

Кодирование информации

Одну и ту же информацию, например, сведения об опасности мы можем выразить разными способами: просто крикнуть; оставить предупреждающий знак (рисунок); с помощью мимики и жестов; передать сигнал «SOS» с помощью азбуки Морзе или используя семафорную и флажковую сигнализацию. В каждом из этих способов мы должны знать правила, по которым можно отобразить информацию. Такое правило назовем кодом.

Код — это набор условных обозначений (или сигналов) для записи (или передачи) некоторых заранее определенных понятий.

Кодирование информации – это процесс формирования определенного представления информации. В более узком смысле под термином «кодирование» часто понимают переход от одной формы представления информации к другой, более удобной для хранения, передачи или обработки.

Обычно каждый образ при кодировании (иногда говорят — шифровке) представляется отдельным знаком.

Знак - это элемент конечного множества отличных друг от друга элементов.

Знак вместе с его смыслом называют символом.

Набор знаков, в котором определен их порядок, называется алфавитом. Существует множество алфавитов:

алфавит кириллических букв {А, Б, В, Г, Д, Е, ...}

алфавит латинских букв {А, В, С, D, E, F,...}

алфавит десятичных цифр {0, 1, 2, 3, 4, 5, 6, 7, 8, 9}

алфавит знаков зодиака {картинки знаков зодиака} и др.

Особенно большое значение имеют наборы, состоящие всего из двух знаков:

пара знаков {+, -}

пара цифр {0, 1}

пара ответов {да, нет}

Алфавит, состоящий из двух знаков, называется двоичным алфавитом. Двоичный знак (англ. binary digit) получил название «бит».

Шифрование - кодирование сообщения отправителя, но такое чтобы оно было не понятно несанкционированному пользователю.

Длиной кода называется такое количество знаков, которое используется при кодировании.

Количество символов в алфавите кодирования и длина кода - совершенно разные вещи. Например, в русском алфавите 33 буквы, а слова могут быть длиной в 1, 2, 3 и т.д. буквы.

Код может быть постоянной и непостоянной длины. Коды различной (непостоянной) длины в технике используются довольно редко. Исключением является лишь троичный код Морзе. В вычислительной технике в настоящее время широко используется двоичное кодирование с алфавитом (0, 1). Наиболее распространенными кодами являются ASCII (American standart code for information interchange - американский стандартный код для обмена информацией) и КОИ-8 (код обмена информации длиной 8 бит).

Одно и то же сообщение можно закодировать разными способами, т. е. выразить на разных языках. В процессе развития человеческого общества люди выработали большое число языков кодирования. К ним относятся:

разговорные языки (русский, английский, хинди и др. — всего более 2000);

язык мимики и жестов;

язык рисунков и чертежей;

язык науки (математические, химические, биологические и другие символы);

язык искусства (музыки, живописи, скульптуры и т. д.);

специальные языки (эсперанто, морской семафор, азбука Морзе, азбука Брайля для слепых и др.).