
P1	Chicken (p)	(6, 6)	(2, 7)
	Dare (1-p)	(7,2)	(0,0)

5. What are the NE & What is the expected utility for the MNE?

If one of the players is going to *Dare*, it is better for the other to *chicken out*. But if one is going to *Chicken out*, it is better for the other to *Dare*. In this situation, everyone wishes to *Dare*, but only given the case that the other one *Chickens out*.

The game has 3 NE's.
There are pure strategy Nash Equilibria/PNE's. Those are the strategy profiles $s^* = (Dare, Chicken)$ and $s^{**} = (Chicken, Dare)$. Within these strategy profiles each player plays his best response to the action of the other.

There is one mixed strategy Nash Equilibrium (MNE).

$$\text{Player 1: } EU_1(Chicken, s_2(q)) = 6 * q + 2 * (1 - q)$$
$$= 4q + 2$$

$$EU_1(Dare, s_2(q)) = 7 * q$$

$$4q + 2 = 7\text{q} \leftrightarrow q = \tfrac{2}{3}$$

$$\text{Player 2: } EU_2(s_1(p), Chicken) = 6 * p + 2 * (1 - p)$$
$$= 4p + 2$$

$$EU_2(s_1(p), Dare) = 7 * p$$

$$4p + 2 = 7\text{p} \leftrightarrow p = \tfrac{2}{3}$$

$\rightarrow$ There is one mixed NE for $p = q = \frac{2}{3}$. In this case, each player *Dares* with probability $\frac{1}{3}$.

The expected utility for each player in this MNE is $\frac{14}{3}$.

6. Expected utilities under action suggestion of third party

P2

		Chicken	Dare
P1	Chicken	$\frac{1}{3}$	$\frac{1}{3}$
	Dare	$\frac{1}{3}$	0

Now there is a third party, that draws $(Dare, Chicken)$, $(Chicken, Dare)$ or $(Chicken, Chicken)$ with probability $\frac{1}{3}$ each. The third party (e.g. traffic lights) is assumed to draw one of these strategy profiles. Next up, the party informs the players about the part of the strategy they are suggested to take. However, the players do not know about the action suggested for their opponent.

Alternative 1: A player is assigned *Dare*. Then, the other player can only be assigned *Chicken*. In this case, the player will not want to deviate, since he gets 7 - being the highest payoff possible - for sure. His expected utility of *Daring* is 7*1 = 7.

Alternative 2: A player is assigned *Chicken*. In this case, his opponent will play *Chicken* with probability $\frac{1}{2}$ and *Dare* also with probability $\frac{1}{2}$. (As there are only two equally possible strategies left)
$\rightarrow$ The expected utility is: $6 * \frac{1}{2} + 7 * \frac{1}{2} = 6.5$.

7. This a NE in a sens that noone would want to deviate.

This game is a correlated Nash Equilibrium, since neither player has an incentive to change his strategy. The expected utility of this CNE is $7 * \frac{1}{3} + 2 * \frac{1}{3} + 6 * \frac{1}{3} = 5$. This expected payoff is higher than that of the MNE, which is $4,67$.

8. Why is the correlated equilibrium a generalisation of the "standard" NE?

Informally spoken, a correlated equilibrium (CE) is a randomized assignment of (potentially correlated) action suggestions to agents in such a way, that none of them would want to deviate.
It is a generalisation of the standard NE in the sense that every standard NE is a special case of the CE, where the agent's probabilities of choosing a strategy are uncorrelated.
E.g. For two agents: $p(s) = p(s_1, s_2) = p(s_1) * p(s_2)$

2.4 Iterated Prisoner's Dilemma

Using the payoffmatrix of the lecture:

		P2	
		Cooperare	Defect
P1	Cooperate	-1,-1	-12,0
	Defect	0,-12	-8,-8

- Both players act rational in a sense that they aim for payoff maximization
- They can get at max 0, at min -12
- They play the same game for n times
  ⇒ They can leaRn from the results of the previous games

Choice C-C is not optimal for the individual, but for the group ⇒ **sharing/coperating** = working together to get more together
Choices D-C / C-D = defecting/betraying = working to get personal gain
Iteration: For repeated games the added up score decides about the strategy.
Reaction strategies:
At first a few baseline methods, that dont consider the history of the game.

1. The random actor.
   The player chooses C and D at random/with a probability of 0.5 again and again no matter of his opponents history.

2. The defector.
   The player chooses C as the pure strategy. He does not learn from his opponents.

3. The angel.
   Some as defector, but always cooperating.

4. The Grudger.
   Cooperates in the first round. Cooperates until opponent chooses D for the first time then only defects afterwards.

5. Tit for Tat.
   Starts cooperating, after that it just copies what the opponent did in te previous round. Even if opponent gets back to cooperating, he defects ince to punish the other guy. It has been found to be the most successfull strategy in Axelrods tournaments. It is forgiving thus always allows to get back to cooperation after defection.

6. Tit for two tats/forgiving tit for tat.
   Strategy requires two defections after another before it retaliates. Prevents echo effets of regular tit for tat, though the opponent theoretically can learn that he can C-D-C-D-C without retaliation.

7. Parlov strategy: Win Stay / Loose Switch.
   Cooperates in the first turn. Then cooperates if and only if both players opted for the same choice in the previous round/move

Tournament of certain length matches

For the first evaluation method, we chose to match these strategies against each other in games of 100 rounds. Then, the mean of each strategie across all matches is looked at. As we understood, that was the original evaluation method of Axelrod. We encountered the following result.

strat	rank	u
grudger	1	-2.43
tit.for.tat	2	-2.61
tit.for.two.tat	3	-2.70
parlov	4	-2.89
always.coop	5	-3.37
random.action	6	-5.58
always.defect	7	-5.70

Mean payoffs for all strategies

- strategy which doesnt forgive one defect wins in this scenario
- As in the original tournament, tit-fortat the best of the rest
- Interesting: Better to always cooperate than always defect.
- Better to always cooperate than random

Tournament of UNcertain length matches

Now, after each match it is decided, whether to continue or not. For a few probability values, we looked at how the results changed. At first, the probability to continue was set to 0.9 with the following end results:

strat	rank	u
grudger	1	-2.59
tit.for.tat	2	-2.64
tit.for.two.tat	3	-2.81
parlov	4	-2.90
always.coop	5	-3.33
random.action	6	-5.05
always.defect	7	-5.23

Mean payoffs for all strategies against each other when roundnumber unsure.

- The ranks of the strategies did not change
- The 'good' strategies got worse, whereas the 'bad' strategies got slightly better

Tournament of UNcertain length matches

strat	rank	u
grudger	1	-2.76
tit.for.tat	2	-2.81
parlov	3	-3.10
tit.for.two.tat	4	-3.14
always.coop	5	-3.28
random.action	6	-4.34
always.defect	7	-4.35

probability to continue =0.7

strat	rank	u
tit.for.tat	1	-3.05
grudger	2	-3.05
tit.for.two.tat	3	-3.31
parlov	4	-3.35
always.coop	5	-3.39
always.defect	6	-3.69
random.action	7	-3.80

probability to continue =0.5

At probability 0.7 to continue, a shift in ranks appeared between parlov and titfor2tat. Other than that have the best values decreased and the worst increased. Same can be seen when it is a coin flip, whether to continue or not. So in general, the less likely it gets to continue or fewer repititions, the smaller the difference among strategies. Titfortat seems to be consistently within the top regions, eventhough the punish strategy GRUDGER is in the same payoff region. These results can not be compared to Axelrods findings, as he had many more strategies against each other and thus possibly other matchups, that had the titfortat more often as a clear winner and thus increasing its payoff.

Noisy Tournament of certain lengths

So, now a probabilty that a strategy sees a C as a D and vice versa is introduced and studied for effects on the strategies successes.

strat	rank	u
tit.for.two.tat	1	-5.04
tit.for.tat	2	-5.07
grudger	3	-5.20
always.defect	4	-5.22
parlov	5	-5.27
random.action	6	-5.59
always.coop	7	-5.94

probability to see error for both C and D =0.15

strat	rank	u
always.defect	1	-4.56
grudger	2	-4.56
tit.for.tat	3	-5.42
parlov	4	-5.44
random.action	5	-5.53
tit.for.two.tat	6	-5.82
always.coop	7	-6.45

probability to see error for both C and D =0.35

It can be seen that even a small probability for errors changes the rankings. Also the overall payoff value decreased substantially leading to values for all between -5 and -6. When there is an even larger chance of getting the previous opponents action wrong (0,35), then the strategies differ entirely. Interestingly, the always defect strategy all of a sudden ends up on top of all strategies.
These results can not be compared to Axelrods findings, as he had many more strategies against each other and thus possibly other matchups, that had the titfortat more often as a clear winner and thus increasing its payoff. Also these results only describe the given strategies for a given number of rounds. So most important are not the numbers but rather the fact that results change, when circumstances differ, even only slightly. Also noticeable is that the titfortat mostly lies within the better strategies, which somewhat coincides with Axelrods findings.