

# *Информация.*

*Николаева Наталья Николаевна*

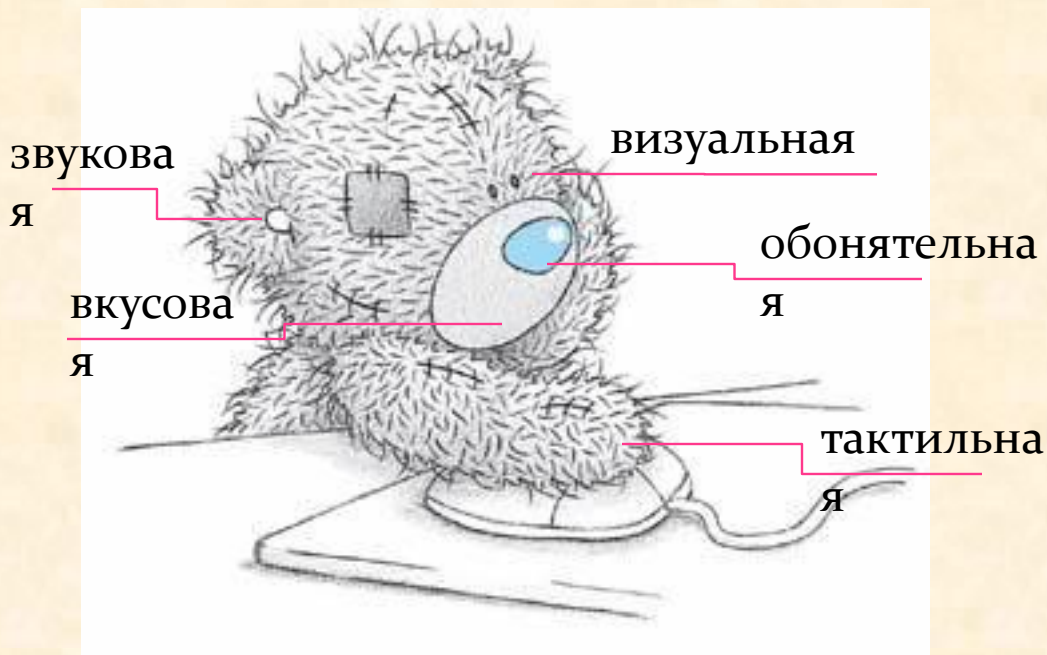
# Виды информации.

Многообразие окружающей нас информации можно сгруппировать по различным признакам. По признаку «область возникновения» информация делится на элементарную, биологическую и социальную. Элементарная информация отражает процессы и явления неодушевлённой природы, биологическая – процессы животного и растительного мира, а социальная – человеческого общества.

Информацию об окружающем мире человек получает с помощью пяти чувств: вкуса, обоняния, осязания, слуха и зрения. Поэтому различают виды информации по способу передачи и восприятию. Запахи и вкусы, воспринимаемые нами, называют органолептической информацией. Информацию, передаваемую звуками, называют звуковой (аудиальной), видимыми образами и символами – визуальной, ощущениями – тактильной, по запаху – обонятельная, по вкусу – вкусовая.

Информацию, создаваемую и используемую человеком, по общественному назначению делят на массовую, на специальную и на личную. К массовой информации относят общественно-политическую и научно-популярную, к специальной – техническую, экономическую и управленческую.

К сожалению, современные компьютеры не умеют воспроизводить запахи, вкусы и ощущения. Компьютер обрабатывает лишь следующие виды информации: текстовую, числовую, графическую и звуковую. Осталось разобраться, каким образом все эти виды информации представляются в современном компьютере.



# Представление звуковой информации.

- Современные персональные компьютеры умеют работать со звуковой информацией. Каждый компьютер, имеющий звуковую плату, микрофон и колонки, может записывать, сохранять и воспроизводить звуковую информацию. Звук представляет собой звуковую волну с непрерывно меняющейся амплитудой и частотой. Чем больше амплитуда, тем он громче для человека, чем больше частота сигнала, тем выше тон.

- Программное обеспечение компьютера в настоящее время позволяет непрерывный звуковой сигнал (аналоговый) преобразовывать в последовательность электрических импульсов, измеряя характеристики звука (период, амплитуду) через равные промежутки времени. Этот процесс выполняется в компьютере аналого-цифровым преобразователем (АЦП) и называется дискретизацией (или оцифровкой).

- Дискретизация – процесс преобразования аналогового сигнала в набор чисел, соответствующих характеристикам звука, измеренным через равные промежутки времени, называемые шагом квантования.

- Для выполнения обратного преобразования служит цифро-аналоговый преобразователь (ЦАП), который получает ступенчатый сигнал и сглаживает его.

Хранение звука в виде цифровой записи занимает достаточно много места на компьютере, поэтому существуют другие способы. Так, с начала 80-х годов XX века разрабатывается MIDI запись, представляющая собой набор команд-указаний MIDI-синтезатору, на каком инструменте и какую ноту играть. Любой MIDI-синтезатор может воспроизводить звучание более 150 различных музыкальных инструментов, имеющих свои номера и названия. Проигрывание идёт одновременно на 16 каналах (треках) в 16 оркестровых группах, состоящих из одинаковых инструментов. На каждом канале одновременно может звучать до 32 голосов – нот, извлекаемых из инструмента. Получается целый оркестр, в котором одновременно может звучать 512 голосов! Число голосов на одном канале характеризует полифонию синтезатора.



# Представление информации в компьютере.

- Когда человек воспринимает информацию с помощью органов чувств, то стремится зафиксировать ее в такой форме, чтобы она была понятна другим людям, например, записать ее. Для этого применяются специальные символы. Например, для представления текстовой информации используют буквы какого-либо языка (русского, английского, немецкого и т.д.). Не всегда эти буквы относятся к естественному языку. Для представления текстовой информации также можно использовать искусственные языки, например, эсперанто. При обработке информации человек дает инструкции компьютеру на искусственном языке – языке программирования. Музыка звучит с помощью нот, числа записывают с помощью цифр. Все это примеры кодирования информации.
- Кодирование – процесс представления данных последовательностью символов.
- Имеются и такие системы знаков, для которых нет какого-то общепринятого порядка, то есть, нет алфавита. К таким системам, например, относятся азбука Брайля (для слепых), язык жестов, язык цветов, знаки генетического кода и т.д.
- В отличие от человека, компьютер может обрабатывать информацию, представленную только в цифровой форме в виде последовательности цифр. Причем «компьютерный алфавит» состоит только из двух символов 0 и 1. Это связано с тем, что устройства, которые могут принимать только два устойчивых состояния: намагничено (1) – не намагничено (0), высокое напряжение (1) – низкое напряжение (0) и т.д., являются самыми надежными и дешевыми. В современной электронике развитие аппаратной базы компьютеров идет именно в этом направлении. Поэтому предварительно вся информация, представленная в компьютере, и текстовая, и графическая, и звуковая должна пройти оцифровку.
- Оцифровка – приведение чего-либо к числовому виду или кодирование чего-либо с помощью чисел.
- Для представления текстовой информации в компьютере или для ее кодирования используют специальные кодовые таблицы. В таких таблицах с каждым символом сопоставляется число. Примером может служить таблица ASCII кодов.
- Кодирование, при котором с каждым символом алфавита сопоставляется код, называется алфавитным кодированием.
- Заметим, что таблица ASCII кодов стандартизирует только первые коды с 32 по 127. Эти коды являются стандартными и обязательными для всех стран и всех компьютеров, а во второй половине (128–255) каждая страна может создавать свой стандарт – национальный.
- В конце 80-х годов прошлого века многие осознали необходимость создания единого стандарта на кодирование символов, учитывающего национальные стандарты, что и привело к появлению Unicode. Этот стандарт использует для кодирования 65536 кодов символов, в отличие от 256, принятого в стандарте ASCII, что оказывается достаточным для всех существующих языков, математических символов и других знаков. Первые 256 индексов используется для совместимости со стандартом ASCII.



# Представление графики

## №1.

- Любое изображение на мониторе компьютера представляет собой набор светящихся точек (пикселей). Качество картинка, которую мы видим на мониторе, напрямую зависит от разрешения и количества цветов. На рисунке вы видите, как изменяется картинка, с уменьшением разрешения.
- Пиксель (pixel) – сокращение от «picture element» – «элемент рисунка». Пиксель это минимальный блок информации на экране монитора, наименьшая из составляющих изображения.
- Разрешение – количества пикселей по горизонтали и вертикали. Рисунок 1. Изображение с различным разрешением.



Рисунок 1. Изображение с различным разрешением.

# Представление графики

## №2.

- Так происходит потому, что при оцифровке часть информации теряется. Для того чтобы понять, почему так происходит. Рассмотрим процесс оцифровки подробнее. Пусть черно-белое изображение, представленное на рисунке, нужно показать на гипотетическом мониторе, имеющем по 8 пикселей по оси X и 8- по оси Y соответственно. Тогда исходную картинку (слева) нужно разбить на 64 областей, как показано на рисунке 2 (в центре). Далее, цвет пикселей в каждой из областей усредняется (с учетом двух цветов: белого и черного), в результате чего, мы получим «размытый» рисунок меньшего качества. Каждая область, показанная на рисунке 2 (справа), и представляет собой пиксель на нашем гипотетическом мониторе.



Рисунок 2. Искажение изображения при оцифровке.

- Процесс усреднения происходит следующим образом, если большинство пикселей в области черные, то цвет всей области будет черным и наоборот.

# Представление графики

## №3.

- Цветные картинки на мониторе мы видим благодаря смешению трех составляющих цвета: красной, зеленой и голубой. Яркость каждой составляющей меняется от 0 до 255, благодаря чему при их смешении мы можем видеть более 16 миллионов цветов.
- Когда графическое изображение представлено в виде таблицы, в каждой ячейке которой хранится информация о яркости пикселя, составляющего картинку, то говорят о растровой графике.
- На рисунке 3 представлено, как искажаются цвета картинки, если в палитре около 16 миллионов цветов (левая картинка), если 256 цветов (центральная картинка), и если 16 цветов (правая картинка).



Рисунок 3. Искажение цвета при оцифровке.

# Представление числовой информации.

- Существуют два основных формата представления чисел в памяти компьютера. Один из них используется для кодирования целых чисел, а второй для кодирования действительных чисел. Целые числа представляются в формате с фиксированной запятой, а действительные – в формате с плавающей точкой. Мы с вами уже говорили о том, что алфавит компьютера состоит всего лишь из двух знаков: 0 и 1. Все числа в компьютере и представляются с помощью этих двух знаков. Если говорить о том, как мы с вами записываем числа, то легко заметить, что мы это делаем всего лишь с помощью десяти цифр: 0, 1, 2, 3, 4, 5, 6, 7, 8 и 9. Например, число 234 означает, что в числе 2 сотни, 3 десятка и 4 единицы:  
$$234 = 2 \cdot 10^2 + 3 \cdot 10^1 + 4 \cdot 10^0$$
- Другими словами, мы с вами работаем с числами в позиционной системе счисления и представляем число как сумму цифр умноженных на соответствующую степень десяти, то есть в десятичной системе счисления. Компьютер же работает с числами, записанными в двоичной системе. В этой системе счисления всего лишь две цифры: 0 и 1, а числа записываются по степеням двойки. Например, число пять кодируется как 101, потому что
- Система счисления – искусственная информационная система, созданная с целью выработки наиболее удобного способа записи чисел.
- Позиционными называют системы счисления, в которых вклад каждой цифры в величину числа зависит от её положения (позиции) в последовательности цифр, изображающей число.
- В десятичной системе счисления используют 10 цифр: 0, 1, 2, 3, 4, 5, 6, 7, 8 и 9. Вклад каждой цифры в число определяется ее местом расположения. Самая крайняя цифра вносит вклад пропорциональный 100, вторая с конца цифра – 10, третья – 102 и т.д. Подробнее о системах счисления Вы можете прочитать во второй теме нашего учебника.
- По словам знаменитого математика и физика XVIII – XIX веков П.Лапласа: «Мысль выражать все числа десятью знаками, придавая им, кроме значения по форме, еще значение по месту, настолько проста, что именно из-за этой простоты трудно понять, насколько она удивительна».
- $5 = 1 \cdot 2^2 + 0 \cdot 2^1 + 1 \cdot 2^0 = 101_2$
- Каждый регистр арифметического устройства компьютера, каждая ячейка памяти представляет собой совокупность однородных элементов, каждый из которых может принимать два состояния и используется для записи соответствующего разряда числа. Нумерацию разрядов в ячейке принято вести справа налево.



Рисунок 4. Восьмиразрядная ячейка.

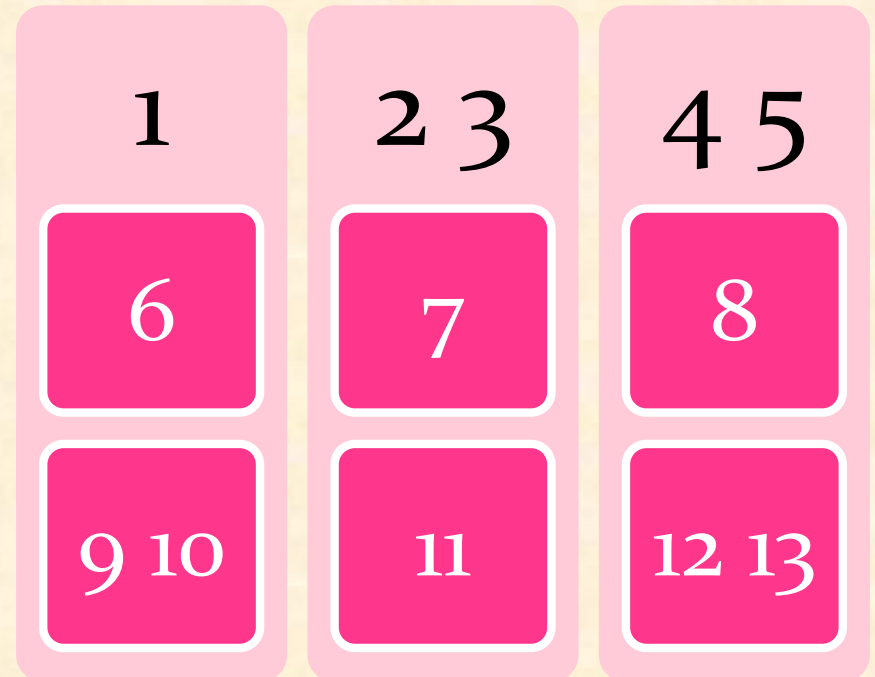


# Единицы измерения объема информации.

- Впервые объективный подход к измерению информации был предложен американским инженером Р. Хартли в 1928 году, затем в 1948 году обобщен американским учёным К. Шенноном. Хартли рассматривал процесс получения информации как выбор одного сообщения из конечного наперед заданного множества из  $N$  равновероятных сообщений, а количество информации  $I$ , содержащееся в выбранном сообщении, определял как двоичный логарифм  $N$ .
- Вероятность - численная мера достоверности случайного события, которая при большом числе испытаний близка к отношению числа случаев, когда событие осуществилось с положительным исходом, к общему числу случаев. Два события называют равновероятными, если их вероятности совпадают.
- Примеры равновероятных событий
  - 1. при бросании монеты: «выпала решка», «выпал орел»;
  - 2. на странице книги: «количество букв чётное», «количество букв нечётное»;
  - 3. при бросании игральной кости: «выпала цифра 1», «выпала цифра 2», «выпала цифра 3», «выпала цифра 4», «выпала цифра 5», «выпала цифра 6».
- Неравновероятные события
- Определим, являются ли равновероятными сообщения «первой из дверей здания выйдет женщина» и «первым из дверей здания выйдет мужчина». Однозначно ответить на этот вопрос нельзя. Во-первых, как известно количество мужчин и женщин неодинаково. Во-вторых, все зависит от того, о каком именно здании идет речь. Если это военная казарма, то для мужчины эта вероятность значительно выше, чем для женщины.
- Поскольку каждый регистр арифметического устройства и каждая ячейка памяти состоит из однородных элементов, а каждый элемент может находиться в одном из двух устойчивых состояний (которые можно отождествить с нулем и единицей), то К. Шенноном была введена единица измерения информации – бит.
- Бит – слишком мелкая единица измерения. На практике чаще применяется более крупная единица – байт, равная восьми битам. Именно восемь битов требуется для того, чтобы закодировать любой из 256 символов алфавита клавиатуры компьютера ( $256=2^8$ ).
- Широко используются также еще более крупные производные единицы информации:
  - 1 Килобайт (Кбайт) =  $1024$  байт =  $2^{10}$  байт,
  - 1 Мегабайт (Мбайт) =  $1024$  Кбайт =  $2^{20}$  байт,
  - 1 Гигабайт (Гбайт) =  $1024$  Мбайт =  $2^{30}$  байт.
- В последнее время в связи с увеличением объемов обрабатываемой информации входят в употребление такие производные единицы, как:
  - 1 Терабайт (Тбайт) =  $1024$  Гбайт =  $2^{40}$  байт,
  - 1 Петабайт (Пбайт) =  $1024$  Тбайт =  $2^{50}$  байт.
- За единицу информации можно было бы выбрать количество информации, необходимое для различения, например, десяти равновероятных сообщений. Это будет не двоичная (бит), а десятичная (дит) единица информации.

# Задачи 1.

- Текст состоит из 20-ти символов (каждый символ кодируется 8 битами). Сколько места в памяти он занимает?
- Решение:
- Раз каждый символ занимает 8 бит, то 20 символов будут занимать  $20 \times 8 = 160$  бит = 20 байт



## Задача 2.

Какое максимально возможное количество цветов в изображении, если цвет кодируется 24 битами?

Решение:

Для того чтобы вычислить максимальное количество цветов, достаточно возвести два в степень 24.

$2 \text{ в степ. } 24 = 16\,777\,216$

Ответ: 24 битами можно представить  $16\,777\,216$  цветов .  
текст содержит 20 байт



## Задача 3.

### Задача 3.

Пусть имеется колода из 32 карт (в колоде отсутствуют шестерки). Задумана одна карта (например, туз пик). Сколько двоичных вопросов нужно задать, чтобы отгадать задуманную карту.

Двоичным называют вопрос, который предполагает только два взаимоисключающих ответа «да» или «нет».

Решение:

Сколько бит информации несет сообщение о задуманной карте? Давайте рассуждать. Сначала, все карты нам нужно разбить на два подмножества так, чтобы вероятности принадлежности задуманной карты к ним, были равными. Разобьем карты по цветам масти: красная и черная.

Первый вопрос: Задумана карта черной масти?

Сколько бит информации мы получили? 1 бит информации. Почему? Потому что уменьшили неопределенность ровно вдвое.

Далее, описанную процедуру следует повторить.

Второй вопрос: Масть карты пика? (1 бит информации)

Третий вопрос: Задумана карта-картинка? (1 бит информации)

Четвертый вопрос: Задумана карта старше дамы? (1 бит информации)

Пятый вопрос: Это король? (1 бит информации)

Ответ: задуманная карта несет 5 бит информации, поэтому нужно задать пять двоичных вопросов.





# Задача 4.

- В корзине 8 шаров. Все шары разного цвета. Сколько информации несет сообщение о том, что из корзины достали красный шар.

.....

Решение:

- $N=8$ ,  $b=?$   $N=2^{\text{в степ.} b}$   
 $b=3$
- Ответ: 3 бита.



# Задача 5.

- Для записи текста использовали 64-х символьный алфавит. Каждая страница содержит 30 строк по 40 символов в строке. Какой объем информации несет 4 страницы текста?
- Решение:  $N=64$  30 стр. 40 симв
- $N=2$  в степ.  $b$   $N=64$   $b=6$
- $30 \times 40 = 1200$  симв. В 1 стр.
- $1200 \times 4 = 4800$  на 5 стр.
- $4800 \times 7 = 33600$   $V$  текста.
- Ответ: 33600



# Задача 6.

- В коробке лежат 16 цветных карандашей. Какое количество информации содержит сообщение, что из коробки достали синий карандаш?
- Решение:  
 $N=16$   $b=?$   
 $N = 2$  в степ.  $b$   $b=3$
- Ответ: 3 бит



# Задача 7.

- Какое кол-во информации (в Unicode и в ASCII) несет в себе слово орхидея?  
Решение: орхидея- 7 знаков;  $7 \times 8 = 56$  бит в ASCII;  $7 \times 16 = 112$  бит в Unicode.
- Ответ: 56 бит, 112 бит.





# Задача 8.

- Два текста содержат одинаковое кол-во символов, 1 текст мощностью 256 символов, 2 текст 512. Во сколько раз мощность 1 меньше мощности 2?

Решение:

$$N_1=256 \quad N_2=512 \quad V_1:V_2=?$$

$$N=2 \text{ в степ.} b \quad N_1=256$$

$$N_2=512; \quad b_1=8 \quad b_2=9;$$

$$V_1:V_2=8:9.$$

$$\text{Ответ: } V_1:V_2=8:9.$$

