

§2. Понятие информации

Термин «**информация**» происходит от латинского слова «informatio», что означает **сведения, разъяснения, изложение**. Несмотря на широкое распространение этого термина, понятие информации является одним из самых дискуссионных в науке. В настоящее время наука пытается найти общие свойства и закономерности, присущие многогранному понятию информация, но пока это понятие во многом остается интуитивным и получает различные смысловые наполнения в различных отраслях человеческой деятельности:

- **в обиходе** информацией называют любые данные или сведения, которые кого-либо интересуют. Например, сообщение о каких-либо событиях, о чьей-либо деятельности и т.п. «Информировать» в этом смысле означает «сообщить нечто, неизвестное раньше»;
- **в технике** под информацией понимают сообщения, передаваемые в форме знаков или сигналов;
- **в кибернетике** под информацией понимает ту часть знаний, которая используется для ориентирования, активного действия, управления, т.е. в целях сохранения, совершенствования, развития системы.

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

Информатика рассматривает информацию как связанные между собой сведения, изменяющие наши представления о явлении или объекте окружающего мира. С этой точки зрения информацию можно рассматривать как совокупность знаний о фактических данных и зависимостях между ними.

Люди обмениваются информацией в форме сообщений. **Сообщение** – это форма представления информации в виде речи, текстов, жестов, взглядов, изображений, цифровых данных, графиков, таблиц и т.п.

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 – в зависимости от их предшествующих 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 может быть высокоинформативным для человека, владеющего японским. Никакой новой информации не содержит и сообщение, изложенное на знакомом языке, 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.

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.

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

В процессе обработки информация может менять структуру и форму. Признаком структуры являются элементы информации и их взаимосвязь. Формы представления информации могут быть различны. Основными из них являются:

- символная (основана на использовании различных символов);
- текстовая (текст – это символы, расположенные в определенном порядке);
- графическая (различные виды изображений);
- световых или звуковых сигналов;
- радиоволн;
- электрических и нервных импульсов;
- магнитных записей;
- жестов и мимики;
- запахов и вкусовых ощущений;
- хромосом, посредством которых передаются по наследству признаки и свойства организмов и т.д.

В повседневной практике такие понятия, как информация и данные, часто рассматриваются как синонимы. На самом деле между ними имеются существенные различия.

Данными называется информация, представленная в удобном для обработки виде. Данные могут быть представлены в виде текста, графики, аудио-визуального ряда. Представление данных называется языком информатики, представляющим собой совокупность символов, соглашений и правил, используемых для общения, отображения, передачи информации в электронном виде.

Одной из важнейших характеристик информации является ее адекватность. **Адекватность информации** – это уровень соответствия образа, создаваемого с помощью информации, реальному объекту, процессу, явлению. От степени адекватности информации зависит правильность принятия решения. Адекватность информации может выражаться в трех формах: синтаксической, семантической и прагматической.

Синтаксическая адекватность отображает формально-структурные характеристики информации, не затрагивая ее смыслового содержания. На синтаксическом уровне учитываются тип носителя и способ представления информации, скорость ее передачи и обработки, размеры кодов представления информации, надежность и точность преобразования этих кодов и т.д. Информацию, рассматриваемую с таких позиций, обычно называют **данными**.

Семантическая адекватность определяет степень соответствия образа объекта самому объекту. Здесь учитывается смысловое содержание информации. На этом уровне анализируются сведения, отражаемые информацией, рассматриваются смысловые связи. Таким образом, семантическая адекватность проявляется при наличии единства информации и пользователя. Эта форма служит для формирования понятий и представлений, выявления смысла, содержания информации и ее обобщения.

Прагматическая адекватность отражает соответствие информации цели управления, реализуемой на ее основе. Прагматические свойства информации проявляются при наличии единства информации, пользователя и цели управления. На этом уровне анализируются потребительские свойства информации, связанные с практическим использованием информации, с соответствием ее целевой функции деятельности системы.

Каждой форме адекватности соответствует своя **мера** количества информации.

Синтаксическая мера информации оперирует с обезличенной информацией, не выражающей смыслового отношения к объекту. На этом уровне объем данных в сообщении измеряется количеством символов в этом сообщении. В качестве минимальной единицы измерения данных информации Клод Шеннон предложил принять один бит (англ. **bit** – **binary digit** – двоичная цифра). **Бит** в теории информации – количество информации, необходимое для различения двух равновероятных сообщений (типа: «чет» или «нечет», «орел» или «решка», «да» или «нет» и т.п.).

В вычислительной технике битом называют наименьшую «порцию» памяти компьютера, необходимую для хранения одного из двух знаков «0» и «1», используемых для внутримашинного представления данных и команд.

Бит – слишком мелкая единица измерения. На практике чаще применяется более крупная единица – **байт**, равная восьми битам. Именно восемь битов требуется для того, чтобы закодировать любой из 256 символов алфавита клавиатуры компьютера $256 = 2^8$.

Широко используются также еще более крупные производные единицы информации:

- 1 Килобайт (Кбайт) = 1024 байт,
- 1 Мегабайт (Мбайт) = 1024 Кбайт,
- 1 Гигабайт (Гбайт) = 1024 Мбайт.

В последнее время в связи с увеличением объемов обрабатываемой информации входят в употребление такие производные единицы, как:

- 1 Терабайт (Тбайт) = 1024 Гбайт,
- 1 Петабайт (Пбайт) = 1024 Тбайт.

За единицу информации можно было бы выбрать количество информации, необходимое для различения, например, десяти равновероятных сообщений. Это будет не двоичная (бит), а десятичная (дит) единица информации.

Семантическая мера информации используется для измерения смыслового содержания информации. Наибольшее распространение здесь получила тезаурусная мера, связывающая семантические свойства информации со способностью пользователя принимать поступившее сообщение. **Тезаурус** – это совокупность сведений, которыми располагает пользователь или система. Максимальное количество семантической информации потребитель получает при согласовании ее смыслового содержания со своим тезаурусом, когда поступившая информация понятна пользователю и несет ему ранее не известные сведения. С семантической мерой количества информации связан коэффициент содержательности, определяемый как отношение количества семантической информации к общему объему данных.

Прагматическая мера информации определяет ее полезность, ценность для процесса управления. Обычно ценность информации измеряется в тех же единицах, что и целевая функция управления системой.

Какое количество информации содержится, к примеру, в тексте романа «Война и мир», во фресках Рафаэля или в генетическом коде человека? Ответа на эти вопросы наука не дает и, по всей вероятности, даст не скоро. А возможно ли объективно измерить количество информации? Важнейшим результатом в теории информации является следующий вывод:

В определенных, весьма широких условиях можно пренебречь качественными особенностями информации, выразить ее количество числом, а также сравнить количество информации, содержащейся в различных группах данных.

В настоящее время получили распространение подходы к определению понятия «количество информации», основанные на том, что информацию, содержащуюся в сообщении, можно нестрого трактовать в смысле ее новизны или, иначе, уменьшения неопределенности наших знаний об объекте. Эти подходы используют математические понятия вероятности и логарифма.

Информация передается в форме сообщений от некоторого **источника** информации к ее **приемнику** посредством **канала связи** между ними. Источник посылает передаваемое сообщение, которое кодируется в передаваемый сигнал. Этот сигнал посылается по каналу связи. В результате в приемнике появляется принимаемый сигнал, который декодируется и становится принимаемым сообщением. Передача информации по каналам связи часто сопровождается воздействием **помех**, вызывающих **искажение** и **потерю информации**.