# Multi-Agent Systems

## Homework Assignment 1

## MSc AI, VU

E.J. Pauwels

Deadline: Wednesday Nov 15, 2023 (23h59)

# 1 Game Theory: Optimality Concepts and Nash Equilibrium

## 1.1 Odd or even game

Two players A and B each can draw either numbers "1" or "2", if the sum of the drawn numbers are even player A pays player B the sum. If the sum is odd, then player B pays player A that sum.

#### Questions

1. Write down the pay-off matrix for this game.
2. What are the regret minimisation strategies in terms of **pure** strategies?
3. What are the regret minimisation strategies in terms of **mixed** strategies?
4. What are the safety strategies for each player in terms of **pure** strategies?
5. What are the safety strategies for each player in terms of **mixed** strategies?

## 1.2 Travelers' dilemma: Discrete version

An airline severely damages identical antiques purchased by two different travelers. The management is willing to compensate them for the loss of the antiques, but they have no idea about the actual value, After some research, the management is convinced that the value is either 1, 2 or 3 (in some appropriate unit), and they ask the two travelers to separately choose the appropriate number. To incentivise the travelers to come up with the correct number they introduce the following rules for the procedure:

1. If both travelers pick the same number, they will be reimbursed that amount.
2. If they pick different numbers, management will assume that the lowest number is the correct one, and pay both of them the lower figure, but the person with the lower number will get a bonus of size $a$ for honesty (where $0 < a < 1/2$), while the one who picked the higher number will get a penalty of size $a$.

**Questions**

1. Write down the pay-off matrix for this game.
2. Determine the pure Nash equilibriua (PNE, there might be none, one or multiple ones).
3. Are there mixed Nash equilibria in which all three strategies are mixed?
4. Does this knowledge about the NE help the travelers in their decision?
5. Write down the regret matrix. This matrix is similar to the pay-off matrix, but now specifies the regret (rather than the pay-off) for each action profile.
6. What are the (pure) regret minimisation strategies?

## 1.3 Cournot's Duopoly (continuous version)

Cournot's duopoly is a game that models economic competition with strategic substitutes. Strategies are called *strategic substitutes* if an increase in your strategy will cause the competitor to decrease the use of his strategy ( *"if you increase your part, I will reduce mine"*). This contrasts with *strategic complements* when an increase in the strategy of one player causes the other player to follow suit (*"if you increase your part, I will do the same."*.)

Two companies make an interchangeable product (e.g. bottled water). Both need to determine (simultaneously!) the quantity they will produce (say for next week). Call these quantities $q_1$ and $q_2$, respectively. The unit price $p$ (price of each unit, e.g. one bottle) of the product in the market depends on the total produced quantity $q_1 + q_2$. Specifically

$$p(q_1, q_2) = \alpha - \beta(q_1 + q_2) \qquad (\alpha, \beta > 0).$$

Firm 1 can produce each unit at a unit-cost $c_1$, whereas the unit-cost for firm 2 equals $c_2$.

1. What is the best response for each company given the quantity the other company will produce?
2. Suppose the companies need **not** decide on their quantity at the same time, but can react to one another (an unlimited number of times). What will be the outcome? (Provide a diagram.)

## 1.4 Ice cream time!

Three competing ice-cream vendors (Alice, Bob and Charlize) are trying to sell their refreshments to tourists on the beach. We are making the following assumptions:

- the beach has total length of 1, while its width is uniform and much smaller than its length. So the beach can be represented as a line segment of length 1, and each position on the beach can be represented by the position parameter $0 \leq x \leq 1$.
- Tourists are uniformly distributed along the total length of the beach and will buy their ice-cream at the stall that is closest to their location;

**Questions:**

1. On a beautiful summer morning Charlize makes her way to the beach and upon arrival finds that her two competitors have already set up their stalls: Alice at location $a = 0.1$ and Bob at location $b = 0.8$. Discuss what Charlize's best response is: i.e. what location should she choose, given $a = 0.1$ and $b = 0.8$?

2. Same question as above, but now assume that all we know is that $a = 0.1$ and $a < b \leq 1$.

3. Earlier that morning, Bob arrived and discovered that Alice had already set up her stall at $a = 0.1$, while Charlize hadn't shown up yet. But Bob knows that Charlize will arrive before too long, and that she will try to position herself in such a way as to maximize her revenue. What location should Bob pick in order to maximize his own expected revenue?

4. At sunrise that morning, Alice arrived before both Bob and Charlize, and set up her stall at location $a = 0.1$. However she knows for sure that the other two vendors will show up soon. Where **should she have set up** her stall in order to maximise her expected revenue?

# SOLUTIONS

# 1 Game Theory: Concepts and Nash Equilibrium

## 1.1 Odd or even game

1. Pay-off matrix (zero-sum game):

$$
\begin{array}{cc|c|c|}
 & & \multicolumn{2}{c}{\text{Pl B}} \\
 & & 1 & 2 \\
\hline
\text{Pl.A} & 1 & -2 \quad 2 & 3 \quad -3 \\
\hline
 & 2 & 3 \quad -3 & -4 \quad 4 \\
\hline
\end{array}
$$

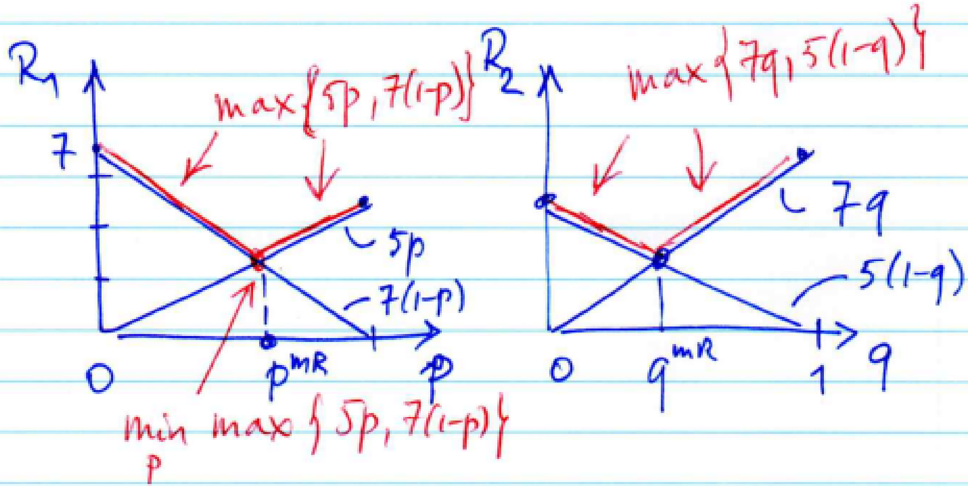
2. Regret matrix and max regret minimisation (pure) strategies:

Regret

$$
\begin{array}{c|c|c|l}
 & 1 & 2 & \text{max} \\
\hline
1 & 5 \quad 0 & 0 \quad 5 & \longrightarrow \boxed{5} \leftarrow \text{min max } R_1 \\
\hline
2 & 0 \quad 7 & 7 \quad 0 & \longrightarrow 7 \\
\hline
\text{max} & \downarrow & \downarrow & \\
 & 7 & \boxed{5} & \leftarrow \text{min max } R_2
\end{array}
$$

3. Regret minimisation for mixed strategies

| | $q$ | $1-q$ | |
|---|---|---|---|
| $p$ | 5   0 | 0   5 | $5(1-q)$ |
| $1-p$ | 0   7 | 7   0 | $7q$ |
| | $5p$ | $7(1-p)$ | |

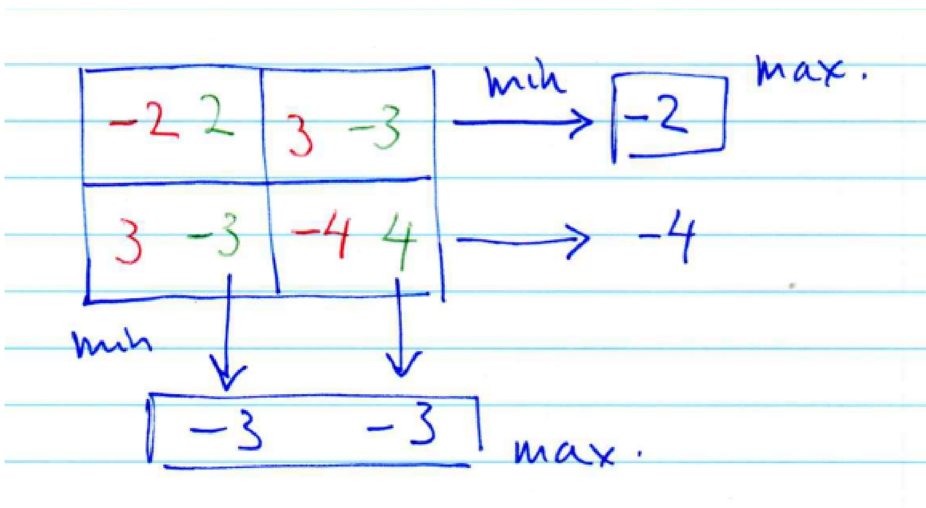



$p^{mR}$ ? $\quad 7(1-p) = 5p \qquad\qquad 7q = 5(1-q)$

$$\boxed{p^{mR} = \frac{7}{12}} \qquad\qquad \boxed{q^{mR} = \frac{5}{12}}$$

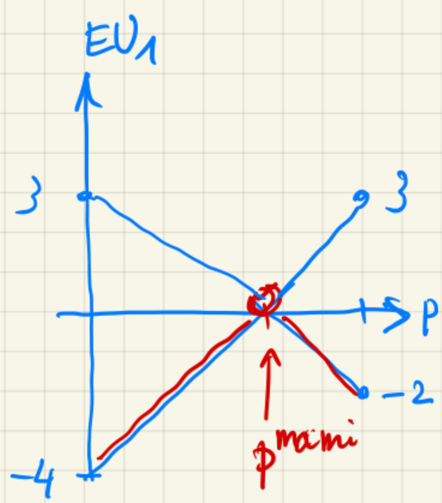
4. Safety strategies for each player in terms of pure strategies?

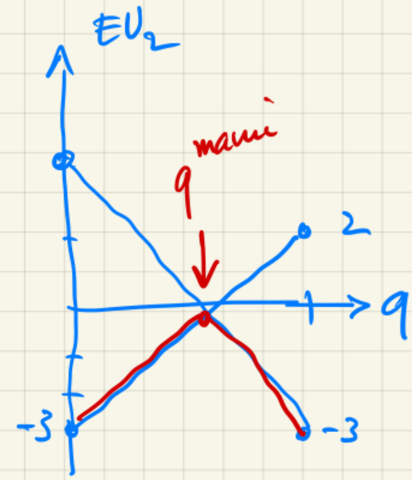
- Player A: Safety strategy = play "1" (corresponding $v_A^{mami} = -2$)
- Player B: Indifferent between both strategies (corresponding $v_B^{mami} = -3$)



5. Safety strategies for each player in terms of mixed strategies:

| | $q$ | $1-q$ | |
|---|---|---|---|
| $p$ | $-2 \quad 2$ | $3 \quad -3$ | $2q-3(1-q)$ <br> $5q-3$ |
| $1-p$ | $3 \quad -3$ | $-4 \quad 4$ | $-3q+4(1-q)$ <br> $=-7q+4$ |
| | $-2p+3(1-p)$ <br> $=-5p+3$ | $3p-4(1-p)$ <br> $=7p-4$ | |




$$-5p+3 = 7p-4$$

$$\boxed{\begin{aligned} p^{mami} &= 7/12 \\ v_A^{mami} &= -5\tfrac{7}{12}+3 \\ &= 1/12 \end{aligned}}$$

$$5q-3 = -7q+4$$

$$\boxed{\begin{aligned} q^{mami} &= 7/12 \\ v_B^{mami} &= 5.\tfrac{7}{12}-3 \\ &= -1/12 \end{aligned}}$$

## 1.2 Traveler's dilemma: discrete version

Notice that since $0 < a < 1/2$ it follows that $0 < a < 1 - a < 1$.

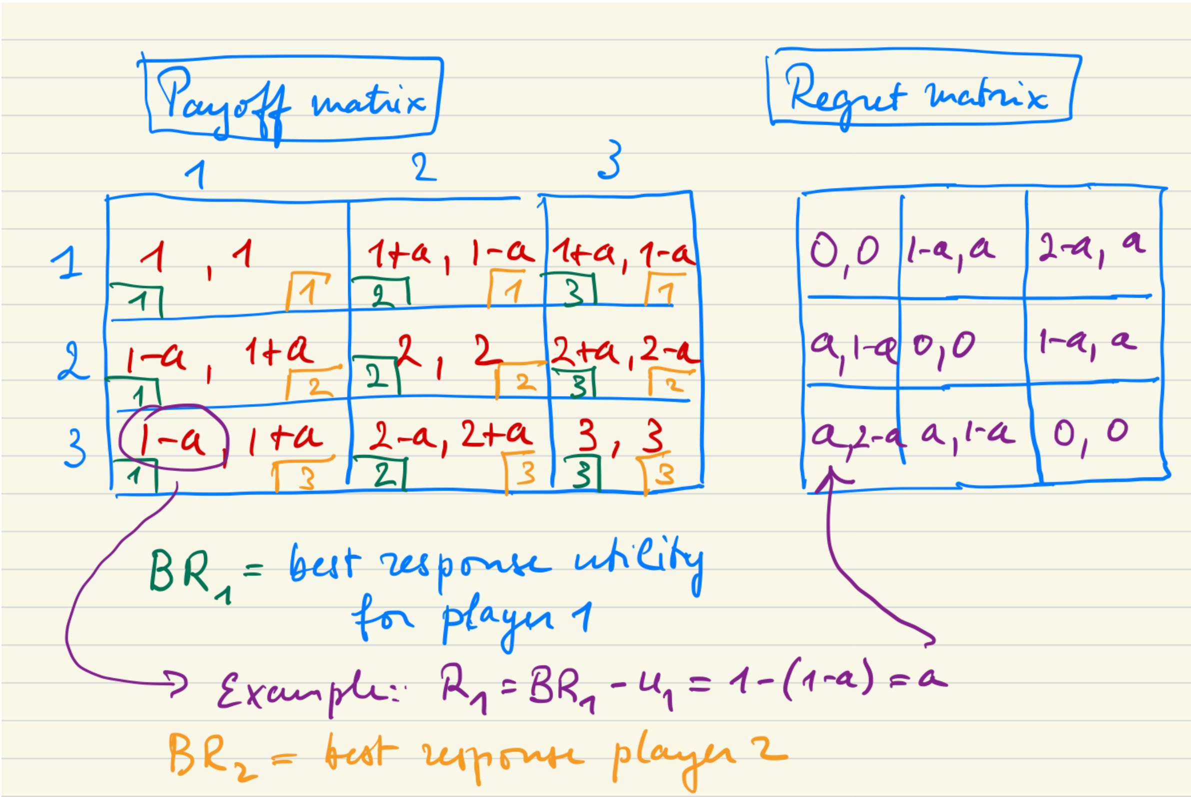



Figure 1: Pay-off and regret matrix for both players.

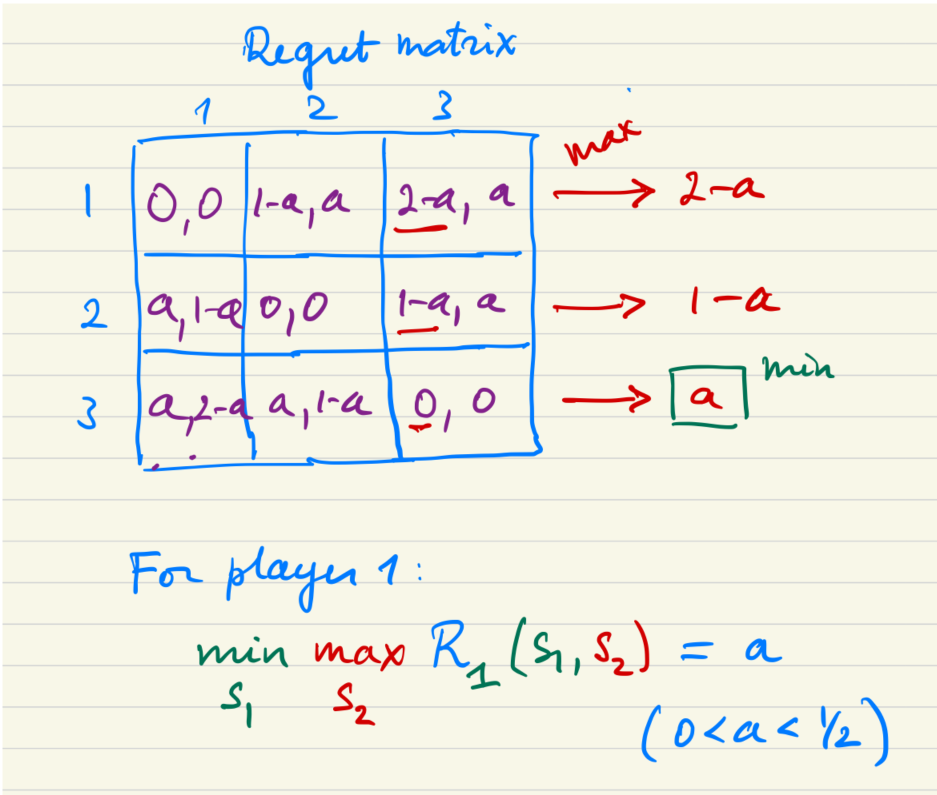



Figure 2: Regret minimisation for player 1. Based on this computation player 1 will opt for 3. Similar conclusion for player 2.

**Computation of mixed Nash eq. (mixing all three strategies**

| | $q_1$ | $q_2$ | $q_3$ |
|---|---|---|---|
| $p_1$ | $1, 1$ | $1+a, 1-a$ | $1+a, 1-a$ |
| $p_2$ | $1-a, 1+a$ | $2, 2$ | $2+a, 2-a$ |
| $p_3$ | $1-a, 1+a$ | $2-a, 2+a$ | $3, 3$ |

Player 2 has to choose his mixing vector $q=(q_1,q_2,q_3)$ such that player 1 becomes indifferent regarding his choice of actions, i.e

$$EU_1(1,q) = EU_1(2,q) = EU_1(3,q):$$

$$u_1 := EU_1(1,q) = 1.q_1 + (1+a)q_2 + (1+a)q_3$$

$$u_2 := EU_1(2,q) = (1-a)q_1 + 2q_2 + (2+a)q_3$$

$$u_3 := EU_1(3,q) = (1-a)q_1 + (2-a)q_2 + 3q_3$$

Figure 3: Computation of mixed Nash equilibrium:

We therefore get a system of 3 equations in 3 unknows $q_1, q_2, q_3$:

$$u_1 - u_2: \quad a q_1 - (1-a) q_2 - q_3 = 0 \qquad (3)$$

$$u_2 - u_3: \quad a q_2 - (1-a) q_3 = 0 \qquad (2$$

$$q_1 + q_2 + q_3 = 1 \qquad (1)$$

$$\rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & a & a-1 & 0 \\ a & a-1 & -1 & 0 \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & a & a-1 & 0 \\ 0 & \underbrace{a-1-a}_{-1} & -1-a & -a \end{array}\right)$$

$$\longrightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & a & a-1 & 0 \\ 0 & 1 & a+1 & a \end{array}\right) \rightarrow \left(\begin{array}{ccc|c} 1 & 1 & 1 & 1 \\ 0 & a & a-1 & 0 \\ 0 & 0 & a^2+1 & a^2 \end{array}\right)$$

$$\Rightarrow q_3 = \frac{a^2}{a^2+1}, \quad q_2 = \frac{1-a}{a} \cdot q_3 = \frac{(1-a)a}{a^2+1}$$

$$q_1 = 1 - \frac{a^2 + a(1-a)}{a^2+1} = \frac{a^2+1-a}{a^2+1} = \frac{a^2-a+1}{a^2+1}$$

Figure 4: Computation of mixed Nash equilibrium. Notice that since $a > 0$ and $a - 1 > 0$, the solutions $q_i > 0$ and therefore constitute a probability distribution.

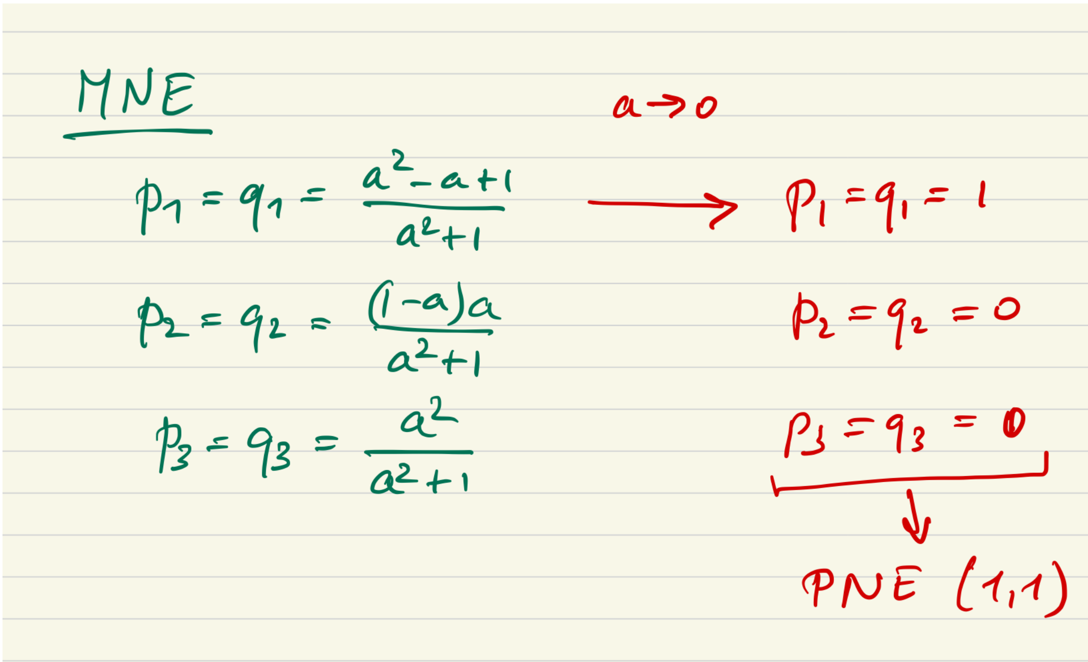



Figure 5: Evolution of the mixed NE to a PNE as the bonus/penalty $a$ is reduced, i.e. $a \to 0$.

**Conclusion**

- There are 3 pure and one mixed NE, so this is not very helpful in selecting a strategy.
- Regret minimisation suggests for both players to pick 3.

## 1.3 Cournot's Duopoly

Each company needs to decide which quantity to produce, and this means that each of the two players has an infinite number of possible actions to choose from. Let's introduce some notation: Company $i$ produces quantity $q_i$ at unit costs $c_i$ ($i = 1, 2$). The price setting (per unit) depends linearly on the total quantity produced (i.e. $q_1 + q_2$):

$$\text{unit-price:} \qquad p = \alpha - \beta(q_1 + q_2) \qquad \text{where } \alpha, \beta > 0.$$

The expected utility (profit) for company 1 assuming that quantities $q_1$ and $q_2$ are produced, equals:

$$u_1(q_1, q_2) = pq_1 - c_1 q_1 = (p - c_1)q_1 = ((\alpha - c_1) - \beta(q_1 + q_2))q_1$$

Similarly:

$$u_2(q_1, q_2) = pq_2 - c_2 q_2 = (p - c_2)q_2 = ((\alpha - c_2) - \beta(q_1 + q_2))q_2.$$

**Best response** Assuming that company 2 produces quantity $q_2$, the best response for company 1 by maximising utility $u_1$. To this end we compute the (partial) derivative and set it to zero:

$$\frac{\partial u_1}{\partial q_1} = -2\beta q_1 + (\alpha - c_1 - \beta q_2) = 0$$

From this it follows that the best response $BR_1(q_2)$ for company 1 (given company 2 produces $q_2$) equals:

$$q_1^* := BR_1(q_2) = \frac{\alpha - c_1 - \beta q_2}{2\beta} = \frac{\alpha - c_1}{2\beta} - \frac{q_2}{2}.$$

Notice how $q_1^*$ is a decreasing function in $q_2$ as expected for strategic substitutes.

Applying the same logic for company 2 yields:

$$q_2^* := BR_2(q_1) = \frac{\alpha - c_2 - \beta q_1}{2\beta} = \frac{\alpha - c_2}{2\beta} - \frac{q_1}{2}.$$

The corresponding dynamics is illustrated in Fig 6.

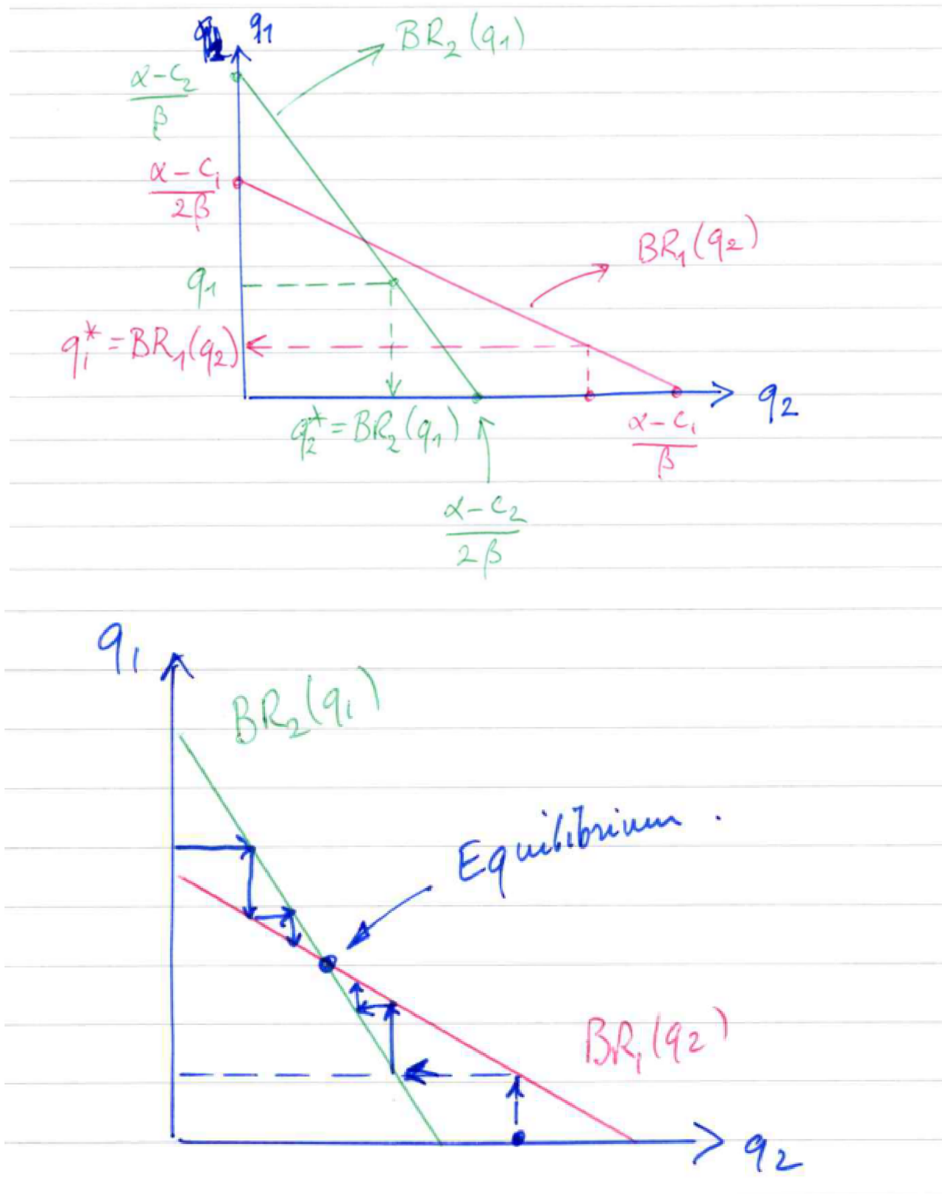



Figure 6: TOP: Best response functions for both companies. BOTTOM: Best response dynamics: one company starts with initial random quantity, the other company reacts, and this is iterated until the (Nash) equilibrium is reached.

## 1.4 Ice cream time!

### 1.4.1 Expected utiity (as function of position $x$) for Charlize $(EU_C(x))$

The expected utility of each player is equal to the number of customers the vendor will attract, or equivalently: what fraction of the $x$-positions is closes to the vendor. This result is graphed in Fig. 7.

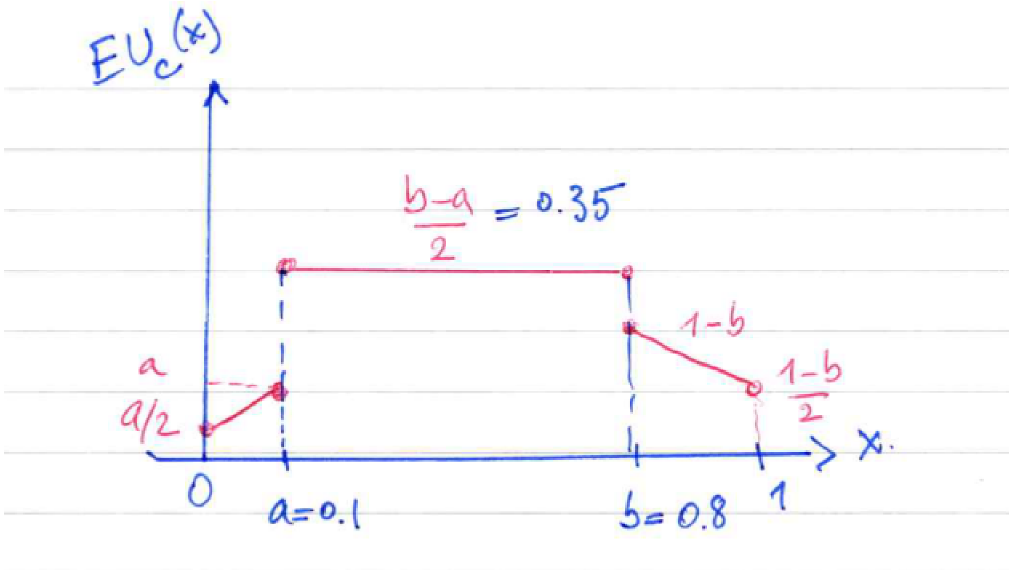



Figure 7: Expected utility $EU_C$ for Charlize for given positions of Alice and Bob.

### 1.4.2 Same question as above, but now for variable $b$-value

When the $b$-position is no longer fixed but can vary ($a < b < 1$) we can express the expected utility for Charlize as follows:

$$EU_C(x;b) = \begin{cases} \dfrac{x+a}{2} & \text{if} \quad 0 < x < a \\ \dfrac{b-a}{2} & \text{if} \quad a < x < b \\ (1-x) + \dfrac{x-b}{2} = 1 - \dfrac{x+b}{2} & \text{if} \quad b < x < 1 \end{cases} \tag{1}$$

### 1.4.3 Optimal positioning for Bob

When Charlize arrives she will select a position $x = c$ that maximizes her expected utility. Referring to Fig 8 we observe the following:

- CENTRE panel: when the expected utility $EU_C$ is continuous at $x = b$, Bob is indifferent about C's decision. Even if Charlize positions herself immediately to the left $c = b^-$ or to the right $c = b^+$ of Bob, Bob's utility will be the same. Continuity requires that

$$\frac{b-a}{2} = 1 - b \quad \Longrightarrow \quad b = \frac{a+2}{3}$$

with corresponding utility:

$$1 - b = 1 - \frac{(a+2)}{3} = \frac{1-a}{3}$$

- LEFT panel: if $(b-a)/2 > 1-b$ or equivalently $b > 0.7$, Charlize will take a position to the left of Bob, and in the worst case this could be right next to him ($c = b^-$). In that case Bob would be left with expected utility $1-b < 0.3$.
- RIGHT panel: if $(b-a)/2 < 1-b$, or $b < 0.7$ Charlize will position herself just to the right of Bob (i.e. $c = b^+$) and Bob will get utility $(b-a)/2 < 0.3$

So from the above we conclude that the optimal choice (best guaranteed minimum outcome) for Bob is (assuming $b^* > a$):

$$b^* = \frac{a+2}{3} \quad \text{with corresponding guaranteed minimal utility } u_B^* = \frac{1-a}{3}.$$

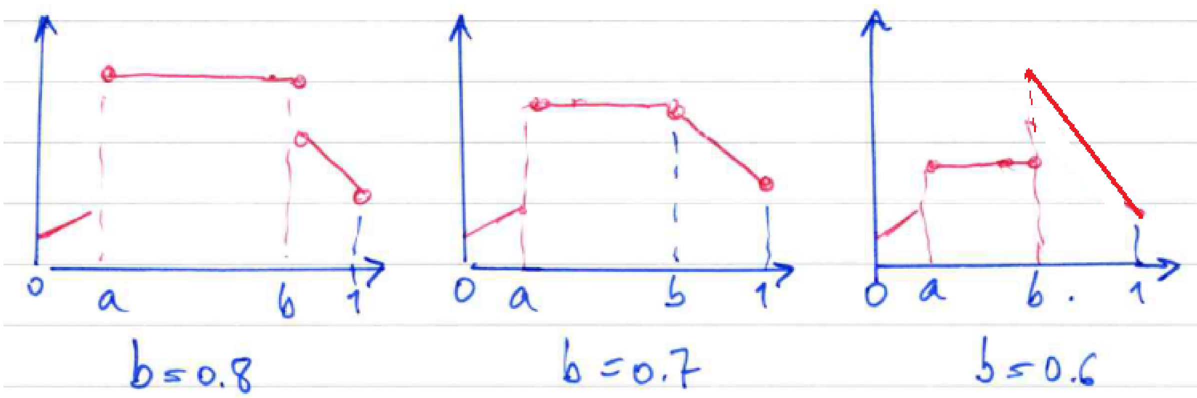



Figure 8: Expected utility $EU_C(x)$ for different choices of Bob's position $b$.

### 1.4.4 Optimal positioning for Alice

Alice knows that whatever her position, Bob will pick a position that maximizes his utility. This is either immediately to the left of Alice (i.e. $b = a^-$, with a utility $a$) or to the right of $a$ at $b^* = (a+2)/3$ with utility $(1-a)/3$. Alice will be indifferent to Bob's choice if these two utilities agree: So the optimal position needs to satisfy the following equation:

$$a^* = \frac{1-a^*}{3} \quad \Longrightarrow \quad a^* = 1/4.$$

This in turn implies that $b^* = 3/4$.

### 1.4.5 Summary of decision process

Recall that this is a constant sum game (the total reward is constant): hence revenue for one vendor can only happen at the expense of another. Suppose that Alice and Bob have already taken positions $0 < a < b < 1$. This divides the total interval in 3 subintervals: $[0, a], [a, b]$ and $[b, 1]$. These intervals should be such that Charlize is indifferent about which one she chooses. If that wasn't the case, then at least Alice or Bob could have chosen a more lucrative position. Charlize's utility in each of the intervals should therefore be the same, and this entails that:

$$a = \frac{b-a}{2} = 1 - b$$

From this we immediately deduce that $a = 1/4$ and $b = 3/4$.