

Раздел 2: Кодирование текстовой, графической и звуковой информации

КОЛИЧЕСТВО ИНФОРМАЦИИ.
ВЕРОЯТНОСТНЫЙ И
АЛФАВИТНЫЙ ПОДХОД К
ИЗМЕРЕНИЮ ИНФОРМАЦИИ.

ГАПОУ МО МСК им. Н.Е. Момота

Преподаватель информатики: Зенченко О.С.

КОЛИЧЕСТВО ИНФОРМАЦИИ

Мы с вами уже говорили, что в основе нашего мира лежат 3 составляющие -
вещество, энергия, информация.

А как много в мире вещества, энергии и информации?

- Можно ли измерить количество вещества? В каких единицах?
- Можно ли определить количество энергии?
- Можно ли измерить количество информации? Как это сделать?

Оказывается, информацию тоже можно измерить и находить ее количество.

Существуют 2 подхода к измерению количества информации

ВЕРОЯТНОСТНЫЙ ПОДХОД

Определите количество информации в следующих сообщениях с позиции «много» или «мало»?

- Столица России - Москва
- Сумма квадратов катетов равна квадрату гипотенузы
- Эйфелева башня имеет высоту 300 метров и вес 9000 тонн

Сообщение несет больше информации, если в нем содержатся Новые понятные сведения. Такое сообщение называется **ИНФОРМАТИВНЫМ**

ВЫВОД: количество информации зависит от информативности.

Рассмотрим понятие информативности с другой стороны .

Если некоторое сообщение является информативным, следовательно , оно пополняет нас знаниями или уменьшает нашу неопределенность знаний. Другими словами сообщение содержит информацию, если оно приводит нас к уменьшению неопределенности наших знаний.

Пример: подбрасывание монеты

Количество информации - мера уменьшения неопределенности знаний при получении информационных сообщений.

Единица количества информации - **бит**

1 бит - количество информации в сообщении, которое уменьшает неопределенность в 2 раза.

Формула Хартли

$$N = 2^i$$

Где **N**- Количество вариантов исхода

i- количество информации, которое несет сообщение (в битах)

Связь количества информации и вариантов исхода

I (бит)	N (варианты исхода)
1	$2^1 = 2$
2	$2^2 = 4$
3	$2^3 = 8$
4	$2^4 = 16$
5	$2^5 = 32$
6	$2^6 = 64$
7	$2^7 = 128$
8	$2^8 = 256$
9	$2^9 = 512$
10	$2^{10} = 1024$

Примеры решения задач по новой теме:

№ 1

«Вы выходите на следующей остановке?» - спросили человека в автобусе. «Нет», - ответил он. Сколько информации содержит ответ?

Решение: человек мог ответить только «Да» или «Нет», т.е. выбрать один ответ из двух возможных. Поэтому $N = 2$. Значит $I = 1$ бит.

Ответ: 1 бит.

№2

«Петя! Ты пойдешь сегодня в кино?» - спросил я друга. «Да», - ответил Петя. Сколько информации я получил?

Решение: Петя мог ответить только «Да» или «Нет», т.е. выбрать один ответ из двух возможных. Поэтому $N = 2$, значит $I = 1$ бит.

Ответ: 1 бит.

№3

Сколько информации содержит сообщение, уменьшающее неопределенность знаний в 8 раз?

Решение: так как неопределенность знаний уменьшается в 8 раз, следовательно, она было равна 8, т.е. существовало 8 равновероятных событий. Сообщение о том, что произошло одно из них, несет 3 бита информации.

Ответ: 3 бита.

№4

Какой объем информации содержит сообщение, уменьшающее неопределенность в 4 раза?

Решение: так как неопределенность знаний уменьшается в 4 раз, следовательно, она было равна 4, т.е. существовало 4 равновероятных событий. Сообщение о том, что произошло одно из них, несет 2 бита информации ($4 = 2^2$).

Ответ: 2 бита.

№5

Вы подошли к светофору, когда горел желтый свет. После этого загорелся зеленый. Какое количество информации вы при этом получили?

Решение: из двух сигналов (желтого и зеленого) необходимо выбрать один - зеленый. Поэтому $N = 2$, а $I = 1$ бит.

Ответ: 1 бит.

№6

Группа школьников пришла в бассейн, в котором 4 дорожки для плавания. Тренер сообщил, что группа будет плавать на дорожке номер 3. Сколько информации получили школьники из этого сообщения?

Решение: из 4 дорожек необходимо выбрать одну, т.е. $N = 4$. Значит по формуле $I = 2$, т.к. $4 = 2^2$.

Пояснение: номер дорожки (3) не влияет на количество информации, так как вероятности событий в этих задачах мы приняли считать одинаковыми.

Ответ: 2 бита.

№7

На железнодорожном вокзале 8 путей отправления поездов. Вам сообщили, что ваш поезд прибывает на четвертый путь. Сколько информации вы получили?

Решение: из 8 путей нужно выбрать один. Поэтому $N = 8$, а $I = 3$, т.к. $8 = 2^3$.

Пояснение: номер пути (4) не влияет на количество информации, так как вероятности событий в этих задачах мы приняли считать одинаковыми.

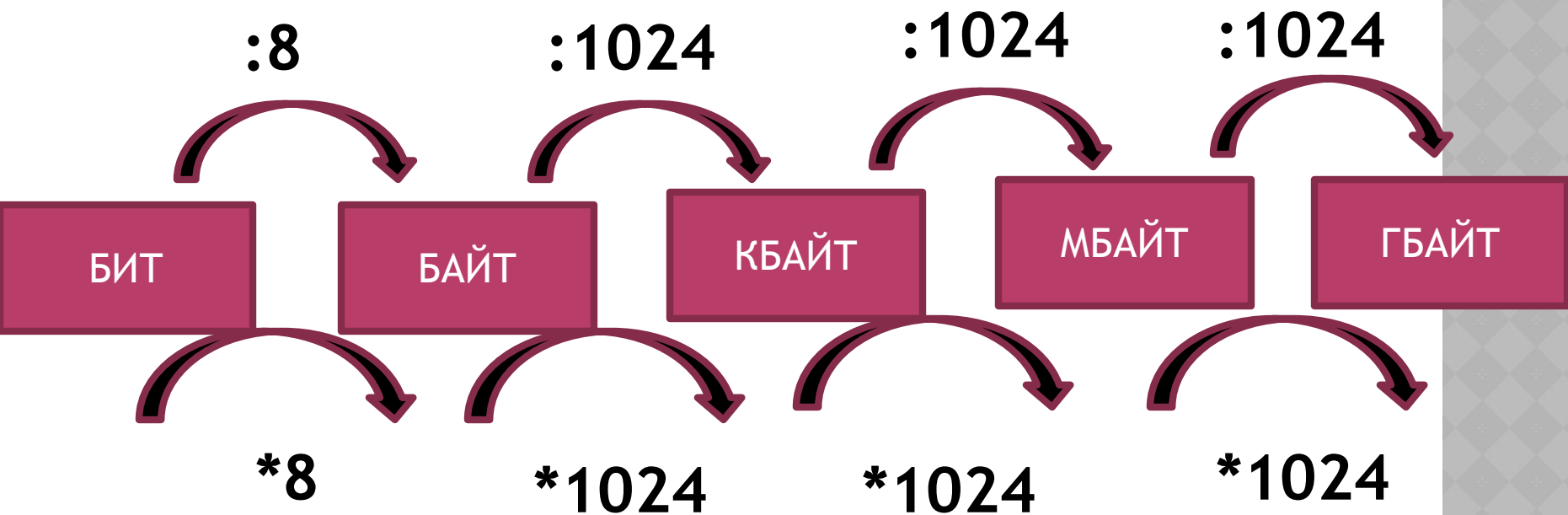
Ответ: 3 бита.

№8

В коробке лежат 16 кубиков. Все кубики разного цвета. Сколько информации несет сообщение о том, что из коробки достали красный кубик?

Решение: из 16 равновероятных событий нужно выбрать одно. Поэтому $N = 16$, следовательно, $I = 4$ ($16 = 2^4$).

ЕДИНИЦЫ ИЗМЕРЕНИЯ КОЛИЧЕСТВА ИНФОРМАЦИИ



АЛФАВИТНЫЙ ПОДХОД К ОПРЕДЕЛЕНИЮ КОЛИЧЕСТВА ИНФОРМАЦИИ

С точки зрения ЭВМ текст состоит из отдельных символов. К числу символов принадлежат не только буквы (заглавные или строчные, латинские или русские), но и цифры, знаки препинания, спецсимволы типа "=", "(", "&" и т.п. и даже (обратите особое внимание!) пробелы между словами. Да, не удивляйтесь: пустое место в тексте тоже должно иметь свое обозначение.

Вспомним некоторые известные нам факты: Множество символов, с помощью которых записывается текст, называется алфавитом. Число символов в алфавите – это его мощность. В алфавит мощностью 256 символов можно поместить практически все необходимые символы. Такой алфавит называется достаточным. Т.к. $256 = 2^8$, то вес 1 символа – 8 бит. Единице измерения 8 бит присвоили название 1 байт: 1 байт = 8 бит. Двоичный код каждого символа в компьютерном тексте занимает 1 байт памяти.

Кодирование текстовой информации

$$N = 2^i$$

Где: N - мощность алфавита
(количество символов в алфавите)
 i - информационный вес символа
(количество бит)

$$I = K \cdot i$$

I - количество информации в тексте
(в битах)
 K - количество символов в тексте
 i - вес одного символа
(количество бит)

$$K = a \cdot b \cdot c$$

a - количество символов в строке
 b - количество строк
 c - количество страниц

ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ

Решение задач [Кодирование символов]

Задача №1.

Считая, что каждый символ кодируется одним байтом, определите, чему равен информационный объём следующего высказывания Жан-Жака Руссо:

Тысячи путей ведут к заблуждению, к истине – только один.

- 1) 92 бита 2) 220 бит 3) 456 бит 4) 512 бит

Дано:

$i = 1$ байт

$k = 57$

Найти:

$I = ?$

Решение:

$$I = k \cdot i \Rightarrow$$

1) $i = 1$ байт = 8 бит

2) $I = 57 \cdot 8 = 456$ (бит)

Ответ: 456 бит [Номер ответа: 3]

Вопросы и задания



Задание 2. В 15-м издании энциклопедии Britannica 32 тома, в каждом из которых порядка 1000 страниц. На одной странице размещается в среднем 70 строк по 120 символов (включая пробелы) в каждой. Найдите объем текстовой информации в энциклопедии, если при записи используется кодировка Unicode («один символ — два байта»).

Дано:

$i = 2$ байта

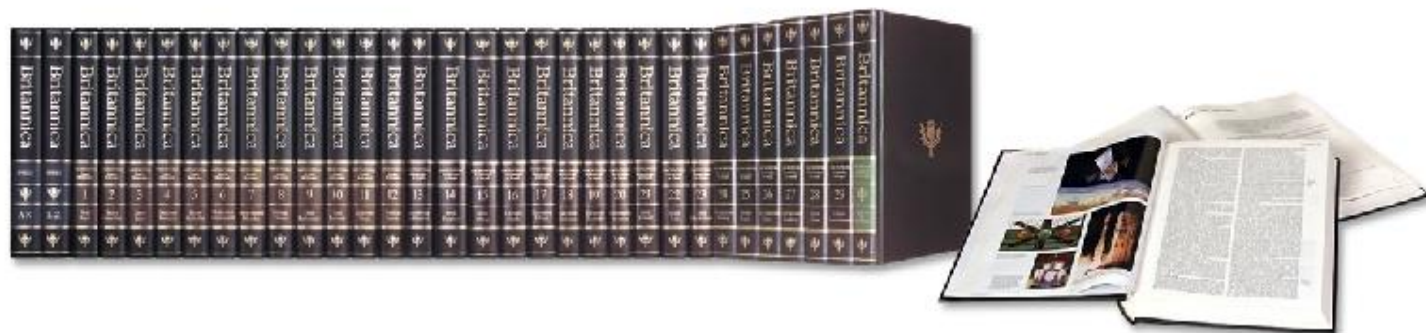
$K = 32 \cdot 1000 \cdot 70 \cdot 120$

$I = ?$

$I = K \cdot i$

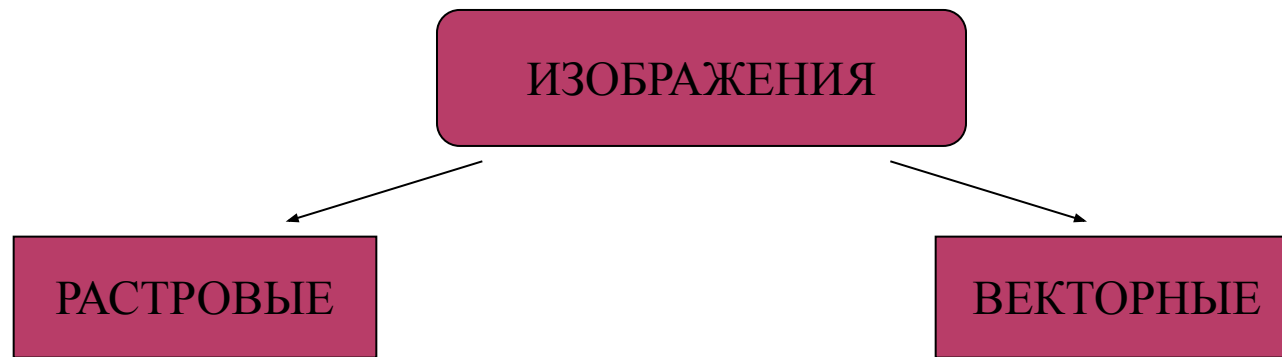
$$I = \frac{32 \cdot 1000 \cdot 70 \cdot 120 \cdot 2}{1024 \cdot 1024} \text{ Мб} \approx 513 \text{ Мб}$$

Ответ: 513 Мбайт



КОДИРОВАНИЕ ГРАФИЧЕСКОЙ ИНФОРМАЦИИ

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.

Информационный объем одной точки в битах (глубина цвета - i) зависит от количества возможных цветов - N (палитры).

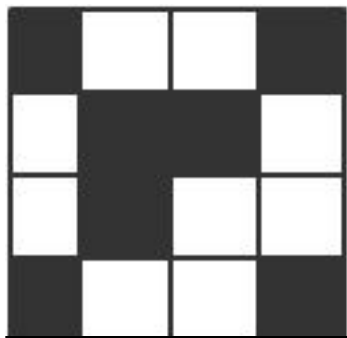
$$N = 2^i$$

Объем растрового изображения определяется как произведение количества точек пикселей и информационного объема одной точки:

$$I = x \cdot y \cdot i \cdot c$$

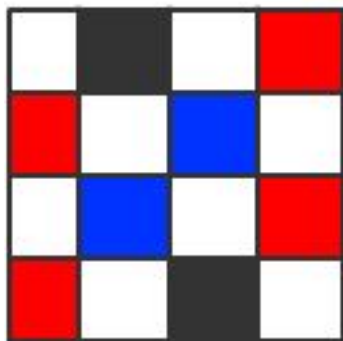
I - объем графического файла (в битах)
 i - глубина цвета (количество бит)
 x, y - разрешение монитора
(по горизонтали, по вертикали)
 c - количество изображений

ДЛЯ КОДИРОВАНИЯ ЧЕРНО-БЕЛОГО ИЗОБРАЖЕНИЯ
ГЛУБИНА ЦВЕТА СОСТАВЛЯЕТ 1 БИТ.



```
1 0 0 1
0 1 1 0
0 1 0 0
1 0 0 1
```

Для кодирования четырехцветного
изображения глубина цвета составляет 2
бита.



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

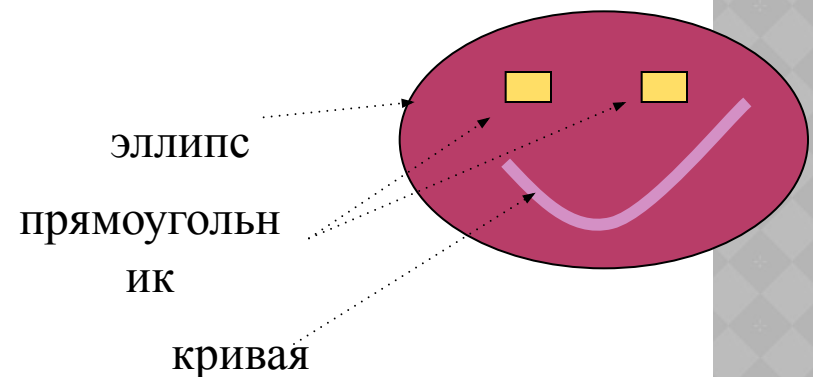

Цветное изображение на экране монитора формируется за счет смешивания трех базовых цветов: красного, зеленого, синего (модель RGB).

Для получения богатой палитры базовым цветам могут быть заданы различные интенсивности.

4 294 967 296 цветов (True Color) - 32 бита (4 байта).

КОДИРОВАНИЕ ВЕКТОРНЫХ ИЗОБРАЖЕНИЙ

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



ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ

Решение задач [Кодирование рисунков]

Задача №1.

Сколько памяти компьютера требуется для двоичного кодирования цветного рисунка (256 цветов) размером 10×10 точек?

Дано:

$$N = 256$$

$$x \cdot y = 10 \times 10$$

Найти:

$$I = ?$$

Решение:

$$I = x \cdot y \cdot i, N = 2^i \Rightarrow$$

$$1) 256 = 2^i \Rightarrow i = 8 \text{ бит}$$

$$2) I = 10 \cdot 10 \cdot 8 = 800 \text{ бит} \\ = 100 \text{ байт}$$

Ответ: 100 байт

В цветовой модели RGB для кодирования одного пикселя используется 3 байта. Фотографию размером 2048x1536 пикселей сохранили в виде несжатого файла с использованием RGB-кодирования. Определите размер получившегося файла.

- 1) 3 килобайта
- 2) 3 мегабайта
- 3) 9 килобайт
- 4) 9 мегабайт

Дано:

$$x * y = 2048 * 1536$$

$$i = 3 \text{ байта}$$

$$V = ?$$

Решение:

$$V = x * y * i = 2048 * 1536 * 3 \text{ байта} =$$

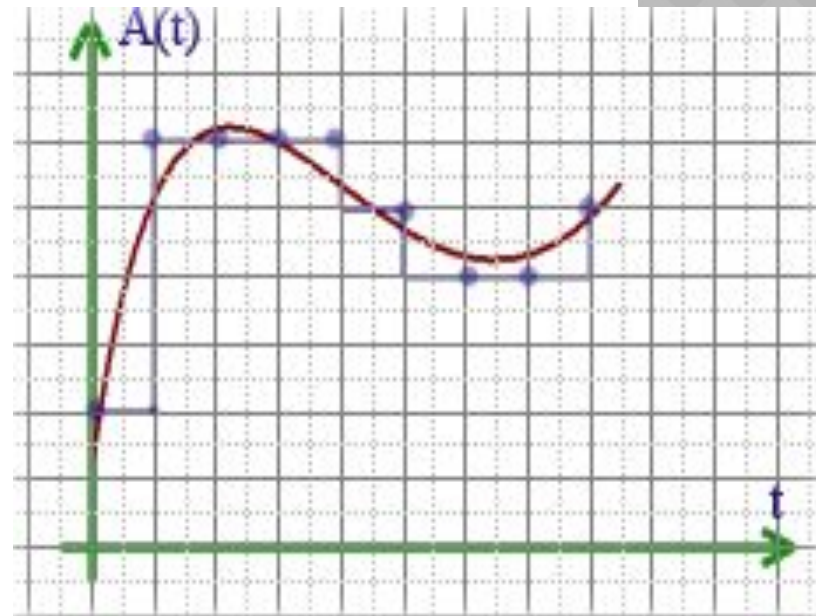
$$= 9437184 \text{ байта} = 9216 \text{ Кбайт}$$

$$= 9 \text{ Мбайт}$$

ДВОИЧНОЕ КОДИРОВАНИЕ ЗВУКА

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

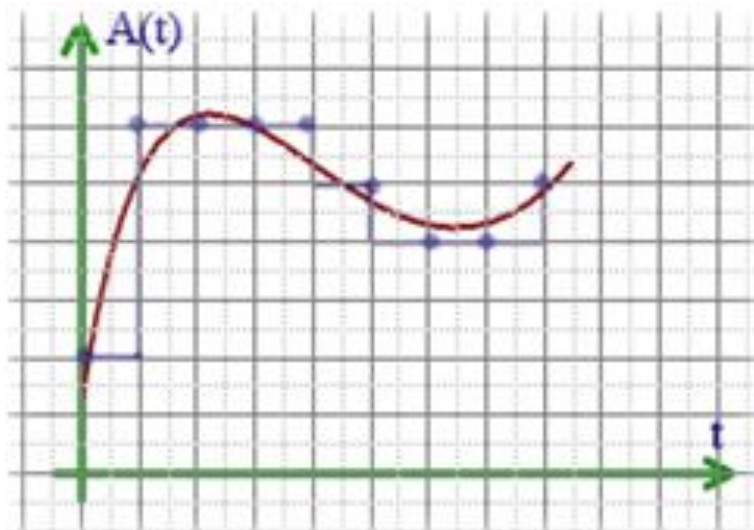
В процессе кодирования звукового сигнала производится его временная дискретизация - непрерывная волна разбивается на отдельные маленькие временные участки.



Кодирование звуковой информации.

Звуковой сигнал – это непрерывная волна с изменяющейся амплитудой и частотой.

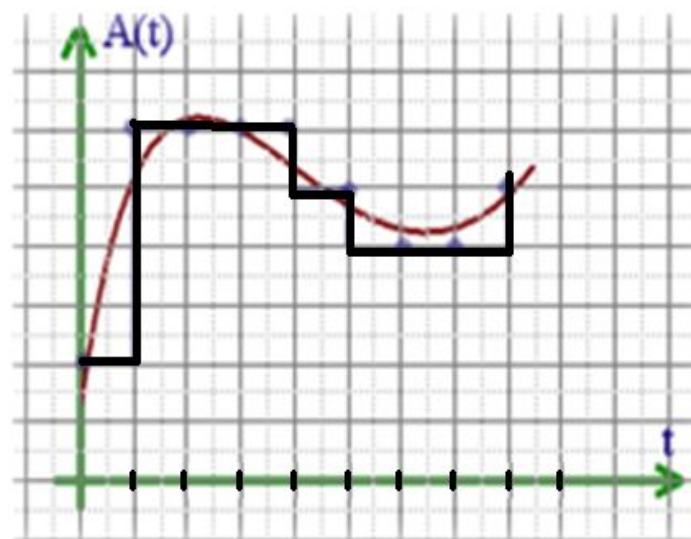
Чтобы компьютер мог обрабатывать звук, звуковой сигнал должен быть преобразован в цифровую дискретную форму с помощью **временной дискретизации**.

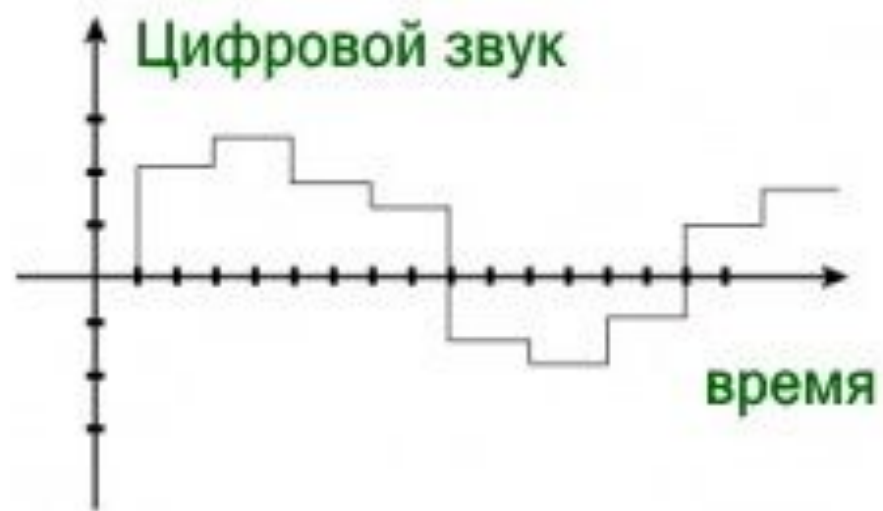


Непрерывная звуковая волна разбивается на отдельные маленькие временные участки

Причем для каждого такого участка устанавливается определенный уровень громкости

Происходит замена *гладкой кривой* на последовательность «ступенек»





Качество полученного цифрового звука зависит от частоты дискретизации.

Частота дискретизации – количество измерений громкости звука за 1 секунду.

ν - частота дискретизации (измеряется в Гц)

N – уровни громкости

i – глубина кодирования звука (количество бит)

$$N = 2^i$$

$$I = \nu \cdot t \cdot i \cdot k$$

I - объем звукового файла (в битах)

t – длительность звукового файла (сек)

k - режим

стерео- и **моно-**режимы

$$k = 1$$

$$k = 2$$

ПРИМЕРЫ РЕШЕНИЯ ЗАДАЧ

Вопросы и задания



Задание 1. В Новый год Петя Иванов записал бой курантов. Перезвон длился 20 секунд, а бой курантов – еще 40 секунд. Определите объём полученного моноаудио-файла (в килобайтах), записанного с глубиной кодирования 8 бит и частотой дискретизации 8 кГц.

Дано:

$$t = 20 \text{ с} + 40 \text{ с} = 60 \text{ с}$$

$$i = 8 \text{ бит} = 1 \text{ байт}$$

$$v = 8 \text{ кГц} = 8000 \text{ Гц}$$

I - ?

$$I = \frac{60 \cdot 1 \cdot 8000}{1024} \text{ Кб} = 468,75 \text{ Кб}$$

Ответ: 468,75 Кбайта

Оценка объёма звукового файла

Определить объем памяти для хранения моноаудиофайла, время звучания которого составляет пять минут при частоте дискретизации 44 кГц и глубине кодирования 16 бит.

Дано:

$$v = 44 \text{ кГц}$$

$$i = 16 \text{ бит}$$

$$t = 5 \text{ мин}$$

$$k = 1$$

Найти:

I

Решение:

$$44 \text{ кГц} = 44000 \text{ Гц}$$

$$5 \text{ мин} = 300 \text{ с}$$

$$I = k v i t$$

$$I = 1 \cdot 44000 \text{ Гц} \cdot 16 \text{ бит} \cdot 300 \text{ с} = 26400000 \text{ бит} \approx$$

$$\approx 25781,25 \text{ Кб} \approx 25,2 \text{ Мб}$$

Ответ: $I = 25,2 \text{ Мб}$