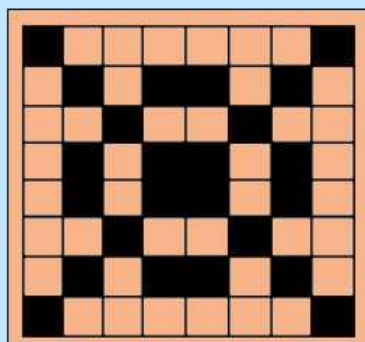


Двоичное кодирование.
Неравномерный код.



Двоичное кодирование



А
Б
В

254



1000 0001
0101 1010
0010 0100
0101 1010
0101 1010
0010 0100
0101 1010
1000 0001



1100 0000
1100 0001
1100 0010



1111 1110



1 0 1 1

10101001100
00010011101
11100011100



| символ | 10- в код | 2-в код | символ | 10- в код | 2-в код | символ | 10-в код | 2-в код | символ | 10-в код | 2-в код |
|--------|-----------------|----------|--------|-----------------|----------|--------|-------------|----------|--------|-------------|----------|
| | 32 | 00100000 | 8 | 56 | 00111000 | P | 80 | 01010000 | h | 104 | 01101000 |
| ! | 33 | 00100001 | 9 | 57 | 00111001 | Q | 81 | 01010001 | i | 105 | 01101001 |
| " | 34 | 00100010 | : | 58 | 00111010 | R | 82 | 01010010 | j | 106 | 01101010 |
| # | 35 | 00100011 | ; | 59 | 00111011 | S | 83 | 01010011 | k | 107 | 01101011 |
| \$ | 36 | 00100100 | < | 60 | 00111100 | T | 84 | 01010100 | l | 108 | 01101100 |
| % | 37 | 00100101 | = | 61 | 00111101 | U | 85 | 01010101 | m | 109 | 01101101 |
| & | 38 | 00100110 | > | 62 | 00111110 | V | 86 | 01010110 | n | 110 | 01101110 |
| ' | 39 | 00100111 | ? | 63 | 00111111 | W | 87 | 01010111 | o | 111 | 01101111 |
| (| 40 | 00101000 | @ | 64 | 01000000 | X | 88 | 01011000 | p | 112 | 01110000 |
|) | 41 | 00101001 | A | 65 | 01000001 | Y | 89 | 01011001 | q | 113 | 01110001 |
| * | 42 | 00101010 | B | 66 | 01000010 | Z | 90 | 01011010 | r | 114 | 01110010 |
| + | 43 | 00101011 | C | 67 | 01000011 | [| 91 | 01011011 | s | 115 | 01110011 |
| , | 44 | 00101100 | D | 68 | 01000100 | \ | 92 | 01011100 | t | 116 | 01110100 |
| - | 45 | 00101101 | E | 69 | 01000101 |] | 93 | 01011101 | u | 117 | 01110101 |
| . | 46 | 00101110 | F | 70 | 01000110 | ^ | 94 | 01011110 | v | 118 | 01110110 |
| / | 47 | 00101111 | G | 71 | 01000111 | _ | 95 | 01011111 | w | 119 | 01110111 |
| 0 | 48 | 00110000 | H | 72 | 01001000 | ` | 96 | 01100000 | x | 120 | 01111000 |
| 1 | 49 | 00110001 | I | 73 | 01001001 | a | 97 | 01100001 | y | 121 | 01111001 |
| 2 | 50 | 00110010 | J | 74 | 01001010 | b | 98 | 01100010 | z | 122 | 01111010 |
| 3 | 51 | 00110011 | K | 75 | 01001011 | c | 99 | 01100011 | { | 123 | 01111011 |
| 4 | 52 | 00110100 | L | 76 | 01001100 | d | 100 | 01100100 | | 124 | 01111100 |
| 5 | 53 | 00110101 | M | 77 | 01001101 | e | 101 | 01100101 | } | 125 | 01111101 |
| 6 | 54 | 00110110 | N | 78 | 01001110 | f | 102 | 01100110 | ~ | 126 | 01111110 |
| 7 | 55 | 00110111 | O | 79 | 01001111 | g | 103 | 01100111 | □ | 127 | 01111111 |

| | | |
|-----------|-------------|---------------|
| А • — | Л • — • • | Ц — • — • |
| Б — • • • | М — — | Ч — — — • |
| В • — — | Н — • | Ш — — — — |
| Г — — • | О — — — | Щ — — • — |
| Д — • • | П • — — • | Ъ • — — • — • |
| Е • | Р • — • | Ы — • — — |
| Ж • • • — | С • • • • | Ь — • • • — |
| З — — • • | Т — | Э • • • — • • |
| И • • | У • • — | Ю • • • — — |
| Й • — — — | Ф • • • — • | Я • — • — |
| К — • — | Х • • • • • | |

Условие Фано

Не всякий неравномерный код можно однозначно декодировать.

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;

Основной смысл заданий на
неравномерный код – подобрать такое
кодирование, которое удовлетворяет
условию Фано и будет занимать меньше
всего места.

Пример задания

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.

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

Решение

1. Отбросим варианты, в которых не выполняется условие Фано

1) A:0, Б:10, В:110, Г:111

~~2) A:0, Б:10, В:01, Г:11~~

~~3) A:1, Б:01, В:011, Г:001~~

4) A:00, Б:01, В:10, Г:11

2. Посчитаем размер сообщения в оставшихся вариантах

1) $16 \cdot 1 + 8 \cdot 2 + 4 \cdot 3 + 4 \cdot 3 = 58$ бит

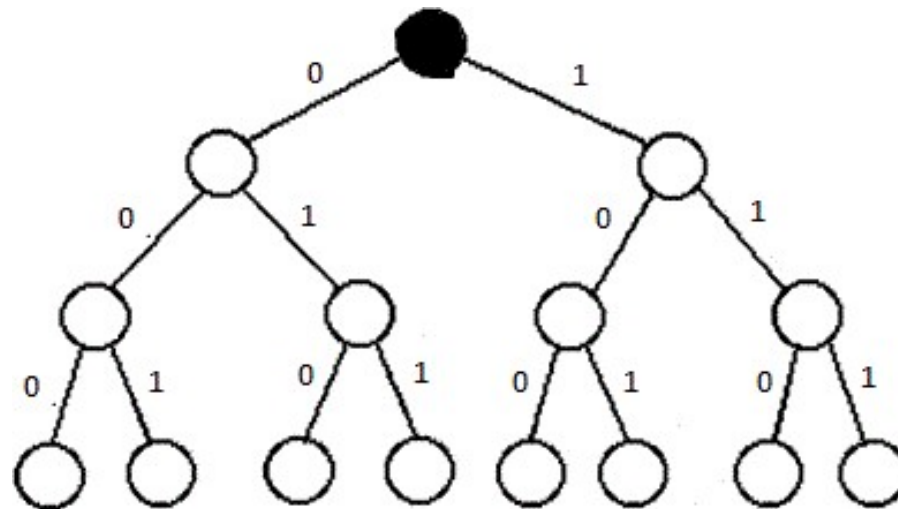
2) $16 \cdot 2 + 8 \cdot 2 + 4 \cdot 2 + 4 \cdot 2 = 64$ бита

Пример задания

Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы А использовали кодовое слово 0, для буквы Б – кодовое слово 110. Какова наименьшая возможная суммарная длина всех четырёх кодовых слов?

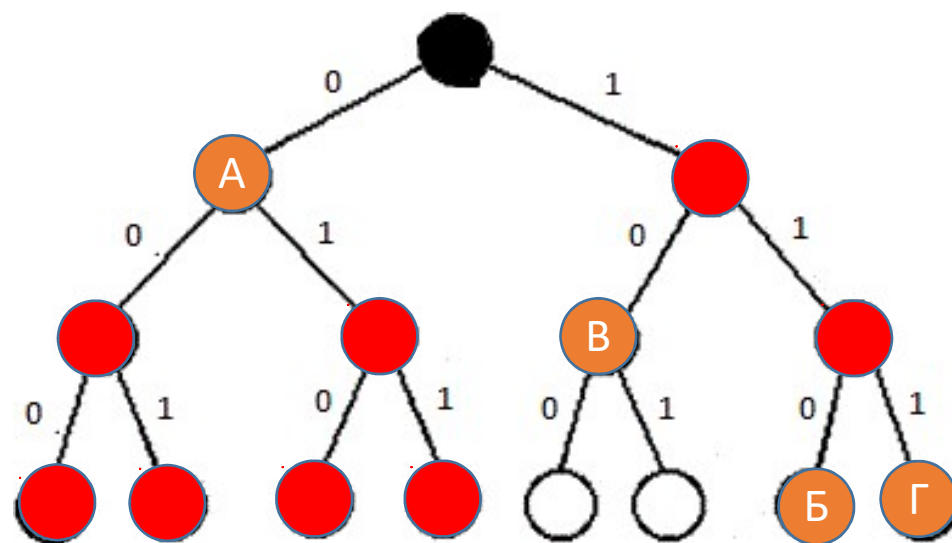
Решение

1. Построим дерево, в котором рассмотрим возможные варианты кодирования букв.



Решение

2. Расположим на нём наши буквы

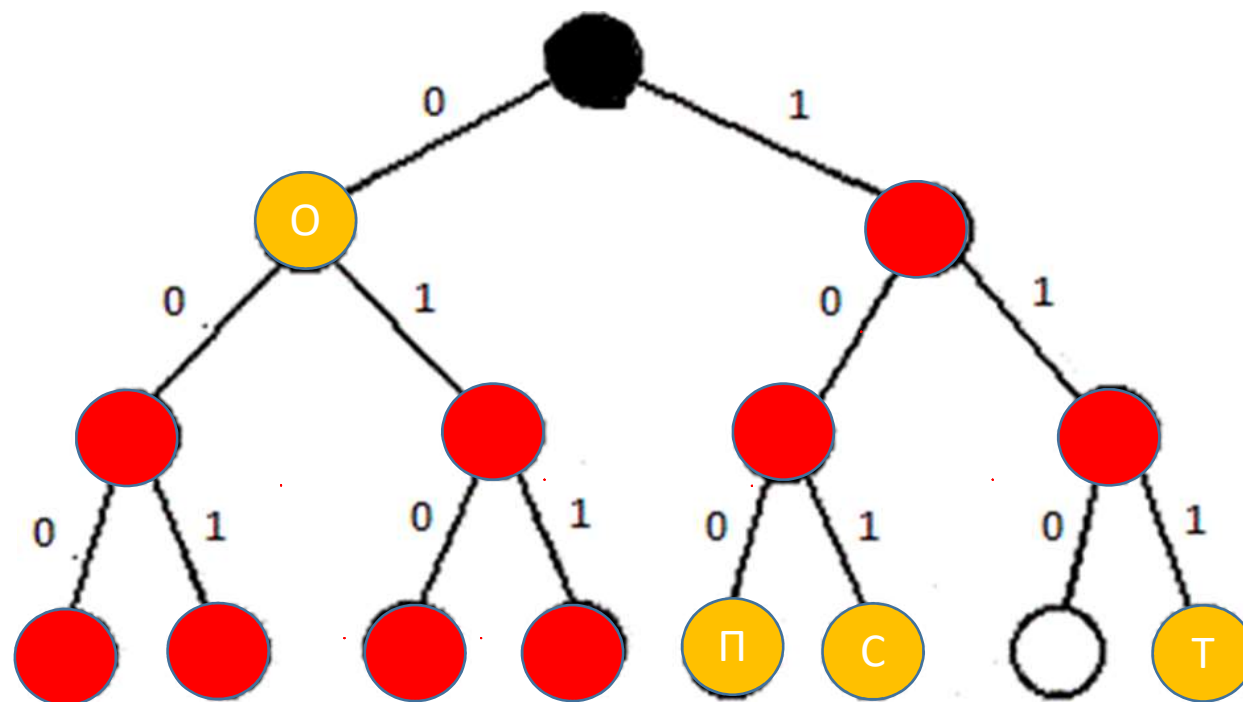


Пример задания

По каналу связи передаются сообщения, содержащие только 4 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.

Решение



Пример задания

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 |

Решение

Проверим на выполнение условия Фано все 4 варианта.

1) A – 0; Б – 100; В – 101; Г – 111; Д – 110

~~2) A – 0, Б – 100, В – 010, Г – 111, Д – 110~~

~~3) A – 0, Б – 10, В – 1010, Г – 111, Д – 110~~

Пример задания

По каналу связи передаются сообщения, содержащие только 4 буквы: Е, Н, О, Т. Для кодирования букв Е, Н, О используются 5-битовые кодовые слова: Е - 00000, Н - 00111, О - 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 Т, чтобы указанное свойство выполнялось для всех четырёх кодовых слов?

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

Пример задания

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

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 |

Определите, какой набор букв закодирован двоичной строкой 1100000100110

1) baade 2) badde 3) bacde 4) bacdb

2. Для 6 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 | F |
| 00 | 100 | 10 | 011 | 11 | 101 |

Определите, какая последовательность из 6 букв закодирована двоичной строкой 011111000101100.

1) DEFBAC 2) ABDEFC 3) DECAFB 4) EFCABD

3. Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный по длине код: А=1, Б=01, В=001. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?

1) 0001 2) 000 3) 11 4) 101

4. Для передачи по каналу связи сообщения, состоящего только из букв А, Б, В, Г, решили использовать неравномерный по длине код: А=0, Б=100, В=101. Как нужно закодировать букву Г, чтобы длина кода была минимальной и допускалось однозначное разбиение кодированного сообщения на буквы?

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

5. 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–001, В–0001, Г–110, Д–111. Можно ли сократить для одной из букв длину кодового слова так, чтобы код по-прежнему можно было декодировать однозначно?

1) для буквы Г – 11 2) это невозможно

3) для буквы В – 000 4) для буквы Б – 00

6. 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–111, Б–110, В–100, Г–0. Укажите, каким кратчайшим кодовым словом может быть закодирована буква Д.

1) 001 2) 00 3) 101 4) 10

7. 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

8. Сообщение содержит только буквы А, Б и В, которые кодируются следующими кодовыми словами:

А — 11010, Б — 00110, В — 10101.

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. 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 00111 11110 11000 10111. Декодируйте это сообщение – выберите правильный вариант.

1) БААх 2) БААВ 3) хААх 4) хххх

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

А — 10010, Б — 11111, В — 00101.

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. 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 10000 10101 11001 10111. Декодируйте это сообщение – выберите правильный вариант.

1) АВББ 2) хххх 3) АВхБ 4) АххБ

10. По каналу связи передаются сообщения, содержащие только 4 буквы Э, Ю, Я, Ы. Для передачи сообщений нужно использовать неравномерный двоичный код, допускающий однозначное декодирование. В любом сообщении больше всего букв Я, следующая по частоте буква — Ю, затем — Э. Буква Ы встречается реже, чем любая другая. Какой из перечисленных ниже кодов нужно использовать, чтобы передаваемые закодированные сообщения были как можно более короткими?

- 1) Э — 0, Ю — 1, Я — 00, Ы — 11
- 2) Я — 1, Ю — 0, Э — 01, Ы — 10
- 3) Э — 1, Ю — 01, Я — 001, Ы — 000
- 4) Я — 0, Ю — 11, Э — 101, Ы — 100

11. Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, Д и Е, используется неравномерный двоичный код, допускающий однозначное декодирование. Даны кодовые слова для четырёх букв: А — 011, Б — 010, В — 001, Г — 000. Какие кодовые слова из приведённых ниже вариантов подходят для букв Д и Е? Если подходит более одного варианта, укажите тот, для которого сумма длин кодовых слов меньше.

- 1) Д — 100, Е — 110 3) Д — 10, Е — 11
- 2) Д — 100, Е — 11 4) Д — 10, Е — 1

12. По каналу связи передаются сообщения, содержащие только 5 букв А, Д, Р, Т, К. Для кодирования букв используется неравномерный двоичный код с такими кодовыми словами:
А — 1, Д — 00, Р — 10, Т — 110, К — 111.

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

1) АКР 2) РАД 3) ТАРА 4) ни одно из сообщений не подходит

13. По каналу связи передаются сообщения, содержащие только 5 букв А, К, Л, Р, Т. Для кодирования букв используется неравномерный двоичный код с такими кодовыми словами:
А — 01, К — 010, Л — 011, Р — 11, Т — 101.

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

1) РАК 2) ЛАК 3) ТАРА 4) ни одно из сообщений не подходит

14. Для кодирования некоторой последовательности, состоящей из букв А, Б, В, Г, решили использовать неравномерный двоичный код, удовлетворяющий условию Фано. Для буквы А использовали кодовое слово 1, для буквы Б – кодовое слово 011. Какова наименьшая возможная суммарная длина всех четырёх кодовых слов?

1) 7 2) 8 3) 9 4) 10

15. В сообщении встречается 7 разных букв. При его передаче использован неравномерный двоичный код, удовлетворяющий условию Фано. Известны коды трёх букв: 1, 01, 001. Коды остальных четырёх букв имеют одинаковую длину. Какова минимальная суммарная длина всех 7-ми кодовых слов?

16. 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

17. По каналу связи передаются сообщения, каждое из которых содержит 10 букв А, 5 букв Б, 20 букв В и 5 букв 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:00, Б:01, В:10, Г:11
- 3) А:0, Б:10, В:11, Г:111
- 4) А:10, Б:111, В:0, Г:110

18. 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

19. 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 – 010. 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) это невозможно 2) для буквы А – 0 3) для буквы В – 00
- 4) для буквы Д – 01

20. 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?

21. 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?