

## 1.7 Обработка информации

Множество сообщений  $N$  представляет интерес только тогда, когда ему соответствует (по крайней мере одно) множество сведений  $I$  и определено соответствующее правило интерпретации  $\varphi: N \rightarrow I$  (см. п. 1.2). Так как множеству сообщений  $N'$  тоже соответствует некоторое множество сведений  $I'$  (и правило интерпретации  $\varphi'$ ), то любое правило обработки сообщений  $\nu: N \rightarrow N'$  (см. п. 1.6) приводит к следующей диаграмме:

$$\begin{array}{ccc} N & \xrightarrow{\varphi} & I \\ \nu \downarrow & & \downarrow \sigma \\ N' & \xrightarrow{\varphi'} & I' \end{array} \quad (*)$$

Эта диаграмма определяет соответствие между множествами  $I$  и  $I'$ . Так как согласно диаграмме (\*) каждому сообщению  $n \in N$  соответствует пара сведений  $i = \varphi(n) \in I$  и  $i' = \varphi(\nu(n)) \in I'$ , построенное соответствие между  $I$  и  $I'$  (обозначим его через  $\sigma$ ), вообще говоря, не является отображением. В самом деле, если правило интерпретации  $\varphi$  не является однозначным (инъективным, когда разные переходят в разные), т. е. если существуют два различных сообщения  $n_1, n_2 \in N$ ,  $n_1 \neq n_2$ , передающих одинаковую информацию  $i = \varphi(n_1) = \varphi(n_2)$ , то может оказаться, что  $\varphi'(\nu(n_1)) \neq \varphi'(\nu(n_2))$  и, следовательно, одной информации  $i \in I$  будут соответствовать (по крайней мере) две различных информации  $i'_1 = \varphi'(\nu(n_1))$  и  $i'_2 = \varphi'(\nu(n_2))$ .

Если отображение  $\varphi$  обратимо, т. е. если существует отображение  $\varphi^{-1}$  (для нас достаточно, чтобы  $\varphi$  было инъективным отображением), то можно построить отображение  $\sigma$ , определяющее обработку информации  $\sigma: I \rightarrow I'$  в виде  $\sigma = \varphi' \circ \nu \circ \varphi^{-1}$  так, что  $i' = \varphi'(\nu(\varphi^{-1}(i)))$ .

Во всех случаях, когда соответствие  $\sigma$  является отображением, правило обработки сообщений  $\nu$  называется **сохраняющим информацию**. Если правило обработки сообщений  $\nu$  сохраняет информацию, то диаграмма

$$\begin{array}{ccc} N & \xrightarrow{\varphi} & I \\ \nu \downarrow & & \downarrow \sigma \\ N' & \xrightarrow{\varphi'} & I' \end{array} \quad (**)$$

коммутативна:  $\nu \circ \varphi' = \varphi \circ \sigma$ . Отображение  $\sigma$  называется в этом случае **правилом обработки информации**.

Обычно обработку информации сводят к обработке сообщений, т. е., исходя из требуемого правила обработки информации  $\sigma$ , пытаются определить отображения  $\nu$ ,  $\varphi$  и  $\varphi'$  таким образом, чтобы диаграмма (\*\*) была коммутативной.

Если  $\sigma$  — обратимое (взаимно однозначное) отображение, т. е. если информация при обработке по правилу  $\sigma$  не теряется, то соответствующую обработку сообщений  $\nu$  называют **перешифровкой**.

Пусть  $\nu$  — обратимая перешифровка. Тогда по сообщению  $n' = \nu(n)$  можно восстановить не только исходную информацию, но и само исходное сообщение  $n$ . Иными словами, в этом случае  $n'$  *кодирует*  $n$  (см. п. 1.4). Обратимая перешифровка  $\nu$  называется **перекодировкой**.

Пусть перешифровка  $\nu$  не является обратимой, т. е. пусть несколько сообщений из  $N$  кодируются одним и тем же сообщением из  $N'$ . Но так как при перешифровке информация

не теряется, это означает, что исходное множество сообщений  $N$  является избыточным: некоторые сообщения из  $N$  содержат одну и ту же информацию (дублируют друг друга). В  $N'$  таких дублирующих сообщений меньше, чем в  $N$ , так как при обработке по правилу  $\nu$  некоторые из дублирующих друг друга сообщений «сливаются» в одно сообщение. Перешифровка  $\nu$ , которая не является обратимой, называется **сжимающей**. Сжатию подвергается множество сообщений. То есть в результате необратимой перешифровки сообщений их количество уменьшается, а информация может либо сохраняться, либо теряться.

**Пример 1.7.1.** Пусть сообщения  $(a, b)$ , составленные из пар целых чисел (например, в десятичной позиционной записи), передают информацию «рациональное число  $r$ , представленное дробью  $\frac{a}{b}$ ». Тогда  $N = \mathbb{Z} \times \mathbb{N}$  (где  $\mathbb{Z}$  — множество целых чисел,  $\mathbb{N}$  — множество натуральных чисел),  $I = \mathbb{Q}$  ( $\mathbb{Q}$  — множество рациональных чисел). Отображение  $\varphi: N \rightarrow I$  не является обратимым, так как при любом целом  $n$  парам  $(a, b)$  и  $(na, nb)$  соответствует одно и то же рациональное число  $r$ . Пусть  $N'$  — множество пар  $(p, q)$  взаимно простых целых чисел и пусть  $\nu: N \rightarrow N'$  переводит все  $(np, nq)$  в  $(p, q)$ . Тогда  $\nu$  — сжимающее отображение, а  $\varphi': N' \rightarrow I$  — обратимое отображение (мы считаем  $I' = I$ ). Такое отображение  $\nu$  называется **вполне сжимающей перешифровкой**, поскольку после обработки сообщений соответствие между сообщениями и информацией биективно. Здесь информация не теряется.

Если  $\sigma$  — необратимое отображение, т. е. если разные сведения из  $I$  отображаются в одну и ту же информацию  $i' \in I'$ , то соответствующую обработку сообщений называют **избирательной**. Особенно часто встречается случай, когда  $I' \subset I$ , а  $\sigma$  — тождественное отображение для всех  $i' \in I'$ . В этом случае производится *выбор* из данного множества сведений.

Таким образом, «обработка информации» — это, как правило, сокращение количества информации. Во всяком случае, верно утверждение: обработка информации никогда **не добавляет** информацию, она состоит в том, что **извлекает** интересную информацию из той, которая содержится в сообщении.

## Лекция 4

### 1.8 Автоматизация обработки информации

Вернемся к рассмотрению диаграммы (\*\*). Если заменить на ней отображение  $\varphi$  обратным отображением  $\psi = \varphi^{-1}$ , получим новую диаграмму:

$$\begin{array}{ccc} N & \xleftarrow{\psi} & I \\ \nu \downarrow & & \downarrow \sigma \\ N' & \xrightarrow{\varphi'} & I' \end{array}$$

Автоматизация обработки информации заключается в выполнении  $\sigma$  или  $\varphi^{-1} \circ \nu \circ \varphi'$  при помощи физических устройств. Однако в программировании изучаются методы автоматического выполнения только отображения  $\nu$ , т. е. обработки сообщений. Программно-аппаратная реализация отображений  $\psi, \varphi'$  изучается в другом разделе информатики, который называется «Искусственный интеллект» — и потому выходят за рамки нашего курса.