

Дистанционное задание по информатике

1. Прочитать и законспектировать нижеследующий материал урока.
 2. Выполнить задания для самостоятельной работы (в конце файла).
 3. Ответить на контрольные вопросы (в конце файла).
 4. Результат работы (фото своего конспекта) прислать для проверки на электронный ящик: **aktstudy@yandex.ru**
 5. Срок сдачи работы: **два дня** с момента данного занятия по расписанию.
-

Кодирование и единицы информации

Языки и алфавиты

Различают естественные и искусственные (формальные) языки. *Естественные языки* развивались веками и служат для общения людей между собой. *Формальные языки* разрабатываются для специальных применений. Примером формальных языков могут служить языки программирования, языки кодирования информации для ее передачи, хранения и т. п.

Каждый язык имеет свой *алфавит*. Под алфавитом языка понимают набор используемых символов. Под *мощностью алфавита* понимают количество составляющих алфавит символов.

Кодом называют совокупность знаков (символов), предназначенных для представления информации в соответствии с определенными правилами. Такое представление называют *кодированием*. Кодируют информацию с целью ее передачи, хранения, преобразования.

Одно и то же понятие на различных языках может кодироваться различными способами. Например, слово "стол" — это код в русском алфавите всем известного предмета мебели. В других языках, в других алфавитах этот предмет кодируется иначе.

Звук *a*, издаваемый человеком, кодируется в некоторых языках буквой А. Буква А в азбуке Морзе кодируется так: ". —" (точка, тире). В компьютере

буква А латинского алфавита в привычной для нас десятичной системе кодируется числом 65. В свою очередь, число 65 в "привычной" для компьютера двоичной системе (цифры только 0 и 1) кодируется так: 01000001.

В естественных языках традиционно сложились некоторые неоднозначности. Например, под словом "коса" понимается и девичья коса, и речная отмель, и инструмент для скашивания травы. Подобные неоднозначности в формальных языках, как правило, недопустимы.

Различны и алфавиты языков кодирования. Это буквы А, В, С, D,... и буквы А, Б, В, Г, Д,... в латинском и русском алфавитах; точка и тире в азбуке Морзе, арабские цифры 0, 1, 2, 3, ... , 9, с помощью которых записываются числа, кодирующие различные количества; красный, желтый и зеленый цвет в светофоре; цифры 0 и 1 в компьютере и т. п.

Количество и графическое изображение символов в алфавитах естественных языков определяются характерными особенностями языка, историей его развития, традициями. Например, русский алфавит имеет 33 буквы, латинский — 26, итальянский — 21, армянский — 39, арабский — 28, китайский — несколько тысяч иероглифов. Количество и изображение знаков в формальных языках могут также существенно различаться.

Наименьший по числу знаков алфавит имеет только один знак. Пусть этот знак **1** (единица). Тогда три цвета светофора можно закодировать, например, так: красный — 1, желтый — 11, зеленый — 111. Такой алфавит самый неэкономичный по записи кодов. В этом легко убедиться, если попытаться записать в этом алфавите, например, число десять: 1111111111.

Двоичный алфавит

В информатике и вычислительной технике широко используется алфавит, имеющий два знака, например, 1 и 0. Этим знакам в логике и технике приводят в соответствие понятия — да и нет, истина и ложь, включено и выключено. Такой алфавит называют *двоичным* или *бинарным*. В соответствии с этим введена и наименьшая единица информации — *бит* (англ. *bit*, от *binary* — двоичный и *digit* — знак).

Одного бита информации достаточно, чтобы передать слово "да" или "нет", закодировать, например, состояние электролампочки. Кстати, на некоторых выключателях пишут "1 — включено" и "0 — выключено". Взгляд на выключатель снимает для нас неопределенность в его состоянии. При этом мы получаем количество информации, равное одному биту.

Пример 1

Пусть требуется отгадать задуманное число из набора натуральных чисел от 1 до 16, задавая вопрос: "В первой половине находится задуманное число?" и получая в ответ "Да" или "Нет".

Задав первый вопрос и услышав первый ответ, получаем один бит информации. При этом неопределенность уменьшается в два раза (из 16 чисел остается только 8). После второго вопроса получаем еще один бит информации, уменьшающий неопределенность еще в два раза (из 8 чисел осталось только 4). После получения еще одного бита информации неопределенность снова уменьшилась вдвое (из 4 чисел осталось 2 числа). Последний, третий вопрос позволяет отгадать число.

Для отгадывания потребовалось получить четыре ответа, четыре бита информации. Столько же бит требуется для оптимального кодирования каждого из шестнадцати чисел.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Где число?	_____															
Где число?	1	2	3	4	5	6	7	8								
Где число?					5	6	7	8								
Где число?					5	6										
Где число?					5											

Рис. 3.2. Пример процедуры отгадывания числа

На рис. 3.2 показана последовательность шагов отгадывания числа 5. На вопросы "Где число?" дается ответ в виде жирной линии. При решении задачи использован широко применяемый, быстрый прием поиска путем деления всего множества чисел пополам.

Двоичное слово. Байт

Если нужно закодировать в двоичном алфавите красный, желтый и зеленый цвет светофора, то требуется уже два бита. Закодировать три цвета можно, например, так: 00, 01 и 10. Сообщение о том, что включен, например, красный цвет светофора, содержит информации больше одного бита.

Для кодирования четырех сторон света (север, восток, юг и запад) требуется также два бита: 00, 01, 10, 11. Поэтому сообщение о том, какая выбрана сторона света, содержит ровно два бита информации.

При кодировании восьми углов куба потребуется три бита: 000, 001, 010, 011, 100, 101, 110, 111. При кодировании от 9 до 16 объектов потребуется уже четыре бита, от 17 до 31 — 5 бит, от 32 до 63 — 6 бит, от 64 до 127 — 7 бит.

Последовательность символов называют *словом*. Можно сделать вывод: чем больше требуется закодировать объектов, тем длиннее требуемое двоичное слово. Восьмибитовое двоичное слово называется *байтом*. С помощью байта можно закодировать 256 различных объектов.

До недавнего времени байта было достаточно, чтобы закодировать все символы текста в русском и латинском алфавите: буквы, цифры, знаки препинания, управляющие сигналы — все то, что передавалось компьютеру с клавиатуры.

С развитием информатики байт начал сдерживать возможность увеличения количества используемых символов. В настоящее время завершается переход на двухбайтовое кодирование символов. 16-битовое двоичное слово позволяет закодировать 65 536 символов и команд.

Единицы информации

На практике используются более объемные, производные единицы количества (объема) информации:

1 байт = 8 бит;

1 килобайт (Кбайт) = 2^{10} байтов = 1024 байтов;

1 мегабайт (Мбайт) = 2^{10} килобайтов = 1024 килобайтов;

1 гигабайт (Гбайт) = 2^{10} мегабайтов = 1024 мегабайтов;

1 терабайт (Тбайт) = 2^{10} гигабайтов = 1024 гигабайтов,

и т. д.

Пример 2

Подсчитаем объем памяти, требуемый для записи и хранения в памяти компьютера книги "Программируем на языке Quick BASIC 4.5" Г. А. Зельднера. В книге 420 страниц. На каждой странице в среднем по 40 строк. В каждой строке по 60 символов. Итого в книге $60 \times 40 \times 420 \approx 1\,000\,000$ знаков. Каждый знак требует для записи 1 байт памяти. Поэтому для записи всей книги нужно около 1 Мбайт. А это всего одна гибкая дискета с объемом памяти 1,44 Мбайт.

Часто ошибочно отождествляют понятия количества и объема информации. Поясним их различие на примере. Пусть имеется большой энциклопедический словарь в 1400 страниц и такого же объема книга, на каждой странице которой все строки заполнены вопросительными знаками. Объем словаря и книги одинаков — 1400 страниц текста. Поэтому для хранения словаря и книги требуется одинаковый объем памяти. Количество же информации в словаре и книге существенно различается. Всю информацию книги можно записать одной фразой: "1400 страниц с вопросительными знаками". Но даже эта фраза имеет сомнительную ценность и для многих людей не несет никакой полезной информации.

В целях экономии памяти при записи информации на хранение (при архивации) широко используют процедуру ее сжатия. Так, информацию словаря

можно сжать примерно в десять раз. При этом следует понимать, что путем сжатия уменьшается не количество информации, а только ее объем.

Кодирование графики и звука

С помощью компьютера рисуют, сочиняют музыку; компьютер понимает человеческий голос и сам может отвечать пользователю. Так как вся информация в компьютере обрабатывается и сохраняется в цифровом виде, то и графическая, и звуковая информация должны быть представлены в цифровом виде. С этой целью ее кодируют.

Кодирование растровой графики

Весь экран дисплея делится на точки — пиксели, подсвечивание которых создает видимое отображение текста и рисунков. Различают растровую и векторную графику.

При растровом способе создания, хранения и отображения графических объектов приходится иметь дело со всеми точками, входящими в изображение. Например, если цветная фотография, полученная с помощью цифровой фотокамеры, имеет размеры 1600×1200 точек, то в памяти компьютера необходимо хранить информацию (цвет, яркость) о каждой из точек фотографии.

Информационный объем растрового изображения (Q) определяется как произведение числа входящих в изображение точек (N) на информационный объем одной точки (q), который зависит от количества возможных цветов, т. е. $Q = N \cdot q$.

При черно-белом изображении $q = 1$ бит (например, 1 — точка подсвечивается и 0 — точка не подсвечивается). Поэтому для хранения черно-белого (без оттенков) изображения размером 100×100 точек требуется 10 000 бит. Если между черным и белым цветом имеется еще шесть оттенков серого (всего 8), то информационный объем точки равен 3 бита ($\log_2 8 = 3$). Информационный объем такого изображения увеличивается в три раза: $Q = 30\,000$ бит.

Цветное изображение получается за счет различной яркости трех основных цветов — красного, синего и зеленого. Цветные изображения могут отображаться в различных режимах, соответственно изменяется и информационный объем точки:

Режим	Информационный объем точки
16 цветов	$q = \log_2 16 = 4$ бита
256 цветов	$q = \log_2 256 = 8$ бит = 1 байт
65 536 цветов	$q = \log_2 65536 = 16$ бит = 2 байта
16 777 216 цветов	$q = \log_2 16\,777\,216 = 24$ бит = 3 байта

Умножение информационного объема точки на количество точек дает значительную величину. Поэтому актуальной становится задача расчета объема памяти (видеопамяти), необходимой для хранения изображения (битовой карты) со всего экрана монитора. Рассмотрим два примера.

Пример 3

При размерах экрана (разрешении) 640×480 точек и количестве цветов 256 ($q = 1$ байт) объем видеопамяти должен быть не менее $Q = 640 \cdot 480 \cdot 1$, т. е. около 300 Кбайт.

Пример 4

Для экрана размером 1280×1024 точки и количестве цветов 16 777 216 ($q = 3$) имеем $Q = 1280 \cdot 1024 \cdot 3$, что примерно равно 3,75 Мбайта.

Современные компьютеры имеют объем видеопамяти 8, 16 и более Мбайт, что позволяет хранить несколько графических изображений размером во весь экран.

Кодирование векторной графики

Векторное изображение состоит из набора элементарных деталей — графических примитивов. Это линии, прямоугольники, окружности, дуги и т. п. Положение этих элементов на экране определяется координатами точек. Например, отрезок прямой задается координатами концов, окружность — координатами центра и радиусом и т. д. Кроме координат задаются цвет отображаемых деталей, толщина линий и другие характеристики.

Информация о векторном рисунке кодируется обычным способом, как хранятся тексты, формулы, числа, т. е. хранится не графическое изображение, а только координаты и характеристики изображения его деталей. Поэтому для хранения векторных изображений требуется существенно меньше памяти, чем для изображений растровых.

При запуске программы с векторным рисунком он создается каждый раз вновь, после чего в растровом виде может сохраняться в видеопамяти. При просмотре графических изображений они считываются компьютером из видеопамяти на экран дисплея с частотой 50—60 и более раз в секунду.

Кодирование звука

Человек слышит звук с частотами от 16 герц (колебаний в секунду) до 20 килогерц. До недавнего времени звуковые сигналы передавались, записывались на граммофонных пластинках и магнитных лентах и воспроизводились в виде амплитудных колебаний. При передаче радио- и телевизионных программ использовалась амплитудная или частотная модуляция так называе-

мой несущей частоты радиосигнала низкочастотными звуковыми сигналами. На приемных устройствах звуковые сигналы усиливались и подавались на репродукторы и динамики. В процессе передачи по каналам связи на звуковые сигналы могут накладываться различной природы внешние электромагнитные воздействия, которые искажают сигналы.

Значительное увеличение быстродействия компьютеров и других цифровых устройств, увеличение памяти компьютеров позволяют переходить на цифровую передачу и воспроизведение аудио и телевизионных передач, сохранять их в памяти компьютера, хранить на лазерных CD- и DVD-дисках. При этом существенно повышается надежность и качество воспроизведения.

Задания для самостоятельной работы

1. Придумать примеры: а) информации, имеющей различную ценность для различных людей, что подтверждает субъективность в оценках информации; б) информации, ценность которой существенно зависит от ее оперативности; в) информации, ценность которой зависит от ее полноты.
2. Определить число бит, потребное для кодирования в двоичном алфавите: а) всех пальцев на руках и ногах; б) всех номеров квартир стапятидесятиквартирного дома.
3. Сколько бит информации нужно получить, чтобы отгадать загаданное число из: а) 16 последовательно расположенных чисел натурального ряда; б) 32 чисел.
4. Выбрать любую книгу и подсчитать потребный для ее записи объем памяти.

Контрольные вопросы:

- Что такое мощность алфавита?
- Что такое Код, Кодирование?
- Объясните понятия Бит, Байт.
- Перечислите известные Вам единицы измерения информации.
- В чем отличие кодирования растровой и векторной графики?

Вывод (кратко перечислить то, что делали на занятии).