

Кодирование информации

Кодирование текстовой информации. Кодировка ASCII.
Основные кодировки кириллицы.

Что нужно знать:

- все символы кодируются одинаковым числом бит (алфавитный подход)
- чаще всего используют кодировки, в которых на символ отводится 8 бит (8-битные) или 16 бит (16-битные)
- 1 байт = 8 бит, 1 килобайт = 1024 байта, 1 мегабайт = 1024 Кбайта
- после знака препинания внутри (не в конце!) текста ставится пробел
- чтобы найти информационный объем текста, нужно умножить количество символов на число бит на символ.

Пример задания:

Определите информационный объем текста

Бамбарбия! Кергуду!

1) 38 бит

2) 144 бита

3) 152 бита

4) 19 бит

Ответ: 3

Пример задания:

Автоматическое устройство осуществило перекодировку информационного сообщения на русском языке, первоначально записанного в 16-битном коде Unicode, в 8-битную кодировку КОИ-8. При этом информационное сообщение уменьшилось на 480 бит. Какова длина сообщения в символах?

- 1) 30 2) 60 3) 120 4) 480

Ответ: 2

Еще пример задания:

В таблице ниже представлена часть кодовой таблицы ASCII:

Символ	1	5	A	B	Q	a	b
Десятичный код	49	53	65	66	81	97	98
Шестнадцатеричный код	31	35	41	42	51	61	62

Каков шестнадцатеричный код символа «q» ?

- 1) 71_{16} 2) 83_{16} 3) $A1_{16}$ 4) $B3_{16}$

Ответ: 1

Тема: Кодирование и декодирование информации.

Что нужно знать:

- кодирование – это перевод информации с одного языка на другой (запись в другой системе символов, в другом алфавите)
- кодирование может быть *равномерное* и *неравномерное*;
при равномерном кодировании все символы кодируются кодами равной длины;
при неравномерном кодировании разные символы могут кодироваться кодами разной длины, это затрудняет декодирование

Пример задания:

Для кодирования букв А, Б, В, Г решили использовать двухразрядные последовательные двоичные числа (от 00 до 11, соответственно). Если таким способом закодировать последовательность символов БАВГ и записать результат шестнадцатеричным кодом, то получится

- 1) $4B_{16}$ 2) 411_{16} 3) $BACD_{16}$ 4) 1023_{16}

Ответ: 1

Пример задания:

Для 5 букв латинского алфавита заданы их двоичные коды (для некоторых букв – из двух бит, для некоторых – из трех). Эти коды представлены в таблице:

A - 000

B - 01

C - 100

D - 10

E - 011

*Определить, какой набор букв закодирован двоичной строкой
0110100011000*

- 1) EBCBA 2) BDDEA 3) BDCEA 4) EBAEA

Ответ: 3

Пример задания:

Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный по длине код: А=0, Б=10, В=110. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?

- 1) 1 2) 1110 3) 111 4) 11

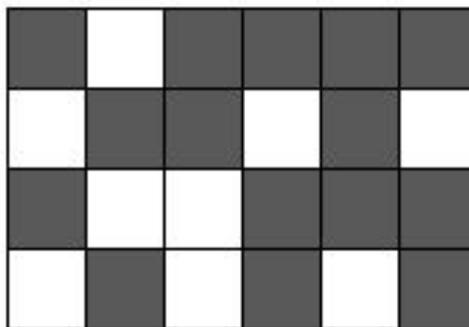
Решение:

1) для того, чтобы сообщение, записанное с помощью неравномерного по длине кода, однозначно декодировалось, требуется, чтобы *никакой код не был началом другого* (более длинного) кода; это условие называют *условием Фано*

Ответ: 3

Пример задания:

Черно-белое растровое изображение кодируется построчно, начиная с левого верхнего угла и заканчивая в правом нижнем углу. При кодировании 1 обозначает черный цвет, а 0 – белый.



Для компактности результат записали в шестнадцатеричной системе счисления. Выберите правильную запись кода.

- 1) BD9AA5 2) BDA9B5 3) BDA9D5 4) DB9DAB

Ответ: 3

Пример задания:

*Для передачи чисел по каналу с помехами используется код проверки четности. Каждая его цифра записывается в двоичном представлении, с добавлением ведущих нулей до длины 4, и к получившейся последовательности дописывается сумма её элементов по модулю 2 (например, если передаём **23**, то получим последовательность **0010100110**). Определите, какое число передавалось по каналу в виде **01010100100111100011**?*

- 1) 59143 2) 5971 3) 102153 4) 10273

Ответ: 2

Тема: Кодирование и обработка графической информации.

Что нужно знать:

- графическая информация может храниться в растровом и векторном форматах
- *векторное* изображение – это набор геометрических фигур, которые можно описать математическими зависимостями;
- *растровое* изображение хранится в виде набора пикселей, для каждого из которых задается свой цвет, независимо от других
- *глубина цвета* – это количество бит на пиксель (обычно от 1 до 24 бит на пиксель)

- в режиме истинного цвета (*True Color*) информация о цвете каждого пикселя растрового изображения хранится в виде набора его RGB-составляющих (*Red, Green, Blue*); каждая из RGB-составляющих – целое число (яркость) в интервале $[0,255]$ (всего 256 вариантов), занимающее в памяти 1 байт или 8 бит (так как $2^8 = 256$); таким образом, на каждый пиксель отводится 3 байта = 24 бита памяти (глубина цвета – 24 бита); нулевое значение какой-то составляющей означает, что ее нет в этом цвете, значение 255 – максимальная яркость; в режиме истинного цвета можно закодировать $256^3 = 2^{24} = 16\,777\,216$ различных цветов

- *палитра* – это ограниченный набор цветов, которые используются в изображении (обычно не более 256); при кодировании с палитрой выбираются N любых цветов (из полного набора 16 777 216 цветов), для каждого из них определяется RGB-код и уникальный номер от 0 до $N-1$; тогда информация о цвете пикселя – это *номер* его цвета в палитре; при кодировании с палитрой количество бит на 1 пиксель (K) зависит от количества цветов в палитре N , они связаны формулой: $N = 2^k$; объем памяти на все изображение вычисляется по формуле $M = Q * k$, где k – число бит на пиксель, а Q – общее количество пикселей

· цвет на Web-страницах кодируется в виде RGB-кода в шестнадцатеричной системе: #RRGGBB, где RR, GG и BB – яркости красного, зеленого и синего, записанные в виде двух шестнадцатеричных цифр; это позволяет закодировать 256 значений от 0 (00_{16}) до 255 (FF_{16}) для каждой составляющей; коды некоторых цветов:

#FFFFFF – белый, #000000 – черный,

#CCCCCC и любой цвет, где $R = G = B$, – это серый разных яркостей

#FF0000 – красный, #00FF00 – зеленый, #0000FF – синий,

#FFFF00 – желтый, #FF00FF – фиолетовый, #00FFFF – цвет морской волны

- чтобы получить **светлый оттенок** какого-то «чистого» цвета, нужно одинаково увеличить нулевые составляющие; например, чтобы получить светло-красный цвет, нужно сделать максимальной красную составляющую и, кроме этого, одинаково увеличить остальные – синюю и зеленую: #FF9999 (сравните с красным – #FF0000)
- чтобы получить **темный оттенок** чистого цвета, нужно одинаково уменьшить все составляющие, например, #660066 – это темно-фиолетовый цвет (сравните с фиолетовым #FF00FF)

Пример задания:

Для хранения растрового изображения размером 32×32 пикселя отвели 512 байтов памяти. Каково максимально возможное число цветов в палитре изображения?

- 1) 256 2) 2 3) 16 4) 4

Ответ: 3

Пример задания:

Для кодирования цвета фона страницы Интернет используется атрибут `bgcolor="#XXXXXX"`, где в кавычках задаются шестнадцатеричные значения интенсивности цветовых компонент в 24-битной RGB-модели. Какой цвет будет у страницы, заданной тэгом `<body bgcolor="#FFFFFF">`?

- 1) белый 2) зеленый 3) красный 4) синий

Ответ: 1

Тема: Вычисление информационного объема сообщения

с помощью K бит можно закодировать 2^K различных вариантов

чтобы найти информационный объем сообщения I , нужно умножить количество символов N на число бит на символ

$$K: \quad I=K*N$$

мощность алфавита M – это количество символов в этом алфавите

если алфавит имеет мощность M , то количество всех возможных «слов» длиной N равно M^N ; для двоичного кодирования (мощность алфавита $M = 2$ символа) получаем формулу: 2^N

Пример задания:

В велокроссе участвуют 119 спортсменов. Специальное устройство регистрирует прохождение каждым из участников промежуточного финиша, записывая его номер с использованием минимально возможного количества бит, одинакового для каждого спортсмена. Каков информационный объем сообщения, записанного устройством, после того как промежуточный финиш прошли 70 велосипедистов?

1) 70 бит

2) 70 байт

3) 490 бит

4) 119 байт

Ответ: 3.

Пример задания:

Объем сообщения, содержащего 4096 символов, равен 1/512 части Мбайта. Какова мощность алфавита, с помощью которого записано это сообщение?

1) 8 2) 16 3) 4096 4) 16384

Ответ: 2

Пример задания:

В зоопарке 32 обезьяны живут в двух вольерах, А и Б. Одна из обезьян – альбинос (вся белая). Сообщение «Обезьяна-альбинос живет в вольере А» содержит 4 бита информации. Сколько обезьян живут в вольере Б?

1) 4 2) 16 3) 28 4) 30

Ответ: 4

Пример задания:

*В корзине лежат 32 клубка шерсти, из них 4 красных.
Сколько бит информации несет сообщение о том, что
достали клубок красной шерсти?*

1) 2 2) 3 3) 4 4) 32

Ответ: 2

Пример задания:

В некоторой стране автомобильный номер длиной 7 символов составляется из заглавных букв (всего используется 26 букв) и десятичных цифр в любом порядке. Каждый символ кодируется одинаковым и минимально возможным количеством бит, а каждый номер – одинаковым и минимально возможным количеством байт. Определите объем памяти, необходимый для хранения 20 автомобильных номеров.

1) 20 байт 2) 105 байт 3) 120 байт 4) 140 байт

Ответ: 3

Пример задания:

Какое наименьшее число символов должно быть в алфавите, чтобы при помощи всевозможных трехбуквенных слов, состоящих из символов данного алфавита, можно было передать не менее 9 различных сообщений?

1) 1 2) 2 3) 3 4) 4

Ответ: 3

Пример задания:

Каждая ячейка памяти компьютера, работающего в троичной системе счисления, может принимать три различных значения (-1, 0, 1). Для хранения некоторой величины отвели 4 ячейки памяти. Сколько различных значений может принимать эта величина?

Ответ: 81

Пример задания:

В школьной базе данных хранятся записи, содержащие информацию об учениках:

<Фамилия> – 16 символов: русские буквы (первая прописная, остальные строчные),

<Имя> – 12 символов: русские буквы (первая прописная, остальные строчные),

<Отчество> – 16 символов: русские буквы (первая прописная, остальные строчные),

<Год рождения> – числа от 1992 до 2003.

Каждое поле записывается с использованием минимально возможного количества бит. Определите минимальное количество байт, необходимое для кодирования одной записи, если буквы е и ё считаются совпадающими.

1) 28 2) 29 3) 46 4) 56

Ответ: 1