Клод Шеннон предложил **формулу определения количества информации, учитывающую возможную неодинаковую вероятность сообщений в наборе**.

Формула Шеннона: $I = -(p_1 \log_2 p_1 + p_2 \log_2 p_2 + \ldots + p_N \log_2 p_N)$,
где p_i — вероятность того, что именно *i*-е сообщение выделено в наборе из N сообщений.

Формула Шеннона:

$$x = \log_2(1/p)$$

- x — количество информации в сообщении о событии
- p — вероятность события

Формула вероятности случайного события:

$$p(A) = m / n$$

- m — количество благоприятных исходов (число случаев, способствующих событию A)
- n — количество общих исходов (общее число равновозможных случаев)

Вывод: *количество информации в сообщении о некотором событии зависит от вероятности этого события. Чем меньше вероятность, тем больше информации*.

Информация как и в алфавитном подходе измеряется в битах, но по формуле может получиться нецелое количество бит. В случае необходимости оценить длину двоичного кода, кодирующего заданную информацию, это число округляется в большую сторону.

1. В корзине лежат 8 черных шаров и 24 белых. Сколько бит информации несет сообщение о том, что достали черный шар?
2. За четверть Василий Пупкин получил 20 оценок. Сообщение о том, что он вчера получил четверку, несет 2 бита информации. Сколько четверок получил Василий за четверть?
3. В колоде 36 карт. Из них 12 карт с «портретами» валетов, дам и королей. Какое количество информации содержит сообщение о том, что из колоды была взята карта с портретом (I1), туз (I2), любая карта от шестерки до десятки (I3), туз пик (I4)?
4. В коробке лежат 64 цветных карандаша. Сообщение о том, что достали белый карандаш, несёт 4 бита информации. Сколько белых карандашей было в корзине?
5. В корзине лежат 32 клубка красной и черной шерсти. Среди них 4 клубка красной шерсти. Сколько информации несет сообщение, что достали клубок красной шерсти? Сколько информации несет сообщение, что достали клубок шерсти любой окраски?

6. В озере живут 8000 карасей, 2000 щук и 4000 пескарей. Сколько информации содержится в сообщениях о том, что рыбак поймал карася, щуку, пескаря, поймал рыбу?
7. В коробке 50 шаров, из них 40 белых и 10 чёрных. Определить количество информации в сообщении о вытаскивании наугад белого шара и чёрного шара.
8. В непрозрачном мешочке хранятся 10 белых, 20 красных, 30 синих и 40 зеленых шариков. Какое количество информации будет содержать зрительное сообщение о цвете вынутого шарика?
9. За контрольную работу по информатике получено 8 пятерок, 13 четверок, 6 троек и 2 двойки. Какое количество информации получил Васечкин при получении тетради с оценкой?
10. В корзине лежат белые и черные шары. Среди них 18 черных шаров. Сообщение о том, что из корзины достали белый шар, несет 2 бита информации. Сколько всего в корзине шаров?
11. В концертном зале 270 девушек и несколько юношей. Сообщение «Первым из зала выйдет юноша» содержит 4 бита информации. Сколько юношей в зале.

Условие Фано.

В чем смысл прямого условия Фано?

Сформулировать данное условие можно следующим образом: «*ни одно кодовое слово не может выступать в качестве начала любого другого кодового слова*».

В чем смысл обратного условия Фано?

Существует также и обратное правило Фано, формулировка которого звучит следующим образом: «*ни одно кодовое слово не может выступать в качестве окончания любого другого кодового слова*».

12. 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?

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

13. Для кодирования растрового рисунка, напечатанного с использованием шести красок, применили неравномерный двоичный код. Для кодирования цветов используются кодовые слова.

Цвет	Кодовое слово
Белый	0
Зелёный	11111
Красный	1110

Цвет	Кодовое слово
Синий	
Фиолетовый	11110
Чёрный	10

ege.sdamgia.ru

Укажите кратчайшее кодовое слово для кодирования синего цвета, при котором код будет удовлетворять условию Фано. Если таких кодов несколько, укажите код с наименьшим числовым значением.

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

14. По каналу связи передаются шифрованные сообщения, содержащие только десять букв: А, Б, Е, И, К, Л, Р, С, Т, У. Для передачи используется неравномерный двоичный код. Для девяти букв используются кодовые слова. Для буквы А − 00, Е — 010, И — 011, К — 1111, Л — 1101, Р — 1010, С — 1110, Т — 1011, У — 100.

Укажите кратчайшее кодовое слово для буквы Б, при котором код будет удовлетворять условию Фано. Если таких кодов несколько, укажите код с наименьшим числовым значением.

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

15. По каналу связи передаются сообщения, содержащие только восемь букв: А, Б, Г, Е, И, М, Р, Т. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны:

Буква	Кодовое слово
А	11
Б	0010
Г	1011
Е	0011

Буква	Кодовое слово
И	
М	01
Р	000
Т	1010

ege.sdamgia.ru

Укажите кратчайшее кодовое слово для буквы И. Если таких кодов несколько, укажите код с **наименьшим** числовым значением.

Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

16. Для передачи данных используется двоичный код. Сообщение содержит только буквы А, Б, В или Г, для букв А, Б и В используются следующие кодовые слова:

А – 0, Б – 101, В – 111.

Найдите кодовое слово минимальной длины для Г при котором сохраняется прямое условие Фано. Если таких кодовых слов несколько, укажите кодовое слово с минимальным двоичным значением.

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

17. По каналу связи передаются сообщения, содержащие только семь букв: А, Б, Г, И, М, Р, Я. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: А — 010, Б — 00, Г — 101. Какое **наименьшее** количество двоичных знаков потребуется для кодирования слова ГРАММ?

Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

18. По каналу связи передаются сообщения, содержащие только семь букв: А, Б, В, Д, О, Р, Т. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: Б — 01, Д — 001, Р — 100. Какое наименьшее количество двоичных знаков потребуется для кодирования слова ВОДОВОРОТ? **Примечание.** Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

19. По каналу связи передаются сообщения, содержащие только заглавные русские буквы. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: А — 000, Б — 01, В — 1101, Г — 111, Д — 0010, Е — 100. Какое наименьшее количество двоичных знаков потребуется для кодирования слова КОКОС?

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

20. По каналу связи передаются сообщения, содержащие только семь букв: А, Б, Г, И, М, Р, Я. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: А — 010, Б — 011, И — 10. Какое **наименьшее** количество двоичных знаков потребуется для кодирования слова ГРАММ?

Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

21. По каналу связи передаются сообщения, содержащие только семь букв: А, Б, В, Д, Е, И, Н. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: А — 110, Б — 01, И — 000. Какое наименьшее количество двоичных знаков потребуется для кодирования слова ВВЕДЕНИЕ?

Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

22. 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

ege.sdamgia.ru

Какой набор букв закодирован двоичной строкой 1100000100110?

23. 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 каждому представлению дописывается сумма его элементов по модулю 2 (например, если передаём 23, то получим последовательность 0010100110). Определите, какое число передавалось по каналу в виде 01100010100100100110.

24. 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

ДОМАШНЕЕ ЗАДАНИЕ:

1. В составе поезда 16 вагонов. Среди них есть вагоны купейные и плацкартные. Сообщение о том, что ваш знакомый приезжает в купейном вагоне несет 2 бита информации. Определите, сколько в поезде купейных вагонов?
2. В корзине 8 белых грибов и 24 подосиновика. Сколько информации в сообщении, что из корзины взяли наугад белый гриб?
3. Соревновались в стрельбе по мишеням два стрелка – новичок и мастер спорта. Болельщики делали на них ставки 1 к 15. Сколько информации получили болельщики, когда победил новичок?
4. В классе 30 человек. За контрольную работу по математике получено 6 пятёрок, 15 четвёрок, 8 троек и 1 двойка. Какое количество информации в сообщении о том, что Иванов получил четвёрку?
5. В коробке 10 белых шаров, 20 красных, 30 синих и 40 зеленых. Определить количество информации в сообщении о вытаскивании наугад белого шара, красного шара, синего шара, зеленого шара, шара любого цвета.
6. В озере живут караси и окуни. Подсчитано, что карасей 1500, а окуней 500. Сколько информации содержится в сообщениях о том, что рыбак поймал карася, окуня, поймал рыбу?

7. В корзине лежат 32 клубка красной и черной шерсти. Среди них 4 клубка красной шерсти. Сколько информации содержится в сообщении, что достали один из клубков?
8. В классе 30 человек. За контрольную работу по информатике получено 15 пятерок, 6 четверок, 8 троек и 1 двойка. Какое количество информации несет сообщение о том, что Андреев получил пятерку?
9. В пруду живут 8000 карасей, 2000 щук и 40 000 пескарей.
    1. Определите вероятность улова каждого вида рыб;
    2. Определите количество информации, полученной рыбаком при улове карася;
    3. Определите количество информации, полученной рыбаком при улове щуки;
    4. Определите количество информации, полученной рыбаком при улове пескаря.
10. В ящике лежат 36 красных и несколько зеленых яблок. Сообщение «Из ящика достали зеленое яблоко» несет 2 бита информации. Сколько яблок в ящике?
11. По каналу связи передаются сообщения, содержащие только пять букв: П, И, Л, О, Т. Для передачи используется двоичный код, удовлетворяющий условию Фано. Для буквы И используется кодовое слово 1; для буквы О используется кодовое слово 01.

    Какова минимальная общая длина кодовых слов для всех пяти букв? Примечание: условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.
12. По каналу связи передаются сообщения, содержащие только четыре буквы: Р, Е, К, А; для передачи используется двоичный код, удовлетворяющий условию Фано. Для букв А, Р, Е используются такие кодовые слова: А: 111, Р: 0, Е: 100.

Укажите кратчайшее кодовое слово для буквы К. Если таких кодов несколько, укажите код с **наименьшим** числовым значением.

Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

13. По каналу связи передаются сообщения, содержащие только восемь букв: А, Б , Г , Е , И , М, Р , Т . Для передачи и спользуется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны:

Буква	Кодовое слово
А	0101
Б	1000
Г	
Е	011

Буква	Кодовое слово
И	00
М	0100
Р	11
Т	1001

ege.sdamgia.ru ege.sdamgia.ru

Укажите кратчайшее кодовое слово для буквы Г. Если таких кодов несколько, укажите код с **наименьшим** числовым значением.

Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

14. По каналу связи передаются сообщения, содержащие только четыре буквы: А, Б, В, Г; для передачи используется двоичный код, удовлетворяющий условию Фано. Для букв А, Б, В используются такие кодовые слова: А — 0; Б — 110; В — 101.

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

Примечание. Условие Фано означает, что никакое кодовое слово не является началом другого кодового слова. Это обеспечивает возможность однозначной расшифровки закодированных сообщений.

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

16. По каналу связи передаются сообщения, содержащие только семь букв: А, Б, Г, И, М, Р, Я. Для передачи используется двоичный код, удовлетворяющий условию Фано. Кодовые слова для некоторых букв известны: А — 010, Б — 011, Г — 100. Какое **наименьшее** количество двоичных знаков потребуется для кодирования слова МАГИЯ?

    Примечание. Условие Фано означает, что ни одно кодовое слово не является началом другого кодового слова.

17. 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
100	110	011	01	10

ege.sdamgia.ru

Какой набор букв закодирован двоичной строкой 1000110110110? Все буквы в последовательности — разные.