

КОДИРОВАНИЕ И ДЕКОДИРОВАНИЕ ИНФОРМАЦИИ

Малаховская М.Г.,
учитель информатики высшей квалификационной
категории, заместитель директора по УВР
МБОУ лицей №3

Сургут, 2019 г

Задание ЕГЭ №5, базовый уровень

Типы заданий

- Кодирование в различных системах счисления
- Расшифровка сообщений
- Передача информации. Выбор кода

Знать:

- **Кодирование** — это представление информации в форме, удобной для её хранения, передачи и обработки.
- Правило преобразования информации к такому представлению называется **кодом**.
- Код называется однозначно декодируемым, если любое сообщение, составленное из кодовых слов, можно декодировать единственным способом.

Пример

A – 1 B – 0 C – 110

Тогда сообщение **01110110**

0 1 1 1 0 1 1 0 ВАААВААВ

0 1 1 1 0 110 ВАААВC

0 1 110 110 ВАСС

...

Сообщение декодируется несколькими способами

Знать:

- Кодирование бывает **равномерным** и **неравномерным**

А	Б	В	Г
00	01	10	11

А	Б	В	Г
000	10	01	001

- при равномерном кодировании всем символам соответствуют коды одинаковой длины;
- равномерное кодирование всегда *однозначно декодируемо*;
- при неравномерном кодировании разным символам соответствуют коды разной длины, это затрудняет декодирование.

Условие ФАНО

- **Прямое условие Фано:** ни одно кодовое слово не должно являться началом другого кодового слова (что *обеспечивает однозначное* декодирование сообщений с начала)
- **Префиксный код** — это неравномерный код, для которого выполняется прямое условие Фано.

А - 00

А - 00

Б - 01

Б - 01

В - 100

В - 100

Условие ФАНО

- **Обратное условие Фано:** никакое кодовое слово не является окончанием другого кодового слова
- **Постфиксный код** - это неравномерный код, для которого выполняется обратное условие Фано. Сообщения при использовании такого кода декодируются однозначно и только с конца.

А - 00

Б - 10

В - 001

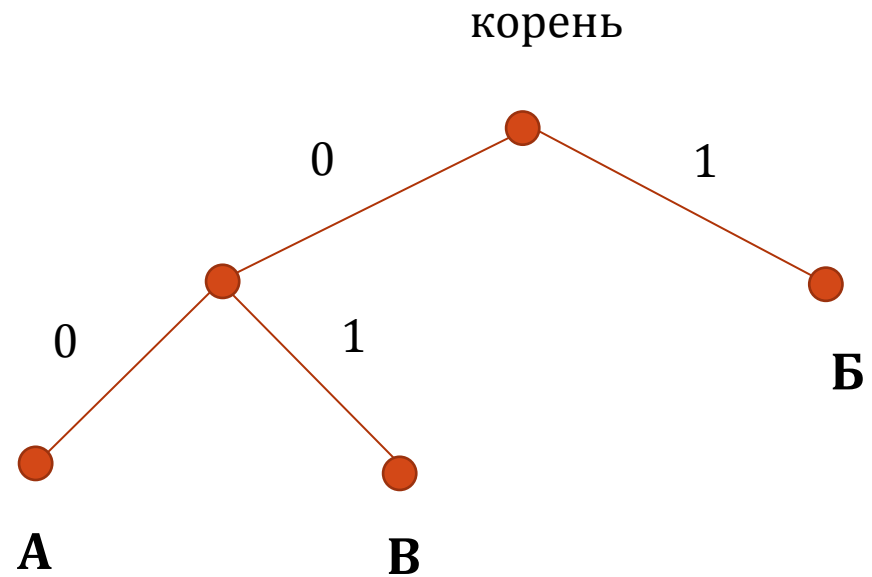
А - 00

Б - 10

В - 001

Уметь: строить двоичное дерево

A - 00 B - 1 В - 01



Задание 1

Пусть заданы следующие коды букв: **A – 100, B – 101, C – 1011.**

Требуется определить, выполняются ли для этого набора кодов условия Фано.

A – 100

A – 100

B – 101

B – 101

C – 1011

C – 1011

Ответ: выполняется обратное условие Фано

Задание 2

Для кодирования букв О, В, Д, П, А решили использовать двоичное представление чисел 0, 1, 2, 3 и 4 соответственно (с сохранением одного незначащего нуля в случае одноразрядного представления). Закодируйте последовательность букв **ВОДОПАД** таким способом и результат запишите восьмеричным кодом.

О	В	Д	П	А
0	1	2	3	4
00	01	10	11	100

ВОДОПАД — 01000100001101100010

Ответ: 2 2 1 6 2

Задание 3

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

A=0, B=10, V=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

Задание 3

1) 1

2) 1110

3) 111

4) 11

A	Б	В	Г
0	10	110	?

Г=1

A	Б	В	Г
0	10	110	1

Г=11

A	Б	В	Г
0	10	110	11

Г=111

A	Б	В	Г
0	10	110	111

Ответ: 3

Задание 4 (№44, Поляков К.)

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-10, Б-11, В-000, Г-001, Д-011**. 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) это невозможно | 2) для буквы Б – 1 |
| 3) для буквы Г – 00 | 4) для буквы Д – 01 |

Задание 4 (№44, Поляков К.)

A-10, Б-11, В-000, Г-001, Д-011

- 1) это невозможно 2) для буквы Б – 1
3) для буквы Г – 00 4) для буквы Д – 01

Б=1 **A-10 Б-1 В-000 Г-001 Д-011**

Нарушение прямого и обратного условия Фано

Г=00 **A-10 Б-11 В-000 Г-00 Д-011**

Нарушение прямого и обратного условия Фано

Д=01 **A-10 Б-11 В-000 Г-001 Д-01**

Выполняется прямое условие Фано

Ответ: 4

Задание 5

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, Д, Е, используется неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы А использовано кодовое слово 0, для буквы Б – кодовое слово 10.

Какова наименьшая возможная сумма длин всех шести кодовых слов?

Задание №5

Найти наименьшую возможную сумму длин всех
шести кодовых слов

A = 0, Б = 10

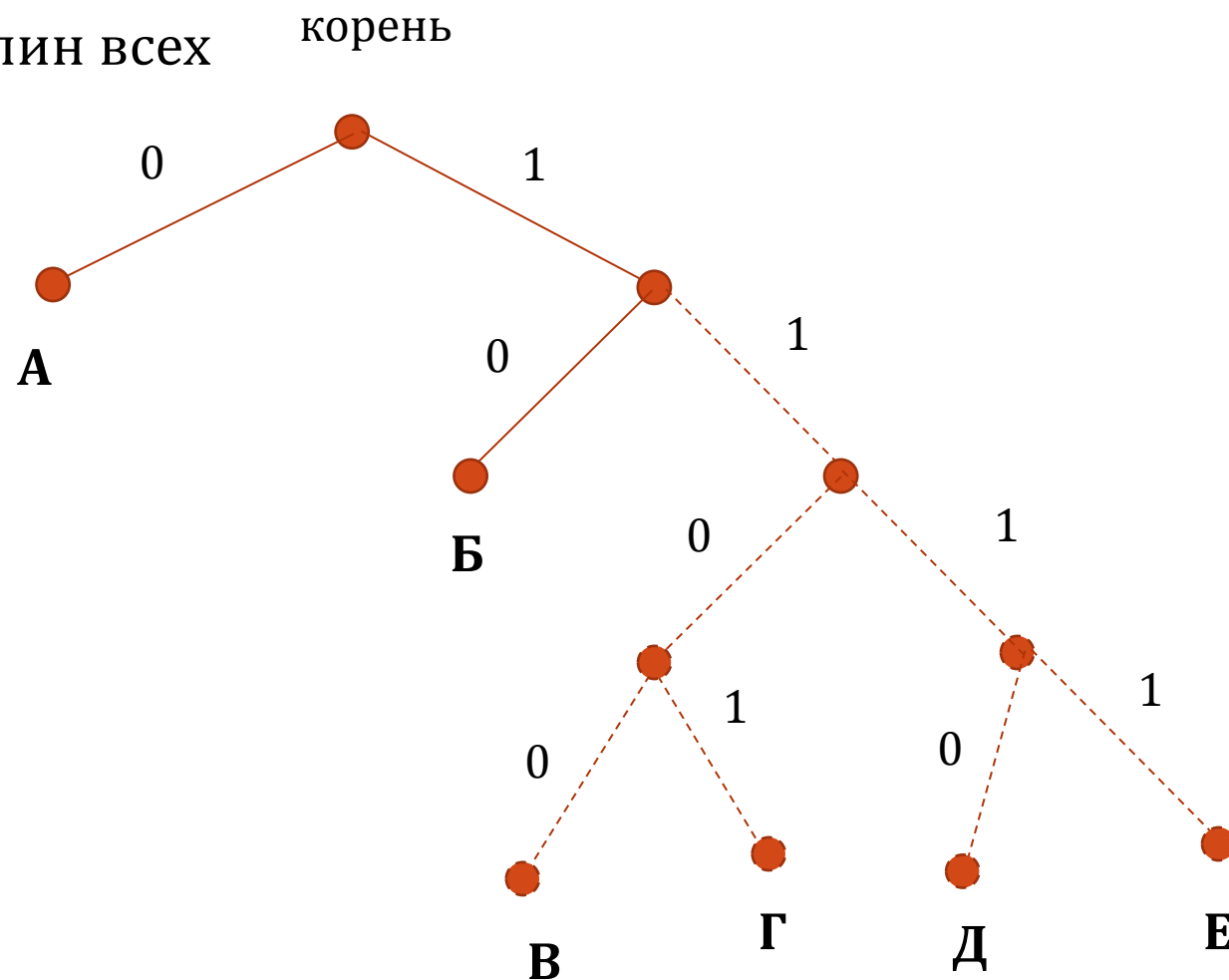
В, Г, Д, Е - ?

В = 1100 Г = 1101

Д = 1110 Е = 1111

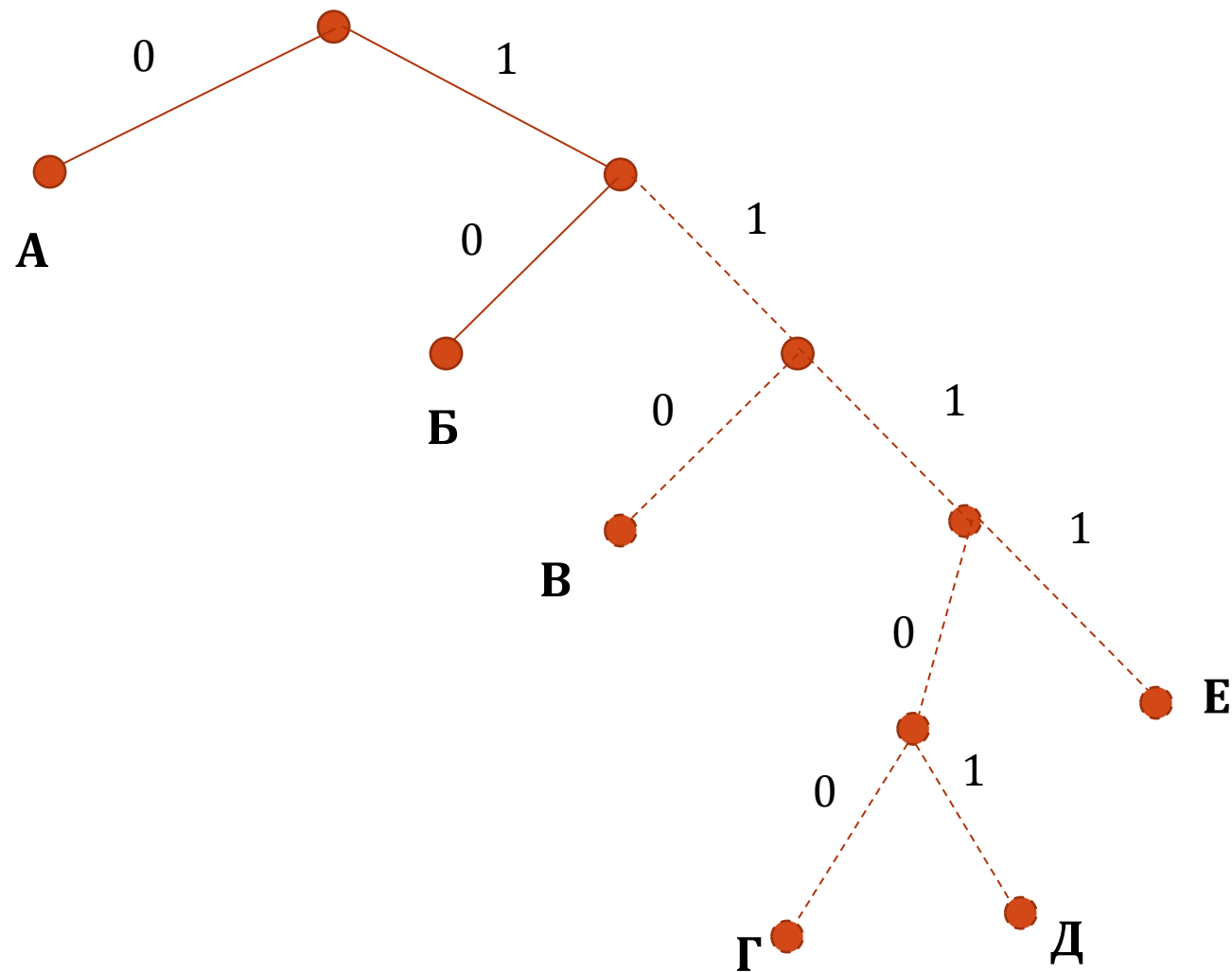
$$1 + 2 + 4 + 4 + 4 + 4 = 19$$

Ответ: 19



Задание №5 (2 вариант)

корень



$$A = 0, B = 10$$

$$B = 110$$

$$D = 11101$$

$$Г = 11100$$

$$E = 1111$$

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

Задание №6

По каналу связи передаются сообщения, содержащие только шесть букв: А, В, С, D, Е, F. Для передачи используется неравномерный двоичный код, удовлетворяющий условию Фано. Для букв А, В, С используются такие кодовые слова: **А – 11, В – 101, С – 0.**

Укажите кодовое слово наименьшей возможной длины, которое можно использовать для буквы **F**. Если таких слов несколько, укажите то из них, которое соответствует наименьшему возможному двоичному числу.

Задание №6

По каналу связи передаются сообщения, содержащие только шесть букв: А, В, С, D, Е, F. Для передачи используется неравномерный двоичный код, удовлетворяющий условию Фано. Для букв А, В, С используются такие кодовые слова: А – 11, В – 101, С – 0.

Укажите кодовое слово **наименьшей возможной длины**, которое можно использовать для буквы F. Если таких слов несколько, укажите то из них, которое соответствует **наименьшему возможному двоичному числу**.

Задание №6

A – 11, B – 101, C – 0

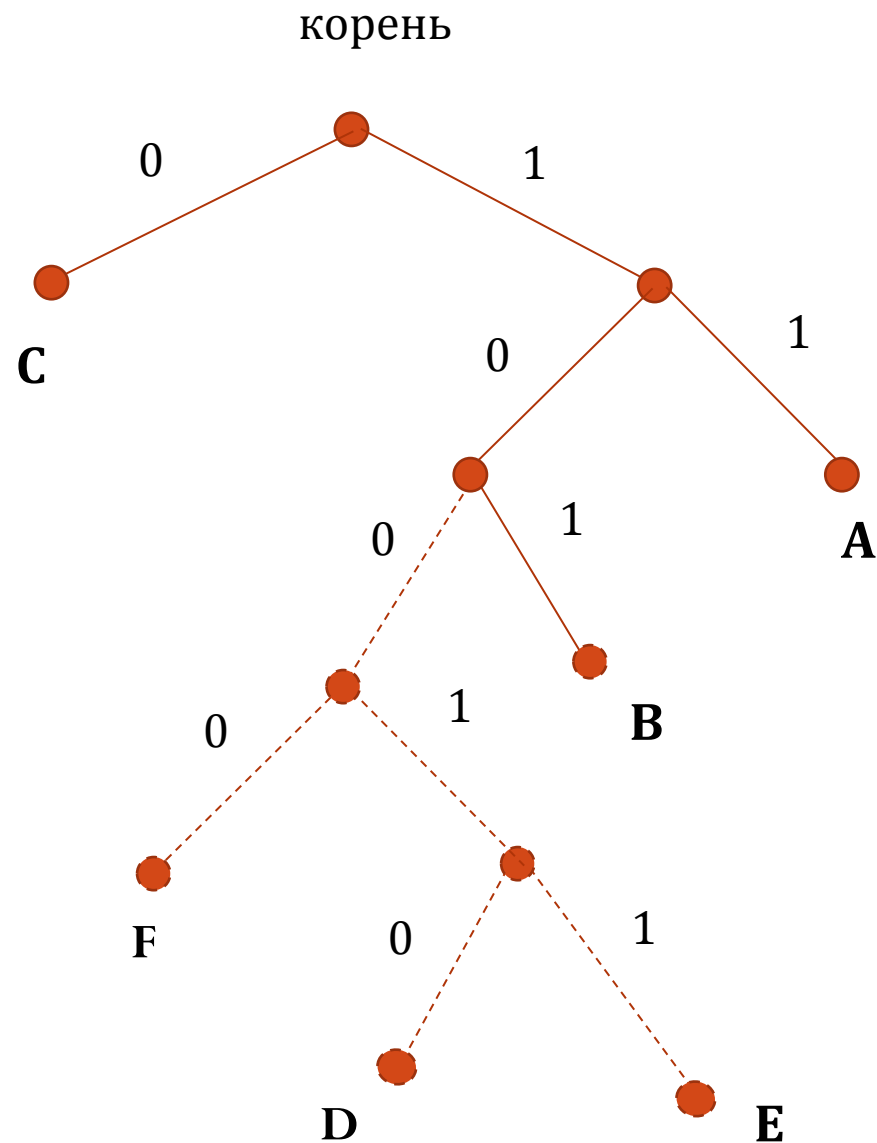
F - ? D - ? E - ?

F - 1000

D - 10010

E - 10011

Ответ: 1000



Задание №7 (104, Поляков К.)

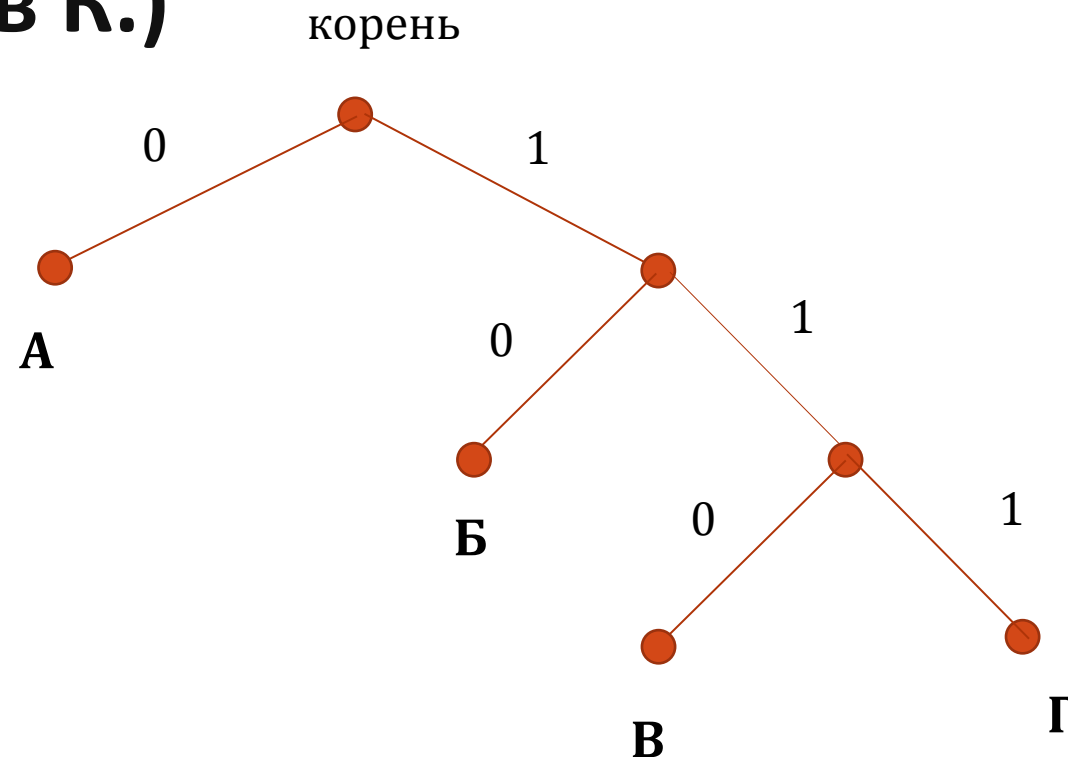
В сообщении встречается 50 букв А, 30 букв Б, 20 букв В и 5 букв Г. При его передаче использован неравномерный двоичный *префиксный* код, который позволил получить минимальную длину закодированного сообщения. Какова она в битах?

Задание №7 (104, Поляков К.)

**50 букв А, 30 букв Б,
20 букв В, 5 букв Г**

А – 0, Б – 10, В – 110, Г – 111

$$50 * 1 + 30 * 2 + 20 * 3 + 5 * 3 = 50 + 60 + 60 + 15 = 185 \text{ бит}$$



Ответ: 185

Задание №8 (105, Поляков К.)

По каналу связи передаются сообщения, каждое из которых содержит 15 букв А, 10 букв Б, 6 букв В и 4 буквы Г (других букв в сообщениях нет). Каждую букву кодируют двоичной последовательностью. При выборе кода учитывались два требования:

- а) ни одно кодовое слово не является началом другого (это нужно, чтобы код допускал однозначное декодирование);
- б) 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:1, Б:01, В:001, Г:111
- 2) А:1, Б:01, В:10, Г:111
- 3) А:00, Б:01, В:10, Г:11
- 4) А:100, Б:101, В:11, Г:0

**! Проверка:
условие Фано**

Задание №8 (105, Поляков К.)

1) A:1, Б:01, В:001, Г:111

букв А - 15

букв Б - 10

2) A:1, Б:01, В:10, Г:111

букв В - 6

буквы Г - 4

3) A:00, Б:01, В:10, Г:11

4) A:100, Б:101, В:11, Г:0

Вычисляем:

3) A:00, Б:01, В:10, Г:11

$$15 * 2 + 10 * 2 + 6 * 2 + 4 * 2 = 70$$

4) A:100, Б:101, В:11, Г:0

$$15 * 3 + 10 * 3 + 6 * 2 + 4 * 1 = 91$$

Ответ: 3

Задание 9 (№119, Поляков К.)

Для кодирования букв А, Б, В, Г, Д, Е, Ж, З, И, использован неравномерный троичный код, удовлетворяющий условию Фано.

Для буквы А используется кодовое слово 0; для буквы Б используется кодовое слово 10; для буквы В используется кодовое слово 11; для буквы Г используется кодовое слово 21; для буквы Д используется кодовое слово 22.

Какова минимальная общая длина кодовых слов для букв Е, Ж, З, И?

Задание 9 (№119, Поляков К.)

А - 0; Б - 10; В - 11; Г - 21; Д - 22.

Какова минимальная общая длина кодовых слов для букв Е, Ж, З, И?

Коды:

Е - 12

Ж - 200

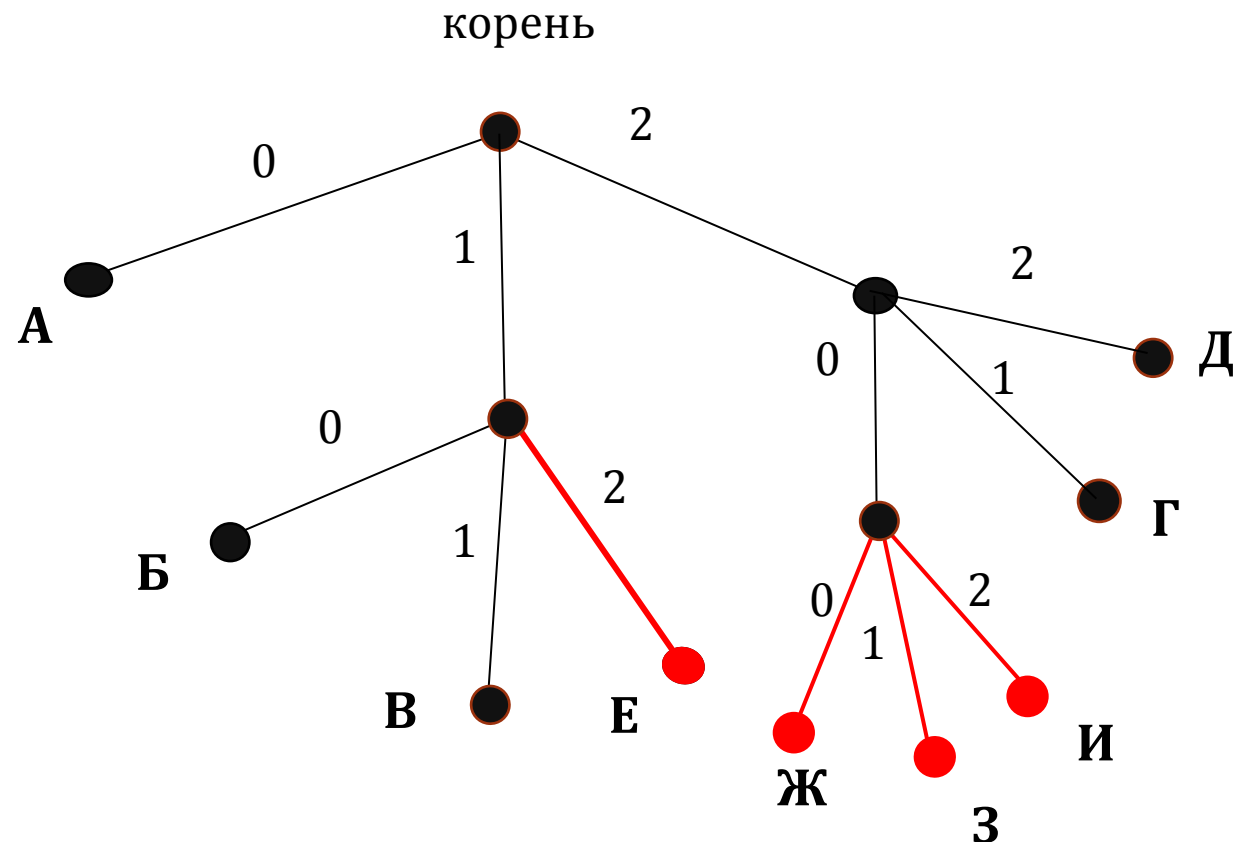
З - 201

И - 202

Суммарная длина кодов:

$$2 + 3 + 3 + 3 = 11$$

Ответ: 11



Задание 10 (№58, Поляков К.)

По каналу связи передаются сообщения, содержащие только 4 буквы:

А, Б, В, Г. Для кодирования букв А, Б, В используются 5-битовые кодовые слова: А - 00110, Б - 11000, В - 10011. 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 Г, чтобы указанное свойство выполнялось для всех четырёх кодовых слов?

- 1) 01101 2) 01001
- 3) 00011 4) не подходит ни одно из указанных выше слов

Задание 10 (№58, Поляков К.)

1) 01101

2) 01001

3) 00011

4) не подходит ни одно из указанных выше слов

А – 00110 Б – 11000 В – 10011

01101 01101 01101

не менее 3 отличий

А – 00110 Б – 11000 В – 10011

01001 01001 01001

для буквы Б два отличия

А – 00110 Б – 11000 В – 10011

00011 00011 00011

для буквы А два отличия, для
буквы В одно отличие

Ответ: 1

Задание 11 (№ 74, Поляков К.)

По каналу связи передаются сообщения, содержащие только 4 буквы: Е, Н, О, Т. В любом сообщении больше всего букв О, следующая по частоте буква – Е, затем – Н. Буква Т 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) Е – 0, Н – 1, О – 00, Т – 11

2) О – 1, Н – 0, Е – 01, Т – 10

3) Е – 1, Н – 01, О – 001, Т – 000

4) О – 0, Н – 10, Е – 111, Т – 110

Задание 11 (№ 74, Поляков К.)

Количество букв: $O > E > H > T$

1) $E - 0, H - 1, O - 00, T - 11$

2) $O - 1, H - 0, E - 01, T - 10$

3) $E - 1, H - 01, O - 001, T - 000$

4) $O - 0, H - 10, E - 111, T - 110$

**! Проверка:
условие Фано**

Задание 11 (№ 74, Поляков К.)

Количество букв: $O > E > H > T$

1) E – 0, H – 1, O – 00, T – 11

2) O – 1, H – 0, E – 01, T – 10

3) E – 1, H – 01, O – 001, T – 000

4) O – 0, H – 10, E – 111, T – 110

Для 1 и 2 нарушается условие Фано

Ответ: 4

Задание 12 (№75, Поляков К.)

Для передачи данных по каналу связи используется 5-битовый код. Сообщение содержит только буквы А, Б и В, которые кодируются следующими кодовыми словами: **А - 10111, Б - 00000, В - 11010**

При передаче возможны помехи. Однако некоторые ошибки можно попытаться исправить. Любые два из этих трёх кодовых слов отличаются друг от друга не менее чем в трёх позициях. Поэтому если при передаче слова произошла ошибка не более чем в одной позиции, то можно сделать обоснованное предположение о том, какая буква передавалась. (Говорят, что «код исправляет одну ошибку».) Например, если получено кодовое слово 00100, считается, что передавалась буква Б. (Отличие от кодового слова для Б только в одной позиции, для остальных кодовых слов отличий больше.) Если принятое кодовое слово отличается от кодовых слов для букв А, Б, В более чем в одной позиции, то считается, что произошла ошибка (она обозначается 'х').

Получено сообщение 10101100001111010010. Декодируйте это сообщение

1) АБВВ

2) хххх

3) АБхх

4) АБхВ

Задание 12 (№75, Поляков К.)

А - 10111, Б - 00000, В – 11010

Сообщение: 10101100001111010010

10101 10000 11110 10010

101**0**1 - отличается от символа А (10111) в одной позиции (передавалась буква А)

10000 - отличается от символа Б (00000) в одной позиции (передавалась буква Б)

11**1**10 - отличается от символа В (11010) в одной позиции (передавалась буква В)

1**0**010 - отличается от символа В (11010) в одной позиции (передавалась буква В)

Ответ: **АБВВ**

Задание 13 (№118, Поляков К.)

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, Д, Е, Ж решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для букв А, Б, В, Г использовали соответственно кодовые слова 000, 1, 010, 011. Укажите кратчайшее возможное кодовое слово для буквы Д, при котором код будет допускать однозначное декодирование. Если таких кодов несколько, укажите код с наименьшим числовым значением.

Задание 13 (№118, Поляков К.)

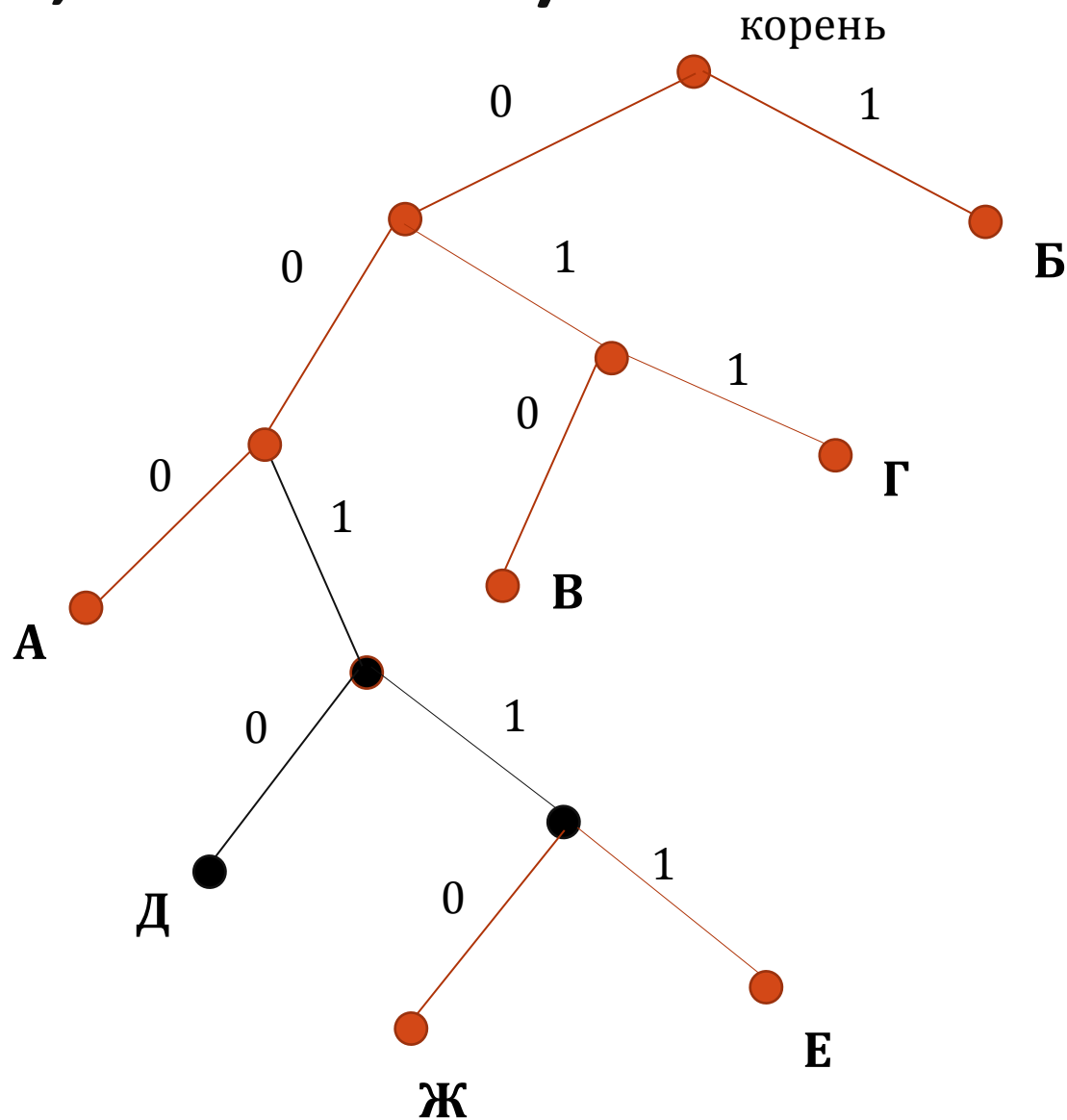
А, Б, В, Г, Д, Е, Ж

А-000 Б-1

В-010 Г-011

Д-?

Ответ: 0010



Задание 14 (№149, Поляков К.)

По каналу связи передаются сообщения, содержащие только семь букв: А, Б, И, Л, О, С, Ц. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: Б – 00, О – 010, Л – 111. Какое наименьшее количество двоичных знаков потребуется для кодирования слова АБСЦИССА?

Задание 14 (№118, Поляков К.)

А, **Б**, И, **Л**, **О**, С, Ц

Б – 00, О – 010, Л – 111

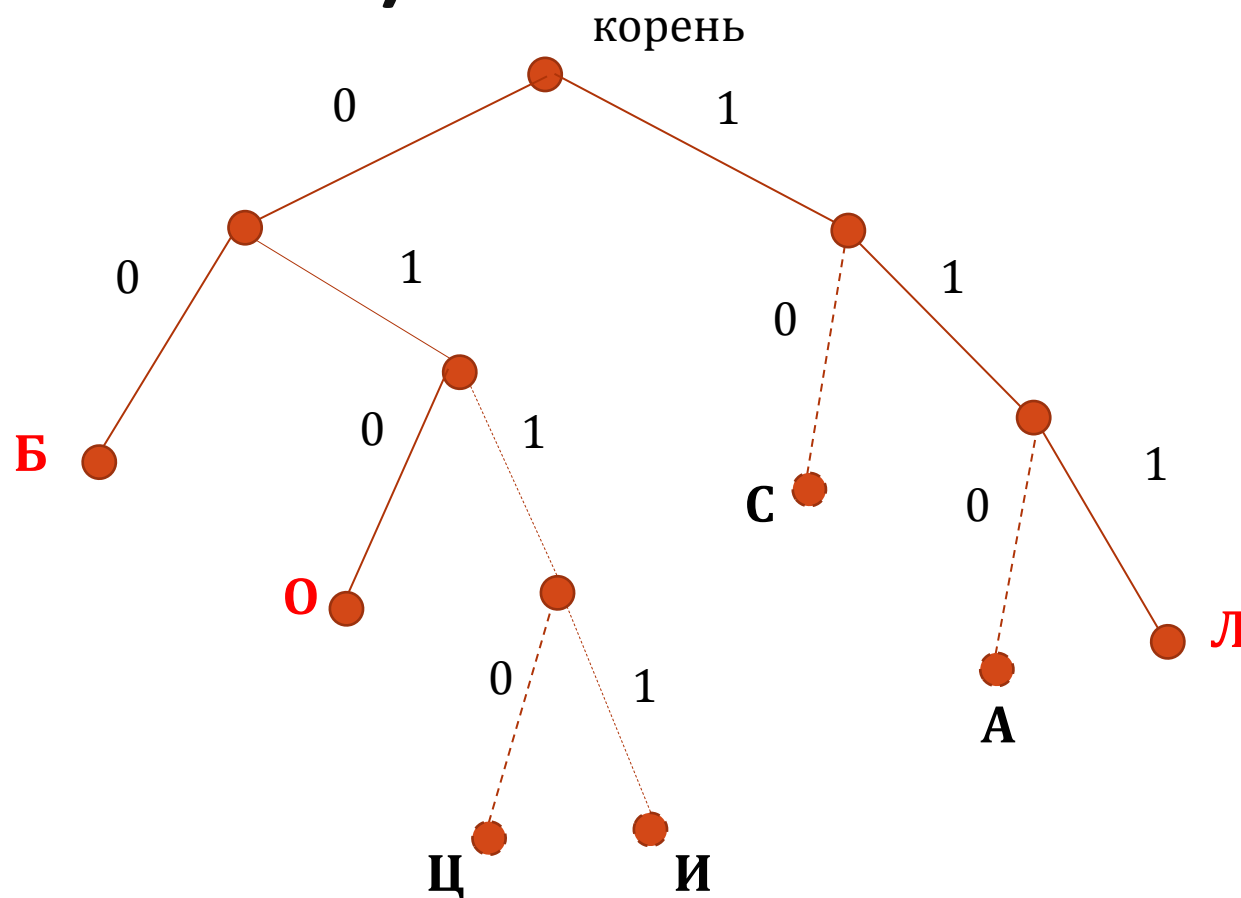
АБСЦИССА

Найти коды: С А И Ц

С - 10 А - 110 И - 0111 Ц - 0110

АБСЦИССА = $2*3+2+3*2+4+4=22$

Ответ: 22



СПАСИБО ЗА ВНИМАНИЕ!