

0**1**1**1**0 – 5-ти
значное число.

L

0 A B C D E F G
H I J K L M

N O P Q R S T
U V W X Y Z

A B C D E F G

1 H I J K L M

H I J

1 K L M

K

1 L M

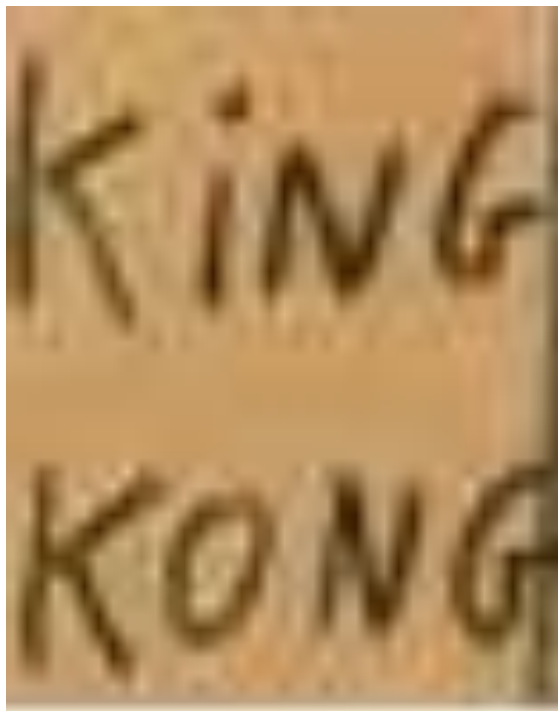
0 L

M

$2^{\text{количество_вопросов}} = \text{размер_сообщения}$

$$2^? = 26$$

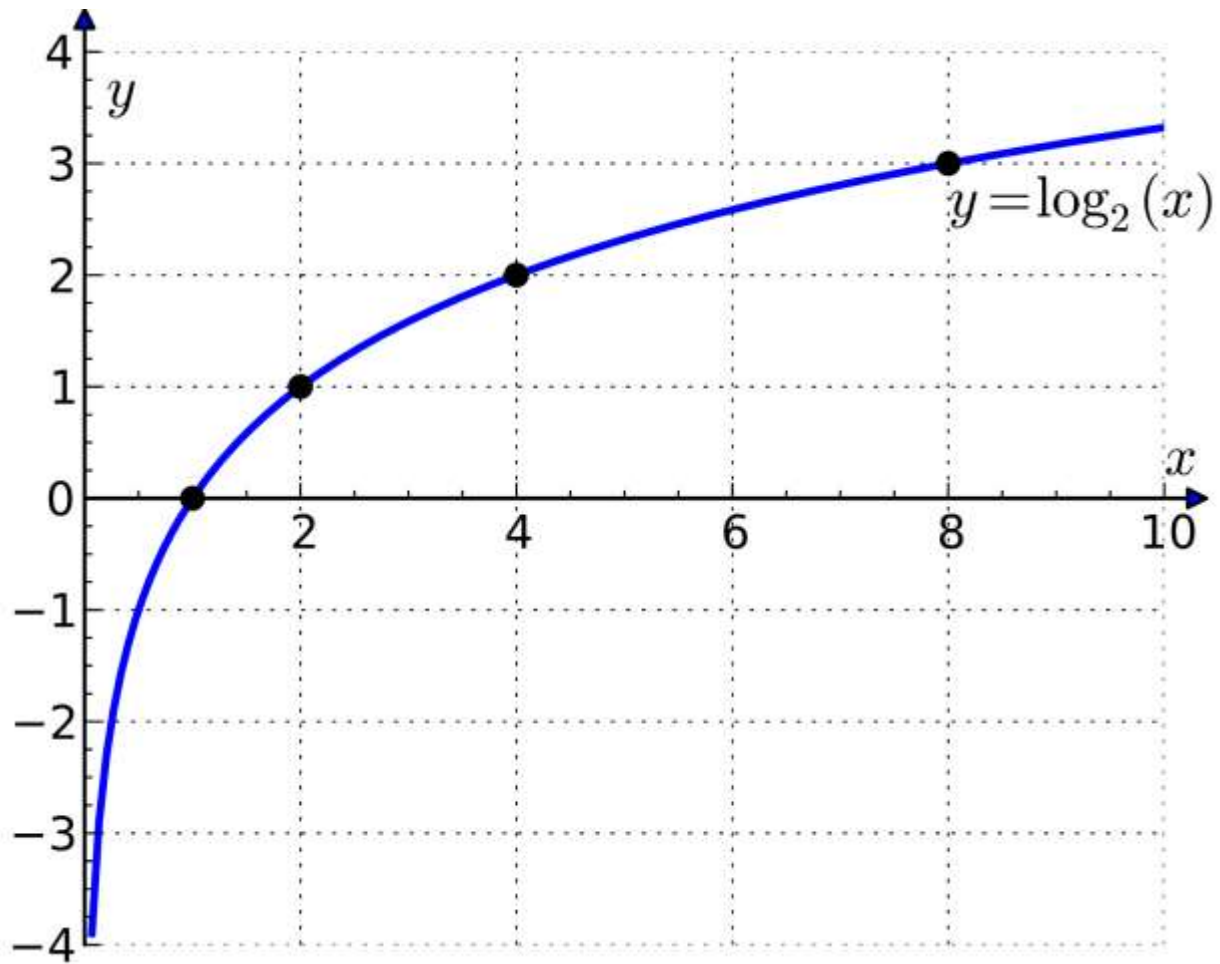
$$? = \log_2 26 = 4,7$$



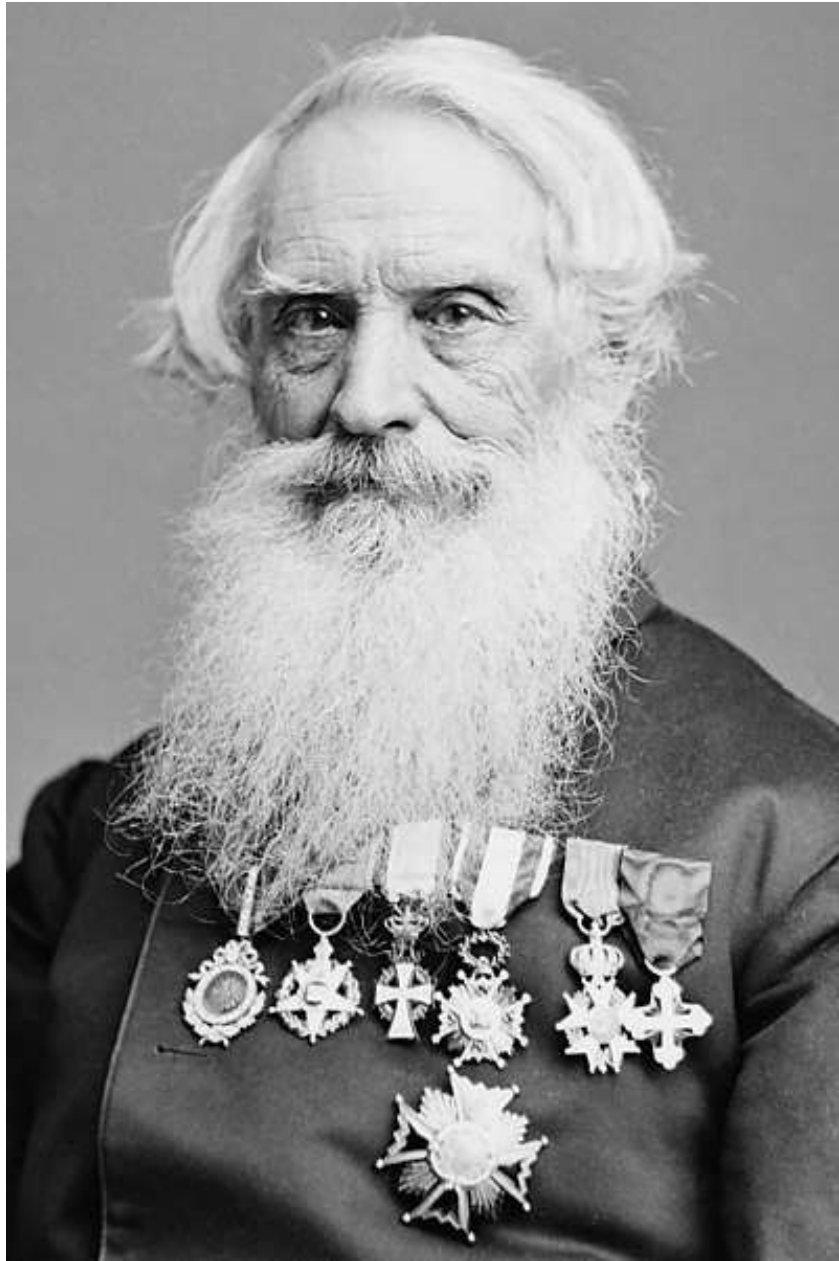
$$4,7 * 8 = 37.$$

6





$$\lg(26) = 4,7$$

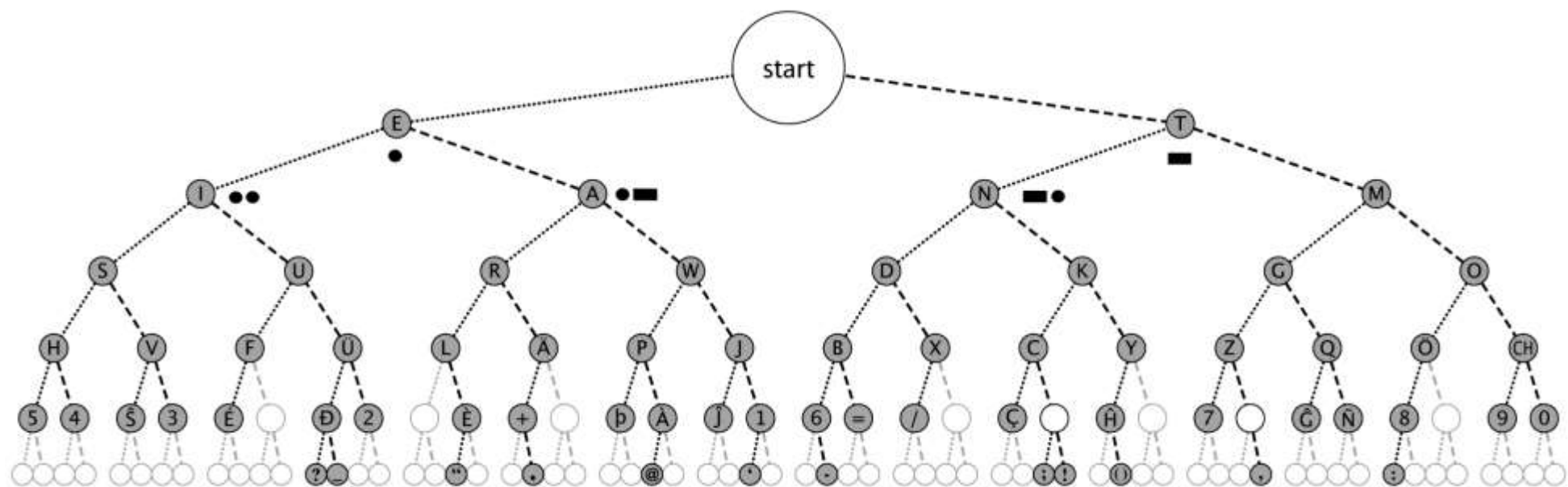


A	● ■■	N	■■ ●
B	■■ ● ● ●	O	■■ ■■ ■■
C	■■ ● ■■ ●	P	● ■■ ■■ ●
D	■■ ● ●	Q	■■ ■■ ● ■■
E	●	R	● ■■ ●
F	● ● ■■ ●	S	● ● ●
G	■■ ■■ ●	T	■■
H	● ● ● ●	U	● ● ■■
I	● ●	V	● ● ● ■■
J	● ■■ ■■ ■■	W	● ■■ ■■
K	■■ ● ■■	X	■■ ● ● ■■
L	● ■■ ● ●	Y	■■ ● ■■ ■■
M	■■ ■■	Z	■■ ■■ ● ●

1838 г. Сэмюэл Морзе

Буква	E	T	A	O	I	N	S	H	R	D	L	C
Частота, %	12,7	9,06	8,17	7,51	6,97	6,75	6,33	6,09	5,99	4,25	4,03	2,78
	●	■	●■	■■■■	●●	■●	●●●	●●●●	●■■●	■■●●		

U	M	W	F	G	Y	P	B	V	K	X	J	Q	Z
2,76	2,41	2,36	2,23	2,02	1,97	1,93	1,49	0,98	0,77	0,15	0,15	0,1	0,05
					■■●■■■	●■■■■●	■■●●●	●●●■■	■■●■■	■■●●■■	●■■■■■	■■■■●■■	■■■■●●



Энтропийное кодирование.



Mark
Zuckerberg

$P(1)=$

50%



Sergey
Brin

$P(2)=$

25%



Stefan
Quandt

$P(3)=$

12,5%



Prince Al
Saud

$P(4)=$

12,5%

A Method for the Construction of Minimum-Redundancy Codes*

DAVID A. HUFFMAN⁺, ASSOCIATE, IRE

Summary—An optimum method of coding an ensemble of messages consisting of a finite number of members is developed. A minimum-redundancy code is one constructed in such a way that the average number of coding digits per message is minimized.

will be defined here as an ensemble code which, for a message ensemble consisting of a finite number of members, N , and for a given number of coding digits, D , yields the lowest possible average message length. In order to avoid the use of the lengthy term “minimum-

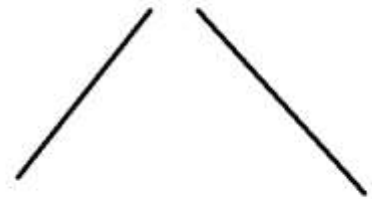
INTRODUCTION

TABLE I
OPTIMUM BINARY CODING PROCEDURE

Original Message Ensemble	Message Probabilities											
	1	2	3	4	5	6	7	8	9	10	11	12
	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20	0.20
	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18	0.18
	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10	0.10
	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06
	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04	0.04
	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03	0.03
	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01



25%



50%



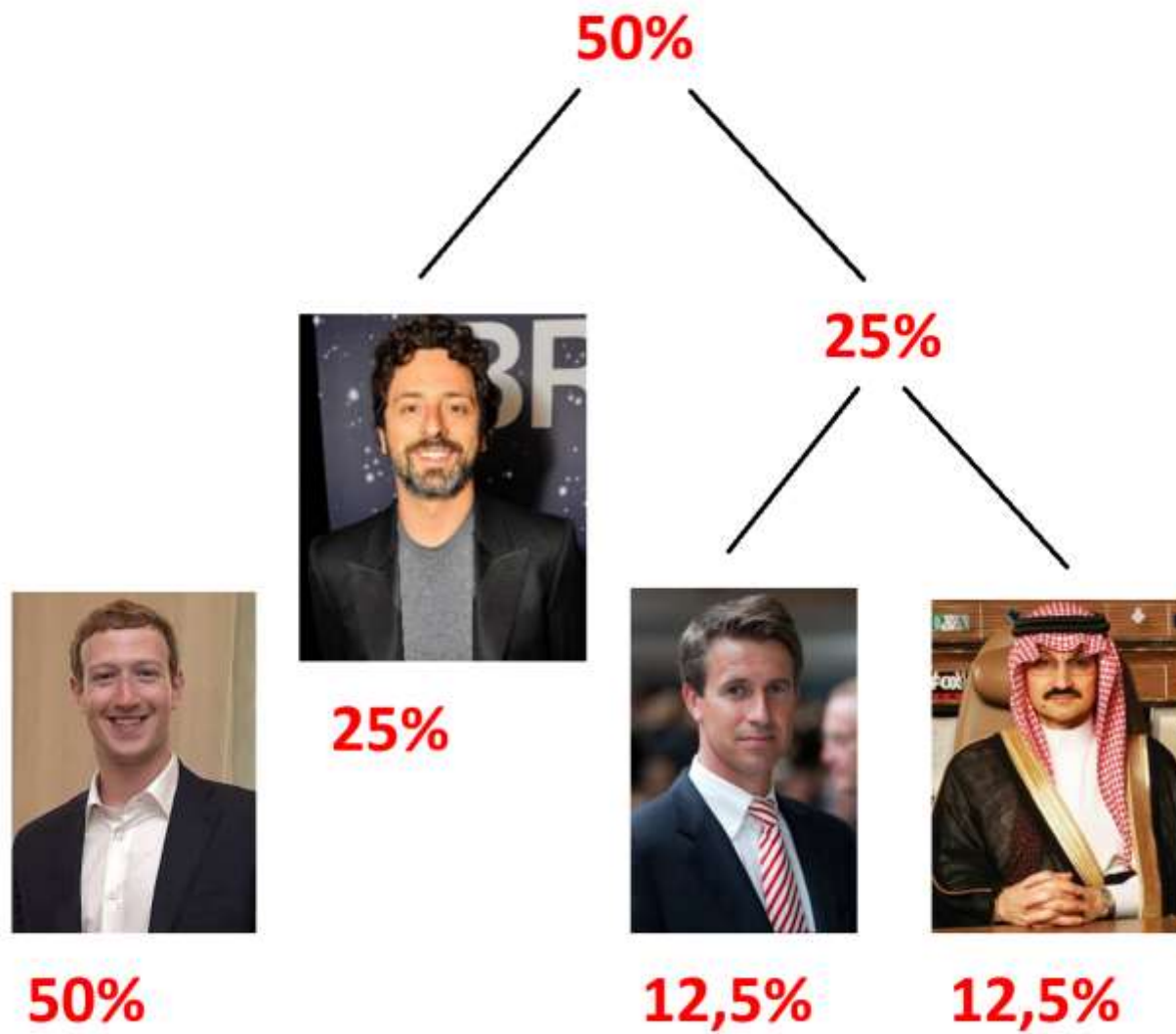
25%

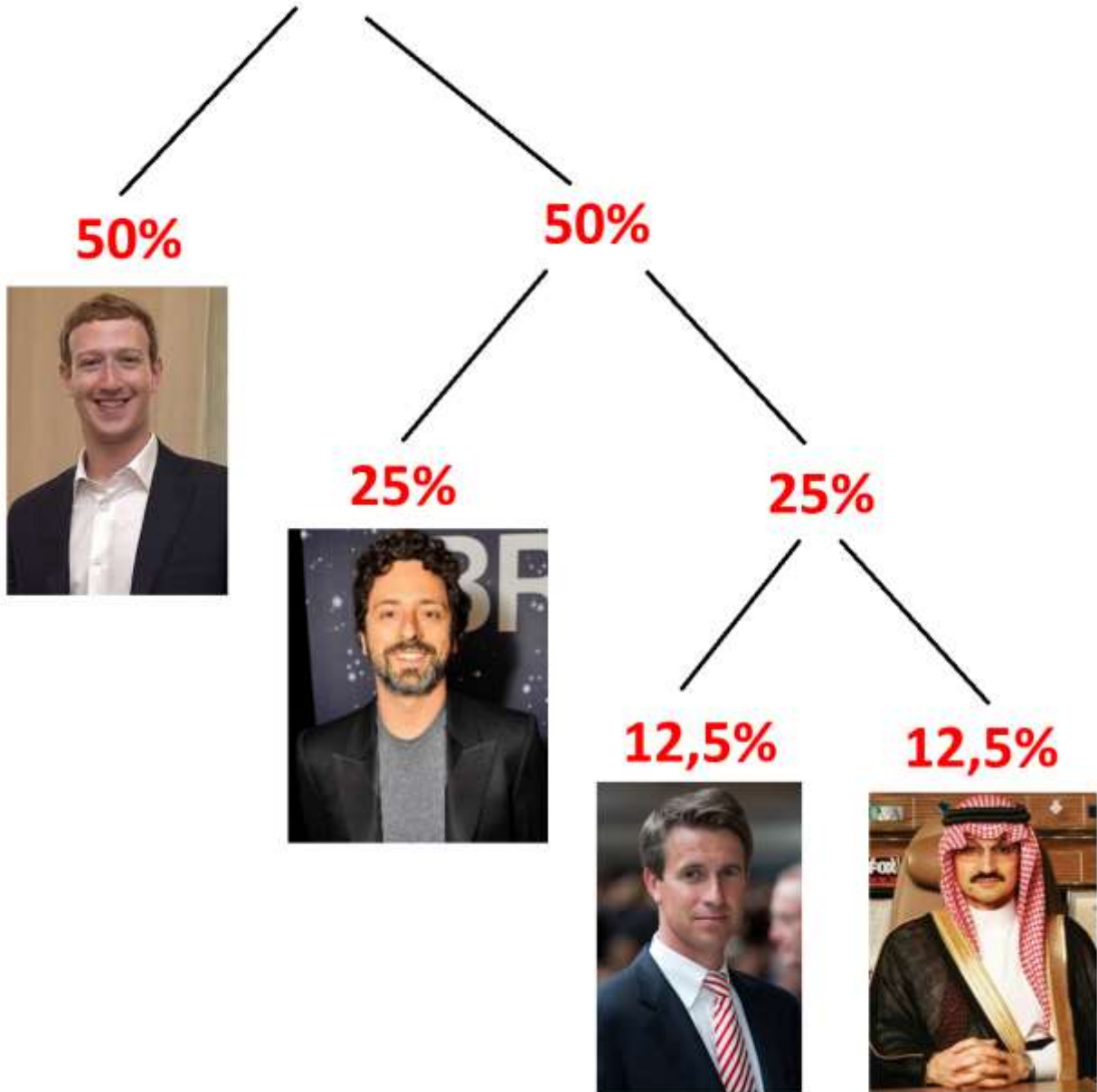


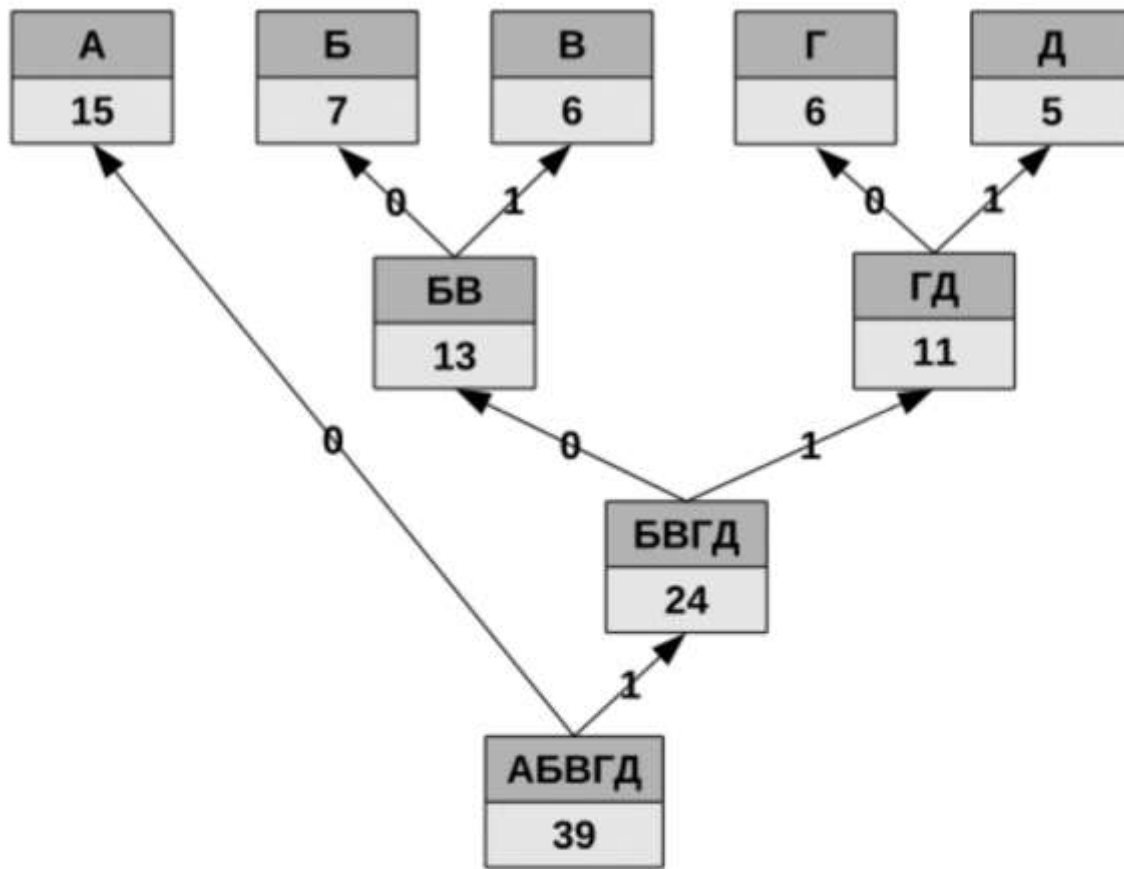
12,5%



12,5%







Итого:

А	Б	В	Г	Д
0	100	101	110	111

https://upload.wikimedia.org/wikipedia/commons/0/07/Huffmantree_ru_animated.gif - на буквах.

Шеннон: предел сжатия не меньше энтропии источника сообщений.

Кодирование Хаффмана широко применяется при сжатии данных, в том числе при сжатии фото- и видеоизображений ([JPEG](#), [MPEG](#)), в популярных архиваторах ([PKZIP](#), [LZH](#) и др.), в протоколах передачи данных HTTP ([Deflate](#)), MNP5 и MNP7 и других.