Mas asg 1 - Multi agent system assignment 1

Multi-agent systems (Vrije Universiteit Amsterdam)

Multi Agent Systems

Homework Assignment 1, MSc AI

1.1 Iterated Best response dynamics for a simple matrix game

1) There are no dominated (weakly or strictly) strategies for player 1 and for player 2. We can also deduce that therefore there are no dominant strategies.

2) To use "best response" to find the *Nash equilibrium* means to turn the simultaneous game into a sequential game to solve it. For each strategy played by the other player, we look at what would be best for the player. In the end, we find that the only mutual best response is <middle, center> which results in a payoff of (7,6). This is the single Nash Equilibrium for this Game.

3) There are simple games like "Rock-Paper-Scissors" which would lead to different outcomes. It has no Pure Strategy Nash Equilibrium (PSNE) contrary to the pay-off matrix in this exercise.
We could have multiple NEs like in the Selten's game which has many equilibria (pure and mixed strategies).
This game has a Pure Strategy Nash Equilibrium and is not a zero-sum game. An example of a zero-sum game would be the Matching Pennies game which has a mixed strategies Nash Equilibrium.

1.2. Travellers' dilemma: Discrete version

1)

	1	2	3
1	<u>1</u>, <u>1</u>	1+a,1-a	1+a, 1-a
2	1-a, 1+a	<u>2</u>, <u>2</u>	2+a, 2-a

3	1-a, 1+a	2-a, 2+a	3, 3

2)

From the matrix we can see that the pure Nash equilibria are both players picking 1, 2, or 3 simultaneously.

3)

Yes, since every strategy for either player contains one best response against different strategies for the opposing player; no strategy is dominated by any other strategy.

Calculating the mixed Nash equilibrium would go something like this:

We equate the expected utility of player 2 choosing one with player 2 choosing two

$EU_2(s_1, 1) = EU_2(s_1, 2) \Rightarrow p1 + p2(1 + a) + p3(1 + a) = p1(1 - a) + 2p2 + p3(2 + a)$

$\Rightarrow ap1 - p2 + ap2 = p3$

We substitute p3 with (1-p1-p2)

$\Rightarrow ap1 + ap2 - p3 = 1 - p1 - p2$

$\Rightarrow ap1 + p1 + ap2 = 1$

Similarly equate the expected utility of player 2 choosing one with player 2 choosing three

$EU_2(s_1, 1) = EU_2(s_1, 3) \Rightarrow p1 + p2(1 + a) + p3(1 + a) = p1(1 - a) + p2(2 - a) + 3p3$

$\Rightarrow - p2 + 2ap2 + 2p3 + ap3 =- ap1$

We substitute p1 with (1-p2-p3)

$\Rightarrow - p2 + 2ap2 + 2p3 + ap3 =- a + ap2 + ap3$

$\Rightarrow - p2 + ap2 + 2p3 =- a$

$\Rightarrow p2 - ap2 - 2p3 = a$

Lastly we can equate the expected utility of choosing two with three

$EU_2(s_1, 2) = EU_2(s_1, 3) \Rightarrow p1(1 - a) + 2p2 + p3(2 + a) = p1(1 - a) + p2(2 - a) + 3p3$

$\Rightarrow - p3 + ap3 =- ap2$

We substitute p2 with (1-p1-p3)

$\Rightarrow - p3 + ap3 =- a + ap1 + ap3$

$\Rightarrow - ap1 - p3 =- a$

$\Rightarrow ap1 + p3 = a$

We convert these three equations into reduced row echelon form

p1	p2	p3	=
1+a	a	0	1
0	1-a	-2	a
a	0	1	a

By subtracting the last row from the first row we get:

p1	p2	p3	=
1	a	-1	1-a
0	1-a	-2	a
a	0	1	a

Subtracting the first row multiplied by *a* from the last row gives:

p1	p2	p3	=
1	a	-1	1-a
0	1-a	-2	a
0	$-a^2$	1+a	a^2

Adding the last row divided by *a* to the first row gives:

p1	p2	p3	=
1	0	$\frac{1}{a}$	1
0	1-a	-2	a
0	$-a^2$	1+a	a^2

Subtracting the last row divided by *a* from the second row results in:

p1	p2	p3	=
1	0	$\frac{1}{a}$	1
0	1	$-\frac{1}{a}-3$	0

0	$-a^2$	1+a	a^2

Adding the second row multiplied by a^2 to the last row gives:

p1	p2	p3	=
1	0	$\frac{1}{a}$	1
0	1	$-\frac{1}{a}-3$	0
0	0	1-3a^2	a^2

Next divide the last row by 1-3a^2 giving:

p1	p2	p3	=
1	0	$\frac{1}{a}$	1
0	1	$-\frac{1}{a}-3$	0
0	0	1	$\frac{a^2}{1-3a^2}$

Dividing the last row by *a* and subtracting it from the first row:

p1	p2	p3	=
1	0	0	$1-\frac{a}{1-3a^2}$
0	1	$-\frac{1}{a}-3$	0
0	0	1	$\frac{a^2}{1-3a^2}$

Multiplying the last row by $\frac{1}{a}+3$ and adding it to the second row gives:

p1	p2	p3	=
1	0	0	$1-\frac{a}{1-3a^2}$

0	1	0	$\frac{a}{1-3a^2} + \frac{3a^2}{1-3a^2}$
0	0	1	$\frac{a^2}{1-3a^2}$

This gives the probabilities for p1, p2, and p3, but also for q1, q2, and q3, because of the symmetrical nature of the matrix.
Due to the difficulty of this question, these probabilities are not accurate to the Nash Equilibrium, but our best effort.

4)

Not really in this case, since the strategy for the mixed Nash equilibrium is dependent on the value of *a* which is unknown, and the pure Nash equilibria all correspond to different strategies. The pareto optimal strategy, however, would be to always choose 3.

5)

	1	2	3	max regret
1	0, 0	1-a, a	2-a, a	2-a
2	a, 1-a	0, 0	1-a, a	1-a
3	a, 2-a	a, 1-a	0, 0	a
max regret	2-a	1-a	a	

6)

As can be seen from the max regret row and column, both players picking the number 3 will result in the least regret.

1.3 Cournot's Duopoly (continuous version)

In Cournot's duopoly game, two companies are producing an interchangeable product which behave the same on the market which make Company 1 C_1 and Company 2 C_2 in competition. Because the product is interchangeable, its price depends on the global quantity there is on the market and not only on one's company quantity which means price is unique p. Price is defined as a decreasing function of quantity 1 q_1 and quantity 2 q_2 because the more there is quantity offered with a fixed demand the less the price is.

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

In order to keep the highest price possible, both companies have an interest in finding a balanced quantity that will neither offer more market share to the other companies nor make the price go down because of too many offers compared to the demand. Their strategies are thus substitutes : an increase in one's strategy will cause the competitor to decrease the use of his strategy.

C_1 and C_2 have respective unit-costs c_1 and c_2.

$Players:\ C_1\ and\ C_2$

$Utility\ function\ or\ Profit:\ \Pi_i = p * q_i - c_i * q_i = (p - c_i) * q_i$

$Best\ response:\ For\ a\ given\ input\ q_1\ of\ player\ 2,\ what\ input\ q_2\ maximises\ the\ latter's\ utility\,?$

To find out we need to calculate the maximum utility Π_2 for a given q1.

By replacing p we have :

$$\Pi_2 = [\alpha - \beta(q_1 + q_2) - c_2] * q_2$$
$$\Pi_2 = \alpha * q_2 - \beta * q_2(q_1 + q_2) - c_2 * q_2$$
$$\Pi_2 = \alpha * q_2 - \beta * q_2 * q_1 - \beta * q_2^2 - c_2 * q_2$$

And we are actually trying to find the maximum of the function :

$$\frac{\partial \Pi_2}{\partial q_2} = 0$$
$$\frac{\partial \Pi_2}{\partial q_2} = \alpha - \beta * q_1 - 2 * \beta * q_2 - c_2$$

We want to resolve :

$$\alpha - \beta * q_1 - 2 * \beta * q_2 - c_2 = 0$$
$$\alpha - \beta * q_1 - c_2 = 2 * \beta * q_2$$

So :

$$q_2 = \frac{\alpha - \beta^* q_1 - c_2}{2^*\beta}$$

Equivalently we have :

$$q_1 = \frac{\alpha - \beta^* q_2 - c_1}{2^*\beta}$$

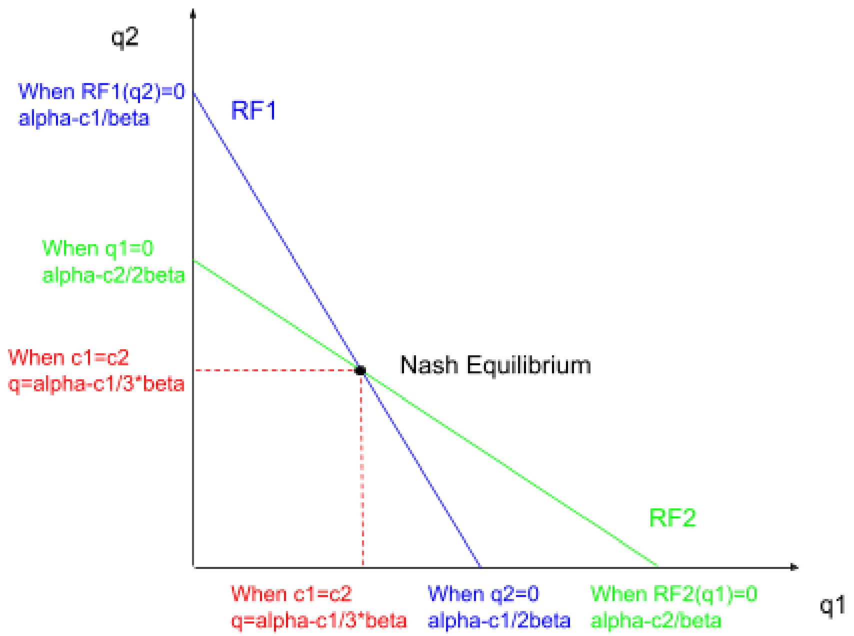
Both quantities are affine functions of the form $ax + b$ of the other quantity where :

$a = \frac{-\beta}{2\beta}$ and $b = \alpha - \frac{c}{2\beta}$

We can make the graph of these two reaction functions :

$$RF_i(q_j) = \frac{-\beta}{2\beta} q_j + \frac{\alpha - c_i}{2\beta}$$

If $q_j = 0$ then $RF_i(0) = \frac{\alpha - c_i}{2\beta}$, if $RF_i(q_j) = 0$, then $q_j = \frac{\alpha - c_i}{\beta}$



<u>Graph 1 : Continuous state space for C1 and C2 with a NE</u>

Because both of the companies want to maximise their profits and because they both presume the other company wants to maximise its profit as well, they both take into account the other strategy. This has to lead to the crossing of the two Responses Functions lines. Indeed, it's where the two lines cross each other that we'll find a Nash equilibrium e.g. somewhere none of the companies would change their strategy.

We want to know the point where both lines cross each other.

$$RF_1(q_2) \ = RF_2(q_1)$$

$$\frac{-\beta}{2\beta}q_2 + \frac{\alpha - c_1}{2\beta} = \frac{-\beta}{2\beta}q_1 + \frac{\alpha - c_2}{2\beta}$$

There are 3 possibilities :

- $c_1 = c_2$
- $c_1 \geq c_2$
- $c_1 \leq c_2$

First possibility: $c_1 = c_2$

Then :

$$\frac{-\beta}{2\beta}q_2 + \frac{\alpha - c}{2\beta} = \frac{-\beta}{2\beta}q_1 + \frac{\alpha - c}{2\beta}$$

$$\frac{-\beta}{2\beta}q_2 = \frac{-\beta}{2\beta}q_1$$

$$q_2 = q_1$$

Which means that if the two costs are the same, the quantity has to be the same as the other company. In such a case $q_2 = q_1 = \ q$ and :

$$q \ = \frac{-\beta}{2\beta}q \ + \ \frac{\alpha - c}{2\beta}$$

$$\frac{3*\beta}{2*\beta}q \ = \frac{\alpha - c}{2\beta}$$

$$q \ = \frac{\alpha - c}{3\beta}$$

In such a configuration, the quantity is then $\frac{\alpha - c}{3\beta}$.

Second possibility $c_1 \geq c_2$:

Then :

$$c_1 \geq c_2$$

$$-c_1 \leq -c_2$$

$$\frac{-\beta}{2\beta}q_2 + \frac{\alpha - c_1}{2\beta} \leq \frac{-\beta}{2\beta}q_1 + \frac{\alpha - c_2}{2\beta}$$

$$q_1 \leq q_2$$

Then :

$$-\ \beta q_2 + \ \alpha - \ c_1 \leq -\ \beta q_1 + \ \alpha - \ c_2$$

$$-\ \beta q_2 +\ \beta q_1 \leq c_1 -\ c_2$$
$$\beta * (q_1 - q_2) \leq c_1 -\ c_2$$
$$q_1 - q_2 \leq \frac{c_1 - c_2}{\beta}$$
$$q_1 \leq q_2 + \frac{c_1 - c_2}{\beta}$$

And because the company wants to maximise its profit it will choose the q_1 the closest to $q_2 + \frac{c_1 - c_2}{\beta}$ while remaining inferior to it. Symmetrically for $c_1 \leq c_2$, the company will choose q_1 the closest to $q_2 + \frac{c_2 - c_1}{\beta}$.

The best response for each company given the quantity of the other company is :

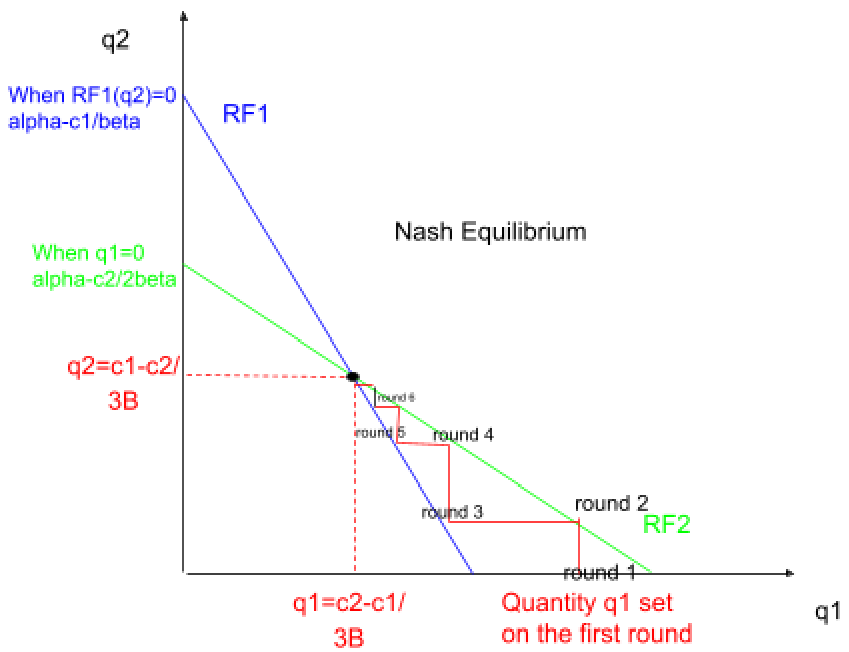
- $c_1 = c_2$ $\quad q = \frac{\alpha - c}{3\beta}$
- $c_1 \geq c_2$ $\quad max\ q_1 \leq q_2 + \frac{c_1 - c_2}{\beta}$ $\quad => q_1 = q_2 + \frac{c_1 - c_2}{\beta}$

or $\quad min\ q_2 \geq q_1 + \frac{c_1 - c_2}{\beta}$ $\quad => q_2 = q_1 + \frac{c_1 - c_2}{\beta}$

- $c_1 \leq c_2$ $\quad min\ q_1 \geq q_2 + \frac{c_2 - c_1}{\beta}$ $\quad => q_1 = q_2 + \frac{c_2 - c_1}{\beta}$

or $\quad max\ q_2 \leq q_1 + \frac{c_2 - c_1}{\beta}$ $\quad => q_2 = q_1 + \frac{c_2 - c_1}{\beta}$

If we suppose the companies need not decide on their quantity at the same time, but can react to one another (an unlimited number of times). Then the outcome can be compared to a series. In series, when N → +inf, then we focus on the convergence.



Graph 2 : Continuous state space for C_1 and C_2 with a NE

In such a case, if C_1 chooses a quantity q_1 in the first round, then C_2 will set its quantity based on profit maximisation strategy (Response Function 2). Then C_1 will choose its quantity based on RF1, and then again C_2 with RF2, etc. On an infinite number of reactions, the result will always converge toward the **Nash equilibrium**.

Indeed, firms are in equilibrium when each does not want to change what it is doing given the other firm's strategy. We know that $q_1 = \frac{\alpha - \beta * q_2 - c_1}{2 * \beta}$ and $q_2 = \frac{\alpha - \beta * q_1 - c_2}{2 * \beta}$. For each round, the value of the other quantity is based on the previous quantity the other firm offered. But on an infinite number of parties the series can be generalised as a system of two equations with two unknown variables:

$$q_1 = \frac{-\beta}{2\beta} q_2 + \frac{\alpha - c_1}{2\beta} \ ; \ q_2 = \frac{-\beta}{2\beta} q_1 + \frac{\alpha - c_2}{2\beta}$$

Which does, if we replace q_2 in the first equation :

$$q_1 = \frac{-\beta}{2\beta} \left(\frac{-\beta}{2\beta} q_1 + \frac{\alpha - c_2}{2\beta} \right) + \frac{\alpha - c_1}{2\beta}$$

...

$$q_1 = \frac{c_2 - c_1}{3\beta}$$

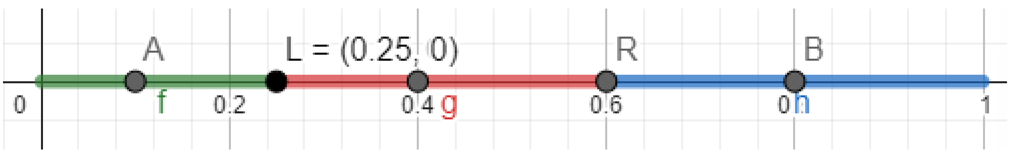
And equivalently :

$$q_2 = \frac{c_1 - c_2}{3\beta}$$

The outcome is that the two quantities will converge to these values. The more the number of times (the longer the sequence of the game), the nearer to the quantities around these values.

1.4. Ice cream time!

The three utilities of the three vendors correspond to the length of their space on the beach. We consider $0 \leq x \leq y \leq z \leq 1$ but the player playing last is not forced to be on z place if y place is better. These formulas can be seen as the utilities of the people on the left, in the middle, and on the right side of the beach respectively. We can define L the middle of x and y, and R the middle between y and z like the example below :



$$L = x + \frac{y-x}{2} = \frac{x+y}{2}$$
$$R = y + \frac{z-y}{2} = \frac{y+z}{2}$$

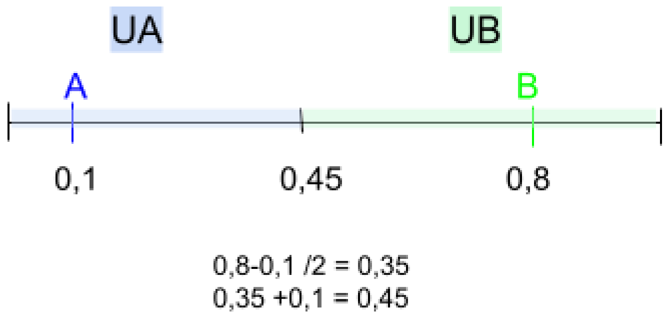
We can then define the Utilities from the left, the middle and the right which actually are the length of [0,L], [L,R] and [R,1].

$$Uleft = L = \frac{x+y}{2}$$
$$Umiddle = R - L = \frac{z-x}{2}$$
$$Uright = 1 - R = 1 - \frac{y+z}{2}$$

1)

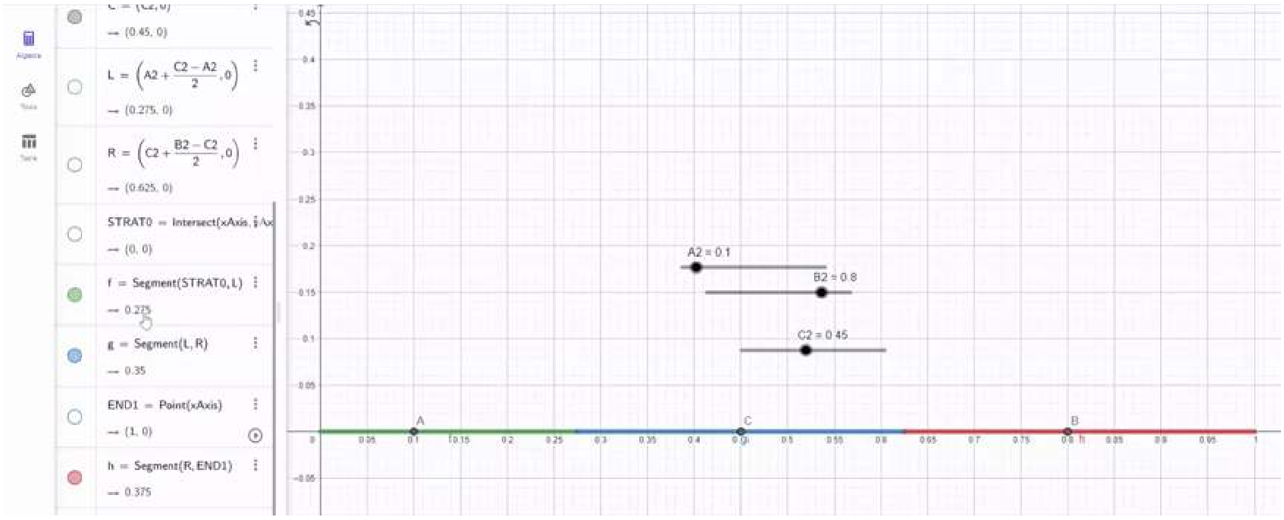
We know that tourists are uniformly distributed along the total length of the beach and will buy their ice-cream at the stall that is closest to their location. To help solve the problems we can make a quick drawing.

When Charlize (let's call everybody by letters A,B,C) arrives she is in front of the situation depicted right below in the first schema. In such a situation we can see that UB>UA because 0,55>0,45. We could have arrived at the same results by using the equations we set before.



Schema of the setting when C arrives

There are three possibilities for C. If C goes *lower* than A then she can appropriate ~0,1 of the length of the beach. Similarly, the maximum utility for choosing to go *higher* than Bob would be ~0,2 (1-0,8). The third possibility could be to go in between A and B. In such a configuration, C wouldn't benefit from a particular position because it would have a utility of 0,35 wherever in between A and B. As we can see on the simulation we've done on GeoGebra, the length of g (the Utility of C) given x and z is never changing. It's logical because we define g as the distance between the point in the middle of A and C, and the point in the middle of C and B. This distance reduce itself to $\frac{z-x}{2}$ which is independent from y and depend only on z and x coordinates.



In particular, we can also calculate the Utility of C in such a position :

$$Umiddle = \frac{0,8+0,1}{2} = 0,35$$

2)

We know that $a < b \leq 1$

There are three possibilities, either b is below 0,7, or above, or on it. Indeed, 0,7 is the point where C will not find any more utility being on the left of it than on the right. To calculate this point we have to find the place where the Utility given from being on the left side of B is equal to the utility given by being on the right side of B. In other terms :

$$\frac{z-x}{2} = 1 - z$$

$$\frac{3z-0,1}{2} = 1$$

$$3z - 2,1 = 0$$

$$z = 2,1/3$$

$$z = 0,7$$

- If b ≤ 0,7 then C should go on the right of B (the closest possible to B so ~b).
- If b = 0,7b = 0,7 then there is no distinction between being just at the right of B or wherever in between A and B.
- If b ≥ 0,7 then C should go wherever in between A and B.

3)

As we saw in the last question, 0,7 would be the place where C could not maximise her Utility by taking more from B. It's actually a **Nash Equilibrium** for B (x being given). Of course, because 0,9 / 2 > 0,1, B will want to choose a location where b >a. By choosing 0,7, B is thus minimising C utility by maximising his. Bob will always get a utility of at least 0,3 no matter what location Charlize chooses.

4)

Following the previous question, A knows how B will choose his location. Because of this, assuming A is going to choose a location where *a* < *b* will be true, A will want to choose *a* so that $\frac{1-a}{3} = a$; if *a* is greater than $\frac{1-a}{3}$, C will choose a location infinitely close to *a* where *c* < *a*, on the other hand if *a* is smaller, Alice will not have the maximum utility she could. From this follows that Alice should choose a location of $1 - a = 3a \Rightarrow 1 = 4a \Rightarrow a = 0,25$.