

МОДУЛЬ 3

3. ОСНОВНЫЕ ПОДХОДЫ И МЕТОДЫ ОПИСАНИЯ ИНФОРМАЦИОННЫХ ПРОЦЕССОВ И ЯВЛЕНИЙ

3.1. Информация и данные

Понятие «**информация**» является одним из наиболее сложных и многогранных понятий. Уяснить сущность этого понятия, которому посвящено огромное количество литературных источников, в отрыве от таких понятий как система, цель и технология не представляется возможным. Но в этом то и состоит сложность феномена информации, восходящего к вышеназванным понятиям, в сложности которых мы убедились. Иная точка зрения (т.е., что информация первична) представляется, мягко говоря, неконструктивной уже хотя бы потому, что информация должна иметь физический носитель и необходимый интерфейс, но это опять же есть система.

Информационная технология реальной системы представляет собой вещественно-энергетические информационные процессы, реализующие сбор, хранение, преобразование, обработку, передачу и использование информации для достижения определенной цели. Поэтому ценность информации и другие ее свойства преломляются (оцениваются) именно через цель системы, в этом то и сложность, поскольку надо иметь описание (модель) системы.

Ввиду разнообразия трактовок понятия «информация», равно как и других родственных понятий, далее, по возможности, будем придерживаться дефиниций, приведенных в стандартах.

Информация (для процесса обработки данных) — содержание, присваиваемое данным, посредством соглашений, распространяющихся на эти данные (ИСО 2389/1-84).

Данные — событие, понятие или команда, представленные в формализованном виде, позволяющем передачу, интерпретацию или обработку, как вручную, так и с помощью средств автоматизации (ИСО 2382/1-84).

Информацию классифицируют по многочисленным признакам. По физической природе восприятия информацию подразделяют на зрительную, слуховую, тактильную; по метрическим свойствам — на параметрическую, топологическую и абстрактную. Параметрическая информация может быть нульмерной (событие), одномерной (величина), двумерной (функция), трехмерной (комплекс), n-мерной (n-мерное пространство). Топологическую информацию составляют геометрические образы, карты местности, различные изображения и объемные объекты. К абстрактной информации относятся обобщенные образы и понятия, математические соотношения.

Наиболее сложными и разнообразными видами информации являются биологическая (обеспечивает жизнедеятельность отдельно взятого живого организма) и социальная информация (связана с практической деятельностью человека).

Социальная информация делится на массовую (для всех членов общества) и специальную (для определенных социальных групп). Специальная социальная информация подразделяется на политическую, экономическую, технологическую, научно-техническую и т.д.

Социальная информация характеризуется тремя основными аспектами: синтаксическим, семантическим и прагматическим.

Синтаксический аспект характеризуют количественные и структурные свойства информации. Эти свойства связаны с формированием носителей информации и организацией информационных потоков в ИТ-системах.

Семантический аспект характеризуют смысловое содержание информации и отношением между ее элементами. Семантические свойства информации связаны с организацией процесса обработки информации.

Прагматический аспект связан с ценностью информации для ее получателя. Прагматические свойства информации проявляются на стадии принятия решений (при разработке и эксплуатации ИТ-систем).

3.2. Сигналы и знаки

Под сигналом в широком смысле слова будем понимать материальный носитель информации. В этом смысле сигнал – это некоторый физический объект или процесс, содержащий информацию. Сигналы в широком смысле можно подразделить на сигналы в узком смысле и знаки.

Вот определение сигнала в узком смысле. Сигнал – изменение физической величины, используемой для передачи данных (ИСО 2382/1-84). Здесь по умолчанию под передачей данных понимается перенос информации в пространстве. Но под передачей данных можно понимать и перенос данных во времени, т.е. хранение информации. Такое расширение обсуждаемого понятия вполне оправдано, поскольку и при передаче и при хранении информации по существу используются одни и те же алгоритмы сжатия данных и помехоустойчивого кодирования. Забегая вперед, заметим, что разработка этих алгоритмов была инициирована основными теоремами Шеннона для дискретных каналов связи.

Сигналы классифицируют по многочисленным признакам. По происхождению сигналы делятся на естественные, которые не

связанны с деятельностью человека и искусственные, являющиеся результатом прямой или косвенной деятельности человека.

При математическом моделировании сигнал, независимо от его физической природы, заменяется математической моделью в виде функции одного или нескольких аргументов.

По пространственно-временным свойствам математические модели сигналов подразделяются на:

- непрерывные (непрерывная функция непрерывного аргумента);
- непрерывно-дискретные (непрерывная функция дискретного аргумента);
- дискретно-непрерывные (дискретная функция непрерывного аргумента);
- дискретные (дискретная функция дискретного аргумента).

По степени знания исходных (априорных) данных о сигнале сигналы делятся на:

- детерминированные (регулярные), которые подразделяются на периодические и непериодические;
- случайные (стохастические), подразделяющиеся на стационарные, нестационарные, с априорной неопределенностью (параметрической и непараметрической неопределенностью).

Изучением сигналов в явном виде занимается теория сигналов, которая рассматривает методы построения математических моделей сигналов, а также изучает свойства и характеристики сигналов.

Знаки делятся на символы и диакритики. Форма символа отражает его значение. Например, стрелка как указатель направления. Диакритики не имеют прямой связи между формой и значением или такая связь утеряна в результате переноса его значения на другие сущности или в результате изменения формы знаков. Диакритиками являются знаки математических операции, буквы русского языка и т.д.

Различают знаковые системы (естественные и искусственные языки) и внесистемные знаки (остатки некогда существовавших систем; знаки, созданные временно в небольших коллективах людей; междометия, восклицания, жесты). Диакритики обычно входят в состав знаковых систем, а знаки-одиночки, как правило, являются символами. Имеются и исключения. Например, система знаков дорожного движения.

Изучением знаков и знаковых систем во всех их проявлениях занимается **семиотика** (основатель семиотики Ч.Пирс), которая имеет следующие направления:

Биосемиотика, посвящена изучению знаковой коммуникации в животном мире (танцы пчел, поведенческие акты животных).

Этносемиотика изучает незнаковые знаки в человеческом обществе (жесты, позы, ритуалы, обряды).

Лингвосемиотика изучает человеческие языки, рассматриваемые как знаковые системы.

Семиотика делится на три основных области: синтактику семантику и прагматику.

Синтактика знаковых систем занимается изучением их структуры и правил соединений отдельных знаков.

Семантика рассматривает отношения между знаками и тем, что они обозначают (замещают). Для уяснения этих отношений обратимся к семантическому треугольнику (треугольнику Фреге), см. рис. 3.1

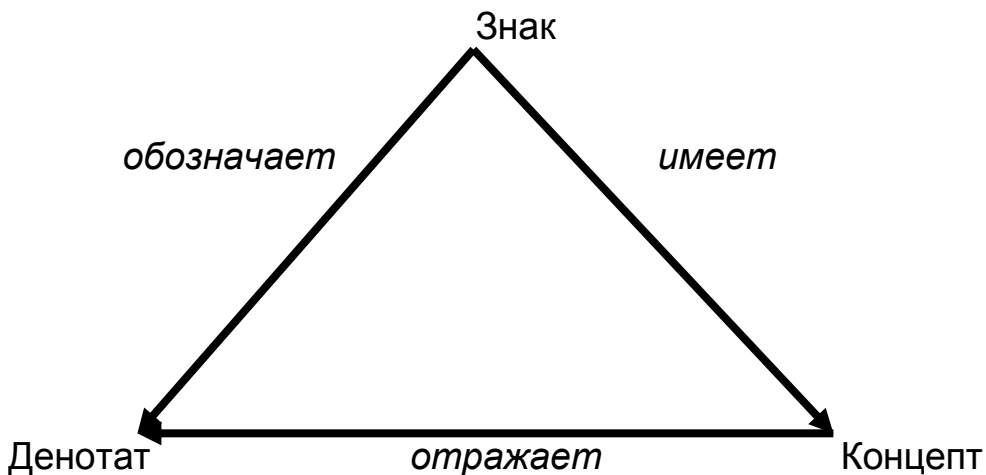


Рис. 3.1. Треугольник Фреге.

Денотатом принято называть то, заместителем чего является знак, а представление, которое вызвано этим знаком, называют концептом или смыслом. Существуют знаки, которые имеют концепт, но не имеют денотата, например флогистон, баба Яга.

Прагматика связана с изучением полезности, ценности знаков для пользователя.

3.3. Информационные подходы, меры и теории

Рассмотрим основные информационные подходы к описанию информационных явлений и процессов, развитие которых было инициировано возникновением теории информации, связанной с именем Клода Шеннона. Эту теорию иногда называют шенноновской, классической или статистической теорией информации. Она по существу сыграло роль катализатора, стимулировавшего исследования различных аспектов феномена информации, которые образовали три, возрастающих по сложности формализации, направления: синтаксическое, семантическое и прагматическое. Сущность этих направлений (аспектов) рассмотрена выше. Ниже приведены основные направления и меры информации

Направление	Мера информации
Синтаксическое	Хартли, Шеннона, Колмогорова, Кульбака, Реньи
Семантическое	Карнапа и Бар-Хиллела, Шрейдера
Прагматическое	Харкевича, Бонгарда, Стратоновича.

Р.Хартли в 1928 г. заложил первый кирпич в «здание» теории информации, введя логарифмическую меру количества информации для равновероятных событий. Основу здания теории информации построил в 1948 г. Клод Шеннон. В создании теории информации, которая сыграла выдающуюся роль в развитии теории и практики информационных систем значительный вклад сделали и русские ученые А.Н. Колмогоров, А.Я. Хинчин и др.. Основные положения теории информации будут рассмотрены в следующем разделе этого курса.

Информационные меры С. Кульбака и А. Реньи, связаны с мерой Шеннона и в основном используются в математической статистике.

Мера А.Н. Колмогорова (колмогоровская сложность), величайшего русского математика (создателя современной теории вероятностей), предназначена для оценки сложности объекта. Сложность некоторого объекта оценивается как минимальная длина программы, которая необходима для повторного воспроизведения этого объекта.

Семантическая мера информации Р. Карнапа и У. Бар-Хиллела является одной из первых попыток оценки смыслового содержания информации. Сущность этой меры сводится к вычислению логарифма вероятности истинности оцениваемого предложения.

Более широкое признание получил подход А. Шрейдера, в котором мера количества семантической информации оценивается как степень изменения тезауруса получателя при приеме некоторого сообщения. Тезаурус — это словарь, в котором указаны семантические отношения между лексическими единицами.

Одним из первых меру ценности информации, как логарифм отношения вероятности достижения цели после получения информации к вероятности достижения цели до получения информации, ввел А. Харкевич. Связь между полезностью информации и мерой шенновского количества установил М. Бонгард. Р. Стратонович установил взаимосвязь между шенновской теорией информации и теорией статистических решений. Ценность информации по Стратоновичу определяется как максимальная польза, которую данное количество информации может принести в деле уменьшения среднего риска.