

## ОБРАЗЕЦ РЕШЕНИЯ ЗАДАЧИ № 4

### *Код Хаффмана*

При построении эффективных кодов с основанием  $m$  ( $m > 2$ ) неопределенность становится еще больше. От указанного недостатка свободна методика Хаффмана. Она гарантирует однозначное построение кода с наименьшим для данного распределения вероятностей средним числом символов на букву. Для двоичного кода методика сводится к следующему.

Буквы алфавита сообщений вписываются в основной столбец в порядке убывания вероятностей. Две последние буквы объединяются в одну, которой приписывается суммарная вероятность. Снова выбираем две минимальных вероятности из вероятностей букв, не участвовавших в объединении, и полученной суммарной вероятности и объединяем их. Процесс продолжается до тех пор, пока не получим единственную вероятность равную 1.

Мы построили кодовое дерево. Из точки, соответствующей вероятности 1, направляются две ветви, причем ветви с большей вероятностью присваивается символ 1, а с меньшей - 0. Используя такое последовательное ветвление, получим код каждого символа.

#### ***Особенности систем эффективного кодирования:***

1. Различие в длине кодовых комбинаций. Если моменты снятия информации с источника не управляемы (например при непрерывном приеме информации с ЗУ на магнитной ленте), кодирующее устройство через равные промежутки времени выдает комбинации различной длины. Поскольку линия связи используется эффективно только в том случае, когда символы поступают в нее с постоянной скоростью, то на выходе кодирующего устройства должно быть предусмотрено буферное устройство ("упругая" задержка). Оно запасает символы по мере их поступления и выдает в линию связи с постоянной скоростью. Аналогичное устройство необходимо и на приемной стороне.

2. Возникновение задержки при передаче информации. Наибольший эффект достигается при кодировании длинными блоками, а это приводит к необходимости накапливать буквы, прежде чем сопоставить им определенную последовательность символов. При декодировании задержка возникает снова. Общее время задержки может быть велико, особенно при появлении блока, вероятность которого мала. Это следует учитывать при выборе длины кодируемого блока.

3. Специфическое влияние помех на достоверность приема. Одиночная ошибка может перевести передаваемую кодовую комбинацию в другую, не равную ей по длительности. Это повлечет за собой неправильное декодирование целого ряда последующих комбинаций, которые называют треком ошибки. Специальными методами трек ошибки стараются свести к минимуму.

### ***Примеры решения задач***

**110)** Простейший дискретный источник ( $n=5$ ) описывается схемой:

$X_5$	$X_4$	$X_2$	$X_3$	$X_1$
0,341	0,289	0,187	0,171	0,012

Закодировать сообщения источника кодом Хаффмана. Найти среднюю и минимальную длину кодового слова.

Решение. Для этого, среди данных  $n$  сообщений выбираем два с минимальными вероятностями и объединяем их в одно с суммарной вероятностью. Далее рассматриваем  $n-1$  сообщение. Среди них также выделяем два сообщения с наименьшей вероятностью и находим их





***Опять подтвердился тот факт, что при укрупнении источника средняя длина ближе к минимальной.***

Кроме того, можно сделать вывод, что коды Шеннона-Фано по оптимальности уступают кодам Хаффмана, т.к. средняя длина слова, закодированного по Хаффману несколько меньше, чем по Шеннону-Фано: 2,030 и 2,046 соответственно для укрупненных источников. Для не укрупненных источников эти значения получились одинаковы -2,183.