Чиркин А. С., Введение в статистическую радиофизику и оптику, М., 1981. *К. Н. Драбович.*

КОДИРОВАНИЕ ИНФОРМАЦИИ — установление соответствия между элементами сообщения и сигналами, при помощи к-рых эти элементы могут быть зафиксированы.

Пусть B, $b_i \in B$, $i=\overline{1,\ n}$ — множество элементов *сообщения*, A — алфавит с символами $a_j \in A$, $j=\overline{1,\ m}$. Пусть конечная последовательность символов наз. словом в данном алфавите. Множество слов в алфавите A наз. кодом, если оно поставлено во взаимно однозначное соответствие с множеством B. Каждое слово, входящее в код, наз. кодовым словом. Число символов в кодовом слове наз. длиной слова. Кодовые слова могут иметь одинаковую или разл. длину. В соответствии с этим код наз. равномерным или неравномерным.

Цели К. и.: представление входной информации в ЭВМ, согласование источников информации с каналом передачи, обнаружение и исправление ошибок при передаче и обработке данных, сокрытие смысла сообщения (криптография) и т. д. Информационные свойства объекта, как правило, таковы, что код может быть представлен наиболее экономным образом. Эту задачу решает кодер источника, удаляя из сообщений избыточность. Дальнейшие этапы прохождения данных — передача по каналу передачи и (или) хранение в запоминающих устройствах — требуют обнаружения и(или) исправления ошибок, возникающих в них вследствие помех. Эти цели достигаются путём корректирующего кодирования, осуществляемого кодером канала. Наконец, защита информации от искажений при обработке в ЭВМ осуществляется применением арифметич. кодов.

Кодирование значений. Натуральное число N представлено в позиционной весомозначной системе счисления, если имеет место соотношение

$$N=\sum_{l=0}^{n-1} a_l p_l, \qquad (1)$$

где $A=\{a_0,\ \ldots,\ a_l,\ \ldots,\ a_{n-1}\}$ — цифровой алфавит с n цифрами, $P=\{p_0,\ \ldots,\ p_l,\ \ldots,\ p_{n-1}\}$ — веса разрядов, $l=\overline{0,\ n-1}$ — номера разрядов. Термин «позиционная» означает, что в кодовом представлении (или просто коде) числа, выражаемом условным равенством

$$N=a_{n-1},\ \ldots,\ a_l,\ \ldots,\ a_0,$$

количественный эквивалент, сопоставляемый цифре a_l, зависит и от её расположения в коде. Термин «весомозначная» означает, что каждый разряд имеет вес p_l. Вес младшего разряда p_0 в цифровой измерительной технике отождествляется с разрешающей способностью аналого-цифрового преобразования. Выбор алфавита A и системы весов P задаёт классификацию позиционных систем счисления (кодирование значений).

В естественных системах

$$p_l=np_{l-1}=n^l p_0 \qquad (2)$$

и, если n — основание системы счисления — натуральное число, любое число X может быть представлено как

$$X=\sum_{l=-\infty}^{\infty} a_l n^l. \qquad (3)$$

Выбор алфавита смещённым: $A=(0,\ 1,\ \ldots,\ n-1)$, $A=(-n-1,\ \ldots,\ 1,\ 0)$, или симметричным: $A=(-n-1,\ \ldots,\ -1,\ 0,\ 1,\ \ldots,\ n-1)$ позволяет представлять соответственно положительные, отрицательные или любые числа. Симметричная система должна обладать нечётным основанием.

В ЭВМ почти исключительно используется позиционная двоичная смещённая система ($n=2$) с цифрами (0, 1) и естественным соотношением весов, представляющих ряд чисел

$$\ldots\ 2^l,\ 2^{l-1},\ \ldots,\ 1,\ 0,\ \ldots,\ 2^{-l}.$$

Возможно применение и иного набора цифр, напр. (−1, 1), дающего нек-рые специфические преимущества.

Развиваются двоичные системы, веса разрядов к-рых находятся не в естественном (2), а в более сложном соотношении, образуя, напр., ряд Фибоначчи (или «золотую пропорцию») [1]. Число N в коде Фибоначчи представляется соотношением

$$N=a_{n-1}\varphi(n-1)+a_{n-2}\varphi(n-2)+\ldots+a_0\varphi(0), \quad (4)$$

где $\varphi(n)$ — числа Фибоначчи, связанные соотношением

$$\varphi(n)=\varphi(n-1)+\varphi(n-2),\ \ \varphi(0)=1,\ \ \varphi(1)=2.$$

Разложение (4) числа N неоднозначно. Для любого N существует код, в к-ром не встречается двух следующих подряд нулей, а также код, в к-ром не соседствуют единицы. Эти, а также др. структурные особенности кодов Фибоначчи и «золотых» кодов делают их удобными для построения самокорректирующихся преобразователей, запоминающих и вычислит. устройств, сервоприводов с цифровым управлением и т. п.

Троичные системы счисления наиб. экономичны в том смысле, что именно в троичном коде определ. кол-вом знаков может быть выражено наибольшее разнообразие чисел. Есть основание полагать, что в будущем именно в силу указанного свойства троичная симметричная система кодирования с цифрами (−1, 0, 1) займёт в вычислит. технике доминирующее место. Проблемой остаётся создание элементов, реализующих ф-ции базиса в троичной логике: троичный инвертор и троичные НЕ—И или троичные НЕ—ИЛИ (см. *Логические схемы*).

Непозиционные коды применяют в специализированных измерит. и вычислит. устройствах [2]. Простейший из непозиционных — унитарный код можно получить, положив в (2) $n=1$ и $p_0=1$. В нём число N представляется как $N=n+1$ — последовательно суммируемые единицы. Так работают, напр., счётчики импульсов.

Среди систем непозиционного кодирования выделяется система счисления в остаточных классах (СОК). Число N в СОК представляется в виде упорядоченного набора остатков (вычетов) по взаимно простым основаниям $p_1,\ \ldots,\ p_n$: $N=(\alpha_1,\ \ldots,\ \alpha_i,\ \ldots,\ \alpha_n)$, где α_i — наименьший вычет N по модулю p_i. Система оснований $p_1,\ p_2,\ \ldots,\ p_n$ определяет диапазон представления чисел $P=p_1,\ p_2,\ \ldots,\ p_n$. В СОК арифметич. операции производятся независимо по каждому основанию и это позволяет существенно увеличить скорость их выполнения. В СОК удобен контроль операций, т. к. ошибки локализованы в пределах оснований. Специфичным для вычислит. устройств, работающих в СОК, является применение табличной арифметики: значения ф-ции, подлежащей вычислению, заранее заносятся в таблицу, а затем извлекаются при поступлении значений операндов.

Эффективное кодирование источника информации [3] имеет целью согласование информационных свойств источника информации (ИИ) и канала передачи. Предполагается, что ИИ выдаёт на выходе сообщение, состоящее из букв m-буквенного алфавита

$$A_m=\{a_1,\ \ldots,\ a_m\},$$

причём появление букв статистически независимо и подчинено распределению

$$P=\{p_1,\ \ldots,\ p_m\},\ \ p_i>0,\ \ \sum_i p_i=1.$$

Источник характеризуется энтропией на символ

$$H(P)=\sum_{i=1}^{m} p_i \log_2(1/p_i).$$

Энтропия $0 \leqslant H(P) \leqslant \log_2 m$ имеет смысл меры неопределённости относительно появления на выходе ИИ очередного символа. Равенство $H(P)=0$ достигается при вырожденном распределении P, т. к. сообщение