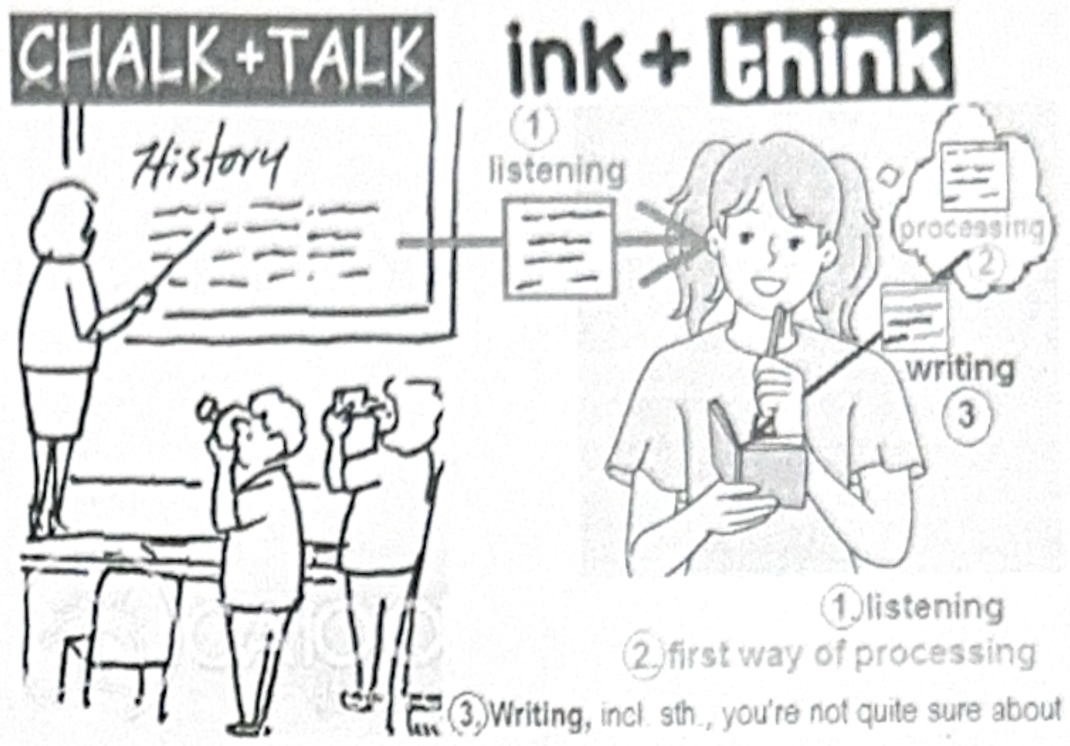
1.

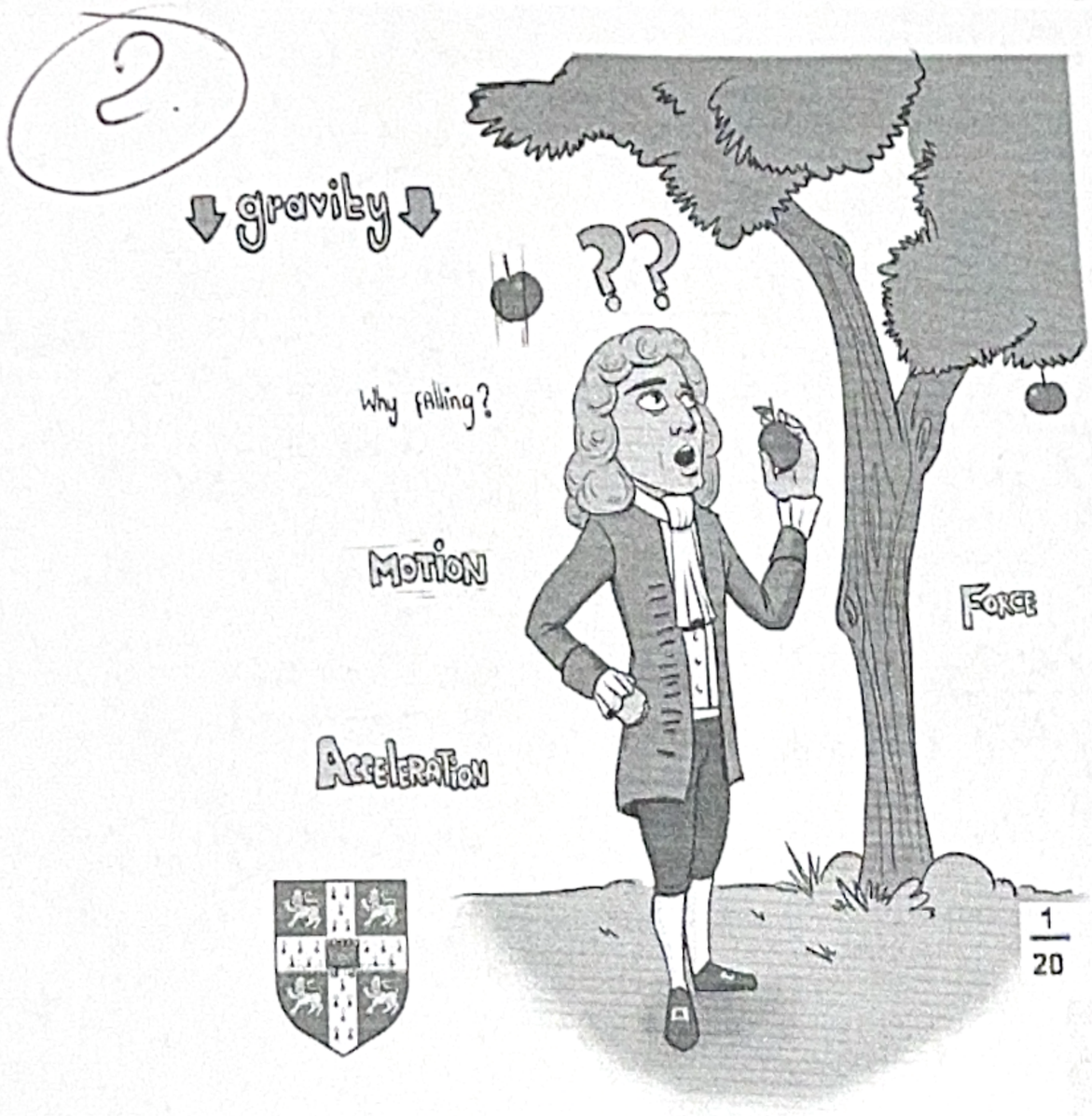




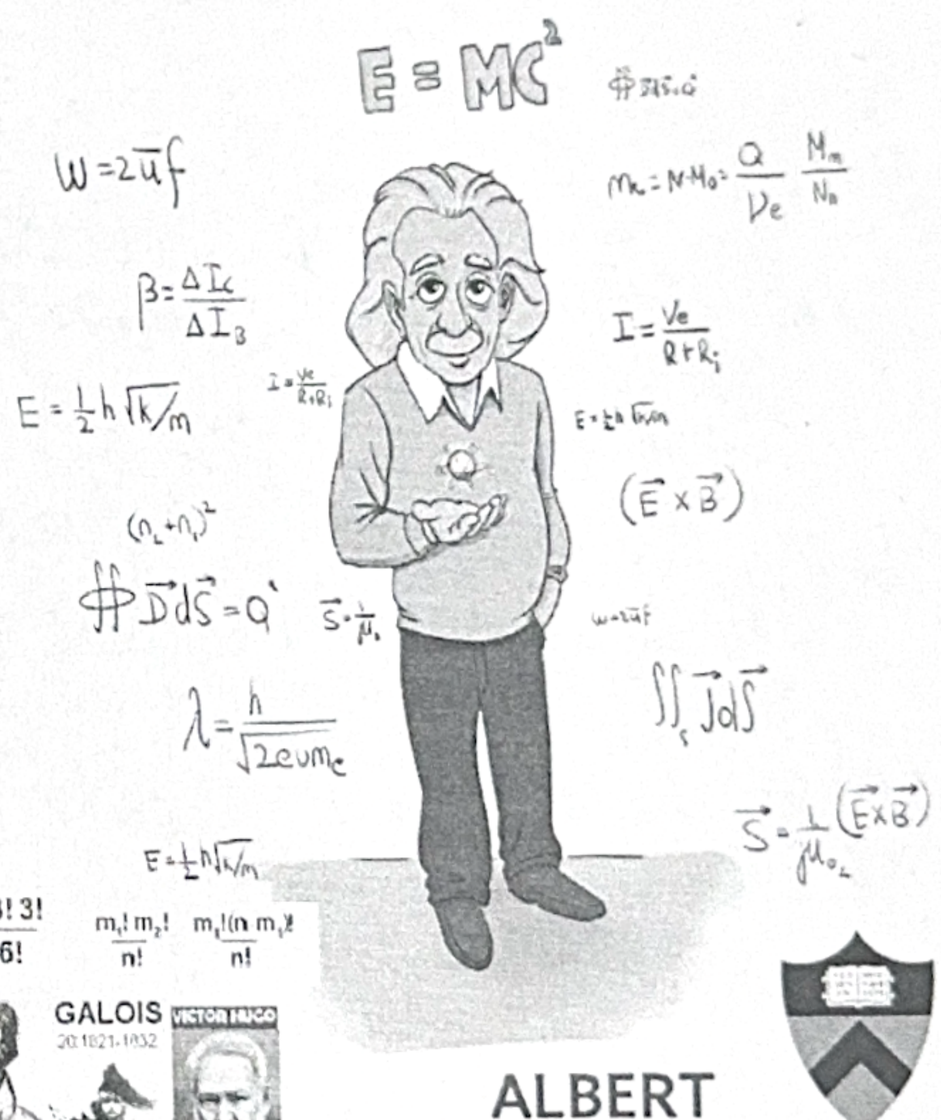
School ⬇ gravity ⬇ MOTION ==formalism==> University $E=MC^2$ $\oint \vec{D}d\vec{S}$ $\iint_S \vec{J}d\vec{S}$

CONCRETE AND ABSTRACT THINKING

2.



ISAAC NEWTON



ALBERT EINSTEIN

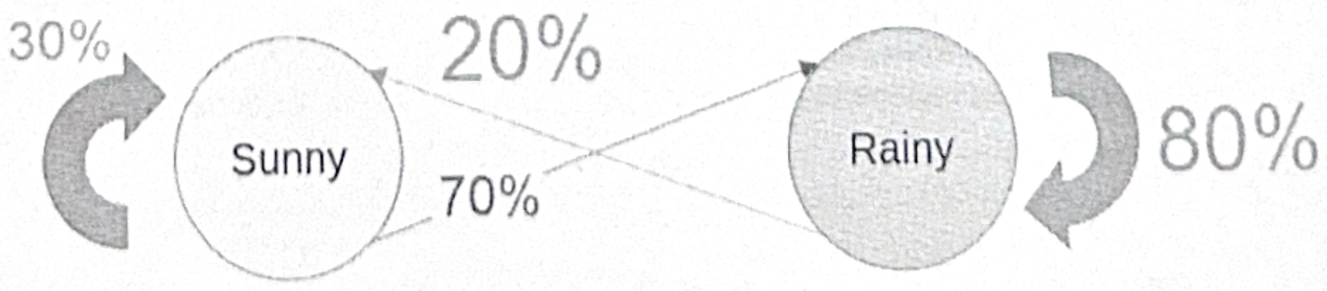
3.

Motivation: 80% chance of rain

Let A_j be the event of rain at 9am on day j of this term, $1 \le j \le n$

Suppose the events A_j each have probability p, independently

Markoff Chain Probability Model for Oxford Weather



(1) • «Chalk and Talk» (Мел и доска)
↓ ↓
классическое обучение: лектор пишет
студентов слушают

• «Ink & Think» (Пиши и думай)
↓
показывает процесс самост. обучения

• «School → Motion ⇒ Formalism ⇒ Uni»
↓
переход от простого наблюдения к
формальной науке

(2). Ньютон ← конкретное мышление
– наблюдал за конкретным
физ. явлением (падением яблока) и вывел закон
всемирного тяготения.

Эйнштейн ← абстрактное мышление
– создавал теории, мыслил
абстрактными категориями, формулами,
моделями, которые опис. Вселенную.

(3) Пример с погодой показывает, как
абстрактная математика помогает
предсказать погоду.

Цепь Маркова:

• Sunny: стрелка на себя (30%) ⇒ если
сегодня солнечно, то с вероятностью 30%
снова будет солнечно.

• Rainy: стрелка на себя (80%) =//= ↙

• Переходы: из ☀ в 🌧 – 70%; наоборот – 20%

Цепь Маркова – мат. модель, которая
предсказывает будущее состояние, основываясь
только на текущем состоянии, без учета всей истории.

→ пример абстр. модели, которая исп. в жизни

Треугольник Паскаля



Pascal's triangle

фундаментальная структура в комбинаторике

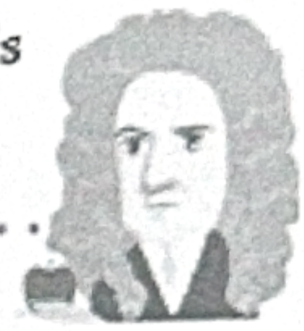
1
1 1
1 2 1 — 4
1 3 3 1 — 8
1 4 6 4 1 — 16
1 5 10 10 5 1 — 32
1 6 15 20 15 6 1 — 64
1 7 21 35 35 21 7 1 — 128
1 8 28 56 70 56 28 8 1

Дерево вероятностей



$$(a+b)^0 = 1$$
$$(a+b)^1 = a+b$$
$$(a+b)^2 = a^2+2ab+b^2$$
$$(a+b)^3 = a^3+3a^2b+3ab^2+b^3$$
$$(a+b)^4 = a^4+4a^3b+6a^2b^2+4ab^3+b^4$$
$$(a+b)^5 = a^5+5a^4b+10a^3b^2+10a^2b^3+5ab^4+b^5$$
$$(a+b)^6 = a^6+6a^5b+15a^4b^2+20a^3b^3+15a^2b^4+6ab^5+b^6$$

Newton's Binomial



Бином Ньютона

Треугольник Паскаля - числовая таблица, где каждое число равно сумме двух чисел над ним.

Дерево вероятностей показывает все возможные исходы последовательных случайных событий.

* на примере монетки

(1 бросок)

I шаг - бросок → 2 варианта О/Р,
вероятность 1:2

II шаг - возможные комбинации:
(2 броска) ОО → 1:4 → 2 строка Δ Паскаля
ОР → 2:4
РО → 2:4
РР → 1:4

III шаг — 1:8 / 3:8 / 3:8 / 1:8 → 4 строки Δ Паскаля
(3 броска)

* Комбинаторика — от латинского «combina»

↙

наука о расположении элементов в определенном порядке и подсчете числа способов такого расположения.

Треугольник Паскаля в математике показывает, сколько существ. способов определить результат при подбрасывании монетки.

Это связь теории вероятностей и чисел.

Massachusetts Institute of Technology (MIT)

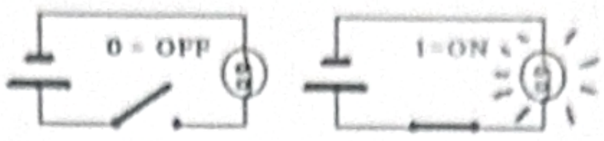


Lecture by Pr. Bob Gallagher

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



Лог. операции

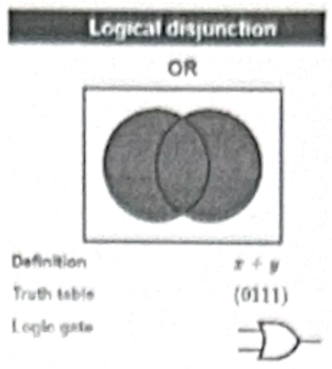




1. Logical addition (disjunction)

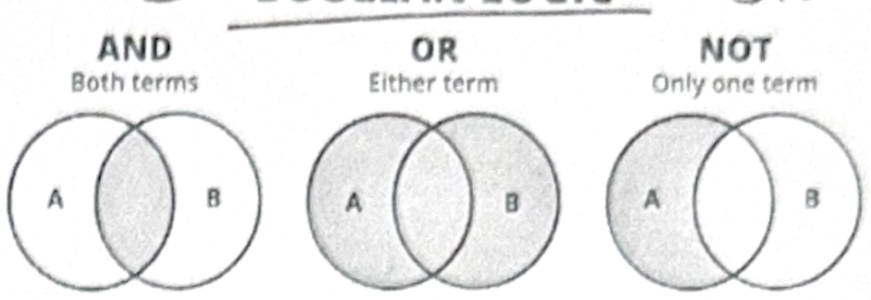
A	B	F=A∨B
0	0	0
0	1	1
1	0	1
1	1	1

A	B	$A \vee B$
True	True	True
True	False	True
False	True	True
False	False	False

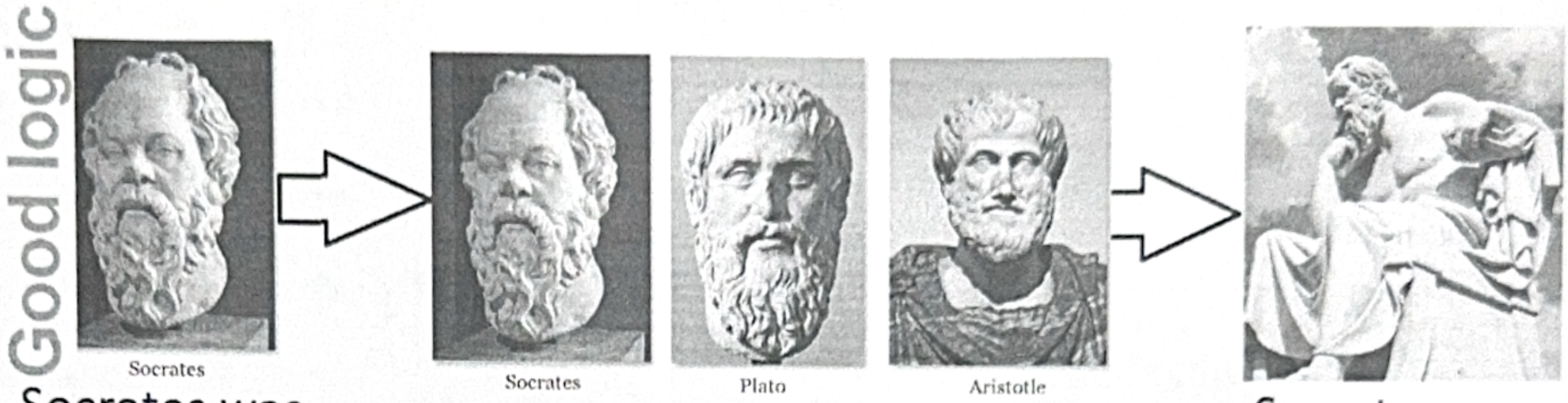


2. BOOLEAN LOGIC

Булевая логика

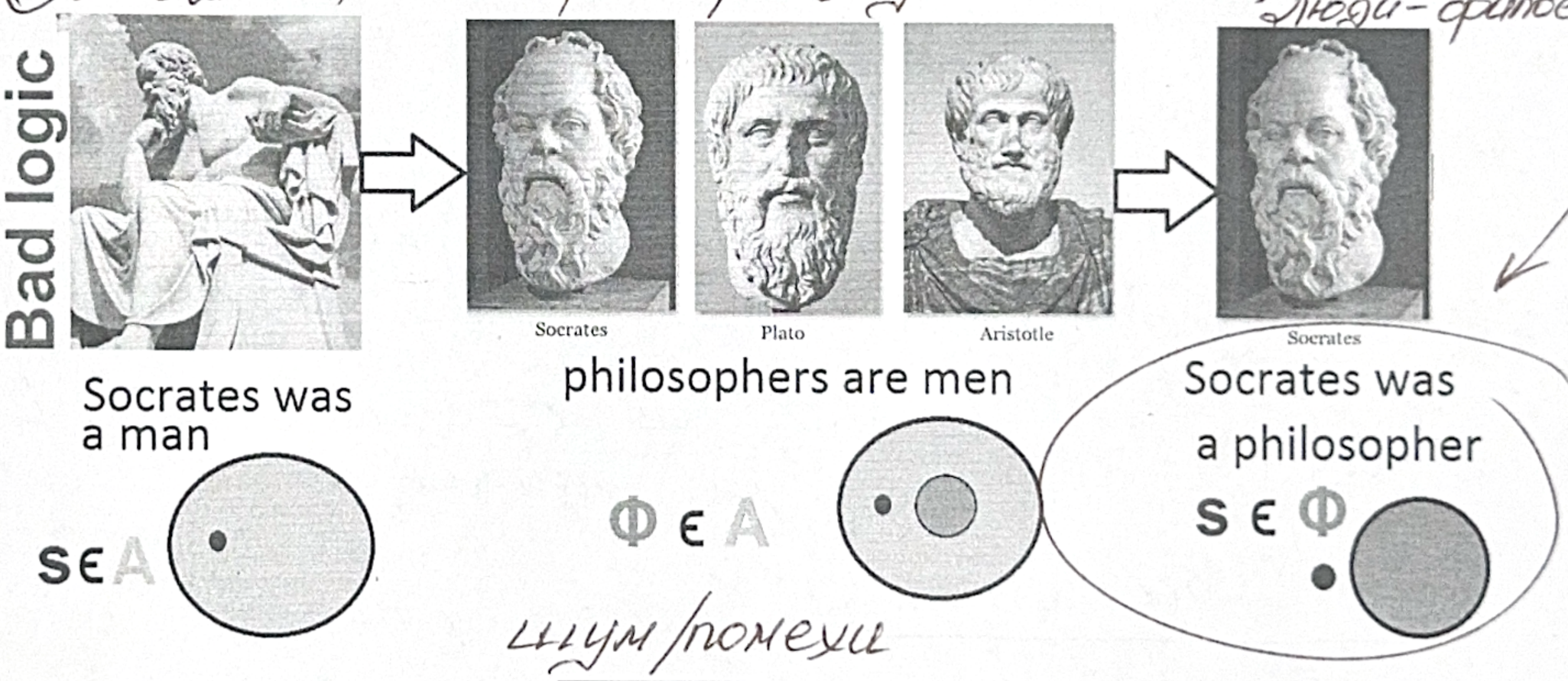


3. Силлогизмы

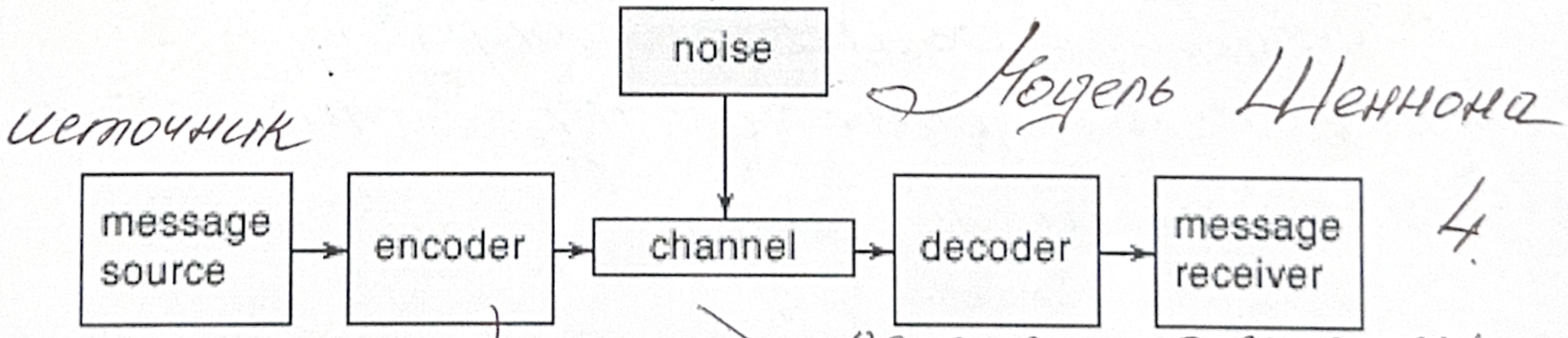


Логически верно

Ошибка: то, что все философы – люди, не означает, что все люди – философы



шум / помехи



источник

Модель Шеннона

4.

кодировщик

декодер

получатель

канал связи

1. Лог. сложение (дизъюнкция):

A ∨ B (A или B) даёт <u>false</u> только в одном случае: когда оба высказывания ложны.

Во всех остальных случаях — <u>true</u>

2. Булева логика → AND - оба истинны
→ OR - достаточно одного (или обоих)
→ NOT - инверсия

*Инверсия → отрицание четко даёт ложь, и наоборот.

3. Силлогизм - умозаключение, в котором из 2 исходных суждений с необходимостью выводится 3-е суждение (вывод).
°основал Аристотель

4. Главная цель <u>модели Шеннона</u> — максимизировать скорость и точность передачи информации по каналам связи при минимальных затратах.

<u>Если кратко</u>: цель модели — сделать так, чтобы сообщение отправленное Источником, дошло до адресата быстро, без искажений и самым дешевым, способом.

Эффективным

Resume of Lecture by Pr. Bob Gallagher from MIT

Massachusetts Institute of Technology (MIT)

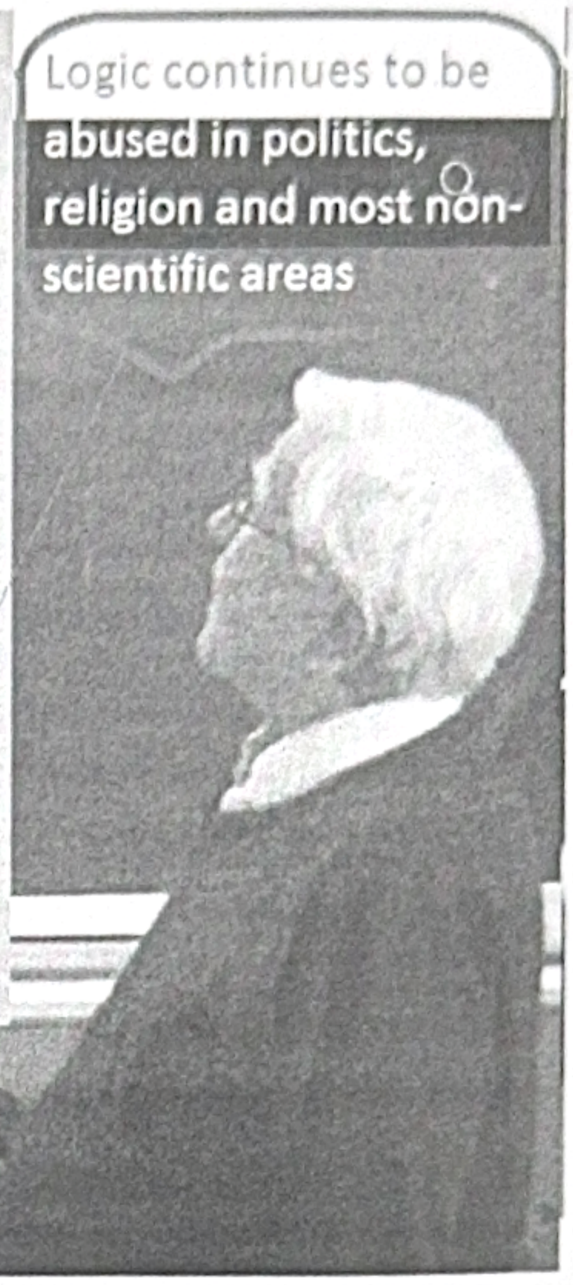
George Boole (1815-1864) developed Boolean logic

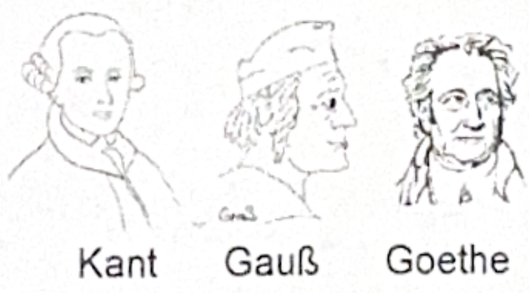
The principles of logical thinking have been understood (and occasionally used) since the Hellenic era.

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

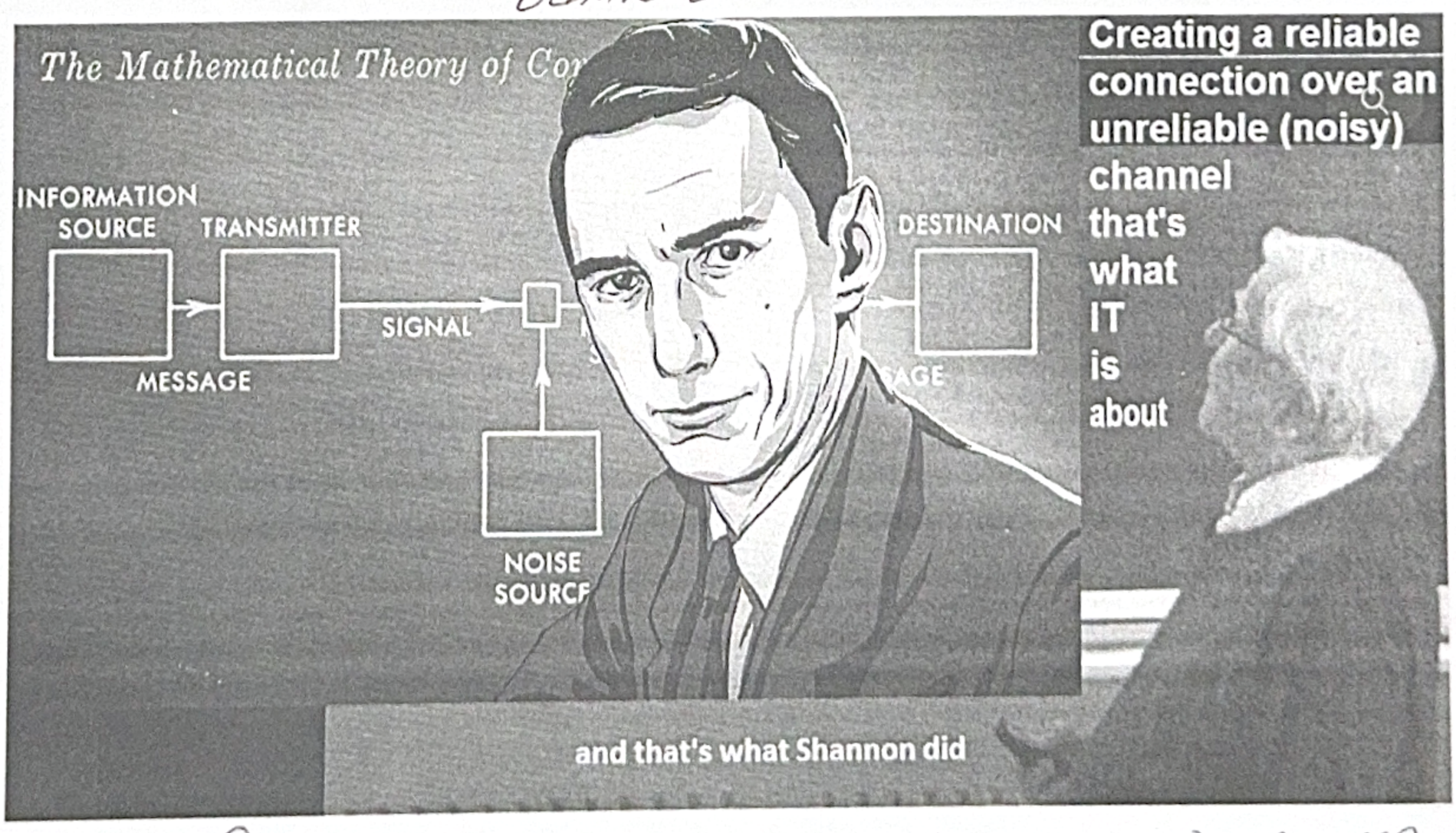
Kant, Gauss, Goethe are great → Kant, Gauss, Goethe - Germans → Germany Great

Великие → немцы → немцы великие



Bad logic (abuse of logic)

Логика часто использ. неправильно в политике, религии и некоторых сферах



Цель IT и связи — создать надежное соединение поверх ненадежного (шумного) канала.

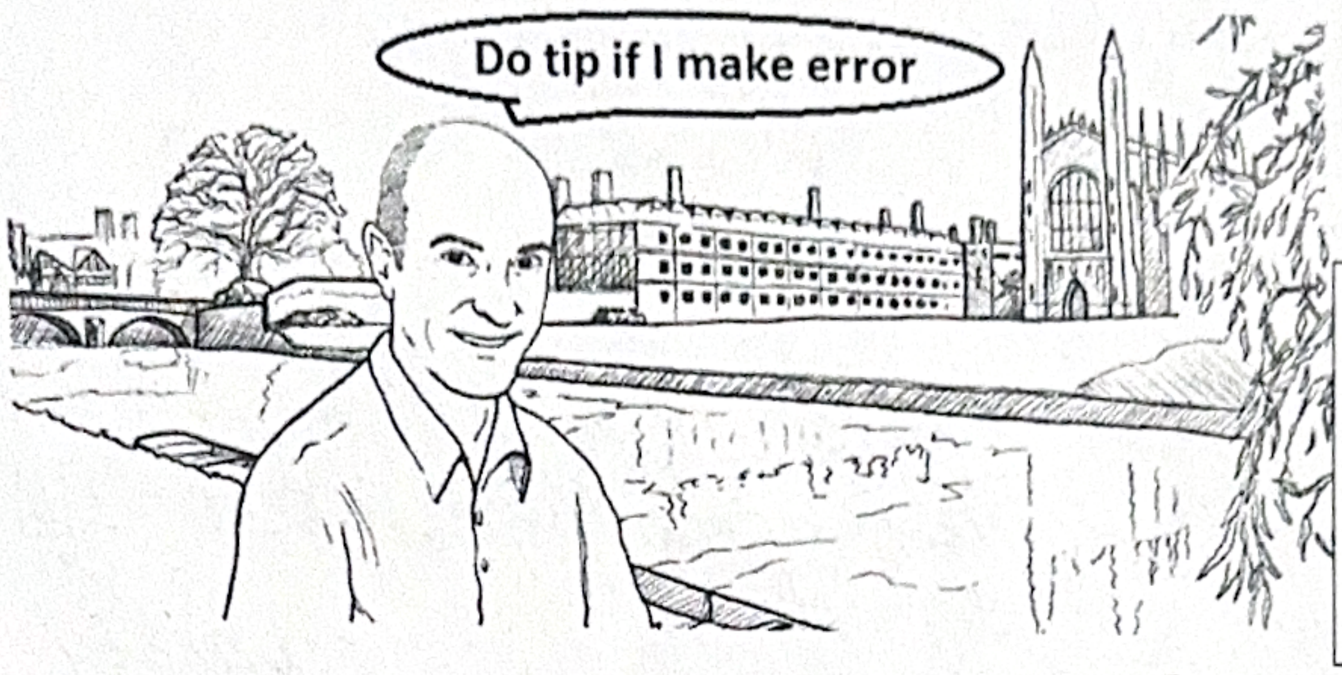
Буль (1815-1864) - систематизировал принципы логического мышления и перевел их в уравнения (булеву алгебру). До него логика была словесной, Буль сделал её математ.

Шеннон (1916-2001) - показал, как использ. булеву алгебру в качестве основы для переключающих схем. Это в свою очередь стало фундаментом для проектирования компьютеров и систем связи.

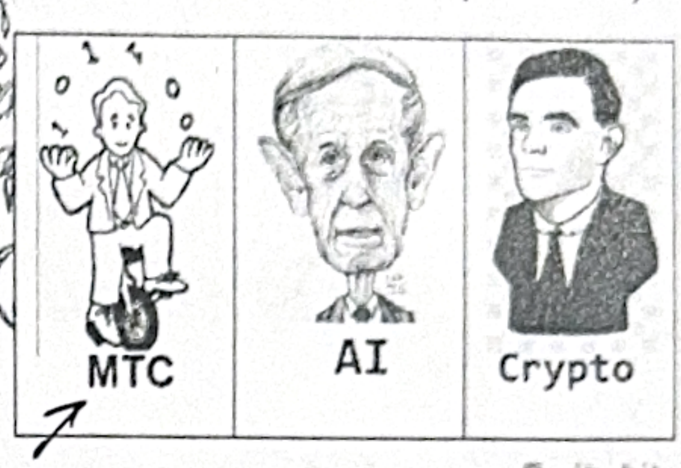
→ Главная мысль

создать надежное соед. поверх ненадежного (шумного) канала.

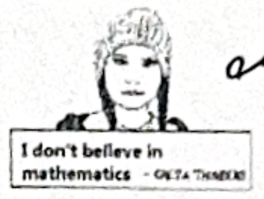
* он придумал, как это делать математически



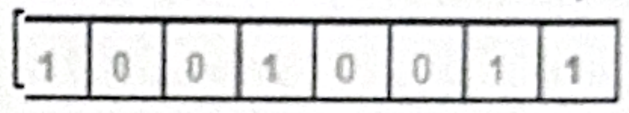
Sir **Dr. D. MacKay**,
University of Cambridge
(22 April 1967 – 14 April 2016)



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

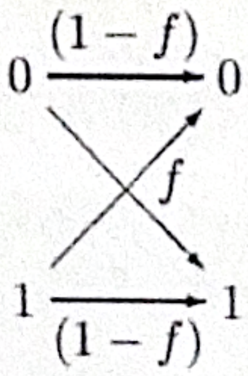


машинное обучение



(2)







(1)

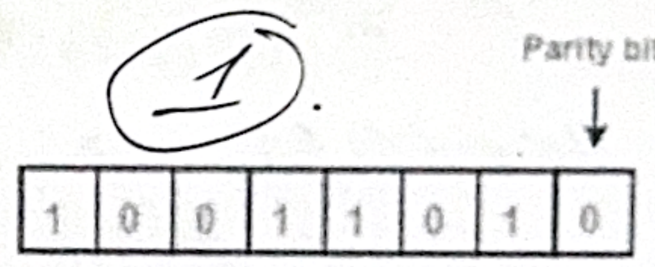


Схема передачи

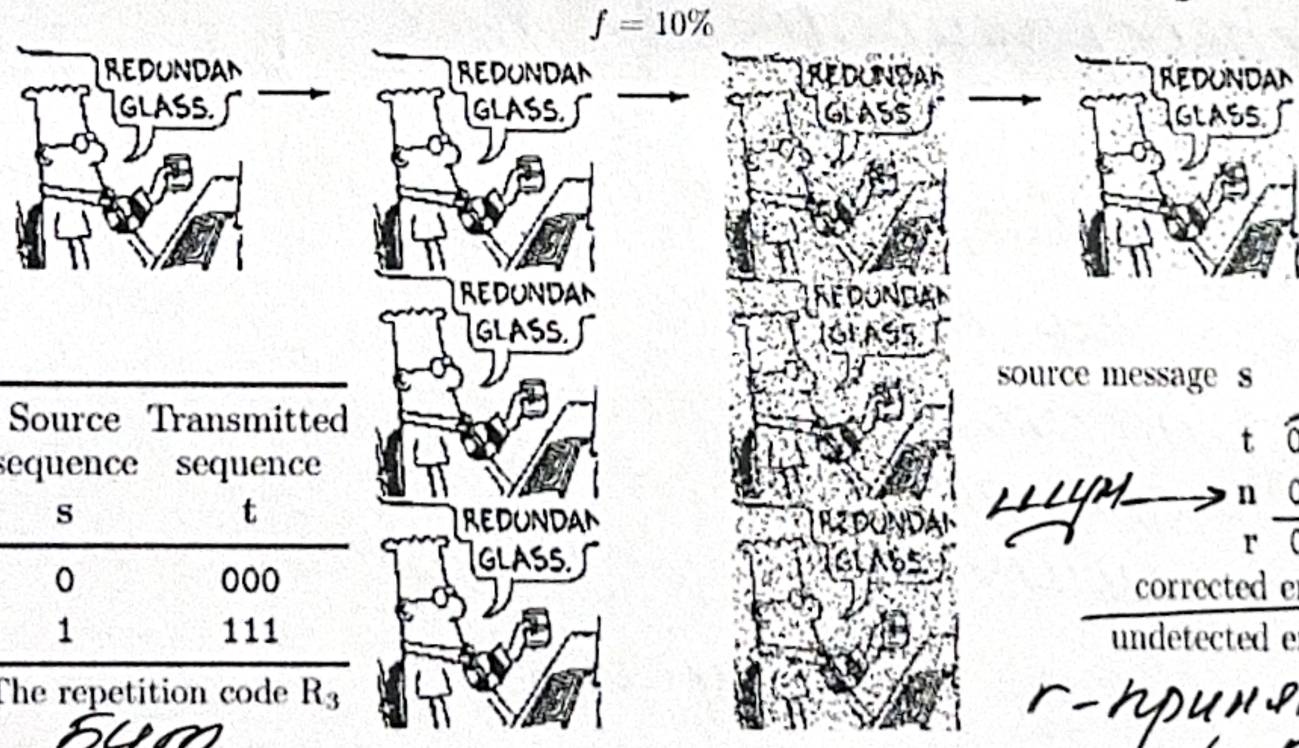
s ENCODER t CHANNEL r DECODER ŝ

$f = 10\%$

обнаруж. ошибки

декодер начел исправл

голосование большинством



(3)

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
шум →	n	000	001	000	000	101	000	000
	r	000	001	111	000	010	111	000
			corrected errors ✓					
			undetected errors *					

* ← необнаруж. ошибки

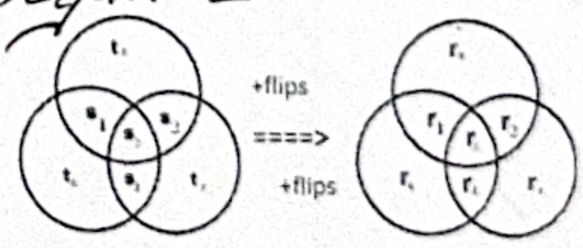
название метода

каждый бит повторяется 3 раза

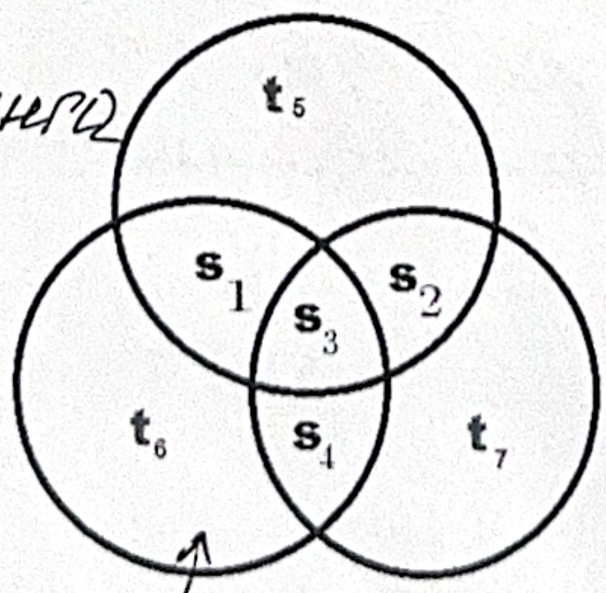
r - принятое
t - переданное

7.4. Hamming code.

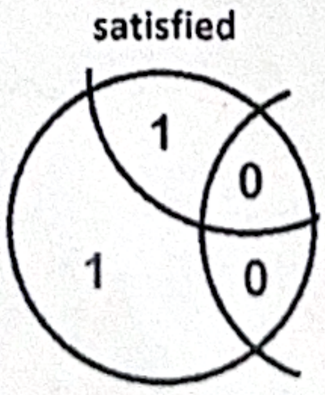
$\frac{4}{\Sigma} \rightarrow \frac{7}{t}$

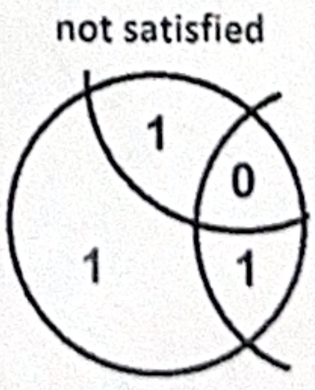


Код Хэмминга

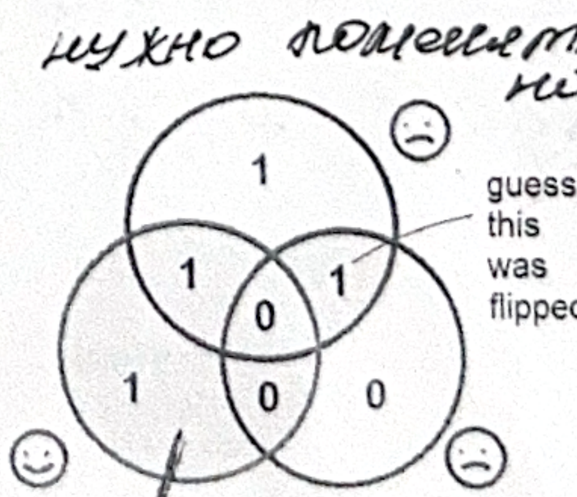


круги Венна





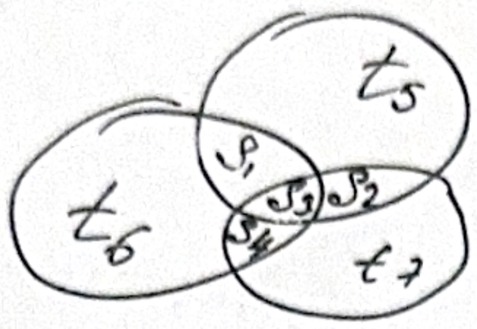
нужно поменять на 0



проверка на четность

кол-во 1 четное ⇒ сумма четная (satisfied)

* каждый круг проверка на четность



• t - внешние биты – биты, которые входят только в 1 проверку

• s - пересечения - биты, которые входят в две или три проверки.

(1.) Передаваемый блок данных – это строки из 8 бит. В нашем примере: 10010011 (1)

Parity bit (бит четности): parity bit последний бит справа – специальный бит, который добавляется к данным, чтобы общее кол-во единиц (или нулей) было четным.

1001011 → 4 единицы → четное

* первые 7 бит - полезное данные

(2.) 1. В реальном мире всегда есть шум ($f > 0$)

2. Без спец. кодов каждый десятый бит будет терться или переворачиваться.

3. Чтобы бороться с этим, мы добавляем избыточность (Parity bit, код Хэмминга) – это и есть «лишний стакан» на картинке.

4. Математики позволяет вычислить, сколько именно избыточных битов нужно добавить, чтобы при уровне ошибок 10% получатель смог восстановить исходное сообщение.

(3.) R_3 - используется, т. к. если в книге произойдет ошибка (1 из 3 перевернется), получатель все равно сможет понять, что было передано, по принципу «голосования большинством».

«Скажи нет 3-ему»

1.



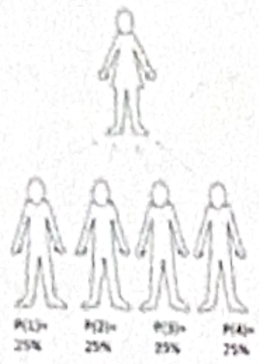
Say **NO** to the **first**

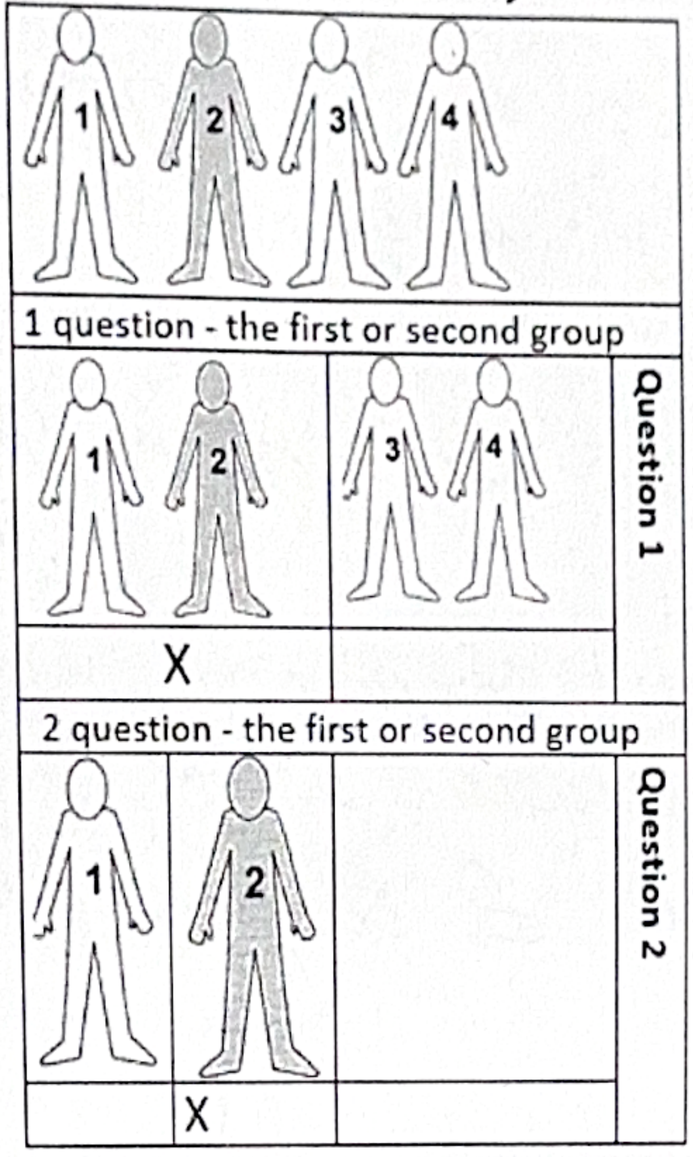


Say YES to the **second** if it is better than the first

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

Игра в вопросы





2.

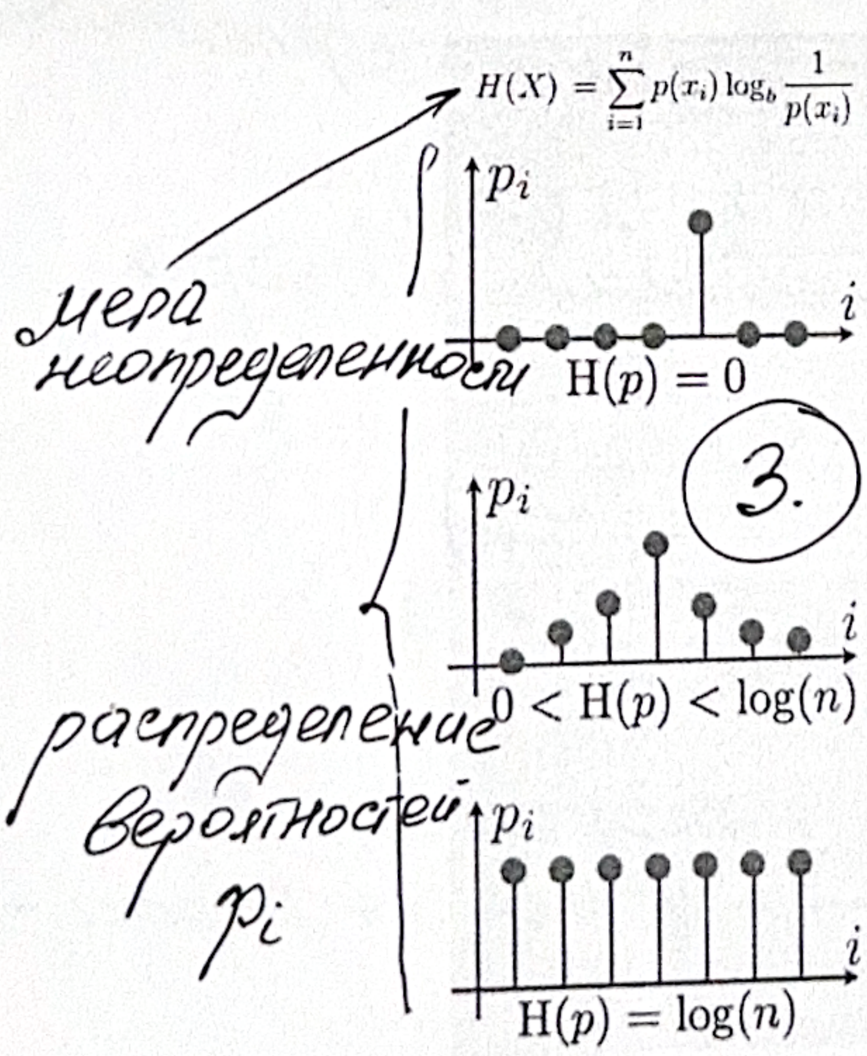
Average number of questions =

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
			

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		



мера неопределенности

3.

распределение вероятностей p_i

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

Энтропия

Quantifying information

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

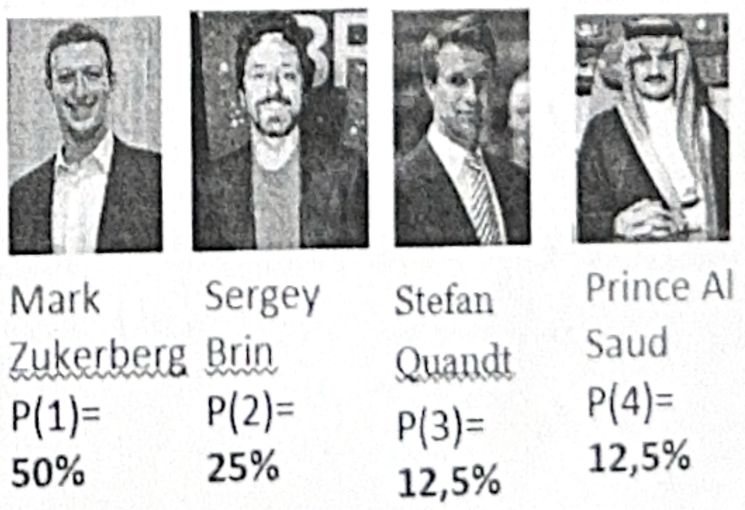
number of bits required to encode choice

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

среднее кол-во информации

кол-во информации





1. Принцип оптимального бинарного поиска — отсекать наименее вероятные варианты, чтобы минимизировать среднее число вопросов.

2. Задаётся вопрос «Это первая или вторая группа?» (то есть делим на 2 пары). После ответа остаётся пара, и задаётся вопрос: «Это первый или второй человек в группе?»

Расчёт среднего числа вопросов при стратегии «делить пополам». Каждый из 4 равновероятных вариантов требует ровно 2 вопроса (так как $2^2=4$). Среднее = 2.

3. Классические формулы Шеннона:

- Энтропия $H(X) = \sum_i^n p(x_i) \log_b \frac{1}{p(x_i)}$ - мера неопределенности
- Кол-во информации в отдельном событии $I(x_i) = \log_2\left(\frac{1}{p_i}\right)$ - число бит необх. для кодирования события.
- Среднее кол-во информации $\sum p(x_i) I(x_i)$ — это и есть энтропия

Графики:

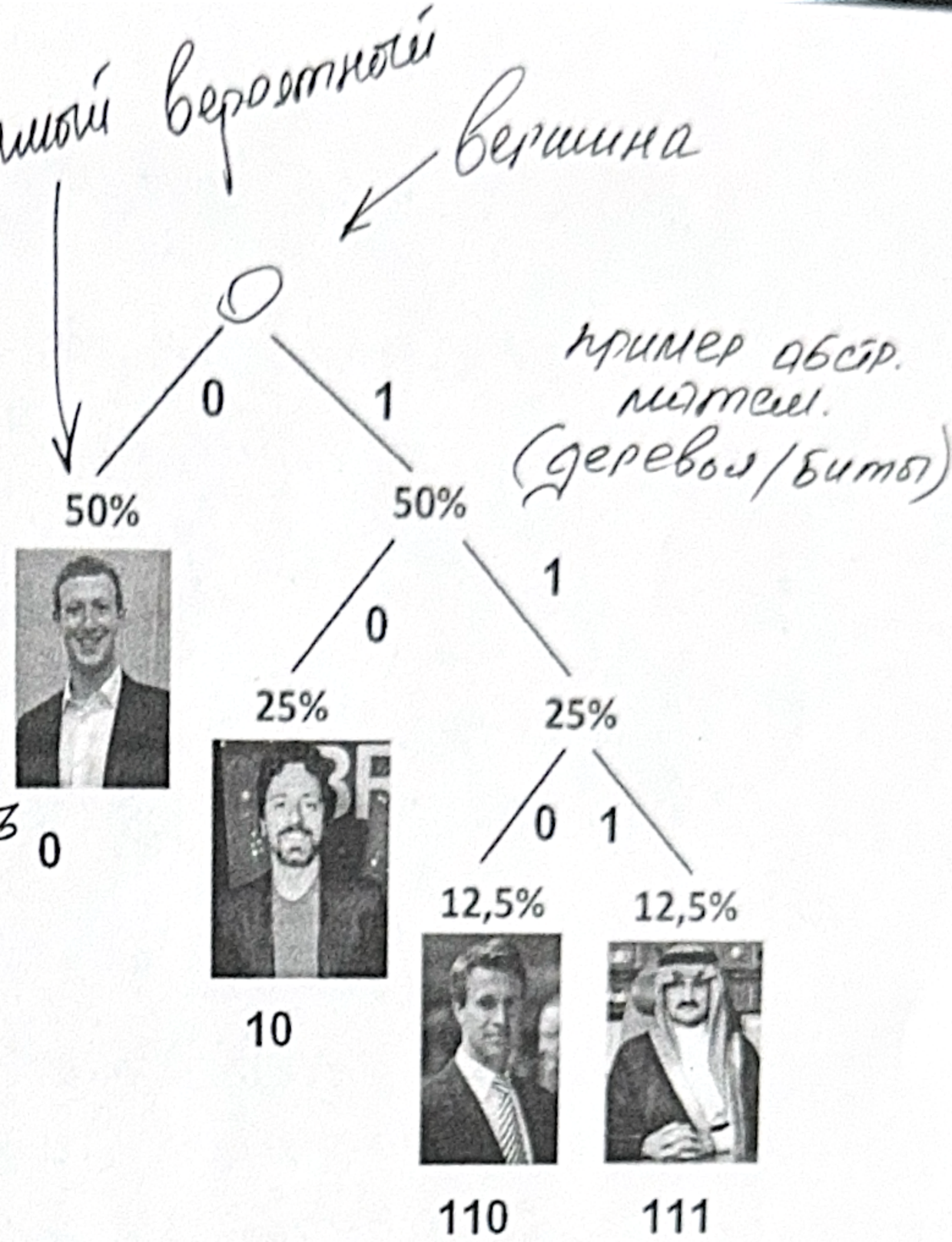
— Вверху: вся вероятность сосредоточена в одном исходе → $H(p) = 0$ (неопределенности нет, информация не нужна)

— В середине: неравномерное распределение → $0 < H(p) \leq \log(n)$

— Внизу: равномерное распределение → $H(p) = \log(n)$ (максимальная неопределенность)

Классический пример кода Хаффмана (дерева решений):

чем чаще встречается объект (~~тем~~ выше p), тем короче его код. Это позволяет минимизировать среднюю длину сообщения



пример обсер. матем. (деревом/битом)

Марковская цепь первого порядка

Таблицы

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		

лламмамар

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



Кодирование и разбиение

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

из цифровой последов. собираются слова

0.467319167351

ам ыл ла ам ма ра ма ыл ла ам мы

мылла рама

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

Суть: 1. Сжатие данных. 2. Вероятностная модель

1.) Вместо того, чтобы хранить каждую букву, мы храним слоги или их коды (цифры). Если слог «ма» встречается довольно часто (30%), ему дают короткий код (например, цифру 4), а редкому слогу - длинный код, это экономит память.

2) Демонстрация того, как язык можно разбить на вероятностные цепочки. Зная, какая буква или слог идет с какой вероятностью, можно предсказывать текст или сжимать его (как в алгоритмах ZIP или в предсказательном наборе текста на смартфоне)

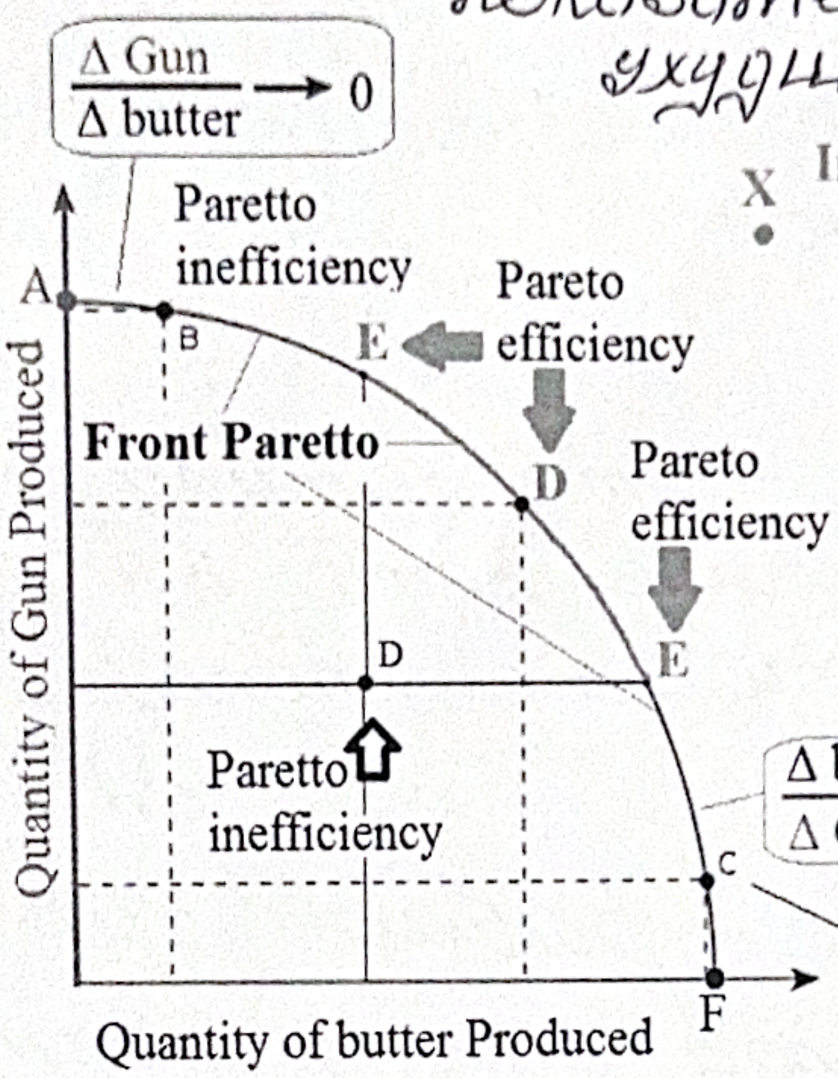
↓

Переход от абстрактной математики к оптимальному кодированию и статическому анализу языка.

Это базовые принципы работы современных алгоритмов сжатия и обработки естественного языка (NLP).

Эффективность по Парето

Состояние системы, при котором улучшение 1 показателя невозможно без ухудшения 2



X Impossible point

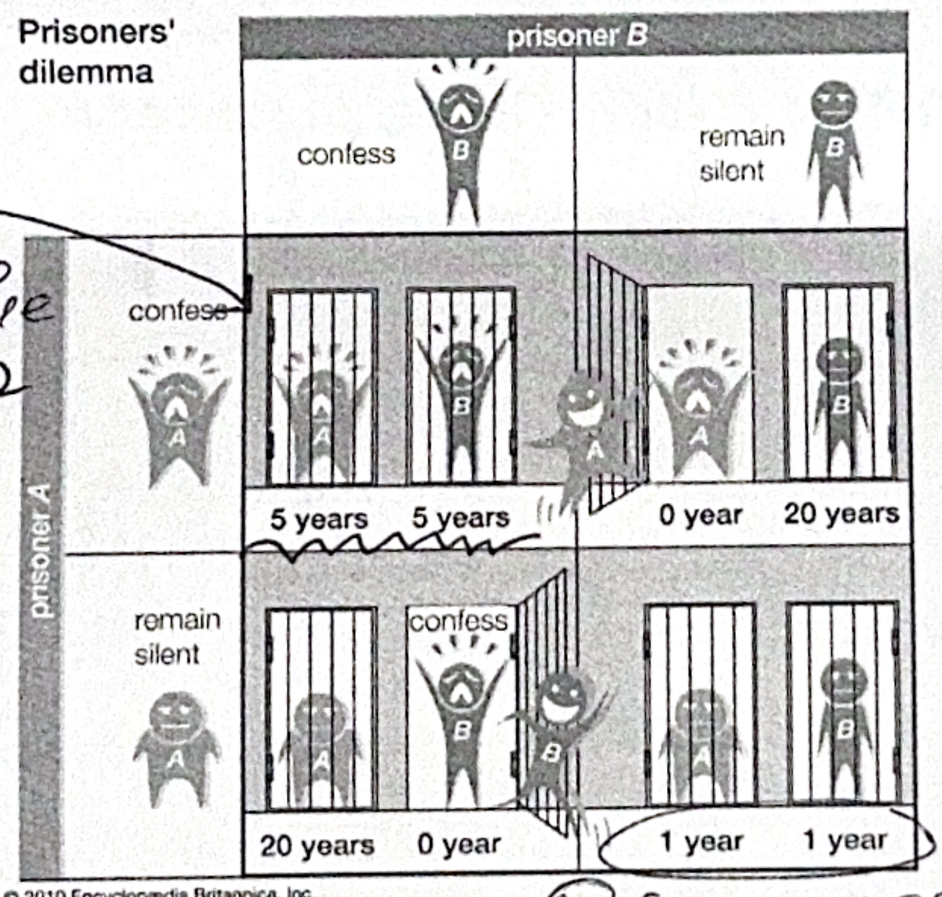


by Vilfredo Pareto

1848-1923

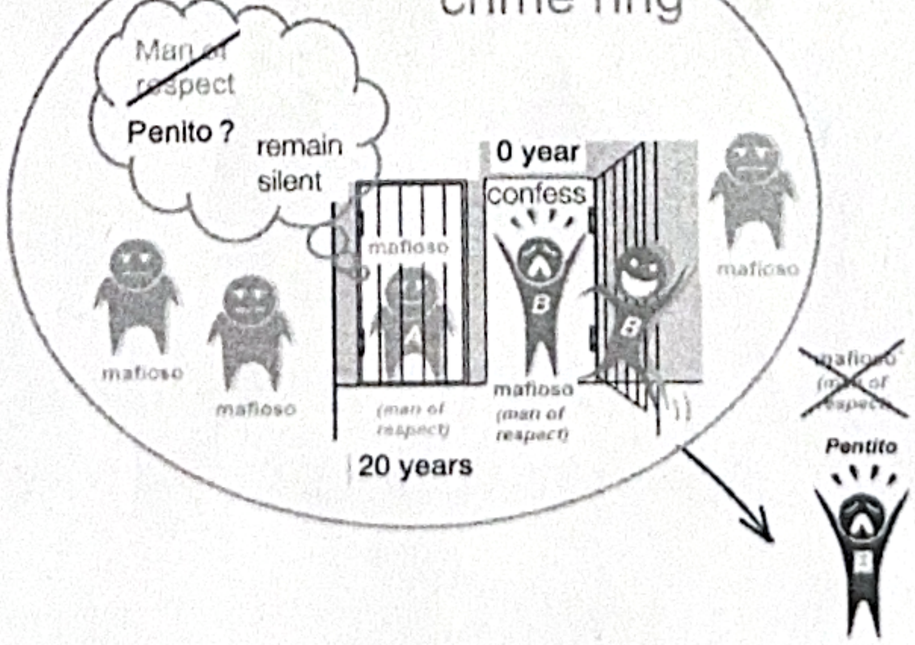
The orange sector E-D-E is the most Pareto efficient - since an increase in one indicator leads to a decrease in another.

*описывает оптимальность а не равенство



Равновесие Нэша

Оптимальность Парето



Game Theory

Nash Equilibrium



** => 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

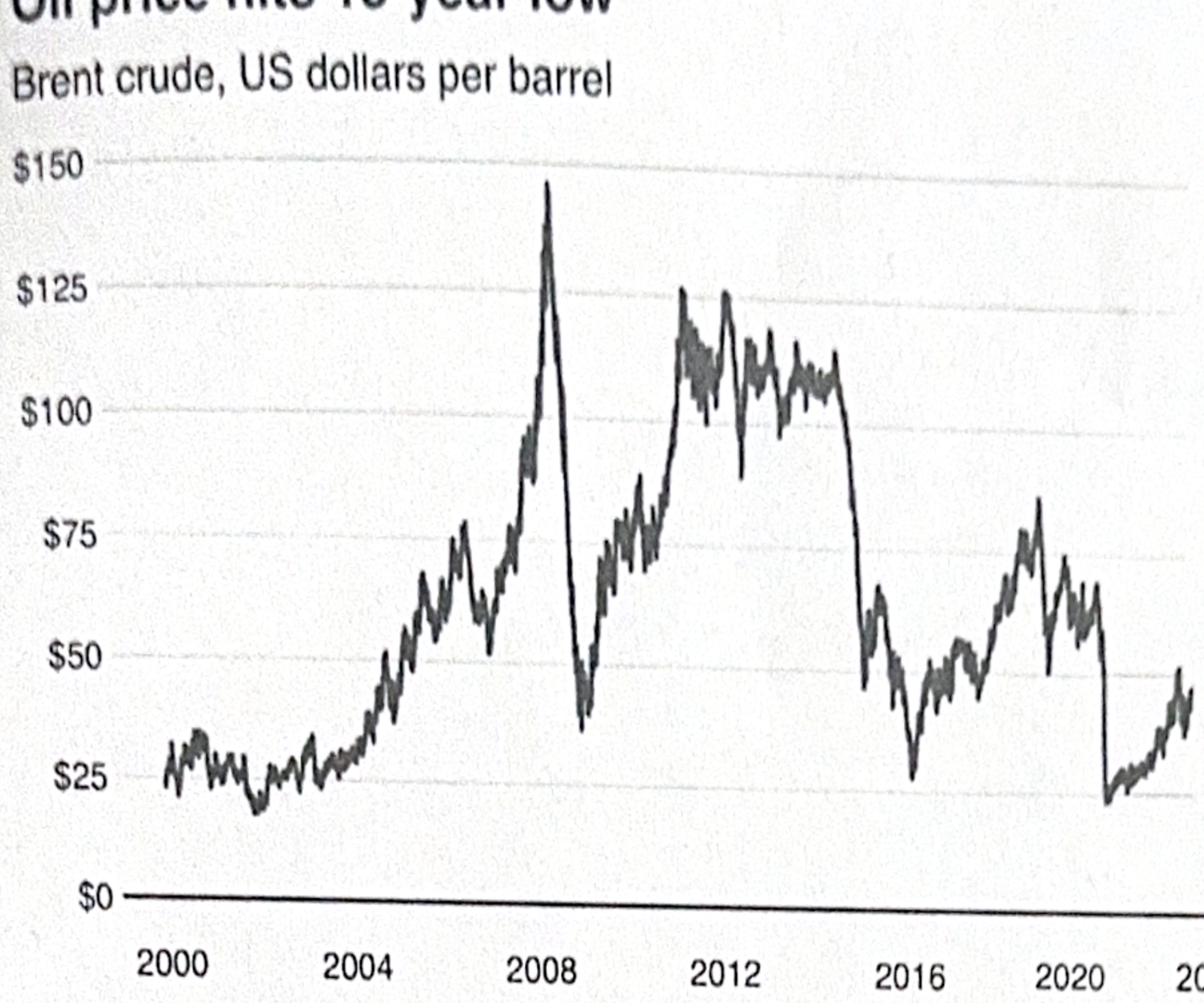
решение принимаемое одним, влияет на решение другого (и наоборот) что в итоге определяет конечный исход

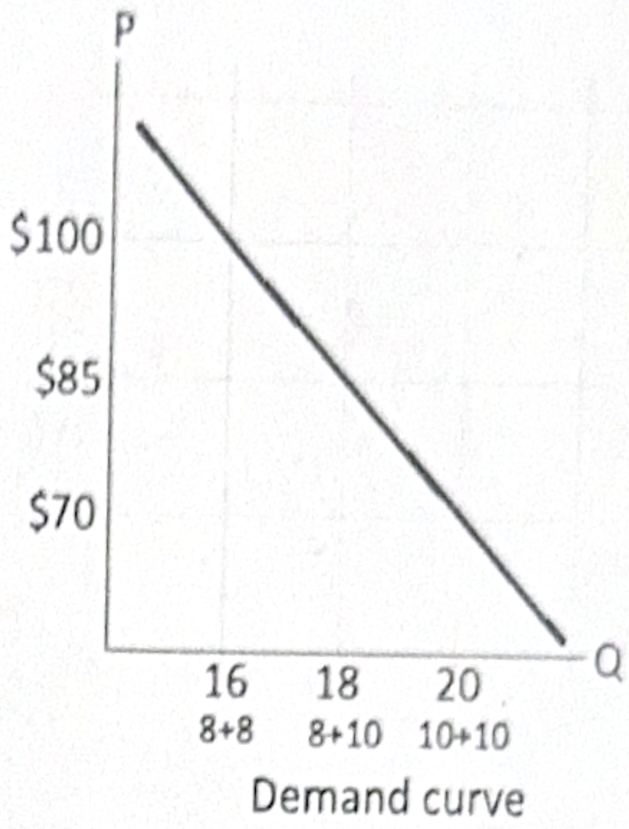
Равновесие Нэша - ситуация, когда стратегия игрока (каждого) явл. наилучшей реакцией на стратегии других игроков и ни одному из них невыгодно в отдельности менять свою стратегию.

Оптимум по Парето — это ситуация, когда нельзя улучшить положение хотя бы одного игрока, не ухудшив при этом положение другого.

Главный парадокс — равновесие Нэша может быть Парето - неэффективным.

Индивидуально рациональное решение приводит к коллективно худшему результату





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

Оптимум по Парето



Равновесие Нэша

Равновесие Нэша здесь - это в нашем случае комбинация 2-2 — $10 \cdot 10^6$

Парето - оптимальный исход обе вабирают $8 \cdot 10^6$