



ТОМСКИЙ
ПОЛИТЕХНИЧЕСКИЙ
УНИВЕРСИТЕТ



Юргинский
технологический институт

Направление 09.03.03

Информатика 1.2

Лекция «Кодирование данных»

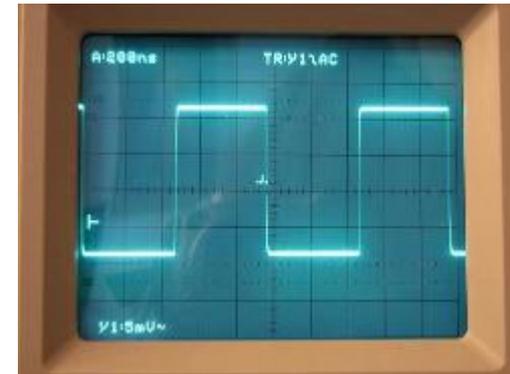
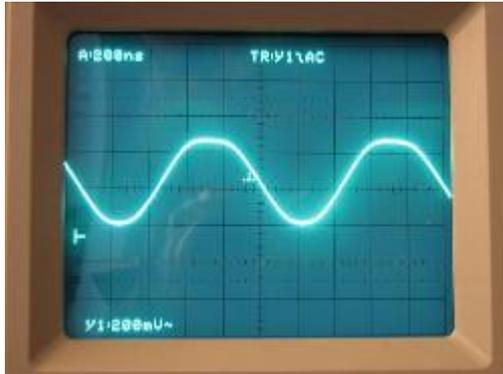
Лектор Молнина Елена Владимировна
Старший преподаватель кафедры Информационных
систем,
ауд.9, гл.корпус.
mail: molnina@list.ru

Кодирование информации

Заголовок фрагмента лекции	Ссылка на запись вебинара
Информация и ее кодирование	http://www.youtube.com/watch?v=IgRmt2JnuWM
Технологии обработки и хранения информации. Часть 1.	http://www.youtube.com/watch?v=gnCzEKPfOFs
Технологии обработки и хранения информации. Часть 2.	http://www.youtube.com/watch?v=G9KE7Rc5vWs
Комбинаторика и теория информации	http://www.youtube.com/watch?v=Da8zz1hhDP0

Цифровой сигнал

Объёмный способ измерения информации



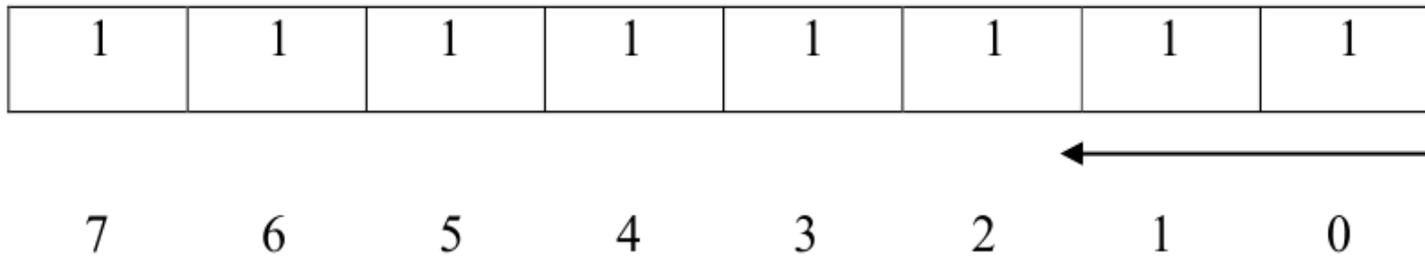
1 символ – 1 бит. Бит принимает значения ноль или единица. Поэтому один бит кодирует только два значения: 0 или 1.

1 байт	2^3 битов					
1 Кбайт	2^{13} битов	2^{10} байтов				
1 Мбайт	2^{23} битов	2^{20} байтов	2^{10} Кбайт			
1 Гбайт	2^{33} битов	2^{30} байтов	2^{20} Кбайт	2^{10} Мбайт		
1 Тбайт	2^{43} битов	2^{40} байтов	2^{30} Кбайт	2^{20} Мбайт	2^{10} Гбайт	
1 Пбайт	2^{53} битов	2^{50} байтов	2^{40} Кбайт	2^{30} Мбайт	2^{20} Гбайт	2^{10} Тбайт

Зависимость между количеством бит и количеством кодируемых значений

Количество бит	1	2	3	4	5	6	7	8
Количество кодируемых значений	2	4	8	16	32	64	128	256

Байт – это название группы из восьми бит



Кодирование текстовой информации

ASCII (American Standard Code for Information Interchange – Американский стандартный код для обмена информацией).

С 1997 г. введён новый международный стандарт ЮНИКОД (Unicode), который отводит на символ не один, а два байта. Эта кодировка используется для кодировки любых символов мира в Microsoft Windows, Microsoft Office.

ASCII Table

Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char	Dec	Hex	Oct	Char
0	0	0		32	20	40	[space]	64	40	100	@	96	60	140	`
1	1	1		33	21	41	!	65	41	101	A	97	61	141	a
2	2	2		34	22	42	"	66	42	102	B	98	62	142	b
3	3	3		35	23	43	#	67	43	103	C	99	63	143	c
4	4	4		36	24	44	\$	68	44	104	D	100	64	144	d
5	5	5		37	25	45	%	69	45	105	E	101	65	145	e
6	6	6		38	26	46	&	70	46	106	F	102	66	146	f
7	7	7		39	27	47	'	71	47	107	G	103	67	147	g
8	8	10		40	28	50	(72	48	110	H	104	68	150	h
9	9	11		41	29	51)	73	49	111	I	105	69	151	i
10	A	12		42	2A	52	*	74	4A	112	J	106	6A	152	j
11	B	13		43	2B	53	+	75	4B	113	K	107	6B	153	k
12	C	14		44	2C	54	,	76	4C	114	L	108	6C	154	l
13	D	15		45	2D	55	-	77	4D	115	M	109	6D	155	m
14	E	16		46	2E	56	.	78	4E	116	N	110	6E	156	n
15	F	17		47	2F	57	/	79	4F	117	O	111	6F	157	o
16	10	20		48	30	60	0	80	50	120	P	112	70	160	p
17	11	21		49	31	61	1	81	51	121	Q	113	71	161	q
18	12	22		50	32	62	2	82	52	122	R	114	72	162	r
19	13	23		51	33	63	3	83	53	123	S	115	73	163	s
20	14	24		52	34	64	4	84	54	124	T	116	74	164	t
21	15	25		53	35	65	5	85	55	125	U	117	75	165	u
22	16	26		54	36	66	6	86	56	126	V	118	76	166	v
23	17	27		55	37	67	7	87	57	127	W	119	77	167	w
24	18	30		56	38	70	8	88	58	130	X	120	78	170	x
25	19	31		57	39	71	9	89	59	131	Y	121	79	171	y
26	1A	32		58	3A	72	:	90	5A	132	Z	122	7A	172	z
27	1B	33		59	3B	73	;	91	5B	133	[123	7B	173	{
28	1C	34		60	3C	74	<	92	5C	134	\	124	7C	174	
29	1D	35		61	3D	75	=	93	5D	135]	125	7D	175	}
30	1E	36		62	3E	76	>	94	5E	136	^	126	7E	176	~
31	1F	37		63	3F	77	?	95	5F	137	_	127	7F	177	

Кодирование чисел в ЭВМ

Числа в компьютере могут быть представлены двумя способами:

- как целые числа для целых чисел;
- как числа с плавающей точкой для вещественных (действительных) чисел.

Целые числа могут быть представлены в компьютере без знака или со знаком.

Целые числа без знака занимают в памяти один или два байта:

- в однобайтовом формате целые числа принимают значения от 0 до 255;
- в двухбайтовом формате целые числа принимают значения от 0 до 65535.

Кодирование чисел со знаком в ЭВМ

Целые числа со знаком могут занимать в памяти один, два или четыре байта. Самый левый (старший) разряд содержит информацию

о знаке числа. Знак «плюс» кодируется нулём, а знак «минус» кодируется единицей.

В компьютерах применяются три формы записи целых чисел со знаком:

- прямой код;
- обратный код;
- дополнительный код.

Кодирование чисел со знаком в ЭВМ

Разр. 7 6 5 4 3 2 1 0
12 0 0 0 0 1 1 0 0

Такая запись соответствует 8-ми разрядному **прямому коду** числа двенадцать. А теперь проинвертируем все разряды регистра, т.е. заменим 0 на 1 и 1 на 0. и получим **обратный код**.

Разр. 7 6 5 4 3 2 1 0
12_{обр} 1 1 1 1 0 0 1 1

Прибавив к числу в обратном коде единицу, получаем искомый **дополнительный код**.

(красным цветом показаны переносы в соответствующий разряд)

Разр. с 7 6 5 4 3 2 1 0
12_{обр} 1 1 1 1 0 0 1 1
+ 1
12_{доп} 1 1 1 1 0 1 0 0



Число со знаком в вычислительной технике представляется путем представления старшего разряда числа в качестве **знакового**. Принято считать, что 0 в знаковом разряде означает знак «плюс» для данного числа, а 1 – знак «минус».

Кодирование вещественных чисел

Для записи вещественных чисел используется форма записи чисел с порядком основания системы счисления.

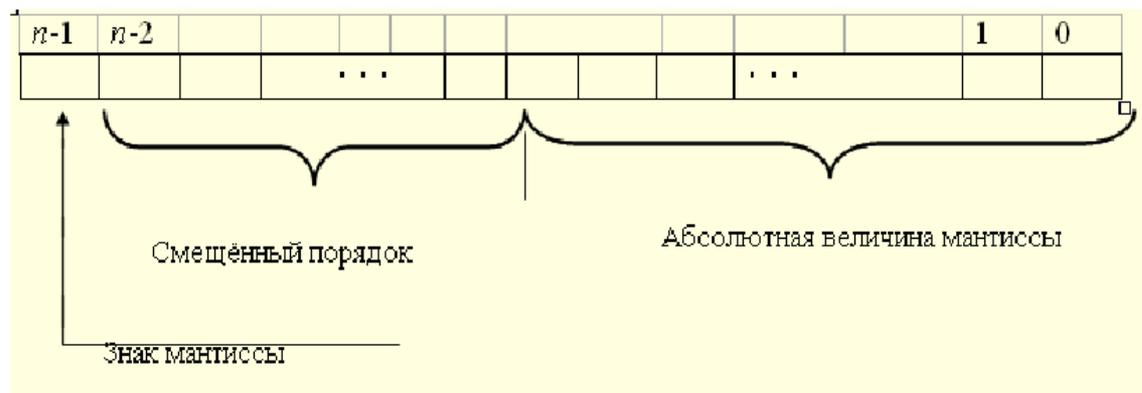
Способ записи чисел называется представлением числа с плавающей точкой, если любое число N в системе счисления с основанием q записано в виде $N = M * q^P$, где q – основание системы счисления; M – множитель, который содержит все цифры числа и называется мантиссой; P – целое число, которое называется порядком.

Пример. Записать десятичные числа как нормализованные:

$$1234.5 = 0.12345 * 10^4;$$

$$-0.000012345 = -0.12345 * 10^{-4}.$$

Вещественные числа записываются в компьютерах по международному стандартному формату. На рис. 10 приведена структура вещественного числа по международному стандартному формату.



Контрольные вопросы и задания

1. Что такое бит?
2. Сколько значений кодируют 6 бит?
3. Что такое байт?
4. Сколько значений кодирует один байт?
5. Что хранится в байте?
6. Что такое килобайт, мегабайт, гигабайт?
7. Какая таблица называется таблицей кодов?
8. Какие способы представления чисел в компьютере вы знаете?
9. Как представляются целые числа в компьютере?
10. Какие три формы записи целых чисел со знаком вы знаете?
11. Как формируется обратный код двоичного отрицательного числа?
12. Как формируется дополнительный код двоичного отрицательного числа?
13. Какие преимущества даёт использование дополнительного кода?
14. Какой способ записи чисел называется представлением числа с плавающей точкой?
15. Какое вещественное число называется нормализованным?
16. Запишите двоичное число 1 0000 0000 в десятичном формате.
17. Укажите, можно ли записать десятичное число (-255) в прямом коде в однобайтовом представлении.
18. Запишите десятичное число -29 в прямом коде, в обратном коде и в дополнительном коде.
19. Запишите число 0,0101333 в нормализованном представлении.

СПАСИБО ЗА ВНИМАНИЕ