Код Хаффмана

Некоторое сообщение содержит только буквы А, Б, В, Г, Д, причём известно их количество: **А — 179, Б — 89, В — 72, Г — 53 и Д — 50.**

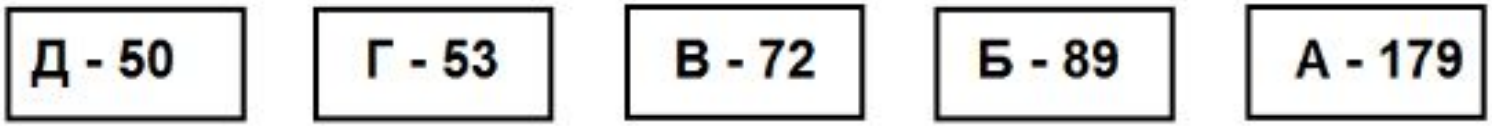
Сколько бит содержит оптимальный префиксный код данного сообщения?

Решение.

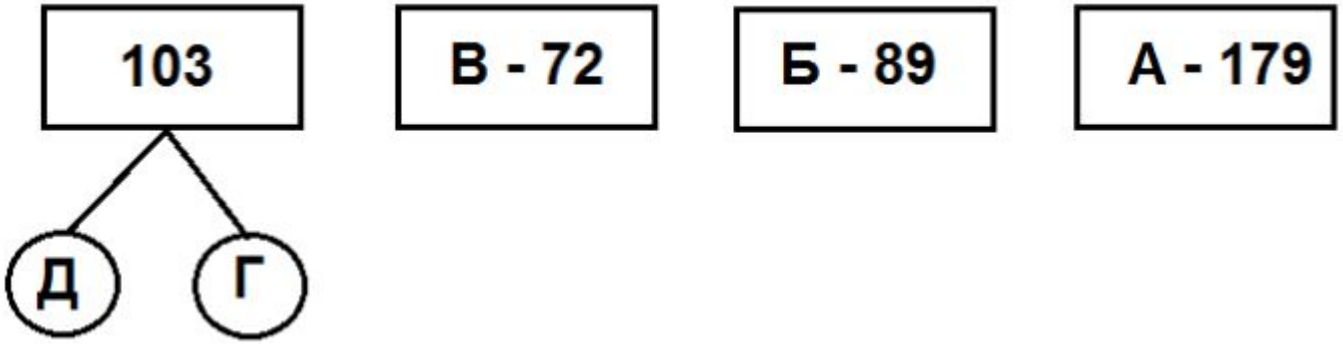
В 1952 году создал алгоритм префиксного кодирования с минимальной избыточностью (известный как алгоритм или код Хаффмана).

Эффективное кодирование по Хаффману состоит в представлении наиболее вероятных (часто встречающихся) букв двоичными кодами наименьшей длины, а менее вероятных -- кодами большей длины (если все кодовые слова меньшей длины уже исчерпаны). Это делается таким образом, чтобы средняя длина кода на букву исходного сообщения была минимальной.

1. Сначала расположим буквы по увеличению их количества в сообщении.



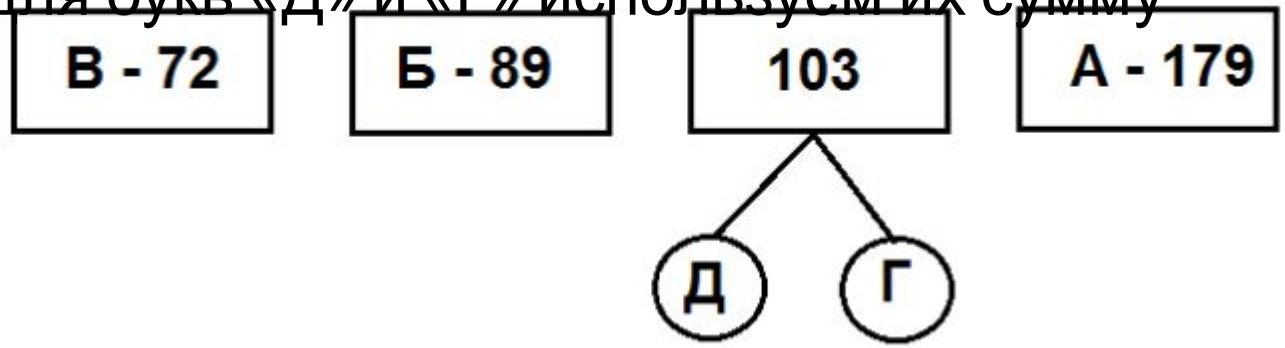
2. Затем берём две первые – это листья дерева, а в узел, с которым они св



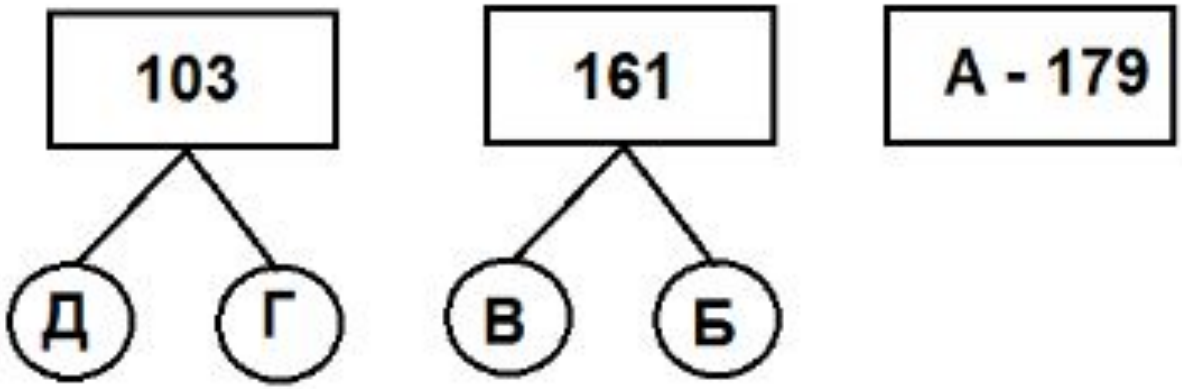
Т.е. буквы, которые встречаются реже всего , получают самый длинный код.

3. Снова располагаем буквы по возрастанию, **но**

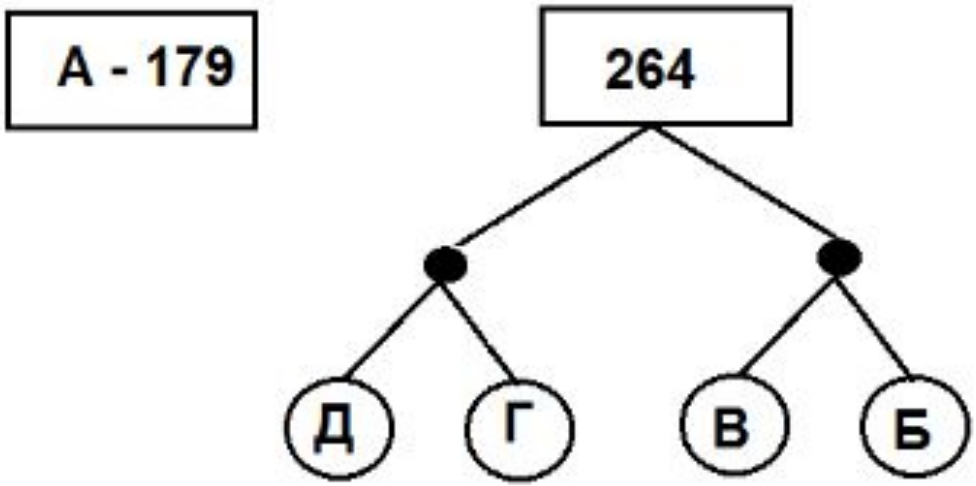
для букв «Д» и «Г» используем их сумму



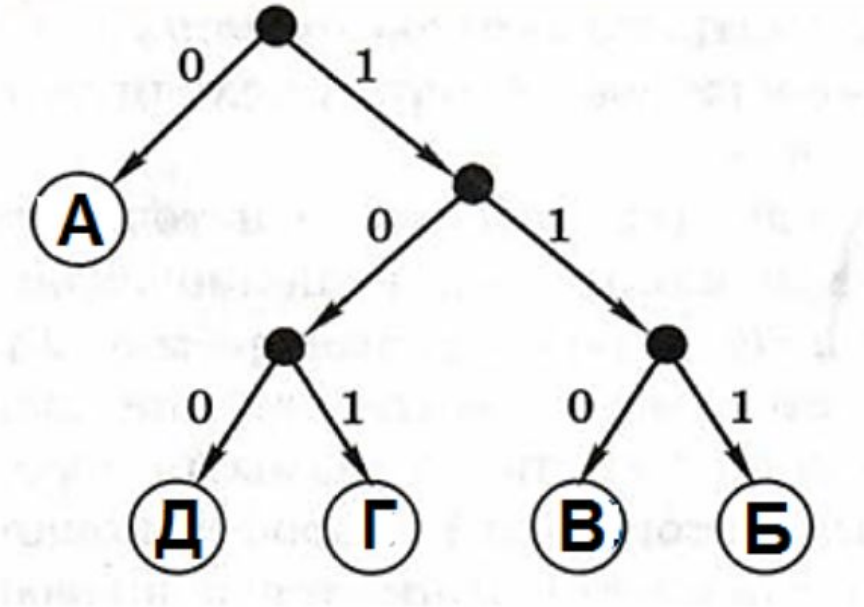
4. Повторяем ту же процедуру для букв «В» и «Б»,



5. Объединяем пары и снова с



6. Объединяем букву «А» с



У стрелок, которые идут влево от узла, пишем код 0,

а у идущих вправо – код 1.

7. Считаем сумму: 1*179 + 3*264 = 179+792 = 971

ОТВЕТ:
971

Самостоятельно

1. Некоторое сообщение содержит только буквы **Ю, В, Е, Л, И, Р,**

известно их количество: Ю—9, В—15, Е—21, Л—17, И—20, Р—18.

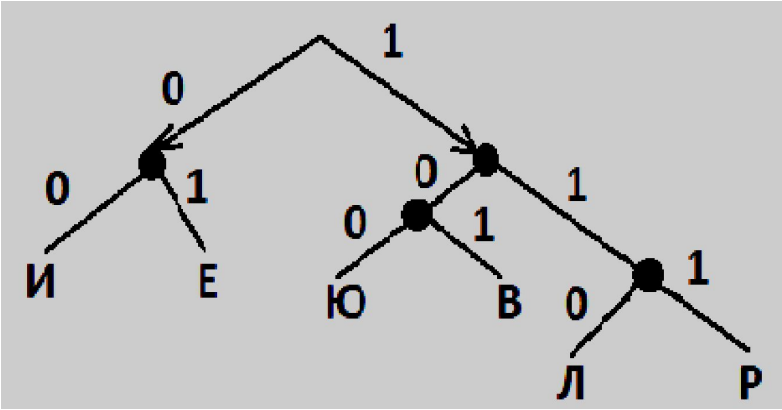
Сколько бит содержит оптимальный префиксный код данного сообщения?

2. Некоторое сообщение содержит только буквы **Л, А, Ф, Е, Т,** известно их количество: Л—18, А—26, Ф—1, Е—31, Т—24. Сколько бит содержит оптимальный префиксный код данного сообщения?

ОТВЕТЫ

№1

.

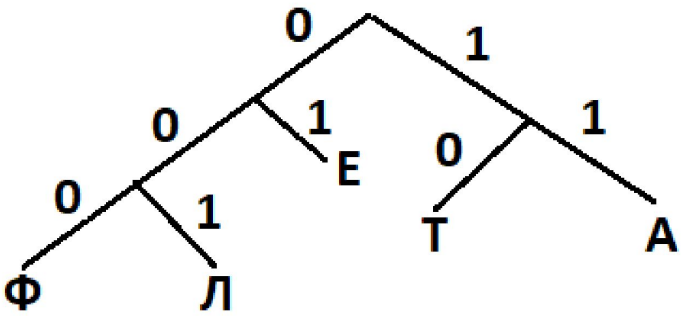


Е-01
И-00
Ю-10
0
В -101
Л- 110
Р- 111

$$\sum = 3 * (9 + 15 + 17 + 18) + 2*(21 + 20) = 3*59 + 2*41 = \mathbf{259}$$

№2

.



Е-01
Т-10
А-11
Ф-000
Л- 001

$$\sum = (3*(1+18) + 2*(31+24+26) = 57+162 = \mathbf{219}$$