



B5

**Кодирование и
декодирование информации.**

Что нужно знать:

- 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
- закодированное сообщение можно однозначно декодировать с начала, если выполняется *условие Фано*: **никакое кодовое слово не является началом другого кодового слова**;
- закодированное сообщение можно однозначно декодировать с конца, если выполняется *обратное условие Фано*: **никакое кодовое слово не является окончанием другого кодового слова**

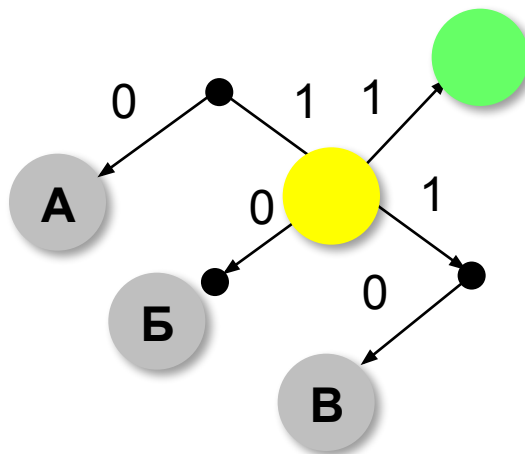
Задача №1

Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный по длине код: **A=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

1

11



Задача №2

Сообщения, содержат буквы П, О, С, Т; используется двоичный код, допускающий однозначное декодирование. Кодовые слова:

Т: **111**, О: **0**, П: **100**.

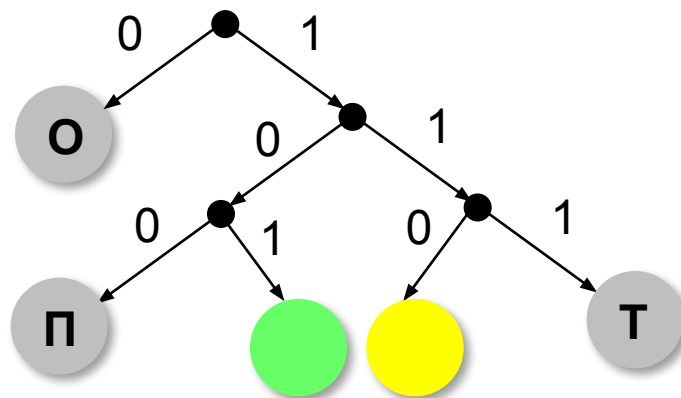
Укажите кратчайшее кодовое слово для буквы С, при котором код будет допускать однозначное декодирование. Если таких кодов несколько, укажите код с наименьшим числовым значением.

1

0x 10 11

0xx 101

110



Задача №3

Сообщения содержат три гласные буквы: А, Е, И – и пять согласных букв: Б, В, Г, Д, К. Буквы кодируются префиксным кодом. Известно, что все кодовые слова для согласных имеют **одну и ту же длину**, и

А – **1**, Е – **01**, И – **001**.

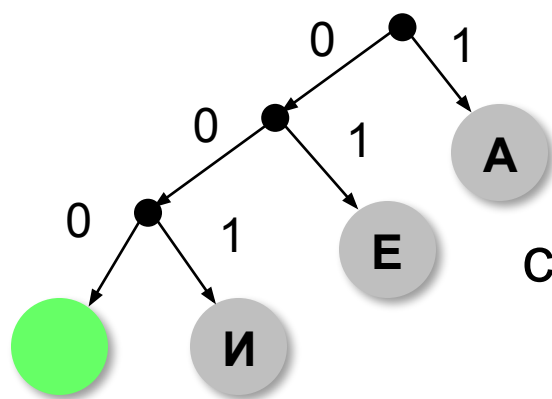
Какова наименьшая возможная длина кодовых слов для согласных букв?

5 согласных букв $\Rightarrow \geq 3$ бита

4 бита

5 бит

6 бит



4: **1xx**

2: **01x**

1: **001**

свободны: **000**

1

000x

2

000xx

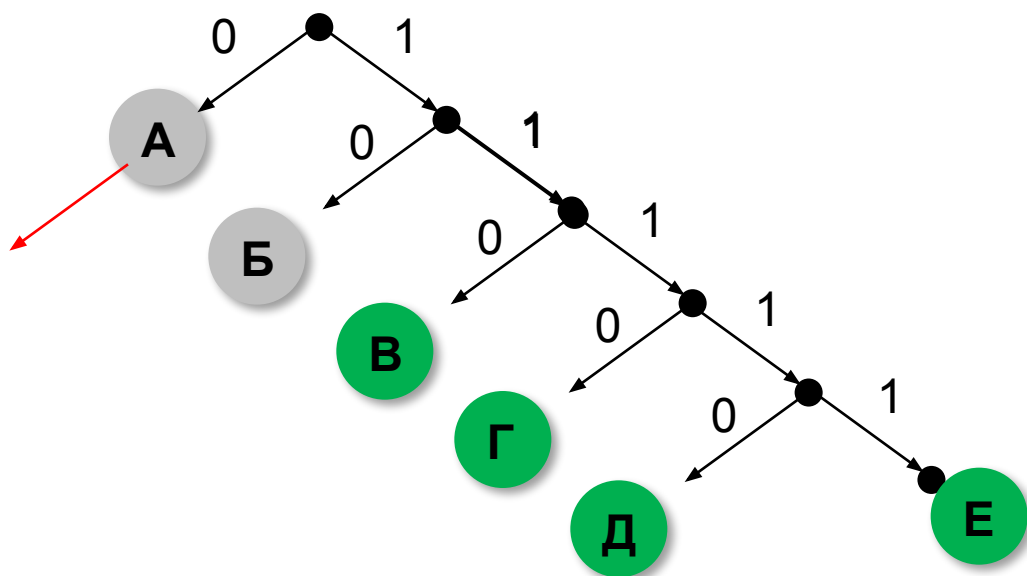
4

000xxx

8

Задача №4

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, Д, Е, решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы **А** использовали кодовое слово **0**; для буквы **Б** – кодовое слово **10**. Какова наименьшая возможная сумма длин всех шести кодовых слов? Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.



0xx

В: 110

Г: 1110

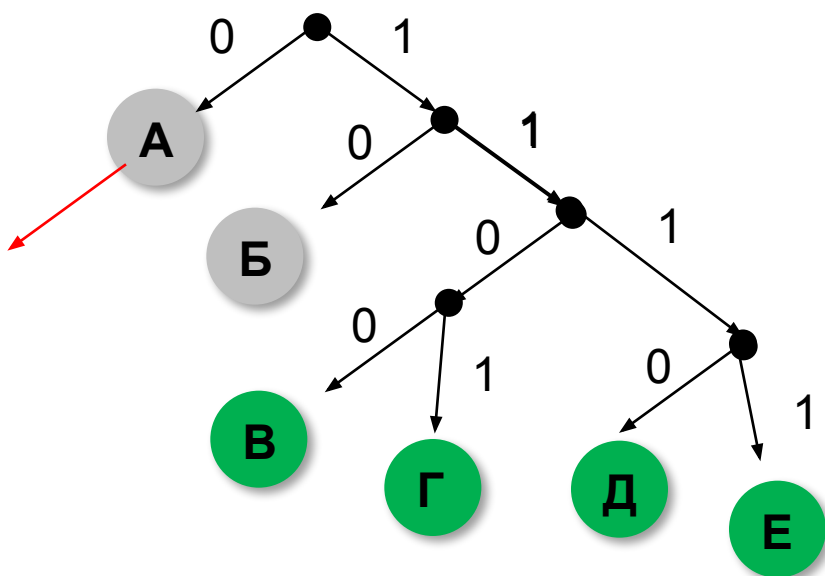
Д: 11110

Е: 11111

$$1 + 2 + 3 + 4 + 2 \cdot 5 = 20$$

Задача №4 (продолжение)

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, Д, Е, решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы **А** использовали кодовое слово **0**; для буквы **Б** – кодовое слово **10**. Какова наименьшая возможная сумма длин всех шести кодовых слов? Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.



0xx

В: 1100

Г: 1101

Д: 1110

Е: 1111

$$1 + 2 + 4 \cdot 4 = 19$$

Задача №5

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

Решение:

1) Код равномерной длины (цифра - 5 бит):

$2 \rightarrow 00101$ и $3 \rightarrow 00110$

2) 4 первых бита – это 2-ый код цифры, а пятый бит (бит четности)

рассчитывается как остаток от деления суммы битов на 2;

3) $2 = 0010_2$, бит четности $(0 + 0 + 1 + 0) \bmod 2 = 1$

4) $3 = 0011_2$, бит четности $(0 + 0 + 1 + 1) \bmod 2 = 0$

5) пятый бит в каждой пятерке **можно отбросить!**

6) $0101\mathbf{0}$, $1001\mathbf{0}$, $0111\mathbf{1}$, $0001\mathbf{1} \Rightarrow 0101, 1001, 0111, 0001$.

$0101_2 = 5$, $1001_2 = 9$, $0111_2 = 7$, $0001_2 = 1$.

Ответ: 2

Задача № 6

По каналу связи передаются сообщения, каждое из которых содержит 16 букв А, 8 букв Б, 4 буквы В и 4 буквы Г (других букв в сообщениях нет). Каждую букву кодируют двоичной последовательностью. При выборе кода учитывались два требования:

- ни одно кодовое слово не является началом другого (это нужно, чтобы код допускал однозначное декодирование);
- общая длина закодированного сообщения должна быть как можно меньше.

Какой код из приведённых ниже следует выбрать для кодирования букв А, Б, В и Г?

- 1) А:0, Б:10, В:110, Г:111
- 2) А:0, Б:10, В:01, Г:11
- 3) А:1, Б:01, В:011, Г:001
- 4) А:00, Б:01, В:10, Г:11

$$16 \cdot 1 + 8 \cdot 2 + 4 \cdot 3 + 4 \cdot 3 = 56 \text{ битов}$$

Условие «а» не выполняется

$$16 \cdot 2 + 8 \cdot 2 + 4 \cdot 2 + 4 \cdot 2 = 64 \text{ бита}$$

Ответ: 1

Задачи:

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	01	100	10	011

Определить, какой набор букв закодирован двоичной строкой 0110100011000

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

2. 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}

Задачи:

3. Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г и Д, решили использовать неравномерный двоичный код, позволяющий однозначно декодировать двоичную последовательность, появляющуюся на приёмной стороне канала связи. Использовали код: А–1, Б–000, В–001, Г–011. Укажите, каким кодовым словом должна быть закодирована буква Д. Длина этого кодового слова должна быть наименьшей из всех возможных. Код должен удовлетворять свойству однозначного декодирования.

- 1) 00 2) 01 3) 11 4) 010

4. 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–00, Б–010, В–011, Г–101, Д–111. Можно ли сократить для одной из букв длину кодового слова так, чтобы код по-прежнему можно было декодировать однозначно? Коды остальных букв меняться не должны. Выберите правильный вариант ответа.

- 1) для буквы Б – 01 2) это невозможно
3) для буквы В – 01 4) для буквы Г – 01

5. По каналу связи передаются сообщения, содержащие только 4 буквы: E, H, O, T. Для кодирования букв E, H, O используются 5-битовые кодовые слова: E - 00000, H - 00111, O - 11011. Для этого набора кодовых слов выполнено такое свойство: любые два слова из набора отличаются не менее чем в трёх позициях. Это свойство важно для расшифровки сообщений при наличии помех. Какое из перечисленных ниже кодовых слов можно использовать для буквы T, чтобы указанное свойство выполнялось для всех четырёх кодовых слов?

1) 11111 2) 11100 3) 00011 4) не подходит ни одно из указанных выше слов

6. По каналу связи передаются сообщения, содержащие только 4 буквы: A, И, С, T.

В любом сообщении больше всего букв A, следующая по частоте буква – С, затем – И. Буква T встречается реже, чем любая другая. Для передачи сообщений нужно использовать неравномерный двоичный код, допускающий однозначное декодирование; при этом сообщения должны быть как можно короче. Шифровальщик может использовать один из перечисленных ниже кодов. Какой код ему следует выбрать?

1) A – 0, И – 1, С – 00, T – 11

2) С – 1, И – 0, A – 01, T – 10

3) A – 1, И – 01, С – 001, T – 000

4) С – 0, И – 11, A – 101, T – 100

7. 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 – 0; Б – 100; В – 1010; Г – 111; Д – 110. 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 В – 101 | 2) это невозможно |
| 3) для буквы В – 010 | 4) для буквы Б – 10 |

8. По каналу связи передаются сообщения, содержащие только 5 букв А, И, К, О, Т. Для кодирования букв используется неравномерный двоичный код с такими кодовыми словами:

А — 0, И — 00, К — 10, О — 110, Т — 111.

Среди приведённых ниже слов укажите такое, код которого можно декодировать только одним способом. Если таких слов несколько, укажите первое по алфавиту.

- 1) КАА 2) ИКОТА 3) КОТ 4) ни одно из сообщений не подходит