

ПОДХОДЫ К ИЗМЕРЕНИЮ ИНФОРМАЦИИ. ИНФОРМАЦИОННЫЕ ОБЪЕКТЫ РАЗЛИЧНЫХ ВИДОВ. ДИСКРЕТНОЕ (ЦИФРОВОЕ) ПРЕДСТАВЛЕНИЕ ТЕКСТОВОЙ, ГРАФИЧЕСКОЙ, ЗВУКОВОЙ И ВИДЕОИНФОРМАЦИИ

Рассматриваемые вопросы:

1. Подходы к измерению информации.
2. Информационные объекты различных видов.
3. Дискретное (цифровое) представление информации.
4. Представление информации в двоичной системе счисления.

1. ПОДХОДЫ К ИЗМЕРЕНИЮ ИНФОРМАЦИИ

Информация, которую обрабатывает компьютер, представлена двоичным кодом с помощью двух цифр – 0 и 1. Эти два символа 0 и 1 принято называть *битами* (от английского *binary digit* – двоичный знак).

Бит – наименьшая единица измерения объема информации. Следующая по величине единица – *байт*. Остальные единица измерения информации являются производными от байта – *килобайт*, *мегабайт*, *гигабайт*, *терабайт*. Ниже в таблице 1 представлены единицы измерения информации и соотношение между ними.

Таблица 1

Единицы измерения информации

Название	Условное обозначение	Соотношение
Байт	Байт	1 Байт = 2 ³ Бит = 8 Бит
Килобайт	Кбайт	1 Кбайт = 2 ¹⁰ Байт = 1024 байт
Мегабайт	Мбайт	1 Мбайт = 2 ¹⁰ Кбайт = 1024 Кбайт
Гигабайт	Гбайт	1 Гбайт = 2 ¹⁰ Мбайт = 1024 Мбайт
Терабайт	Тбайт	1 Тбайт = 2 ¹⁰ Гбайт = 1024 Гбайт

В информатике используются следующие подходы к измерению информации: *содержательный* и *алфавитный*.

1. *Содержательный подход* к измерению информации – сообщение, уменьшающее неопределенность знаний человека в два раза, несет для него 1 бит информации. Количество информации, заключенное в сообщении, определяется по формуле *Ральфа Хартли (формула Хартли)*, которую он ввел в 1928 г: $I = \log_2 N$, где I – количество информации (бит), заключенное в сообщении, а N – количество равновероятных событий (количество вариантов). Из данной формулы также следует формула: $N = 2^I$.

2. *Алфавитный (технический) подход* к измерению информации – основан на подсчете числа символов в сообщении. Если допустить, что все символы в сообщении вычисляются по формуле: $I_c = i \cdot N$, где I_c – информационный объем сообщения, N – количество символов (мощность алфавита: $N = 2^i$), i – информационный объем 1 символа.

Далее рассмотрим примеры решения задач.

Пример 1. Переведите в биты 12 байт.

Решение: так как 1 байт = 8 битам, то 12 байт = $12 \cdot 8 = 96$ бит.

Ответ: 96 бит.

Пример 2. Переведите в байты 72 бит.

Решение: так как 1 байт = 8 битам, то 72 бит = $72 : 8 = 9$ байт.

Ответ: 9 байт.

Пример 3. Определите информационный объем сообщения «Информатика».

Решение: в слове «Информатика» 11 символов, по формуле $I_c = i \cdot N$, где i – информационный объем 1 символа, который равен 1 байту (1 символ = 1 байту), а N – количество символов, то получаем, что $I_c = 11 \cdot 1 = 11$ байт = $11 \cdot 8 = 88$ бит.

Ответ: 88 бит.

Пример 4. Какова мощность алфавита, если слово длиной 10 символов несет 30 бит информации.

Решение: мощность алфавита вычисляется по формуле $N = 2^i$, где i – информационный объем 1 символа. Так как в слове 10 символов, а количество информации равно 30 битам, то 1 символ = $\frac{30}{10} = 3$ бит, тогда мощность алфавита равна $N = 2^3 = 8$ символам.

Ответ: 8 символов.

Пример 5. Объем сообщения – 7,5 Кбайт. Известно, что данное сообщение содержит 7680 символов. Какова мощность алфавита?

Решение: воспользуемся формулой $I_c = i \cdot N$, i – информационный объем одного символа, $i = \frac{7,5 \text{ Кбайт} = 7680 \text{ байт}}{7680} = 1$ байт = 8 бит. Тогда мощность алфавита равна $N = 2^8 = 256$ символам.

Ответ: 256 символов.

2. ИНФОРМАЦИОННЫЕ ОБЪЕКТЫ РАЗЛИЧНЫХ ВИДОВ

Информационный объект – обобщающее понятие, описывающее различные виды объектов; это предметы, процессы, явления материального или нематериального свойства, рассматриваемые с точки зрения их информационных свойств.

Информационные объекты могут быть *простыми* и *комплексными*. К простым информационным объектам относятся: звук, изображение, текст, число. Комплексные (структурированные) информационные объекты – элемент, база данных, таблица, гипертекст, гипермедиа.

Стремление зафиксировать, сохранить надолго свое восприятие информации было всегда свойственно человеку и обществу в целом. Мозг человека хранит множество информации, и использует для хранения ее свойств способы, основа которых – *двоичный код*, как и у компьютеров. *Двоичный код* – это способ представления данных в виде кода,

в котором каждый разряд принимает одно из двух возможных значений 0 или 1.

Человек всегда стремился иметь возможность поделиться своей информацией с другими людьми и найти надежные средства для ее передачи и долговременного хранения. Для этого в настоящее время изобретено множество способов хранения информации на внешних носителях и ее передачи на огромные расстояния.

Существуют основные виды информации по способам ее кодирования и хранения:

1) *графическая информация* – первый вид, для которого был реализован способ хранения информации об окружающем мире в виде наскальных рисунков, а позднее в виде картин, фотографий, схем, чертежей и т.п.;

2) *звуковая информация* – мир вокруг нас полон звуков, и задача их хранения и тиражирования была решена с изобретением звукозаписывающих устройств в 1877 г. Разновидностью звуковой информации является музыкальная информация – для этого вида был изобретен способ кодирования с использованием специальных символов, что делает возможным хранение ее аналогично графической информации;

3) *текстовая информация* – способ кодирования речи человека специальными символами – буквами, причем разные народы имеют разные языки и используют различные наборы букв (алфавиты) для отображения речи. Особенно большое значение этот способ приобрел после изобретения бумаги и книгопечатания;

4) *числовая информация* – количественная мера объектов и их свойств в окружающем мире. Аналогично текстовой информации для ее отображения используется метод кодирования специальными символами – цифрами, причем системы кодирования (счисления) могут быть разными;

5) *видеоинформация* – способ сохранения «живых» картин окружающего мира, появившийся с изобретением кино.

Для передачи информации на большие расстояния первоначально использовались закодированные световые сигналы, с изобретением электричества – передача закодированного определенным образом сигнала по проводам, позднее – с использованием радиоволн.

Создатель общей теории информации и основоположник цифровой связи Клод Шеннон впервые обосновал возможность применения *двоичного кода* для передачи информации. С появлением компьютеров вначале появилось средство для обработки числовой информации. Однако в дальнейшем, особенно после широкого распространения персональных компьютеров, компьютеры стали использоваться для хранения, обработки, передачи и поиска текстовой, числовой, графической, звуковой и видеоинформации. Хранение информации при использовании компьютеров осуществляется на магнитных дисках или лентах, на лазерных дисках (CD и DVD), специальных устройствах энергонезависимой памяти (флэш-память и пр.). Эти методы постоянно совершенствуются, изобретаются новые устройства и носители информации.

Особым видом информации в настоящее время можно считать информацию, представленную в глобальной сети Интернет. Здесь используются особые приемы хранения, обработки, поиска и передачи распределенной информации больших объемов и особые способы работы с различными видами информации. С помощью компьютера возможно создание, обработка и хранение информационных объектов любых видов, для чего служат специальные программы. Ниже в таблице 2 приведены основные виды программ и соответствующие информационные объекты, которые с их помощью создаются и обрабатываются.

Таблица 2

Список программ, использующихся для обработки и хранения информационных объектов различных видов

Программы	Информационные объекты
Текстовые редакторы и процессоры	Числовая и текстовая информация
Графические редакторы и пакеты компьютерной графики	Графические объекты (чертежи, рисунки и т.п.)
Табличные процессоры	Числовые данные
СУБД – системы управления базами данных	Базы данных
Пакеты мультимедийных презентация	Компьютерные презентации
Клиент-программа электронной почты	Электронные письма, архивы, адресные с
Интернет браузер	Web-страницы, файлы из архива интерне

Данный список не является полным, так как развитие в сфере информатизации общества постоянно развивается и вносит в жизнедеятельность человека все новые и новые программы для обработки, передачи, хранения информационных объектов.

3. ДИСКРЕТНОЕ (ЦИФРОВОЕ) ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ

Дискретизация – это преобразование непрерывных изображений и звука в набор дискретных значений в форме кодов.

В настоящее время все чаще данные, изначально имеющие аналоговую форму (речь, телевизионное изображение), передаются по каналам связи в *дискретном виде*, то есть в виде последовательности единиц и нулей. Процесс представления аналоговой информации в дискретной форме называется *дискретной модуляцией*. В отличие от аналоговых машин, оперирующих непрерывной информацией, современные компьютеры имеют дело с дискретной информацией, на входе и выходе которых в качестве такой информации могут выступать любые последовательности десятичных цифр, букв, знаков препинания и других символов. Внутри системы эта информация кодируется в виде последовательности сигналов, принимающих лишь два различных значения. В то время, как возможности аналоговых машин ограничены преобразованиями строго ограниченных типов сигналов, современные компьютеры обладают свойством универсальности, иными словами, компьютер может производить преобразования любых буквенно-цифровых данных благодаря программе, составленной для выполнения той или иной задачи. Свойство универсальности компьютера не ограничивается возможностью оперирования одной лишь буквенно-цифровой информацией.

В данном виде может быть представлена (закодирована) любая дискретная информация.

Таким образом, компьютеры могут рассматриваться как универсальные преобразователи информации.

4. ПРЕДСТАВЛЕНИЕ ИНФОРМАЦИИ В ДВОИЧНОЙ СИСТЕМЕ СЧИСЛЕНИЯ

К достоинству двоичной системы счисления относится – простота совершаемых операций, возможность автоматической обработки информации с использованием двух состояний элементов ПК.

Кодирование – это операция преобразования знаков или групп знаков одной знаковой системы в знаки или группы знаков другой знаковой системы. *Декодирование* – расшифровка кодированных знаков, преобразование кода символа в его изображение. *Двоичное кодирование* – кодирование информации в виде 0 и 1.

Способы кодирования и декодирования информации в компьютере, в первую очередь, зависит от вида информации, а именно, что должно кодироваться: числа, символьная информация (буквы, цифры, знаки), графические изображения, звук.

1. *Двоичное кодирование чисел* – для записи информации о количестве объектов используются числа. Числа записываются с использованием особых знаковых систем, которые называют системами счисления.

2. *Двоичное кодирование текста* – присвоение символу определенного кода – это вопрос соглашения, которое фиксируется в кодовой таблице. В качестве международного стандарта была принята кодовая таблица ASCII (American Standard Code for Information Interchange) – на 1 символ отводится 1 байт (8 бит), всего можно закодировать 256 символов. С 1997 г. появился новый международный стандарт Unicode, который отводит для кодировки одного символа 2 байта (16 бит), и можно закодировать 65536 различных символов.

3. *Двоичное кодирование графики* – пространственная дискретизация – перевод графического изображения из аналоговой формы в цифровой компьютерный формат путем разбиения изображения на отдельные маленькие фрагменты (точки) где каждому элементу присваивается код цвета. Пиксель – минимальный участок изображения на экране, заданного цвета. Растровое изображение формируется из отдельных точек – пикселей, каждая из которых может иметь свой цвет. Двоичный код изображения, выводимого на экран храниться в видеопамяти. Кодирование рисунка растровой графики напоминает – мозаику из квадратов, имеющих определенный цвет. Для хранения черно-белого изображения используется 1 бит, цветные изображения формируются в соответствии с двоичным кодом цвета, который хранится в видеопамяти. Цветное изображение на экране формируется за счет смешивания трех базовых цветов – красного, зеленого и синего.

4. *Двоичное кодирование звука* – в аналоговой форме звук представляет собой волну с непрерывно меняющейся амплитудой и частотой. На компьютере работать со звуковыми файлами начали с начала 90-х гг. В основе

кодирования звука с использованием ПК лежит – процесс преобразования колебаний воздуха в колебания электрического тока и последующая дискретизация аналогового электрического сигнала. Кодирование и воспроизведение звуковой информации осуществляется с помощью специальных программ (редактор звукозаписи). Временная дискретизация – способ преобразования звука в цифровую форму путем разбивания звуковой волны на отдельные маленькие временные участки, где амплитуды этих участков квантуются (им присваивается определенное значение).

5. *Представление видеоинформации* – в последнее время компьютер все чаще используется для работы с видеоинформацией. Простейшей такой работой является просмотр кинофильмов и видеоклипов. Следует четко представлять, что обработка видеоинформации требует очень высокого быстродействия компьютерной системы. Что представляет собой фильм с точки зрения информатики? Прежде всего, это сочетание звуковой и графической информации. Кроме того, для создания на экране эффекта движения используется дискретная по своей сути технология быстрой смены статических картинок.