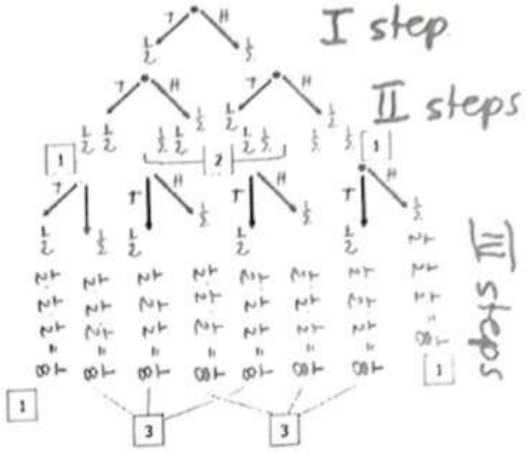
Pascal's triangle

1								1
1	1							2
1	2	1						4
1	3	3	1					8
1	4	6	4	1				16
1	5	10	10	5	1			32
1	6	15	20	15	6	1		64
1	7	21	35	35	21	7	1	128



Newton's Binomial

$(a+b)^0 = 1$

$(a+b)^1 = a + b$

$(a+b)^2 = a^2 + 2ab + b^2$

$(a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3$

$(a+b)^4 = a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4$

$(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$

$(a+b)^7 = a^7 + 7a^6b + 21a^5b^2 + 35a^4b^3 + 35a^3b^4 +$

$+ 21a^2b^5 + 7ab^6 + b^7$

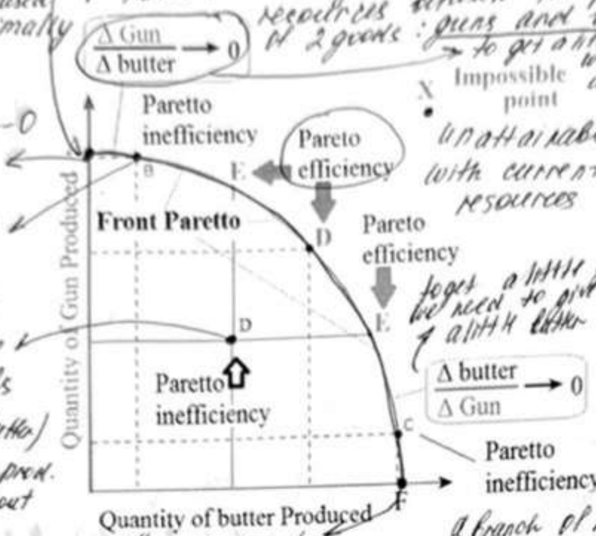
$(a+b)^8 = a^8 + 8a^7b + 28a^6b^2 + 56a^5b^3 + 70a^4b^4 +$

$+ 56a^3b^5 + 28a^2b^6 + 8ab^7 + b^8$

5 O-we can't increase the production of a one good, not decrease production of another. Resources are used fully and optimally

curve line represents the production possibilities. It illustrates the maximum achievable combinations of 2 goods produced when all available resources are fully and efficiently utilized.

classic economic model / shows how society allocates scarce resources between the production of 2 goods: guns and butter



to get a little butter, we need to give a little gun

unattainable with current resources

all resources to guns, butter – 0

resources are used not fully

can produce more goods (guns and butter)

! can increase prod. of both without loss

to get a little gun, we need to give a little butter

all resources to butter, guns – 0

by Vilfredo Pareto describes the Pareto frontier on a graph with 2 axes (payoff for player 1 and 2)

The orange sector E-D-E is the most Pareto efficient - since an increase in one indicator leads to a decrease in another.

! This reflects the law of increasing opportunity costs: the more of 1 good we produce, the higher the cost of producing each other addit. unit

оба сознаются – по 5 лет

если оба молчат – по 1 году

раз-бол-нем. – сознались

a branch of math studying strategic interactions, where each participant's outcome depends on the actions of others

Game Theory

Nash Equilibrium – a state where no player can improve their payoff by changing their own strategy alone, assuming all other players keep their strategy unchanged

This is the main concept in game theory ** => **Nash equilibrium**

$H_2(x)$ / $H_1(x)$	Player 2	Recognition;	Non-recognition;
Player 1		1	2
Recognition;	1	* -5 / -5 *	-20 / 0
Non-recognition;	2	0 / -20	-1 / -1

-1-1 Pareto Optimality

Индив. рациональность не всегда приводит к коллективно оптимальному результату.

член семьи

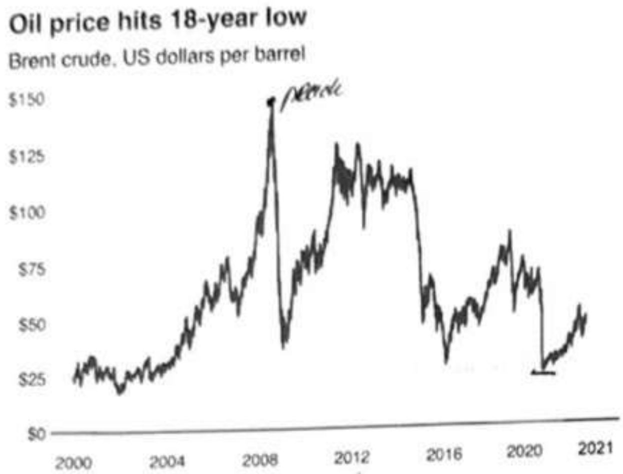
рецидивист / предатель

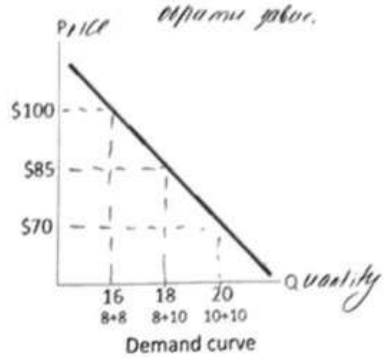
должны молчать (omerta - кодекс молчания)

! для мафии. по убьют за предательство.

если 1 сознается (какой Penito), он выходит на свободу, а 2 лет зонат.

• Mafia creates external incentives (threat of death for betrayal) that force members to remain silent, even of confessing is beneficial from the point of view of the law. This changes the game — now staying silent becomes a rational choice.





Все по 8 млн
цена: 100 $
Р: 800 в день
А: 800 в день
Итог: 1600

100 $ → 85 $ (из-за увел. предложения)
Р: 850
А: 680

(Дол. сокр. добычи) [illegible] спроса из-за [illegible]. В реальности: в [illegible] году Россия отказ. сократ. добычу, что привело к краху. В ответ Аравия начала ценовую войну, резко увеличив добычу и снизив цены на нефть. Это вызвало обвал цен на 30% за 1 день.

Barrel	1. $8 \cdot 10^6$ barrel/day	2. $10 \cdot 10^6$ barrel/day
1. $8 \cdot 10^6$ barrel (объём добычи 8 млн)	$800 millions per day; $100; $800 millions per day	$850 millions per day; $85; $680
2. $10 \cdot 10^6$ barrel/day (10 млн)	$680 millions per day; $85; $850 millions per day	$700 millions per day; $70; $700 millions per day

один из участн. переговоров

Р: 680
А: 850

↓ → 70 $
Р: 700
А: 700
Итог: 1400 млн The worst

- Оптимально для обеих: сотрудничать (производить по 8 млн) и пол. по 800 $ млн.
- Но у каждой есть стимул увел. добычу до 10 млн, чтобы зар. больше
- В результате обе увел. → цены пад. → обе получ. меньше

Это класс. пример неэффект. равновесия Нэша в теории игр: индивид. рациональность приводит к коллективно худш. рез.

4 The main idea: only discard the 3rd method if it fails to beat every previous method (not the 1st one). If it beats the worst one, it has minimal value — worth as a trade-off.

The 1st method (simple sequential guessing) is bad.

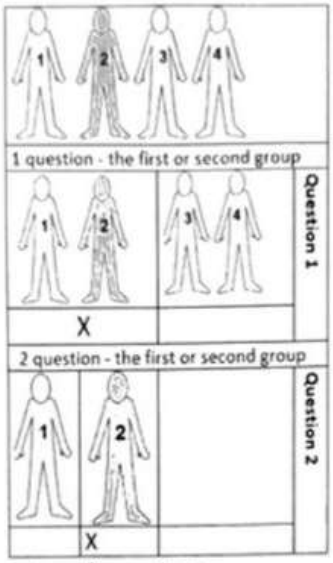
The 2nd method (optimal questions) is better if it reduces the average number of questions

a comprom

If repetition rule prevents us from rejecting a strategy too early

adaptive probability based

a uniform distribution
All 4 people are equally likely (prob 0,25 each)
If you ask optimal questions (halving the set each time), any person requires exactly 2 questions.
Average 2



Average number of questions =
2*0.25+ 2*0.25+ 2*0,25+ 2*0,25 =2

prob.

Average number of questions =

1*0.5+	2*0.25+	3*0,125+	3*0,125

in exp. of the number of quest

It is implied that there are 4 possible people with dif p

Question 1. Is this Zuckerberg?	50%	1*0.5
Question 2. Is this Sergey Brin?	25%	2*0.25
Question 3. Is this Stefan from BMW?	12,5%	3*0,125
So Prince Saud	12,5%	3*0,125
Average number of questions =1,75		

If one option is very likely (Zuckerberg with 50%) it's best to ask about it first this gives a gain in the average

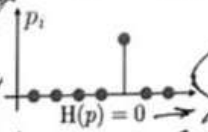
Shannon entropy

$$\sum_{i=1}^{n} p(i) \log_2 \frac{1}{p(i)}$$

$H = 0,5 \cdot \log_2(1/0,5) + 0,25 \cdot \log_2(1/0,25) + 0 = 0,5 \cdot 1 + 0,25 \cdot 2 + 0,125 \cdot 3 + 0,125 \cdot 3 = 0,5 + 0,5 + 0,375 + 0,375 = 1,75$ bits

not the average number of quest, but the lower bound

- $H(X)$ - entropy of random var. X. this is the average amount of inform
- $p(x_i)$ - probab of the i-th event
- $\log_2 \frac{1}{p(x_i)}$ - inf. content of the event (how many bits are needed to encode it optimally.



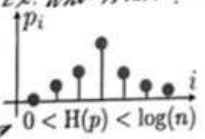
Quantifying information

when one event happ. 1 all others have probability 0.

Ex: Who is here? - Always Zuckerberg. - No quest. needed (0 bits of inform.)

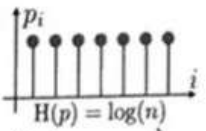
$I(x_i) = \log_2\left(\frac{1}{p_i}\right)$ → self-inform of a specific event x_i.

the lower the probability the more informat.



number of bits required to encode choice

Mark Zukerberg	Sergey Brin	Stefan Quandt	Prince Al Saud
P(1)= 50%	P(2)= 25%	P(3)= 12,5%	P(4)= 12,5%



$$\sum_{i=1}^{n} p(x_i)\, I(x_i)$$

the definition of entropy - the weighted average of information

$H(p) \neq 0$ - if not all probab are 0 or 1 (there is uncertainty)

- $H(p) < \log(n)$ - if the distribution is non-uniform
- $H(p) = \log(n)$ - if all probabilities are equal ($1/n$ - max uncertainty)

$n = 4$, $\log_2(4) = 2$ Entropy = $1,75 < 2$ - the distribution is non-uniform.

Entropy measures uncertainty.
If one outcome is certain ($p=1$, others 0) → $H = 0$

— Entropy is the minimum average number of bits per symbol under optimal encoding

the maximum possible entropy for a random variable with n possible outcomes

! it happens when all outcomes are equally probable. $p_1 = p_2 = \ldots = p_n = \frac{1}{n}$

If outc. are equal → max uncert. more outc.

→ quant. of bites in code   binary tree

$L = \sum_{i=1} p_i \cdot length(code_i)$

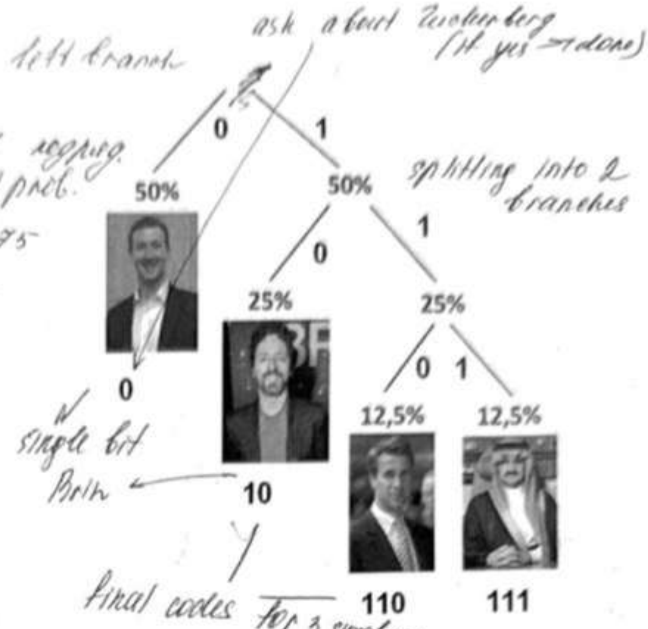
↓ probab. of symbol i

средняя длина кода (сколько битов в среднем нужно, чтобы закодировать один символ, если кодировать много символов подряд)

total prob.

$L = 0{,}5 \cdot 1 + 0{,}25 \cdot 2 + 0{,}125 \cdot 3 + 0{,}125 \cdot 3 = 1{,}75$

частоты символов   в среднем на один символ



First-order approximation (symbols independent but with frequencies of Belarusian txt).

Мама мыла ра

		index ranges
М - 3	30%	1-3 М
а - 4	40%	4-7 а
ы - 1	10%	8 - ы
л - 1	10%	9 - л
р - 1	10%	10 - р
10		

лла мам ма р

Second-order approximation (digram (2-symbols) structure as in Belarusian)

Мама мыла ра

		assigning numbers to diagrams	
Ма - 2	22%	1-2	ма
ам - 2	22%	3-4	ам
мы - 1	11%	5	мы
ыл - 1	11%	6	ыл
ла - 1	11%	7	ла
а р - 1	11%	8	ар
ра - 1	11%	9	ра
9			

0. 4 6 7 3 1 9 1 6 7 3 5

ам ыл ла ам ма ра ма ыл ла ам мы

мылла   рама

better frequency → probab. → entropy of text

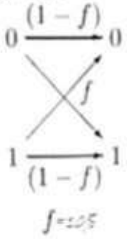
3

Sir Dr. D. MacKay,
University of Cambridge
(22 April 1967 – 14 April 2016)

Richard Hamming (1915 - 1998)

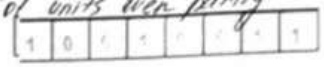
"I believe in clean energy,
but I also believe in mathematics"

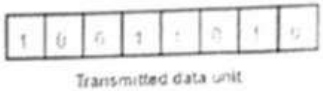
emphasizes that even in applied things you need strict math

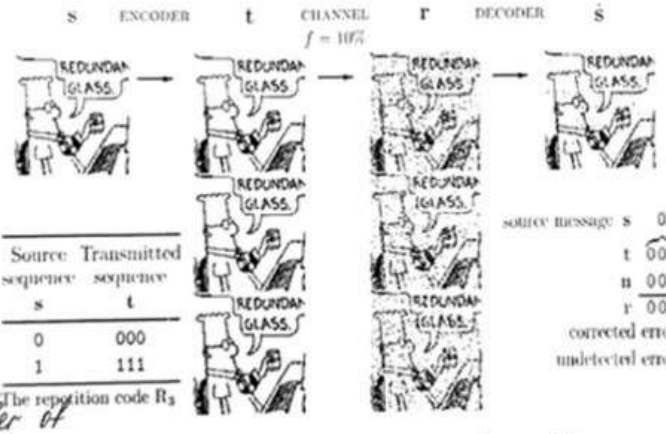


$f = 10\%$

add to the block with data to make the quint 0 units even parity







Source sequence s	Transmitted sequence t
0	000
1	111

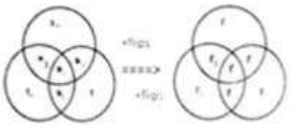
The repetition code R_3

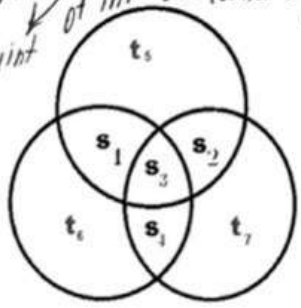
source message	s	0	0	1	0	1	1	0
	t	000	000	111	000	111	111	000
	n	000	001	000	000	101	000	000
	r	000	001	111	000	010	111	000
corrected errors			*					
undetected errors							*	

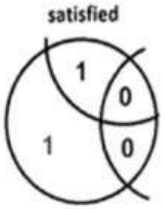
total number of bits in the code word (4 data + 3 parity control bits) quint

7,4. Hamming code.

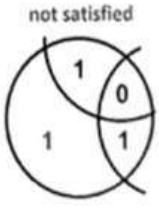
$\frac{4}{\Sigma} \rightarrow \frac{7}{t}$ of int. bits (data we give)



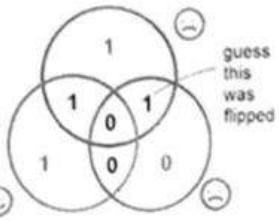




1+1+0+0=**0**



1+1+0+1=1



+0.1