



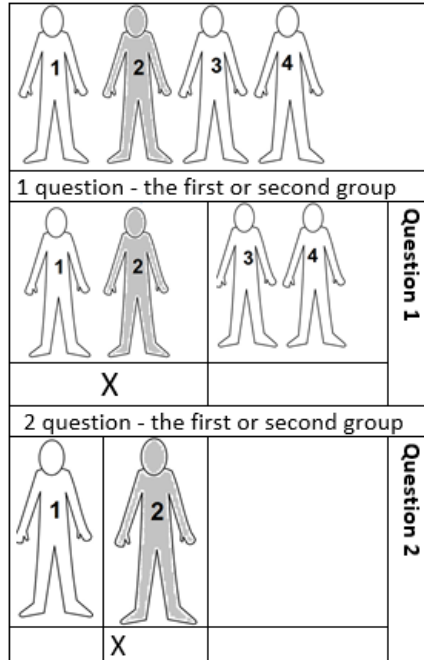
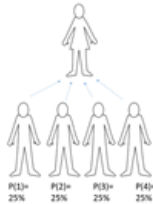
Say **NO** to the first



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

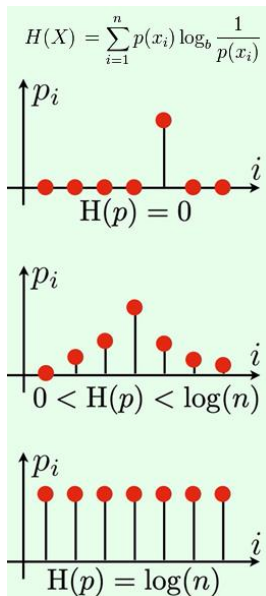


Average number of questions = $2 * 0.25 + 2 * 0.25 + 2 * 0.25 + 2 * 0.25 = 2$

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



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

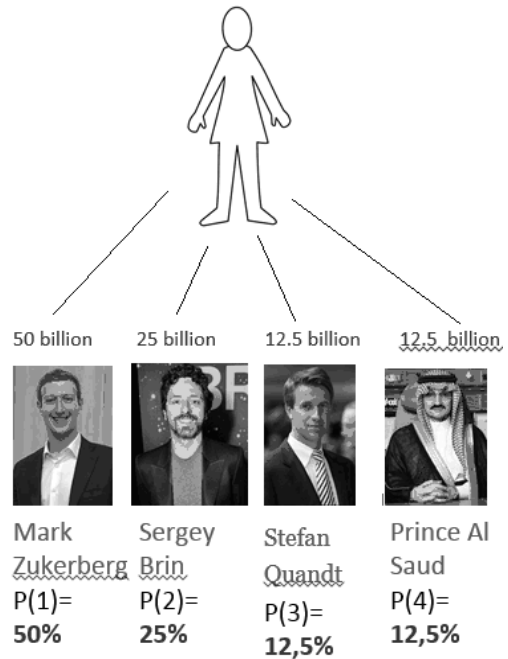
$$\sum_{i=1}^n p(i) \log_2 \frac{1}{p(i)}$$

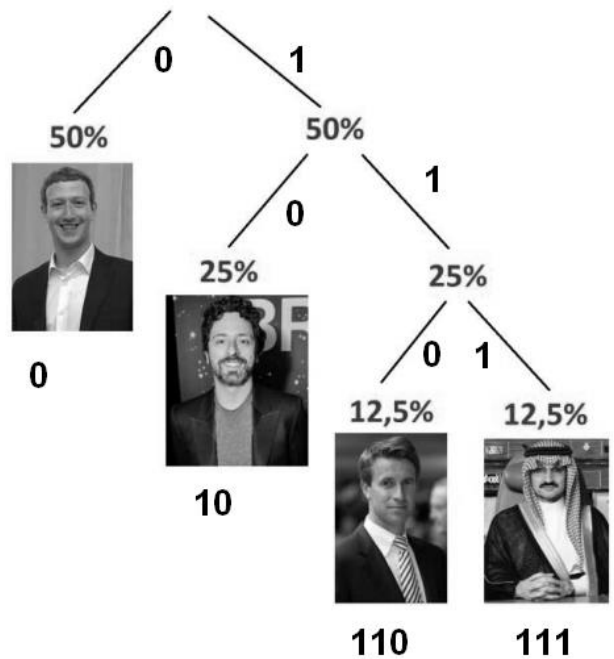
Quantifying information

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

number of bits required to encode choice

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





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	мама	р

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		

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



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

