



Кодирование текстовой, графической и звуковой информации

Количество информации как мера уменьшения неопределённости знания

Определение

За единицу количества информации принимается такое количество информации, которое содержит сообщение, уменьшающее неопределённость знания в два раза. Такая единица названа бит.

Единицы измерения количества информации

- 1 байт = 2^3 бит = 8 бит;
- 1 Кбайт = 2^{10} байт = 1024 байт;
- 1 Мбайт = 2^{10} Кбайт = 1024 Кбайт;
- 1 Гбайт = 2^{10} Мбайт = 1024 Мбайт.

Формула

Существует формула, которая связывает между собой количество возможных событий **N** и количество информации **I**

$$N = 2^I$$

Задача 1

- В донорский пункт пришли 56 человек. Количество миллилитров крови, взятых у каждого донора (от 50 до 500), записали минимально возможным количеством бит. Определите информационный объём этих результатов.

Варианты ответа:

- 1) 525 бит 2) 500 бит 3) 256 байт 4) 63 байта

Решение

- $N = 500 - 49 = 451$ возможное количество вариантов.
- $N = 2^l$ $512 = 2^9$ Каждое число можно записать девятью битами,
т. е. 9 – информационный вес каждого числа.
- $56 * 9 = 504$ бита
- $504 / 8 = 63$ байта

Алфавитный подход к определению количества информации

Задача 2 Считая, что каждый символ кодируется 16 битами, оцените объём следующего предложения в кодировке Unicode:

Дураков на свете мало, но они расставлены так грамотно, что встречаются на каждом шагу.

Варианты ответа:

1) 87 байт

2) 174 бита

3) 174 байта

4) 87 бит

Двоичное кодирование текстовой информации в компьютере

- **Задача 3** Для записи текста использовался 256-символьный алфавит. Каждая страница содержит 30 строк по 70 символов в строке. Какой объём информации содержат 5 страниц текста?

Решение

$256 = 2^8$ $I = 8$ бит – информационный вес одного символа.

Подсчитаем количество символов в тексте

$30 \times 70 \times 5 = 10500$ символов в тексте.

Найдём объём всей информации в тексте

$10500 \times 8 = 84000$ бит = 10500 байт = $10,25$ Кбайт.

Формула Шеннона

Существует множество ситуаций, когда возможные события имеют различные вероятности реализации.

$$I = - \sum_{i=1}^N p_i \log_2 p_i$$

где I – количество информации,
 N – количество возможных событий,
 p_i – вероятности отдельных событий.

Для частного, но широко распространённого случая, когда события равновероятны, величину количества информации I можно рассчитать по формуле

$$I = \log_2 N$$

Двоичное кодирование графической информации в компьютере

- **Пиксель** – минимальный участок изображения, которому независимым образом можно задать цвет.
- **Разрешающая способность** растрового изображения определяется количеством точек по горизонтали и вертикали на единицу длины изображения.
- Количество информации, которое используется при кодировании цвета точек изображения, называется **глубиной цвета**.

Задача 4

- Укажите минимальный объём памяти (в килобайтах) достаточный для хранения любого растрового изображения размером 128×128 пикселей, если известно, что в изображении используется палитра на 256 цветов.
- **Варианты ответа:**
1) 256 2) 8 3) 16 4) 128

Решение

$N = 2^l$ $256 = 2^8$ $l = 8$ – глубина цвета.

$128 \times 128 = 16384$ – количество точек (пикселей) в изображении.

$16384 \times 8 = 131072$ бит = 16384 байта = 16 Кбайт.

Двоичное кодирование звуковой информации в компьютере

- Частота дискретизации звука – это количество измерений громкости звука за одну секунду.
- Глубина кодирования звука – это количество информации, которое необходимо для кодирования дискретных уровней громкости цифрового звука.

Задача 5

- Оцените информационный объём высококачественного стереоаудиофайла длительностью звучания 1 минута, если «глубина» кодирования 16 бит, а частота дискретизации 48 кГц.

Решение

Информационный объём звукового файла длительностью в 1 секунду равен:

$$16 \text{ бит} \times 48000 \text{ Гц} \times 2 = 1536000 \text{ бит} = 96000 \text{ байт} = 187,5 \text{ Кбайт.}$$

Информационный объём звукового файла длительностью в 1 минуту равен:

$$187,5 \text{ Кбайт} \times 60 \text{ с} = 11250 \text{ Кбайт} = 11 \text{ Мбайт}$$