

ПОСИМВОЛЬНОЕ КОДИРОВАНИЕ (№13)

ОСНОВНЫЕ ФОРМУЛЫ

-

$$N = 2^i$$

$$I = i \cdot k$$

ЗАДАНИЕ №1

В некоторой стране автомобильный номер длиной 6 символов составляют из заглавных букв (используются только 33 различных буквы) и десятичных цифр в любом порядке. Каждый такой номер в компьютерной программе записывается минимально возможным и одинаковым целым количеством байтов (при этом используют посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством битов). Определите объём памяти, отводимый этой программой для записи 125 номеров. (Ответ дайте в байтах.)

РЕШЕНИЕ №1

Согласно условию, в номере могут быть использованы 10 цифр (0..9) и 33 буквы, всего $10 + 33 = 43$ символов. Известно, что с помощью N бит можно закодировать 2^N различных вариантов. Поскольку $2^5 < 43 < 2^6$, то для записи каждого из 43 символов необходимо 6 бит.

Для хранения всех 6 символов номера нужно $6 * 6 = 36$ бит, а т. к. для записи используется целое число байт, то берём ближайшее не меньшее значение, кратное восьми, это число $40 = 5 * 8$ бит (5байт).

Тогда 125 номеров занимают $5 * 125 = 625$ байт.

ЗАДАНИЕ №2

При регистрации в компьютерной системе каждому пользователю выдаётся пароль, состоящий из 11 символов и содержащий только символы А, Б, В, Г, Д, Е. Каждый такой пароль в компьютерной программе записывается минимально возможным и одинаковым целым количеством байт, при этом используют посимвольное кодирование и все символы кодируются одинаковым и минимально возможным количеством бит. Определите, сколько байт необходимо для хранения 20 паролей.

РЕШЕНИЕ №2

Согласно условию, в пароле могут быть использованы 6 символов. Известно, что с помощью N бит можно закодировать 2^N различных вариантов. Поскольку $2^2 < 6 < 2^3$, то для записи каждого из 6 символов необходимо 3 бита.

Для хранения всех 11 символов пароля нужно $3 \cdot 11 = 33$ бита, а т. к. для записи используется целое число байт, то берём ближайшее не меньшее значение, кратное восьми, это число $40 = 5 \cdot 8$ бит = 5 байт.

Тогда для записи двадцати паролей необходимо $5 \cdot 20 = 100$ байт.

САМОСТОЯТЕЛЬНОЕ ЗАДАНИЕ

При регистрации в компьютерной системе каждому пользователю выдаётся пароль, состоящий из 15 символов и содержащий только символы из 12-символьного набора: А, В, С, D, E, F, G, H, K, L, M, N. В базе данных для хранения сведений о каждом пользователе отведено одинаковое и минимально возможное целое число байт. При этом используют посимвольное кодирование паролей, все символы кодируют одинаковым и минимально возможным количеством бит. Кроме собственно пароля, для каждого пользователя в системе хранятся дополнительные сведения, для чего отведено 12 байт на одного пользователя.

Определите объём памяти (в байтах), необходимый для хранения сведений о 50 пользователях. В ответе запишите только целое число — количество байт.

РЕШЕНИЕ

На кодирование одного символа из 12-буквенного алфавита требуется 4 бита. Тогда на один пароль необходимо бит. Минимальное количество байт, вмещающее 60 бит — 8. Итого на одного пользователя необходимо байт. А на 50 пользователей нужно байт.

**ПОДСЧЕТ
ПРОМЕЖУТОЧНОГ
О КОЛИЧЕСТВА
ИНФОРМАЦИИ
(№13)**

ЗАДАНИЕ №1

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?

РЕШЕНИЕ №1

1 символ в коде Unicode кодируется 16-ю битами, 1 символ в коде КОИ-8 — 8-ю битами. Количество символов при перекодировке не меняется, поэтому обозначим его за x .

Составим уравнение: $16x - 8x = 480$

Решая его, получим $x = 60$.

ЗАДАНИЕ №2

В скачках участвуют 20 лошадей. Специальное устройство регистрирует прохождение каждой лошадью финиша, записывая ее номер с использованием минимально возможного количества бит, одинакового для каждой лошади. Каков информационный объем сообщения, записанного устройством, если до финиша добрались только 15 из 20 участвовавших в скачках лошадей? (Ответ дайте в битах.)

РЕШЕНИЕ №2

Известно, что с помощью N бит можно закодировать 2^N различных чисел. Поскольку $2^4 < 20 < 2^5$, то для записи каждого из 20 номеров необходимо 5 бит памяти. Поскольку до финиша добрались только 15, то информационный объем сообщения составит $15 \cdot 5 = 75$ бит.

**КОДИРОВАНИЕ
ЗВУКА, ГРАФИКИ
(№9)**

ОСНОВНЫЕ ФОРМУЛЫ

- $$I_3 = R \cdot D \cdot i \cdot t$$

R – тип файла (моно/стерео/квадро)

D – частота дискретизации

i – глубина кодирования

t – длительность звуковой дорожки

ЗАДАНИЕ №1

Производилась четырехканальная (квадро) звукозапись с частотой дискретизации 24 кГц и 16-битным разрешением. В результате был получен файл размером 1800 Мбайт, сжатие данных не производилось. Определите приблизительно, сколько минут производилась запись.

В качестве ответа укажите ближайшее к времени записи целое число минут.

РЕШЕНИЕ №1

- $1800 \text{ (Мб)} = 4 \cdot 24 \text{ (кГц)} \cdot 16 \text{ (бит)} \cdot t \text{ (с)}$

$$1800 \cdot 2^{20} = 4 \cdot 24000 \cdot 2 \cdot t$$

$$75 \cdot 2^{20} = 4 \cdot 1000 \cdot 2 \cdot t$$

$$75 \cdot 2^{17} = 1000 \cdot t$$

$$3 \cdot 2^{17} = 40 \cdot t$$

$$3 \cdot 2^{14} = 5 \cdot t \text{ (с)}$$

$$t \text{ (м)} = \frac{3 \cdot 2^{14}}{60 \cdot 5} = \frac{2^{14}}{20 \cdot 5} = \frac{2^{12}}{25} = \frac{4096}{25} = 163,84$$

ОТВЕТ: 164

ЗАДАНИЕ №2

Производится двухканальная (стерео) звукозапись с частотой дискретизации 32 кГц и 32-битным разрешением. Запись длится 3 минуты, её результаты записываются в файл, сжатие данных не производится.

Определите приблизительно размер полученного файла (в Мбайт). В качестве ответа укажите ближайшее к размеру файла целое число, кратное пяти.

РЕШЕНИЕ №2

•

$$I_3 = \frac{2 \cdot 32000 \cdot 32 \cdot 180}{2^{23}}$$

$$I_3 = \frac{2 \cdot 2^5 \cdot 1000 \cdot 2^5 \cdot 2^2 \cdot 45}{2^{23}}$$

$$I_3 = \frac{1000 \cdot 45}{2^{10}} = \frac{125 \cdot 45}{128} = \frac{5625}{128}$$

$$I_3 = 43,94$$

ОТВЕТ: 45

ЗАДАНИЕ №3

Какой минимальный объём памяти (в Кбайт) нужно зарезервировать, чтобы можно было сохранить любое растровое изображение размером 128×128 пикселей при условии, что в изображении могут использоваться 256 различных цветов? В ответе запишите только целое число, единицу измерения писать не нужно.

РЕШЕНИЕ №3

Один пиксель кодируется 8 битами памяти, так как $2^8 = 256$.

Всего $128 * 128 = 2^7 \cdot 2^7 = 2^{14}$ пикселей.

Тогда объем памяти, занимаемый изображением $2^{14} * 8 = 2^{17}$ бит = 2^{14} байт = 2^4 Кбайт = 16 Кбайт.

ЗАДАНИЕ №4

Автоматическая фотокамера производит растровые изображения размером 300 на 200 пикселей. При этом объём файла с изображением не может превышать 30 Кбайт, упаковка данных не производится. Какое максимальное количество цветов можно использовать в палитре?

РЕШЕНИЕ №4

Объём растрового изображения находится как произведение количества пикселей в изображении на объём памяти x , необходимый для хранения цвета одного пиксела:

$$300 \cdot 200 \cdot x < 30 \cdot 2^{13} \text{ бит, откуда } x < 4,096 \text{ бит}$$

$= 4 \text{ бит}$. Значит, в изображении можно использовать не более $2^4 = 16$ цветов.

Ответ: 16.

ЗАДАНИЕ №5

Документ объёмом 80 Мбайт можно передать с одного компьютера на другой двумя способами.

А. Сжать архиватором, передать архив по каналу связи, распаковать.

Б. Передать по каналу связи без использования архиватора.

Какой способ быстрее и на сколько, если

— скорость передачи данных по каналу связи составляет 2^{25} бит в секунду;

— объём сжатого архиватором документа равен 35% исходного;

— время, требуемое на сжатие документа, – 15 секунд, на распаковку — 3 секунды?

РЕШЕНИЕ №5

Способ А. Общее время складывается из времени сжатия, распаковки и передачи. Время передачи t рассчитывается по формуле $t = Q / q$, где Q — объём информации, q — скорость передачи данных.

Найдём сжатый объём: $80 * 0,35 = 28$ Мбайт

Переведём Q из Мбайт в биты: $28 \text{ Мбайт} = 28 * 2^{20} \text{ байт} = 28 * 2^{23} \text{ бит}$.

Найдём общее время: $t = 15 \text{ с} + 3 \text{ с} + 28 * 2^{23} \text{ бит} / 2^{25} \text{ бит/с} = 18 + 7 \text{ с} = 25 \text{ с}$.

Способ Б. Общее время совпадает с временем передачи: $t = 80 * 2^{23} \text{ бит} / 2^{25} \text{ бит/с} = 20 \text{ с}$.

Видно, что способ Б быстрее на $25 - 20 = 5 \text{ с}$.

Ответ: Б5.

ЗАДАНИЕ №6

У Толи есть доступ к сети Интернет по высокоскоростному одностороннему радиоканалу, обеспечивающему скорость получения информации бит в секунду. У Миши нет скоростного доступа в Интернет, но есть возможность получать информацию от Толи по низкоскоростному телефонному каналу со средней скоростью бит в секунду. Миша договорился с Толей, что тот будет скачивать для него данные объемом 5 Мбайт по высокоскоростному каналу и ретранслировать их Мише по низкоскоростному каналу.

Компьютер Толи может начать ретрансляцию данных не раньше, чем им будут получены первые 512 Кбайт этих данных. Каков минимально возможный промежуток времени (в секундах) с момента начала скачивания Толей данных до полного их получения Мишей?

РЕШЕНИЕ №6

Нужно определить, сколько времени будет передаваться файл объемом 5 Мбайт по каналу со скоростью передачи данные 2^{15} бит/с; к этому времени нужно добавить задержку файла у Толи (пока он не получит 512 Кбайт данных по каналу со скоростью 2^{19} бит/с).

Переведём объём информации в Мб в биты: $Q = 5$ Мб = $5 * 2^{20}$ байт = $5 * 2^{23}$ бит.

Время задержки: = $512 \text{ кб} / 2^{19} \text{ бит/с} = 2^{(9 + 10 + 3) - 19} \text{ с} = 2^3 \text{ с}$.

Время скачивания данных Мишей: = $5 * 2^{23} \text{ бит} / 2^{15} \text{ бит/с} = 5 * 2^8 \text{ с}$.

Полное время: = $5 * 2^8 \text{ с} + 2^3 \text{ с} = (256 * 5 + 8) \text{ с} = 1288 \text{ с}$.

Ответ: 1288.

НЕРАВНОМЕРН ОЕ КОДИРОВАНИЕ

ЗАДАНИЕ №1

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	b	c	d	e
000	110	01	001	10

Какой набор букв закодирован двоичной строкой 1100000100110?

РЕШЕНИЕ №1

Мы видим, что выполняется условие Фано: никакое кодовое слово не является началом другого кодового слова, поэтому однозначно можем раскодировать сообщение с начала.

Разобьём код слева направо по данным таблицы и переведём его в буквы:

110 000 01 001 10 — b a c d e.

ЗАДАНИЕ №2

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г и Д, используется неравномерный двоичный код, позволяющий однозначно декодировать полученную двоичную последовательность. Вот этот код: А — 1; Б — 0100; В — 000; Г — 011; Д — 0101. Требуется сократить для одной из букв длину кодового слова так, чтобы код по-прежнему можно было декодировать однозначно. Коды остальных букв меняться не должны. Каким из указанных способов это можно сделать?

- 1) для буквы Г — 11
- 2) для буквы В — 00
- 3) для буквы Г — 01
- 4) это невозможно

РЕШЕНИЕ №2

Для однозначного декодирования получившееся в результате сокращения кодовое слово не должно быть началом никакого другого. Первый вариант ответа не подходит, поскольку код буквы А является началом кода буквы Г. Второй вариант ответа подходит. Третий вариант ответа не подходит, т. к. в таком случае код буквы Г является началом кода буквы Д.

Правильный ответ указан под номером: 2.

ЗАДАНИЕ №3

Для кодирования некоторой последовательности, состоящей из букв И, К, Л, М, Н, решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы Н использовали кодовое слово 0, для буквы К – кодовое слово 10. Какова наименьшая возможная суммарная длина всех пяти кодовых слов?

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

РЕШЕНИЕ №3

Нельзя использовать кодовые слова, которые начинаются с 0 или с 10. 11 также не можем использовать, поскольку тогда мы больше не сможем взять никакое другое кодовое слово, а нам их нужно пять. Поэтому берём трёхзначное 110. 111 опять же не можем использовать, потому что понадобится ещё одно кодовое слово, а вместе с этим не останется больше свободных. Теперь осталось взять всего два слова и это будут 1110 и 1111. Итого имеем 0, 10, 110, 1110 и 1111 — 14 символов.

САМОСТОЯТЕЛЬНОЕ ЗАДАНИЕ

По каналу связи передаются сообщения, содержащие только четыре буквы: П, О, С, Т; для передачи используется двоичный код, допускающий однозначное декодирование. Для букв Т, О, П используются такие кодовые слова: Т: 111, О: 0, П: 100.

Укажите кратчайшее кодовое слово для буквы С, при котором код будет допускать однозначное декодирование. Если таких кодов несколько, укажите код с наименьшим числовым значением.

САМОСТОЯТЕЛЬНОЕ ЗАДАНИЕ

Буква С не может кодироваться строкой, которая начинается с 0, поскольку О имеет код 0.

Буква С не может кодироваться как 1, так как кодирование буквы Т начинается с 1.

Буква С не может кодироваться как 10, так как кодирование буквы П начинается с 10.

Буква С не может кодироваться как 11, так как кодирование буквы Т начинается с 11.

Буква С может кодироваться как 101 – это наименьшее возможное значение.

Ответ: 101.