

Кодирование информации

**Информация и
информационные процессы**

Кодирование и декодирование

Для обмена информацией с другими людьми человек использует естественные языки. Наряду с *естественными* языками были разработаны *формальные* языки для профессионального применения их в какой-либо сфере. Представление информации с помощью какого-либо языка часто называют кодированием.

Код — набор символов (условных обозначений) для представления информации.

Код — система условных знаков (символов) для передачи, обработки и хранения информации(со общения).

Кодирование — процесс представления информации (сообщения) в виде кода.

Все множество символов, используемых для кодирования, называется *алфавитом кодирования*. Например, в памяти компьютера любая информация кодируется с помощью двоичного алфавита, содержащего всего два символа: 0 и 1.

Декодирование- процесс обратного преобразования кода к форме исходной символьной системы, т.е. получение исходного сообщения. Например: перевод с азбуки Морзе в письменный текст на русском языке.

В более широком смысле декодирование — это процесс восстановления содержания закодированного сообщения. При таком подходе процесс записи текста с помощью русского алфавита можно рассматривать в качестве кодирования, а его чтение — это декодирование.

Способы кодирования информации

Для кодирования одной и той же информации могут быть использованы разные способы; их выбор зависит от ряда обстоятельств: цели кодирования, условий, имеющихся средств.

Если надо записать текст в темпе речи — используем стенографию; если надо передать текст за границу — используем английский алфавит; если надо представить текст в виде, понятном для грамотного русского человека, — записываем его по правилам грамматики русского языка.

«Здравствуй, Саша!»

«Zdravstvuy, Sasha!»

Способы кодирования информации

Выбор способа кодирования информации может быть связан с предполагаемым способом ее обработки.

Покажем это на примере представления чисел — количественной информации. Используя русский алфавит, можно записать число "**тридцать пять**". Используя же алфавит арабской десятичной системы счисления, пишем «**35**». Второй способ не только короче первого, но и удобнее для выполнения вычислений. Какая запись удобнее для выполнения расчетов: "**тридцать пять умножить на сто двадцать семь**" или "**35 x 127**"? Очевидно — вторая.

Шифрование сообщения

В некоторых случаях возникает потребность засекречивания текста сообщения или документа, для того чтобы его не смогли прочитать те, кому не положено. Это называется **защитой от несанкционированного доступа**.

В таком случае секретный текст ***шифруется***.

В давние времена шифрование называлось *тайнописью*.

Шифрование представляет собой процесс превращения открытого текста в зашифрованный, а дешифрование — процесс обратного преобразования, при котором восстанавливается исходный текст.

Шифрование — это тоже кодирование, но с засекреченным методом, известным только источнику и адресату.

Методами шифрования занимается наука под названием **криптография**.

Оптический телеграф Шаппа

В 1792 году во Франции Клод Шапп создал систему передачи визуальной информации, которая получила название «**Оптический телеграф**».

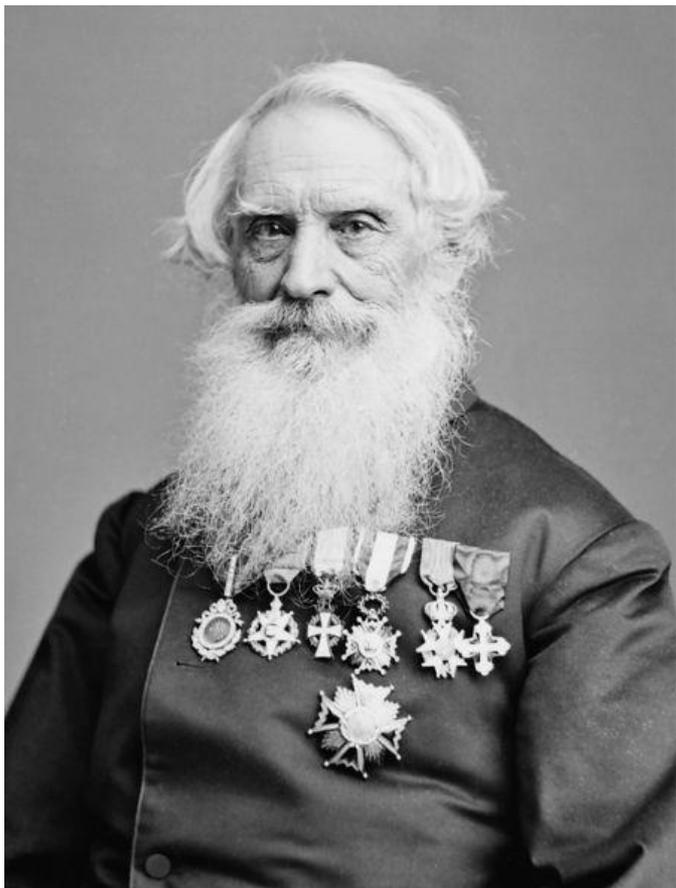
В простейшем виде это была цепь типовых строений, с расположенными на кровле шестами с подвижными поперечинами, которая создавалась в пределах видимости одно от другого. Шесты с подвижными поперечинами — семафоры — управлялись при помощи тросов специальными операторами изнутри строений.

Шапп создал специальную таблицу кодов, где каждой букве алфавита соответствовала определенная фигура, образуемая Семафором, в зависимости от положений поперечных брусьев относительно опорного шеста.

Система Шаппа позволяла передавать сообщения на скорости два слова в минуту и быстро распространилась в Европе. В Швеции цепь станций оптического телеграфа действовала до 1880 года.



Первый телеграф



Первым техническим средством передачи информации на расстояние стал **телеграф**, изобретенный в 1837 году американцем Сэмюэлем Морзе.

Телеграфное сообщение — это последовательность электрических сигналов, передаваемая от одного телеграфного аппарата по проводам к другому телеграфному аппарату.

Изобретатель Сэмюэль Морзе изобрел удивительный код (Азбука Морзе, код Морзе, «Морзянка»), который служит человечеству до сих пор. **Информация кодируется тремя «буквами»: длинный сигнал (тире), короткий сигнал (точка) и отсутствие сигнала (пауза)** для разделения букв. Таким образом, кодирование сводится к использованию набора символов, расположенных в строго определенном порядке.

Самым знаменитым телеграфным сообщением является сигнал бедствия "**SOS**" (Save Our Souls - спасите наши души). Вот как он выглядит: «•••

Азбука Морзе

А	• —	И	• •	Р	• — •	Ш	— — — —
Б	— • • •	Й	• — — —	С	• • •	Щ	— — • —
В	• — —	К	— • —	Т	—	Ъ	• — — • — •
Г	— — •	Л	• — • •	У	• • —	Ь	— • • —
Д	— • •	М	— —	Ф	• • — •	Ы	— • — —
Е	•	Н	— •	Х	• • • •	Э	• • — • •
Ж	• • • —	О	— — —	Ц	— • — •	Ю	• • — —
З	— — • •	П	• — — •	Ч	— — — •	Я	• — • —

Азбука Морзе

1 ● — — — —

2 ● ● — — —

3 ● ● ● — —

4 ● ● ● ● —

5 ● ● ● ● ●

6 ● ● ● ●

7 — — ● ● ●

8 — — — ● ●

9 — — — — ●

0 — — — — —

Точка ● ● ● ● ● ●

Запятая ● — ● — ● —

/ — ● ● — ●

? ● ● — — ● ●

! — — ● ● — —

@ ● — — ● — ●

Неравномерность кода



Характерной особенностью азбуки Морзе является *переменная длина кода разных букв*, поэтому код Морзе называют *неравномерным кодом*.

Буквы, которые встречаются в тексте чаще, имеют более короткий код, чем редкие буквы. Это сделано для того, чтобы сократить длину всего сообщения. Но из-за переменной длины кода букв возникает проблема отделения букв друг от друга в тексте. Поэтому для разделения приходится использовать паузу (пропуск). Следовательно, телеграфный алфавит Морзе является троичным, т.к. в нем используются три знака: точка, тире, пропуск.

Первый беспроводной телеграф (радиоприемник)



7 мая 1895 года российский ученый Александр Степанович Попов на заседании Русского Физико-Химического Общества продемонстрировал прибор, названный им "грозоотметчик", который был предназначен для регистрации электромагнитных волн.

Этот прибор считается **первым в мире аппаратом беспроводной телеграфии, радиоприемником**. В 1897 году при помощи аппаратов беспроводной телеграфии Попов осуществил прием и передачу сообщений между берегом и военным судном.

В 1899 году Попов сконструировал модернизированный вариант приемника электромагнитных волн, где прием сигналов (**азбукой Морзе**) осуществлялся на головные телефоны оператора.

В 1900 году благодаря радиостанциям, построенным на острове Гогланд и на российской военно-морской базе в Котке под руководством Попова, были успешно осуществлены аварийно-спасательные работы на борту военного корабля "Генерал-адмирал Апраксин", севшего на мель у острова Гогланд. В результате обмена сообщениями, переданным методом беспроводной телеграфии, экипажу российского ледокола Ермак была своевременно и точно передана информация о финских рыбаках,

Телеграфный аппарат Бодо



Равномерный телеграфный код был изобретен французом Жаном Морисом Бодо в конце XIX века. В нем использовалось всего два разных вида сигналов. Не важно, как их назвать: точка и тире, плюс и минус, ноль и единица. Это два отличающихся друг от друга электрических сигнала. **Длина кода всех символов одинаковая и равна пяти.** В таком случае не возникает проблемы отделения букв друг от друга: каждая пятерка сигналов — это знак текста. Поэтому пропуск не нужен.

Код называется равномерным, если длина кода всех символов равна.

Код Бодо — это первый в истории техники способ *двоичного кодирования*, информации. Благодаря этой идее удалось создать буквопечатающий телеграфный аппарат, имеющий вид пишущей машинки. Нажатие на клавишу с определенной буквой вырабатывает соответствующий пятиимпульсный сигнал, который передается по линии связи.

В честь Бодо была названа единица скорости передачи информации — бод.

В современных компьютерах для кодирования текста также применяется равномерный двоичный код.

Двоичное кодирование в компьютере

Вся информация, которую обрабатывает компьютер должна быть представлена двоичным кодом с помощью двух цифр: **0** и **1**. *Эти два символа принято называть двоичными цифрами или битами.*

С помощью двух цифр 0 и 1 можно закодировать любое сообщение. Это явилось причиной того, что в компьютере обязательно должно быть организовано два важных процесса: кодирование и декодирование.

Кодирование – преобразование входной информации в форму, воспринимаемую компьютером, т.е. двоичный код.

Декодирование – преобразование данных из двоичного кода в форму, понятную человеку.

Почему двоичное кодирование

С точки зрения технической реализации использование двоичной системы счисления для кодирования информации оказалось намного более простым, чем применение других способов. Действительно, удобно кодировать информацию в виде последовательности нулей и единиц, если представить эти значения как два возможных устойчивых состояния электронного элемента:

0 – отсутствие электрического сигнала;

1 – наличие электрического сигнала.

Эти состояния легко различать. *Недостаток* двоичного кодирования – *длинные коды*. Но в технике легче иметь дело с большим количеством простых элементов, чем с небольшим числом сложных.

Способы кодирования и декодирования информации в компьютере, в первую очередь, зависит от вида информации, а именно, что должно кодироваться: числа, текст, графические изображения или звук.

Вопросы:

- Что такое код?
- Приведите примеры кодирования информации, используемой в физике, биологии, географии, математике?
- Придумайте свои способы кодирования русских букв.
- Закодируйте сообщение «информатика» с помощью кода Морзе.