

КОДЫ И КОДИРОВАНИЕ ОСНОВНЫЕ ПОНЯТИЯ

Кодирование - преобразование дискретного сообщения в дискретный сигнал, осуществляемое по определенному правилу.

Восстановление дискретного сообщения по сигналу на выходе дискретного канала, осуществляемое с учетом правил кодирования, называется **декодированием**.

Код - есть совокупность условных сигналов, обозначающих дискретные сообщения.

Кодовая последовательность- комбинация представление дискретного сигнала.

Целями кодирования сообщений обычно являются:

- 1) передача по общему каналу связи нескольких или многих сообщений для кодового разделения сигналов;
- 2) повышение помехоустойчивости и достоверности передачи сообщений;
- 3) более экономное использование полосы частот канала связи, т.е. уменьшение избыточности;
- 4) уменьшение стоимости передачи и хранения сообщений;
- 5) обеспечение скрытности передачи и хранения информации;
- 6) преобразование любой информации независимо от ее происхождения и назначения в единую систему символов;
- 7) приведение исходных символов в соответствие с характеристиками канала связи.

Основание кода X - это количество признаков или число букв (цифр).

Кодовая комбинация, составленная из n символов или n элементов, называется кодовым словом, имеющим длину n или число разрядов n .

Если длина всех кодовых комбинаций одинакова, то такие коды называют равномерными **комплектными**.

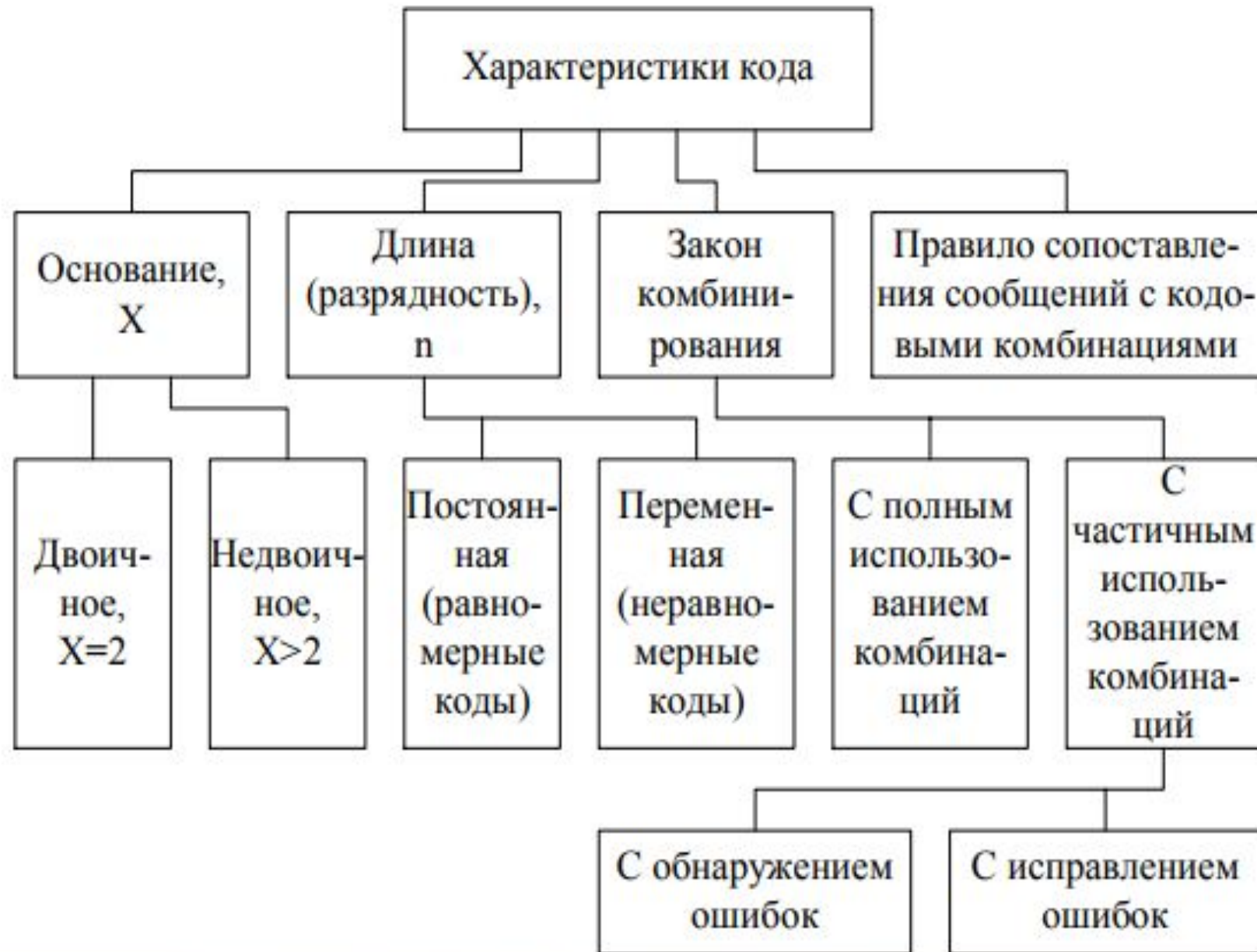
Например, код 001, 011, 101 является комплектным, а код 1, 11, 101 - **некомплектным**.

По способу образования кодовых комбинаций коды разделяются на числовые и нечисловые.

В **числовых кодах**, получивших название цифровых, кодовые комбинации образуют ряд возрастающих по весу чисел, определяемый системой счисления. Они применяются в системах измерений, контроля, ЭВМ.

Нечисловые (невзвешенные) коды не имеют систем счисления и применяются в системах управления и телеуправления, где команды и сигналы независимы.

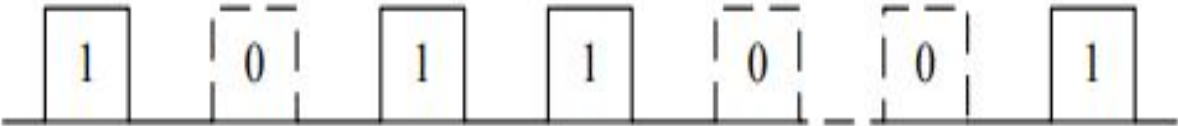
ХАРАКТЕРИСТИКИ КОДА



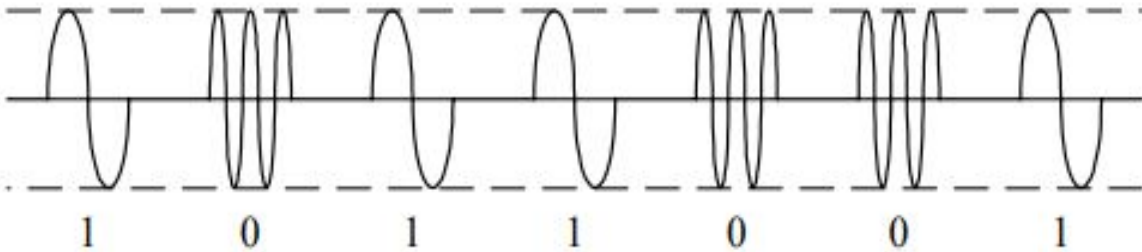
Импульсные признаки, используемые для передачи двоичных кодов

Символ	Амплитудные			Временные			Полярные	Частотные		Фазовые
1										
0										

Последовательная передача кодовой комбинации
видеоимпульсами

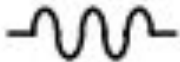
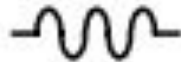


Последовательная передача кодовой комбинации
радиоимпульсами



Для передачи кодовых комбинаций параллельно во времени каждому разряду присваивается своя частота. Однако признаки у каждого разряда должны быть не частотными, а амплитудными или временными.

Параллельная передача кодовых комбинаций

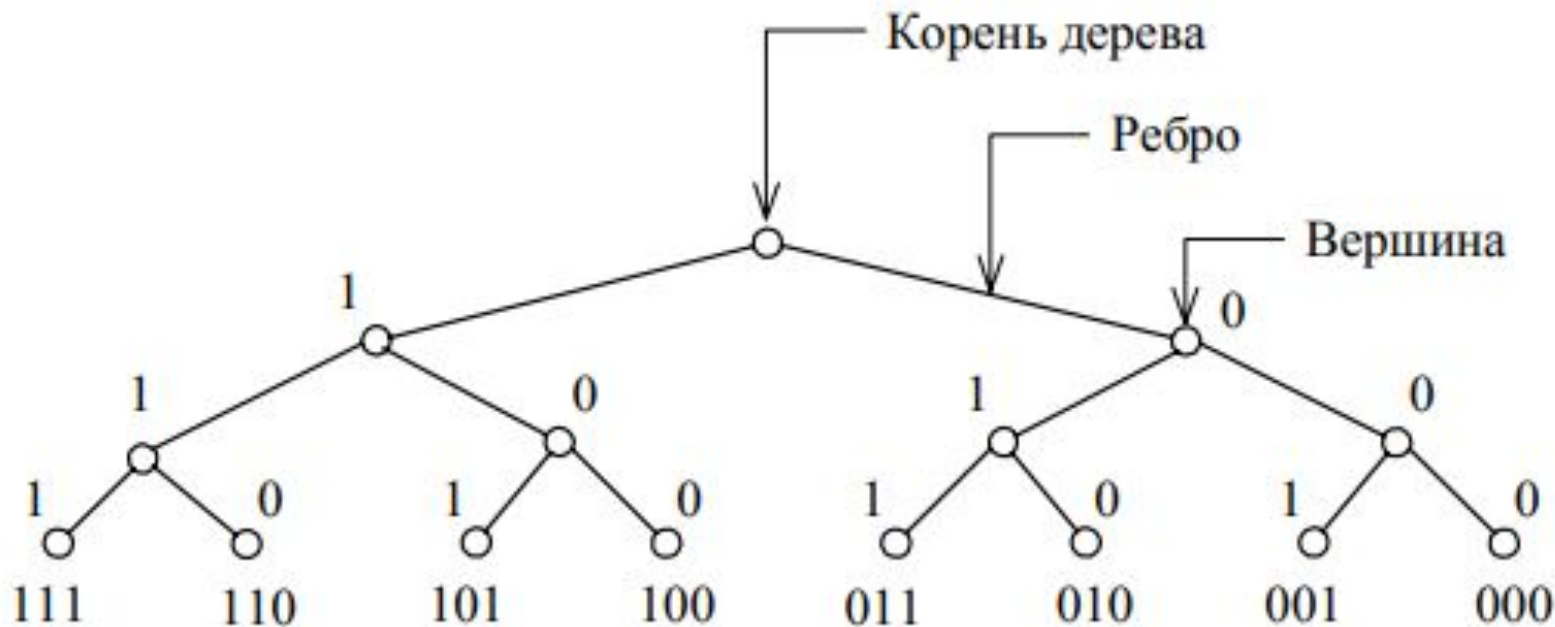
Номер разряда	Частота	Номер кодовой комбинации и время её передачи	
		1- t_1	2- t_2
1	f_1 	1 	1 
2	f_2 	0 —	1 
3	f_3 	0 —	1 
4	f_4 	1 	0 —

Первая кодовая комбинация 1001 передается в течение первого интервала времени t_1 частотами f_1 и f_4 , посылаемыми одновременно, а вторая 1110 передается в течение второго интервала времени одновременной посылкой частот f_1 , f_2 и f_3 .

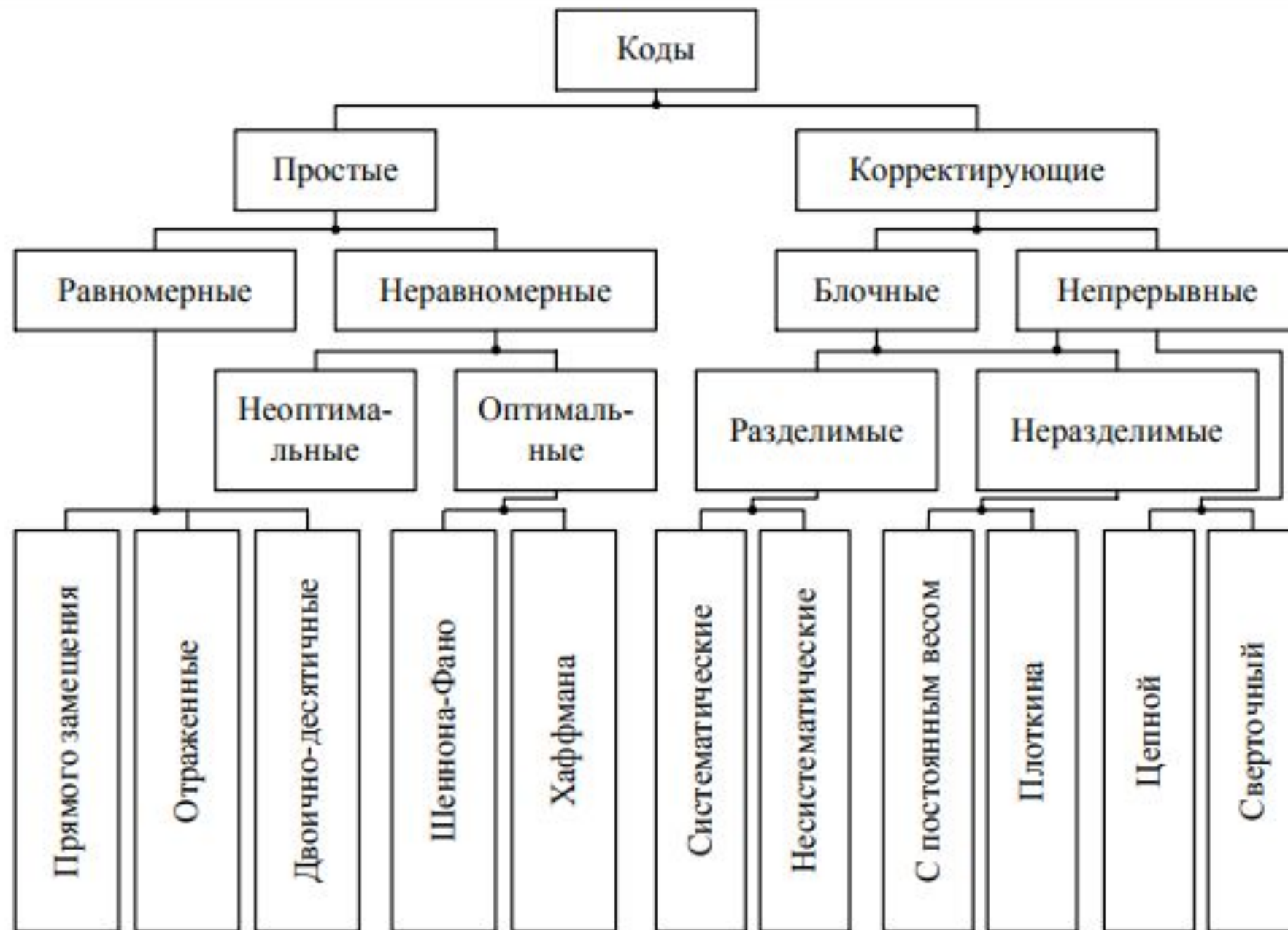
ГРАФИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ КОДА

Точки графа называются **вершинами**, а соединяющие их линии **ребрами**. Начальная вершина, от которой начинается расхождение ребер, называется **корнем дерева**, а число ребер, которое надо пройти от корня к некоторой вершине **порядком этой вершины**. Максимальное число ребер, которые могут выходить из каждой вершины дерева, равно основанию кода, а максимальный порядок вершин, которое оно содержит, равен максимальной длине кодовых комбинаций. Значения разрядов комбинации, приписываемой каждой вершине, соответствующей направлениям движения по ребрам от корня дерева к данной вершине. Ребра, ведущие от корня к вершинам первого порядка, определяют значение первого слева разряда комбинации; ребра, соединяющие вершины первого и второго порядков, дают значение второго разряда комбинации, и т.д.

Кодовое дерево для двоичного трехразрядного кода



КЛАССИФИКАЦИЯ ДВОИЧНЫХ КОДОВ



Корректирующий код называется **блочным**, если каждая его комбинация имеет ограниченную длину, и **непрерывным**, если его комбинация имеет неограниченную, а точнее, полубесконечную длину.

Коды в зависимости от методов внесения избыточности подразделяются на **разделимые** и **неразделимые**.

В **разделимых кодах** четко разграничена роль отдельных символов. Одни символы являются **информационными**, другие являются **проверочными** и служат для обнаружения и исправления ошибок.

Разделимые блочные коды называются обычно n, k - кодами, где n - длина кодовых комбинаций, k - число информационных символов в комбинациях.

Неразделимые коды не имеют четкого разделения кодовой комбинации на информационные и проверочные символы.

Разделимые блочные коды делятся на **систематические** и **несистематические**.

Несистематические коды строятся таким образом, что проверочные символы определяются как сумма подблоков длины 1, на которые разделяется блок информационных символов.

У **систематических кодов** проверочные символы определяются в результате проведения линейных операций над определенными информационными символами.

ОСНОВНЫЕ ХАРАКТЕРИСТИКИ ДВОИЧНЫХ КОДОВ

Весом кода w называется количество единиц в кодовой комбинации. Например, для кодовой комбинации 1011110 вес кода $w = 5$.

Кодовое расстояние - это минимальное число элементов, которыми любая кодовая комбинация отличается от другой (по всем парам кодовых слов).

Например, для кода, состоящего из комбинаций 1100, 1000, 1011, 1101, $d_{\min}=1$, так как $\overline{1100} \oplus 1101 = 0001$.

Весовая характеристика кода $F(w)$ - число кодовых комбинаций определенного веса w . Например, для кода, представленного комбинациями 00001 ($w = 1$), 11010 ($w = 3$), 10110 ($w = 3$), 11110 ($w = 4$), имеем $F(1) = 1$, $F(3) = 2$, $F(4) = 1$, т.е. код состоит из одного кодового слова веса 1, двух слов веса 3 и одного слова веса 4.

Корректирующие коды имеют и некоторые дополнительные характеристики.

Абсолютная избыточность кода определяется числом проверочных символов (r), т.е. количеством разрядов, отводимых для коррекции ошибок.

Относительная избыточность кода (R) есть отношение числа проверочных символов к длине кода: $R = r / n$. В общем случае относительную избыточность рассчитывают по формуле $R = 1 - \log_2 N_p / \log_2 N$, где N_p - число кодовых комбинаций, используемых для передачи сообщений (рабочая мощность кода); N - полное число кодовых комбинаций (мощность кода).

ПРОСТЫЕ ДВОИЧНЫЕ КОДЫ

Непомехозащищенным кодом называется код, в котором искажение одного разряда кодовой комбинации не может быть обнаружено

Двоичный код на все сочетания

Кодовые комбинации этого кода соответствуют записи натурального ряда чисел в двоичной системе счисления. Вес разряда кода определяется из выражения:

$$q_i = 2^{i-1},$$

где $i = 1, 2, 3, \dots, n$.

Общее число комбинаций

$$N = 2^n.$$

Единично-десятичный код

Каждый разряд десятичного числа записывается в виде соответствующего числа единиц. При этом разряды разделяются интервалами. Например, $24 \rightarrow 11\ 1111$. Этот код неравномерный. Для преобразования в равномерный необходимо в каждом разряде слева дописать столько нулей, чтобы общее число символов в каждом десятичном разряде было равно 9. Например, $24 \rightarrow 000000011\ 000001111$.

Двоично-десятичный код

Каждый разряд десятичного числа записывается в виде комбинации двоичного кода.

Например в коде 8-4-2-1 (код на все сочетания) $576 \rightarrow 010101110110$

Числоимпульсный

Иногда его называют единичным (или унитарным) кодом. Кодовые комбинации отличаются друг от друга числом единиц. Например $25 \rightarrow 11\ 1111$ или $14 \rightarrow 1\ 1111$

Код Грея

Эти коды применяются в устройствах, преобразующих линейные и угловые перемещения в кодовые комбинации.

Начиная с младшего разряда веса разрядов запишутся следующим образом: 1, 3, 7, 15, 31,....

Чтобы прочесть число в коде Грея, под каждым разрядом записывают его десятичный эквивалент, старший значащий разряд берется со знаком плюс, перед остальными значащими разрядами знаки чередуются. Например, перевод комбинации кода Грея 101111 и 010011 в десятичный код производится следующим образом:

1	0	1	1	1	1
63	31	15	7	3	1

$$= 63 - 15 + 7 - 3 + 1 = 53;$$

0	1	0	0	1	1
63	31	15	7	3	1

$$= 31 - 3 + 1 = 29.$$

Преобразование кода Грея в двоичный код

$$a_i = \sum_{l=n}^i b_l (\text{mod } 2).$$

a_i -число в двоичном коде, b_i -число в коде Грея

Пример: преобразование кодовой комбинации 101111, записанной в коде Грея, в двоичный код:

$$\begin{aligned} 101111 &= b_6 b_5 b_4 b_3 b_2 b_1 \rightarrow a_6 a_5 a_4 a_3 a_2 a_1 = b_6 (b_6 \oplus b_5)(b_6 \oplus b_5 \oplus b_4) \times \\ &\times (b_6 \oplus b_5 \oplus b_4 \oplus b_3)(b_6 \oplus b_5 \oplus b_4 \oplus b_3 \oplus b_2)(b_6 \oplus b_5 \oplus b_4 \oplus b_3 \oplus b_2 \oplus b_1) = \\ &= 1(1 \oplus 0)(1 \oplus 0 \oplus 1)(1 \oplus 0 \oplus 1 \oplus 1)(1 \oplus 0 \oplus 1 \oplus 1 \oplus 1)(1 \oplus 0 \oplus 1 \oplus 1 \oplus 1 \oplus 1) = 110101 \end{aligned}$$

Проверка правильности преобразования

$$\begin{array}{cccccc} 1 & 0 & 1 & 1 & 1 & 1 \\ 63 & 31 & 15 & 7 & 3 & 1 \end{array} \rightarrow \begin{array}{cccccc} 1 & 1 & 0 & 1 & 0 & 1 \\ 32 & 16 & 8 & 4 & 2 & 1 \end{array}$$

$$63 - 15 + 7 - 3 + 1 \rightarrow 32 + 16 + 4 + 1$$

$$53 = 53$$

2 способ преобразования в код Грея.

Обычный двоичный код преобразуется в код Грея путем суммирования по модулю 2 данной комбинации с такой же, но сдвинутой вправо на один разряд.

Например, преобразование двоичных чисел 110011 и 111011 в код Грея

$$\begin{array}{r} \oplus 110011 \\ 110011 \\ \hline 101010 \end{array}$$

$$\begin{array}{r} \oplus 111011 \\ 111011 \\ \hline 100110 \end{array}$$

$$1\ 1\ 0\ 0\ 1\ 1 \rightarrow 1\ 0\ 1\ 0\ 1\ 0; 32+16+2+1 \rightarrow 63-15+3; 51=51.$$

$$\begin{array}{cccccc} 32 & 16 & 8 & 4 & 2 & 1 \\ 1 & 1 & 0 & 0 & 1 & 1 \end{array} \quad \begin{array}{cccccc} 63 & 31 & 15 & 7 & 3 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{array}$$

$$1\ 1\ 1\ 0\ 1\ 1 \rightarrow 1\ 0\ 0\ 1\ 1\ 0; 32+16+8+2+1 \rightarrow 63-7+3; 59=59.$$

$$\begin{array}{cccccc} 32 & 16 & 8 & 4 & 2 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \end{array} \quad \begin{array}{cccccc} 63 & 31 & 15 & 7 & 3 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{array}$$

ОПТИМАЛЬНЫЕ КОДЫ

Оптимальные по длине коды относятся к неравномерным непомехозащищенным кодам. Оптимальным кодом считается код, имеющий минимальную среднюю длину кодового слова

$$L = \sum_{i=1}^n \mu_i P(x_i),$$

где μ_i — длина кодового слова, сопоставляемая сообщению x_i ; $P(x_i)$ — вероятность появления этого сообщения.

В оптимальном коде энтропия на символ должна быть максимальной, а это возможно в том случае, когда вероятности появления единиц и нулей $P(1)$ и $P(0)$ приблизительно одинаковы.

Код Шеннона-Фано

Все сообщения выписываются в порядке убывания их вероятностей

Записанные так сообщения затем разбиваются на две по возможности равновероятные подгруппы. Всем сообщениям первой подгруппы присваивают цифру 1 в качестве первого кодового символа, а сообщениям второй подгруппы цифру 0. Аналогичное деление на подгруппы продолжается до тех пор, пока в каждую подгруппу не попадает по одному сообщению.

Построение кода Шеннона-Фано

x_i	$P(x_i)$	Разбиение сообщений на подгруппы					Код	μ_i	L_{x_i}
x_1	0,35	1	1				11	2	0,70
x_2	0,15	1	0				10	2	0,30
x_3	0,13	0	1	1			011	3	0,39
x_4	0,09	0	1	0			010	3	0,27
x_5	0,09	0	0	1	1		0011	4	0,36
x_6	0,08	0	0	1	0		0010	4	0,32
x_7	0,05	0	0	0	1		0001	4	0,20
x_8	0,04	0	0	0	0	1	00001	5	0,20
x_9	0,02	0	0	0	0	0	00000	5	0,10

Энтропия сообщений

$$H(X) = -\sum_{i=1}^9 P(x_i) \log P(x_i) = -(0,35 \log 0,35 + 0,15 \log 0,15 + 0,13 \log 0,13 + \\ + 0,09 \log 0,09 + 0,09 \log 0,09 + 0,08 \log 0,08 + 0,05 \log 0,05 + 0,04 \log 0,04 + \\ + 0,02 \log 0,02) \cong 2,75 \frac{\text{бит}}{\text{сообщение}}.$$

Средняя длина кодового слова

$$L = \sum_{i=1}^9 L_{x_i} = 0,70 + 0,30 + 0,39 + 0,27 + 0,36 + 0,32 + 0,20 + 0,20 + 0,10 = 2,84.$$

Среднее число нулей

$$L(0) = 0,15 + 0,13 + 0,18 + 0,18 + 0,24 + 0,15 + 0,16 + 0,10 = 1,29.$$

Среднее число единиц

$$L(1) = 0,70 + 0,15 + 0,26 + 0,09 + 0,18 + 0,08 + 0,05 + 0,04 = 1,55.$$

Вероятность появления нулей

$$P(0) = \frac{L(0)}{L} = \frac{1,29}{2,84} = 0,455.$$

Вероятность появления единиц

$$P(1) = \frac{L(1)}{L} = \frac{1,55}{2,84} = 0,545.$$

Таким образом, получен код, близкий к оптимальному.

Код Хаффмана

Для получения кода Хаффмана все сообщения выписывают в порядке убывания вероятностей. Две наименьшие вероятности объединяют скобкой и одной из них присваивают символ 1, а другой 0. Затем эти вероятности складывают, результат записывают в промежутке между ближайшими вероятностями. Процесс объединения двух сообщений с наименьшими вероятностями продолжают до тех пор, пока суммарная вероятность двух оставшихся сообщений не станет равной единице. Код для каждого сообщения строится при записи двоичного числа справа налево путем обхода по линиям вверх направо, начиная с вероятности сообщения, для которого строится код.

Построение кода Хаффмана

x_i	$P(x_i)$	Объединение сообщений							Код
x_1	0,35	0,35	0,35	0,35	0,35	0,35	0,37	0,63	11
x_2	0,15	0,15	0,15	0,17	0,20	0,28	0,35	0,37	101
x_3	0,13	0,13	0,13	0,15	0,17	0,20	0,28		100
x_4	0,09	0,09	0,11	0,13	0,15	0,17			010
x_5	0,09	0,09	0,09	0,11	0,13				001
x_6	0,08	0,08	0,09	0,09					000
x_7	0,05	0,06	0,08						0110
x_8	0,04	0,05							01111
x_9	0,02								01110

КОРРЕКТИРУЮЩИЕ КОДЫ

Помехоустойчивыми (корректирующими) называются коды, позволяющие обнаружить и исправить ошибки в кодовых комбинациях. Отсюда и деление кодов на две большие группы:

1) коды с обнаружением ошибок; 2) коды с обнаружением и исправлением ошибок.

Коды с обнаружением ошибок

Общее число кодовых комбинаций в данном коде

$$N = C_n^m = \frac{n!}{m!(n-m)!},$$

где m — число единиц в слове длиной n .

Код C_4^2 правильность принятых кодовых комбинаций определяется путем подсчета количества единиц, и если их число отличается от m , то в передаче произошла ошибка. Необнаруженная ошибка имеет место, если произошло искажение типа смещения, т. е. когда единица переходит в нуль, а нуль в единицу.

Распределительный код

Это также разновидность кода с постоянным весом, равным единице.

Число кодовых комбинаций в данном коде

$$N = C_n^l = n.$$

Код с постоянным числом единиц

Номер кодовой комбинации	Код C_4^2	Код C_6^1
1	0011	000001
2	0119	000010
3	1100	000100
4	1001	001000
5	1010	010000
6	0101	100000

Кодовая комбинация при n=6 представлена в табл. (столбец 3). Сложение по модулю 2 двух комбинаций показывает, что они отличаются друг от друга на кодовое расстояние d=2. В системах телемеханики этот код нашел широкое применение из-за простой реализации.

КОД С ПРОВЕРКОЙ НА ЧЕТНОСТЬ

Код с проверкой на четность образуется путем добавления к передаваемой комбинации одного контрольного символа (0 или 1), так чтобы общее количество единиц в передаваемой комбинации было четным.

Код с проверкой на чётность

Информационные символы k	Контрольные символы r	Код с проверкой на чётность $n=k+r$
01101	1	011011
01111	0	011110

Такой код состоит из $N = 2^k$ комбинаций и имеет минимальное кодовое расстояние $d_{\min} = 2$. Коэффициент избыточности кода с проверкой на четность зависит от числа информационных символов:

$$K_{\text{изб}} = 1 - \frac{k}{k+1} = 1 - \frac{k}{n}.$$

Обнаружение ошибок на приемной стороне осуществляется подсчетом количества единиц в принятой комбинации, и если оно четное, считается что искажений нет. Тогда контрольный символ отбрасывается и исходная k -разрядная комбинация выдается получателю информации. В противном случае кодовая комбинация бракуется. Данный код может обнаружить любое нечетное число искажений.

КОД С ПРОВЕРКОЙ НА НЕЧЕТНОСТЬ

Особенностью кода является то, что каждая комбинация содержит нечетное число единиц (табл.). К проверке этого факта и сводится обнаружение ошибок в кодовых комбинациях. Другие основные характеристики кода такие же, как и у кода с одной проверкой на четность.

Код с проверкой на нечётность

Информационные символы k	Контрольные символы r	Полная кодовая комбинация $n=k+r$
10101	0	101010
11101	1	111011

КОД С ДВУМЯ ПРОВЕРКАМИ НА ЧЕТНОСТЬ

Данный код является разновидностью кода с проверкой на четность и образуется путем добавления к передаваемой комбинации двух контрольных символов (табл.). Первый символ добавляет 0 или 1 так, чтобы общее количество единиц в передаваемой комбинации было четным, а второй символ добавляет 0 или 1 так, чтобы количество единиц в нечетных разрядах передаваемой комбинации было четным.

Обнаружение ошибок осуществляется подсчетом количества единиц в информационной части кодовой комбинации и первом контрольном разряде, а также в нечетных разрядах информационной части и втором контрольном символе, и если оно четное в первом и втором случае, то считается, что искажений нет. В противном случае принятая кодовая комбинация бракуется. Данный код позволяет обнаруживать все нечетные искажения и искажения в смежных разрядах, т.е. стоящих рядом.

Код с двумя проверками на чётность

Информационные символы k	Контрольные символы		Полная кодовая комбинация
	$r1$	$r2$	
101011	0	1	10101101
111101	1	0	11110110
100010	0	0	10001000
101010	1	1	10101011

КОД С ПОВТОРЕНИЕМ

Этот код имеет две разновидности. В одной из них имеет место m - кратное повторение комбинаций простого кода

$$a_1, a_2 \dots a_k :$$

$$\underbrace{a_1 a_2 \dots a_k}_1 \underbrace{a_1 a_2 \dots a_k}_2 \dots \underbrace{a_1 a_2 \dots a_k}_m$$

Например, при $m = 3$ кодовая комбинация 1011 в коде с повторением комбинаций будет 1011 1011 1011. Вторая разновидность кода с повторением характеризуется m -кратной передачей каждого разряда (код с повторением элементов кода):

$$\underbrace{a_1 \ a_1 \dots a_1}_{m \text{ раз}} \quad \underbrace{a_2 \ a_2 \dots a_2}_{m \text{ раз}} \quad \dots \quad \underbrace{a_k \ a_k \dots a_k}_{m \text{ раз}}$$

Например, при $m = 3$ кодовая комбинация 1011 в коде с m -кратной передачей каждого разряда будет 111 000 111 111.

Код с повторением имеет длину $n = m \times k$, число контрольных разрядов $r = k(m - 1)$. Избыточность этих кодов равна $(m - 1) m$. Весьма высокая избыточность является недостатком кодов с повторением. Даже при двукратном повторении она составляет 0,5:

$$K_{изб} = 1 - \frac{k}{n} = 1 - \frac{k}{2k} = 0,5.$$

Код имеет минимальное кодовое расстояние $d_{min} = m$ и может использоваться как для обнаружения, так и для исправления ошибок. Для обнаружения ошибок применяют, как правило, код с четным d_{min} , для исправления с нечетным d_{min} . Правильность принятой информации определяется при проведении по-элементного сравнения информационных и контрольных символов, и при наличии хотя бы одного несовпадения вся принятая комбинация бракуется.

Код с повторением позволяет обнаруживать ошибки любой кратности за исключением случаев, когда искажается один информационный символ и все соответствующие ему контрольные, два информационных символа и соответствующие им контрольные. При исправлении ошибок в комбинациях обычно применяется мажоритарный принцип исправления для каждого информационного символа, т.е. за истинное значение информационного символа принимается то, которое большее число раз встречается в этом информационном и соответствующих ему контрольных символах.

При трехкратном повторении мажоритарный принцип реализуется как решение по двум символам из трех, при пятикратном как решение по трем из пяти и т.д.

При увеличении числа повторений увеличивается минимальное кодовое расстояние, соответственно улучшаются корректирующие свойства кода, но значительно увеличивается и избыточность. Поэтому кратность повторений больше трех практически не используется. В условиях коррелированных ошибок обычно применяют первую разновидность кода с повторением (код с повторением комбинаций), имеющую в этом случае более высокую помехоустойчивость. Это обусловлено тем, что входящие в одну проверку на четность разряды достаточно далеко отстоят друг от друга и с малой вероятностью поражаются одним пакетом ошибок.

КОД С ЧИСЛОМ ЕДИНИЦ КРАТНЫМ ТРЕМ

Этот код образуется добавлением к k информационным символам двух дополнительных контрольных символов ($r = 2$), которые должны иметь такие значения, чтобы сумма единиц, посылаемых в линию кодовых комбинаций, была кратной трем. Примеры комбинаций такого кода представлены в табл.

Код с числом единиц, кратным трем

Информационные символы k	Контрольные символы		Полная кодовая комбинация
	$r1$	$r2$	
001000	1	1	00100011
011000	1	0	01100010
011001	0	0	01100100

Он позволяет обнаружить все одиночные ошибки и любое четное количество ошибок одного типа (например, только переход 0 в 1) не обнаруживаются двойные ошибки разных типов (смещения) и ошибки одного типа, кратные трем. На приемной стороне полученную комбинацию проверяют на кратность трем. При наличии такой кратности считают, что ошибок не было, два контрольных знака отбрасывают и записывают исходную комбинацию. Данный код обладает дополнительной возможностью обнаруживать ошибки: если первый контрольный символ равен нулю, то и второй тоже должен быть равен нулю.

ИНВЕРСНЫЙ КОД (КОД С ПОВТОРЕНИЕМ ИНВЕРСИИ)

Это разновидность кода с двукратным повторением. При использовании данного кода комбинации с четным числом единиц повторяются в неизменном виде, а комбинации с нечетным числом единиц 0 в инвертированном. Примеры представления кодовых комбинаций в инверсном коде приведены в табл.

Инверсный код

Информационные символы k	Контрольные символы r	Инверсный код $n=k+r$
111100	111100	111100111100
011100	100011	011100100011
110111	001000	110111001000
111010	111010	111010111010

Прием инверсного кода осуществляется в два этапа. На первом этапе суммируются единицы в первой половине кодовой комбинации. Если их количество окажется четным, то вторая половина кодовой комбинации принимается без инверсии, а если нечетным то с инверсией. На втором этапе обе зарегистрированные комбинации поэлементно сравниваются, и при обнаружении хотя бы одного несовпадения комбинация бракуется. Это поэлементное сравнение эквивалентно суммированию по модулю 2. При отсутствии ошибок в обеих группах символов их сумма равна нулю.

Рассмотрим процесс обнаружения ошибок на следующем примере. Пусть передана последняя кодовая комбинация из табл. Ниже показано суммирование для трех вариантов приема переданной комбинации:

$$1) \oplus \begin{array}{r} 111010 \\ 111010 \\ \hline 000000 \end{array}$$

$$2) \oplus \begin{array}{r} 1\dot{0}1010 \\ 000101 \\ \hline 101111 \end{array}$$

$$3) \oplus \begin{array}{r} 111010 \\ 1\dot{0}1010 \\ \hline 010000 \end{array}$$

В первом варианте принята комбинация 111010111010. В первой половине кодового слова (информационных символах) четное количество единиц, поэтому производится ее суммирование по модулю 2 с не инвертируемыми контрольными символами r , что в результате дает нулевую сумму, т.е. комбинация принята без искажений.

Во втором варианте принята комбинация 101010111010. Подсчитывая количество единиц в информационных символах и замечая, что оно нечетное, контрольные символы инвертируют и суммируют с информационными символами. Присутствие единиц в результате свидетельствует о наличии ошибки, а нуль в этой сумме показывает ее место.

В третьем варианте принята комбинация 111010101010. Поскольку в информационной последовательности четное количество единиц, при проверке контрольные символы суммируются с информационными без инверсии. В этом случае в итоге появляется одна единица. Ее место указывает номер искаженной позиции в принятой последовательности контрольных символов. Таким образом, если при суммировании в результате среди единиц появляется один нуль - ошибка появилась в первой половине принятой кодовой комбинации (в информационных символах) и нуль указывает ее место.

Если в результате среди нулей появляется одна единица - ошибка во второй половине кодовой комбинации (в контрольных символах) и ее место указывает единица. Если в результате суммирования имеется несколько единиц или нулей, это означает, что комбинация принята с несколькими искажениями. Кодовое расстояние инверсного кода равно количеству разрядов исходного кода при $k < 4$ и равно 4 при $k \geq 4$.

Например, при $d=4$ код может обнаруживать двойные ошибки и исправлять одиночные. Обычно этот код используется только для обнаружения ошибок. Он позволяет обнаруживать ошибки любой кратности за исключением таких, когда искажены 2 информационных символа и соответствующие им 2 контрольных, 4 информационных и соответствующие им 4 контрольных и т.д.

Коэффициент избыточности инверсного кода равен 0,5.

КОРРЕЛЯЦИОННЫЙ КОД - КОД С УДВОЕНИЕМ ЧИСЛА ЭЛЕМЕНТОВ

В рассматриваемом коде символы исходного кода кодируются повторно. Правило вторичного кодирования таково: если в исходном кодовом слове на какой-либо позиции стоит 0, в новом помехоустойчивом коде на эту позицию записывается пара символов 01, а если в исходном коде была 1, она записывается как 10.

Например, кодовое слово 1001 в корреляционном коде будет выглядеть следующим образом: 10010110.

Корреляционный код будет всегда иметь вдвое больше элементов, чем исходный. Поэтому его коэффициент избыточности всегда равен 0,5:

$$K_{изб} = 1 - \frac{k}{n} = 1 - \frac{k}{2k} = 0,5.$$

На приеме ошибка обнаруживается в том случае, если в парных элементах содержатся одинаковые символы, т.е. 11 или 00 (вместо 10 и 01).

При правильном приеме вторые (четные) элементы отбрасываются и остается первоначальная комбинация. Код обладает сравнительно высокой помехоустойчивостью, поскольку ошибка не будет обнаружена только в том случае, если будут искажены два рядом стоящие элемента, соответствующие одному элементу исходного кода, т.е. 0 перейдет в 1, а 1 в 0.

Наибольшая эффективность корреляционного кода проявляется при применении его на каналах, у которых вероятность искажения элементов (единиц и нулей) непрерывно меняется и в отдельные интервалы времени существенно различна.

КОД БЕРГЕРА

Контрольные символы в этом коде представляют разряды двоичного числа в прямом или инверсном виде количества единиц или нулей, содержащихся в исходной кодовой комбинации. Число контрольных символов определяется из выражения

$$r = E \log(k + 1),$$

где E - знак округления в большую сторону. Примеры составления комбинаций в коде Бергера из обычного шести- разрядного двоичного кода представлены в табл.

Информационные символы	Контрольные символы				Полная кодовая комбинация n=k+r
	Количество единиц в двоичном коде		Количество нулей в двоичном коде		
	прямом	инверсном	прямом	инверсном	
101011	100	011	010	101	101011100
101011					101011011
101011					101011010
101011					101011101

На приемной стороне подсчитывается число единиц (нулей) в информационной части и сравнивается с контрольной кодовой комбинацией (складывается по модулю 2). При отсутствии ошибок в обеих комбинациях их сумма равна нулю. Ниже показана проверка для шести вариантов приема переданной комбинации. Искаженные символы отмечены точкой.

101011100	$100 \oplus 100 = 000$ - искажений нет.
100̇011100	$011 \oplus 100 = 111$ - искажение обнаружено.
11̇0011100	$100 \oplus 100 = 000$ - искажение не обнаружено.
11̇1̇111100	$110 \oplus 100 = 010$ - искажение обнаружено.
101011101̇	$100 \oplus 101 = 001$ - искажение обнаружено.
1010110̇1̇0	$100 \oplus 010 = 110$ - искажение обнаружено.

Данный код обнаруживает все одиночные и большую часть многократных ошибок

КОДЫ С ОБНАРУЖЕНИЕМ И ИСПРАВЛЕНИЕМ ОШИБОК

Если кодовые комбинации составлены так, что отличаются друг от друга на кодовое расстояние $d \geq 3$, то они образуют корректирующий код, который позволяет по имеющейся в кодовой комбинации избыточности не только обнаруживать, но и исправлять ошибки. Большую группу кодов, исправляющих ошибки, составляют систематические коды.

СИСТЕМАТИЧЕСКИЕ КОДЫ

Систематическими кодами называются блочные (n, k) коды, у которых k (обычно первые) разрядов представляют собой двоичный избыточный код, а последующие r -контрольные разряды сформированные путем линейных комбинаций над информационными. Основное свойство систематических кодов: сумма по модулю 2 двух и более разрешенных кодовых комбинаций также дает разрешенную кодовую комбинацию. Правило формирования кода обычно выбирают так, чтобы при декодировании имелась возможность выполнить ряд проверок на четность для некоторых определенным образом выбранных подмножеств информационных и контрольных символов каждой кодовой комбинации. Анализируя результаты проверок, можно обнаружить или исправить ошибку ожидаемого вида. Информацию о способе построения такого кода содержит проверочная матрица, которая составляется на базе образующей матрицы.

Образующая матрица M состоит из единичной матрицы размерностью $k \times k$ и приписанной к ней справа матрицы дополнений размерностью $k \times r$:

$$r_{d=3} \geq E \log(n + 1) . \quad (2.4)$$

$$r_{d=3} \geq E \log((k + 1) + E \log(k + 1)) . \quad (2.5)$$

Разрядность матрицы дополнений выбирается из выражения (2.4) или (2.5). Причем вес (число ненулевых элементов) каждой строки матрицы дополнений должен быть не меньше чем $d_{\min} - 1$. Проверочная матрица N строится из образующей матрицы следующим образом. Строками проверочной матрицы являются столбцы матрицы дополнений образующей матрицы. К полученной матрице дописывается справа единичная матрица размерностью $g \times g$. Таким образом, проверочная матрица размерностью $r \times k$ имеет вид:

$$M = \left[\begin{array}{cccc|cccc} 1 & 0 & 0 & \dots & 0 & b_{11} & b_{12} & \dots & b_{1r} \\ 0 & 1 & 0 & \dots & 0 & b_{21} & b_{22} & \dots & b_{2r} \\ 0 & 0 & 1 & \dots & 0 & b_{31} & b_{32} & \dots & b_{3r} \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots \\ 0 & 0 & 0 & \dots & 1 & b_{k1} & b_{k2} & \dots & b_{kr} \end{array} \right].$$

$$N = \left[\begin{array}{cccc|cccc} b_{11} & b_{21} & b_{31} & \dots & b_{k1} & 1 & 0 & 0 & \dots & 0 \\ b_{12} & b_{22} & b_{32} & \dots & b_{k2} & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots \\ b_{1r} & b_{2r} & b_{3r} & \dots & b_{kr} & 0 & 0 & 0 & \dots & 1 \end{array} \right].$$

Единицы, стоящие в каждой строке, однозначно определяют, какие символы должны участвовать в определении значения контрольного разряда. При- чем единицы в единичной матрице определяют номера контрольных разрядов.

Получить алгоритм кодирования в систематическом коде всех четырехразрядных кодовых комбинаций, позволяющего исправлять единичную ошибку. Таким образом, задано число информационных символов $k = 4$ и кратность исправления $S = 1$. По выражению (2.5) определим число контрольных символов:

$$r \geq E \log((4 + 1) + E \log(4 + 1)) = 3.$$

Минимальное кодовое расстояние определим из выражения

$$d_{\min} \geq 2S + 1.$$

$$d_{\min} \geq 2 \cdot 1 + 1 = 3.$$

Строим образующую матрицу

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{bmatrix}.$$

Проверочная матрица будет иметь вид

$$N = \begin{matrix} & a_1 & a_2 & a_3 & a_4 & a_5 & a_6 & a_7 \\ \begin{matrix} 0 \\ 1 \\ 1 \end{matrix} & \begin{vmatrix} 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 1 & 0 & 0 & 1 \end{vmatrix} \end{matrix}.$$

Обозначим символы, стоящие в каждой строке, через $a_i (a_1 a_2 a_3 a_4 a_5 a_6 a_7)$. Символы a_5 , a_6 и a_7 примем за контрольные, так как они будут входить только в одну из проверок.

Составим проверки для каждого контрольного символа. Из первой строки имеем

$$a_5 = a_2 \oplus a_3 \oplus a_4. \quad (2.12)$$

Из второй строки получим алгоритм для формирования контрольного символа a_6 :

$$a_6 = a_1 \oplus a_2 \oplus a_4. \quad (2.13)$$

Аналогично из третьей строки получим алгоритм для формирования контрольного символа a_7 :

$$a_7 = a_1 \oplus a_3 \oplus a_4. \quad (2.14)$$

Нетрудно убедиться, что все результаты проверок на четность по выражениям (2.12)–(2.14) дают нуль, что свидетельствует о правильности составления образующей и проверочной матриц.

Запись кодовых комбинаций в виде многочлена

Любое число в системе счисления с основанием X можно представить в виде многочлена.

$$F(X) = \sum_{i=0}^{n-1} A_i X^i = A_{n-1} X^{n-1} + A_{n-2} X^{n-2} + \dots + A_0 X^0,$$

где A - цифровые коэффициенты, имеющие значения от 0 до $X - 1$.

$$G(X) = 1 \cdot x^5 + 0 \cdot x^4 + 1 \cdot x^3 + 1 \cdot x^2 + 1 \cdot x^1 + 1 \cdot x^0 = x^5 + x^3 + x^2 + x^1 + 1.$$

Таким образом, члены многочленов записываются только при наличии коэффициента единицы.

Сложение

$$x^6 + x^5 + x^3 + x^2 + 1; x^5 + x^4 + x^2; x^6 + x^5 + x^4 + x^3 + x^2 + x^1 + 1.$$

$$x^6 + x^5 + 0 + x^3 + x^2 + 0 + 1 \rightarrow 1101101 \rightarrow 109_{10}$$

$\oplus$

$$0 + x^5 + x^4 + 0 + x^2 + 0 + 0 \rightarrow 0110100 \rightarrow 52_{10}$$

$\oplus$

$$x^6 + x^5 + x^4 + x^3 + x^2 + x + 1 \rightarrow 1111111 \rightarrow 127_{10}$$

$$0 + x^5 + 0 + 0 + x^2 + x + 0 \rightarrow 0100110 \rightarrow 38_{10}$$

Умножение. Для того чтобы при умножении многочленов не увеличилась разрядность степени многочлена выше заданной, производят так называемое символическое умножение, или умножение в конечном поле двоичных чисел, состоящее из двух этапов. Первый этап заключается в умножении многочленов по обычным правилам алгебры, за исключением сложения, которое производится по модулю 2.

$$\begin{array}{r}
 x^6+x^5+x^3+x^2+1= \\
 \times \\
 \hline
 x^5+x^4+x^2= \\
 \hline
 x^8+x^7+0+x^5+x^4+0+x^2 \\
 x^{10}+x^9+0+x^7+x^6+0+x^4 \\
 x^{11}+x^{10}+0+x^8+x^7+0+x^5 \\
 \hline
 x^{11}+0+x^9+0+x^7+x^6+0+0+0+x^2
 \end{array}$$

$$\begin{array}{r}
 1101101 \\
 \times \\
 \hline
 110100 \\
 \hline
 1101101 \\
 1101101 \\
 1101101 \\
 \hline
 101011000100
 \end{array}$$

=

Произведем теперь умножение многочлена на n .

$$(x^5+x^4+x^2)\times x^3=x^8+x^7+x^5.$$

В результате умножения степень каждого члена многочлена повышалась на n. В двоичной форме записи 110100 1000= 110100000. Таким образом, умножение многочлена на X n означает приписывание справа n нулей.

Деление. При делении в двоичной записи делитель умножается на частное и подписывается под делимым так, чтобы совпадали старшие разряды. В частное записывается единица. Для нахождения первого остатка из делимого вычитается делитель (что эквивалентно их сложению по модулю 2) и к остатку справа сносится очередной разряд делимого. Далее под первым остатком снова подписывается делитель и в частное приписывается еще одна единица, если число разрядов в остатке равно числу разрядов делителя. В противном случае в частном записывается нуль и к остатку подписывается очередной член делимого. Деление продолжается до тех пор, пока степень остатка не станет меньше степени делителя, т.е. число разрядов остатка не окажется меньше числа разрядов делителя.

$$\begin{array}{r}
 110110101 \mid 101011 \\
 \underline{101011} \\
 111011 \\
 \underline{101011} \\
 100000 \\
 \underline{101011} \\
 10111 \rightarrow x^4 + x^2 + x + 1 \text{ — остаток}
 \end{array}
 \quad
 \begin{array}{l}
 1110 \rightarrow x^4 + x^2 + x \text{ — частное } Q(X)
 \end{array}$$

Циклические коды относятся к числу блочных систематических кодов, в которых каждая комбинация кодируется самостоятельно (в виде блока) таким образом, что информационные k и контрольные r символы всегда находятся на определенных местах

При составлении циклических кодов необходимо уметь находить только остатки без определения частного. Ниже дается пример нахождения нескольких остатков при делении единицы с нулями на случайно выбранный многочлен $P(X) = 1011$. Следует помнить, что число разрядов у остатков на единицу меньше, чем у делителя

Решение задачи

100000000	1011	
1011		
01100		— первый остаток $R_1(X) = 011$
1011		— второй остаток $R_2(X) = 110$
		— третий остаток $R_3(X) = 111$
1110		
1011		
1010		— четвертый остаток $R_4(X) = 101$
1011		
0010		— пятый остаток $R_5(X) = 001$
		шестой остаток $R_6(X) = 010$
		седьмой остаток $R_7(X) = 100$

Любая разрешенная кодовая комбинация циклического кода может быть получена в результате умножения образующего многочлена на некоторый другой многочлен с приведением результата по модулю X^{n+1} .

При соответствующем выборе образующего многочлена, любой многочлен циклического кода будет делиться на него без остатка. Ни один многочлен, соответствующий запрещенной кодовой комбинации, на образующий многочлен без остатка не делится. Это свойство позволяет обнаружить ошибку. По виду остатка можно определить и вектор ошибки. Умножение и деление многочленов весьма просто осуществляется на регистрах сдвига с обратными связями и сумматорах по модулю 2. В основу циклического кодирования положено использование неприводимого многочлена $P(X)$, который применительно к циклическим кодам называется образующим, генераторным или производящим многочленом (полиномом). Многочлен в поле двоичных чисел называется неприводимым, если он делится без остатка только на себя или на единицу

МЕТОДЫ ПОСТРОЕНИЯ ЦИКЛИЧЕСКОГО КОДА

Существует несколько различных способов кодирования. Принципиально наиболее просто комбинации циклического кода можно получить, умножая многочлены $G(X)$, соответствующие комбинациям безызбыточного кода (информационным символам), на образующий многочлен кода $P(X)$. Такой способ легко реализуется, однако он имеет тот существенный недостаток, что получающиеся в результате умножения комбинации кода не содержат информационных символов в явном виде. После исправления ошибок такие комбинации для выделения информационных символов приходится делить на образующий многочлен кода. Ситуацию можно значительно упростить, если контрольные символы переписать в конце кода, т.е. после информационных символов. Для этой цели прибегают к следующему искусственному приему.

Умножаем кодовую комбинацию $G(X)$, которую мы хотим закодировать, на многочлен X^r , имеющий ту же степень, что и образующий многочлен $P(X)$. Делим произведение $G(X)X^r$ на образующий полином $P(X)$:

$$\frac{G(X) \cdot X^r}{P(X)} = Q(X) + \frac{R(X)}{P(X)},$$

где $Q(X)$ - частное от деления; $R(X)$ - остаток.

Умножая выражение на $P(X)$ и перенося $R(X)$ в другую часть равенства, согласно правилам алгебры двоичного поля, т.е. без перемены знака на обратный, получаем:

$$F(X) = Q(X) \cdot P(X) = G(X) \cdot X^r + R(X).$$

Закодировать кодовую комбинацию циклическим кодом

$$G(X) = 1111 = x^3 + x^2 + x + 1$$

Умножая $G(X) \times X^r$, получаем:

$$G(X) \cdot X^3 = (x^3 + x^2 + x + 1)x^3 = x^6 + x^5 + x^4 + x^3 \rightarrow 1111000.$$

От умножения степень каждого члена повысилась, что равносильно приписыванию трех нулей к многочлену, выраженному в двоичной форме.

Разделив на $G(X) \cdot X^r$ на $P(X)$, согласно (2.28) получим

$$\frac{x^6 + x^5 + x^4 + x^3}{x^3 + x + 1} = (x^3 + x^2 + 1) + \frac{x^2 + x + 1}{x^3 + x + 1},$$

или в двоичном эквиваленте

$$1111000/1011 = 1101 + 111/1011.$$

Таким образом, в результате деления получаем частное $Q(X) = 1101$ той же степени, что и $G(X) = 1111$, и остаток $R(X) = 111$. В итоге комбинация двоичного кода, закодированная циклическим кодом, согласно () примет вид:

$$F(X) = 1101 \cdot 1011 = 1111000 \oplus 111 = 1111111.$$

Действительно, умножение 1101×1011 (первый способ) дает тот же результат, что и сложение 1111000 (по модулю 2)
111 (второй способ).

ЦИКЛИЧЕСКИЕ КОДЫ ОБНАРУЖИВАЮЩИЕ ОДИНОЧНУЮ ОШИБКУ

Код, образованный генераторным полиномом $P(X) = X + 1$, обнаруживает любое нечетное число ошибок.

Закодируем сообщение $G(X) = 1101$ с помощью многочлена $P(X) = 11$.

$$F(X) = G(X) \cdot X^r + R(X) = 11010 + 1 = 11011,$$

т.е. на первых четырех позициях находятся разряды исходной комбинации $G(X)$, а на пятой - контрольный символ.

Сообщение 1101 является одной из 16 комбинаций четырехразрядного кода. Если требуется передать все эти сообщения в закодированном виде, то каждое из них следует кодировать так же, как и комбинацию $G(X) = 1101$. Однако проделывать дополнительно 15 расчетов (в общем случае 2^r расчетов) нет необходимости. Это можно сделать проще, путем составления образующей матрицы. Образующая матрица составляется из единичной транспонированной и матрицы дополнений, составленной из остатков от деления единицы с нулями на образующий многочлен $P(X)$, выраженный в двоичном эквиваленте.

$$M = \begin{array}{cccc|c} 0 & 0 & 0 & 1 & 1 & a_1 \\ 0 & 0 & 1 & 0 & 1 & a_2 \\ 0 & 1 & 0 & 0 & 1 & a_3 \\ 1 & 0 & 0 & 0 & 1 & a_4 \end{array}$$

Единичная транспонированная матрица

Матрица дополнений

Четыре кодовые комбинации, из которых состоит образующая матрица, являются первыми кодовыми комбинациями циклического кода. Пятая комбинация нулевая, а так как в четырехразрядном непомехозащищенном коде всего $N = 2^4 = 16$ комбинаций, то остальные 11 ненулевых комбинаций находят суммированием по модулю 2 всевозможных комбинаций строк матрицы M :

5. 0000,	9. $a_2 \oplus a_3 = 01100$,	13. $a_2 \oplus a_3 \oplus a_4 = 11101$,
6. $a_1 \oplus a_2 = 00110$,	10. $a_2 \oplus a_4 = 10100$,	14. $a_1 \oplus a_3 \oplus a_4 = 11011$,
7. $a_1 \oplus a_3 = 01010$,	11. $a_3 \oplus a_4 = 11000$,	15. $a_1 \oplus a_2 \oplus a_4 = 10111$,
8. $a_1 \oplus a_4 = 10010$,	12. $a_1 \oplus a_2 \oplus a_3 = 01111$,	16. $a_1 \oplus a_2 \oplus a_3 \oplus a_4 = 11110$.

Рассмотрение полученных комбинаций показывает, что все они имеют четное число единиц. Таким образом, циклический код с обнаружением одиночной ошибки $d = 2$ является кодом с проверкой на четность.

Циклический код с $d=3$

Эти коды могут обнаруживать одиночные и двойные ошибки или обнаруживать и исправлять одиночные ошибки. Можно предложить следующий порядок кодирования кодовых комбинаций в циклическом коде с $d = 3$.