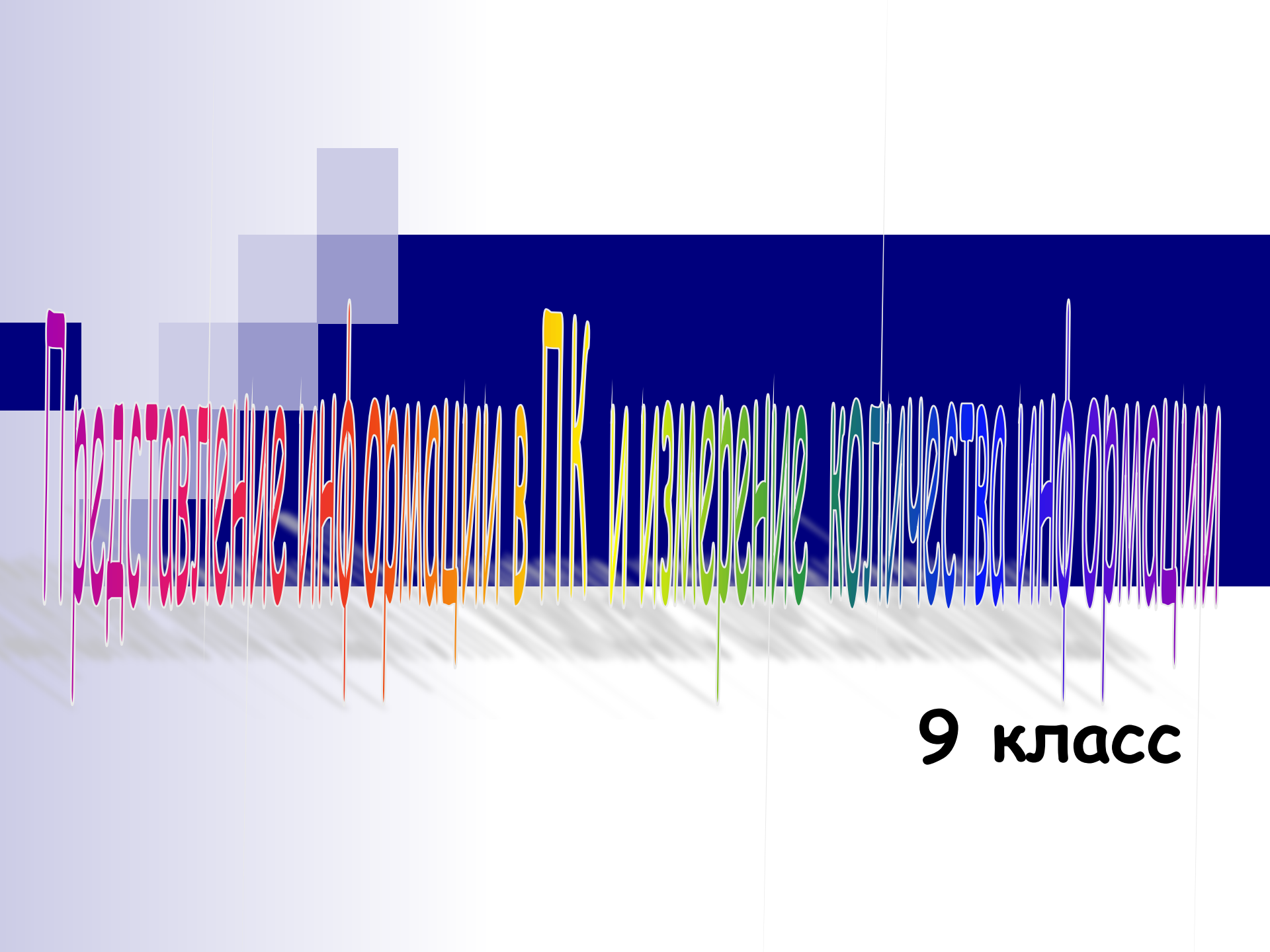




представление информации в ПК и измерение количества информации

9 класс

ДВОИЧНЫЙ КОД

- Вся информация, которую обрабатывает компьютер, представлена **ДВОИЧНЫМ КОДОМ** помощью двух цифр – **0 и 1**
- Эти два символа 0 и 1 принято называть битами
(от англ. **binary digit** – двоичный знак)

Кодирование – преобразование входной информации в машинную форму (в двоичный код)

Декодирование – преобразование двоичного кода в понятную человеку

Почему двоичное кодирование

С точки зрения технической реализации использование двоичной системы счисления для кодирования информации оказалось намного более простым, чем применение других способов. Действительно, удобно кодировать информацию в виде последовательности нулей и единиц, если представить эти значения как два возможных устойчивых состояния электронного элемента:

0 – отсутствие электрического сигнала;

1 – наличие электрического сигнала.

Эти состояния легко различать.

Недостаток двоичного кодирования – *длинные коды*. Но в технике легче иметь дело с большим количеством простых элементов, чем с небольшим числом сложных.

Способы кодирования и декодирования информации в компьютере, в первую очередь, зависит от вида информации, а именно, что должно кодироваться: числа, текст, графические изображения или звук.

Измерение количества информации

Бит – наименьшая единица информации
измерения объема информации

Название	Усл. обозн.	Соотношение
Байт	Байт	1 байт = 2^3 бит = 8 бит
Килобит	Кбит	1 Кбит = 2^{10} бит = 1024 бит
КилоБайт	Кб	1 Кб = 2^{10} байт = 1024 байт
МегаБайт	Мб	1 Мб = 2^{10} Кб = 1024 Кб
ГигаБайт	Гб	1 Гб = 2^{10} Мб = 1024 Мб
ТераБайт	Тб	1 Тб = 2^{10} Гб = 1024 Гб

Как измерить информацию?

Вопрос:

«Как измерить информацию?»

очень непростой.

Ответ на него зависит от того, что понимать под информацией. Но поскольку определять информацию можно по-разному, то и **способы измерения тоже могут быть разными.**



ИНФОРМАЦИЯ

Измерение информации

Содержательный подход

Сообщение, уменьшающее неопределенность знаний человека в два раза, несет для него **1 бит** информации.

Количество информации, заключенное в сообщении, определяется по формуле Хартли:

$$I = \log_2 N$$

$$N = 2^I$$

где **N** – количество равновероятных событий;

I – количество информации (бит), заключенное в сообщении об одном из событий.

Измерение информации

Алфавитный (технический) подход

Основан на подсчете числа символов в сообщении

Если допустить, что все символы алфавита встречаются в тексте с одинаковой частотой, то количество информации, заключенное **сообщении** вычисляется по формуле:

$$I_c = i * K$$

$$N = 2^i$$

I_c – информационный объем сообщения

K – количество символов

N – мощность алфавита (количество символов)

i - информационный объем 1 символа

Способы кодирования

Способы кодирования и декодирования информации в компьютере, в первую очередь, зависит от вида информации, а именно, что должно кодироваться:

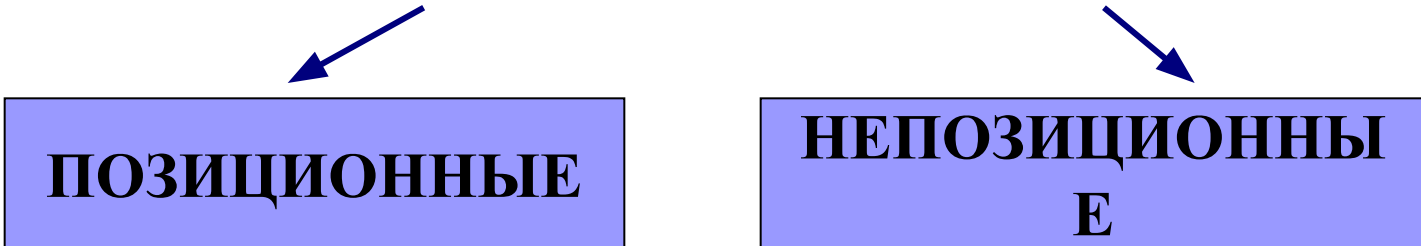
- числа
- символьная (буквы, цифры, знаки)
- графические изображения
- звук



Двоичное кодирование числовой информации

Позиционные и непозиционные системы счисления

Все **системы счисления** делятся на две большие группы:



ПОЗИЦИОННЫЕ

Количественное значение каждой цифры числа зависит от того, в каком месте (позиции или разряде) записана та или иная цифра.

0,7
7
70

НЕПОЗИЦИОННЫЕ

Количественное значение цифры числа не зависит от того, в каком месте (позиции или разряде) записана та или иная цифра.

XIX

Представление чисел

Для записи информации о количестве объектов используются числа

Числа записываются с использованием особых знаковых систем, которые называют системами счисления
 $100 \rightarrow 1100100_2$

Система счисления – совокупность приемов и правил записи чисел с помощью определенного набора СИМВОЛОВ



Двоичное кодирование текстовой информации

Кодирование текстовой информации

Процесс кодирования текстовой информации состоит в том, что **каждому символу присваивается уникальный десятичный (или шестнадцатеричный) код**, который затем представляется в виде двоичного. Данный код называется кодом символа.

Конкретное соответствие между символами и их кодами называется **системой кодировки**.

Каждая кодировка задается своей собственной **кодовой таблицей**. Одному и тому же коду в разных таблицах поставлены в соответствие разные символы.

Таблицы кодировки

ASCII - Стандартной в этой таблице является только первая половина, т.е. символы с номерами от 0 (00000000) до 127 (0111111). Сюда входят буква латинского алфавита, цифры, знаки препинания, скобки и некоторые другие символы.

Остальные 128 кодов используются в разных вариантах. В русских кодировках размещаются символы русского алфавита.

(256 символов – 1 байт (8 бит) каждый символ)

В настоящее время существует

5 разных кодовых таблиц для русских букв
(KOI8, **CP1251 (WIN-1251)**, CP866, Mac, ISO).

и **одна** универсальная кодовая таблица
для **латинского алфавита**

В настоящее время получил широкое распространение новый международный стандарт **Unicode**, который отводит на каждый символ 2 байта (16 бит). С его помощью можно закодировать 65536 ($2^{16} = 65536$) различных символов.

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 двоичный код (по правилам систем счисления)

Возьмем число **57**

При использовании в тексте каждая цифра будет представлена своим кодом

в соответствии с таблицей ASCII

Код передаваемого числа будет – **00110101 00110111**

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



Измерение информации: алфавитный подход

Алфавитный подход к измерению информации

Познакомимся с способом измерения информации, который **не связывает** количество информации **с содержанием сообщения**, и называется он **алфавитным** подходом.

При алфавитном подходе к определению количества информации **отвлекаются от содержания** информации и рассматривают информационное сообщение как **последовательность знаков** определенной знаковой системы.

Применение алфавитного подхода удобно прежде всего при использовании технических средств работы с информацией. В этом случае теряют смысл понятия «новые — старые», «понятные — непонятные» сведения.

Алфавитный подход является объективным способом измерения информации в отличие от субъективного содержательного подхода.

Алфавит и его мощность

Все множество используемых в языке **символов** будем традиционно называть алфавитом.

Обычно под алфавитом понимают только буквы, но поскольку в тексте могут встречаться знаки препинания, цифры, скобки, то мы их тоже включим в алфавит. В алфавит также следует включить и пробел, т.е. пропуск между словами.

Полное количество символов алфавита принято называть мощностью алфавита.

Будем обозначать эту величину буквой **N**. Например, мощность алфавита из заглавных русских букв и отмеченных дополнительных символов равна 54.

АБВГДЕЁЖЗИЙКЛМНОПРСТУФХЦШЩЪЭЮЯ0123456789(),!/?«»:-; (пробел)

Сколько информации несет один символ в русском языке

Представьте себе, что текст к вам поступает последовательно, по одному знаку, словно бумажная ленточка, выползающая из телеграфного аппарата. Предположим, что каждый появляющийся на ленте символ с одинаковой вероятностью может быть любым символом алфавита.

В каждой очередной позиции текста может появиться любой из N символов. Тогда, согласно известной нам формуле $2^I = N$, каждый такой символ несет I бит информации, которое можно определить из решения уравнения: $2^I = 54$.

Получаем: $I = 5.755$ бит или $I = 6$ бит

Вот сколько информации несет один символ в русском тексте!



Информационный объем символа:

$$N = 2^I$$

- В 2-х символьном алфавите каждый символ весит 1 бит.
- В 4-х символьном алфавите каждый символ весит 2 бита.
- В 256-ти символьном алфавите каждый символ весит 8 бит (ASCII)
- В кодировке Unicode 1 символ весит 16 бит (16)

Информационный объем текста

- Если весь текст состоит из K символов, то при алфавитном подходе размер содержащейся в ней информации равен:

$$I = i * K$$

где i – информационный вес
одного символа

K – количество символов

Количество информации в тексте

Чтобы найти информационный объем всего текста, нужно посчитать число символов в нем (k) и умножить на информационный объем 1 символа (i)

Определить количество информации на одной странице книги (50 строк по 60 символов на русском языке – 54 символа)



$$I = k * i$$

$$k = ?$$

$$50 \times 60 = 3000 \text{ знаков}$$

$$i = ?$$

$$N = 54 \quad 2^6 = 64 \quad i = 6$$

Объем информации

$$6 \times 3000 = 18000 \text{ бит} = 2250 \text{ байт} = 2,1 \text{ Кб}$$

При алфавитном подходе к измерению информации количество информации зависит не от содержания, а от размера текста и мощности алфавита

Вопросы и задания:

- Что такое «алфавит»?
- Что такое «мощность алфавита»?
- Как определяется количество информации в сообщении с алфавитной точки зрения?
- Что больше 1 Кбайт или 1000 байт?
- Расположите единицы измерения информации в порядке возрастания:

Гигабайт; Байт; Мегабайт; Килобайт.

- Сколько информации содержится в сообщении, если для кодирования одного символа использовать 1 байт:
Компьютер – универсальный прибор
- Два текста содержат одинаковое количество символов. Первый текст составлен в алфавите мощностью 32 символа, второй – мощностью 64 символа. Во сколько раз отличается количество информации в этих текстах?

Вопросы и задания:

- В чем заключается кодирование текстовой информации в компьютере?
- Как изменится объем текста при его преобразовании из стандартной кодировки в кодировку Unicode?
- Сколько существует кодировок латинского алфавита?
- Сколько существует кодировок русского алфавита?
- Считая, что каждый символ кодируется одним байтом, оцените информационный объем следующего предложения из пушкинского четверостишия:

**Певец-Давид был ростом мал,
Но повалил же Голиафа!**

Задание 1:

- Определите информационный объем страницы книги, если для записи текста использовались только заглавные буквы русского алфавита, кроме буквы Ё.

Решение:

1. $N = 32$
2. $2^i = N$
3. $2^i = 32$
4. $i = 5$
5. На странице 3000 знаков, тогда объем информации = $3000 * 5 = 15000$ бит.



Задание 2:

Племя Мумбу-Юмбу использует алфавит из букв: $\alpha\beta\gamma\delta\epsilon\zeta\eta\theta\lambda\mu\xi\sigma\varphi\psi$, точки и для разделения слов используется пробел.

Сколько информации несет свод законов племени, если в нем 12 строк и в каждой строке по 20 символов?

Задание 3:

Вычислите какова мощность алфавита, с помощью которого записано сообщение, содержащее 2048 символов, если его объем составляет 1.25 Кбайта



Аналоговая и дискретная форма представления информации

Аналоговая и дискретная форма представления информации

Человек способен воспринимать и хранить информацию в форме образов (зрительных, звуковых, осязательных, вкусовых и обонятельных). Зрительные образы могут быть сохранены в виде изображений (рисунков, фотографий и так далее), а звуковые — зафиксированы на пластинках, магнитных лентах, лазерных дисках и так далее.

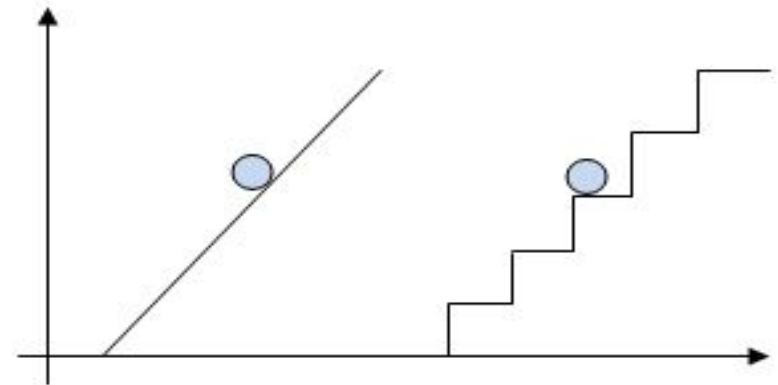
Информация, в том числе графическая и звуковая, может быть представлена в *аналоговой* или *дискретной* форме. При аналоговом представлении физическая величина принимает бесконечное множество значений, причем ее значения изменяются непрерывно. При дискретном представлении физическая величина принимает конечное множество значений, причем ее величина изменяется скачкообразно.

Аналоговая и дискретная форма представления информации

Приведем пример аналогового и дискретного представления информации.

Положение тела на наклонной плоскости и на лестнице задается значениями координат X и Y .

При движении тела по наклонной плоскости его координаты могут принимать **бесконечное множество непрерывно изменяющихся значений** из определенного диапазона, а при движении по лестнице — только **определенный набор значений**, причем меняющихся скачкообразно.



Дискретизация

Примером аналогового представления графической информации может служить, например, живописное полотно, цвет которого изменяется непрерывно, а дискретного — изображение, напечатанное с помощью струйного принтера и состоящее из отдельных точек разного цвета. Примером аналогового хранения звуковой информации является виниловая пластинка (звуковая дорожка изменяет свою форму непрерывно), а дискретного — аудиокомпакт-диск (звуковая дорожка которого содержит участки с различной отражающей способностью).

Преобразование графической и звуковой информации из аналоговой формы в дискретную производится путем *дискретизации, то есть разбиения непрерывного графического изображения и непрерывного (аналогового) звукового сигнала на отдельные элементы*. В процессе дискретизации производится кодирование, то есть присвоение каждому элементу конкретного значения в форме кода.

Дискретизация
это преобразование непрерывных изображений и звука в набор дискретных значений в форме кодов.



Двоичное кодирование графической информации

Кодирование векторной графики

Одним из способов представления изображений на компьютере является **векторная графика**.

Основным ее элементом является **линия**, которая задается в виде формулы, то есть в виде нескольких параметров.

Для описания прямой линии достаточно всего два параметра:
 $y = kx + b$.

Кривые второго порядка (параболы, гиперболы) задаются пятью параметрами.

Кривые третьего порядка имеют точку перегиба.
Для их описания необходимо 9 параметров.

Кодирование растровой графики

Графическая информация на экране монитора представляется в виде **растрового** изображения, которое формируется из точек (**пикселей**), причем каждой точке присваивается значение его цвета (код цвета). Чем меньший размер имеет точка, тем, соответственно, больше точек на экране и тем выше качество изображения. Количество точек на экране, называется **разрешающей способностью** монитора.

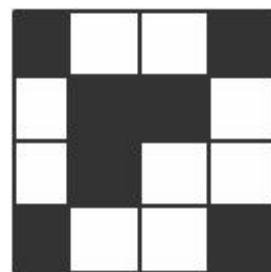
Растр - разложение изображения на отдельные точки с помощью специальной сетки.

Кодирование растровой графики

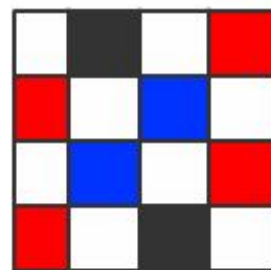
Растровое изображение представляет собой совокупность точек (пикселей) разных цветов. Пиксель - минимальный участок изображения, цвет которого можно задать независимым образом.

В процессе кодирования изображения производится его **пространственная дискретизация**.

Пространственную дискретизацию изображения можно сравнить с построением изображения из мозаики (большого количества маленьких разноцветных стекол). *Изображение разбивается на отдельные маленькие фрагменты (точки), причем каждому фрагменту присваивается значение его цвета, то есть код цвета* (красный, зеленый, синий и так далее).



1 0 0 1
0 1 1 0
0 1 0 0
1 0 0 1

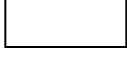


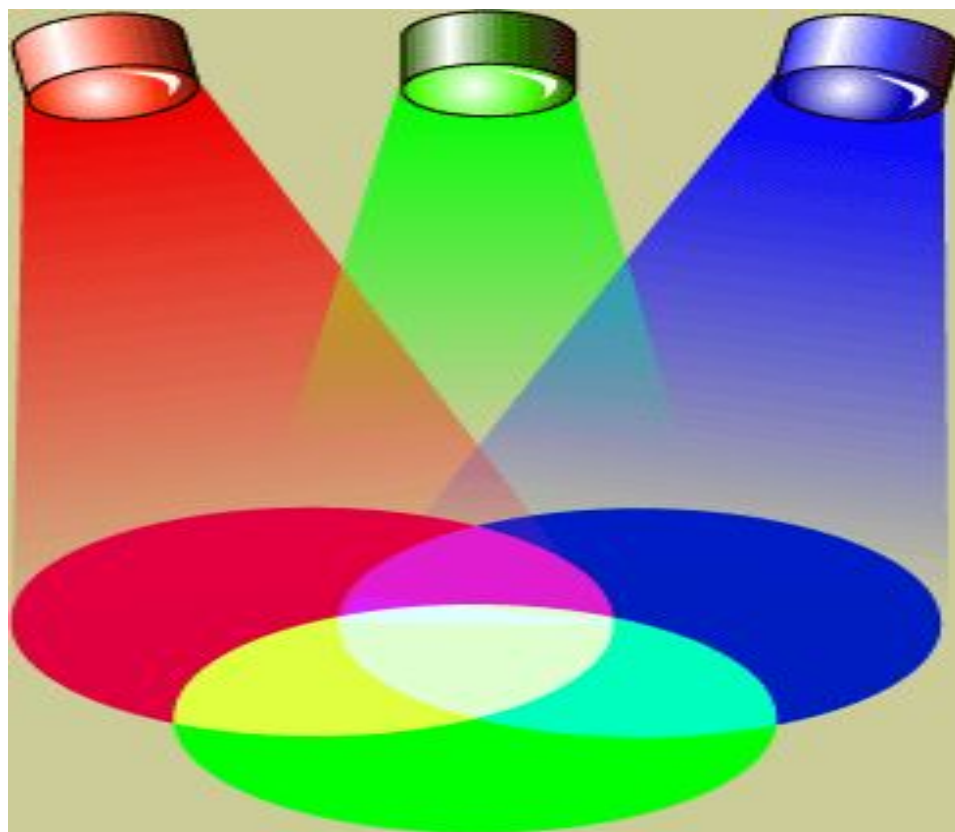
00 11 00 01
01 00 10 00
00 10 00 01
01 00 11 00



Кодирование растровой графики

При формировании изображения на экране монитора
используется **RGB- модель**

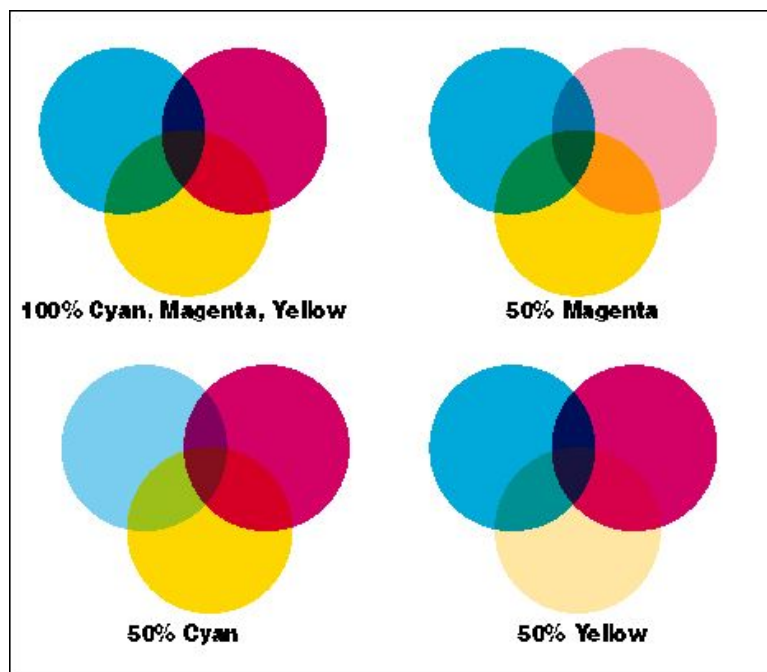
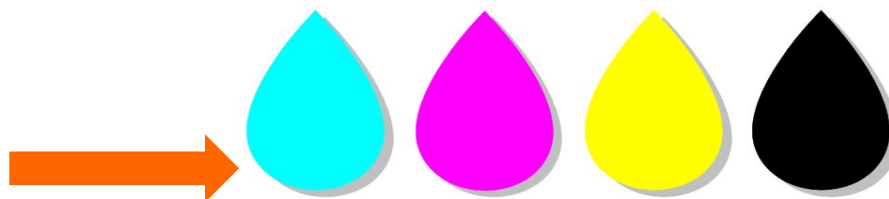
черный		000000
красный		FF0000
зеленый		00FF00
синий		0000FF
белый		FFFFFF



Кодирование растровой графики

При подготовке печатных
изображений используется

СМУК - модель.



Кодирование графической информации

Количество цветов, которые используются для воспроизведения изображения, называется **цветовой палитрой**.

Качество двоичного кодирования изображения определяется разрешающей способностью экрана и **глубиной цвета**.

$$N = 2^i$$

N – мощность палитры

i – глубина цвета

Измерение объема графической информации

- Если изображение состоит из K точек, объем графической информации равен:

$$I = i * k$$

где i – глубина цвета

k – размер изображения (пиксел)

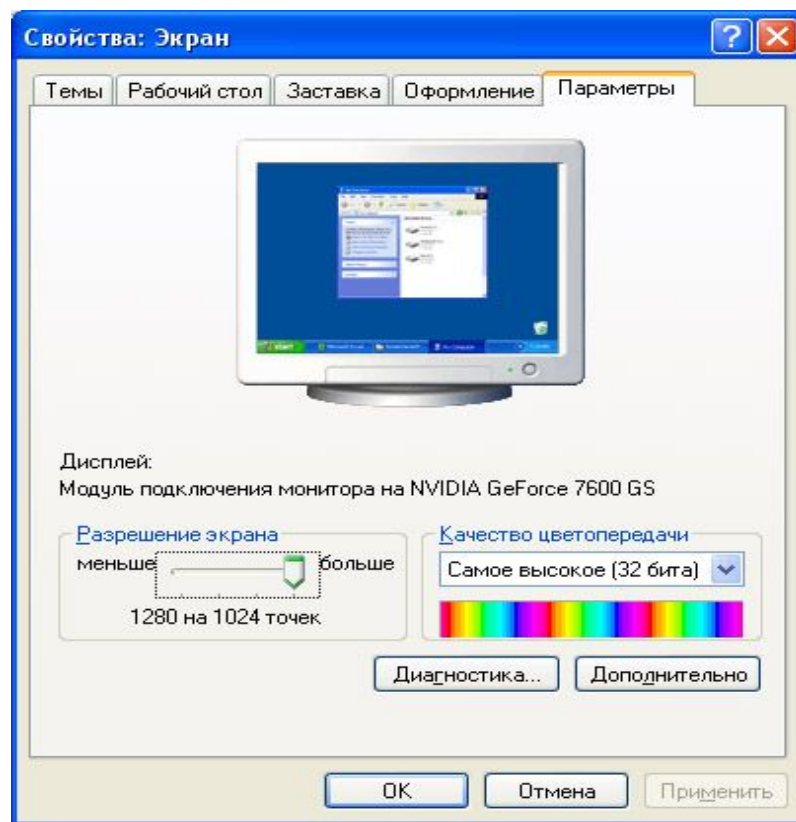
Вычислим объем видеопамати

Для того чтобы на экране монитора формировалось изображение, информация о каждой точке (код цвета точки) должна храниться в видеопамати компьютера.

Рассчитаем необходимый объем видеопамати для одного из графических режимов.

В современных компьютерах разрешение экрана обычно составляет 1280x1024 точек. Т.е. всего
 $1280 * 1024 = 1310720$ точек.

При глубине цвета 32 бита на точку необходимый объем видеопамати:
 $32 * 1310720 = 41943040$ бит =
 5242880 байт = 5120 Кб = 5 Мб.



Вопросы и задания:

- Какие виды компьютерных изображений вы знаете?
- В растровом графическом редакторе минимальным объектом, цвет которого можно изменить, является ...
- В векторном графическом редакторе минимальным объектом, размер которого можно изменить, является ...
- Двоичный код изображения, выводимого на экран дисплея ПК, хранится в...
- Какое максимальное количество цветов может быть использовано в изображении, если на каждую точку отводится 3 бита?
- Что вы знаете о цветовой модели RGB?
- Цветное (с палитрой 256 цветов) растровое графическое изображение имеет размер 10×10 точек. Какой объем памяти займет это изображение?
- Рассчитайте необходимый объем видеопамати для графического режима: разрешение экрана 800×600 , качество цветопередачи 16 бит.



Двоичное кодирование звука. Представление видеоинформации

Кодирование звуковой информации

Звуковая волна - это непрерывная волна с меняющейся амплитудой и частотой. Чем больше амплитуда сигнала, тем он громче для человека, чем больше частота сигнала, тем выше тон.

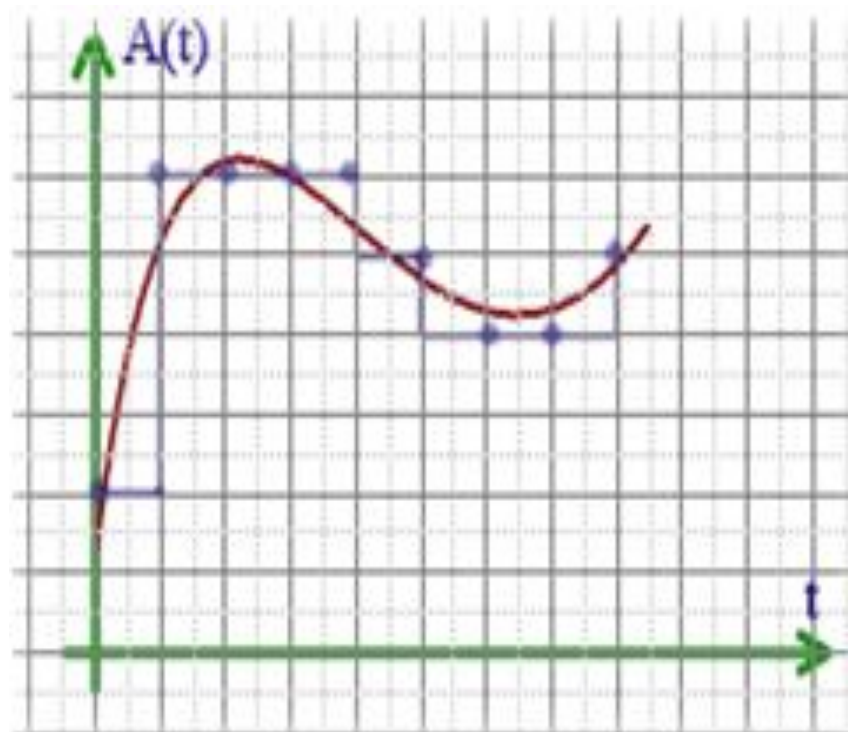


Человеческий слух воспринимает звуковые колебания с частотой от 20 до 20 000 раз в секунду – от 20 Гц до 20 кГц.

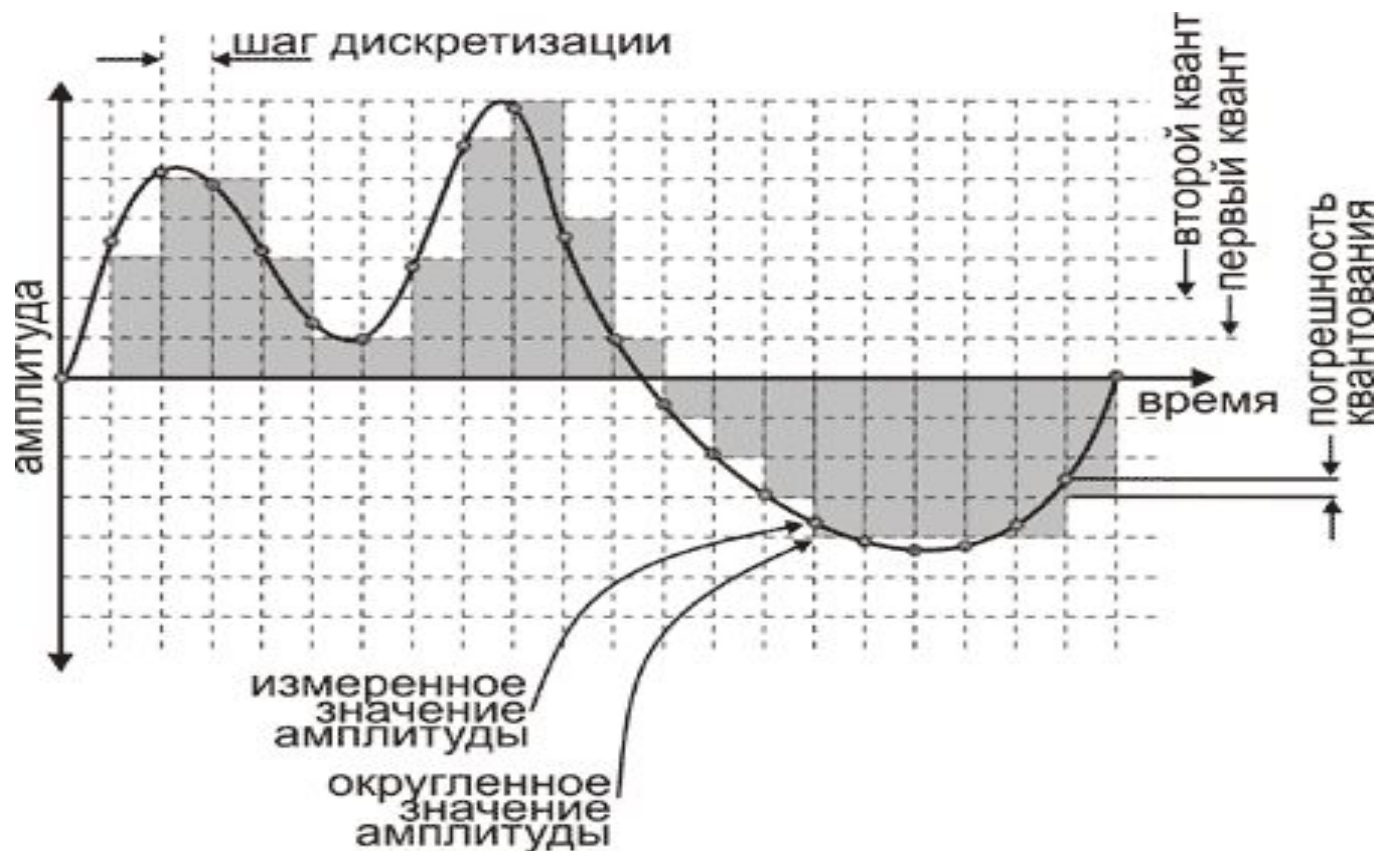
Временная дискретизация звука

В процессе кодирования звукового сигнала производится его временная дискретизация – непрерывная волна разбивается на отдельные **маленькие временные участки** и для каждого такого участка устанавливается **определенная величина амплитуды**.

Таким образом непрерывная зависимость амплитуды сигнала от времени заменяется на дискретную последовательность уровней громкости.



Кодирование звуковой информации



Дискретизация звука

Глубина кодирования звука

Качество двоичного кодирования звука определяется глубиной кодирования и частотой дискретизации.

Частота дискретизации – количество измерений уровня сигнала в единицу времени.

Количество уровней громкости определяет глубину кодирования. Современные звуковые карты обеспечивают 16-битную глубину кодирования звука. При этом количество уровней громкости равно $N = 2^l = 2^{16} = 65536$.

Кодирование звуковой информации

Качество двоичного кодирования звука определяется
глубиной кодирования и частотой дискретизации.

$$N = 2^i$$

N – количество различных уровней сигнала

i – глубина кодирования звука

Информационный объем звуковой информации

- объем звуковой информации равен:

$$I = i * K * t$$

где i – глубина звука (бит)

K – частота вещания

(качество звука) (Гц) (48 кГц – аудио CD)

t – время звучания (сек)

Представление видеоинформации

В последнее время компьютер все чаще используется для работы с видеоинформацией. Простейшей такой работой является просмотр кинофильмов и видеоклипов. Следует четко представлять, что обработка видеоинформации требует очень высокого быстродействия компьютерной системы.

Что представляет собой *фильм* с точки зрения информатики? Прежде всего, это *сочетание звуковой и графической информации*. Кроме того, для создания на экране эффекта движения используется дискретная по своей сути технология быстрой смены статических картинок. Исследования показали, что если за одну секунду сменяется более 10-12 кадров, то человеческий глаз воспринимает изменения на них как непрерывные.