**Кодирование информации с помощью знаковых систем**

**Знаки: форма и значение**

С древних времен знаки используются человеком для долговременного хранения информации и ее передачи на большие расстояния.

**Форма знаков.** В соответствии со способом восприятия знаки можно разделить на зрительные, слуховые, осязательные, обонятельные и вкусовые, причем в человеческом общении используются знаки первых трех типов.

К зрительным знакам, воспринимаемым с помощью зрения, относятся буквы и цифры, которые используются в письменной речи, знаки химических элементов, музыкальные ноты, дорожные знаки и т. д.

К слуховым знакам, воспринимаемым с помощью слуха, относятся звуки, которые используются в устной речи, а также звуковые сигналы, которые производятся с помощью звонка, колокола, свистка, гудка, сирены и т. д.

Для слепых разработана азбука Брайля, которая использует осязательный способ восприятия текстовой информации.

В коммуникации многих видов животных особую роль играют обонятельные знаки. Например, медведи и другие дикие животные помечают место обитания клочьями шерсти, сохраняющей запах, чтобы отпугнуть чужака и показать, что данная территория уже занята.

Для долговременного хранения знаки записываются на носители информации.

Для передачи информации на большие расстояния используются знаки в форме **сигналов**. Всем известны световые сигналы светофора, звуковые сигналы школьного звонка оповещают о начале или конце урока, электрические сигналы передают информацию по телефонным и компьютерным сетям, электромагнитные волны передают сигналы радио и телевидения.

**Значение знаков.** Знаки отображают объекты окружающего мира или понятия, т. е. имеют определенное значение (смысл).

Знаки различаются по способу связи между их формой и значением. **Иконические знаки** позволяют догадаться об их смысле, так как они имеют форму, похожую на отображаемый объект. Примером таких знаков являются значки на *Рабочем столе* операционной системы компьютера, например значок *Мой компьютер*.

**Символами** называются знаки, для которых связь между формой и значением устанавливается по общепринятому соглашению. Примером таких знаков являются символы химических элементов, отображающие атомы химических веществ (табл. 2.1).

Если неизвестно соглашение о связи формы и значения символов, то ничего нельзя сказать о смысле информации, записанной такими знаками. Существуют найденные археологами и до сих пор нерасшифрованные тексты на древних языках, так как неизвестно значение знаков, которыми они записаны.

|  |
| --- |
| Таблица 2.1. Иконические знаки и символы |
| http://www.5byte.ru/9/images/predinfo1.gif |

В современном мире широко применяется шифрование, которое использует секретный ключ в качестве соглашения о связи формы символов с их значениями. Если секретный ключ неизвестен, то содержание передаваемого текста понять невозможно.

Один и тот же символ может иметь различное значение в разных знаковых системах. Например, знак "О" используется в качестве:

- буквы "О" в русском алфавите;
- буквы "О" [ou] в английском алфавите;
- цифры 0 в системах счисления;
- символа химического элемента "О" (кислорода) в таблице Д. И. Менделеева.

Контрольные вопросы

1. Приведите примеры зрительных, слуховых, осязательных, обонятельных и вкусовых знаков. Какие типы знаков применяются в человеческом общении?

2. Приведите примеры знаков в форме сигналов.

3. В чем состоит различие между иконическими знаками и символами?

4. Приведите примеры символов, которые имеют различное значение в нескольких знаковых системах.

**Знаковые системы**

Знаковые системы являются наборами знаков определенного типа. С некоторыми знаковыми системами вы хорошо знакомы и постоянно ими пользуетесь (языки и системы счисления), с другими познакомитесь в этом пункте.

*Каждая* ***знаковая система*** *строится на основе определенного алфавита (набора знаков) и правил выполнения операций над знаками.*

**Естественные языки.** Человек широко использует для представления информации знаковые системы, которые называются **языками**. Естественные языки начали формироваться еще в древнейшие времена в целях обеспечения обмена информацией между людьми. В настоящее время существуют сотни естественных языков (русский, английский, китайский и др.).

В устной речи, которая используется как средство коммуникации при непосредственном общении людей, в качестве знаков языка используются различные звуки (**фонемы**).

В основе письменной речи лежит **алфавит**, т. е. набор знаков (букв), которые человек различает по их начертанию. В большинстве современных языков буквы соответствуют определенным звукам устной речи. Алфавит русского языка называется кириллицей и содержит 33 знака, английский язык использует латиницу и содержит 26 знаков.

На основе алфавита по правилам **грамматики** образуются основные объекты языка - слова. Правила, согласно которым из слов данного языка строятся предложения, называются **синтаксисом**. Необходимо отметить, что в естественных языках грамматика и синтаксис языка формулируются с помощью большого количества правил, из которых существуют исключения, так как такие правила складывались исторически.

**Формальные языки.** В процессе развития науки были разработаны формальные языки (системы счисления, алгебра, языки программирования и др.), основное отличие которых от естественных языков состоит в существовании строгих правил грамматики и синтаксиса.

Например, десятичную систему счисления можно рассматривать как формальный язык, имеющий алфавит (цифры) и позволяющий не только именовать и записывать объекты (числа), но и выполнять над ними арифметические операции по строго определенным правилам.

Существуют формальные языки, в которых в качестве знаков используют не буквы и цифры, а другие символы, например музыкальные ноты, изображения элементов электрических или логических схем, дорожные знаки, точки и тире (код азбуки Морзе).

Физическая реализация знаков в естественных и формальных языках может быть различной. Например, текст и числа могут быть напечатаны на бумаге, высвечены на экране монитора компьютера, записаны на магнитном или оптическом диске.

**Генетический алфавит.** Генетический алфавит является "азбукой", на которой строится единая система хранения и передачи наследственной информации живыми организмами.

Как слова в языках образуются из букв, так и гены состоят из знаков генетического алфавита. В процессе эволюции от простейших организмов до человека количество генов постоянно возрастало, так как было необходимо закодировать все более сложное строение и функциональные возможности живых организмов.

Генетическая информация хранится в клетках живых организмов в специальных молекулах. Эти молекулы состоят из двух длинных скрученных друг с другом в спираль цепей, построенных из четырех различных молекулярных фрагментов (рис. 1.6). Фрагменты образуют генетический алфавит и обычно обозначаются латинскими прописными буквами {A, G, С, Т}.

|  |
| --- |
| http://www.5byte.ru/9/images/predinfo2.gif |
| Рис 1.6 Модель молекулы генетического кода |

**Двоичная знаковая система.** В процессах хранения, обработки и передачи информации в компьютере используется двоичная знаковая система, алфавит которой состоит всего из двух знаков {0, 1}. Физически знаки реализуются в форме электрических импульсов (нет импульса - 0, есть импульс - 1), а также состояний ячеек оперативной памяти и участков поверхностей носителей информации (одно состояние - 0, другое состояние - 1).

Именно двоичная знаковая система используется в компьютере, так как существующие технические устройства могут надежно сохранять и распознавать только два различных состояния (знака).

В 60-е годы XX века в СССР учеными Московского государственного университета была разработана и запущена в производство ЭВМ "Сетунь" (всего было произведено 50 экземпляров) (рис. 1.7). "Сетунь" использовала троичное кодирование информации и, соответственно, состояла из устройств, способных находиться в одном из трех возможных состояний.

|  |
| --- |
| http://www.5byte.ru/9/images/predinfo3.gif |
| Рис. 1.7. ЭВМ "Сетунь" |

Контрольные вопросы

1. Приведите примеры знаковых систем. Какова может быть физическая природа знаков?

2. В чем состоит различие между естественными и формальными языками?

3. Обладают ли генетическим кодом растения? Животные? Человек?

4. Почему в компьютерах используется двоичная знаковая система для кодирования информации?

Задания для самостоятельного выполнения

1.1. Задание с развернутым ответом. Заполните нижеприведенную таблицу: введите алфавит и перечислите возможную физическую природу знаков для различных знаковых систем.

|  |  |  |
| --- | --- | --- |
| Знаковая система | Алфавит | Физическая природа знаков |
| Русский язык (письменный) |   |   |
| Русский язык (устный)  |   |   |
| Английский язык (письменный) |   |   |
| Десятичная система счисления |   |   |
| Генетический алфавит |   |   |
| Двоичный компьютерный код |   |   |

**Кодирование информации**

В процессах восприятия, передачи и хранения информации живыми организмами, человеком и техническими устройствами происходит ее кодирование.

**Код. Длина кода.** В процессе представления информации с помощью знаковой системы производится ее кодирование. Результатом кодирования является последовательность символов данной знаковой системы, то есть информационный код. Примерами кодов являются последовательности букв в тексте, цифр в числе, генетический код, двоичный компьютерный код и т. д.

Код состоит из определенного количества знаков (например, текстовое сообщение состоит из определенного количества букв, число - из определенного количества цифр и т. д.), т. е. имеет определенную длину.

*Количество знаков в коде называется* ***длиной кода****.*

Так, длина кода текста данного учебника составляет около 300 тысяч знаков, а генетический код человека в 10 тысяч раз длиннее, так как состоит из 3 миллиардов знаков генетического алфавита.

**Перекодирование информации из одной знаковой системы в другую.** Информация, представленная с помощью естественных и формальных языков, может быть выражена в форме устной речи или в письменном виде. Каждая форма представления использует особую знаковую систему, ориентированную на способ ее восприятия. Устная речь использует в качестве знаков набор звуков (фонем) и рассчитана на слуховое восприятие. В основе письменной речи лежит алфавит, т. е. набор знаков (букв), которые человек воспринимает с помощью зрения.

В процессе обмена информацией между людьми часто приходится переходить от одной формы представления информации к другой. Так, в процессе чтения вслух производится переход от письменной формы представления информации к устной и, наоборот, в процессе диктанта или записи объяснения учителя происходит переход от устной формы к письменной. В процессе преобразования информации из одной формы представления (знаковой системы) в другую происходит перекодирование информации.

***Перекодирование*** *- это операция преобразования знаков или групп знаков одной знаковой системы в знаки или группы знаков другой знаковой системы.*

Средством перекодирования служит таблица соответствия знаковых систем (таблица перекодировки), которая устанавливает взаимно однозначное соответствие между знаками или группами знаков двух различных знаковых систем. Ниже приведена табл. 2.2, которая устанавливает соответствие между гласными буквами русского алфавита и фонемами.

|  |
| --- |
| Таблица 2.2 Соответствие букв и звуков |
|

|  |  |
| --- | --- |
| Буквы | Звуки (фонемы) |
| а | [а] |
| о | [о] |
| у | [у] |
| и | [и] |
| ы | [ы] |
| э | [э] |
| е | [i]+[э] |
| ё | [i]+[о] |
| ю | [i]+[у] |
| я | [i]+[а] |

 |

В русской письменной речи только шесть гласных букв могут быть озвучены в устной речи соответствующими звуками. Для озвучивания остальных четырех гласных букв используются составные звуки, которые начинаются со звука [j].

Контрольные вопросы

1. Приведите примеры кодов и определите их длины.

2. Приведите примеры перекодирования информации из одной знаковой системы в другую. Какие в этих случаях используются таблицы перекодировки?

Задания для самостоятельного выполнения

1.2. Задание с кратким ответом. Перекодируйте с русского письменного языка на русский устный имя Юля.