REPEATED GAMES – PRISONER'S DILEMMA

☛ *Example* – **Prisoner's Dilemma 1**

One of the interpretations:

It is 1930's. In the Soviet Union at that time a conductor travels by train to Moscow, to the symphony orchestra concert. He studies the score and concentrates on the demanding performance. Two KGB agents are watching him, who – in their ignorance – think that the score is a secret code. All conductor's efforts to explain that it is yet Tchajkovskij are absolutely hopeless. He is arrested and imprisoned. The second day our couple of agents visit him with the words: "You have better speak. We have found your comrade Tchajkovskij and he is already speaking . . ."

Two innocent people, one because he studied a score and the second because his name was coincidentally Tchajkovskij, find themselves in prison, faced the following problem: if both of them bravely keep denying, despite physical and psychical torture, they will be sent to Gulag for three years, then they will be released. If one of them confesses the fictive espionage crime of them both, and the second one keeps denying, then the first one will get only one year in Gulag, while the second one 25. If both of them confess, they will be sent to Gulag for 10 years.

The situation can be described by the bimatrix:

Tchajkovskij

Conductor		Deny	Confess
	Deny	$(-3, -3)$	$(-25, -1)$
	Confess	$(-1, -25)$	$(-10, -10)$

Dilemma – commonly it would be the most convenient for both to keep denying and go to Gulag for three years.
The problem: they have no chance to make a deal – and even if they had a chance to make a deal, there is a danger of comrade's confessing – whatever under a press or a temptation to take advantage of a shorter sentence. And even if both of them were solidary, each of them can think about the other that he falls prey to the temptation or a torture and confesses – hence he is in the danger of 25 year sentence which is even much worse than 10 years. Both therefore choose the second strategy and confess.

The strategy "confess" dominates the strategy "deny"

the pair

(confess, confess)

is the only equilibrium point in the game

☞ *Example* – Prisoner's Dilemma 2

More generally, prisoner's dilemma is a name for every situation of the type:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	(*reward, reward*)	(*sucker, temptation*)
	Defect	(*temptation, sucker*)	(*punish., punish.*)

where   *sucker $<$ punishment $<$ reward $<$ temptation.*

Cooperation can express whatever – the strategy pair
(*cooperate, cooperate*) corresponds to mutually solidary action

Examples of Occurence of Prisoner's Dilemma

- **Building the Sewage Water Treatment Plant**
  (two big hotels by one mountain lake):

  - *Cooperate* = build the purify facility
  - *Defect* = do not build it
  - *Reward* = pure water attracts tourists – customers, profits increase, nevertheless, we had to invest a certain sum of money
  - *Temptation* = take advantage of the purify facility of the second hotel and save on the investment
  - *Punishment* = polluted water discourages tourists, the profit decreases to zero

- **Duopolists:**

  - *Cooperate* = collude on the optimal total production (corresponding to monopoly)
  - *Defect* = break the deal
  - *Reward* = the highest total profit
  - *Temptation* = produce somewhat more at the expense of the second duopolist
  - *Punishment* = less profit for both

- **Removing the Parasites:**

  - *Cooperate* = mutual removing of parasites
  - *Defect* = have removing done by the comrade but do not return the favor
  - *Reward* = free of par., paying by removing other's
  - *Temptation* = free of par. without paying it back
  - *Punishment* = all are full of parasites which is much worse than a slight effort to remove the other's par.

- **Public Transportation:**

  - *Cooperate* = pay the fare
  - *Defect* = do not pay
  - *Reward* = public transportation runs, I can use it, nevertheless I have to pay a certain sum every month.
  - *Temptation* = use the public transportation, don't pay
  - *Punishment* = (almost) nobody pays, the public transportation is dissolved, I have to pay a taxi which is much more expensive than the original fare payment

- **Television Licence Fee:**

  - *Cooperate* = pay
  - *Defect* = do not pay
  - *Reward* = public service broadcast works, I can watch it, but I have to pay some small sum of money
  - *Temptation* = do not pay and watch
  - *Punishment* = (almost) nobody pays, the broadcast is dissolved

- **Battle:**

  - *Cooperate* = fight
  - *Defect* = hide
  - *Reward* = victory but also a risk of injury
  - *Temptation* = victory without a risk of injury
  - *Punishment* = the enemy wins without any fighting

- **Nuclear Armament:**

  - *Cooperate* = disarm
  - *Defect* = arm
  - *Reward* = the world without nuclear threat
  - *Temptation* = to be the only one armed
  - *Punishment* = all arm, pay much money for it, moreover a danger threats

Repeated Prisoner's Dilemma

In the case of infinite or indeterminate time horizon, *cooperate* is not necessarily irrational:

☛ *Example* – **Prisoner's Dilemma 3**

Consider the following variant of Prisoner's dilemma:

		Player 2	
		Cooperate	Defect
Player 1	Cooperate	$(3,3)$	$(0,5)$
	Defect	$(5,0)$	$(1,1)$

Imagine that the game will be repeated with the probability of $2/3$ in each round that the next round occurs, too.

When both players cooperate, the expected payoff for each:

$$\pi_C = 3 + 3 \cdot \tfrac{2}{3} + 3 \cdot \left(\tfrac{2}{3}\right)^2 + 3 \cdot \left(\tfrac{2}{3}\right)^3 + \cdots + 3 \cdot \left(\tfrac{2}{3}\right)^n + \cdots$$

Strategy in repeated game = a complete plan how the player will act in the whole course of the game in all possible situations in which he can find himself.

For example: **Grudger strategy:** *Cooperates until the second has defected, after that move defects forever.*

When two Grudgers meet in a game, they cooperate all the time and each of them receives the value $\pi_G = \pi_C$.

It can easily be proven that the pair of strategies

$$(Grudger,\ Grudger)$$

is an **equilibrium point** of the game in question.

Consider a **Deviant** who deviates from the **Grudger** strategy played with Grudger. In some round this Deviant defects, although the Grudger has cooperated (this can also happen in the first round). Let this deviation occurs first in the round $n+1$. Since the Deviant plays with the Grudger, in the next round the opponent chooses his strategy **defect** and holds on it forever. The Deviant can not therefore obtain more than

$$\pi_D = 3 + 3 \cdot \tfrac{2}{3} + \cdots + 3 \cdot \left(\tfrac{2}{3}\right)^{n-1} + 5 \cdot \left(\tfrac{2}{3}\right)^n + 1 \cdot \left(\tfrac{2}{3}\right)^{n+1} + \cdots$$

Since $\quad \pi_G - \pi_D =$

$$= (3-5) \cdot \left(\tfrac{2}{3}\right)^n + (3-1) \cdot \left(\tfrac{2}{3}\right)^{n+1} + \cdots + (3-1) \cdot \left(\tfrac{2}{3}\right)^{n+k} + \cdots$$
$$= -2 \cdot \left(\tfrac{2}{3}\right)^n + 2 \cdot \left(\tfrac{2}{3}\right)^{n+1} + \cdots + 2 \cdot \left(\tfrac{2}{3}\right)^{n+k} + \cdots$$
$$= \left(\tfrac{2}{3}\right)^n \left(-2 + 2 \cdot \tfrac{2}{3} \cdot \frac{1}{1-\tfrac{2}{3}}\right) = \left(\tfrac{2}{3}\right)^n \cdot 2 > 0\,,$$

it does not pay to deviate.

Similarly, we can consider the strategy **Tit for Tat**, which begins with cooperation and then plays what its opponent played in the last move. The pair

$$(Tit\ for\ Tat,\ Tit\ for\ Tat)$$

is an equilibrium point, too.

Examples of Strategies in Repeated Prisoner's Dilemma

Always Cooperates

Always Defects

Grudger, Spiteful: Cooperates until the second has defected, after that move defects forever (he does not forgive).

Tit for Tat: begins with cooperation and then plays what its opponent played in the last move (if the opponent defects in some round, Tit for Tat will defect in the following one; to cooperation it responds with cooperation).

Mistrust Tit for Tat: In the first round it defects, than it plays opponent's move.

Naive Prober: Like Tit for Tat, but sometimes, after the opponent has cooperated, it defects (e.g. at random, in one of ten rounds in average).

Remorseful Prober: Like Naive Prober, but he makes an effort to end cycles C–D caused by his own double-cross: after opponent's defection that was a reaction to his unfair defection, he cooperates for one time.

Hard Tit for Tat: Cooperates unless the opponent has defected at least once in the last two rounds.

Gradual Tit for Tat: Cooperates until the opponent has defected. Then, after the first opponent's defection it defects once and twice it cooperates, after the second

defection it defects in two subsequent rounds and twice it cooperates, . . . , after the n-th opponent's defection it defects in n subsequent rounds and twice it cooperates, etc.

Gradual Killer: In the first five rounds it defects, than it cooperates in two rounds. If the opponent has defected in rounds 6 and 7, than the Gradual Killer keeps defecting forever, otherwise he keeps cooperation forever.

Hard Tit for 2 Tats: Cooperates except the case when the opponent has defected at least in two subsequent rounds in the last three rounds.

Soft Tit for 2 Tats: Cooperates except the case when the opponent has defected in the last two subsequent rounds.

Slow Tit for Tat: Plays C–C, then if opponent plays two consecutive times the same move, plays its move.

Periodically DDC: Plays periodically: Defect–Defect–Coop.

Periodically SSZ: Plays periodically: Coop.–Coop.–Defect

Soft Majority: Cooperates, than plays opponent's majority move, if equal then cooperates.

Hard Majority: Cooperates, than plays opponent's majority move, if equal then defects.

Pavlov: Cooperates if and only if both players opted for the same choice in the previous move, otherwise it defects.

Pavlov P_n: Adjusts the probability of cooperation in units of $1/n$ according to the previous round: when it cooperated with the probability p in the last round, the probability of cooperation in the next round is

$$p \oplus \frac{1}{n} = \min(p + \frac{1}{n}, 1) \text{ if it obtained } R = \textit{reward}\,;$$

$$p \ominus \frac{1}{n} = \max(0, p - \frac{1}{n}) \text{ if it obtained } P = \textit{punishment}\,;$$

$$p \oplus \frac{2}{n} \text{ if it obtained } T = \textit{temptation}\,;$$

$$p \ominus \frac{2}{n} \text{ if it obtained } S = \textit{sucker}\,.$$

Random: Cooperates with the probability 0.5.

Hard Joss: Plays like Tit for Tat, but it cooperates only with the probability 0.9.

Soft Joss: Plays like Tit for Tat, but it defects only with the probability 0.9.

Generous Tit for Tat: Plays like Tit for Tat, but it after the defection it cooperates with the probability

$$g(R, P, T, S) = \min\left(1 - \frac{T - R}{R - S}, \frac{R - P}{T - P}\right).$$

Better and Better In n-th round it defects with the probability $(1000 - n)/1000$, i.e. the probability of defection is lesser and lesser.

Worse and Worse: In n-th round it defects with the probability $n/1000$, i.e. the probability of defection is greater and greater.

Occurrences of Repeated Prisoner's Dilemma (further examples)

- **Front Linie – Live and Let Live:**

  - *Cooperate* = live and let live

  - *Defect* = kill every man from the opposite side when the opportunity knocks

  - *Reward* = survival of long war years

  - *Temptation* = take advantage of the situation that the opponent is an easy chased and earn for example a medal – it is afterall better to remove the enemy

  - *Punishment* = all are upon the guard all the time . . .

- **Fig Tree and Chalcidflies:**

  - *Cooperate* = balanced ratio of pollinated flowers and flowers with layed eggs inside the fig

  - *Defect* = lay eggs to a greater number of flowers

  - *Reward* = genes spread

  - *Temptation* = lay eggs to a greater number of flowers and hence to encrease the number of offspring

  - *Punishment* = the fig hosting the treacherous Chalcidfly family is thrown down and the whole family dies out

- **Mutual Help of Males of Baboon Anubi:**

  - *Cooperate* = help the other male drive an enemy away during his mating

  - *Defect* = do not pay the help back

  - *Reward* = successful mating, offspring

  - *Temptation* = take advantage of help but do not pay it back and save the time and effort

  - *Punishment* = less offspring

Baboon Anubi

In the nature: the more often a male A supports a male B, the more the male B supports A.

- **Sexual Roles Alternating by Hermaphrodite Grouper:**

  - *Cooperate* = if I am a male now, I will became a female the next time

  - *Defect* = became a male again after acting a male

  - *Reward* = living together in harmony, many offspring

  - *Temptation* = repeat an easy male role

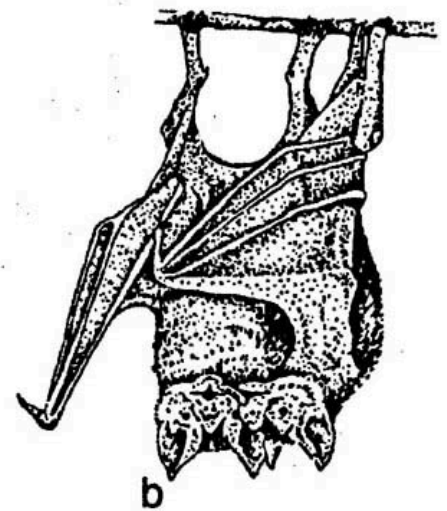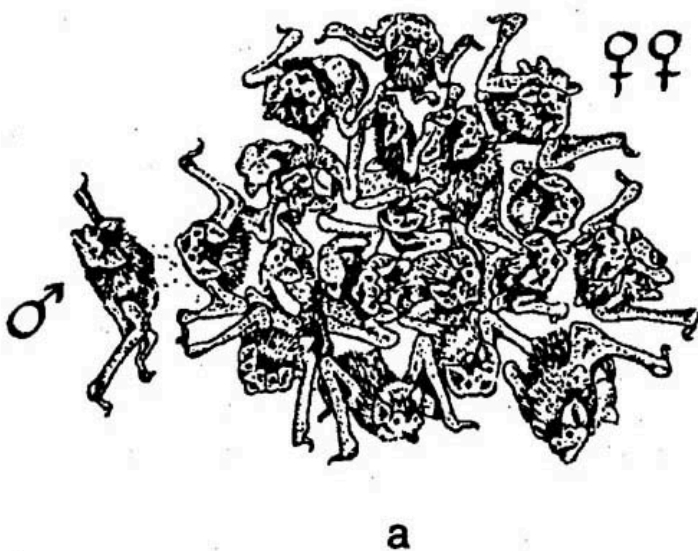 - *Punishment* = the relation breaks down
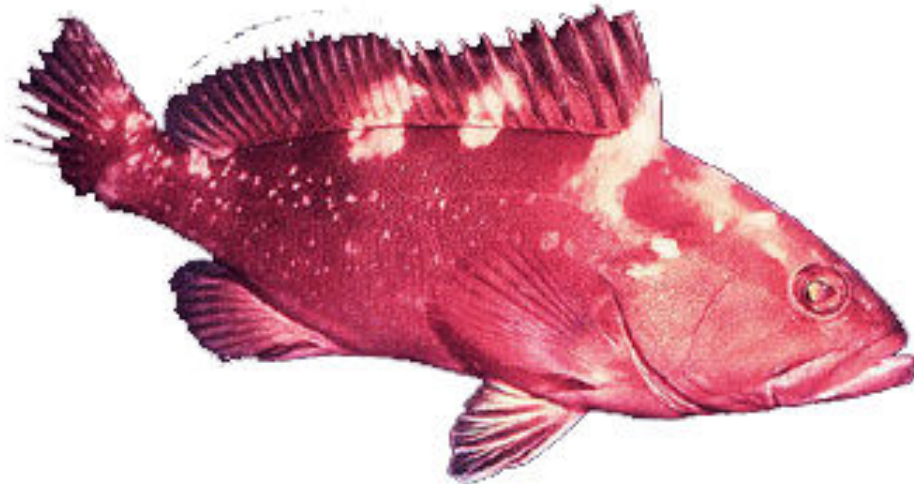
Red grouper (Epinephelus morio)

- **Desmodus Rotundus Vampire (a bat sucking mammal blood) – feeding hungry individuals:**

  – *Cooperate* = after a successful hunt, feed unsuccessful "colleagues"

  – *Defect* = keep all blood

  – *Reward* = long-run successful survival

  – *Temptation* = in the case of need, let the colleagues to feed me, do not share the catch with the others

  – *Punishment* = in the case of unsuccessful hunt, starving out

  In the nature: the individuals that have returned from a unsuccessful hunt are feeded by successful ones, even non-relatives; they recognize each other.

Desmodus Rotundus Vampires