



Путь свободен

Сколько же
там бит?

Содержательный подход



Путь свободен

Наступает одно из двух событий:
«путь свободен» или «путь
закрыт».

Получен 1 бит.

А	Б	В	Г
001	010	011	100

1 бит это один разряд двоичного
кода

ГАВ = 100001011 ← 9 бит



Двоичное кодирование — это кодирование
с помощью двух знаков (например 1 и 0).

Пусть длина кода равна $i = 2$, сколько же букв русского алфавита N можно закодировать с помощью них?

А — 00, Б — 01, В — 10, Г — 11

$N = 4$

Если $i = 4$, то

А — 0000, Б — 0001 ... О — 1111

$N = 16$

$$N = 2^i$$

В русском алфавите 32 буквы (если считать, что е = ё). Тогда длина кода будет равна 5.

Если учитывать регистр (большие и маленькие буквы), то уже нужен код длиной 6 ($64 = 2^6$).

В обычном тексте мы используем ещё знаки препинания, цифры, математические знаки и латинские буквы.

В стандартную кодовую **таблицу ASCII** мощностью 256 знаков помещаются все символы, необходимые для стандартного текста. $256 = 28$. Каждый символ имеет информационный вес 8 бит или 1 байт, поэтому эту таблицу называют **однобайтовой**.

символ	10- в код	2-Б код	символ	10- в код	2-Б код	символ	10- в код	2-Б код	символ	10- в код	2-Б код
	32	00100000	8	56	00111000	P	80	01010000	h	104	01101000
!	33	00100001	9	57	00111001	Q	81	01010001	i	105	01101001
"	34	00100010	:	58	00111010	R	82	01010010	j	106	01101010
#	35	00100011	;	59	00111011	S	83	01010011	k	107	01101011
\$	36	00100100	<	60	00111100	T	84	01010100	l	108	01101100
%	37	00100101	=	61	00111101	U	85	01010101	m	109	01101101
&	38	00100110	>	62	00111110	V	86	01010110	n	110	01101110
'	39	00100111	?	63	00111111	W	87	01010111	o	111	01101111
(40	00101000	@	64	01000000	X	88	01011000	p	112	01110000
)	41	00101001	A	65	01000001	Y	89	01011001	q	113	01110001
*	42	00101010	B	66	01000010	Z	90	01011010	r	114	01110010
+	43	00101011	C	67	01000011	[91	01011011	s	115	01110011
,	44	00101100	D	68	01000100	\	92	01011100	t	116	01110100
-	45	00101101	E	69	01000101]	93	01011101	u	117	01110101
.	46	00101110	F	70	01000110	^	94	01011110	v	118	01110110
/	47	00101111	G	71	01000111	_	95	01011111	w	119	01110111
0	48	00110000	H	72	01001000	~	96	01100000	x	120	01111000
1	49	00110001	I	73	01001001	a	97	01100001	y	121	01111001
2	50	00110010	J	74	01001010	b	98	01100010	z	122	01111010
3	51	00110011	K	75	01001011	c	99	01100011	{	123	01111011
4	52	00110100	L	76	01001100	d	100	01100100		124	01111100
5	53	00110101	M	77	01001101	e	101	01100101	}	125	01111101
6	54	00110110	N	78	01001110	f	102	01100110	~	126	01111110
7	55	00110111	O	79	01001111	g	103	01100111	□	127	01111111

Информационный объём теста
в однобайтовой кодировке равен
количеству знаков в тексте.

Для кодирования большего
количества символов используют
двухбайтовую таблицу Unicode.





Для измерения количества информации в сообщении в двоичной кодировке нужно:

1. Определить количество знаков в алфавите или мощность алфавита N .
2. Определить количество битов информации i , приходящихся на один символ, — информационная ёмкость (объём) символа из формулы $N = 2^i$.
3. Количество информации в сообщении I длиной k знаков будет равно $I = k \cdot i$.

Определите объём информации в слове ВАСЯ, для кодирования которого использовался алфавит русского языка, состоящий из 32 прописных букв.

АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦЧШЩЪЫЬ
ЭЮЯ

$N = 32$

ВАСЯ

$k = 4$

$I = 4 * 5 = 20$ бит.

$i = 5$ бит ($32 = 2^5$)

Единицы измерения информации

Бит — это наименьшая единица информации.

Байт = 8 бит. Это связано с тем, что ячейки памяти первых компьютеров (регистры) были 8 разрядными.

1 байт = 8 бит

1 Кбайт = 1024 байта

1 Мбайт = 1024 Кбайт

1 Гбайт = 1024 Мбайт

1 Тбайт = 1024 Гбайт

1 Пбайт = 1024 Тбайт

$$1024 = 2^{10} = 10000000000_2$$