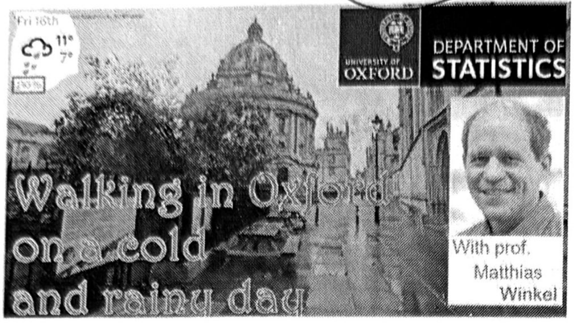
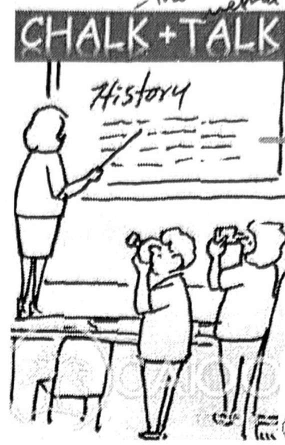
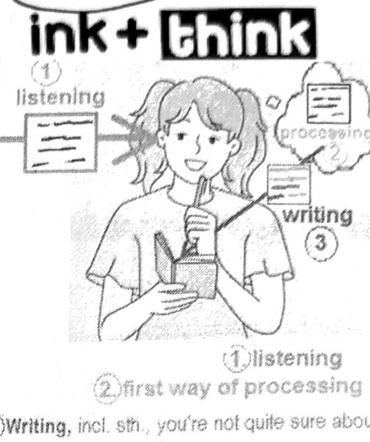
+0.1 +0.1 +0.1 11.2.26 +0.1 16.2.2026

+0.1 5.3.26 +0.1 12.2.2026



Traditional lecture method





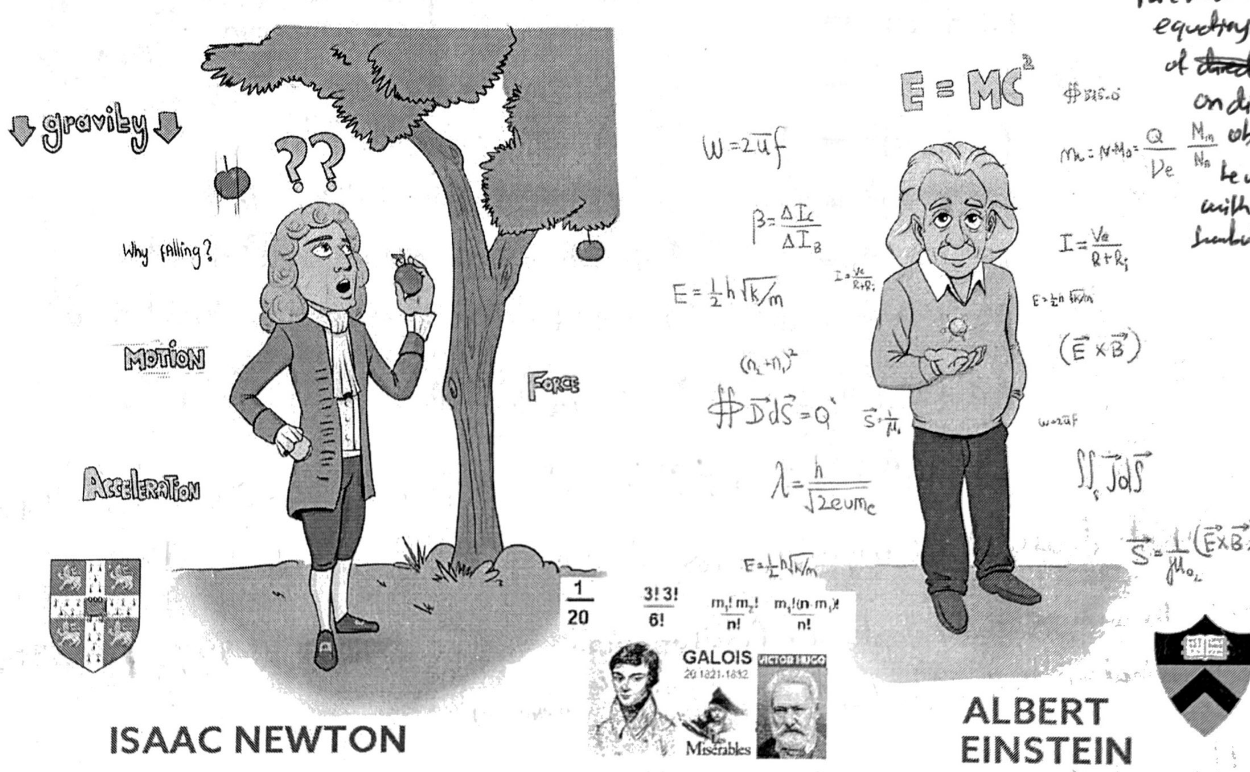
Better learning method

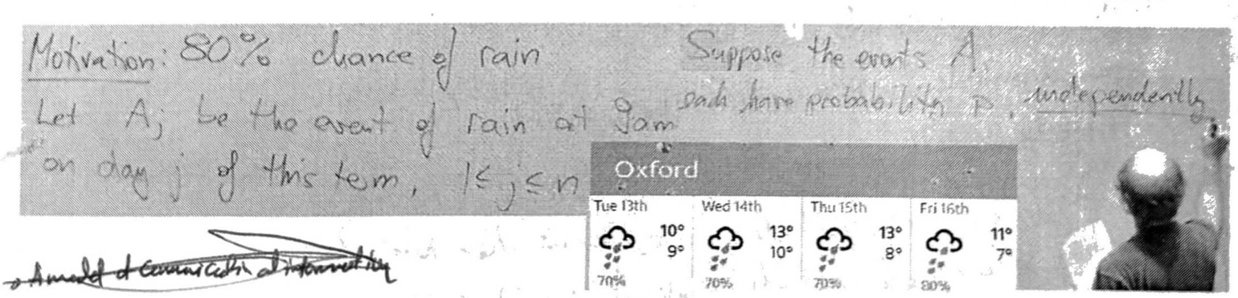
focused on observable, physical phenomena.

School ⇩ gravity ⇩ MOTION ==formalism==> University $E = MC^2$ $\oiint \vec{D} d\vec{S} = Q$ $\iint_S \vec{J} d\vec{S}$

CONCRETE AND ABSTRACT THINKING

focused on pure mathematical equations. Instead of ~~directly~~ relying on direct sensory observation, he worked with abstract symbols, equations.

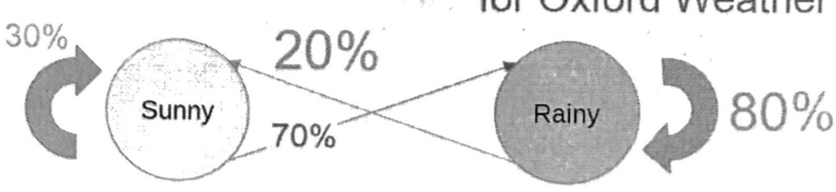




~~A model of communication and interactivity~~

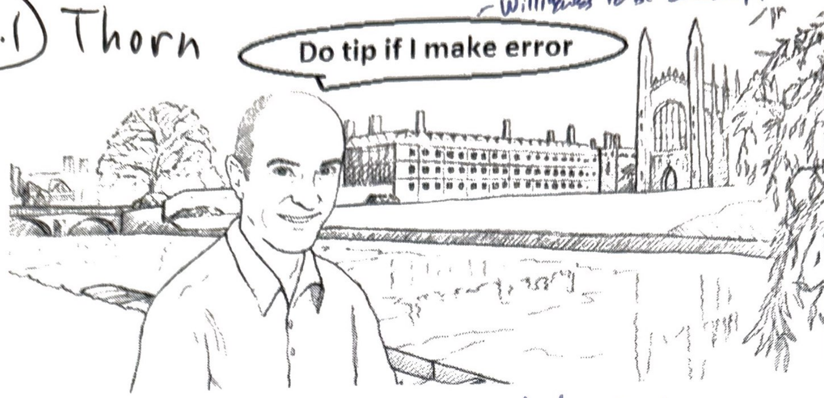
Markoff Chain Probability Model for Oxford Weather

If it is raining today, there is an 80% chance it will be raining.



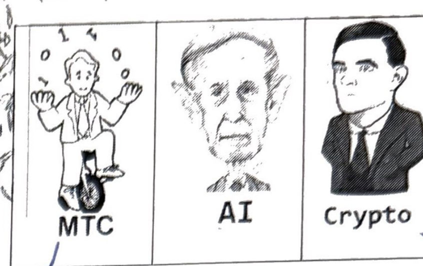
+0.1 7.4 +0.1 +0.1 +0.1 +0.1 +0.1 (+0.1) 7.4

(+0.1) Thorn



– Willingness to be corrected.

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

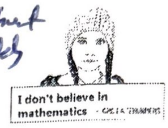


Machine translation and coding

Crypto cryptography

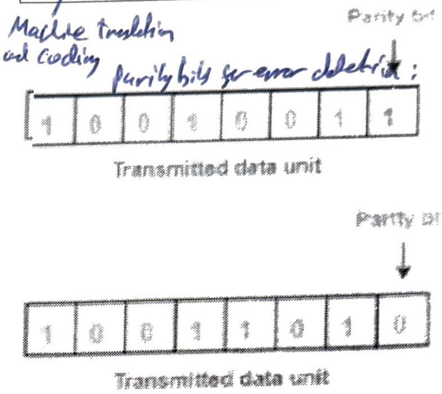
"I believe in clean energy, but I also believe in mathematics"

dual commitment to both fields

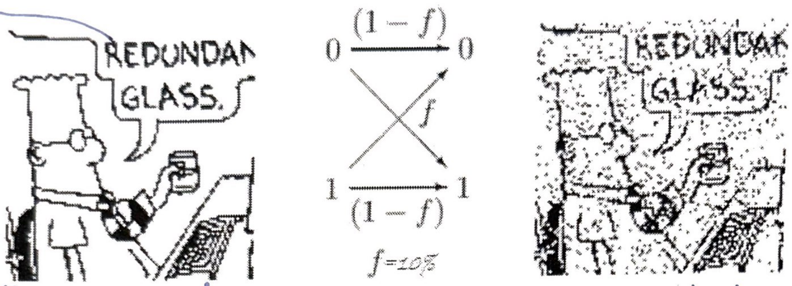


Redundant data helps ensure accuracy and reliability in the presence of errors.

Parity bits for error detection:



0.1

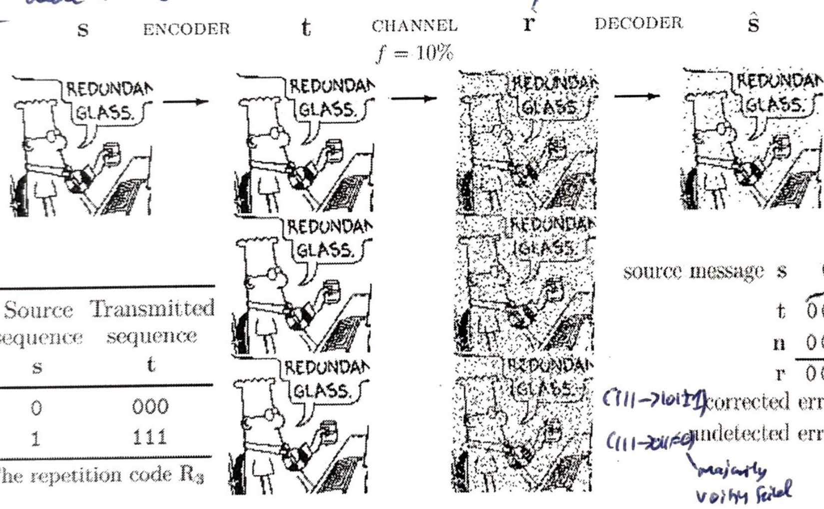


1) Source code: original data either 0 or 1

2) encoder: each bit is repeated.

reception

decoding



Repeat each bit multiple times to create redundancy.

Source sequence s	Transmitted sequence t
0	000
1	111

The repetition code R_3

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

(111→101→1) corrected errors *

(111→010→0) undetected errors *

majority voting failed

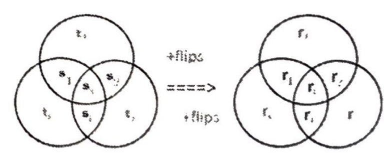
1 error

* −2 error

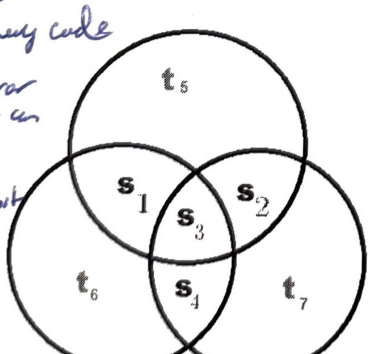
7.4. Hamming code.

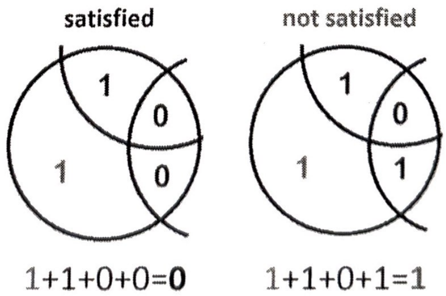
type of binary code

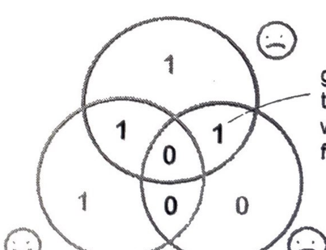
$$\frac{4}{\Sigma} \rightarrow \frac{7}{t}$$



using for error correction that can detect and correct single bit errors.

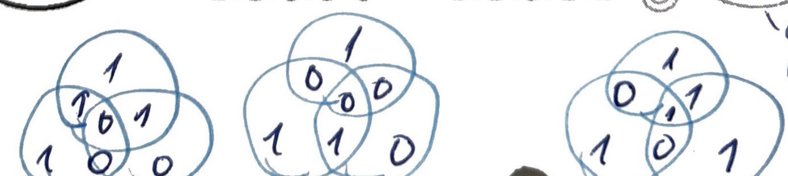






overlaps help in points the errors.

Represents different sets of parity bits. The intersecting the data bits being protected by the parity bits.

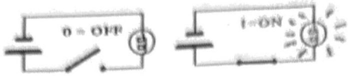


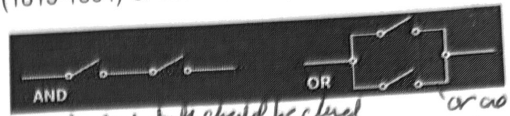
– Flips

Boolean Algebra

Lecture by Pr. Bob Gallagher
Boole (1815-1864) & **Shannon** (1916-2001)

Massachusetts Institute of Technology (MIT)





And-both should be closed

or one should be closed.

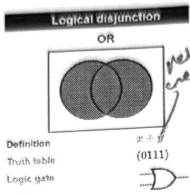
BOOLEAN LOGIC

Logical addition (disjunction)

A	B	F=A∨B
0	0	0
0	1	1
1	0	1
1	1	1

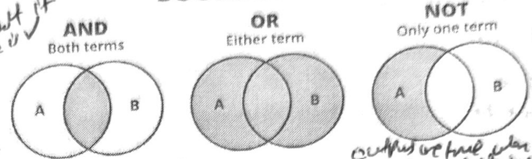
AvB is true if A or B or Both are true

A	B	$A \lor B$
True	True	True
True	False	True
False	True	True
False	False	False



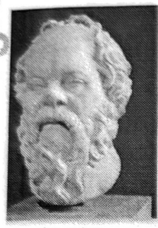
result if one is v

OR operation includes all elements that are in either set A, B, or both



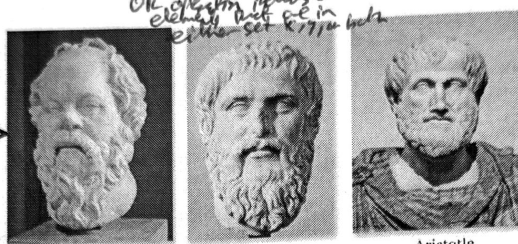
outputs are true when at least one of the inputs are true

Good logic



Socrates

Socrates was a philosopher

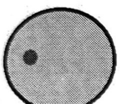


Socrates Plato Aristotle

philosophers are men

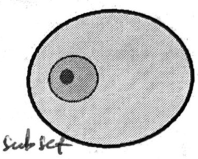
Socrates was a man

S ∈ Φ



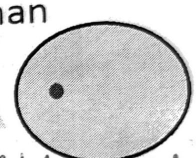
Socrates belongs to the set of philosophers (Φ)

Φ ∈ A



Philosophers are a subset of men (A)

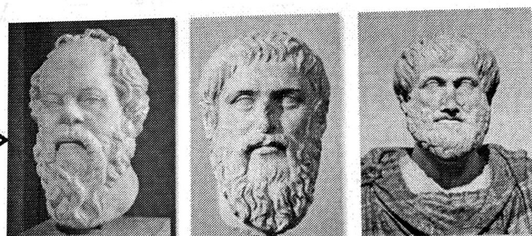
S∈A



Socrates belongs to the set of men.

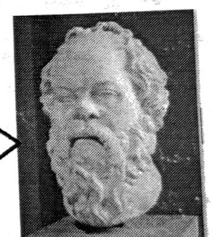
Bad logic

Socrates was a man



Socrates Plato Aristotle

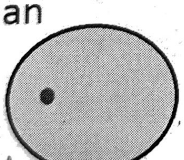
philosophers are men



Socrates

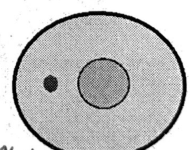
Socrates was a philosopher

S∈A



Socrates belongs to the set of men(A)

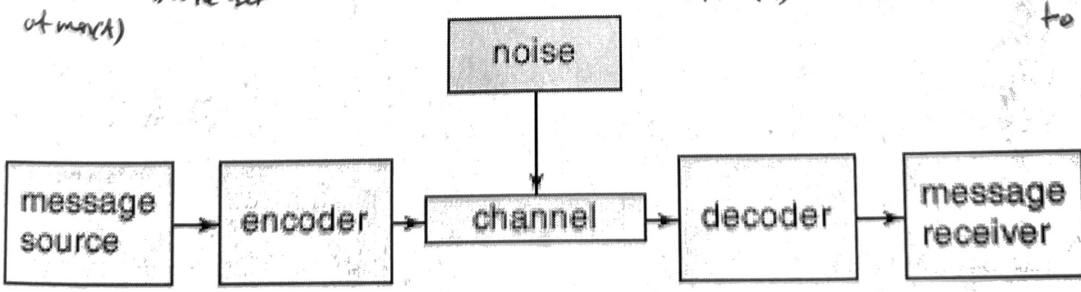
Φ ∈ A



Philosophers are a subset of men(A)

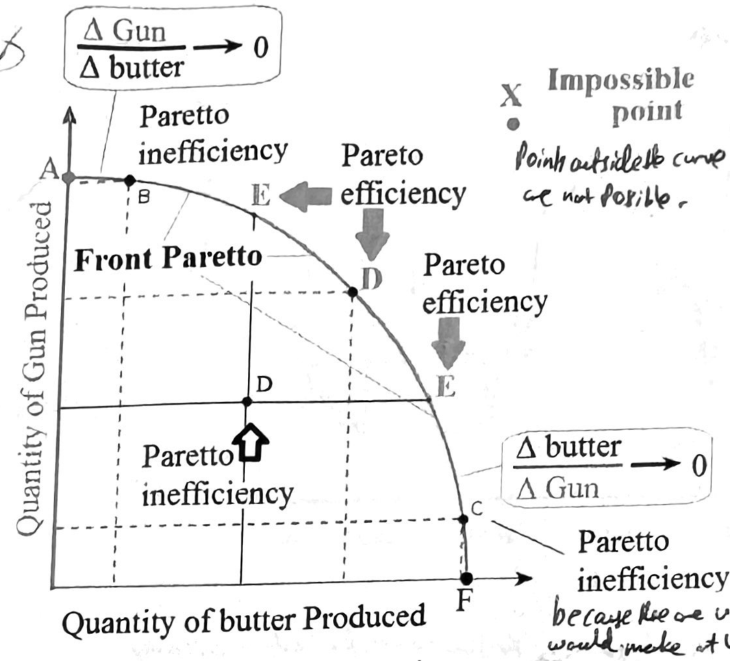
S ∈ Φ

Socrates belongs to philosophers



Trade off between two goods guns and butter. Points on the curve show efficient allocations of resources.

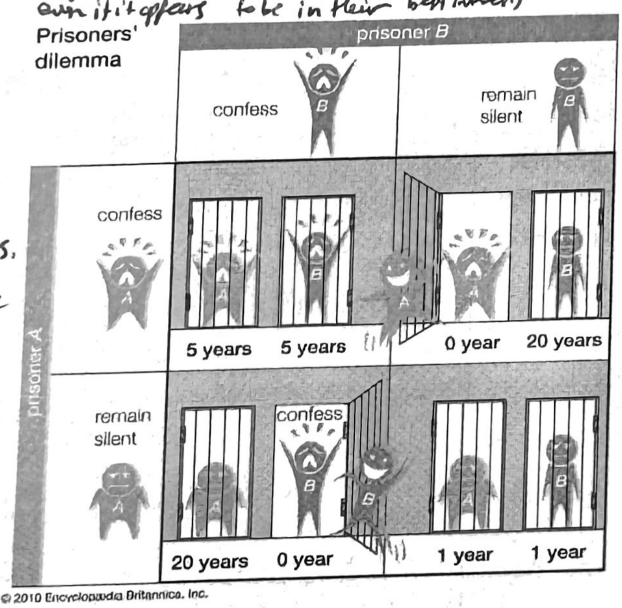
A state where resources are allocated in the most efficient manner, such that any change to make one individual better off would make another worse off.



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.

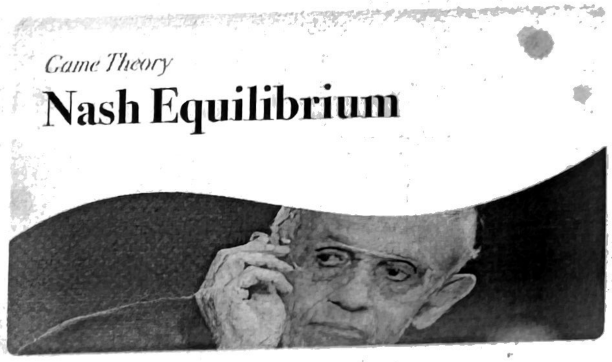
The orange sector E-D-E is the most Pareto efficient - since an increase in one indicator leads to a decrease in another.

Why two rational individuals might not cooperate even if it appears to be in their best interests



© 2010 Encyclopædia Britannica, Inc.

confess, get 5 yrs. and the the gets 0 and gets 20 yrs. Silent,



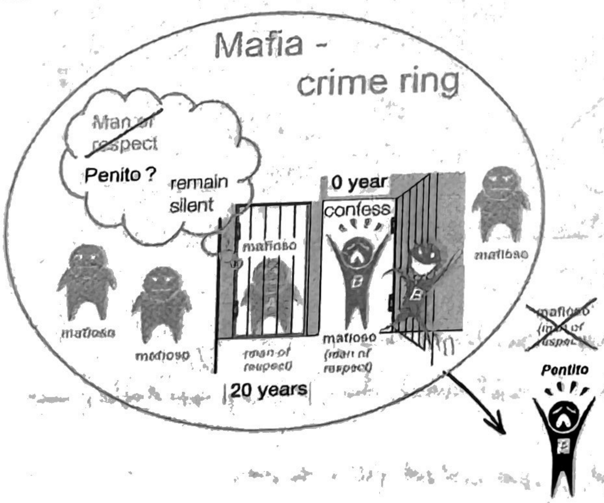
** => Nash equilibrium

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

-1-1

Pareto Optimality

- If both gets recognized, payoff is 5 each.
- If one recognizes and the other doesn't, payoffs vary
- If neither gets recognized, the payoff is -1 each.

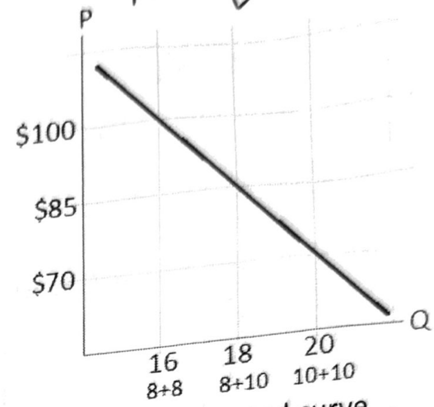


Nash equilibrium - a concept where no player can benefit by changing their strategy while the other players keep theirs unchanged.

Significant fluctuation of oil prices, Peak 2008.

Oil price hits 18-year low

Brent crude, US dollars per barrel



Payoff Matrix used to analyze the strategies and potential outcomes for two oil-producing entities

8 mill

10 mill

– 2 Players & Strategies.

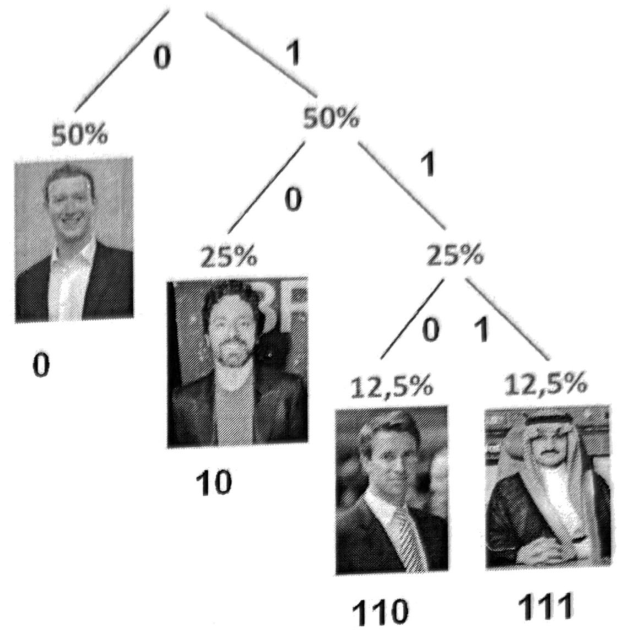
Barrel	1. $8 \cdot 10^6$ / day	2. $10 \cdot 10^6$ / day
1. $8 \cdot 10^6$ / day	$800 millions per day; $100 per barrel; $800 millions per day	$850 millions per day; $85 per barrel; $680
2. $10 \cdot 10^6$ / day	$680 millions per day; $85 per barrel; $850 millions per day	$700 millions per day; $70 per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

John Nash, work in game theory / won Nobel Peace Prize.

Optimal Strategies:

- If both players cooperate and limit their production to 8m. Joint revenue 1600m per day.
- If one player increased production to 10, other stays 8, total revenue decreases, but player producing more oil gains more market share.
- If both increase to 10 million, prices will drop, revenue will be 1400 per day

Huffman tree:
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.



This assumes each symbol is independent of each other.

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			

лла мам ма р

breakdown of individual symbols into sentences.

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

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

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
а р	- 1	11%	8	ар
ра	- 1	11%	9	ра
	9			

numbers correspond to 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

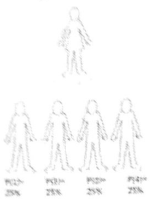
+0.1 +0.1

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

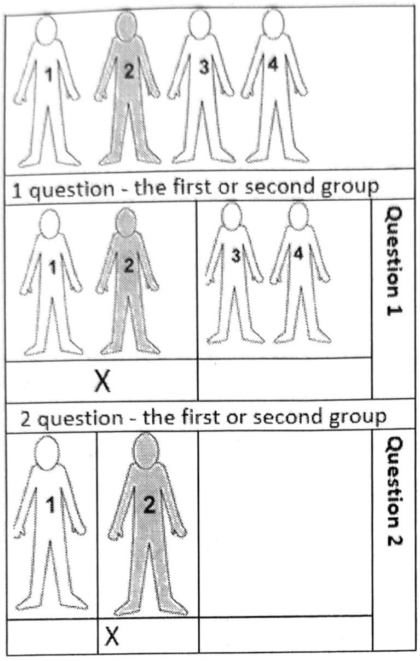
Say **YES** to the **second** if it is better than the first

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

The rules at the top describe a secretary problem. Goal is to choose the best option from the candidates. Strategy is you must reject a baseline number of initial candidates, after that, then see who is better.



Demonstration on how to optimize a search using binary decision trees, if all 4 choices have an equal chance 25%, then it takes 2 yes/no questions on average to find an answer.



Average number of questions =

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

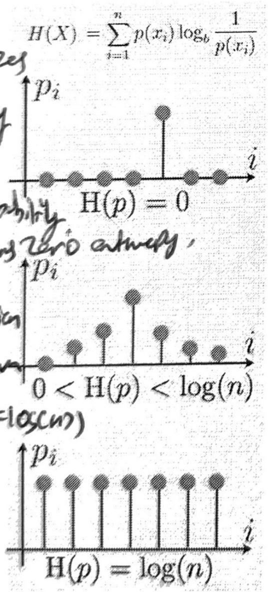
Average number of questions =

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

Here, probabilities are shown and skewed 50, 25, 12.5, 12.5.

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**		

Asking targeted questions about the most likely candidates lowers the average to 1.75 questions. Huffman coding is used.



This visualizes how entropy changes based on probability. Certainty means zero entropy, equal distribution creates maximum entropy ($H(P)=\log(n)$)

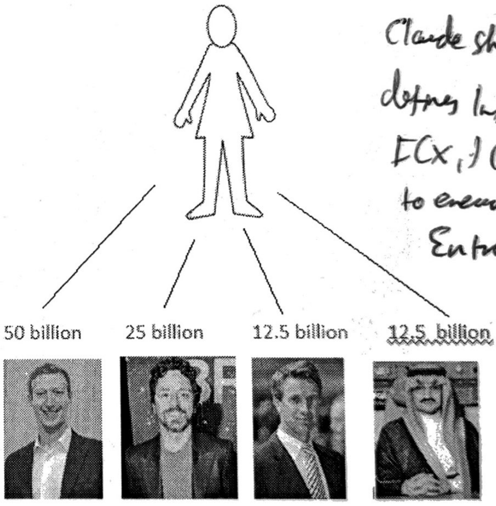
$$\sum_{i=1}^{n} p(i) \log_2 \frac{1}{p(i)}$$

Quantifying information

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

number of bits required to encode choice

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



Claude Shannon, defines Information Content $I(x_i)$ (number of bits needed to encode a choice) and Entropy $H(x)$ Average uncertainty or surprise.

♤♤♤♤ = ~~0.5~~ $\frac{4}{8}$

♡ = $\frac{1}{8}$ Entropy ?

++ = $\frac{2}{8} \cdot \log_2(4) = 0.125$ ~~0.52~~

◇ = $\frac{1}{8} \cdot \log_2(8) = 0.375$

$0.5 \times \log_2(2) = 0.5$

$0.125 \times \log_2(8) = 0.125 \times 3 = 0.375$

$\sum P(0.5) \log_2\left(\frac{1}{0.5}\right) + \sum P(\frac{1}{8}) \cdot \log_2\left(\frac{1}{\frac{1}{8}}\right) +$

$\sum P(\frac{2}{8}) \log_2\left(\frac{1}{\frac{2}{8}}\right) + \sum P(\frac{1}{8}) \log_2\left(\frac{1}{\frac{1}{8}}\right)$