

## **Основной теоретический материал по теме «Информация и ее кодирование», необходимый для решения задач.**

**Код** – это система условных знаков для представления информации.

**Кодирование** – операция преобразования символов или группы символов одного кода в символы или группы символов другого кода. Обратное преобразование – **декодирование**.

### **Алфавитный подход к измерению информации.**

Множество используемых символов – алфавит, размер алфавита – его мощность.

*Обозначения:*

Мощность алфавита обозначается через  $N$ .

Информационный объем одного символа –  $I$ .

### **Алгоритм нахождения количества информации с точки зрения алфавитного подхода.**

1. Найти мощность алфавита  $N$ .
2. Найти информационный объем одного символа –  $I = \log_2 N$ .
3. Найти количество символов в сообщении –  $K$ .
4. Найти информационный объем всего сообщения –  $K \cdot I$ .

### **Единицы измерения информации.**

1 байт = 8 бит

1 килобайт = 1024 байт

1 мегабайт = 1024 килобайт

1 гигабайт = 1024 мегабайт

**Задача №1.** Считая, что каждый символ кодируется одним байтом, оцените информационный объем следующего предложения из пушкинских строк:

*Певец Давид был ростом мал, Но повалил же Голиаф!*

1) 400 битов; 2) 50 битов; в) 400 байтов; г) 5 байтов.

Решение:

Поскольку в тексте содержится 50 символов (считая все пробелы и знаки препинания), а каждый символ кодируется одним байтом, то получаем  $50 \text{ символов} * 1 \text{ байт} = 50 \text{ байт}$ . Пользуясь формулой  $1 \text{ байт} = 8 \text{ бит}$ , получаем, что сообщение кодируется 400 битами.

**Ответ №1.**

**Задача №2.** Сколько существует различных последовательностей из символов «плюс» и «минус», длиной ровно в 5 символов?

1) 64; 2) 50; 3) 32; 4) 20.

Решение.

Различных комбинаций из символов «плюс» и «минус» существует ровно столько же, сколько и соответствующих двоичных кодов той же длины, то есть  $2^5 = 32$ . **Ответ №3.**

**Задача №3.** Обычный дорожный светофор без дополнительных секций подает шесть видов сигналов (непрерывные красный, желтый и зеленый, мигающий желтый, мигающий зеленый, мигающие красный и желтый одновременно). Электронное устройство управления светофором последовательно воспроизводит записанные сигналы. Подряд записано 100 сигналов светофора. В байтах данный информационный объем составляет: 1) 37; 2) 38; 3) 50; 4) 100.

Решение.

Для кодирования шести различных состояний достаточно 3-х битов (при этом две комбинации даже остаются не востребованными). К этому выводу можно прийти, воспользовавшись формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита.

В нашем случае  $N=6$ , т.к. светофор подает 6 различных сигналов.

$I = \log_2 6 \approx 3$ .

Таким образом, 100 сигналов кодируется 300 битами. Делим это число на 8 (1 байт = 8 бит) и округляем в большую сторону (дробных байтов не бывает).

Получаем 38

байтов. **Ответ №2.**

**Задача №6.** Для букв латинского алфавита заданы их двоичные коды (для некоторых букв из двух битов, для некоторых – из трех). Эти коды представлены в таблице:

A	B	C	D	E
000	01	100	10	011

Определите, какой набор букв закодирован двоичной строкой 0110100011000.

1) EBCEA; 2) BDDEA; 3) BDCEA; 4) EBAEA.

*Решение.*

При кодировании текста кодом переменной длины правильная комбинация кодов символов однозначна. Выполним разделение комбинации на коды отдельных символов (разбиение целесообразно начать в этом примере с конца цепочки): 01 10 100 011 000.

Таким образом, получаем: BDCEA. Ответ №3.

16.0 | \_\_\_\_\_  
12.0 | \_\_\_\_\_  
- - | \_\_\_\_\_

**Задача №7.** В корзине лежат шары. Все разного цвета. Сообщение о том, что достали синий шар, несет 5 битов информации. Сколько всего шаров в корзине?

1) 5; 2) 10; 3) 16; 4) 32.

*Решение.*

Вспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита. В нашем примере известно  $I$  – информационный объем сообщения о том, что достали шар синего цвета,  $I = 5$  бит. Найдем  $N$  – мощность алфавита, то есть количество шаров в корзине.  $N = 2^i = 2^5 = 32$  бит. **Ответ №4.**

**Задача №8.** В лотерее разыгрывается 64 шара. Выигрышная комбинация состоит из  $X$  шаров, и сообщение о ней несет 42 бита информации. Чему равно  $X$ ?

1) 7; 2) 2; 3) 42; 4) 64.

*Решение.*

В данной задаче идет речь о лотерее, в которой из 64 шаров вытягивается какое-то количество шаров, которые являются выигрышной комбинацией. Известно, что сообщение о выигрышной комбинации шаров несет 42 бита. Необходимо определить количество шаров в выигрышной комбинации.

Вспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита. Мощность алфавита по условию равна 64 (шара).

Найдем информационный объем сообщения о вытягивании 1 шара:

$I = \log_2 N = \log_2 64 = 6$  бит. (подобрать в какую степерь нужно возвести двоку, чтобы получить 64):  $2^6 = 64 \dots$

6 бит – информационный объем сообщения о вытягивании 1 шара, а информационный объем о вытягивании  $X$  шаров – 42 бит. Нетрудно догадаться, что количество шаров определяется как  $42/6 = 7$  шаров. **Ответ №1.**

**Задание 1.** Исследователь наблюдает изменение параметра, который может принимать одно из семи значений. Значения записываются при помощи минимального количества бит. Исследователь зафиксировал 120 значений.

Определите информационный объем результатов наблюдения.

**Решение.**

Нам известно максимальное количество значений, которые требуется закодировать с помощью одинаково количества знаков алфавита. Это семь. В качестве алфавита используется бит, который может принимать всего два значения (0 и 1).

Для определения минимального количества бит, необходимых для кодирования одного значения, воспользуемся **формулой Хартли:  $k = \log_2 N$** . Здесь  $k$  – это количество бит, а  $N$  – кодируемое количество значений.

В нашем случае:  $k = \log_2 7$ . Другими словами, в какую степень нужно возвести двойку, чтобы получить семерку? Мы знаем, что  $2^2 = 4$ , а  $2^3 = 8$ . Следовательно, значение  $k$  находится между 2 и 3 и является дробью. Но количество бит не может быть дробным числом. Поэтому в данном случае, для кодирования одного значения требуется 3 бита.

Поскольку исследователь зафиксировал 120 значений, то общий информационный объем наблюдения равен ( $3 * 120 =$ ) 360 битам или ( $360 / 8 =$ ) 45 байтам.

**Ответ.** Информационный объем 120 наблюдений, принимаемых семь различных значений, равен 45 байтам.

---

**Задание 2.** Если каждый символ кодируется двумя байтами, то каков информационный объем следующего предложения в коде Unicode:

*Сегодня 35 градусов тепла.*

**Решение.**

Посчитаем общее количество символов в предложении с учетом пробелов, цифр и знаков препинания. В данном случае, всего 26 символов.

Каждый символ кодируется двумя байтами. Значит информационный объем предложения равен ( $26 * 2 =$ ) 52 байта или ( $52 * 8 =$ ) 416 бита.

**Ответ.** Информационный объем предложения равен 416 бит.

**Задание 3.** Каждое показание датчика, фиксируемое в памяти компьютера, занимает 20 бит. Записано 54 показания датчика. Каков информационный объем снятых значений в байтах?

**Решение.**

Информационный объем сообщения в битах равен  $(20 * 54 =) 1080$ , что в переводе в байты равно  $(1080 / 8 =) 135$ .

**Ответ.** Информационный объем снятых значений равен 135 байт.

**Задача №9.** Сообщение, записанное буквами 64-символьного алфавита, содержит 20 символов. Какой объем информации оно несет?

1) 64 бита; 2) 20 байтов) 3) 120 битов; 4) 64 байта.

Решение.

Вспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита. По условию мощность алфавита равна 64.

Найдем информационный объем одного символа:  $I = \log_2 N = \log_2 64 = 6$  (бит). Поскольку в задаче говорится о 20 символах, то информационный объем находим как произведение:  $6 * 20 = 120$  бит. **Ответ №3.**

**Задача №10.** Сколько информации несет сообщение о том, что было угадано число в диапазоне целых чисел от 684 до 811?

1) 6 битов; 2) 7 битов; 3) 127 битов; 4) 128 битов.

Решение.

Определим диапазон:  $811 - 684 + 1 = 128$ .

Вспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита. Мощность алфавита равна 128 различным значениям. Определим информационный объем  $I = \log_2 N = I = \log_2 128 = 7$  бит. **Ответ №2.**

**Задача №11.** Объем информационного сообщения 12 288 битов (учитывая, что 1 байт = 8 битов), можно выразить как:

1) 1536 Кбайт; 2) 1,5 Мбайт; 3) 1,5 Кбайт; 4) 1,2 Кбайт.

Решение.

Для решения задачи необходимо воспользоваться таблицей:

1 байт = 8 бит

1 килобайт = 1024 байт

1 мегабайт = 1024 килобайт

1 гигабайт = 1024 мегабайт

$12288 \text{ бит} = 12288 / 8 \text{ байт} = 1536 \text{ байт} = 1536 / 1024 \text{ килобайт} = 1,5 \text{ килобайт}$ . **Ответ №3.**

**Задача №12.** Книга состоит из 64 страниц. На каждой странице 256 символов. Какой

объем информации содержится в книге, если используемый алфавит состоит из 32 символов?

1) 81 920 байт; 2) 40 Кбайт; 3) 16 Кбайт; 4) 10 Кбайт.

Решение.

Воспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита. По условию задачи мощность алфавита равна 32 символам.

Найдем информационную емкость одного символа  $I = \log_2 N = \log_2 32 = 5$  (бит). Определим информационную емкость одной страницы: поскольку на странице 256 символов, то имеем  $256 * 5 = 1280$  (бит).

Определим информационную емкость всей книги:  $64 * 1280 = 81920$  (бит).

Воспользуемся таблицей (для того, чтобы перевести в другие единицы измерения информации):

1 байт = 8 бит

1 килобайт = 1024 байт

1 мегабайт = 1024 килобайт

1 гигабайт = 1024 мегабайт

$81920$  бит =  $81920/8$  байт =  $10240$  байт =  $10240/1024$  килобайт =  $10$  килобайт. **Ответ №4.**

**Задача №13.** В слове **информатика** содержится следующее количество битов (используется система кодировки ASCII):

1) 1; 2) 11; 3) 44; 4) 88.

Решение.

Слово **информатика** состоит из 11 символов, каждый символ в кодировке ASCII кодируется 8 битами. Таким образом, получаем:  $11 * 8 = 88$  бит. **Ответ №4.**

**Задача №15.** В кодировке Unicode каждый символ кодируется 16 битами. Каков информационный объем следующего сообщения?

**2+2=4, а 5+5=10.**

1) 16 битов; 2) 256 битов; 3) 12 байтов; г) 16 байтов.

Решение.

Сообщение **2+2=4, а 5+5=10.** состоит из 16 символов. Информационный объем всего сообщения находим как  $16 * 16 = 256$  бит. **Ответ №2.**

**Задача №16.** Два текста содержат одинаковое количество символов. Первый текст составлен в алфавите мощностью 16 символов. Второй текст в алфавите мощностью 256 символов. Во сколько раз количество информации во втором тексте больше, чем в первом?

1) 2; 2) 4; 3) 8; 4) 4.

Решение.

Воспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита.

Для первого текста  $N = 16$ , отсюда  $I = \log_2 N = \log_2 16 = 4$  (бит).

Для второго текста  $N = 256$ , отсюда  $I = \log_2 N = \log_2 256 = 8$  (бит).

Поскольку оба текста содержат одинаковое количество символов, то во втором тексте количество информации больше, чем в первом в  $8/4 = 2$  раза. **Ответ №1**

**Задача №17.** Какое количество информации содержит сообщение, уменьшающее неопределенность знаний в 8 раз?

1) 1 бит; 2) 2 бита; 3) 3 бита; 4) 4 бита.

Решение.

Количество информации, уменьшающее неопределенность знаний в 8 раз, найдем из уравнения  $8 = 2^x$ , откуда  $x = 3$  бита. Ответ №3.

**Задача №18.** «Вы выходите на следующей остановке?» - спросили человека в автобусе. «Нет», - ответил он. Сколько информации содержит ответ?

Решение.

Воспользуемся формулой  $I = \log_2 N$ , где  $I$  – информационный объем одного символа,  $N$  – мощность алфавита. Человек мог ответить только «да» и «нет», поэтому мощность составляет 2. Откуда информационная емкость сообщения равна  $I = \log_2 N = \log_2 2 = 1$ . Ответ 1 бит.

## Шифрование

**Задание 1.** Для шифрования каждой буквы используются двузначные числа.

Известно, что буква «е» закодирована числом 20.

Среди слов «елка», «полка», «поле», «пока», «кол» есть слова, кодируемые последовательностями цифр:

*11321220, 20121022.*

Выясните код слова «колокол».

**Решение.**

1. Данные последовательности цифр (11321220, 20121022) состоят из восьми символов.
2. По условию задачи каждая буква кодируется двумя символами. Значит, эти последовательности кодируют слова, состоящие из четырех букв.
3. Число 20 кодирует букву «е». В последовательности 11321220 есть число 20 в конце. Из представленного перечня слов подходит только «поле».
4. Отсюда следует, что код «п» - 11, «о» - 32, «л» - 12.
5. В последовательности 20121022 есть число 20 в начале. С буквы «е» начинается только слово «елка».
6. Следовательно, код «л» - 12 (мы это уже знаем), «к» - 10, «а» - 22.
7. Запишем код слова «колокол» - 10|32|12|32|10|32|12.

**Ответ.** Кодом слова «колокол» является комбинация следующих цифр - 10 32 12 32 10 32 12.

**Задание 2.** Для пяти букв 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	010	001	10

Определите, какой набор букв закодирован двоичной строкой *1100000100110*.

### Решение.

1. Ни одна последовательность из двух символов не начинается с *11*, значит первый код *110*. Он соответствует букве «*b*». Будем отбрасывать известную часть кода, оставляя неизвестную - *0000100110*.
2. Ни одна последовательность из двух символов не начинается с *00*, значит следующий код *000*. Это соответствует букве «*a*». Остается *0100110*.
3. Следующие два символа – *01*. Если рассматривать три символа, то *010*, однако такая комбинация ничего не кодирует. Следовательно, мы имеем дело с буквой «*c*». Остаток – *00110*.
4. *001* – это «*d*».
5. *10* – это «*e*».
6. Запишем буквы по порядку: *bacde*.

**Ответ.** Двоичной строкой *1100000100110* закодирован следующий набор букв -*bacde*.

## Информационный объем информации

**Задание 1.** Исследователь наблюдает изменение параметра, который может принимать одно из семи значений. Значения записываются при помощи минимального количества бит. Исследователь зафиксировал 120 значений. Определите информационный объем результатов наблюдения.



## Решение.

Нам известно максимальное количество значений, которые требуется закодировать с помощью одинаково количества знаков алфавита. Это семь. В качестве алфавита используется бит, который может принимать всего два значения (0 и 1).

Для определения минимального количества бит, необходимых для кодирования одного значения, воспользуемся **формулой Хартли:  $k = \log_2 N$** . Здесь  $k$  – это количество бит, а  $N$  – кодируемое количество значений.

В нашем случае:  $k = \log_2 7$ . Другими словами, в какую степень нужно возвести двойку, чтобы получить семерку? Мы знаем, что  $2^2 = 4$ , а  $2^3 = 8$ . Следовательно, значение  $k$  находится между 2 и 3 и является дробью. Но количество бит не может быть дробным числом. Поэтому в данном случае, для кодирования одного значения требуется 3 бита.

Поскольку исследователь зафиксировал 120 значений, то общий информационный объем наблюдения равен ( $3 * 120 =$ ) 360 битам или ( $360 / 8 =$ ) 45 байтам.

**Ответ.** Информационный объем 120 наблюдений, принимаемых семь различных значений, равен 45 байтам.

---

**Задание 2.** Если каждый символ кодируется двумя байтами, то каков информационный объем следующего предложения в коде Unicode:

*Сегодня 35 градусов тепла.*

## Решение.

Посчитаем общее количество символов в предложении с учетом пробелов, цифр и знаков препинания. В данном случае, всего 26 символов.

Каждый символ кодируется двумя байтами. Значит информационный объем предложения равен ( $26 * 2 =$ ) 52 байта или ( $52 * 8 =$ ) 416 бита.

**Ответ.** Информационный объем предложения равен 416 бит.

---

**Задание 3.** Каждое показание датчика, фиксируемое в памяти компьютера, занимает 20 бит. Записано 54 показания датчика. Каков информационный объем снятых значений в байтах?

**Решение.**

Информационный объем сообщения в битах равен  $(20 * 54 =) 1080$ , что в переводе в байты равно  $(1080 / 8 =) 135$ .

**Ответ.** Информационный объем снятых значений равен 135 байт.

# Передача данных

**Задание 1.** Длительность непрерывной передачи данных в сеть Интернет было 12 минут. Определите максимальный размер файла в мегабайтах, который может быть передан за это время, если скорость передачи информации в среднем была 128 килобит/с.

**Решение.**

1. 12 минут – это  $720 (= 12 * 60)$  секунд.
2. 128 килобит/с – это  $128 * 1024$  бит/с
3. Размер файла в битах равен  $720 * 128 * 1024$
4. Размер файла в байтах равен  $(720 * 128 * 1024) / 8 = 90 * 128 * 1024$
5. Размер файла в килобайтах равен  $(90 * 128 * 1024) / 1024 = 90 * 128$
6. Размер файла в мегабайтах равен  $90 * 128 / 1024 = 11,25$

**Ответ.** Размер файла, который можно передать по сети за 12 минут на скорости в 128 килобит/с, составляет 11,25 Мб.

---

**Задание 2.** Сколько времени потребуется сети, работающей со скоростью 56000 бит/с, для передачи 30 страниц текста по 50 строк в 70 символов каждая, при условии, что каждый символ кодируется одним байтом.

**Решение.**

1. Общее количество символов и размер текста в байтах:  $30 * 50 * 70$ .
2. В одном байте 8 бит. Следовательно, размер текста в битах:  $30 * 50 * 70 * 8$
3. Время находится в результате деления количества бит, которые требуется передать, на скорость сети:  $(30 * 50 * 70 * 8) / 56\ 000 = (3 * 5 * 7 * 8) / 56 = 15$

**Ответ.** Для передачи 30 страниц текста по 50 строк, состоящих из 70 символов, со скоростью в 56 000 бит/с потребуется 15 секунд.

---

**Задание 3.** Сколько секунд потребуется модему, передающему сообщения со скоростью 14400 бит/с, чтобы передать сообщение длиной 225 Кбайт.

**Решение.**

1. Выражаем Кбайты в байтах:  $225 * 1024$ .
2. Выражаем байты в битах:  $225 * 1024 * 8$ . Именно такое количество бит потребуется передать модему.
3. Время находится в результате деления количества бит, которые требуется передать, на скорость сети:  $(225 * 1024 * 8) / 14400 = 128$

**Ответ.** Модему потребуется 128 секунд.