

Виды информации. Представление информации. Хранилища информации. Способы кодирования и носители информации. Понятие модели и алгоритма в информатике.

Виды информации

Термин информация происходит от латинского слова *informatio*, что означает "сведения, разъяснения, изложение".

Информация - это настолько общее и глубокое понятие, что его нельзя объяснить одной фразой. В это слово вкладывается различный смысл в технике, науке и в житейских ситуациях.

В обиходе информацией называют любые данные или сведения, которые кого-либо интересуют, к примеру сообщение о каких-либо событиях, о чьей-либо деятельности и т. п. "Информировать" в данном смысле означает "сообщить нечто, неизвестное раньше". Информация — сведения об объектах и явлениях окружающей среды, их параметрах, свойствах и состоянии, которые воспринимают информационные системы (живые организмы, управляющие машины и др.) в процессе жизнедеятельности и работы.

Одно и то же информационное сообщение (статья в газете, объявление, письмо, телеграмма, справка, рассказ, чертеж, радиопередача и т. п.) может содержать разное количество информации для разных людей исходя из их накопленных знаний, от уровня понимания этого сообщения и интереса к нему. Так, сообщение, составленное на японском языке, не несет никакой новой информации человеку, не знающему этого языка, но должна быть высокоинформативным для человека, владеющего японским. Никакой новой информации не содержит и сообщение, изложенное на знакомом языке, в случае если его содержание непонятно или уже известно.

Информация есть характеристика не сообщения, а соотношения между сообщением и его потребителем. Без наличия потребителя, хотя бы потенциального, говорить об информации бессмысленно.

В случаях, когда говорят об автоматизированной работе с информацией посредством каких-либо технических устройств, обычно в первую очередь интересуются не содержанием сообщения, а тем, сколько символов это сообщение содержит.

Применительно к компьютерной обработке данных под информацией понимают некоторую последовательность символических обозначений (букв, цифр, закодированных графических образов и звуков и т. п.), несущую смысловую нагрузку и представленную в понятном компьютеру виде. Каждый новый символ в такой последовательности символов увеличивает информационный объем сообщения.

Основные виды информации по ее форме представления, способам ее кодирования и хранения, что имеет наибольшее значение для информатики, это:

1. *Графическая или изобразительная* — первый вид, для которого был реализован способ хранения информации об окружающем мире в виде наскальных рисунков, а позднее в виде картин, фотографий, схем, чертежей на бумаге, холсте, мраморе и др.
Размещено на реф.рф
материалах, изображающих картины реального мира;
2. *Звуковая* — мир вокруг нас полон звуков и задача их хранения и тиражирования была решена с изобретением звукозаписывающих устройств в 1877 г. ее разновидностью является музыкальная информация — для этого вида был изобретен способ кодирования с использованием специальных символов, что делает возможным хранение ее аналогично графической информации;
3. *Текстовая* — способ кодирования речи человека специальными символами — буквами, причем разные народы имеют разные языки и используют различные наборы букв для отображения речи; особенно большое значение данный способ приобрел после изобретения бумаги и книгопечатания;
4. *Числовая* — количественная мера объектов и их свойств в окружающем мире; особенно большое значение приобрела с развитием торговли, экономики и денежного обмена; аналогично текстовой информации для ее отображения используется метод кодирования специальными символами — цифрами, причем системы кодирования (счисления) бывают разными;
5. *Видеоинформация* — способ сохранения "живых" картин окружающего мира, появившийся с изобретением кино.

Существуют также виды информации, для которых до сих пор не изобретено способов их кодирования и хранения — это тактильная информация, передаваемая ощущениями, органолептическая, передаваемая запахами и вкусами и др.

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

Создателем общей теории информации и основоположником цифровой связи считается Клод Шеннон (Claude Shannon). Всемирную известность ему принес фундаментальный труд 1948 года — "Математическая теория связи" (A Mathematical Theory of Communication), в котором впервые обосновывается возможность применения двоичного кода для передачи информации.

С появлением компьютеров (или, как их вначале называли в нашей стране, ЭВМ — электронные вычислительные машины) вначале появилось средство для обработки числовой информации. При этом в дальнейшем, особенно после широкого распространения персональных компьютеров (ПК), компьютеры стали использоваться для хранения, обработки, передачи и поиска текстовой, числовой, изобразительной, звуковой и видеоинформации. С момента появления первых персональных компьютеров — ПК (80-е годы 20 века) — до 80 % их рабочего времени посвящено работе с текстовой информацией.

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. Обработку информации (воспроизведение, преобразование, передача, запись на внешние носители) выполняет

процессор компьютера. С помощью компьютера возможно создание и хранение новой информации любых видов, для чего служат специальные программы, используемые на компьютерах, и устройства ввода 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 объёмов и особые способы работы с различными видами информации. Постоянно совершенствуется программное обеспечение, обеспечивающее коллективную работу с информацией всех видов

Предметы, процессы, явления материального или нематериального свойства, рассматриваемые с точки зрения их информационных свойств, называются информационными объектами.

Что можно делать с информацией:

- создавать
- принимать
- комбинировать
- хранить
- передавать
- копировать
- обрабатывать
- искать
- воспринимать
- формализовать
- делить на части
- измерять
- использовать
- распространять
- упрощать
- разрушать
- запоминать
- преобразовывать
- собирать и т. д.

Все эти процессы, связанные с определенными операциями над информацией, называются информационными процессами.

Представление информации

Познание начинается с восприятия окружающего мира органами чувств. Зрение, вкус, слух, обоняние, осязание доводят до нашего сознания информацию о самых разнообразных свойствах предметов, а также явлениях и процессах, происходящих вокруг нас.

Язык – это знаковая система представления информации. Ряд основных информационных процессов – передача, сохранение – без языка неосуществимы, другие – получение, обработка, использование – его существенно используют в сфере человеческой информационной деятельности. Коммуникативный язык выступает носителем логической информации, именно с помощью языка человек превращает получаемую информацию в знание и передает это знание другому человеку и обществу в целом.

Первобытный человек для обозначения каждого нового предмета придумывал новое имя. Чтобы получить необходимое разнообразие имен, он начал комбинировать звуки друг с другом, получая слова. Так в ходе развития человека появилась идея конечного алфавита, т.е. некоторого фиксированного набора знаков, из которого можно составить как угодно много слов. Комбинацию знаков алфавита называют словом. Из слов составляются фразы, несущие смысловую нагрузку.

Возникновение многих других алфавитов было продиктовано необходимостью привлечения технических средств для передачи информации. Одним из первых примеров такого алфавита является азбука Морзе. В ней каждый знак обычного алфавита кодируется набором точек и тире.

Человек в своей практике общения использует самые разнообразные языки (знаки дорожного движения, пиктограммы олимпийских видов спорта, ноты). Прежде всего, это языки устной и письменной речи. Это иностранные языки, которыми может пользоваться человек.

Кроме того, человек использует ряд языков профессионального назначения. Сюда относятся языки математических и химических формул, обозначений электроники (например, схема электрической цепи). Отметим, что развитие языков программирования идет по пути замены знаков символами.

Воспринимая информацию с помощью органов чувств, человек стремится зафиксировать ее так, чтобы она стала понятной и другим, представляя ее в той или иной форме.

Музыкальную тему композитор может наиграть на пианино, а затем записать с помощью нот. Образы, навеянные все той же мелодией, поэт может воплотить в виде стихотворения, хореограф выразить танцем, а художник — в картине.

Человек выражает свои мысли в виде предложений, составленных из слов. Слова, в свою очередь, состоят из букв. **Это — алфавитное представление информации.**

Форма представления одной и той же информации может быть различной. Это зависит от цели, которую вы перед собой поставили. С подобными операциями вы сталкиваетесь на уроках математики и физики, когда представляете решение в разной форме. Например, решение задачи: «Найти значение математического выражения $y = 5x + 3$, при $x = -3; -2; -1; 0; 1; 2; 3$ » можно представить в табличной или графической форме. Для этого вы пользуетесь визуальными средствами представления информации: числами, таблицей, рисунком.

Таким образом, информацию можно представить в различной форме:

♦ **знаковой письменной, состоящей из различных знаков, среди которых принято выделять:**

- ◆ *символьную в виде текста, чисел, специальных символов (например, текст учебника);*
- ◆ *графическую (например, географическая карта);*
- ◆ *табличную (например, таблица записи хода физического эксперимента);*
- ◇ *в виде жестов или сигналов (например, сигналы регулировщика дорожного движения);*
- ◇ *устной словесной (например, разговор).*

Форма представления информации очень важна при ее передаче: если человек плохо слышит, то передавать ему информацию в звуковой форме нельзя; если у собаки слабо развито обоняние, то она не может работать в розыскной службе. В разные времена люди передавали информацию в различной форме с помощью: речи, дыма, барабанного боя, звона колоколов, письма, телеграфа, радио, телефона, факса. Независимо от формы представления и способа передачи информации, она всегда передается с помощью какого-либо языка.

На уроках математики вы используете специальный язык, в основе которого — цифры, знаки арифметических действий и отношений. Они составляют алфавит языка математики.

На уроках физики при рассмотрении какого-либо физического явления вы используете характерные для данного языка специальные символы, из которых составляете формулы. Формула — это слово на языке физики.

На уроках химии вы также используете определенные символы, знаки, объединяя их в «слова» данного языка.

Существует язык глухонемых, где символы языка — определенные знаки, выражаемые мимикой лица и движениями рук.

Основ любого языка составляет алфавит — конечный набор знаков (символов) любой природы, из которых формируется сообщение.

Языки делятся на естественные (разговорные) и формальные. Алфавит естественных языков зависит от национальных традиций. Формальные языки встречаются в специальных областях человеческой деятельности (математике, физике, химии и т. д.).

Хранилища информации

Понятие «хранилища данных» (ХД) (data warehouse) появилось сравнительно недавно. Потребность в этом термине возникает, в основном, тогда, когда требуется интеграция тематической, текстовой информации и данных, необходимых для принятия решений. Изначально такие совокупности получили название «хранилища информации» (information warehouse), что представляется более верным, поскольку текстовые документы (книги, статьи, диссертации, документы вторичной информации — реферативная, обзорная, экспресс-информация и т.п.) нельзя считать данными, и в органах, занимающихся регистрацией, хранением и обработкой информации, приняты термины «информационные массивы», а места хранения так и называют — «хранилища». Однако позднее, в 1990-е гг. Б. Инмоном¹ в ряде опубликованных им работ был использован термин «хранилища данных». Нередко этот термин используют для всех видов информации.

Обычные БД предназначены для того, чтобы помочь пользователям выполнять повседневную работу, в то время как ХД предназначены для принятия решений. В отличие от базы данных хранилище является основой для извлечения значимой информации из оперативной базы данных.

В основе концепции ХД лежат следующие положения [30]:

- интеграция и согласование данных из различных источников — обычные системы оперативной обработки базы данных, учетные системы, офисные документы, электронные архивы, внутренние и внешние документы;
- разделение наборов данных на данные, используемые для обработки транзакций^[2], т.е. для обработки повседневной, текущей информации, поступающей из цехов, складов, банков, от поставщиков и т.д., обеспечиваемой системами OLTP, и данных для систем поддержки принятия решений (СПИР), т.е. для анализа динамики продаж за несколько лет, позволяющего планирование работ с поставщиками и т.п.

Принципы организации хранилища определяются в разных источниках по-разному. По всей видимости, это связано с тем, что они определялись на основе опыта создания хранилищ для информационных систем разного вида и назначения.

По определению, сформулированному автором термина Б. Ипмоном, хранилище данных — это предметно-ориентированное, привязанное ко времени и неизменяемое собрание данных для поддержки процесса принятия управляющих решений. С учетом этого, отличительными особенностями ХД по сравнению с обычными БД считаются:

- проблемно-предметная ориентация — данные объединяются в категории и хранятся в соответствии с областями, которые они описывают, а не с приложениями, которые они используют;
- интегрированность — данные объединены так, чтобы они удовлетворяли всем требованиям предприятия в целом, а не единственной функции бизнеса;
- некорректируемость — данные в ХД не создаются, т.е. поступают из внешних источников, не корректируются и не удаляются;
- зависимость от времени — данные в хранилище точны и корректны только в том случае, когда они привязаны к некоторому промежутку или моменту времени.

Для обновления данных предусмотрено два варианта:

- 1) полное обновление данных в хранилище. Сначала старые данные удаляются, потом происходит загрузка новых данных. Процесс происходит с определенной периодичностью;
- 2) инкрементальное обновление — обновляются только те данные, которые изменились в OLTP-системе.

Хранилище относительно стабильно: данные в нем обычно обновляются согласно расписанию (например, еженедельно, ежедневно или ежечасно в зависимости от потребностей). В идеале процесс пополнения представляет собой просто добавление новых данных за определенный период времени без изменения прежней информации, уже находящейся в хранилище.

Обычные БД являются источником данных, попадающих в хранилище. Кроме того, хранилище может пополняться за счет внешних источников, например статистических отчетов.

Основными требованиями к хранилищам данных являются:

- поддержка высокой скорости получения данных из хранилища;
- поддержка внутренней непротиворечивости данных;
- возможность получения и сравнения так называемых срезов данных;
- наличие удобных утилит просмотра данных в хранилище;
- полнота и достоверность хранимых данных;
- поддержка качественного процесса пополнения данных.

Способы кодирования и носители информации

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

Кодирование — процесс представления информации в виде кода.

Большие объемы информации требуют ее постоянной систематизации и кодировки для корректной и точной передачи любого количества и качества полученных данных, необходимых для работы, упрощая их восприятие для одного человека или группы людей. Информационное насыщение и свободная обработка подготовленных блоком требует тщательной кодировки данных, для их сохранения и передачи, поэтому современной наукой разработано несколько эффективных видов и способов кодировки данных.

Широко используемые виды кодировки информации.

- двоичная система исчисления. Используется при эффективном конструировании электронных систем, существуют усовершенствованные версии, представляющие собой восьмеричные и шестнадцатеричные системы исчисления, что упрощает работу с объемными массивами предоставленных числовых данных;
- кодирование текстовых данных. Для применения этого вида кодирования в необходимом и выбранном порядке производится перенумеровывание используемые в ходе письма буквы, цифры и знаки препинания, согласно заданной таблице, при этом появляется возможность обеспечить эффективное хранение информации;
- кодировка изображений. Картинки разбиваются линиями, прочерченными горизонтально вертикально, каждый отдельный двумерный массив, или пиксель как элемент общего изображения кодируется по цвету выделенным числом, благодаря чему создается возможность обрабатывать непрерывные величины.

Условные знаки, или коды, используемые в ходе кодирования в обязательном порядке проходят систематизацию и помогают представлять любые объемы информации для быстрой и эффективной их обработки в ходе дальнейшего применения необходимых данных.

Существуют разные варианты классификации информации по видам:

- по сфере применения информации (экономическая, географическая, социологическая и пр.);
- по характеру источников информации (первичная, вторичная, обобщающая и пр.);
- по характеру носителя информации (информация, «зашифрованная» в молекулах ДНК или в длинах световых волн; информация на бумажном или магнитном носителе и пр.).

НОСИТЕЛИ ИНФОРМАЦИИ

Распространение информации всегда связано с материальным носителем, то есть средой для её записи, хранения, передачи.

Носителем информации может быть:

- любой материальный предмет (бумага, камень, дерево, стол, классная доска, звездная пыль, мусор на полу);
- волны различной природы: акустическая (звук), электромагнитная (свет, радиоволна), гравитационная (давление, притяжение);
- вещество в различных состояниях (жидкий раствор — информацией может быть, в частности, концентрация молекул; газ — информацией в этом случае могут являться его температура и давление).

Для хранения информации, которая будет обрабатываться с помощью компьютера, используются специальные машинные носители информации: электронные схемы, перфокарты (использовались раньше), магнитные ленты, магнитные диски, оптические компакт-диски, флешки и т.п.

Понятие модели и алгоритма в информатике

Известно, что модель - это некоторое упрощенное подобие реального объекта.

Модель - это объект-замениТЕЛЬ, который в определенных условиях может заменять объект-оригинал. Модель воспроизводит интересующие нас свойства и характеристики оригинала.

Модели бывают материальными и информационными. Примером материальной модели является глобус - модель Земли. Предметом изучения информатики являются информационные модели.

В информационной модели отображаются знания человека об объекте моделирования. Информационное моделирование - описание в той или иной форме объекта моделирования.

Информационная модель — модель объекта, представленная в виде информации, описывающей существенные для данного рассмотрения параметры и переменные величины объекта, связи между ними, входы и выходы объекта и позволяющая путём подачи на модель информации об изменениях входных величин моделировать возможные состояния объекта[1].

Информационная модель (в широком, общенаучном смысле) — совокупность информации, характеризующая существенные свойства и состояния объекта, процесса, явления, а также взаимосвязь с внешним миром.

Информационные модели делятся на описательные и формальные.

- *Описательные информационные модели — это модели, созданные на естественном языке (например, русском) в устной или письменной форме.*
- *Формальные информационные модели — это модели, созданные на формальном языке (то есть научном, профессиональном или специализированном).*

Информационная модель в информатике — это представление понятий, связей, ограничений, правил и операций, предназначенное для определения семантики данных для конкретной проблемной области.

Алгоритмом называется точное и понятное предписание исполнителю совершить последовательность действий, направленных на решение поставленной задачи. Слово «алгоритм» происходит от имени математика Аль Хорезми, который сформулировал правила выполнения арифметических действий. Первоначально под алгоритмом понимали только правила выполнения четырех арифметических действий над числами. В дальнейшем это понятие стали использовать вообще для обозначения последовательности действий, приводящих к решению любой поставленной задачи. Говоря об алгоритме вычислительного процесса, необходимо понимать, что объектами, к которым применялся алгоритм, являются данные. Алгоритм решения вычислительной задачи представляет собой совокупность правил преобразования исходных данных в результатные.

Основными свойствами алгоритма являются:

1. *детерминированность (определенность). Предполагает получение однозначного результата вычислительного процесса при заданных исходных данных. Благодаря этому свойству процесс выполнения алгоритма носит механический характер;*
2. *результативность. Указывает на наличие таких исходных данных, для которых реализуемый по заданному алгоритму вычислительный процесс должен через конечное число шагов остановиться и выдать искомый результат;*
3. *массовость. Это свойство предполагает, что алгоритм должен быть пригоден для решения всех задач данного типа;*
4. *дискретность. Означает расчлененность определяемого алгоритмом вычислительного процесса на отдельные этапы, возможность выполнения которых исполнителем (компьютером) не вызывает сомнений.*

Алгоритм должен быть формализован по некоторым правилам посредством конкретных изобразительных средств. К ним относятся следующие способы записи алгоритмов: словесный, формульно-словесный, графический, язык операторных схем, алгоритмический язык.

Наибольшее распространение благодаря своей наглядности получил графический (блок-схемный) способ записи алгоритмов.

Блок-схемой называется графическое изображение логической структуры алгоритма, в котором каждый этап процесса обработки информации представляется в виде геометрических символов (блоков), имеющих определенную конфигурацию в зависимости от характера выполняемых операций. Перечень символов, их наименование, отображаемые ими функции, форма и размеры определяются ГОСТами.

При всем многообразии алгоритмов решения задач в них можно выделить три основных вида вычислительных процессов:

- линейный;
- ветвящийся;
- циклический.

Линейным называется такой вычислительный процесс, при котором все этапы решения задачи выполняются в естественном порядке следования записи этих этапов.

Ветвящимся называется такой вычислительный процесс, в котором выбор направления обработки информации зависит от исходных или промежуточных данных (от результатов проверки выполнения какого-либо логического условия).

Циклом называется многократно повторяемый участок вычислений. Вычислительный процесс, содержащий один или несколько циклов, называется циклическим. По количеству выполнения циклы делятся на циклы с определенным (заранее заданным) числом повторений и циклы с неопределенным числом повторений. Количество повторений последних зависит от соблюдения некоторого условия, задающего необходимость выполнения цикла. При этом условие может проверяться в начале цикла — тогда речь идет о цикле с предусловием, или в конце — тогда это цикл с постусловием.