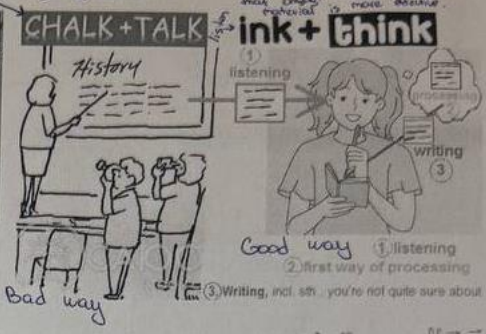
80% chance of rain in Oxford. In IT the information content of a message is $-\log(p)$. Higher probability means less information. The entropy of the source is the average information. This example links probability, statistics, and inf. theory.

+0,1 11.2.26

Represents traditional teaching methods, which might be less effective.

+0.1 18.2.26 ABBA

A better learning method. This suggests that engagin more actively with material is more effective.

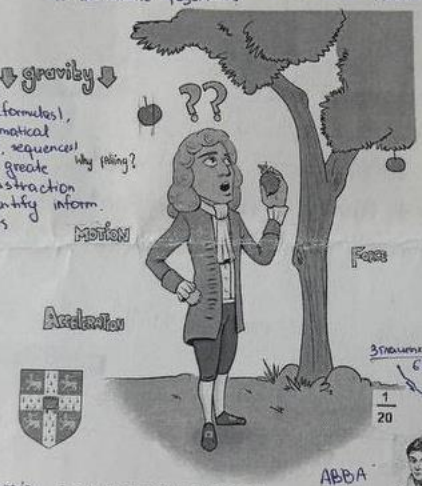


The contrast of school and university education.
School = formalism: memorize formulas and symbols without deep understanding.
University = understanding: know meaning, connections, why

School ⇩ gravity ⇩ MOTION ==formalism==> University

CONCRETE AND ABSTRACT THINKING

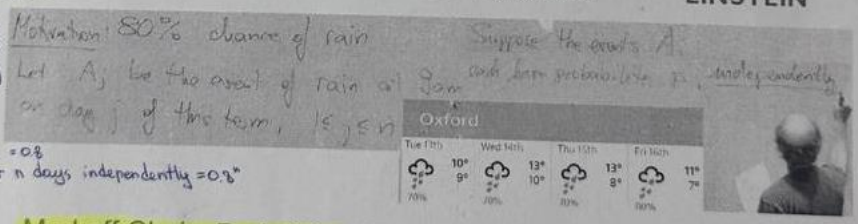
From concrete observations (gravity, motion) to abstract mathematical laws (formulas), then to pure mathematical structures (factorials, sequences) and the names of greate thinkers. In IT, abstraction allows us to quantify inform. independently of its physical form.



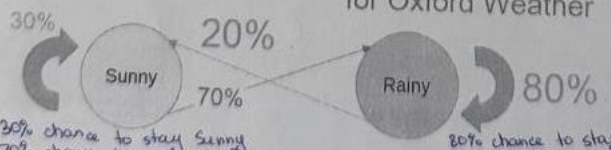
ISAAC NEWTON

ALBERT EINSTEIN

If the probability of rain each day is independent and identical, this implies that the occurrence of rain on one day doesn't influence its occurrence on another day.
Single day prob = 0.8
Prob. of rain for n days independently = 0.8^n



Markoff Chain Probability Model for Oxford Weather



30% chance to stay Sunny
70% chance to switch to Rainy

80% chance to stay Rainy
20% chance to switch to Sun

The next state depends only on the current state. Used in IT for modeling correlated sequences and computing entropy rates.

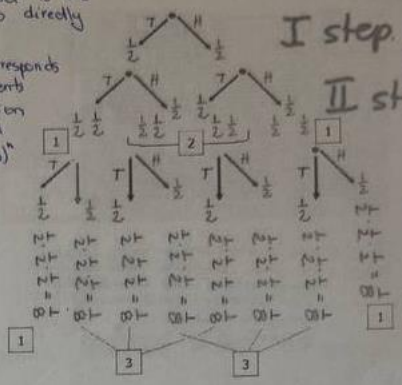
Pascal's triangle

Pascal's Triangle is a triangular array of numbers where each number is the sum of the two directly above it.

Each row corresponds to the coefficients in the expansion of a binomial expression $(a+b)^n$

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
1	8	28	56	70	56	28	8	1	← 256

a tree diagram to illustrate the step-by-step construction of Pascal's Triangle.



Each node splits into two branches, summing the values from the previous step.

The binomial theorem states that $(a+b)^n$ can be expanded using the coefficients from Pascal's Triangle.

The general term in the expansion is given by $\binom{n}{k} a^{n-k} b^k$, where $\binom{n}{k}$ is a binomial coefficient.

Newton's Binomial

$$(a+b)^0 = 1$$

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

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

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

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

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

$$(a+b)^6 = a^6 + 6a^5b + 15a^4b^2 + 20a^3b^3 + 15a^2b^4 + 6ab^5 + b^6$$

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

Every C# program must be inside a class.

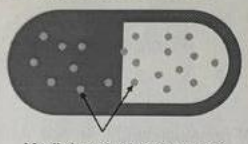
a template that groups data and methods together

This is the program we need to write today

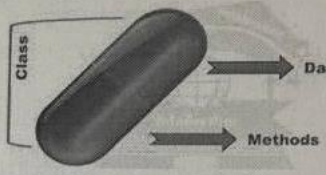
```
class ABBA
 {
 static void Main(string[] args)
                . // Here's a method called Main.
 {
    System.Console.WriteLine("ABBA!");
  }
 }
```

So there's the keyword class. Unlike C++, in C# all code must be placed in a class.

Encapsulated in a class.



Medicines Inside Capsule



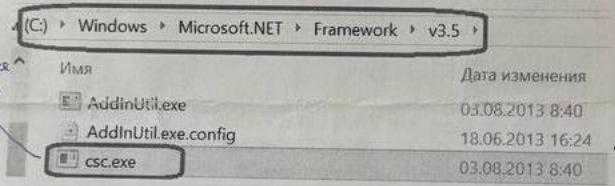
A ≠ a

C# is case sensitive

in + Number;
in + number;
these are 2 different variables

C:\WINDOWS\Microsoft.NET\Framework\v3.5\ csc.exe

ABBA.cs → csc.exe → ABBA.exe



Complier – translates source code (.cs) into executable code (.exe)

ABBA.cs ⇒ csc.exe = ABBA.exe

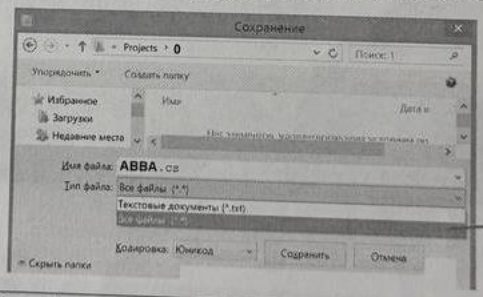
C# Compiler

Step 1. And on my HDD, I also make a folder with the same name D:\ IT

Step 2. In the folder **E:** IT\ we make the folder of the **Projects** - **E:\ IT\Projects**

And in the **Project** folder make folder **0** - **E:\ IT\Projects\0** where our today's practical work will be stored

Step 3. As I mentioned above, C# is a built-in language of Windows. Notepad is enough to write a program



You need to switch from *.txt (Text documents) to *.* (all files)

Otherwise, notepad with *.txt extension

Step 4. Entering command mode

Start=>Run=>cmd

cd E: - After that go to the folder IT/Projects/0/

cd IT -Then go to the folder Projects

cd projects - Then go to the folder 0

cd 0 –

```
E:\>cd IT
E:\IT>cd Projects
E:\IT\Projects>cd 0
E:\IT\Projects\0>
```

1. write code → 2. save (.cs) → 3. compile → 4. errors box → 5. run → 6. best output

$n! = n \cdot (n-1) \cdot \ldots \cdot 1$

$1! = 1$ $4! = 24$

$3! = 6$ $5! = 120$

Step 5.

Now we need to compile the file using the compiler csc.exe which is in the folder

C:\Windows\Microsoft.NET\Framework\v3.5

C:\WINDOWS\Microsoft.NET\Framework\v3.5\csc.exe ABBA.cs

the result is a file ABBA.exe that can already be run (which is located in the same folder). If you have taken the 1st step, then this means that the education-process has begun. This is victory. csc /target:library ABBA.cs - will make ABBA.dll.

Step 6. Modify the file as follows

```
using System;

class ABBA
{
  static int Factorial(int n)
  {
    if (n == 1) return 1;
    return n * Factorial(n - 1);
  }

  static void Main(string[] args)
   // Here's a method called main.
  {
      System.Console.WriteLine("ABBA -"+
                                Factorial(4));
  }
}
```

prevents infinite recursion

```
>C:\WINDOWS\Microsoft.NET\Framework\v3.5\csc.exe ABBA.cs

>ABBA
ABBA – 24
```

V - number of vowels
C - number of consonants

$V! \cdot C!$ - number of possible arrangements

Step 7. Modify the file as follows

```
using System;

class ABBA
{
  static int Factorial(int n)
  {
   if (n == 1) return 1;
    return n * Factorial(n - 1);
   }


public static void Main()
 {
  System.Console.WriteLine("ABBA");
  System.Console.Write("Enter number of Vovels:");
  string s = Console.ReadLine();
  int vovels=int.Parse(s);
  System.Console.Write("Enter number of Consonants:");
  s=System.Console.ReadLine();
  int consonants=int.Parse(s);
  Console.WriteLine(Factorial(vovels)*Factorial(consonants) );
  }
}
```

```
>ABBA
ABBA
Enter number of Vovels:2
Enter number of Consonants:2
4
```

```
>ABBA
ABBA
Enter number of Vovels:3
Enter number of Consonants:3
36
```

1 2 3 4

+0.1 12.9 +0.1 16.2

+0.1 5.3.26

Resume of Lecture by Pr. Bob Gallagher from MIT

Massachusetts Institute of Technology (MIT)

George Boole (1815-1864) developed Boolean logic

The principles of logical thinking have been understood (and occasionally used) since the Hellenic era.

Boole's contribution was to show how to systemize these principles and express them in equations (called Boolean logic or Boolean algebra).

Claude Shannon (1916-2001) showed how to use Boolean algebra as the basis for switching technology. This conribution systemized logical thinking for computer and communication systems, both for the design and programming of the systems and their applications.

Logic continues to be abused in politics, religion, and most non-scientific areas.

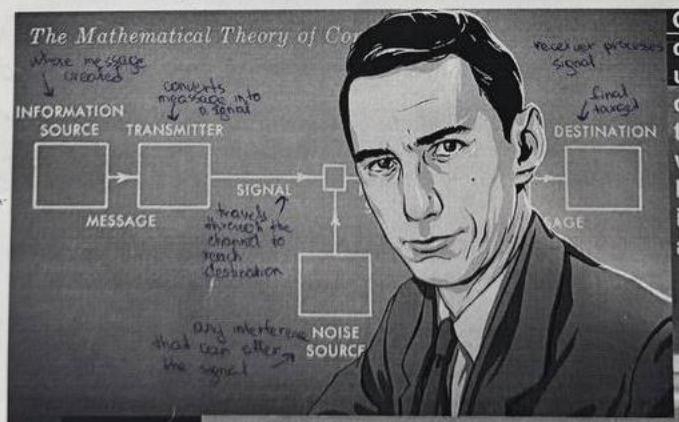
A little nationalistic, but this is an sample of right logic

Kant, Gauss, Goethe are great ⇒ Kant, Gauss, Goethe - Germans ⇒ Germany Great

The logical structure is valid and follows a clear, correct reasoning path.

Bad logic (abuse of logic)

An individual incorrectly generalizes personal attributes based on group identity.



a diagram of a communication system.

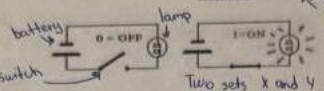
A private land-grant research university in Cambridge, established in 1861 in technology and science

Massachusetts Institute of Technology (MIT)

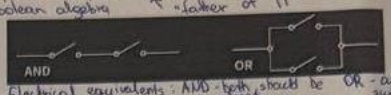
Lecture by Pr. Bob Gallagher

Boole (1815-1864) & **Shannon** (1916-2001)

↑ Boolean algebra ↑ "father of IT"



Two sets X and Y



Electrical equivalents: AND - both should be closed. OR - at least one switch must be closed.

Boolean symbols
∧ - and
∪ - or
¬ - not

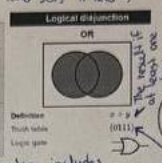
Logical addition (disjunction)

A	B	F=A∨B
0	0	0
0	1	1
1	0	1
1	1	1

A∨B is true if either A or B or both are true.

A	B	A∨B
True	True	True
True	False	True
False	True	True
False	False	False

Logical disjunction



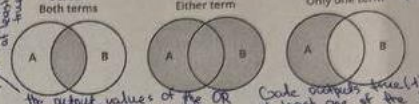
OR operation includes all elements that are in either set X, set Y, or in both.

BOOLEAN LOGIC - Venn Diagrams (visual representation)

AND Both terms

OR Either term

NOT Only one term



the output values of the OR operation for all possible combinations of inputs

Gate outputs true(1) when at least one of the inputs is true.

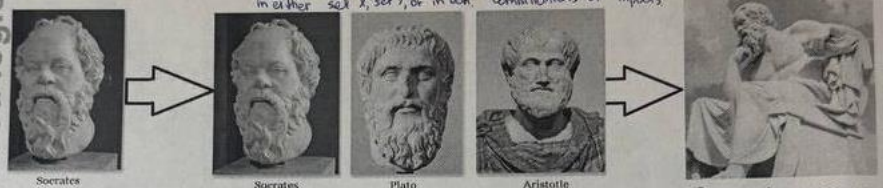
Good logic:

Knowing Socrates is a philosopher and philosophers are men, it correctly that Socrates is a men

Bad logic:

From the fact that Socrates is a man and philosophers are men, it wrong that Socrates is a philosopher.

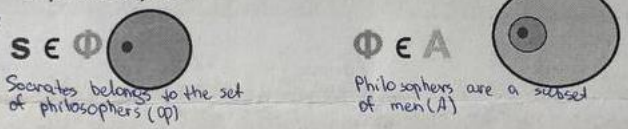
Good logic



Socrates was a philosopher

philosophers are men

Socrates was a man



S ∈ Φ

Socrates belongs to the set of philosophers (Φ)

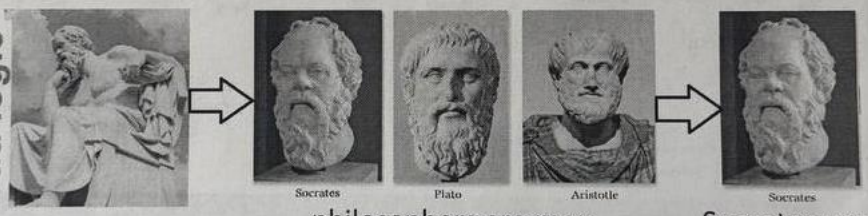
Φ ∈ A

Philosophers are a subset of men (A)

S ∈ A

Socrates belongs to the set of men

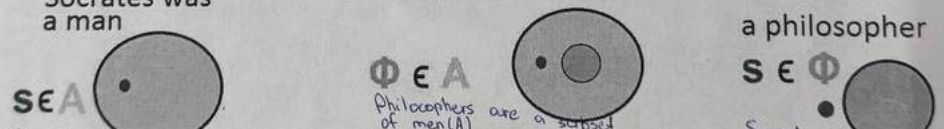
Bad logic



Socrates was a man

philosophers are men

Socrates was a philosopher



S ∈ A

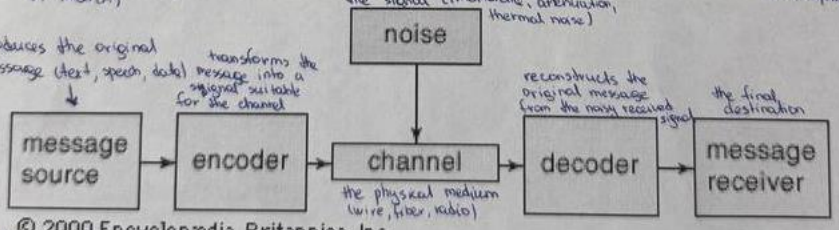
Socrates belongs to the set of men (A)

Φ ∈ A

Philosophers are a subset of men (A)

S ∈ Φ

Socrates belongs to Philosophers.



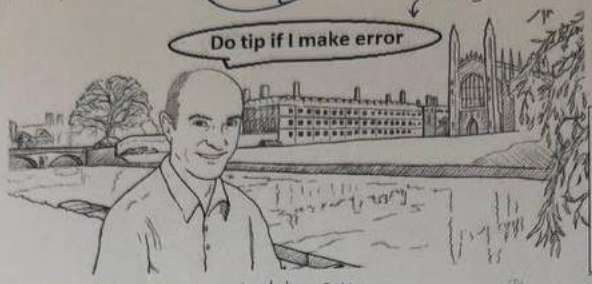
What this model gives us:

1. Channel capacity - the fastest speed you can send inf. without making errors.
2. Source entropy - how much uncertainty or surprise the source has.
   Mutual inf. - how much the output tells you about the input.
3. Shannon's two big theorems:
   - Source coding theorem: you can't compress a message below its entropy (lossless compression limit)
   - Channel coding theorem: you can send reliably up to the channel capacity, even with noise (error-correcting codes exist)

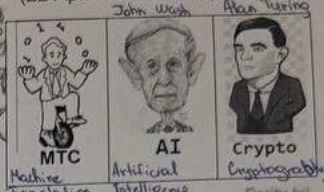
+0.1 +0.1 +0.1 +0.1

Sir Dr. D. MacKay - a noteable figure in IT and error correction.

Willingness to be corrected



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



Redundant data helps ensure accuracy and reliability in the presence of errors.

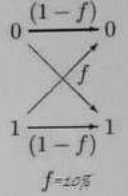
His dual commitment to both fields

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

Parity bits for error detection

1	0	0	1	0	0	1	1

Transmitted data unit



Parity bit

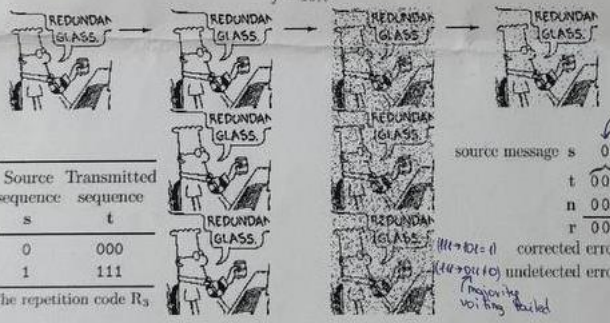
1	0	0	1	1	0	1	0

Transmitted data unit

+0.1

1) Source code: The original data either 0 or 1
2) Encoding: each bit is repeated

s → ENCODER → t → 3] CHANNEL f = 10% → r → 4) reception 5) Decoding DECODER → ŝ



Source sequence s	Transmitted sequence t
0	000
1	111

The repetition code R_3

repeat each bit multiple times to create redundancy

Also Process of error correction:

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
		*			*		

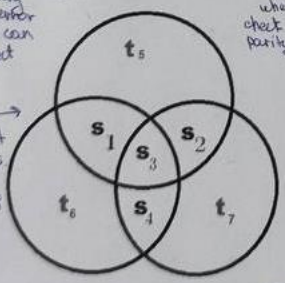
corrected errors
undetected errors
1 error
2 error
majority voting failed

7.4. Hamming code.

$$\frac{4}{\Sigma} \rightarrow \frac{7}{t}$$

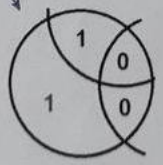
A type of binary code used for error correction that can detect and correct single-bit errors

represent different sets of parity bits. The intersections - the data bits being protected by the parity bits.



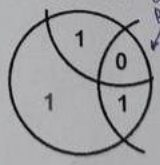
when the parity check results in Even parity (sum of bits is even)

satisfied



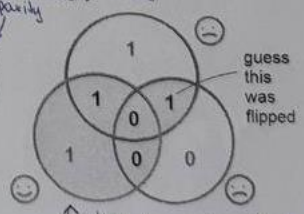
1+1+0+0=0

not satisfied



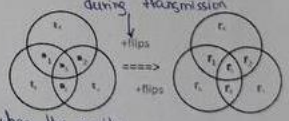
1+1+0+1=1

when the parity check results in odd parity

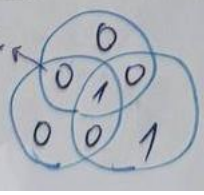


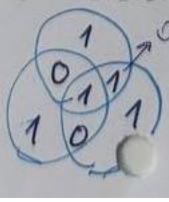
How the overlaps help in pointing the errors.

potential bit flips during transmission



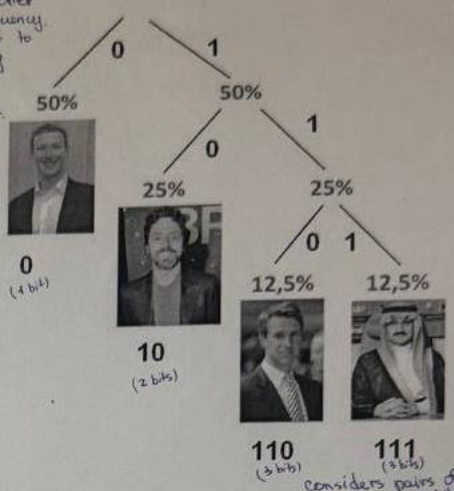
+0.1





+0.1 Flip

Huffman Tree: The tree shows the process of creating a Huffman coding tree. Each character is assigned a binary code based on its frequency. Characters with higher frequencies are closer to the root, resulting in shorter codes. The binary codes are created by traversing the tree: left branches are 0 and right branches are 1.



assumes that each symbol is independent of the others, meaning the prob of each symbol appearing doesn't depend on previous symbols.

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

Мама мыла ра

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

лламмар

breaks down the sentence into individual symbols using 1-order approximation

considers pairs of symbols where the probability of a symbol appearing depends on the previous symbol.

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

Мама мыла ра

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
ар - 1	11%	8 ар
ра - 1	11%	9 ра

9

numbers represent positions in the text

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

using 5-order ap.

Higher-order models better approximate real language entropy. Important for source coding and redundancy estimation.

* Entropy ($H(x)$) - quantifies the average amount of information ($I(x)$) - measures how surprisingly the information event may be.
  - If every outcome is equally likely => Entropy is high / lots of unpredictability.
  - If a outcome is much more likely than others => Entropy is low (more predictability)

+0.1 4.5

Say **NO** to the **first**

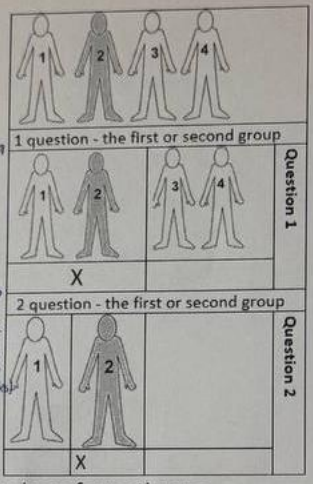
Say **YES** to the **second** if it is better than the first

Say **NO** to the **third** only if it is worse than all the others

explains how to determine the average number of questions needed to identify a specific person from a set of candidates based on their prob.

Average number of questions =

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



The decision-making process involves comparing these candidates in pairs to determine the best option.

Inefficient strategy: always ask the same two questions for every person. Everyone takes 2 questions → average = 2 (regardless of probabilities). Worse than the optimal strategy (right side) which uses adaptive questions and achieves average = 1.75 = entropy.

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

Average number of questions =

2*0.25+ 2*0.25+ 2*0,25+ 2*0,25 =2

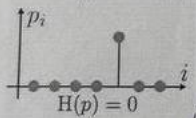
$$H(X) = \sum_{i=1}^{n} p(x_i) \log_b \frac{1}{p(x_i)}$$

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

This formula calculates the entropy which is the measure of uncertainty or inf. content

$p(i)$ - prob. of occurrence of event i.

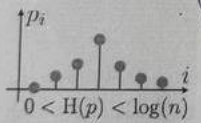
The formula takes into account the prob. of each candidate being the correct one and the number of questions required to identify them.



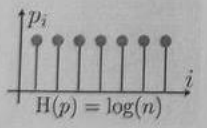
Quantifying information

The formula gives the inf. content $I(x_i)$ of a single event x_i, which is the number of bits required to encode that event.

$$I(x_i) = \log_2\left(\frac{1}{p_i}\right)$$



number of bits required to encode choice



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

50 billion	25 billion	12.5 billion	12.5 billion
			
Mark Zuckerberg P(1)= **50%**	Sergey Brin P(2)= **25%**	Stefan Quandt P(3)= **12,5%**	Prince Al Saud P(4)= **12,5%**
1 bit	2 bits	3 bits	3 bits

I(Mark) = $\log_2(\frac{1}{0.5}) = 1$ bit

00000 = $\frac{5}{10}$ = 0.5

000 = $\frac{3}{10}$ = 0.3

0 = $\frac{1}{10}$ = 0.1

0 = $\frac{1}{10}$ = 0.1

Кол-во	вопросов:
4 - 1	0.5
0 - 2	0.6
9 - 3	0.3
5 - 3	0.3

=> 1.7

32 → f(x) → 5
64 → f(x) → 6

$\log_b(n) = x <=> b^x = n$

argument, base, exponent

base, exponent, argument

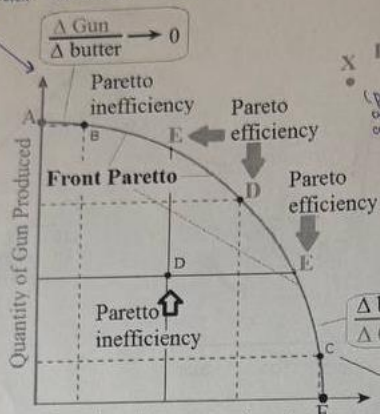
Pareto efficiency (or Paretreto optimality) is a state where resources are allocated in the most efficient manner, such that any change to make one individual better off would make another worse off.

by Vilfredo Pareto

1848-1923

The trade-off between two goods (guns and butter). Points on the curve represent efficient allocations of resources.

Comment



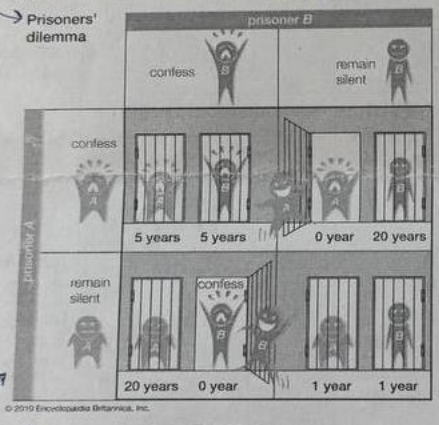
X Impossible point (points outside the curve are unattainable with current resources)

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

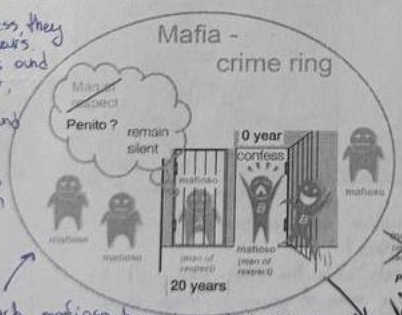
(because there are unutilized resources, and moving to a point on the curve would make at least one person better without making anyone worse off)

A standard example in game theory that shows why two rational individuals might not cooperate, even if it appears to be in their best interest.

Prisoners' dilemma



- if both confess, they both get 5 years.
- if one confess and the other doesn't, the confessor gets 0 years and the silent one gets 20 years.
- if both remain silent, they both get 1 year.



Each mafioso has two choices: remain silent or confess (betray the other)

1) If both remain silent: they are considered "man of respect", receive a lesser punishment (lack of evidence)

2) If one confess, and the other remains silent: the one who confesses is rewarded with freedom and is labeled as "Pentito" who remains silent - receives the full punishment 20 years.

Game Theory

Nash Equilibrium

** => Nash equilibrium

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

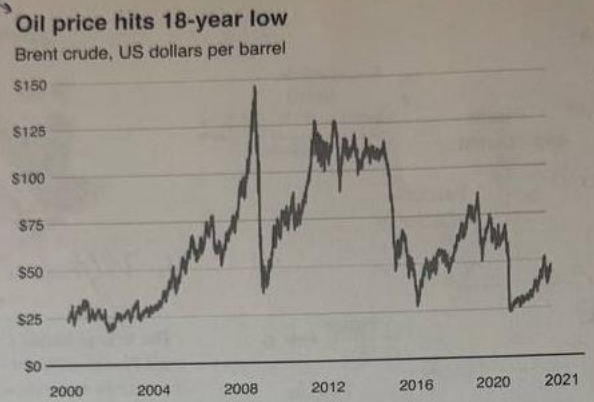
-1-1

- if both recognize, the payoff is 5 each.
- if one recognizes and the other doesn't, payoffs vary (20 for non-recognition, 0 for recognition)
- if neither recognizes, the payoff is 1 each.

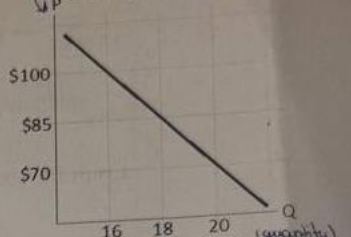
Pareto Optimality

Nash equilibrium - a concept of game theory where no player can benifit by changing their strategy while the other players keep theirs unchanged.

Shows the historical price of Brent crude oil from 2000 to 2021. It highlights significant fluctuations (rarefaction), including a peak around 2008 and subsequent variations.

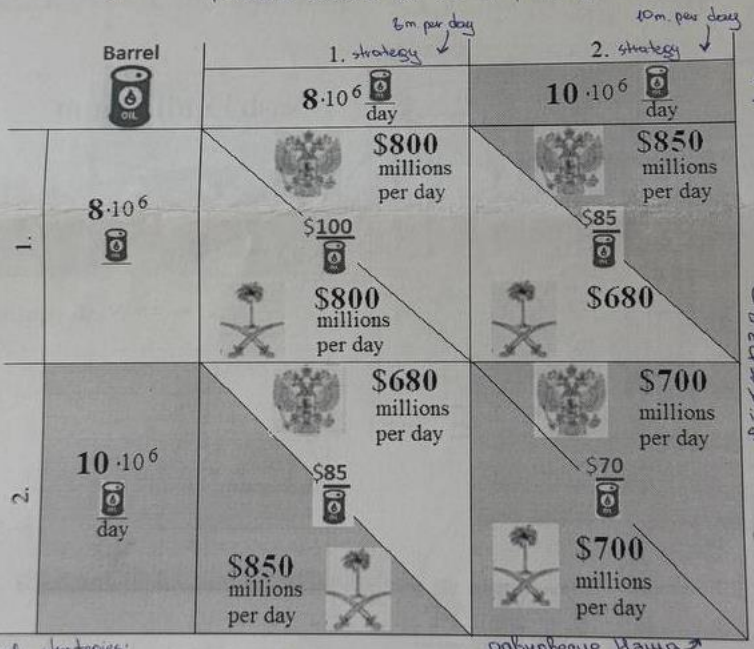


As the quantity demanded increases, the price decreases, following the law of demand.



Demand curve for oil

Game theory matrix: a payoff matrix used to analyze the strategies and potential outcomes for two oil-producing entities.



two players and their strategies.

Caricature of John Nash an american mathematician known for his work in game theory, which earned him the Nobel Prize in Economic Sciences in 1994.

равновесие Нэша

Optimal strategies:

- If both players cooperate and limit their production to 8m. barrels per day, they maximaze their joint revenue (1600 m. per day)
- If one player increases production to 10 m. barrels while the other stays at 8m, the total revenue decreases, but the player producing more oil gains a larger share.
- If both increase production to 10 m barrels per day, the price drops significantly, reducing revenue for both (1400 m total)

Game theory implications: this scenario can be analyzed using game theory to understand the Nash Equilibrium, where each player chooses the best strategy, considering the other player's choices.