

Кодирование информации

- Общие сведения о кодировании информации
- Кодирование числовой информации
- Кодирование текстовой информации
- Кодирование графической информации
- Кодирование звуковой информации
- Кодирование видео информации

Кодирование и декодирование

Для обмена информацией с другими людьми человек использует естественные языки. Наряду с *естественными* языками были разработаны *формальные* языки для профессионального применения их в какой-либо сфере. Представление информации с помощью какого-либо языка часто называют кодированием.

Код — набор символов (условных обозначений) для представления информации.

Код — система условных знаков (символов) для передачи, обработки и хранения информации(сообщения).

Кодирование — процесс представления информации (сообщения) в виде кода.

Все множество символов, используемых для кодирования, называется *алфавитом кодирования*. Например, в памяти компьютера любая информация кодируется с помощью двоичного алфавита, содержащего всего два символа: **0** и **1**.

Способы кодирования информации

Для кодирования одной и той же информации могут быть использованы разные способы; их выбор зависит от ряда обстоятельств: цели кодирования, условий, имеющихся средств.

Если надо записать текст в темпе речи — используем стенографию; если надо передать текст за границу — используем английский алфавит; если надо представить текст в виде, понятном для грамотного русского человека, — записываем его по правилам грамматики русского языка.

«Добрый день, Дима!»

«Dobryi den, Dima»

Способы кодирования информации

Выбор способа кодирования информации может быть связан с предполагаемым способом ее обработки.

Покажем это на примере представления чисел — количественной информации. Используя русский алфавит, можно записать число «**сорок семь**». Используя же алфавит арабской десятичной системы счисления, пишем «**47**». Второй способ не только короче первого, но и удобнее для выполнения вычислений. Какая запись удобнее для выполнения расчетов: "**сорок семь умножить на сто двадцать пять**" или «**47x 125**»? Очевидно — вторая.

Шифрование сообщения

В некоторых случаях возникает потребность засекречивания текста сообщения или документа, для того чтобы его не смогли прочитать те, кому не положено. Это называется **защитой от несанкционированного доступа**.

В таком случае секретный текст *шифруется*.

В давние времена шифрование называлось *тайнописью*.

Шифрование представляет собой процесс превращения открытого текста в зашифрованный, а дешифрование — процесс обратного преобразования, при котором восстанавливается исходный текст.

Шифрование — это тоже кодирование, но с засекреченным методом, известным только источнику и адресату.

Методами шифрования занимается наука под названием **криптография**.

Оптический телеграф

Шаппа

В 1792 году во Франции Клод Шапп создал систему передачи визуальной информации, которая получила название «*Оптический телеграф*».

В простейшем виде это была цепь типовых строений, с расположенными на кровле шестами с подвижными поперечинами, которая создавалась в пределах видимости одно от другого. Шесты с подвижными поперечинами — семафоры — управлялись при помощи тросов специальными операторами изнутри строений.

Шапп создал специальную таблицу кодов, где каждой букве алфавита соответствовала определенная фигура, образуемая Семафором, в зависимости от положений поперечных брусьев относительно опорного шеста.

Система Шаппа позволяла передавать сообщения на скорости два слова в минуту и быстро распространилась в Европе. В Швеции цепь станций оптического телеграфа действовала до 1880 года.



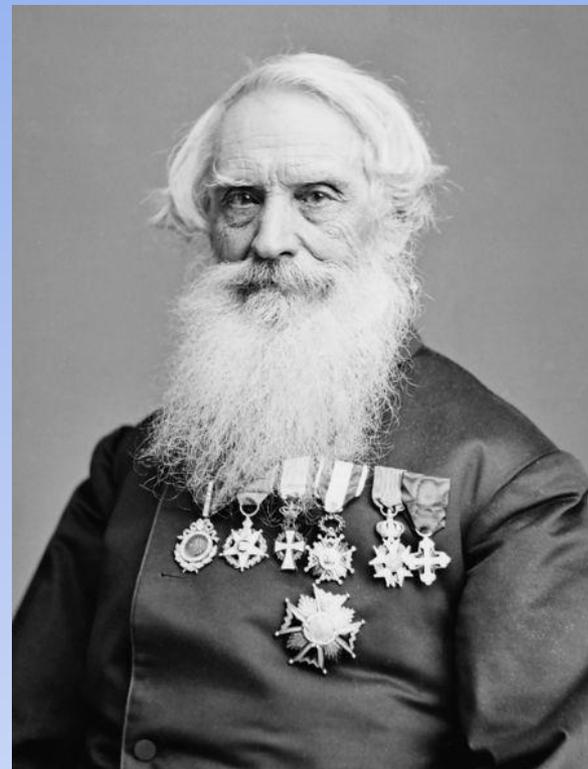
Первый телеграф

Первым техническим средством передачи информации на расстояние стал **телеграф**, изобретенный в 1837 году американцем Сэмюэлем Морзе.

Телеграфное сообщение — это последовательность электрических сигналов, передаваемая от одного телеграфного аппарата по проводам к другому телеграфному аппарату.

Изобретатель Сэмюэль Морзе изобрел удивительный код (Азбука Морзе, код Морзе, «Морзянка»), который служит человечеству до сих пор. **Информация кодируется тремя «буквами»: длинный сигнал (тире), короткий сигнал (точка) и отсутствие сигнала (пауза)** для разделения букв. Таким образом, кодирование сводится к использованию набора символов, расположенных в строго определенном порядке.

Самым знаменитым телеграфным сообщением является сигнал бедствия "**SOS**" (Save Our Souls - спасите наши души). Вот как он выглядит: «••• - - - •••»



Азбука Морзе

А	• —	И	• •	Р	• — •	Ш	— — — —
Б	— • • •	Й	• — — —	С	• • •	Щ	— — • —
В	• — —	К	— • —	Т	—	Ъ	• — — • — •
Г	— — •	Л	• — • •	У	• • —	Ь	— • • —
Д	— • •	М	— —	Ф	• • — •	Ы	— • — —
Е	•	Н	— •	Х	• • • •	Э	• • — • •
Ж	• • • —	О	— — — —	Ц	— • — •	Ю	• • — —
З	— — • •	П	• — — •	Ч	— — — •	Я	• — • —

Азбука Морзе

1	• — — — —	9	— — — — •
2	• • — — —	0	— — — — —
3	• • • — —	Точка	• • • • • •
4	• • • • —	Запятая	• — • — • —
5	• • • • •	/	— • • — •
6	• • • •	?	• • — — • •
7	— — • • •	!	— — • • — —
8	— — — • •	@	• — — • — •

Первый беспроводной телеграф (радиоприемник)

7 мая 1895 года российский ученый Александр Степанович Попов на заседании Русского Физико-Химического Общества продемонстрировал прибор, названный им "грозоотметчик", который был предназначен для регистрации электромагнитных волн.

Этот прибор считается **первым в мире аппаратом беспроводной телеграфии, радиоприемником**. В 1897 году при помощи аппаратов беспроводной телеграфии Попов осуществил прием и передачу сообщений между берегом и военным судном.

В 1899 году Попов сконструировал модернизированный вариант приемника электромагнитных волн, где прием сигналов (**азбукой Морзе**) осуществлялся на головные телефоны оператора.

В 1900 году благодаря радиостанциям, построенным на острове Гогланд и на российской военно-морской базе в Котке под руководством Попова, были успешно осуществлены аварийно-спасательные работы на борту военного корабля "Генерал-адмирал Апраксин", севшего на мель у острова Гогланд. В результате обмена сообщениями, переданным методом беспроводной телеграфии, экипажу российского ледокола Ермак была своевременно и точно передана информация о финских рыбаках, находящихся на оторванной льдине.



Телеграфный аппарат Бодо

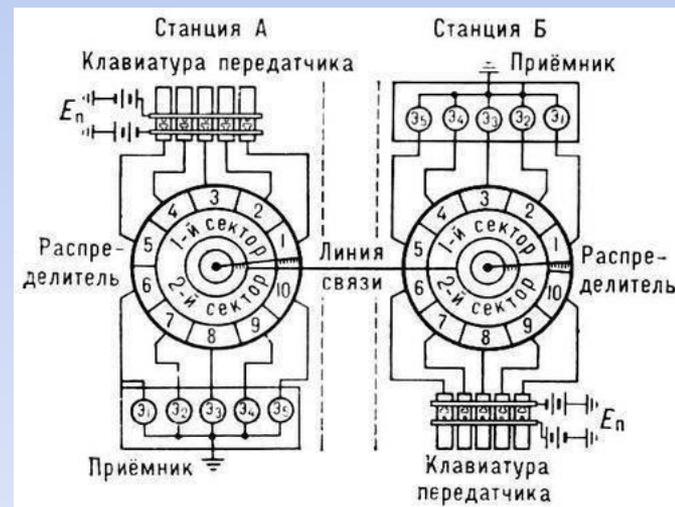
Равномерный телеграфный код был изобретен французом Жаном Морисом Бодо в конце XIX века. В нем использовалось всего два разных вида сигналов. Не важно, как их назвать: точка и тире, плюс и минус, ноль и единица. Это два отличающихся друг от друга электрических сигнала. **Длина кода всех символов одинаковая и равна пяти.** В таком случае не возникает проблемы отделения букв друг от друга: каждая пятерка сигналов — это знак текста. Поэтому пропуск не нужен.

Код называется равномерным, если длина кода всех символов равна.

Код Бодо — это первый в истории техники способ *двоичного кодирования* информации. Благодаря этой идее удалось создать буквопечатающий телеграфный аппарат, имеющий вид пишущей машинки. Нажатие на клавишу с определенной буквой вырабатывает соответствующий пятиимпульсный сигнал, который передается по линии связи.

В честь Бодо была названа единица скорости передачи информации — бод.

В современных компьютерах для кодирования текста также применяется равномерный двоичный код.



Двоичное кодирование в компьютере

Вся информация, которую обрабатывает компьютер должна быть представлена двоичным кодом с помощью двух цифр: **0** и **1**. *Эти два символа принято называть двоичными цифрами или битами.*

С помощью двух цифр 0 и 1 можно закодировать любое сообщение. Это явилось причиной того, что в компьютере обязательно должно быть организовано два важных процесса: кодирование и декодирование.

Кодирование – преобразование входной информации в форму, воспринимаемую компьютером, т.е. двоичный код.

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.

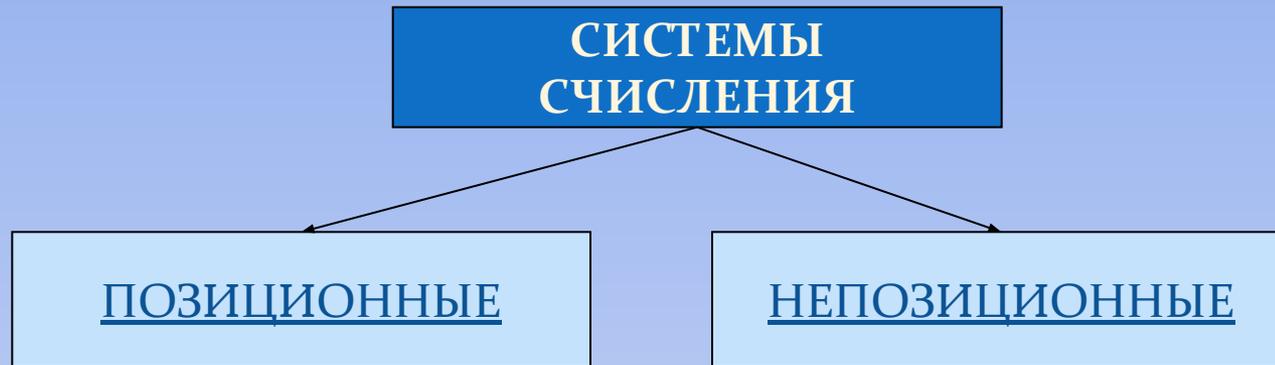
Способы кодирования и декодирования информации в компьютере, в первую очередь, зависит от вида информации, а именно, что должно кодироваться: числа, текст, графические изображения или звук.

Система счисления

Для записи информации о количестве объектов используются числа. Числа записываются с помощью набора специальных символов.

Система счисления — способ записи чисел с помощью набора специальных знаков, называемых цифрами.

Виды систем счисления



В *позиционных* системах счисления величина, обозначаемая цифрой в записи числа, *зависит* от её *положения* в числе (*позиции*).

В *непозиционных* системах счисления величина, которую обозначает цифра, *не зависит* от *положения* в числе.

Непозиционные системы счисления

Каноническим примером фактически *непозиционной системы* счисления является *римская*, в которой в качестве цифр используются латинские буквы:

I обозначает 1, V - 5, X - 10, L - 50, C - 100, D - 500, M -1000.

Натуральные числа записываются при помощи повторения этих цифр.

Например, II = 1 + 1 = 2, здесь символ I обозначает 1 независимо от места в числе.

Для правильной записи больших чисел римскими цифрами необходимо сначала записать число тысяч, затем сотен, затем десятков и, наконец, единиц.

Пример: число 2988. Две тысячи MM, девять сотен CM, восемьдесят LXXX, восемь VIII. Запишем их вместе: MCMLXXXVIII.

$MMCMLXXXVIII = 1000+1000+(1000-100)+(50+10+10+10)+5+1+1+1 = 2988$

Для изображения чисел в непозиционной системе счисления нельзя ограничиться конечным набором цифр. Кроме того, выполнение арифметических действий в них крайне неудобно.

Древнеегипетская десятичная непозиционная система счисления.

- Примерно в третьем тысячелетии до нашей эры древние египтяне придумали свою числовую систему, в которой для обозначения ключевых чисел 1, 10, 100 и т.д. использовались специальные значки — иероглифы.
- Все остальные числа составлялись из этих ключевых при помощи операции сложения. Система счисления Древнего Египта является десятичной, но непозиционной.

Алфавитные системы счисления.

- Более совершенными непозиционными системами счисления были алфавитные системы. К числу таких систем счисления относились греческая, славянская, финикийская и другие. В них числа от 1 до 9, целые количества десятков (от 10 до 90) и целые количества сотен (от 100 до 900) обозначались буквами алфавита.
- В алфавитной системе счисления Древней Греции числа 1, 2, ..., 9 обозначались первыми девятью буквами греческого алфавита, например $\alpha = 1$, $\beta = 2$, $\gamma = 3$ и т.д. Для обозначения чисел 10, 20, ..., 90 применялись следующие 9 букв ($\iota = 10$, $\kappa = 20$, $\lambda = 30$, $\mu = 40$ и т.д.), а для обозначения чисел 100, 200, ..., 900 — последние 9 букв ($\rho = 100$, $\sigma = 200$, $\tau = 300$ и т.д.). Например, число 141 обозначалось $\rho\mu\alpha$.
- У славянских народов числовые значения букв установились в порядке славянского алфавита, который использовал сначала глаголицу, а затем кириллицу. Подробнее с происхождением и развитием русской письменности можно ознакомиться на сайте

Древнерусская алфавитная система счисления

БУКВЫ КИРИЛЛИЦЫ	ЦИФРОВОЕ ЗНАЧЕНИЕ КИРИЛЛИЦЫ	БУКВЫ ГЛАГОЛИЦЫ	ЦИФРОВОЕ ЗНАЧЕНИЕ ГЛАГОЛИЦЫ	КИРИЛЛИЧЕСКОЕ НАЗВАНИЕ
А	1	Ѡ	1	Азь
Б		Ѳ	2	Буки
В	2	Ѵ	3	Вѣди
Г	3	Ѷ	4	Глаголь
Д	4	Ѹ	5	Добро
Е	5	Ѻ	6	Есть
Ж		Ѽ	7	Живѣте
З	6	Ѿ	8	Зѣло
З	7	Ѡ	9	Земля
Н	8	Ѳ	10	Иже
І	10	Ѵ	20	І

Позиционные системы счисления

В позиционных системах счисления величина, обозначаемая цифрой в записи числа, зависит от её положения в числе (позиции).

Количество используемых цифр называется основанием системы счисления.

Например, 11 – это одиннадцать, а не два: $1 + 1 = 2$ (сравните с римской системой счисления). Здесь символ 1 имеет различное значение в зависимости от позиции в числе.

Первые позиционные системы счисления

Самой первой такой системой, когда счетным "прибором" служили пальцы рук, была *пятеричная*.



Некоторые племена на филиппинских островах используют ее и в наши дни, а в цивилизованных странах ее реликт, как считают специалисты, сохранился только в виде школьной пятибалльной шкалы оценок.

Двенадцатеричная система счисления

Следующей после пятеричной возникла *двенадцатеричная* система счисления. Возникла она в древнем Шумере. Некоторые учёные полагают, что такая система возникала у них из подсчёта фаланг на руке большим пальцем.

Широкое распространение получила двенадцатеричная система счисления в XIX веке. На ее широкое использование в прошлом явно указывают названия числительных во многих языках, а также сохранившиеся в ряде стран способы отсчета времени, денег и соотношения между некоторыми единицами измерения. Год состоит из 12 месяцев, а половина суток состоит из 12 часов.

Элементом двенадцатеричной системы в современности может служить счёт дюжинами. Первые три степени числа 12 имеют собственные названия: 1 дюжина = 12 штук; 1 гросс = 12 дюжин = 144 штуки; 1 масса = 12 гроссов = 144 дюжины = 1728 штук.

Английский фунт состоит из 12 шиллингов.

Шестидесятеричная система счисления

Следующая позиционная система счисления была придумана еще в Древнем Вавилоне, причем вавилонская нумерация была *шестидесятеричная*, т.е. в ней использовалось шестьдесят цифр!

В более позднее время использовалась арабами, а также древними и средневековыми астрономами. Шестидесятеричная система счисления, как считают исследователи, является синтезом уже вышеупомянутых пятеричной и двенадцатеричной систем.

Какие позиционные системы счисления используются сейчас?

В настоящее время наиболее распространены десятичная, двоичная, восьмеричная и шестнадцатеричная системы счисления.

Двоичная, восьмеричная (в настоящее время вытесняется шестнадцатеричной) и шестнадцатеричная система часто используется в областях, связанных с цифровыми устройствами, программировании и вообще компьютерной документации.

Современные компьютерные системы оперируют информацией представленной в цифровой форме.

Десятичная система счисления

Десятичная система счисления

— позиционная система
счисления по основанию 10.

Предполагается, что основание
10 связано с количеством
пальцев рук у человека.

Наиболее распространённая
система счисления в мире.

Для записи чисел используются
символы 0, 1, 2, 3, 4, 5, 6, 7, 8, 9,
называемые арабскими
цифрами.

Современные цифры	Арабские цифры	Индийские цифры
0	٠	०
1	١	१
2	٢	२
3	٣	३
4	٤	४
5	٥	५
6	٦	६
7	٧	७
8	٨	८
9	٩	९

Двоичная система счисления

Двоичная система счисления — позиционная система счисления с основанием 2.

Используются цифры 0 и 1.

Двоичная система используется в цифровых устройствах, поскольку является наиболее простой и удовлетворяет требованиям:

- Чем меньше значений существует в системе, тем проще изготовить отдельные элементы.
- Чем меньше количество состояний у элемента, тем выше помехоустойчивость и тем быстрее он может работать.
- Простота создания таблиц сложения и умножения — основных действий над числами

Алфавит десятичной, двоичной, восьмеричной и шестнадцатеричной систем 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

Соответствие десятичной, двоичной, восьмеричной и шестнадцатеричной систем счисления

p=10	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
p=2	0	1	10	11	100	101	110	111	1000	1001	1010	1011	1100	1101	1110	1111	10000
p=8	0	1	2	3	4	5	6	7	10	11	12	13	14	15	16	17	20
p=16	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F	10

Количество используемых цифр называется основанием системы счисления.

При одновременной работе с несколькими системами счисления для их различения основание системы обычно указывается в виде нижнего индекса, который записывается в десятичной системе:

123_{10} — это число 123 в десятичной системе счисления;

11011_2 — то же число, но в двоичной системе.

Двоичное число 11011 можно расписать в виде: $11011_2 = 1 \cdot 2^4 + 1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0$.

Перевод чисел из одной системы счисления в другую

Перевод из десятичной системы счисления в систему счисления с основанием p осуществляется последовательным делением десятичного числа и его десятичных частных на p , а затем выписыванием последнего частного и остатков в обратном порядке.

Переведем десятичное число 99_{10} в двоичную систему счисления (основание системы счисления $p=2$). В итоге получили $1100011_2 = 99_{10}$

$$\begin{array}{r} 99 \mid 2 \\ \hline 49 \mid 2 \\ \hline 1 \quad 1 \quad 24 \mid 2 \\ \hline 1 \quad 24 \mid 12 \mid 2 \\ \hline 0 \quad 12 \mid 6 \mid 2 \\ \hline 0 \quad 6 \mid 3 \mid 2 \\ \hline 0 \quad 3 \mid 1 \\ \hline 0 \quad 1 \end{array}$$

Числа в компьютере

Числа в компьютере хранятся и обрабатываются в двоичной системе счисления.

Последовательность нулей и единиц называют двоичным кодом.

Специфической особенности представления чисел в памяти компьютера рассмотрим на других уроках по теме «**системы счисления**».

Кодирование текстовой информации

- В традиционных кодировках для кодирования одного символа используется 8 бит. Легко подсчитать по формуле 2.3, что такой 8-разрядный код позволяет закодировать 256 различных символов.
- Присвоение символу определенного числового кода – это вопрос соглашения. В качестве международного стандарта принята кодовая таблица ASCII (American Standard Code for Information Interchange), кодирующая первую половину символов с числовыми кодами от 0 до 127 (коды от 0 до 32 отведены не символам, а функциональным клавишам).
- Национальные стандарты кодировочных таблиц включают международную часть кодовой таблицы без изменений, а во второй половине содержат коды национальных алфавитов, символы псевдографики и некоторые математические знаки. К сожалению, в настоящее время существуют пять различных кодировок кириллицы (КОИ8-Р, Windows.MS-DOS, Macintosh и ISO), что вызывает дополнительные трудности при работе с русскоязычными документами.
- Хронологически одним из первых стандартов кодирования русских букв на компьютерах был КОИ8 ("Код обмена информацией, 8-битный"). Эта кодировка применялась еще в 70-ые годы на компьютерах серии ЕС ЭВМ, а с середины 80-х стала использоваться в первых русифицированных версиях операционной системы UNIX.
- Наиболее распространенной в настоящее время является кодировка Microsoft Windows, обозначаемая сокращением CP1251 ("CP" означает "Code Page", "кодовая страница").

Кодирование текстовой информации

- От начала 90-ых годов, времени господства операционной системы MS DOS, остается кодировка CP866. Компьютеры фирмы Apple, работающие под управлением операционной системы Mac OS, используют свою собственную кодировку Mac. Кроме того, Международная организация по стандартизации (International Standards Organization, ISO) утвердила в качестве стандарта для русского языка еще одну кодировку под названием ISO 8859-5.
- В конце 90-ых годов появился новый международный стандарт Unicode, который отводит под один символ не один байт, а два, и поэтому с его помощью можно закодировать не 256, а 65536 различных символов. Полная спецификация стандарта Unicode включает в себя все существующие, вымершие и искусственно созданные алфавиты мира, а также множество математических, музыкальных, химических и прочих символов.
- **Пример 2.46.** Представьте в форме шестнадцатеричного кода слово «ЭВМ» во всех пяти кодировках. Воспользуйтесь CD-ROM для получения кодировочных таблиц CP866, Mac и ISO и компьютерным калькулятором для перевода чисел из десятичной в шестнадцатеричную систему счисления.
- Последовательности десятичных кодов слова «ЭВМ» в различных кодировках составляем на основе кодировочных таблиц:
 - КОИ8-Р: 252 247 237
 - CP1251: 221 194 204
 - CP866: 157 130 140
 - Mac: 157 130 140
 - ISO: 205 178 188
- Переводим с помощью калькулятора последовательности кодов из десятичной системы в шестнадцатеричную:
 - КОИ8-Р: FC F7 ED
 - CP1251: DD C2 CC
 - CP866: 9D 82 8C
 - Mac: 9D 82 8C
 - ISO: CD B2 BC
- Для преобразования русскоязычных текстовых документов из одной кодировки в другую используются специальные программы-конвертеры. Одной из таких программ является текстовый редактор Hieroglyph, который позволяет осуществлять перевод набранного текста из одной кодировки в другую и даже использовать различные кодировки в одном тексте.

Международная кодировка ASCII

sp 32	! 33	" 34	# 35	\$ 36	% 37	& 38	' 39	(40) 41	* 42	+ 43	, 44	- 45	. 46	/ 47
0 48	1 49	2 50	3 51	4 52	5 53	6 54	7 55	8 56	9 57	: 58	; 59	< 60	= 61	> 62	? 63
@ 64	A 65	B 66	C 67	D 68	E 69	F 70	G 71	H 72	I 73	J 74	K 75	L 76	M 77	N 78	O 79
P 80	Q 81	R 82	S 83	T 84	U 85	V 86	W 87	X 88	Y 89	Z 90	[91	\ 92] 93	^ 94	_ 95
` 96	a 97	b 98	c 99	d 100	e 101	f 102	g 103	h 104	i 105	j 106	k 107	l 108	m 109	n 110	o 111
p 112	q 113	r 114	s 115	t 116	u 117	v 118	w 119	x 120	y 121	z 122	{ 123	 124	} 125	~ 126	

Кодировка КОИ8-Р

—		Г	Г	└	└	┌	┌	┐	┐	┘	■	■	■	■	■
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
▬	▬	▬	┌	■	●	√	≈	≤	≥	nbsp	┘	◦	2	•	÷
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
=		F	ё	П	Р	Т	П	П	Е	Ц	Ц	Г	Ц	Ц	Т
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
Т	Т	Т	Ё	П	Р	Т	П	П	Е	Ц	Ц	Г	Ц	Ц	©
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
Ю	А	Б	Ц	Д	Е	Ф	Г	Х	И	Й	К	Л	М	Н	О
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
П	Я	Р	С	Т	У	Ж	В	Ь	Ы	З	Ш	Э	Щ	Ч	Ъ
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
Ю	А	Б	Ц	Д	Е	Ф	Г	Х	И	Й	К	Л	М	Н	О
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
П	Я	Р	С	Т	У	Ж	В	Ь	Ы	З	Ш	Э	Щ	Ч	Ъ
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255

Кодировка CP1251

Á	à	,	è	„	…	†	‡	€	‰	É	<	й	Й	Ó	Ú
128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143
á	‘	’	“	”	•	–	—	ë	™	é	>	ò	й	ó	ú
144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159
nbsp	ỳ	Ы	Э	И	Ы	!	§	Ё	©	Ю	«	¬	shy	®	Я
160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175
°	±	Ы	Э	’	μ	¶	•	ё	№	Ю	»	э	Ю	Я	Я
176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191
А	Б	В	Г	Д	Е	Ж	З	И	Й	К	Л	М	Н	О	П
192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207
Р	С	Т	У	Ф	Х	Ц	Ч	Ш	Щ	Ъ	Ы	Ь	Э	Ю	Я
208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223
а	б	в	г	д	е	ж	з	и	й	к	л	м	н	о	п
224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239
р	с	т	у	ф	х	ц	ч	ш	щ	ъ	ы	ь	э	ю	я
240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255

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.

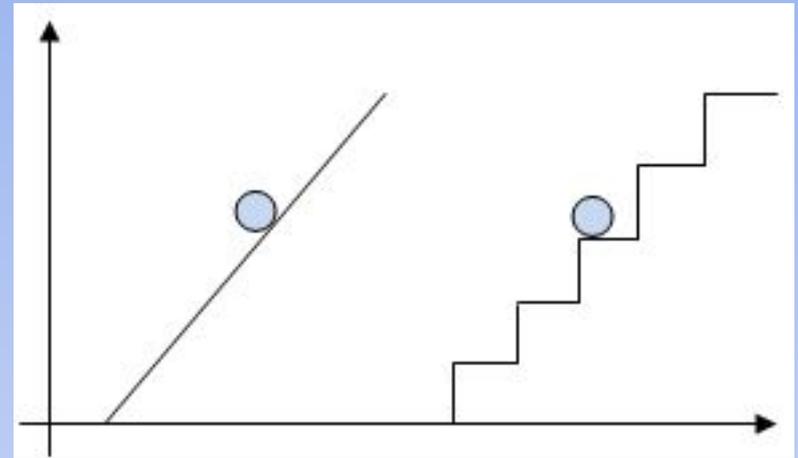
Информация, в том числе графическая и звуковая, может быть представлена в *аналоговой* или *дискретной* форме. При аналоговом представлении физическая величина принимает бесконечное множество значений, причем ее значения изменяются непрерывно. При дискретном представлении физическая величина принимает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



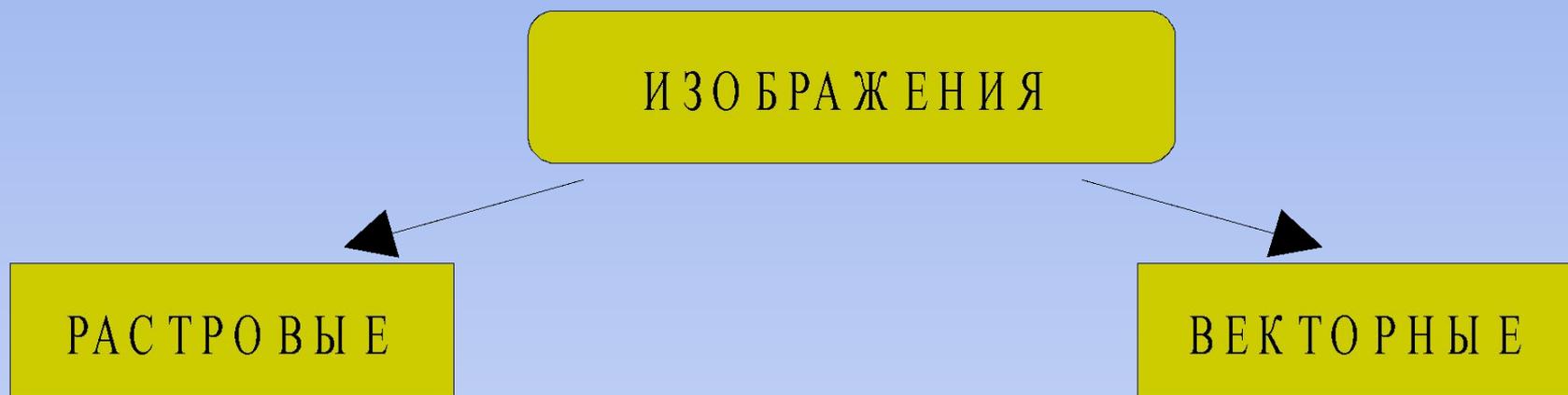
Дискретизация

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.

Виды компьютерных изображений



Создавать и хранить графические объекты в компьютере можно двумя способами – как растровое или как векторное изображение. Для каждого типа изображений используется свой способ кодирования.

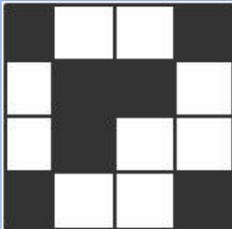
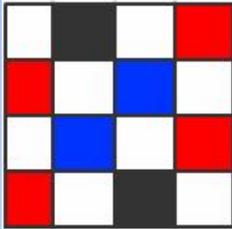
Кодирование растровых изображений

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).*

Качество изображения зависит от количества точек (чем меньше размер точки и, соответственно, больше их количество, тем лучше качество) и количества используемых цветов (чем больше цветов, тем качественнее кодируется изображение).

	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



Цветовые модели

Для представления цвета в виде числового кода используются две обратных друг другу цветовые модели: **RGB** или **CMYK**.

- Модель **RGB** используется в телевизорах, мониторах, проекторах, сканерах, цифровых фотоаппаратах... Основные цвета в этой модели: красный (**Red**), зеленый (**Green**), синий (**Blue**).
- Цветовая модель **CMYK** (Cyan, Magenta, Yellow, black) используется в полиграфии при формировании изображений, предназначенных для печати на бумаге.

Цветовая модель RGB

R	G	B	Цвет
1	1	1	белый
1	1	0	желтый
1	0	1	пурпурный
1	0	0	красный
0	1	1	голубой
0	1	0	зеленый
0	0	1	синий
0	0	0	черный

Цветные изображения могут иметь различную глубину цвета, которая задается количеством битов, используемых для кодирования цвета точки.

Если кодировать цвет одной точки изображения тремя битами (по одному биту на каждый цвет RGB), то мы получим все восемь различных цветов.

True Color

На практике же, для сохранения информации о цвете каждой точки цветного изображения в модели RGB обычно отводится 3 байта (т.е. 24 бита) - по 1 байту (т.е. по 8 бит) под значение цвета каждой составляющей.

Таким образом, каждая RGB-составляющая может принимать значение в диапазоне от 0 до 255 (всего $2^8=256$ значений), а каждая точка изображения, при такой системе кодирования может быть окрашена в один из 16 777 216 цветов.

Такой набор цветов принято называть True Color (правдивые цвета), потому что человеческий глаз все равно не в состоянии различить большего разнообразия.

Кодирование векторных изображений



Векторное изображение представляет собой совокупность графических примитивов (точка, отрезок, эллипс...). Каждый примитив описывается математическими формулами. Кодирование зависит от прикладной среды.

Достоинством векторной графики является то, что файлы, хранящие векторные графические изображения, имеют сравнительно небольшой объем.

Важно также, что векторные графические изображения могут быть увеличены или уменьшены без потери качества.

Графические форматы файлов

Форматы графических файлов определяют способ хранения информации в файле (растровый или векторный), а также форму хранения информации (используемый алгоритм сжатия).

Наиболее популярные растровые форматы:

- BMP
- GIF
- JPEG
- TIFF
- PNG

Графические форматы файлов

- **Bit MaP image (BMP)** — универсальный формат растровых графических файлов, используется в операционной системе Windows. Этот формат поддерживается многими графическими редакторами, в том числе редактором Paint. Рекомендуется для хранения и обмена данными с другими приложениями.
- **Tagged Image File Format (TIFF)** — формат растровых графических файлов, поддерживается всеми основными графическими редакторами и компьютерными платформами. Включает в себя алгоритм сжатия без потерь информации. Используется для обмена документами между различными программами. Рекомендуется для использования при работе с издательскими системами.
- **Graphics Interchange Format (GIF)** — формат растровых графических файлов, поддерживается приложениями для различных операционных систем. Включает алгоритм сжатия без потерь информации, позволяющий уменьшить объем файла в несколько раз. Рекомендуется для хранения изображений, создаваемых программным путем (диаграмм, графиков и так далее) и рисунков (типа аппликации) с ограниченным количеством цветов (до 256). Используется для размещения графических изображений на Web-страницах в Интернете.
- **Portable Network Graphic (PNG)** — формат растровых графических файлов, аналогичный формату GIF. Рекомендуется для размещения графических изображений на Web-страницах в Интернете.
- **Joint Photographic Expert Group (JPEG)** — формат растровых графических файлов, который реализует эффективный алгоритм сжатия (метод JPEG) для отсканированных фотографий и иллюстраций. Алгоритм сжатия позволяет уменьшить объем файла в десятки раз, однако приводит к необратимой потере части информации. Поддерживается приложениями для различных операционных систем. Используется для размещения графических изображений на Web-страницах в Интернете.

Кодирование звука

Использование компьютера для обработки звука началось позднее, нежели чисел, текстов и графики.

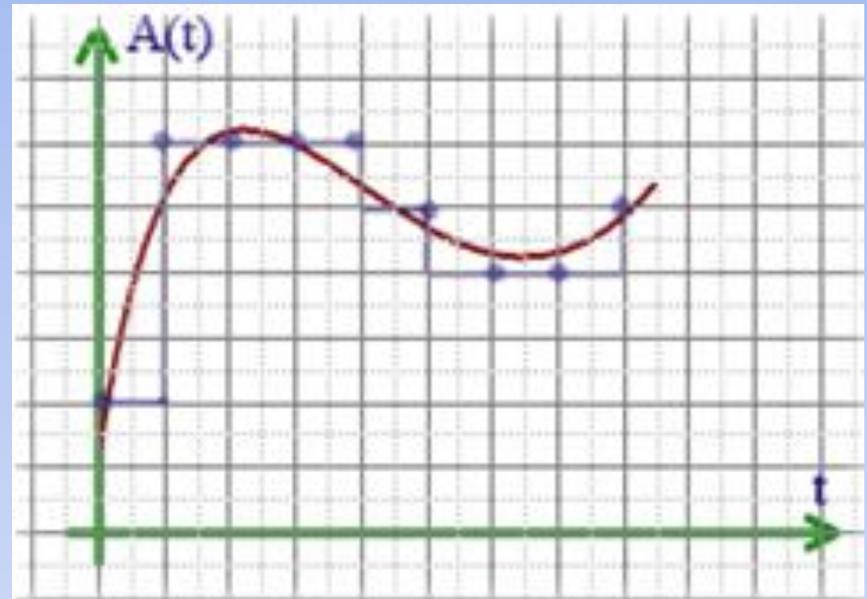
Звук – волна с непрерывно изменяющейся амплитудой и частотой. Чем больше амплитуда, тем он громче для человека, чем больше частота, тем выше тон.

Звуковые сигналы в окружающем нас мире необычайно разнообразны. Сложные непрерывные сигналы можно с достаточной точностью представлять в виде суммы некоторого числа простейших синусоидальных колебаний. Причем каждое слагаемое, то есть каждая синусоида, может быть точно задана некоторым набором числовых параметров – амплитуды, фазы и частоты, которые можно рассматривать как код звука в некоторый момент времени.

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.



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 $N = 2^1 = 2^{16} = 65536$.

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.

Представление видеоинформации

Казалось бы, если проблемы кодирования статической графики и звука решены, то сохранить видеоизображение уже не составит труда.

Но это только на первый взгляд, поскольку при использовании традиционных методов сохранения информации электронная версия фильма получится слишком большой.

Достаточно очевидное усовершенствование состоит в том, чтобы первый кадр запомнить целиком (в литературе его принято называть ключевым), а в следующих сохранять лишь отличия от начального кадра (разностные кадры).

Некоторые форматы видеофайлов

Существует множество различных форматов представления видеоданных.

- В среде Windows, например, уже более 10 лет применяется формат Video for Windows, базирующийся на универсальных файлах с расширением AVI (Audio Video Interleave – чередование аудио и видео).
- Более универсальным является мультимедийный формат Quick Time, первоначально возникший на компьютерах Apple.
- Все большее распространение в последнее время получают системы сжатия видеоизображений, допускающие некоторые незаметные для глаза искажения изображения с целью повышения степени сжатия. Наиболее известным стандартом подобного класса служит MPEG (Motion Picture Expert Group). Методы, применяемые в MPEG, непросты для понимания и опираются на достаточно сложную математику.
- Больше распространение получила технология под названием DivX (Digital Video Express). Благодаря DivX удалось достигнуть степени сжатия, позволившей вместить качественную запись полнометражного фильма на один компакт-диск – сжать 4,7 Гб DVD-фильма до 650 Мб.

Мультимедиа

Мультимедиа (multimedia, от англ. multi - много и media - носитель, среда) - совокупность компьютерных технологий, одновременно использующих несколько информационных сред: текст, графику, видео, фотографию, анимацию, звуковые эффекты, высококачественное звуковое сопровождение.

Под словом «мультимедиа» понимают воздействие на пользователя по нескольким информационным каналам одновременно. Можно еще сказать так: **мультимедиа – это объединение изображения на экране компьютера (в том числе и графической анимации и видеокадров) с текстом и звуковым сопровождением.**

Наибольшее распространение системы мультимедиа получили в области обучения, рекламы, развлечений.

Вопросы:

- Что такое код?
- Приведите примеры кодирования информации, используемые в школьных предметах?
- Придумайте свои способы кодирования русских букв.
- Закодируйте сообщение «информатика» с помощью кода Морзе.
- Что такое система счисления?
- Какие два вида систем счисления вы знаете?
- Что такое основание системы счисления? Что такое алфавит системы счисления? Примеры.
- В какой системе счисления хранятся и обрабатываются числа в памяти компьютера?
- Какие виды компьютерных изображений вы знаете?
- Какое максимальное количество цветов может быть использовано в изображении, если на каждую точку отводится 3 бита?
- Что вы знаете о цветовой модели RGB?

Задания:

- Запишите число 1945 в римской системе счисления.
- Запишите в развернутом виде числа: 2011_{10} , 957_8 , 110110_2 .
- Чему будут равны числа 74_8 , $3E_{16}$, $1010,1011_2$ в десятичной системе счисления?
- Как будет записываться число 143_{10} в двоичной системе счисления? 100_{10} в восьмеричной?
- Рассчитайте необходимый объем видеопамати для графического режима: разрешение экрана 800x600, качество цветопередачи 16 бит.