

Лекция № 1.

Информация: свойства, измерение, кодирование

Элементы теории информации. Основные понятия. Информация, сообщения, данные, сигнал. Основные информационные процессы. Виды и свойства информации. Объемная (техническая) мера информации. Единицы измерения информации. Энтропийный подход в измерении количества информации. Представление и кодирование информации. Двоичное кодирование.

1. Понятие информации с точки зрения различных наук

«Информация – это не материя и не энергия, информация – это информация.»

Норберт Винер¹

В любой науке основные понятия определить достаточно сложно. Так и понятие «информация» невозможно определить через другие, более «простые» понятия. В зависимости от области знания существуют различные подходы к определению этого термина.

Впервые как научное понятие термин «информация» стал применяться в **теории журналистики** в 30-х годах XX века, хотя в исследованиях по **библиотечному делу** он появился еще раньше. Под информацией понимались различные сведения, сообщения. Что соответствует переводу с латинского языка *informatio* – сведение, разъяснение, ознакомление.

[ГОСТ 7.0-99](#)

Информационно-библиотечная деятельность, библиография. Термины и определения.

Информация – сведения, воспринимаемые человеком или специальными устройствами как отражение фактов материального мира в процессе коммуникации.

В **физике** понятие «информация» рассматривается как антиэнтропия или энтропия² с обратным знаком. Поскольку мерой беспорядка термодинамической системы является энтропия системы, то информация (антиэнтропия) является мерой упорядоченности и сложности системы. По мере увеличения сложности системы величина энтропии уменьшается, и величина информации увеличивается. Процесс увеличения информации характерен для открытых, обменивающихся веществом и энергией с окружающей средой, саморазвивающихся систем живой природы (белковых молекул, организмов, популяций животных и т.д.).

В **неживой природе** понятие «информация» связано с понятием отражения, отображения. В некоторых физических и химических теориях информация определяется как отраженное многообразие. *Отражение заключается в таком изменении одного материаль-*

¹ Винер Н. Кибернетика, или управление и связь в животном и машине; или Кибернетика и общество/ 2-е издание. — М.: Наука; Главная редакция изданий для зарубежных стран, 1983. — 344 с.

² Понятие **энтропии** впервые было введено в термодинамике в 1865 году для определения меры необратимого рассеивания энергии, меры отклонения реального процесса от идеального. Определённая как сумма приведённых теплот, она является функцией состояния и остаётся постоянной при замкнутых обратимых процессах, тогда как в необратимых — её изменение всегда положительно. Математически энтропия определяется как функция состояния системы, равная в равновесном процессе количеству теплоты, сообщённой системе или отведённой от системы, отнесённому к термодинамической температуре системы.

ного объекта под воздействием другого, при котором все особенности отражаемого объекта каким-либо образом воспроизводятся отражающим объектом. В процессе отражения и происходит передача информации. Т.е. информация - это результат отражения. В соответствии с этим взглядом информация существовала и будет существовать вечно, она содержится во всех элементах и системах материального мира.

Информация, наряду с веществом и энергией, является неотъемлемым свойством материи.

Под информацией в **технике** понимают сообщение, передаваемое с помощью знаков и символов. В **теории связи**, например, под информацией принято понимать любую последовательность символов, не учитывая их смысл.

В основанной американским ученым Клодом Шенноном³ **математической теории информации** под информацией понимались не любые сведения, а лишь те, которые снимают полностью или уменьшают существующую до их получения неопределенность (неизвестность). Каждому сигналу в теории Шеннона соответствует вероятность его появления. Например, при передаче текста телеграммы вероятность появления буквы «т» равна 1/33. Чем меньше вероятность появления того или иного сигнала, тем больше информации он несет для потребителя. В обыденном понимании это означает, что, чем неожиданнее новость, тем больше ее информативность.

Информация – сведения об объектах окружающей среды, их параметрах, свойствах и состоянии, которые уменьшают имеющуюся о них степень неопределенности, неполноты знаний.

Наш соотечественник математик Ю.А. Шрейдер⁴ оценивал информацию по увеличению объема знаний у человека под воздействием информационного сообщения. Академик А.А. Харкевич⁵ измерял содержательность сообщения по увеличению вероятности достижения цели после получения информации человеком или машиной. Таким образом, под информацией в **семантической теории** понимают сведения, обладающие новизной.

В **кибернетике** – науке об управлении в живых, неживых и искусственных системах – понятие информации связывают воедино с понятием управления (Норберт Винер⁶). Жизнедеятельность любого организма или нормальное функционирование технического устройства зависит от процессов управления, благодаря которым поддерживаются в необходимых пределах значения их параметров. Процессы управления включают в себя получение, хранение, преобразование и передачу информации. Информация является обозначением содержания, полученного из внешнего мира в процессе приспособления к нему

³ **Клод Элвуд Шеннон** (англ. *Claude Elwood Shannon*; 30 апреля 1916, Мичиган, США — 24 февраля 2001, Медфорд, Массачусетс, США) — американский инженер и математик, его работы являются синтезом математических идей с конкретным анализом чрезвычайно сложных проблем их технической реализации. Является основателем **теории информации**, нашедшей применение в современных высокотехнологических системах связи. Шеннон внес огромный вклад в теорию вероятностных схем, теорию автоматов и теорию систем управления — области наук, входящие в понятие «кибернетика». В 1948 году предложил использовать слово «бит» для обозначения наименьшей единицы информации.

⁴ **Юлий Анатольевич Шрёйдер** (28 октября 1927, Днепропетровск — 24 августа 1998) — советский и российский математик, кибернетик и философ.

⁵ **Александр Александрович Харкевич** (21 января (3 февраля) 1904, Санкт-Петербург — 30 марта 1965, Москва) — советский учёный, член-корреспондент АН УССР, действительный член АН СССР, профессор, к 1951 г. — руководитель отдела технической физики киевского Института физики АН УССР.

⁶ **Норберт Винер** (англ. *Norbert Wiener*; 26 ноября 1894, Колумбия, штат Миссури, США — 18 марта 1964, Стокгольм, Швеция) — американский учёный, выдающийся математик и философ, основоположник кибернетики и теории искусственного интеллекта.

наших чувств. Информацию составляет та часть знания, которая используется для ориентирования, принятия решений, активного действия, управления, т.е. в целях сохранения, совершенствования и развития системы. Данная концепция отрицает существование информации в неживой природе, не дает ответа на вопросы: являются ли информацией неиспользованные знания, являются ли информацией неосмысленная информация?

Для преодоления этих противоречий академик В.Г. Афанасьев⁷ ввел понятие *информационных данных*.

Информационные данные – это всякие сведения, сообщения, знания, которые могут храниться, перерабатываться, передаваться, но характер информации они приобретут лишь тогда, когда получают содержание и форму, пригодную для управления, и используются в управлении.

Дальнейшим развитием математического подхода к феномену информация послужили работы Р. Карнапа⁸, И. Бар-Хиллела⁹, А.Н. Колмогорова¹⁰ и др. В них понятие информации не связано с содержанием сообщений, передаваемых по каналу связи. Информация – абстрактная величина, не существующая в физической реальности, подобно тому, как не существует мнимое число или не имеющая линейных размеров материальная точка.

В **биологии**, которая изучает живую природу, понятие «информация» связано с целесообразным поведением живых организмов. Такое поведение строится на основе получения и использования организмом информации об окружающей среде.

Понятие информация используется в связи с исследованием механизмов наследственности. В **генетике** сформулировано понятие генетической информации, которое определяется как программа (код) биосинтеза белков, представленных цепочками ДНК. Реализуется эта информация в ходе развития особи. Последнее обстоятельство позволило проводить научные эксперименты по клонированию, т.е. созданию точных копий организмов из одной клетки.

В **социальных науках** (социологии, психологии, политологии и др.) под информацией понимают сведения, данные, понятия, отраженные в нашем сознании и изменяющие наши представления о реальном мире. Эту информацию, передающуюся в человеческом обществе и участвующую в формировании общественного сознания, называют социальной информацией.

Под информацией в **документалистике** понимают все то, что так или иначе зафиксировано в знаковой форме в виде документов.

С точки зрения индивидуального **человеческого сознания** информация – это то, что поступает в наш мозг из многих источников в разных формах и, взаимодействуя там, образует структуру нашего знания. Под информацией в **быту** (житейский аспект) понимают сведения об окружающем мире и протекающем в нем процессах, воспринимаемые чело-

⁷ **Виктор Григорьевич Афанасьев** (18 ноября 1922 — 10 апреля 1994) — советский философ и деятель КПСС, академик Российской академии наук (1991; академик АН СССР с 1981). В трудах Афанасьева В.Г. исследуются проблемы системности, организационной теории, теории и практики управления, филос. аспекты теории операций, информатики (автор одной из первых кн. по соц. Информатике).

⁸ **Рудольф Карнап** (нем. *Rudolf Carnap*); (18 мая 1891, Вупперталь, Германия — 16 сентября 1970, Санта-Мария, Калифорния, США) — немецко-американский философ и логик, ведущий представитель логического позитивизма и философии науки.

⁹ **Йегошуа Бар-Хиллел** (8 сентября 1915, Вена — 25 сентября 1975, Иерусалим) — израильский философ, математик и лингвист. Известен своими пионерскими работами в области машинного перевода и формальной лингвистики.

¹⁰ **Андрей Николаевич Колмогоров** (урождённый *Катаев*, 12 (25) апреля 1903, Тамбов — 20 октября 1987, Москва) — советский математик, один из крупнейших математиков XX века.

веком или специальными устройствами. Информацией для человека являются не только сухие факты, строгие инструкции, но и то, что радует нас, волнует, печалит, заставляет переживать, восторгаться, презирать, негодовать. Более половины общего объема сведений, полученных в процессе разговора, приходится на так называемую бессмысловую информацию. Эту информацию говорящий по своему желанию, а иногда и непроизвольно, сообщает нам своей тональностью разговора, своей возбужденностью, жестиком, выражением лица, глаз и т.д.

Определения понятия «информация» из международных стандартов:

- знания о предметах, фактах, идеях и т. д., которыми могут обмениваться люди в рамках конкретного контекста (ISO/IEC 10746-2:1996)¹¹;
- знания относительно фактов, событий, вещей, идей и понятий, которые в определенном контексте имеют конкретный смысл (ISO/IEC 2382-1:1993)¹².

Определений информации существует множество, причём академик Н. Н. Моисеев¹³ даже полагал, что в силу широты этого понятия нет и не может быть строгого и достаточно универсального определения информации. Как бы не варьировалось понятие «информация», главное в нем то, что оно уничтожает незнание о чем-либо.

2. Информация и данные

Информация и данные - это базовые понятия, которые используются в информатике. Эта наука занимается вопросами систематизации, хранения, обработки и передачи данных и информации средствами вычислительной техники. Эти понятия зачастую используются как синонимы, но между ними существуют и принципиальные различия.

Мы живем в материальном мире. Все, что нас окружает и с чем мы сталкиваемся ежедневно, относится либо к *физическим телам*, либо к *физическим полям*. Из курса физики мы знаем, что состояния абсолютного покоя не существует и физические объекты находятся в состоянии непрерывного движения и изменения, которое сопровождается обменом энергией и ее переходом из одной формы в другую.

Все виды *энергообмена* сопровождаются появлением *сигналов* (от лат. signum – знак), то есть все сигналы имеют в своей основе материальную энергетическую основу. При взаимодействии сигналов с физическими телами в последних возникают определенные изменения свойств – это явление называется регистрацией сигналов. Такие изменения можно наблюдать, измерять или фиксировать иными способами – при этом возникают и регистрируются новые сигналы, то есть образуются *данные* (рис. 1).

Данные – это зарегистрированные сигналы.

Отметим, что данные несут в себе информацию о событиях, произошедших в материальном мире, поскольку они являются регистрацией сигналов, возникших в результате

¹¹ ISO/IEC 10746-2:1996, Information technology — Open Distributed Processing — Reference Model: Foundations.3.2.5: *knowledge that is exchangeable amongst users about things, facts, concepts, and so on, in a universe of discourse.*

¹² ISO/IEC 2382-1:1993, Information technology — Vocabulary — Part 1: Fundamental terms.01.01.01: *knowledge concerning objects, such as facts, events, things, processes, or ideas, including concepts, that within a certain context has a particular meaning.*

¹³ **Никита Николаевич Моисеев** (23 августа 1917, Москва — 29 февраля 2000, там же) — советский и российский учёный в области общей механики и прикладной математики.

этих событий. Однако данные не тождественны информации. Приведем несколько примеров, иллюстрирующих данное утверждение.

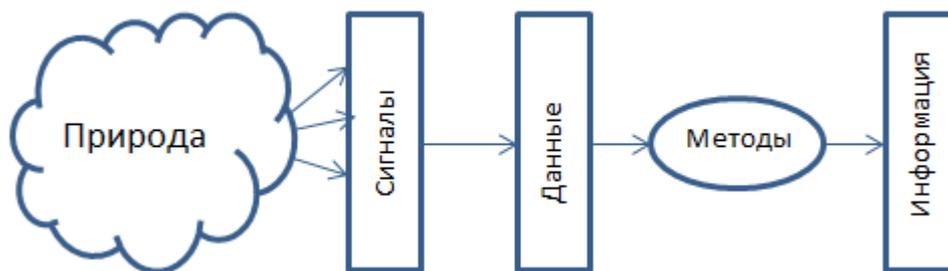


Рис. 1. Схема информационного процесса

Наблюдая за состязанием бегунов, мы с помощью механического секундомера регистрируем начальное и конечное положение стрелки прибора. В итоге мы измеряем величину перемещения за время забега – это регистрация данных. Однако информации о времени преодоления дистанции мы пока не получаем. Для того чтобы данные о перемещении стрелки дали информацию о времени забега, необходимо наличие метода пересчета одной физической величины в другую.

Прослушивая передачу радиостанции на незнакомом языке, мы получаем данные, но не получаем информацию в связи с тем, что не владем методом преобразования данных в известные нам понятия. Если эти данные записать на лист бумаги или на магнитную ленту, изменится форма представления, произойдет новая регистрация и, соответственно, образуются новые данные. Такое преобразование можно использовать, чтобы все-таки извлечь информацию из данных путем подбора метода, адекватного их новой форме. Для обработки данных, записанных на листе бумаги адекватным может быть метод перевода со словарем, а для обработки данных, записанных на магнитной ленте, можно пригласить переводчика, обладающего своими методами перевода.

Данные - это совокупность сведений, зафиксированных на определенном носителе в форме, пригодной для постоянного хранения, передачи и обработки. Преобразование и обработка данных позволяет получить информацию. Например, в базах данных хранятся различные данные, а по определенному запросу система управления базой данных выдает требуемую информацию.

Можно сказать, что **информация** – это продукт взаимодействия данных и адекватных им методов, это результат преобразования и анализа данных.

3. Общая характеристика процессов сбора, передачи, обработки и накопления информации

Сбор информации – это процесс получения информации из внешнего мира и приведение ее к стандарту для данной информационной системы. Обмен информацией между воспринимающей ее системой и окружающей средой осуществляется посредством сигналов. **Сигнал** – средство передачи информации в пространстве и времени. В качестве носителя сигнала могут выступать звук, свет, электрический ток, магнитное поле и т.д. Сбор информа-

ции, как правило, сопровождается ее регистрацией, т.е. фиксацией информации на материальном носителе (документе или машинном носителе).

Передача информации осуществляется различными способами: с помощью курьера, пересылка по почте, доставка транспортными средствами, дистанционная передача по каналам связи. Для осуществления последней необходимы специальные технические средства. Дистанционно может передаваться как первичная информация с мест ее возникновения, так и результатная информация в обратном направлении. Поступление информации по каналам связи осуществляется двумя способами: на машинном носителе и непосредственно в компьютер при помощи специальных программных и аппаратных средств.

Преобразование (обработка) информации – внесение изменений в набор данных, вычисления, информационный поиск, сортировка, построение графиков и т.п.

В современных развитых информационных системах машинная **обработка информации** предполагает последовательно-параллельное во времени решение вычислительных задач. Это возможно при наличии определенной организации вычислительного процесса. Вычислительная задача по мере необходимости обращается с запросами в вычислительную систему. Организация процесса предполагает определение последовательности решения задач и реализацию вычислений. Последовательность решения задается, исходя из информационной взаимосвязи, когда результаты решения одной задачи используются как исходные данные для решения другой.

Технология электронной **обработки информации** – человеко-машинный процесс исполнения взаимосвязанных операций, протекающих в установленной последовательности с целью преобразования исходной информации (первичной) в результатную. Операция представляет собой комплекс совершаемых технологических действий, в результате которых информация **преобразуется**. Технологические операции разнообразны по сложности, назначению, технике реализации, выполняются на различном оборудовании разными исполнителями.

Хранение и накопление информации вызвано многократным ее использованием, применением постоянной информации, необходимостью комплектации первичных данных до их обработки. Хранение осуществляется на машинных носителях в виде информационных массивов, где данные располагаются по установленному в процессе проектирования группировочному признаку.

4. Виды и свойства информации

Классификация информации может быть выполнена по различным критериям. Например,

- 1) *по форме представления*: графическая, текстовая, числовая, звуковая, видео;
- 2) *по способу восприятия*: визуальная, аудиальная, тактильная, обонятельная, вкусовая;
- 3) *по стадии обработки*: первичная, вторичная, промежуточная, результатная;
- 4) *по стабильности*: переменная, постоянная;
- 5) *по функции управления*: плановая, нормативно-справочная, учетная, оперативная.

Качество информации – положительная характеристика информации, отражающая степень ее полезности для пользователя.

Показатель качества – одно из важных положительных свойств информации (с позиции потребителя). К **потребительским показателям качества информации** можно отнести следующие:

- **Репрезентативность** – правильность отбора и формирования информации в целях адекватного отражения свойств объекта.
- **Содержательность** – семантическая емкость. С увеличением содержательности информации растет семантическая пропускная способность информационных систем, так как для получения одних и тех же сведений требуется преобразовать меньший объем данных.
- **Достаточность (полнота)** – свойство информации исчерпывающе (для данного потребителя) характеризовать отображаемый объект и/или процесс.
- **Доступность** – свойство информации, характеризующее возможность ее получения данным потребителем.
- **Актуальность** – определяется степенью сохранения ценности информации для управления в момент ее использования и зависит от динамики изменения ее характеристик и от интервала времени, прошедшего со времени решения поставленной задачи.
- **Своевременность** – означает ее поступление не позже заранее назначенного момента времени, согласованного со временем решения поставленной задачи.
- **Точность** – определяется степенью близости получаемой информации к реальному состоянию объекта, процесса, явления и т. п.
- **Достоверность** – определяется ее свойством отражать реально существующие объекты с необходимой точностью.
- **Устойчивость** – отражает способность информации реагировать на изменение исходных данных без нарушения необходимой точности.
- **Защищенность** – свойство, характеризующее невозможность несанкционированного использования или изменения.
- **Эргономичность** – свойство, характеризующее удобство формы или объем информации с точки зрения данного потребителя.

Внутренние свойства информации:

- Объем информации.
- Способ организации информации:
 - 1) *Данные*, т.е. логически не упорядоченный набор сведений;
 - 2) *Структура данных* – логически упорядоченные наборы данных.

Свойства информации, связанные с процессом хранения

- **Живучесть** – способность информации сохранять свое качество с течением времени.
- **Уникальность** – информация хранится в единственном экземпляре.

Внешние свойства информации по отношению к отражаемому объекту:

Адекватность – свойство информации однозначно соответствовать отображаемому объекту или явлению.

5. Измерение информации. Количество информации: формулы Хартли и Шеннона.

Для информатики информация играет такую же роль, как вещество в физике. И подобно тому, как веществу можно приписывать довольно большое количество характеристик: массу, заряд, объем и т.д., так и для информации имеется свой набор характеристик, для

которых имеются единицы измерения. Это позволяет некоторой порции информации приписывать числа – *количественные характеристики информации*.

В технике часто используют способ определения количества информации называемый *объемным* (или *техническим*). Он основан на подсчете числа символов в *сообщении*, т.е. связан с его длиной и не зависит от содержания.

Сообщение – это информация, представленная в определенной форме и предназначенная для передачи.

В вычислительной технике вся обрабатываемая и хранимая информация вне зависимости от ее природы представлена в двоичной форме (с использованием алфавита, состоящего всего из двух символов 0 и 1). Это позволило ввести две стандартные единицы измерения: **бит** (binary digit) и **байт** (bate).

Бит – минимальная единица измерения информации, которая представляет собой двоичный знак двоичного алфавита {0;1}.

Байт – единица количества информации в СИ, представляющая собой восьмиразрядный двоичный код, с помощью которого можно представить один символ.

Информационный объем сообщения (информационная емкость сообщения) - количество информации в сообщении, измеренное в стандартных единицах или производных от них (Кбайтах, Мбайтах и т. д.)¹⁴.

	1 байт = 8 бит	В
Кило-	1 Кбайт = 2^{10} байт = 1024 байт	КВ
Мега-	1 Мбайт = 2^{10} Кбайт = 2^{20} байт	МВ
Гига-	1 Гбайт = 2^{10} Мбайт = 2^{20} Кбайт = 2^{30} байт	ГВ
Тера-	1 Тбайт = 2^{10} Гбайт = 2^{20} Мбайт = 2^{30} Кбайт = 2^{40} байт	ТВ
Пета-	1 Пбайт = 2^{50} байт	ПВ
Экса-	1 Эбайт = 2^{60} байт	ЕВ
Зетта-	1 Збайт = 2^{70} байт	ЗВ
Йотта-	1 Йбайт = 2^{80} байт	УВ

В соответствии с объемным методом совокупность 100 букв — фраза из 100 букв из газеты, пьесы Шекспира или теории относительности Эйнштейна — имеет в точности одинаковое количество информации.

В теории информации принят и другой подход к измерению информации, называемый *энтропийным*. Энтропийный подход исходит из следующей модели. Получатель сообщения имеет определенные представления о возможных наступлениях некоторых событий. Эти представления в общем случае недостоверны и выражаются вероятностями, с которыми он ожидает то или иное событие. *Общая мера неопределенности* – **энтропия** - характеризуется некоторой математической зависимостью от совокупности этих вероятностей. Количество информации в сообщении определяется тем, насколько уменьшается эта мера после получения сообщения.

Количеством информации называют числовую характеристику сигнала, не зависящую от его формы и содержания, и характеризующую неопределенность, которая исчезнет после получения сообщения в виде данного сигнала.

¹⁴ Измерения в байтах. ГОСТ 8.417-2002

Для абсолютно достоверного события (событие обязательно произойдет, поэтому его вероятность равна 1) количество информации в сообщении о нем равно 0. Чем неожиданнее событие, тем больше информации он несет.

Научный подход к оценке сообщений был предложен в 1928 году **Ральф Хартли**¹⁵. Проиллюстрируем его следующим образом. Пусть имеется алфавит, из букв которого составляется сообщение. Количество букв в алфавите равно m . В этом случае количество возможных равновероятных вариантов разных сообщений длины n будет равно $N = m^n$.

Пример 1. Пусть алфавит состоит из двух букв «А» и «В», длина сообщения 3 буквы. Тогда $m=2$, $n=3$. С таким алфавитом и длине сообщения можно составить $N = 2^3 = 8$ разных сообщений («ААА», «ААВ», «АВА», «АВВ», «ВАА», «ВАВ», «ВВА», «ВВВ»).

Формула Хартли определяет количество информации, содержащейся в сообщении длины n :

$$H = \log_2 N = n \log_2 m,$$

H – количество информации (бит¹⁶). Для приведенного выше примера $H = 3$ бит.

Если N – это количество равновероятных событий, то количество информации в сообщении о том, что произошло одно из N событий, равно $H = \log_2 N$ бит.

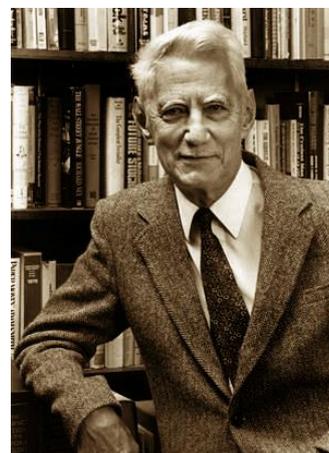
Пример 2. В поезде 16 вагонов. Сообщение о том, что ваш знакомый приехал во 2-м вагоне, несет $H = \log_2 16 = 4$ бита информации.

Более общий подход к вычислению количества информации в сообщении об одном из N , но уже неравновероятных событий, был предложен **Клодом Шенноном**¹⁷ в 1948 году.

Формула Шеннона. Если известна, вероятность p_A наступления некоторого события A , то количество информации в сообщении о том, что произошло событие A , равно $H = -\log_2 p_A$ бит.



Р. Хартли



Кл. Шеннон

¹⁵ **Ральф Винтон Лайон Хартли** (англ. *Ralph Vinton Lyon Hartley*, 30 ноября 1888, Спрус, Невада — 1 мая 1970, Нью-Джерси) — американский учёный-электронщик. Он предложил генератор Хартли, преобразование Хартли и сделал вклад в **теорию информации**, введя в 1928 году логарифмическую меру информации $H = n \log_2 m$, которая называется *хартлиевским количеством информации* или просто *мерой Хартли*.

Источник фотографии: http://upload.wikimedia.org/wikipedia/commons/thumb/0/01/Hartley_ralph-vinton-lyon-001.jpg

¹⁶ Название единицы измерения зависит от значения a основания логарифма $\log_a x$.

a	2	e	3	10	256
Название единицы измерения	бит	нат	трит	дит (хартли)	байт

¹⁷ Источник фотографии: <https://upload.wikimedia.org/wikipedia/ru/3/3d/Shannon.jpg>

Пример 3. В составе поезда 16 вагонов, из них 4 – купейные. Сообщение о том, что ваш знакомый приехал в купейном вагоне несет $H = -\log_2 \frac{4}{16} = 2$ бит информации.

Таким образом, информацию можно измерять длиной сообщения в битах. Такой способ ничего не говорит об информативности сообщения, но зато характеризует объём работы системы связи при передаче. Если же в задаче необходимо учитывать информативность, то следует пользоваться энтропийным подходом к измерению информации. При этом нужно уточнить, о каком множестве событий будет сообщаться, каковы их вероятности.

6. Представление и кодирование информации

Для обмена информации с другими людьми человек использует естественные языки (русский, английский, китайский и др.), т.е. информация представляется с помощью естественных языков. В основе языка лежит **алфавит**, т.е. набор символов (знаков), которые человек различает по их начертанию. Например, в основе русского языка лежит кириллица, содержащая 33 знака, английский язык использует латиницу (26 знаков), китайский язык использует иероглифы (десятки тысяч знаков).

Последовательности символов алфавита в соответствии с правилами **грамматики** образуют основные объекты языка – слова. Правила, согласно которым образуются предложения из слов данного языка, называют **синтаксисом**.

Наряду с естественными языками были разработаны и формальные языки (системы счисления, язык алгебры, языки программирования и др.). При этом алфавит могут составлять цифры, символы, формулы, ноты, изображения элементов электрических или логических схем, дорожные знаки, точки и тире и т.п. Основное отличие формальных языков от естественных состоит в наличии строгих правил грамматики и синтаксиса.

Представление информации может осуществляться с помощью **языков**, которые являются знаковыми системами. Каждая **знаковая система** строится на основе определенного **алфавита** и правил выполнения операций над знаками.

Знаки могут иметь различную физическую природу. Например, для представления информации с использованием языка в письменной форме, используются знаки, которые являются изображениями на бумаге или других носителях, в устной речи в качестве знаков языка используют различные звуки (фонемы), а при обработке текста на компьютере знаки представляются в виде последовательностей электрических импульсов (компьютерных кодов).

Преобразование информации из одной формы представления в другую называют **кодированием**. Средством кодирования служит таблица соответствия знаковых систем, которая устанавливает взаимно-однозначное соответствие между знаками или группой знаков двух различных знаковых систем.

В процессе обмена информацией часто приходится производить операции **кодирования** и **декодирования** информации. Например, при вводе знака алфавита в компьютер путем нажатия соответствующей клавиши на клавиатуре происходит кодирование знака, т.е. преобразование его в компьютерный код. При выводе знака на экран монитора или

принтер происходит обратный процесс – декодирование, когда из компьютерного кода знак преобразуется в его графическое изображение.

Кодирование – это операция преобразования знаков или групп знаков одной знаковой системы в знаки или группы знаков другой знаковой системы.

При кодировании информации ставятся следующие цели:

- 1) удобство физической реализации;
- 2) удобство восприятия;
- 3) высокая скорость передачи и обработки;
- 4) экономичность, т.е. уменьшение избыточности сообщения;
- 5) надежность, т.е. защита от случайных искажений;
- 6) сохранность, т.е. защита от нежелательного доступа к информации.

Эти цели часто противоречат друг другу. Например, стремясь к экономным сообщениям, мы тем самым уменьшаем их надежность и удобство восприятия.

На разных этапах обработки информации достигаются разные цели и поэтому информация неоднократно перекодируется, преобразуется из вида, удобного для восприятия человеком, к виду, удобному для обработки автоматическими системами, и наоборот.

Двоичное кодирование

В компьютере для представления информации используется двоичное кодирование, т.к. используются технические устройства, которые могут сохранять и распознавать не более двух различных состояний (цифр):

- Электромагнитные реле (замкнуто/разомкнуто), широко использовались в конструкциях первых ЭВМ;
- участок поверхности магнитного носителя информации (намагничен/размагничен);
- участок поверхности лазерного диска (отражает/не отражает) и т.д.

Все виды информации в компьютере кодируются на машинном языке, в виде логических последовательностей нулей и единиц.

Информация в компьютере представлена в **двоичном коде**, алфавит которого состоит из **двух цифр** (0 и 1).

Цифры двоичного кода можно рассматривать как два равновероятных состояния (события). При записи двоичной цифры реализуется выбор одного из двух возможных состояний (одной из двух цифр).

Каждая цифра машинного двоичного кода несет количество информации, равное одному биту.

Представление чисел в компьютере

Целые числа кодируются двоичным кодом достаточно просто. При этом используются правила перевода чисел в двоичную систему счисления¹⁸. Целое число надо делить на 2 до тех пор, пока в остатке не образуется ноль или единица. Совокупность остатков от каждо-

¹⁸ **Система счисления** – это знаковая система, в которой числа записываются по определенным правилам с помощью символов некоторого алфавита, называемых цифрами. См. тему для самостоятельного изучения: http://pmik.karelia.ru/user/semenova/Informatika/DOC/Sam_izuch/Notations.pdf

го деления, записанная справа налево вместе с последним остатком, и образует двоичный аналог десятичного числа. Например, для числа 19 имеем:

	19	9	4	2	1
остаток	1	1	0	0	1

Таким образом, $19_{10} = 10011_2$.

Для кодирования целых чисел от 0 до 255 достаточно иметь 8 разрядов двоичного кода (8 бит, или 1 байт).

Количество бит K , которое требуется для кодирования N различных значений равно наименьшему целому числу, большему или равному $\log_2 N$. Или, следует найти целое K , для которого справедливо следующее двойное неравенство: $2^{K-1} < N \leq 2^K$.

Шестнадцать бит позволяют закодировать целые числа от 0 до 65 535, а 24 бита – уже более 16,5 млн. разных значений.

Количество различных значений, которое может быть закодировано с помощью K бит равно.

Целые числа без знака, переведенные в двоичную систему счисления, занимают один, два или большее количество байт, левые незначащие разряды заполняются нулями. Так, например, в однобайтном формате число 19 записывается следующим образом:

0	0	0	1	0	0	1	1
---	---	---	---	---	---	---	---

Целые числа со знаком переводятся в двоичную систему счисления и записываются в **прямом**, **обратном** или **дополнительном** коде. Они могут занимать один, два или большее количество байт, незначащие разряды заполняют нулями, кроме самого левого (старшего). Этот разряд содержит информацию о знаке числа. Знак «плюс» кодируется нулем, а «минус» — единицей.

Положительные числа в прямом, обратном и дополнительном коде имеют одинаковое представление, в знаковом разряде стоит ноль.

Для отрицательных чисел:

- 1) **прямой код** формируется как двоичное представление модуля числа, а в знаковый разряд ставится единица, так для -19_{10} :

1	0	0	1	0	0	1	1
---	---	---	---	---	---	---	---

- 2) **Обратный код** получается инвертированием (0 заменяется на 1, 1 – на 0) всех разрядов прямого кода, кроме знакового разряда, так для -19_{10} :

1	1	1	0	1	1	0	0
---	---	---	---	---	---	---	---

- 3) **Дополнительный код** получается из обратного кода добавлением единицы к младшему разряду, при этом перенос в знаковый разряд игнорируется, так для -19_{10} :

1	1	1	0	1	1	0	1
---	---	---	---	---	---	---	---

Специальные коды (прямой и дополнительный) используют для упрощения выполнения операций алгебраического сложения чисел в ЭВМ.

При кодировании действительного числа его предварительно преобразуют в *нормализованную форму*:

$$3,1415926 = 0,3141596 \cdot 10^1,$$

$$300000 = 0,3 \cdot 10^6,$$

$$123 = 0,123 \cdot 10^3,$$

$$0,0123 = 0,123 \cdot 10^{-1}.$$

Первая часть числа называется *мантиссой*, а вторая – *характеристикой*. Мантисса и порядок числа хранятся как целые числа со знаком.

Кодирование текстовой информации

Если каждому символу алфавита сопоставить определенное целое число (например, порядковый номер), то с помощью двоичного кода можно кодировать и текстовую информацию.

Если для кодирования одного символа использовать 1 байт = 8 бит информации, то с помощью одного байта можно закодировать $N = 2^I = 2^8 = 256$ символов.

Такое количество символов вполне достаточно для представления текстовой информации, включая прописные и строчные буквы русского и латинского алфавита, цифры, знаки основных арифметических операций и некоторые общепринятые специальные символы, например, символ «§». При этом кодирование заключается в установлении однозначного соответствия символа и уникального десятичного кода от 0 до 255 или двоичного кода от 00000000 до 11111111. Если человек различает символы по начертаниям, то компьютер по их кодам.

Присвоение символу конкретного кода – это вопрос соглашения, которое фиксируется в **кодовой таблице**.

Институтом стандартизации США (ANSI – American National Standard Institute) в 1963 году была введена в действие система кодирования ASCII (*American Standard Code for Information Interchange* – *стандартный код информационного обмена США*). В системе ASCII закреплены две таблицы кодирования – *базовая* и *расширенная*. Базовая таблица (рис. 2) закрепляет значения кодов от 0 до 127, а расширенная относится к символам с номерами от 128 до 255.

ASCII Code Chart																
	0	1	2	3	4	5	6	7	8	9	A	B	C	D	E	F
0	NUL	SOH	STX	ETX	EOT	ENQ	ACK	BEL	BS	HT	LF	VT	FF	CR	SO	SI
1	DLE	DC1	DC2	DC3	DC4	NAK	SYN	ETB	CAN	EM	SUB	ESC	FS	GS	RS	US
2		!	"	#	\$	%	&	'	()	*	+	,	-	.	/
3	0	1	2	3	4	5	6	7	8	9	:	;	<	=	>	?
4	@	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
5	P	Q	R	S	T	U	V	W	X	Y	Z	[\]	^	_
6	`	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o
7	p	q	r	s	t	u	v	w	x	y	z	{		}	~	DEL

Рис. 2. Базовая таблица кодировки ASCII¹⁹

¹⁹ <https://ru.wikipedia.org/wiki/ASCII>

Первые 32 кода (0-31) базовой таблицы отданы производителям аппаратных средств (компьютеров и печатающих устройств). В этой области размещаются так называемые *управляющие коды*, которым не соответствуют никакие символы языков, и, соответственно, эти коды не выводятся ни на экран, ни на устройства печати, но ими можно управлять тем, как производится вывод других данных или действия с другими данными (например, код 8 соответствует возврату на один символ назад (BACKSPACE)).

Начиная с кода 32 по код 127 размещены коды символов английского алфавита, знаков препинания, цифр, арифметических действий и некоторых вспомогательных символов.

Поддержка производителей оборудования и программ вывела американский код ASCII на уровень международного стандарта, и национальным системам кодирования пришлось «отступить» во вторую, расширенную часть системы кодирования, определяющую значения кодов со 128 по 255 (рис. 3).

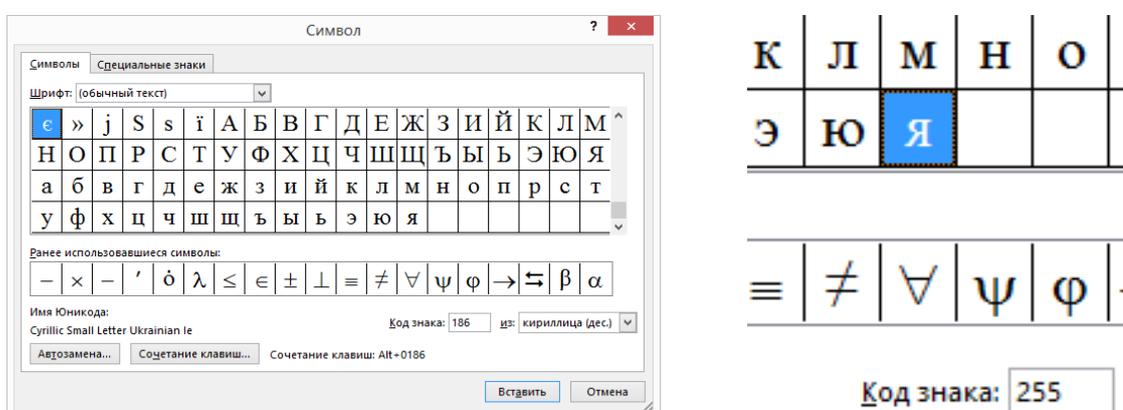


Рис. 3. Расширенная таблица кодировки ASCII (кириллица)

Отсутствие единого стандарта в этой области привело к множественности одновременно действующих кодировок. Только в России можно указать три действующих стандарта кодировки (Windows-1251, КОИ-8²⁰, ISO²¹).

Если проанализировать организационные трудности, связанные с созданием единой системы кодирования текстовых данных, то можно прийти к выводу, что они вызваны ограниченным набором кодов (всего 256). В то же время очевидно, что если, например, кодировать символы не восьмиразрядными двоичными числами, а числами с большим количеством разрядов, то и диапазон возможных значений кодов станет намного большим. Такая система, основанная на 16-разрядном кодировании символов, получила название *универсальной* – UNICODE. Шестнадцать разрядов позволяют обеспечить уникальные коды для $N = 2^{16} = 65\,536$ различных символов – этого количества кодов достаточно для размещения в одной таблице символов большинства языков планеты.

Несмотря на тривиальную очевидность такого подхода, простой механический переход на данную систему долгое время сдерживался из-за недостаточных ресурсов средств

²⁰ **КОИ-8** – Код обмена информацией(8 бит), восьмибитовая кодовая таблица. Ее происхождение относится ко временам действия Совета Экономической Безопасности государств Восточной Европы. Сегодня эта кодировка имеет широкое распространение в компьютерных сетях на территории России и в российском секторе Интернета.

²¹ ISO (International Standard Organization – Международный институт стандартизации) - международный стандарт, в котором предусмотрена кодировка символов русского алфавита.

вычислительной техники (в системе кодирования UNICODE все текстовые документы автоматически становятся вдвое длиннее). Во второй половине 90-х годов прошлого века технические средства достигли необходимого уровня обеспеченности ресурсами, и сегодня идет постепенный перевод документов и программных средств на универсальную систему кодирования. Эту кодировку поддерживают, начиная с 1997 года, Microsoft Windows&Office.

Кодирование графической информации

Существует несколько способов кодирования графической информации.

При рассмотрении черно-белого графического изображения с помощью увеличительного стекла заметно, что в его состав входит несколько мельчайших точек, образующих характерный узор (или *растр*) (рис. 4). Растровое компьютерное изображение состоит из пикселей.

Пиксель, пиксел (англ. **pixel** — сокращение от **pix element**) — физическая точка или наименьший логический элемент двумерного цифрового изображения.

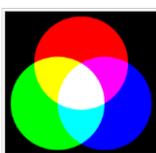
Линейные координаты и индивидуальные свойства каждой из точек изображения можно выразить с помощью целых чисел, поэтому способ *растрового кодирования* базируется на использовании двоичного кода представления графических данных. Общеизвестным стандартом считается приведение черно-белых иллюстраций в форме комбинации точек с 256 градациями серого цвета (рис. 5), т. е. для кодирования яркости любой точки необходимы 8-разрядные двоичные числа.



Рис. 4. Растр – это метод кодирования графической информации



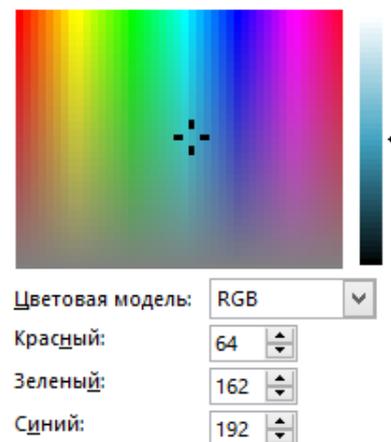
Рис. 5. Градации серого цвета



В основу кодирования цветных графических изображений положен принцип разложения произвольного цвета на основные составляющие (*принцип декомпозиции*), в качестве которых применяются три основных цвета: красный (Red), зеленый (Green) и синий (Blue). На практике при-

нимается, что любой цвет, который воспринимает человеческий глаз, можно получить с помощью механической комбинации этих трех цветов. Такая система кодирования называется RGB (по первым буквам основных цветов).

Если для кодирования яркости каждой из основных составляющих использовать по 256 значений (это 8 бит), то на кодирование цвета одной точки требуется 24 бита. С помощью 24 бит можно закодировать более 16,5 млн. различных цветов ($2^{24} = 16\,777\,216$). Режим представления цветной графики с использованием 24 двоичных разрядов называется *полноцветным* (True Color).



Каждому из основных цветов можно сопоставить дополнительный цвет, т.е. цвет, дополняющий основной цвет до белого. Для любого из основных цветов дополнительным будет являться цвет, который образован суммой пары остальных основных цветов. Соответственно среди дополнительных цветов можно выделить голубой (Cyan, C), пурпурный (Magenta, M) и желтый (Yellow, K). Принцип разложения произвольного цвета на составляющие компоненты используется не только для основных цветов, но и для дополнительных, т.е. любой цвет можно представить в виде суммы голубой, пурпурной и желтой составляющей. Этот метод кодирования цвета применяется в полиграфии, но там используется еще и четвертая краска – черная (Black, K), поэтому эта система кодирования обозначается четырьмя буквами – CMYK. Для представления цветной графики в этой системе применяется 32 двоичных разряда. Данный режим также носит название полноцветного.

При уменьшении количества двоичных разрядов, применяемых для кодирования цвета каждой точки, сокращается объем данных, но заметно уменьшается диапазон кодируемых цветов. Кодирование цветной графики 16-разрядными двоичными числами носит название режима High Color.

При кодировании графической цветной информации с применением 8 бит данных можно передать только 256 оттенков. Данный метод кодирования цвета называется *индексным*. Код каждой точки раstra выражает не цвет сам по себе, а только его номер (индекс) в некоторой справочной таблице, называемой *палитрой*. Эта таблица должна прикладываться к графическим данным, так как без нее нельзя адекватно воспроизвести информацию на экране или принтере.

Двоичное кодирование звуковой информации

В настоящий момент не существует единой стандартной системы кодирования звуковой информации, так как приемы и методы работы со звуковой информацией начали развиваться по сравнению с методами работы с другими видами информации самыми последними. Поэтому множество различных компаний, которые работают в области кодирования информации, создали свои собственные корпоративные стандарты для звуковой

информации. Но среди этих корпоративных стандартов выделяются два основных направления.

В основе *метода FM* (Frequency Modulation) положено утверждение о том, что теоретически любой сложный звук может быть представлен в виде разложения на последовательность простейших гармонических сигналов разных частот. Каждый из этих гармонических сигналов представляет собой правильную синусоиду и поэтому может быть описан числовыми параметрами или закодирован. Звуковые сигналы образуют непрерывный спектр, т. е. являются аналоговыми, поэтому их разложение в гармонические ряды и представление в виде дискретных цифровых сигналов выполняется с помощью специальных устройств – *аналого-цифровых преобразователей* (АЦП). Обратное преобразование, которое необходимо для воспроизведения звука, закодированного числовым кодом, производится с помощью *цифро-аналоговых преобразователей* (ЦАП). Из-за таких преобразований звуковых сигналов возникают потери информации, которые связаны с методом кодирования, поэтому качество звукозаписи с помощью метода *FM* обычно получается недостаточно удовлетворительным и соответствует качеству звучания простейших электромузыкальных инструментов с окраской, характерной для электронной музыки. При этом данный метод обеспечивает вполне компактный код, поэтому он широко использовался в те годы, когда ресурсы средств вычислительной техники были явно недостаточны.

Основная идея *метода таблично-волнового синтеза* (Wave-Table) состоит в том, что в заранее подготовленных таблицах находятся образцы звуков для множества различных музыкальных инструментов. Данные звуковые образцы носят название *сэмплов*. Числовые коды, которые заложены в сэмпле, выражают такие его характеристики, как тип инструмента, номер его модели, высоту тона, продолжительность и интенсивность звука, динамику его изменения, некоторые компоненты среды, в которой наблюдается звучание, и другие параметры, характеризующие особенности звучания. Поскольку для образцов применяются реальные звуки, то качество закодированной звуковой информации получается очень высоким и приближается к звучанию реальных музыкальных инструментов, что в большей степени соответствует нынешнему уровню развития современной компьютерной техники.

Выводы

Все процессы в природе сопровождаются сигналами. Зарегистрированные сигналы образуют данные. Данные преобразуются, транспонируются и потребляются с помощью методов. При взаимодействии данных и адекватных им методов образуется информация.

Информация – это динамический объект, образующийся в ходе информационного процесса.

С целью унификации приемов и методов работы с данными в вычислительной технике применяется универсальная система кодирования, называемая двоичным кодом. Элементарной единицей представления данных в двоичном коде является двоичный разряд (бит)

Литература и Internet-источники

1. Симонович С.В. Информатика. Базовый курс: Учебник для вузов. Стандарт третьего поколения. – СПб.: Питер, 2015. – 640 с.
2. Острейковский В.А. Информатика: Учеб. для вузов. – М.: Высшая школа, 2001. – 511 с.
3. <https://sites.google.com/site/rgpuktnoscience/Home/sr/inf>
4. <https://ru.wikipedia.org/wiki/>
5. <http://www.e-reading.club/book.php?book=97791>
6. <http://u4isna5.ru/konspektlekcii/35-lekciiinformatika/101-informatlekcija3>



Вопросы и задачи на измерение информации

1. Сколько байт в 5 Кбайт? Сколько бит в 5 Кбайт?
Ответ: 5 Кбайт = $5 \times 1024 = 5120$ байт,
 5 Кбайт = $5120 \times 8 = 40960$ бит.
2. Какой будет последняя цифра числа 7896543126710 в двоичной системе счисления?
Решение:
 Поскольку число является четным, то последней цифрой в его двоичной записи будет 0.
3. Какое наибольшее натуральное число можно закодировать 4 битами?
Ответ: $1 \cdot 2^3 + 1 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0 = 15$,
 или $2^4 - 1 = 15$.
4. Сколько целых неотрицательных чисел можно закодировать 5 битами?
Ответ: $2^5 = 32$.
5. Чему равна минимальная длина равномерных двоичных кодов для букв русского алфавита (33 буквы)?
Решение:
 Так как $2^5 < 33 < 2^6$, то минимальная длина двоичного кода равна 6.
6. В урне 8 черных и 56 белых шаров. Чему равно количество бит информации в сообщении "выбранный из урны шар является черным"?
Решение:
 Вероятность того, что выбранный из урны шар является черным равна: $\frac{8}{8+56} = \frac{1}{8}$.
 Тогда количество информации в сообщении «выбранный из урны шар является черным» равна: $-\log_2 \frac{1}{8} = 3$ бита.
7. Чему равно количество различных символов, закодированных полубайтами в сообщении 100110001111100010011110?
Решение:
 Разбив текст сообщения на группы из четырех цифр, получим
 1001 1000 1111 1000 1001 1110
 Полученное разбиение содержит 4 различные группы цифр:
 1001, 1000, 1111, 1110

Таким образом, в сообщении 4 различных символов.

8. Чему равно максимальное количество книг (каждая объемом 200 страниц, на каждой странице 60 строк, 80 символов в строке), полностью размещенных на лазерном диске емкостью 600 Мбайт?

Решение:

Общее количество символов в одной книге равно: $200 \times 60 \times 80 = 960\,000$.

Так как одному символу соответствует 1 байт, то одна книга занимает 960 000 байт.

Емкость лазерного диска составляет: $600 \times 1024 \times 1024 = 629\,145\,600$ байт.

Так как $\frac{629145600}{960000} = 655,36$, то максимально количество книг, которое может быть размещено на лазерном диске, равно 655.

9. Емкость одного условного печатного листа (усл. п. л.) равна приблизительно 96 Кбайтам, а 1 символ занимает 8 бит. Сколько минут потребуется для распечатки текста одной газеты (4 усл. п. л.) на лазерном принтере (скорость печати - 512 символов в секунду) без учета смены бумаги? (ответ округлите до целого числа)

Решение:

Так как один символ занимает один байт, то один условный печатный лист содержит 96×1024 символов.

Для печати 4-х листов потребуется: $\frac{96 \cdot 1024 \cdot 4}{512} = 768$ секунд,

что составляет: $\frac{768}{60} = 12,8 \approx 13$ минут

10. Вариант теста в среднем имеет объем 20 килобайт (на каждой странице теста 40 строк по 64 символа в каждой, 1 символ занимает 8 бит). Чему равно количество страниц в тесте?

Решение:

Каждая страница содержит 40×64 символов, что составляет 2560 байт.

Количество страниц в тексте равно: $\frac{20 \cdot 1024}{2560} = 8$.

11. Сколько байт составляет фраза «Я изучаю информационные технологии», если для кодирования символов используется таблица ASCII?

Решение:

Фраза содержит 34 символа, что в кодировке ASCII соответствует 34 байтам.

12. Для хранения растрового изображения размером 128×128 пикселей отвели 4 Кбайта памяти. Какое максимальное количество цветов в палитре изображения?

Решение:

Определим количество пикселей в изображении: $128 \times 128 = 2^7 \times 2^7 = 2^{14}$.

Отведенный объем памяти в битах составляет: 4 Кбайт = 4×2^{10} байт = $2^2 \times 2^{10} \times 2^3 = 2^{15}$ бит.

Тогда на один пиксель изображения приходится $2^{15}/2^{14} = 2$ бита.

С помощью 2 бит можно закодировать $2^2 = 4$ цвета.

13. Укажите минимальный объем памяти (в Кбайтах), достаточный для хранения любого растрового изображения размером 32×32 пикселя, если известно, что в изображении

используется палитра из 256 цветов (не учитывая объем памяти, необходимый для хранения палитры).

Решение:

Исходя из количества цветов в палитре определим минимальное количество двоичных разрядов, необходимое для хранения одного пикселя. Для представления 256 различных состояний требуется $\log_2 256 = 8$ двоичных разрядов, т.е. 1 байт.

Поэтому для представления изображения размером 32×32 пикселя потребуется:

$32 \times 32 = 2^5 \times 2^5 = 2^{10}$ байт информации, т.е. 1 Кбайт.

Контрольные вопросы

1. Что понимают под термином «информация»?
2. Перечислите свойства информации.
3. Назовите различия в понятиях *информация, сообщения, данные, сигнал*.
4. Что такое бит и байт?
5. Единицы измерения информации. Информационный объем.
6. Характеристика качества информации, заключающаяся в достаточности данных для принятия решений, - это ...
 - A. Полнота;
 - B. Репрезентативность;
 - C. Доступность;
 - D. Актуальность
7. Семантический аспект – это характеристика информации с точки зрения ее ...
 - A. Качества;
 - B. Количества;
 - C. Смысла;
 - D. Полезности
8. Что является наименьшей единицей измерения информации?
9. Расположите единицы измерения количества информации *Гигабайт, Мегабит, Мегабайт, Терабайт* в порядке возрастания.
10. Сколько различных значений можно закодировать, используя четыре двоичных разряда?
 - A. 2;
 - B. 4;
 - C. 8;
 - D. 16
11. Количество чисел, которое можно закодировать нулями и единицами в 10 позициях, равно:
 - A. 1000;
 - B. 1024;
 - C. 10;
 - D. 256;
 - E. 512.
12. Сколько двоичных знаков необходимо и достаточно для того, чтобы закодировать целые числа от 0 до 5?
 - A. 1;
 - B. 2;
 - C. 3;
 - D. 4
13. Сколько символов содержит кодовое пространство, задаваемое ASCII?
 - A. 1024;
 - B. 215;
 - C. 8;
 - D. 256
14. Информационный объем одной точки черно-белого растрового изображения равен:
 - A. 1 биту;
 - B. 2 битам;
 - C. 1 байту;
 - D. 2 байтам
15. Информационный объем одной точки 256-цветного растрового изображения равен:
 - A. 1 биту;
 - B. 2 битам;
 - C. 1 байту;
 - D. 2 байтам
16. С помощью одного байта при двоичном кодировании можно представить целое неотрицательное число от нуля до ...

- A. 1; B. 256; C. 255; D. 1024
17. При подключении к Интернету модем обеспечивает скорость передачи данных, равную 28 800 бит/с. Сколько времени потребуется для передачи файла размером 72 000 байта?
- A. 5 сек; B. 10 сек; C. 20 сек; D. 60 сек
18. Модем передает сообщения со скоростью 28 800 бит/с. Сколько секунд потребуется для передачи 100 страниц текста в 30 строк по 60 символов каждая в кодировке ASCII?
- A. 50; B. 0,02; C. 62,5; D. 60
19. Количество различных символов, закодированных байтами в сообщении 1110010100111100011111011110010101111101, равно:
- A. 3; B. 4; C. 5; D. 1; E. 2.
20. Количество различных символов, закодированных полубайтами в сообщении 10111000101110001001, равно:
- A. 2; B. 4; C. 5; D. 1; E. 3.
21. Скольким битам равно количество информации в слове «ИНФОРМАТИКА» при условии, что для кодирования используется 32-значный алфавит?
- A. 11; B. 32; C. 55; D. 352
22. При перекодировке сообщения из кода Unicode в код ASCII объем сообщения изменился на 1/512 Мб. Сколько символов содержит сообщение?
- A. 1024; B. 256; C. 2048; D. 64
23. Растровый графический файл содержит черно-белое изображение с 2 градациями цвета (черный и белый) размером 800 x 600 точек. Определите необходимый для кодирования цвета точек (без учета информации о формате, авторстве, способах сжатия и пр.) размер этого файла на диске в байтах.
- A. 480 000; B. 60 000; C. 3 840 000; D. 480
24. Какое наибольшее натуральное число можно закодировать 16 битами?
25. Сколько байт, необходимо для записи числа 2^{44} , 2^{82} , 8^{14} ?
26. Вариант теста в среднем имеет объем 20 килобайт (на каждой странице теста 40 строк по 64 символа в каждой, 1 символ занимает 8 бит). Сколько страниц в тесте?
27. Сведения о сотруднике хранятся в виде строки из 2048 символов (один символ занимает 8 бит). На каком минимальном числе дискет емкостью 1.2Мбайт можно разместить сведения обо всех 8192 сотрудниках?
28. Чему равно максимальное количество страниц книги (32 строки по 64 символа, 1 символ занимает 8 бит), которое поместится в файле объемом 640 Кбайт?
29. Чему равен информационный объем одной точки черно-белого растрового изображения?
30. Во сколько раз увеличится информационная емкость файла, содержащего растровое изображение, если повысить глубину его цвета со стандарта «Черно-белое» до стандарта «65 536 цветов»?
31. 256-цветный рисунок содержит 1 Кбайт информации. Из какого количества точек он состоит?