

Модуль #3

Кодирование информации

- В данном модуле изучим задания КЕГЭ #4, #7 и #11, содержащие понятия: кодирование и декодирование, передача информации, а также объём информации



Задание #4

Кодирование. Выбор кода.

Кодирование – это представление информации в форме, удобной для её хранения, передачи и обработки. Правило преобразования информации к такому представлению называется кодом.

Кодирование бывает равномерным и неравномерным:

- при равномерном кодировании всем символам соответствуют коды одинаковой длины;
- при неравномерном кодировании разным символам соответствуют коды разной длины, это затрудняет декодирование.

В заданиях этого модуля информация кодируется битами: 0 и 1.

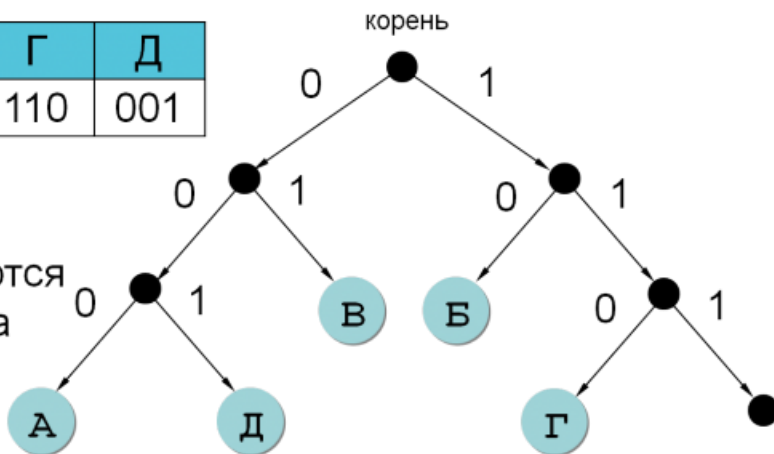
То есть кодовое слово может быть составлено из нулей и единиц.

Условие Фано

Ни одно кодовое слово не должно являться началом другого кодового слова (что обеспечивает однозначное декодирование сообщений с начала).

А	Б	В	Г	Д
000	10	01	110	001

Если все кодовые
слова заканчиваются
на листьях дерева



4

По каналу связи передаются сообщения, содержащие только буквы из набора: А, З, К, Н, Ч. Для передачи используется двоичный код, удовлетворяющий прямому условию Фано, согласно которому никакое кодовое слово не является началом другого кодового слова. Это условие обеспечивает возможность однозначной расшифровки закодированных сообщений. Кодовые слова для некоторых букв известны: Н – 1111, З – 110. Для трёх оставшихся букв А, К и Ч кодовые слова неизвестны. Какое количество двоичных знаков потребуется для кодирования слова КАЗАЧКА, если известно, что оно закодировано **минимально** возможным количеством двоичных знаков?

Ответ: _____.

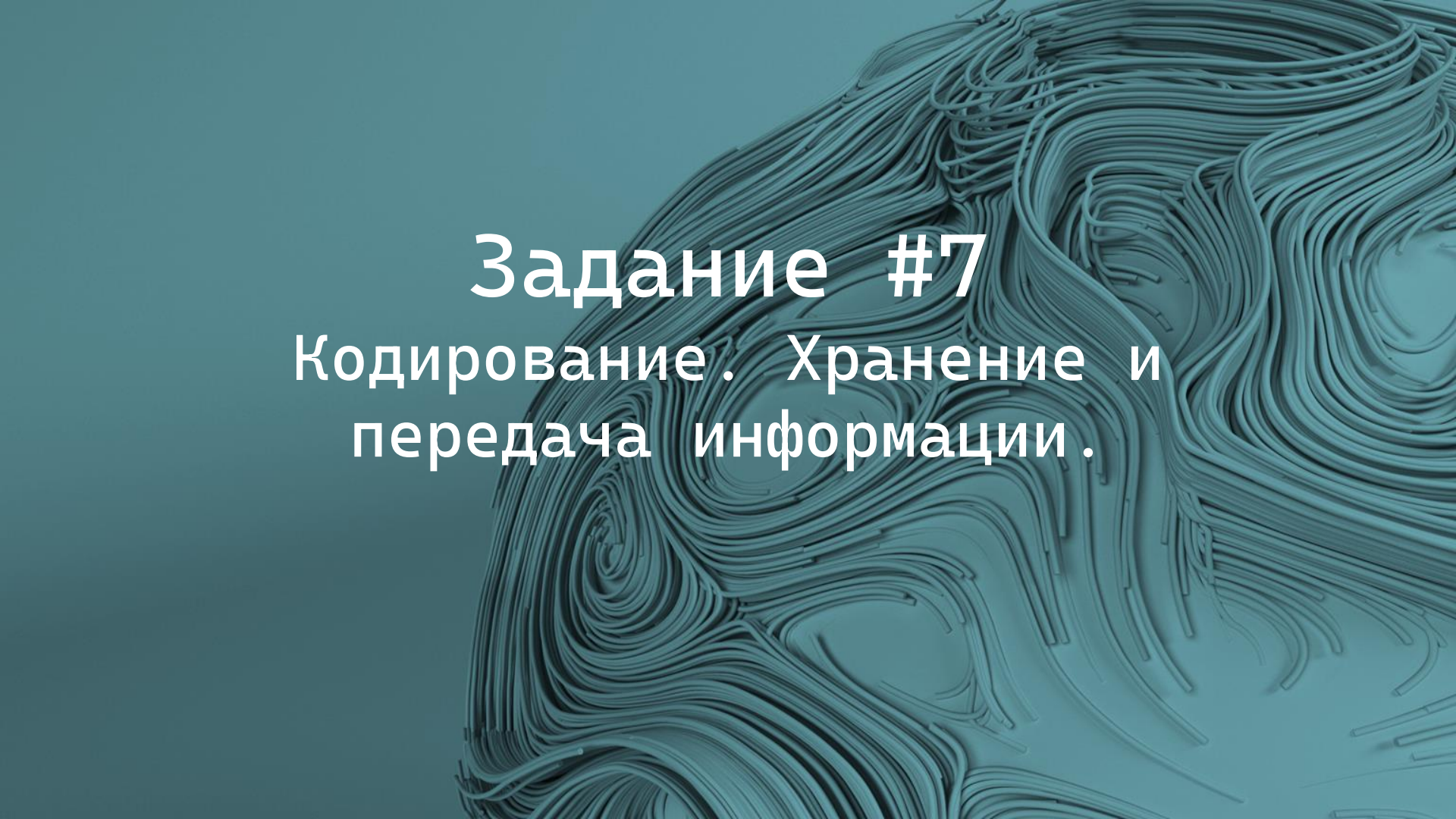
Алгоритм выполнения задания #4

1. Выписать все используемые в наборе буквы и указать коды для некоторых из них из условия (если не оговорено, что кодируется весь алфавит: в таком случае важно просто подобрать коды для всех букв слова и оставить одну ветку дерева свободной)
2. Разобрать слово, которое требуется закодировать: какие буквы повторяются в нём несколько раз: чем больше повторяется буква, тем короче код необходимо подбирать для неё.
3. Расписать дерево, отметив в нём коды, фигурирующие в условии задания.
4. Исходя из пункта 2, выбрать наиболее оптимальные коды для оставшихся букв.

Частые ошибки в задании #4

Если в задании указано, что несколько букв остались без кодовых слов, то кодовое слово для указанной буквы должно быть подобрано таким образом, чтобы осталась возможность найти кодовые слова, удовлетворяющие условию Фано, и для других букв.

Так, например, если мы букву А закодируем нулём, а букву Б единицей, то букву В мы уже никак не сможем закодировать с соблюдением условия Фано, поэтому длину кодового слова для А или Б следует увеличить.



Задание #7

Кодирование. Хранение и
передача информации.

Кодирование текстовой информации

Для определения объёма сообщения I используют формулу:

$$I = n * i$$

Где:

I – объём сообщения

n – количество символов в сообщении (его длина)

i – количество бит, кодирующих один символ

Виды графики

В компьютерной графике изображение может формироваться двумя способами:

- из маленьких квадратов (пикселей) – это растровая графика
- с помощью кривых линий, заданных математической формулой, – это векторная графика

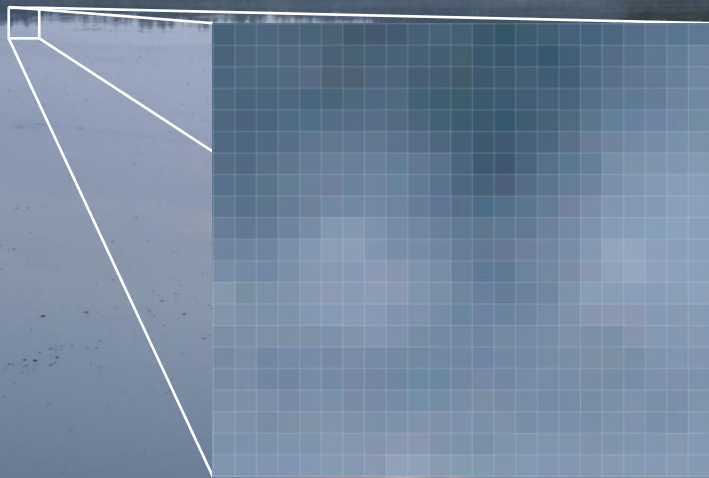
Кодирование графики

Пиксель – это наименьший элемент растрового изображения, который имеет определенный цвет.

Глубина цвета – это количество битов, необходимое для кодирования цвета пикселя.

Если глубина кодирования составляет i битов на пиксель, код каждого пикселя выбирается из 2^i возможных вариантов, поэтому можно использовать не более 2^i различных цветов.

Пиксели в растровом изображении



Количество цветов в палитре

Для определения глубины цвета i используют следующее соотношение между глубиной цвета и количеством цветов изображения:

$$N \leq 2^i$$

Где:

N – количество цветов в палитре

i – количество бит, кодирующих один цвет (глубина цвета)

Кодирование графики

Для определения объёма графического изображения I используют формулу:

$$I = a * b * i$$

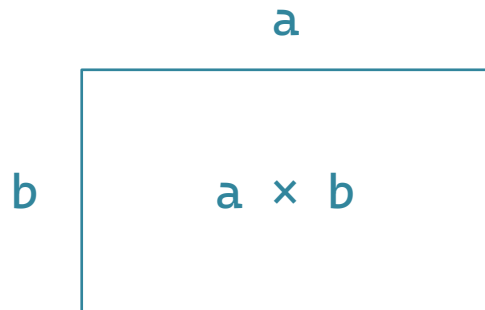
Где:

I – объём изображения

a – количество пикселей в длину изображения

b – количество пикселей в ширину изображения

i – количество бит, кодирующих один цвет



Кодирование звука

Звук сам по себе непрерывен – поэтому естественным образом в нём нельзя выделить какую-то часть, подобную точкам в изображении. Для этого используется дискретизация – это преобразование аналогового сигнала в цифровой код.

Частота дискретизации определяет количество отсчетов, т.е. отдельных значений сигнала, запоминаемых за 1 секунду.

Измеряется в герцах, 1 Гц (один герц) – это один отсчет в секунду, а, например, 7 кГц – это 7000 отсчетов в секунду.

Разрядность кодирования (глубина, разрешение) – это число битов, используемое для хранения одного отсчёта

Кодирование звука

Для определения объёма звукового файла I используют формулу:

$$I = f * i * t * S$$

Где:

I – объём звукового файла

f – частота дискретизации

i – количество бит, кодирующих один отсчёт

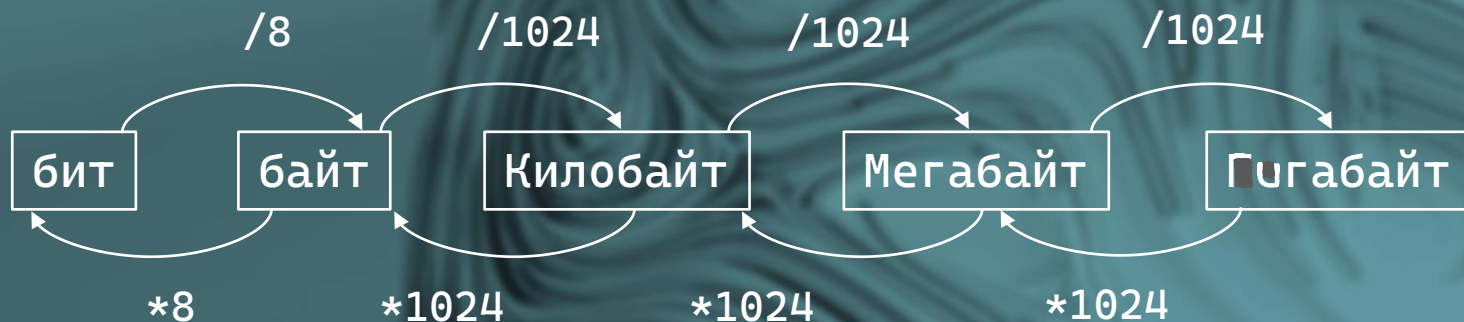
t – длительность звуковой дорожки

S – количество каналов (для моно $S = 1$, для стерео $S = 2$, для квадро $S = 4$)

Единицы измерения информации

1 байт	8 бит	2^3 бит
1 Килобайт (Кб)	1024 байт	2^{10} байт
1 Мегабайт (Мб)	1024 Килобайт	2^{20} байт
1 Гигабайт (Гб)	1024 Мегабайт	2^{30} байт

При переводе из меньшей единицы в большую делим



При переводе из большей единицы в меньшую умножаем

Скорость передачи информации

Канал связи всегда имеет ограниченную пропускную способность (скорость передачи информации), которая зависит от свойств аппаратуры и самой линии связи(кабеля)

$$I = v * t$$

Где:

I – объём информации


v – пропускная способность канала связи (измеряется в битах в секунду)

t – время передачи файла

$$V = I/t$$

Частые ошибки в задании #7

Если вычисления получаются слишком громоздкими, значит, вы неправильно решаете задачу. Удобно выделить во всех множителях степени двойки, тогда умножение сведётся к сложению показателей степеней, а деление – к их вычитанию.



Задание #11

Вычисление объёма информации.

Скорость передачи информации

С помощью K бит можно закодировать N различных символов

$$N \leq 2^i$$

Где:

N – мощность алфавита (количество символов в нём)

i – количество бит для хранения одного символа из N вариантов символов

Для определения объёма сообщения I используют формулу:

$$I = n * i$$

Где:

I – объём сообщения

n – количество символов в сообщении (его длина)

i – количество бит, кодирующих один символ

Частые ошибки в задании #11

Необходимо учитывать, что в заданиях этой линии для кодирования слов обычно отводится одинаковое и минимально возможное целое число байт, а для кодирования символов – одинаковое и минимально возможное целое количество бит