
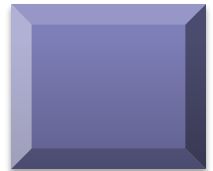




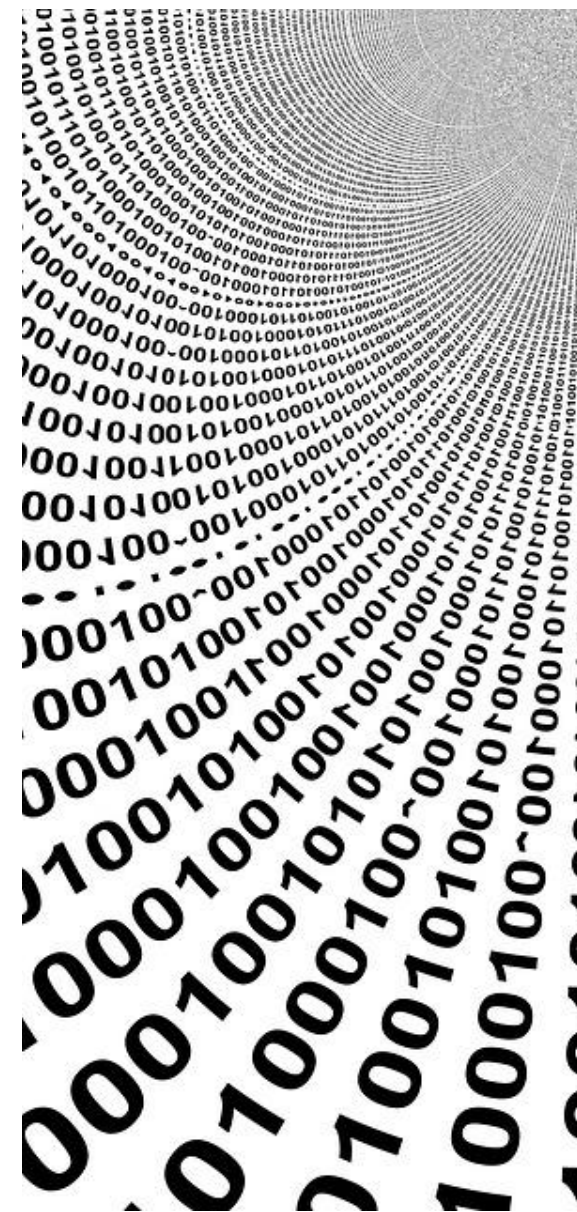
Представление и организация данных в ПК



- 
- Двоичное кодирование
 - Кодирование чисел
 - Кодирование текста
 - Кодирование графики
 - Кодирование звука
 - Организация хранения данных



- **Код** - это система условных знаков для представления информации.
- **Кодирование** - это перевод информации в удобную для передачи, обработки или хранения форму с помощью некоторого кода.
- **Декодирование** - это процесс восстановления содержания закодированной информации.



- Информация для использования обязательно должна быть зашифрована в виде знаков, символов, кодов.
- Для представления информации в ЭВМ нужно было найти способ представить любую ее форму (символьную, текстовую, графическую) в едином виде.



Двоичное кодирование

Двоичная форма представления



• в такой форме информация



ИЛИ



ВСЕ ВИДЫ

• нужны тогда

КОМУСЯ ИЛИ В КАКИХ СИТУАЦИЯХ

• практически не используются в передаче

• компьютеру легче обрабатывать данные
двоичными цифрами или битами



(от англ. *Binary digit* двоичная цифра)
• человеку сложно воспринимать двоичные коды

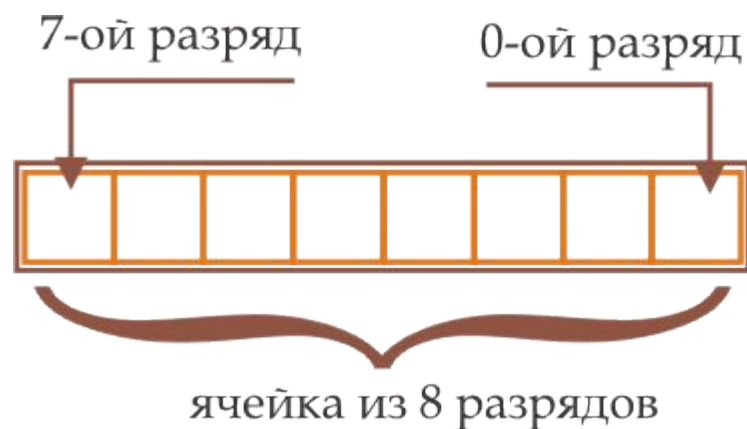
Информационная ёмкость двоичных кодов

Каждый последующий разряд последовательности двоичных цифр имеет уже вдвое больше различных значений, чем предыдущий

Таблица информационной ёмкости

Разряды	1	2	3	4	5	6	7	8	9	10
Кол-во значений	2	4	8							

- При хранении информации в памяти компьютера *каждый бит хранится в одном разряде памяти.*
- Разряды объединяются в ячейки памяти фиксированного размера – 8, 16, 32 разряда и носят специальные названия **байт**, **слово** и **двойное слово**.
- При двоичном кодировании в компьютере нельзя записать ничего, кроме двоичных чисел.

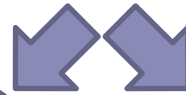


Системы счисления – это способ записи (изображения) чисел.

Непозиционные

Вклад каждой цифры в величину числа **не зависит** от ее положения (позиции), в последовательности цифр его изображающей

XIX



Позиционные

Вклад каждой цифры в величину числа **зависит** от ее положения (позиции), в последовательности цифр его изображающей

0,7; 7; 70

Основание системы счисления

Количество различных символов, используемых для изображения числа в позиционных системах счисления, называется **основанием системы счисления**.

Система счисления	Основание	Алфавит цифр
Десятичная	10	0, 1, 2, 3, 4, 5, 6, 7, 8, 9
Двоичная	2	0, 1
Восьмеричная	8	0, 1, 2, 3, 4, 5, 6, 7
Шестнадцатеричная	16	0, 1, 2, 3, 4, 5, 6, 7, 8, 9, A, B, C, D, E, F

2 1 0

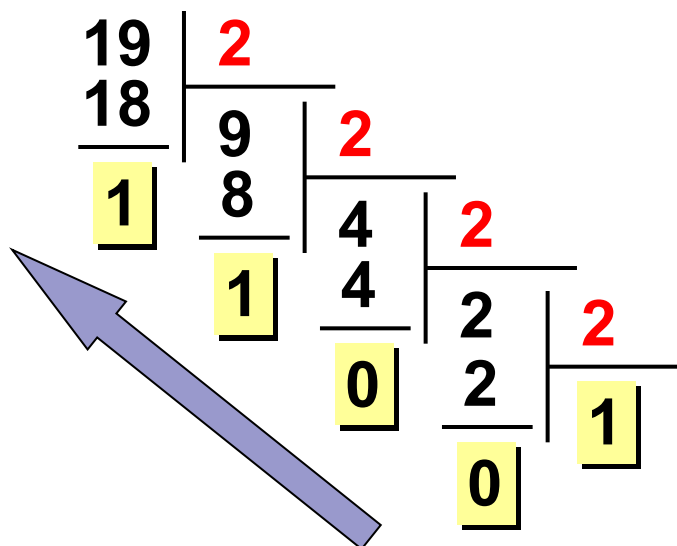
$$123_{10} = 100 + 20 + 3 = 1 \cdot 100 + 2 \cdot 10 + 3 \cdot 1 = \\ = 1 \cdot 10^2 + 2 \cdot 10^1 + 3 \cdot 10^0$$

2 → 10

4 3 2 1 0 разряды

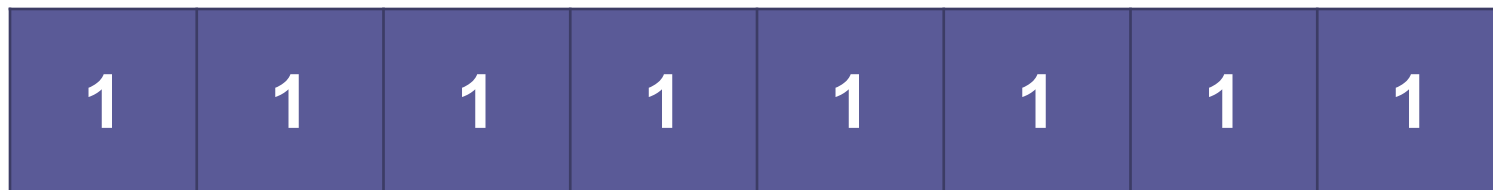
$$10011_2 = 1 \cdot 2^4 + 0 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = \\ = 16 + 2 + 1 = 19_{10}$$

10 → 2



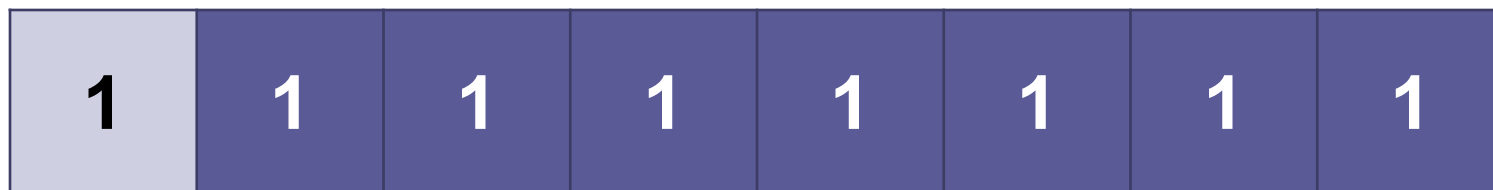
Д/з: добавить к числу года рождения номер месяца и перевести получившееся число в 2-чную, 8-чную и 16-чную системы

Кодирование целых чисел



В 1 байте могут храниться целые неотрицательные числа от 0 до 255

знак



В 1 байте могут храниться целые числа от -128 до -1 и от 0 до 127

Кодирование вещественных чисел

$$0,0123 = 0,123 * 10^{-1}$$

$$123 = 0,123 * 10^3$$

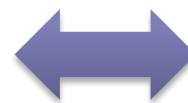


Д/з: Перевести в десятичную систему и вычислить закодированные числа

Кодирование символов

- Каждому символу ставится в соответствие некоторое неотрицательное число, называемое **кодом символа**, и это число записывается в память компьютера в двоичном виде.
- Конкретное соответствие между символами и их кодами называется **системой кодировки**.

1000001_2	1000010_2	1000011_2	1000100_2
65	66	67	68



символ	10-й код	2-й код	символ	10-й код	2-й код	символ	10-й код	2-й код	символ	10-й код	2-й код
	32	00100000	8	56	00111000	P	80	01010000	h	104	01101000

Windows XP Professional

- Microsoft Update
- Выбор программ по умолчанию
- Каталог Windows
- Mail.Ru Агент
- ScreenshotReader
- Far Manager 2
- Программы
 - ArcGIS
 - GIMP
 - Стандартные
 - Служебные
 - ACT-Тест
 - Специальные возможности
 - Dropbox
- Документы
- Настройка
- Найти
- Справка и поддержка
- Выполнить...
- Завершение работы...

Таблица символов

Шрифт: Arial

!	"	#	\$	%	&	'	()	*	+	,	-	.	/	0	1	2	3	4
5	6	7	8	9	:	;	<	=	>	?	@	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	h	i	j	k	l	m	n	o	p
q	r	s	t	u	v	w	x	y	z	{		}	~		ı	ç	£	¤	¥
ı	§	¨	©	ª	«	¬	-	®	¯	°	±	²	³	´	µ	¶	·	,	'
°	»	¼	½	¾	¿	À	Á	Â	Ã	Ä	Å	Æ	Ç	È	É	Ê	Ë	Ì	Í
Î	Ï	Ð	Ñ	Ò	Ó	Ô	Õ	Ö	×	Ø	Ù	Ú	Û	Ü	Ý	Þ	ß	à	á
â	ã	ä	å	æ	ç	è	é	ê	ë	ì	í	î	ï	ð	ñ	ò	ó	ô	õ
ö	÷	ø	ù	ú	û	ü	ý	þ	ÿ	Ā	ā	Ă	ă	Ą	ą	Ć	ć	Ĉ	ĉ

Для копирования:

Дополнительные параметры просмотра

Набор символов: Юникод

Группировка: Все

Поиск:

U+0021: Exclamation Mark

-	45	00101101	E	69	01000101	d	100	01100100		124	01111100
.	46	00101110	F	70	01000110	e	101	01100101	}	125	01111101
/	47	00101111	G	71	01000111	f	102	01100110	~	126	01111110
0	48	00110000	H	72	01001000	g	103	01100111	□	127	01111111
1	49	00110001	I	73	01001001						
2	50	00110010	J	74	01001010						
3	51	00110011	K	75	01001011						
4	52	00110100	L	76	01001100						
5	53	00110101	M	77	01001101						
6	54	00110110	N	78	01001110						
7	55	00110111	O	79	01001111						

Обратите внимание!

Цифры кодируются по стандарту ASCII в двух случаях – при вводе-выводе и когда они встречаются в тексте.

Если цифры участвуют в вычислениях, то осуществляется их преобразование в другой двоичных код.

Например: Возьмем число **57**.

При использовании в тексте каждая цифра будет представлена своим кодом в соответствии с таблицей ASCII. В двоичной системе это – **00110101 00110111**.

При использовании в вычислениях код этого числа будет получен по правилам перевода в двоичную систему и получим – **00111001**.

Решение задач

- Для хранения на диске слова ИНФОРМАТИКА в системе кодирования ASCII необходимо __ БИТ
- В кодируемом английском тексте используется только 26 букв латинского алфавита и еще 6 знаков пунктуации. В этом случае текст, содержащий 1000 символов можно гарантированно сжать без потерь информации до размера
- В текстовом файле хранится текст объемом в 400 страниц. Каждая страница содержит 3200 символов. Если используется кодировка KOI-8 (8 бит на один символ), то размер файла составит

Два типа кодирования рисунков

растровое кодирование

точечный рисунок,
состоит из **пикселей**
(фотографии, размытые
изображения)

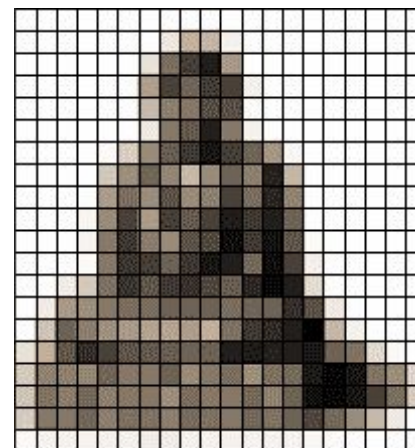
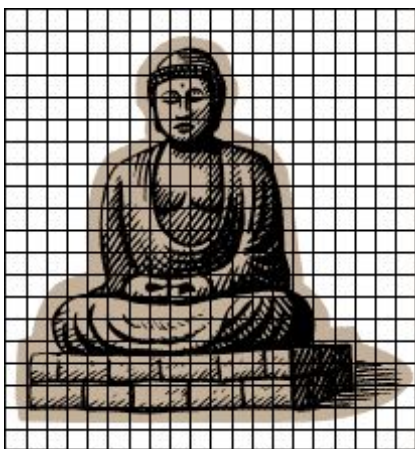


векторное кодирование

рисунок, состоит из
**отдельных
геометрических фигур**
(чертежи, схемы, карты)



Растровое кодирование



Шаг 1. Дискретизация:
разбивка на *пиксели*.

Пиксель – это наименьший элемент рисунка, для которого можно независимо установить цвет.

Разрешение:

dpi (англ. *dots per inch*) — количество точек на дюйм (желаемое)

ppi (англ. *pixels per inch*) — количество пикселей на дюйм
(зависит от устройства)

Шаг 2. Для каждого пикселя определяется **единый цвет**.



Есть потеря информации!

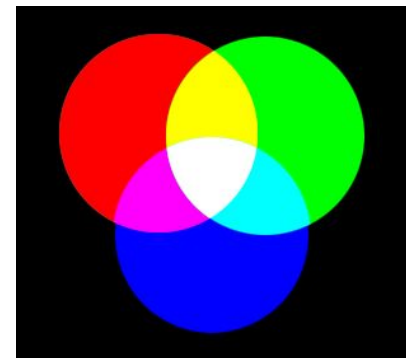
- почему?
- как ее уменьшить?

Растровое кодирование (*True Color*)

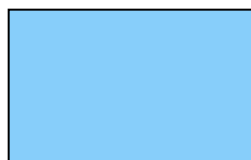
Шаг 3. От цвета – к числам: модель RGB

цвет = **R** + **G** + **B**

<i>red</i>	<i>green</i>	<i>blue</i>
красный	зеленый	синий
0..255	0..255	0..255



R = 218
G = 164
B = 32



R = 135
G = 206
B = 250

Шаг 4. Числа – в двоичную систему.



Сколько разных цветов можно кодировать?

$256 \cdot 256 \cdot 256 = 16\,777\,216$ (*True Color*)

Глубина
цвета



Сколько памяти нужно для хранения цвета 1 пикселя?

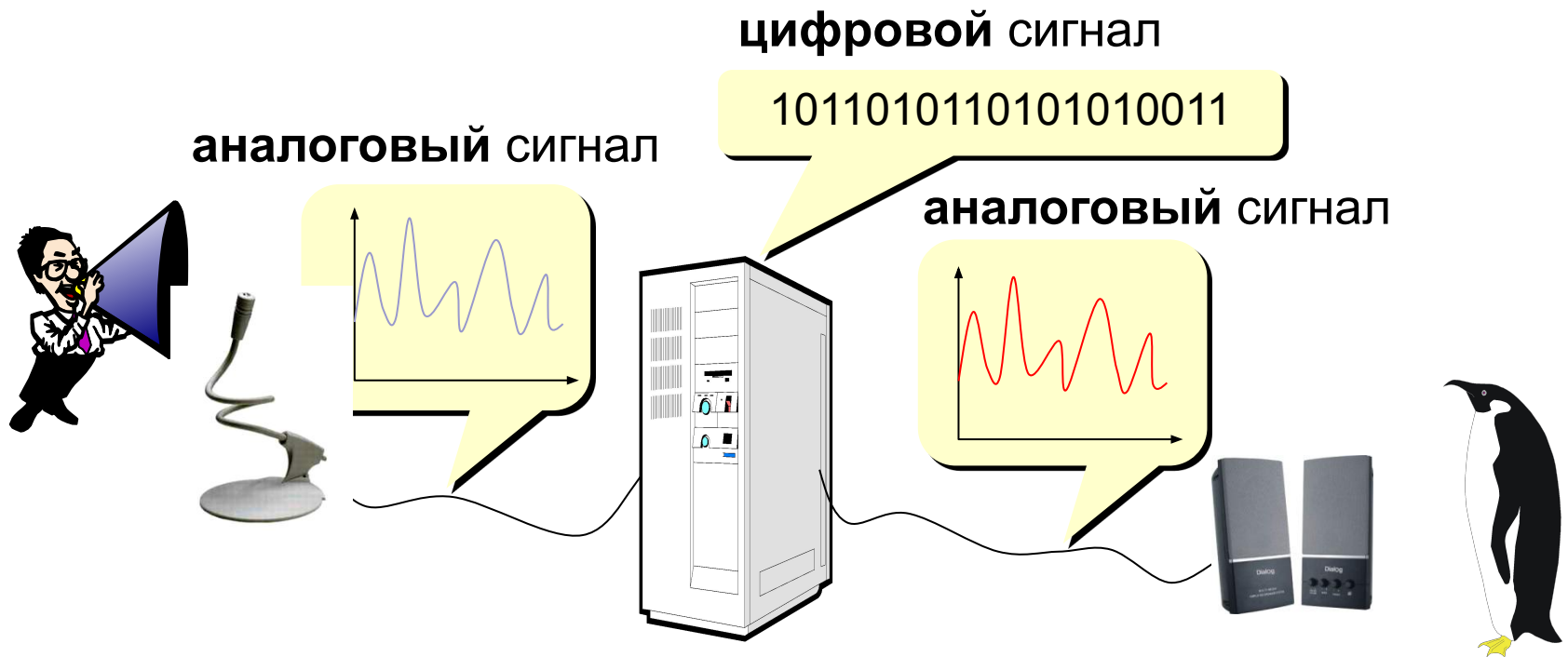
R: $256=2^8$ вариантов, нужно 8 бит = 1 байт
R G B: всего 3 байта

Решение задач

- Растровый графический файл содержит черно-белое изображение с 2 градациями цвета (черный и белый) размером 800х600 точек. Размер этого файла на диске в БАЙТАХ равен ...
- Для цветной картинки, составленной из 256 цветов в графическом режиме монитора 640х480 требуется объем видеопамяти (Кбайт) _____
 - А в цветовой модели True Color?

Кодирование звука и музыки

- Звук представляет собой колебания воздуха.
- Амплитуда колебаний постоянно меняется, так как звук является непрерывным сигналом.



Дискретизация по времени и по уровню

Частота дискретизации: $f = \frac{1}{T}$

$f = 8$ кГц, 11 кГц, 22 кГц, 44 кГц (CD)

Человек слышит 16 Гц ... 20 кГц

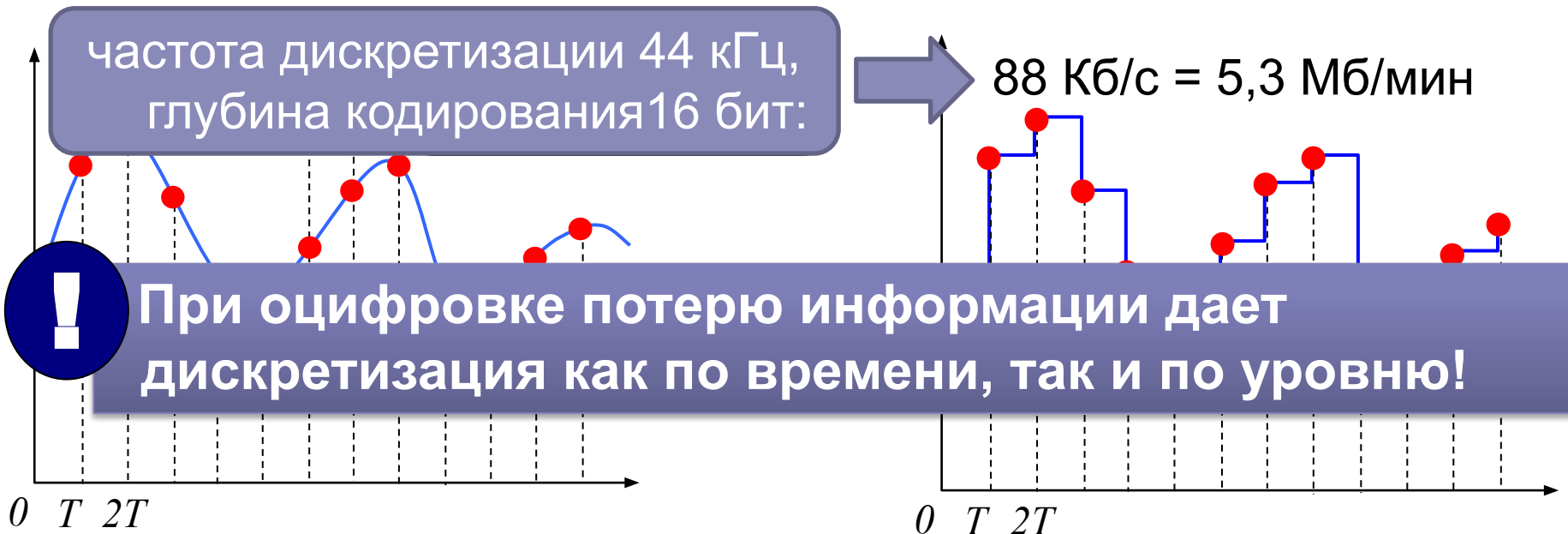
22 кГц $T = \frac{1}{22000} \approx 0,00005$ с

Разрядность квантования («Глубина» кодирования):

8 бит = 256 уровней

16 бит = 65536 уровней

24 бита = 2^{24} уровней

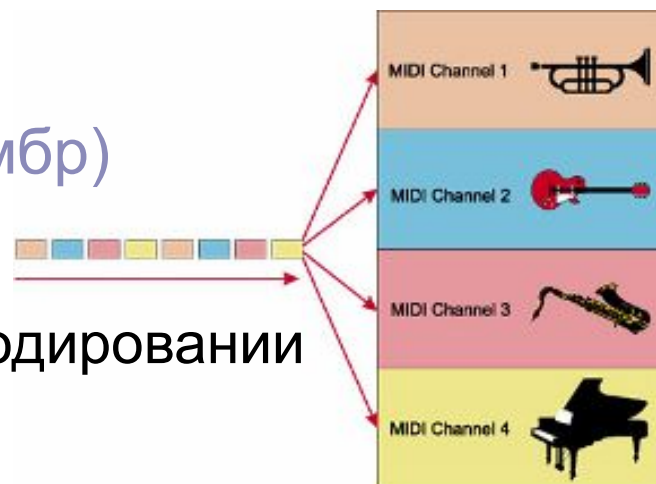


Инструментальное кодирование

MIDI (*Musical Instrument Digital Interface*), файлы ***.MID**

в файле:

- нота (высота, длительность)
- музыкальный инструмент
- параметры звука (громкость, тембр)
- может быть несколько каналов



- **нет потери информации** при кодировании инструментальной музыки
- маленький **размер файлов**



невозможно закодировать нестандартный звук, голос

MIDI-клавиатура:



Единицы представления данных

Двоичные приставки введены Международной электротехнической комиссией (МЭК) в 1999 году. Используются для образования единиц измерения информации, кратных битам и байтам.

Приставка	Аналогичная десятичная приставка	Сокращения по МЭК для битов, байтов	Значение, на которое умножается исходная величина
киби	кило (10^3)	Кибит, КиБ	$2^{10} = 1\ 024$
меби	мега (10^6)	Мибит, МиБ	$2^{20} = 1\ 048\ 576$
гиби	гига (10^9)	Гибит, ГиБ	$2^{30} = 1\ 073\ 741\ 824$
теби	тера (10^{12})	Тибит, ТиБ	$2^{40} = 1\ 099\ 511\ 627\ 776$
пеби	пета (10^{15})	Пибит, ПиБ	$2^{50} = 1\ 125\ 899\ 906\ 842\ 624$
эксби	экса (10^{18})	Эибит, ЭиБ	$2^{60} = 1\ 152\ 921\ 504\ 606\ 846\ 976$
зеби	зетта (10^{21})	Зибит, ЗиБ	$2^{70} = 1\ 180\ 591\ 620\ 717\ 411\ 303\ 424$
йоби	йотта (10^{24})	Йибит, ЙиБ	$2^{80} = 1\ 208\ 925\ 819\ 614\ 629\ 174\ 706\ 176$

Единицы хранения данных

При хранении данных решаются две проблемы:

- как сохранить данные в наиболее компактном виде
- как обеспечить к ним удобный и быстрый доступ

В качестве ***единицы хранения данных*** принят объект переменной длины, называемый **файлом**.

Файл – это последовательность произвольного числа байтов, обладающая уникальным собственным именем.

Имя файла

Полное имя файла:

D:\Студенты\Иванов Иван\Лаб.раб.№2.docx



Путь к
файлу

Собственное
ИМЯ


Расширение


Местонахождение файла
в иерархической системе
каталогов

Название файла
в ОС Windows
до 255 символов

Может состоять
не более чем из
четырёх символов

Отображение


Текстовый
документ.txt


Текстовый
документ.txt

Значок

Ярлык

обозначает
тип
хранящейся
информации
txt, exe, avi

определяется
приложением,
в котором
создан документ
doc, xls, ppt

Файловая система

Файловая система – функциональная часть ОС, организует работу с файлами, отвечает за хранение данных на внешних носителях и обмен данными между внешними носителями.

FAT (File Allocation Table)

FAT32 (File Allocation Table 32)

NTFS (New Technology File System)

HPFS (High Performance File System)

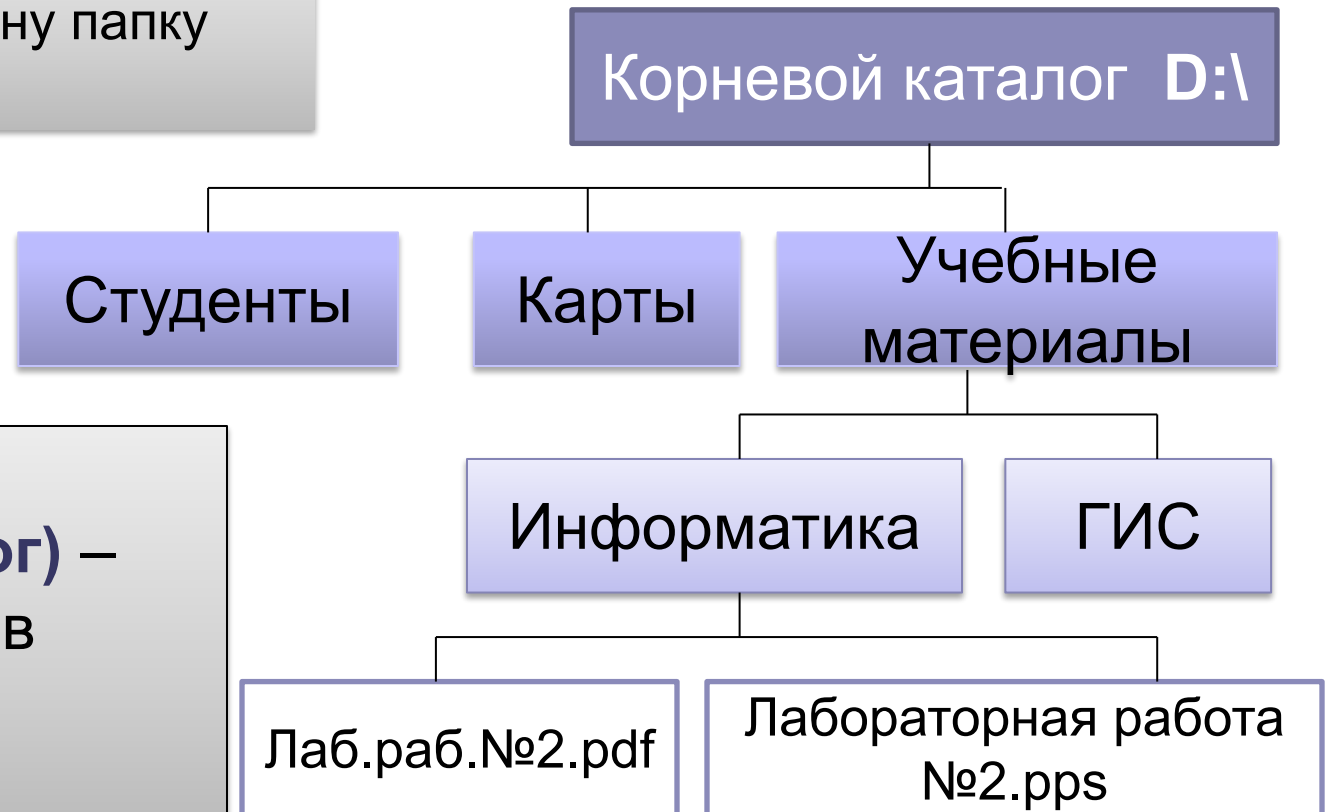
NetWare File System

Linux **Ext2** и Linux **Swap**

Д/з: Найти и
выписать
определения и
основные
характеристики
FAT32 и **NTFS**

Файловая система имеет иерархическую (древовидную) структуру

Папка или файл может входить лишь в одну папку верхнего уровня.



Рабочая папка (текущий каталог) – папка, с которой в данный момент работает

ПОЛЬЗОВАТЕЛЬ



Организация хранения файлов на жестких дисках