

## Введение

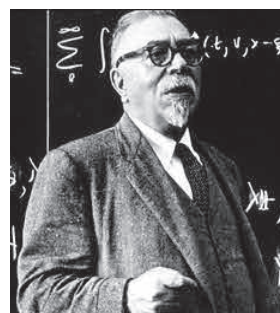
Выделение информатики как самостоятельной области человеческой деятельности связано в первую очередь с развитием компьютерной техники. Слово «*информатика*» (франц. *informatique*) появилось в начале 1960-х годов как гибрид двух французских слов – «*information*» (информация) и «*automatique*» (автоматика) и дословно означает «*информационная автоматика*». Во Франции и других странах Европы данный термин применяется для обозначения области научных знаний, связанных с автоматизацией обработки информации с помощью ЭВМ. Фактическое его использование немного шире, потому как до изобретения электронной вакуумной лампы и транзистора также существовали разнообразные вычислительные машины, скорее механические и электромеханические, нежели электронные.

Возможно, чтобы скрасить сей недочёт, а может и по другим причинам, в 70-е годы в англоязычной литературе наука о переработке информации на основе вычислительной техники утвердилась под названием «*Computer Science*», что буквально означает «*компьютерная наука*».

В нашей стране в 60-е годы XX века вопросы, связанные с разработкой, функционированием и применением автоматизированных систем обработки информации, объединялись термином «*кибернетика*<sup>1</sup>», хотя это было не вполне верно, поскольку, по определению Н. Винера<sup>2</sup>, кибернетика – это наука о законах управления в живой и неживой природе [1], то есть сфера её интересов охватывает лишь часть (хотя обширную и важную) вопросов, связанных с информацией и информационными процессами. К началу 80-х годов термин «*кибернетика*» постепенно ушёл из широкого распространения.

Так что же такое *информатика*?

Академик В. М. Глушков<sup>3</sup> в письме президенту Академии наук СССР по поводу создания нового отделения Академии ссылается на определение, данное Международным конгрессом по информатике, проходившим в Японии в 1978 году: «*Понятие информатики охватывает области, связанные с разработкой, созданием, использованием и материально-техническим обслуживанием систем обработки информации, включая машины, оборудование, математическое обеспечение, организационные аспекты, а также комплекс промышленного, коммерческого, административного, социального и политического воздействия*». Это очень широкое определение, его нужно кон-



Норберт Винер



Виктор Михайлович  
Глушков

<sup>1</sup> От греч. *kiber* – над, *nautus* – моряк, кормчий, управляющий рулём, предположительно, отсюда у Н. Винера появилось «управление». Впервые термин «кибернетика» встречается в работах древнегреческого философа Платона (ок. 427-347 гг. до н. э.), в которых он обозначил правила управления обществом. Через две с лишним тысячи лет французский физик А. И. Ампер (1775–1836) в своей работе «Опыт философских наук» (1834) термин «кибернетика» также применил к науке об управлении обществом. [8]

<sup>2</sup> Норберт Винер (англ. Norbert Wiener; 1894–1964) – американский учёный, выдающийся математик и философ, основоположник кибернетики и теории искусственного интеллекта.

кретизировать, но в нём выделено главное – понятие о вычислительной технике и автоматизации, которые не существуют отдельно друг от друга [3]. После, в 1983 году на сессии годовичного собрания Академии наук СССР, и было организовано новое «Отделение информатики, вычислительной техники и автоматизации». Информатика стала рассматриваться как *«комплексная научная и инженерная дисциплина, изучающая все аспекты разработки, проектирования, создания, оценки, функционирования основанных на ЭВМ систем переработки информации, их применения и воздействия на различные области социальной практики»*. Были созданы институты данного профиля, информатика стала обязательным предметом в системе среднего и высшего образования.

Мы предложим читателям следующее определение, взятое из [4]:

**Информатика** – наука, изучающая общие свойства информации, закономерности и способы её создания, хранения, поиска, преобразования и использования с помощью компьютерных систем.

На сайте Министерства образования и науки Российской Федерации приведена *«Примерная программа по информатике и информационным технологиям»*<sup>4</sup> для среднего (полного) общего образования и *«Примерная программа дисциплины информатика»*<sup>5</sup>, составленная в соответствии с государственными образовательными стандартами высшего профессионального образования для направлений *«Технические науки»*, *«Образование»* и прочих. Обе программы содержат разделы:

- общие сведения об информации,
- аппаратное и программное обеспечение информационных систем,
- технологии обработки текстовой, числовой, графической и мультимедийной информации,
- работа с базами данных,
- алгоритмизация и программирование,
- локальные и глобальные компьютерные сети.

Сравнение двух программ показывает, что школьный и вузовский курсы информатики достаточно близки по содержанию. Следовательно, задача вуза – углубить знания по информатике, полученные в школе, познакомить студентов со всеми новостями в области технического и программного обеспечения информационных процессов, а также с использованием информационных технологий в их будущей профессиональной деятельности. При выполнении практических работ студент должен совершенствовать свои навыки в работе на компьютере в процессах обработки, поиска и хранения всех видов информации.

Смеем предположить, что компьютер как для грамотного специалиста, так и для обычного человека должен стать привычным инструментом автоматизации его интел-

<sup>3</sup> Виктор Михайлович Глушкóв (1923–1982) – советский математик, кибернетик. Решил обобщённую пятую проблему Гильберта. Был инициатором и главным идеологом разработки и создания ОГАС, предназначенной для автоматизированного управления всей экономикой СССР. Для этого им была разработана система алгоритмических алгебр и теория для управления распределёнными базами данных. Также под его руководством в 1966 году была разработана первая персональная ЭВМ «МИР-1» (машина для инженерных расчётов).

<sup>4</sup> [http://window.edu.ru/window\\_catalog/files/r37206/09-1-s.pdf](http://window.edu.ru/window_catalog/files/r37206/09-1-s.pdf).

<sup>5</sup> <http://www.edu.ru/db/portal/spe/prog/htm/mf0201.htm>.

лектуального труда. На удивление, выделенные В. А. Острейковским<sup>6</sup> более десятилетия назад основные признаки информационного общества [4] находят всё большее отражение в нашей повседневной жизни, и вы без труда узнаете их сами:

1. Большинство работающих в информационном обществе (около 80%) занято в информационной сфере, то есть сфере производства информации и информационных услуг.

2. Обеспечены техническая, технологическая и правовая возможности доступа любому члену общества практически в любой точке территории и в приемлемое время к нужной ему информации (за исключением военных и государственных секретов, точно оговорённых в соответствующих законодательных актах).

3. Информация становится важнейшим стратегическим ресурсом общества и занимает ключевое место в экономике, образовании и культуре.

**Информатизация** – необходимое условие научно-технического, социального, экономического и политического прогресса в обществе. Неизбежность информатизации обусловлена следующими причинами:

- беспрецедентным усложнением социально-экономических процессов в результате увеличения масштабов и темпов общественного производства, углубления разделения труда и его специализации в научно-технической революции;
- необходимостью адекватно реагировать на возникающие проблемы в динамично меняющейся обстановке, присущей постоянно развивающемуся обществу;
- повышением степени самоуправления предприятий, территорий, регионов.

Конечно, процесс перехода от индустриального общества к информационному происходит не только не одновременно в различных странах, но и внутри России имеет разную скорость и темпы развития.

Полагаем, что читатели и без авторов заметили, что в нашей жизни буквально за несколько лет появились такие, вещи как «электронные билеты», когда для посадки на самолёт или поезд нужен лишь паспорт. Различные «электронные приёмные» от уровня Президента страны до глав управ и районов. Всё больше появляется терминалов для оплаты пластиковыми картами (особенно в мегаполисах), повсеместно появляются терминалы для приёма денег. Сотовая связь и её сервисы покрывают большинство мест проживания граждан. На рекламах вместо телефонов и адресов мы видим непонятые квадратики с точками (QR-коды). Электронные книги вытесняют бумажные издания.

Надеемся, что наш бумажный учебник окажется вам полезным.

<sup>6</sup> Острейковский Владислав Алексеевич (1932 г. р.) – д. т. н., заслуженный деятель науки и техники РФ, академик Международной академии информатизации. В 1993 году получил медаль Минатома РФ «За заслуги в повышении безопасности атомных станций». В 1995 году по представлению Стокгольмского королевского университета был удостоен диплома Кембриджского университета (Великобритания) «Человек года в области кибернетики». С 1985 года является членом Ядерного общества, редколлегии журнала «Ядерная энергетика», Президиума объединенного учебно-методического совета Министерства образования РФ.

# Глава 1. Общие сведения об информационных процессах

## 1.1. Понятие информации, её виды и свойства

Вся жизнь человека так или иначе связана с накоплением и обработкой информации, которую он получает из окружающего мира, используя пять органов чувств – зрение, слух, вкус, обоняние и осязание. Как научная категория «*информация*» составляет предмет изучения для самых различных дисциплин: *информатики, теории связи, теории информации, кибернетики, философии, физики, биологии, семиотики* (науки о знаках и знаковых системах), *теории массовой коммуникации* (исследование средств массовой информации (СМИ) и их влияния на общество), *соционики* (теории информационного метаболизма индивидуальной и социальной психики), *информодинамики* (науки об открытых информационных системах), *информациологии* (науки о получении, сохранении и передаче информации для различных множеств объектов) и т. д. И всё же общепринятого строгого научного определения информации до настоящего времени не существует, а вместо него обычно используют понятие об информации. В разных областях науки и техники в это понятие вкладывают разный смысл, чтобы этот смысл в наибольшей степени соответствовал задачам соответствующей предметной области. Имеется множество определений понятия информации – от наиболее общего философского (информация есть отражение реального мира) до наиболее частного прикладного (информация есть сведения, являющиеся объектом переработки). Вот некоторые из них:

- сообщение, осведомление о чём-либо;
- сведения об окружающем мире, протекающих в нём процессах и т. д., которые воспринимают живые организмы, управляющие машины и другие информационные системы в процессе жизнедеятельности и работы;
- сведения о лицах, предметах, фактах, событиях, явлениях и процессах независимо от формы их представления;
- результат отражения реальности в сознании человека, представленный на его внутреннем языке;
- сведения, содержащиеся в сообщении и рассматриваемые как объект передачи, хранения и обработки;
- знания, переданные кем-то другим или приобретенные путём собственного исследования и изучения;
- содержание сигнала, сообщения;
- содержательное описание объекта или явления;
- семантика или прагматика синтаксиса языка представления данных;
- продукт научного познания, средство изучения реальной действительности;
- основное содержание отображения;
- результат выбора;

- мера разнообразия;
- отражённое разнообразие;
- сущность, сохраняющаяся при вычислимом изоморфизме;
- мера сложности структур, мера организации;
- вечная категория, содержащаяся во всех без исключения элементах и системах материального мира, проникающая во все «поры» жизни людей и общества;
- бесконечный законопроцесс триединства энергии, движения и массы с различными плотностями кодовых структур бесконечно-беспредельной Вселенной.

Наряду с названными существуют сотни других, иногда противоречащих друг другу или взаимоисключающих определений информации. По мнению профессора Игоря Н. Бекмана [9], к настоящему времени различных определений этого термина накопилось где-то около пятисот, но список не закрыт, так как дать однозначного и чёткого определения информации до сих пор никому не удалось, и вряд ли когда удастся. По крайней мере, в наше время довольно распространено мнение, что информация наряду с материей и энергией является первичным понятием нашего мира и поэтому в строгом смысле не может быть определена. Многообразие этих определений свидетельствует о широте подхода к понятию информации и свидетельствует о том, что становление концепции информации в современной науке ещё не завершено.

В Федеральном законе Российской Федерации от 27 июля 2006 г. № 149-ФЗ «Об информации, информационных технологиях и о защите информации»<sup>7</sup> даётся следующее определение этого термина:

*«информация – сведения (сообщения, данные) независимо от формы их представления».*

Толковый словарь русского языка<sup>8</sup> Сергея Ивановича Ожегова<sup>9</sup> приводит два определения слова *информация*:

- 1) Сведения об окружающем мире и протекающих в нём процессах, воспринимаемые человеком или специальными устройствами.
- 2) Сообщения, осведомляющие о положении дел, о состоянии чего-нибудь. (*Научно-техническая и газетная информация, СМИ – печать, радио, телевидение, кино, интернет – прим. авторов*).

В **информатике** наиболее часто используется следующее определение этого термина:

**Информация** – это осознанные сведения об окружающем мире, которые являются объектом хранения, преобразования, передачи и использования. *Сведения* – это знания, выраженные в сигналах, сообщениях, известиях, уведомлениях и т. д.

Каждого человека в мире окружает море информации различных видов. Стремление зафиксировать, сохранить надолго своё восприятие информации было всегда свойственно человеку. Мозг человека хранит множество информации и использует для хра-

<sup>7</sup> <http://www.rg.ru/2006/07/29/informacia-dok.html>

<sup>8</sup> Ожегов С. И. Словарь русского языка: ок. 57 000 слов / под ред. чл.-корр. АН СССР Н. Ю. Шведовой. – 20-е изд., стереотип. – М.: Рус. яз., 1988. – 750 с. ISBN 5-200-00313-X.

<sup>9</sup> Сергей Иванович Ожегов (1900–1964) – лингвист, лексикограф, доктор филологических наук, профессор. Ударение в фамилии *Ожегов* ставится на первом слоге на букве «о», так как эта фамилия является производной от слова «ожег», что означало палку, которую в старину окунали в расплавленный металл, чтобы определить степень его готовности к разливке.

нения её свои способы, основа которых – скорее всего, двоичный код [4], как и у компьютеров. Человек всегда стремился иметь возможность поделиться своей информацией с другими людьми и найти надёжные средства для её передачи и долговременного хранения. Для этого в настоящее время изобретено множество способов хранения информации на внешних (относительно мозга человека) носителях и её передачи на огромные расстояния.

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

- **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 других материалах, изображающих картины реального мира;
- **звуковая** – мир вокруг нас полон звуков, и задача их хранения и тиражирования была решена с изобретением звукозаписывающих устройств в 1877 г.<sup>10</sup>; её разновидностью является **музыкальная** информация – для этого вида был изобретён способ кодирования с использованием специальных символов (музыкальных нот), что делает возможным хранение её аналогично графической информации;
- **текстовая** – способ кодирования речи человека специальными символами – буквами, причём разные народы имеют разные языки и используют различные наборы букв для отображения речи; особенно большое значение этот способ приобрёл после изобретения бумаги и книгопечатания;
- **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 её отображения используется метод кодирования специальными символами – цифрами, причём системы кодирования (счисления) могут быть разными;
- **видеоинформация** – способ сохранения движущихся картин окружающего мира, появившийся с изобретением кино.

Существуют также виды информации, для которых до сих пор не изобретено эффективных способов их кодирования и хранения, – это **тактильная** информация, передаваемая ощущениями, **органолептическая**, передаваемая запахами и вкусами, и другие виды, для которых современная наука даже не нашла признанных всеми терминов определения (например, экстрасенсорная информация).

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

Хранение информации при использовании компьютеров осуществляется на магнитных дисках или лентах, на лазерных дисках (**CD, DVD, BD**), специальных устройствах энергонезависимой памяти (флэш-память и прочее). Эти методы постоянно совершенствуются, изобретаются новые устройства и носители информации. Обработку информации (воспроизведение, преобразование, передачу, запись на внешние носители) выполняет процессор компьютера. С помощью компьютера возможно создание и

<sup>10</sup> Рекомендуем ознакомиться с экспозицией «История звукозаписи» в музее радио имени А. С. Попова (г. Екатеринбург, ул. Розы Люксембург/Энгельса, 9/11, филиал Свердловского областного краеведческого музея, +7 (343) 371-50-60, <http://uole-museum.ru>).

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 приёмы хранения, обработки, поиска и передачи распределённой информации больших объёмов и особые способы работы с различными видами информации.

Постоянно совершенствуется программное обеспечение, благодаря которому становится возможным не только коллективно работать с информацией, взяв хотя бы *GoogleDocuments*, но и для многих программ появляется возможность сохранения «в облако» (сервисы *Яндекс.Диск*, *Amazon S3*, *Vox.com*, *Copy.com*, *Dropbox*, *DVCS-Autosync*, *Google Drive*, *iCloud*, *iDrive*, *ownCloud*, *Rackspace Cloud Files*, *Selectel Cloud Storage*, *SkyDrive*, *SparkleShare*, *SugarSync*, *Ubuntu One*, *Windows Azure Blob* и др.). Сместем предположить, что чувства, испытываемые нашими внуками при виде «флэшек», будут сродни тем, что испытывает читатель при виде перфокарт сегодня.

## 1.2. Свойства информации

Как и всякий объект, информация обладает свойствами. Характерной отличительной особенностью информации от других объектов природы и общества является дуализм: на свойства информации влияют как свойства исходных данных, составляющих её содержательную часть, так и свойства методов, фиксирующих эту информацию.

С точки зрения информатики наиболее важными представляются следующие **общие качественные свойства**: объективность, достоверность, полнота, точность, актуальность, полезность, ценность, своевременность, понятность, доступность, краткость и прочие.

**1) Объективность информации.** Объективный – существующий вне и независимо от человеческого сознания. Информация – это отражение внешнего объективного мира. Информация объективна, если она не зависит от методов её фиксации, чьего-либо мнения, суждения.

**Пример.** Сообщение «На улице тепло» несёт субъективную информацию, а сообщение «На улице 22 °С» – объективную, но с точностью, зависящей от погрешности средства измерения.

Объективную информацию можно получить с помощью измерительных приборов. Отражаясь в сознании конкретного человека, информация перестаёт быть объективной, так как преобразовывается (в большей или меньшей степени) в зависимости от мнения, суждения, опыта, знаний конкретного субъекта.

**2) Достоверность информации.** Информация достоверна, если она отражает истинное положение дел. Объективная информация всегда достоверна, но достоверная информация может быть как объективной, так и субъективной. Достоверная информация помогает принять нам правильное решение.

Недостоверной информация может быть по следующим причинам:

- преднамеренное искажение (дезинформация) или непреднамеренное искажение субъективного свойства;
- искажение в результате воздействия помех и недостаточно точных средств измерений.

**3) Полнота информации.** Информацию можно назвать полной, если её достаточно для понимания и принятия решений. Неполная информация может привести к ошибочному выводу или решению.

**4) Точность информации** определяется степенью её близости к реальному состоянию объекта, процесса, явления (погрешностью средства измерения).

**5) Актуальность информации** – важность для настоящего времени, злободневность, насущность. Иногда только вовремя полученная информация может быть полезна.

**6) Полезность (ценность) информации** может быть оценена применительно к нуждам конкретных её потребителей и оценивается по тем задачам, которые можно решить с её помощью.

**7) Относимость информации** показывает возможность её использования в узкой области, применимость для решения какой-либо конкретной задачи.

Многие свойства могут как дополнять друг друга, так и быть синонимами. В различных ситуациях важны разные свойства. Немного подумав, читатели без труда вспомнят и добавят к этому списку ещё несколько свойств информации.

Самая ценная информация – объективная, достоверная, полная и актуальная. При этом следует учитывать, что и необъективная, недостоверная информация (например, художественная литература) имеет большую значимость для человека. Важно чтобы об этом было сообщено заранее: сказка, мультфильм, фантастический рассказ или используются спецэффекты, все герои сюжета выдуманы. В противном случае, это будет дезинформация, обман, фейковая новость, со своими последствиями.

Социальная (общественная) информация обладает ещё и дополнительными свойствами:

- имеет семантический (смысловой) характер, то есть понятийный, так как именно в понятиях обобщаются наиболее существенные признаки предметов, процессов и явлений окружающего мира;
- имеет языковую природу (кроме некоторых видов эстетической информации, например изобразительного искусства). Одно и то же содержание может быть выражено на разных естественных (разговорных) языках, записано в виде математических формул и т. д.

### **Процессы и действия в отношении информации: накопление, старение, копирование, размножение, удаление, уничтожение, блокирование, модификация**

С течением времени количество информации растёт, информация накапливается, происходят её систематизация, оценка и обобщение. Это свойство назвали ростом и кумулированием информации. (Кумуляция – от лат. *cumulatio* – увеличение, скопление).

**Старение информации** заключается в уменьшении её ценности с течением времени. Старит информацию появление новой информации, которая уточняет, дополняет или отвергает полностью или частично более раннюю. Научно-техническая информация стареет быстрее, эстетическая (произведения искусства) – медленнее. (Со своей стороны мы постарались наполнить данный учебник по большей части информацией, менее подверженной старению. К сожалению, сделать учебник полностью таковым не удалось.)



Процессы копирования и размножения информации схожи как в теории, так и с точки зрения обывателя, однако, в юридическом аспекте между терминами искусственно внесена разница. Особенно это заметно в отношении компьютерной информации, где копирование – это повторение и устойчивое запечатление её на машинном или ином носителе. Оно может быть осуществлено путём записи содержащегося во внутренней памяти ЭВМ файла на дискету, диск, флешку, его распечатки и т. д. С другой же стороны, из диспозиции (дословного толкования содержания) ст. 272 УК РФ к копированию компьютерной информации не относится копирование от руки, путём фотографирования текста с экрана дисплея, а также считывание информации путём перехвата излучений ЭВМ, расшифровки шумов принтера и пр. (Также см. ниже ещё одно определение копирования.)

Размножение информации (с точки зрения УК РФ), отличается от копирования информации тем, что информация повторяется не на обособленном от оригинального носителе, а на оригинальном носителе (например, в памяти ЭВМ заводится несколько файлов одного и того же содержания), либо на однородном носителе, оставшемся в распоряжении пользователя (например, копия заводится в памяти ЭВМ, образующей с данным компьютером систему, либо на дискете, сознательно оставленной в компьютере).

Как избавиться от информации? Забывание, удаление и уничтожение информации – это одно и то же или нет? А чем информация отличается от данных? Последние понятия зачастую в теории используются как синонимы, но между ними существуют и принципиальные различия.

Данные – это совокупность сведений, которые зафиксированы на каком-либо носителе – бумаге, диске, флешке, плёнке и т.п. Эти сведения должны быть в форме, пригодной для хранения, передачи и обработки. Дальнейшее преобразование данных позволяет получить информацию. Таким образом, информацией можно назвать результат анализа и преобразования данных.

**Уничтожение информации** (*destruction of information*): любое условие, делающее информацию непригодной для использования независимо от причины.<sup>11</sup>

В юридической практике устоялись следующие определения<sup>12</sup>:

а) **уничтожение информации** – это приведение информации или её части в непригодное для использования состояние независимо от возможности её восстановления. Уничтожением информации не является переименование файла, где она содержится, а также само по себе автоматическое «вытеснение» старых версий файлов последними по времени;

б) **блокирование информации** – результат воздействия на компьютерную информацию или технику, последствием которого является невозможность в течение некоторого времени или постоянно осуществлять требуемые операции над компьютерной информацией полностью или в требуемом режиме, то есть совершение действий, приводящих к ограничению или закрытию доступа к компьютерному оборудованию и находящимся на нём ресурсам, целенаправленное затруднение доступа законных пользователей к компьютерной информации, не связанное с её уничтожением;

<sup>11</sup> «Финансовые услуги. Рекомендации по информационной безопасности. ГОСТ Р ИСО/ТО 13569-2007» (утв. Приказом Ростехрегулирования от 27.12.2007 № 514-ст), <http://docs.cntd.ru/document/1200068821>

<sup>12</sup> «Методические рекомендации по осуществлению прокурорского надзора за исполнением законов при расследовании преступлений в сфере компьютерной информации» (утв. Генпрокуратурой России 14 апреля 2014, [https://genproc.gov.ru/upload/iblock/a77/мет\\_рек\\_ПКИ.doc](https://genproc.gov.ru/upload/iblock/a77/мет_рек_ПКИ.doc))

в) **модификация информации** – внесение изменений в компьютерную информацию (или ее параметры). Законом установлены случаи легальной модификации программ (баз данных) лицами, правомерно владеющими этой информацией, а именно: модификация в виде исправления явных ошибок; модификация в виде внесения изменений в программу, базы данных для их функционирования на технических средствах пользователя; модификация в виде частной декомпиляции программы для достижения способности к взаимодействию с другими программами;

г) **копирование информации** – создание копии имеющейся информации на другом носителе, то есть перенос информации на обособленный носитель при сохранении неизменной первоначальной информации, воспроизведение информации в любой материальной форме – от руки, фотографированием текста с экрана дисплея, а также считывания информации путем любого перехвата информации и т.п.

Все аспекты в отношении информации невозможно рассмотреть коротко, поэтому завершим параграф<sup>13</sup> очевидным: логичность, компактность, удобная форма представления облегчают понимание и усвоение информации, не зря говорят: «краткость – сестра таланта».

## 1.3. Характеристики информации

Рассмотрим вопросы связанные с количественными и качественными характеристиками информации, а так же как и в чём их оценивать (измерять).

### 1.3.1. Формальные единицы измерения информации в технике

В теории кодирования и передачи сообщений под количеством информации понимают количество кодируемых, передаваемых или хранимых символов. При этом используют простой способ определения количества информации как число использованных символов. (Подробнее об этом методе оценки см. п. 1.3.3.1.) Он основан на подсчёте числа символов в сообщении, то есть связан с его длиной и не учитывает содержания. Скорее это не измерение количества информации, а потенциально возможная верхняя оценка информационной вместимости носителя или сообщения. Для упрощения и формализации процесса оценки в вычислительной технике символы исходного алфавита сообщения обычно кодируются двоичными числами, то есть с использованием нулей и единиц. Как следствие появились и стандартные единицы измерения: бит (*binary digit*) и байт (*byte*).

**Бит** – минимальная единица измерения информации – величина, которая может принимать одно из двух значений (в математическом представлении 0 или 1).

**Байт** – единица количества информации в системе СИ. Байт – восьмиразрядный двоичный код, с помощью которого наиболее часто представляют один символ текста (о кодировании чисел и текста байтами рассказано ниже).

**Информационный объём сообщения** (информационная ёмкость сообщения) – количество информации в сообщении, измеренное в битах, байтах или производных единицах (Кбайтах, Мбайтах и т. д.).

<sup>13</sup> Отдельные юридические аспекты в отношении информации будут рассмотрены в Главе 2.

### 1.3.2. Обозначение одного байта по ГОСТ 8.417–2002

Обозначение одного байта по ГОСТ 8.417–2002 «Государственная система обеспечения единства измерений. Единицы величин»: «байт» или «Б». При этом:  $2^{10}$  байт = 1024 байта = 1 Кбайт = 1 КБ;  $2^{20}$  байт = 1 048 576 байт = 1 Мбайт = 1 МБ =  $2^{10}$  КБ = 1024 КБ = 1024 Кбайта и т. д.

Обратите внимание, что используется буква «К» (большая, заглавная, прописная), в то время как в международной системе единиц СИ (SI, фр. *Le Système International d'Unités*, не имеет никакого отношения в языке Си (англ. C)) приставка «кило-», означающая  $1000 = 10^3$ , в русском языке обозначается буквой «к» (маленькая, строчная). Чтобы запомнить этот нюанс, приведём короткий анекдот:

*«Физик ошибочно думает, что в одном килобайте (КБ) тысяча (1000) байт, а программист – что в одном килограмме (кг) тысяча двадцать четыре (1024) грамма.»*

В зарубежной практике сегодня используются рекомендации стандарта IEEE 1541–2002, согласно которым бит (англ. bit) обозначается как «b», а байт (анг. byte) – «B». При этом, чтобы не путать «килобайты с килограммами», используются приставки:

- kibi (обозн. «Ki») =  $2^{10} = 1024$ ;
- mebi (обозн. «Mi») =  $2^{20} = 1\,048\,576$ ;
- gibi (обозн. «Gi») =  $2^{30} = 1\,073\,741\,824$  и др.

Надо заметить, что несмотря на цифру 2002 в названии, стандарт был встречен обществом «двояко» и лишь 27 марта 2008 г. был окончательно переутверждён как действующий. Скорее всего, он был «скопирован» с утверждённого ещё в 1999 г. МЭК стандарта IEC 60027-2. Отличие последнего в том, что в нём для обозначения бита вместо «b» предлагается использовать слово «bit».

Поскольку в массе своей покупатели и продавцы не сведущи в обозначениях сокращений по стандартам, благодаря стараниям маркетологов приставки kibi и прочие не используются – занижать визуально наблюдаемый объём невыгодно. Например, диск с этикеткой «1000 у. е.» будет продаваться лучше, чем с «932 ГБ», хотя в абсолютном исчислении  $932 \cdot 1024 \cdot 1024 \cdot 1024$  даже чуть больше, чем  $1000 \cdot 1000 \cdot 1000 \cdot 1000$ . Не техническому специалисту ориентироваться в подобных нюансах сложно, потому как в «сегодняшних» прайс-листах и технических спецификациях (том, что у всех на виду) запросто можно увидеть ёмкости жёстких дисков в «попугаях» (См. мультфильм «38 попугаев»).

### 1.3.3. Методы и модели оценки количества информации

*По мнению В. А. Острейковского [4]:*

Для теоретической информатики информация играет такую же роль, как и вещество в физике. И подобно тому, как веществу можно приписывать довольно большое количество характеристик: массу, заряд, объём, и т. д., – так и для информации имеется пусть и не столь большой, но достаточно представительный набор характеристик единицы измерения, что позволяет некоторой порции приписывать числа – количественные характеристики информации.

На сегодняшний день известно много методов измерения информации, рассмотрим некоторые из них, в том числе наиболее часто встречающиеся:

- объёмный;
- энтропийный;
- алгоритмический.

### 1.3.3.1. Объёмный метод

Объёмный метод – является самым простым и грубым методом измерения информации. Соответствующую количественную оценку информации естественно назвать объёмом информации. Объём информации в сообщении – это количество символов в сообщении. Поскольку, например, одно и то же число может быть записано многими разными способами (с использованием различных алфавитов): «двадцать один», «twenty one», «einundzwanzig», 21, 11001<sub>2</sub>, 15<sub>16</sub>, XXI, то этот способ чувствителен к форме представления (записи) сообщения.<sup>14</sup>

### 1.3.3.2. Энтропийный метод (информация как снятая неопределённость)

В теории информации и кодирования применяется энтропийный метод измерения информации. Этот метод измерения исходит из следующей модели. Получатель информации имеет определённые представления о возможности наступления некоторых событий. Эти представления в общем случае достоверны и выражаются вероятностями, с которыми он ожидает то или иное событие (*например, сигнал – прим. авт.*). Общая мера неопределённости (энтропия) характеризуется математической зависимостью от совокупности этих вероятностей. Количество информации в сообщении определяется тем, насколько уменьшится эта мера после получения сообщения.

Создателем *общей теории информации* и основоположником цифровой связи считается Клод Шеннон<sup>15</sup>. Он ввёл понятие информационной энтропии – меры хаотичности информации, определяющей неопределённость появления какого-либо символа первичного алфавита. При отсутствии информационных потерь численно равна количеству информации на символ передаваемого сообщения.

**Замечание.** Приводимая в конце параграфа формула для информационной энтропии по своему внешнему виду была известна физикам со времён классических исследований Максвелла, Больцмана и Гиббса в области статистической механики. Больцман обозначил такой тип суммирования символом  $H$  и показал, что в идеализированных системах эта сумма пропорциональна термодинамической величине энтропии. Те же исследователи понимали, что существует тесная связь между энтропией и информацией; однако количественной теории информации они не разработали. Они и не могли её разработать, ибо в те времена не приходилось сталкиваться с

<sup>14</sup> В вычислительной технике вся обрабатываемая и хранимая информация вне зависимости от её природы (число, текст, изображение, изображение и прочее) представлена в двоичной форме (с использованием алфавита, состоящего из двух символов 0 и 1). Такая стандартизация позволила ввести две стандартные единицы измерения: бит и байт. Байт – это восемь бит. Подробнее см. раздел «Представление информации в компьютере».

<sup>15</sup> Клод Элвуд Шеннон (англ. *Claude Elwood Shannon*; 1916–2001) – 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 внёс огромный вклад в теорию вероятностей, теорию автоматов и теорию систем управления – области наук, входящие в понятие «кибернетика». В 1948 году в статье «*Математическая теория связи*» (*A Mathematical Theory of Communication*) ввёл термин «*бит*» для обозначения наименьшей единицы информации и обосновал возможность применения двоичного кода для передачи информации, что и принесло ему всемирную известность.

передачей информации техническими средствами.[10] Вполне естественно, что Шеннон использовал те же обозначения, которые сейчас приняты повсеместно.

**Количеством информации** называют числовую характеристику сигнала, не зависящую от его формы и содержания и характеризующую неопределённость, которая исчезает после получения сообщения в виде данного сигнала – в этом случае количество информации зависит от вероятности получения сообщения о том или ином событии. Оценка количества информации основывается на законах теории информации.

Рассмотрим пример [4], предложенный В. А. Острейковским:

Пусть имеется колода карт, содержащая 32 различные карты. Вероятность выбора одной карты из колоды равна  $1/32$ . Априори (доопытно, до произведения выбора) естественно предположить, что наши шансы выбрать некоторую определённую карту одинаковы для всех карт колоды. Произведя выбор, мы устраняем эту априорную неопределённость. Нашу априорную неопределённость можно было бы охарактеризовать количеством возможных равновероятностных выборов. Если теперь определить количество информации как меру устраненной неопределённости, то и полученную в результате выбора информацию можно охарактеризовать числом 32.

Однако в теории информации получила использование другая количественная оценка  $H$ , рассчитываемая по формуле Ральфа Хартли<sup>16</sup>, а именно – логарифм от описанной выше оценки по основанию 2:

$$H = \log_2 m,$$

где  $m$  – число возможных равновероятных выборов (при  $m = 2$ ,  $H = 1$ ). То есть для выбора из колоды имеем следующую оценку количества информации, получаемую в результате выбора:

$$H = \log_2 32 = 5.$$

Полученная оценка имеет интересную интерпретацию (на примере урезанной игровой колоды из 32 карт: «7», «8», «9», «10», «Валет», «Дама», «Король», «Туз» четырёх привычных мастей (♠, ♣, ♥, ♦). Она характеризует число «двоичных» вопросов, ответы на которые позволяют выбрать либо «да», либо «нет». Например, для выбора дамы пик такими вопросами могут быть:

№	Вопрос	Ответ	Интерпретация ответа
1	Карта красной масти?	Нет	0
2	Трефы?	Нет	0
3	Одна из четырёх старших?	Да	1
4	Одна из двух старших?	Нет	0
5	Дама?	Да	1

Таким образом, чтобы устранить неопределённость (сделать выбор нужной карты) в колоде карт, необходимо получить в диалоговом режиме пять вопросов – пять от-

<sup>16</sup> Ральф Хартли (англ. Ralph Vinton Lyon Hartley, 1888–1970) – американский учёный-электронщик. Он предложил генератор Хартли, преобразование Хартли и сделал вклад в теорию информации, введя в 1928 году логарифмическую меру  $H = K \cdot \log_2 N$  (в нашем случае  $K = 1$ ), которая называется хартлиевским количеством информации, содержащейся в сообщении.

ветов (сообщений). По колонке интерпретации ответов этот выбор можно описать последовательностью из пяти двоичных символов 00101. Количество информации такого сообщения составляет 5 бит (равно числу вопросов в нашем случае).

Сообщение, уменьшающее неопределённость знаний в два раза, несёт 1 бит информации.

На первый взгляд может показаться, что эта интерпретация не годится в случае, когда количество выборов не равно степени двойки, так как получается нецелое количество вопросов, к примеру если взять колоду из 36 карт (добавлены шестёрки), то можно заметить, что для того, чтобы выяснить у участника «эксперимента», какую карту он выбрал, в ряде случаев понадобится 5 вопросов, как и в предыдущем случае, а в ряде случаев – и 6 вопросов. Усреднение по случаям и даёт получаемую по формуле нецелую величину.

### Двоечник и отличник

Рассмотрим ещё один пример. Допустим, что имеется двое обучаемых с обоснованно присвоенными общественными характеристиками «двоечник» и «отличник». Мы подсчитали количество полученных ими оценок за учебный год и занесли данные в таблицу.

Оценки	«Двоечник»	«Отличник»
«Отлично» (5)	3	85
«Хорошо» (4)	15	9
«Удовлетворительно» (3)	20	5
«Неудовлетворительно» (2)	60	1
«Плохо» (1)	2	0

Так получилось, что суммарно они получили каждый по 100 оценок (это ни на что не влияет, но зато упростит наши последующие расчёты). Подсчитаем некоторые характеристики, а именно вероятности получения оценок. Приблизённо<sup>17</sup> вероятностью будем считать отношение числа тех или иных оценок, поделённое на общее число выставленных оценок в год. В итоге получим:

Вероятность какого события	«Двоечник»	«Отличник»
вероятность получения пятёрки	$3/100 = 0,03$	$85/100 = 0,85$
вероятность получения четвёрки	$15/100 = 0,15$	$9/100 = 0,09$
вероятность получения тройки	$20/100 = 0,2$	$5/100 = 0,05$
вероятность получения двойки	$60/100 = 0,6$	$1/100 = 0,01$
вероятность получения единицы	$2/100 = 0,02$	$0/100 = 0$

<sup>17</sup> Мы предполагаем, что результаты опыта сводятся к схеме случаев, при которой вероятность события  $A$  вычисляется по формуле  $P(A) = m/n$ , где  $n$  – общее число случаев, а  $m$  – число случаев, благоприятных событию  $A$  (См. стр.5. Вентцель Е. С., Овчаров Л. А. Теория вероятностей. – 2-е изд., стер. – М.: «Наука» Главная редакция физико-математической литературы, 1973 – 368 с.: ил.).

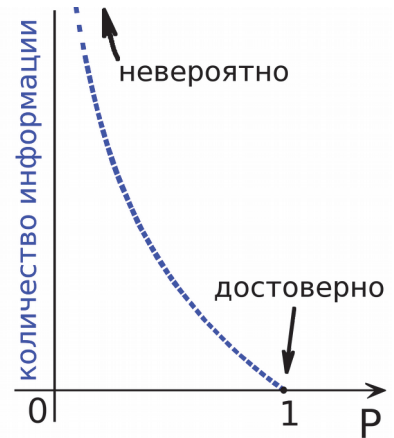
Логично предположить, что сообщение вида «Мама, я получил оценку 5 по информатике» от отличника ожидаемо и несёт мало новой информации для родителей. В то же время, если то же самое скажет двоечник, в его семье будет праздник. То есть:

чем выше вероятность события,  
тем меньше информации оно несёт

и

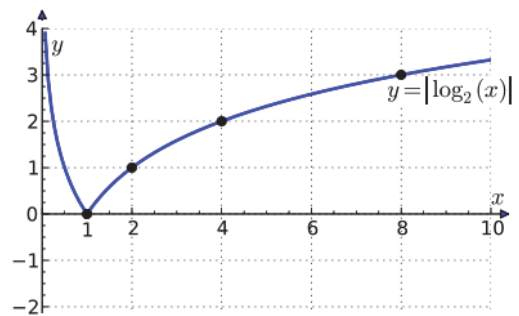
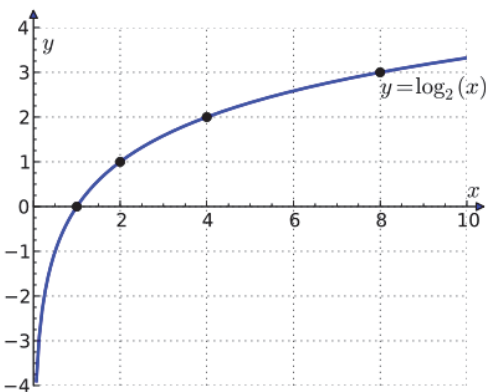
чем ниже вероятность события,  
тем больше информации оно несёт.

Сам факт того, что ребёнком получена оценка в школе обычно родителей не интересует так сильно как то, какая именно. Этому есть объяснение. Если предположить что в какой-то школе «Б1» оценки ставят ежедневно (или еженедельно), то наличие поставленной оценки за отведённый период означает, что учитель обязательно какую-то оценку да поставил. Или, вероятностным языком, произошло одно из пяти событий: «получена пятёрка», «получена четвёрка», «получена тройка», «получена двойка», «получена единица». Как вы понимаете, раз других оценок быть не может, то такое событие (как выбор одной оценки из возможных) будет **достоверным событием**, вероятность которого можно принять за единицу, а информацию от этого сообщения можно считать нулевой.



**Невероятное событие** – это то, у которого вероятность стремится к нулю, а количество полученной при этом информации велико (на графике асимптотически стремиться к бесконечности). Скажем, прилетели инопланетяне и поэтому все уроки отменили.

Мы не берёмся оценить (как и доказать) правдоподобность степени кривизны отображённой нами линии, как и сделать выводы относительно её выпуклости (впуклости). Предположим, что наиболее подходящей по ряду качеств будет функция «модуль логарифма». См. Рисунок 1.1.



как видно,  $y = |\log_2(x)| = -\log_2(x)$ ,  
при  $x \in (0,1]$

Рисунок 1.1. Графики логарифма и модуля логарифма

В защиту нашего предположения о выборе в пользу логарифма вероятности приведём взятые из теории информации следующие доводы:

1. Необходимо, чтобы количество информации было аддитивно. Например если у нас последовательно переданы два независимых сообщения  $x_1$  и  $x_2$ , то количество информации в суммарном сообщении должно равняться сумме количества информации каждого сообщения по отдельности.

2. Количество информации содержащееся в достоверном сообщении равно нулю (новую информацию заранее известное сообщение не несёт).

3. Информационная мера не должна зависеть от степени восприятия информации конкретным получателем (информация должна быть объективной).

Таким образом, наши интуитивные представления можно описать более строго.

Определим количество информации  $h_i$ , содержащееся в каждом сообщении о полученной оценке, исходя из предложенной выше статистики оценок. Для этого напомним свойство логарифма (известное читателям из школьного курса алгебры), что

$$\log_b \frac{1}{a} = \log_b a^{-1} = -\log_b a \quad \text{и воспользуемся формулой} \quad h_i = \log_2 \left( \frac{1}{p_i} \right),$$

где  $p_i$  – вероятность получения  $i$ -й (по значению, а не порядку) оценки.

Таким образом, количество информации <sup>18</sup>:

Событие	«Двоечник»	«Отличник»
получена оценка 5	$\log_2(1/0,03) \approx 5,059$	$\log_2(1/0,85) \approx 0,234$
получена оценка 4	$\log_2(1/0,15) \approx 2,737$	$\log_2(1/0,09) \approx 3,474$
получена оценка 3	$\log_2(1/0,2) \approx 2,322$	$\log_2(1/0,05) \approx 4,322$
получена оценка 2	$\log_2(1/0,6) \approx 0,737$	$\log_2(1/0,01) \approx 6,644$
получена оценка 1	$\log_2(1/0,02) \approx 5,644$	$\log_2(1/0) = \infty$

Как видим, ранее предположенное утверждение, что количество информации в сообщении о некотором событии зависит от вероятности этого события, подтвердилось. Чем меньше вероятность, тем больше информации.

Обычно передаваемое сообщение не состоит из одного передаваемого символа. Оно может состоять из нескольких независимых символов (появление символа – событие), тогда **информационная энтропия** для  $m$  независимых случайных событий с  $n$  возможными состояниями (от 1 до  $n$ ) рассчитывается по формуле Шеннона:

$$H(m) = -m \sum_{i=1}^n p(i) \log_2 p(i),$$

где  $p(i)$  – вероятность  $i$ -го события. Пользоваться этой формулой в отношении оценки количества содержащейся информации в используемых нами словах и предложениях (текстах) невозможно поскольку буквы в них оказываются зависимыми друг от друга.

**Замечание 1.** Откуда в формуле взялся минус? – См. Рисунок 1.1.

**Замечание 2.** В примере выше основание логарифма равно двум, но в ряде случаев оно может не указываться, поскольку это не существенно в целом. Изменение основания приведёт лишь к изменению размерности.

<sup>18</sup> Если у вас под рукой не оказалось калькулятора или логарифмической линейки, то предлагаем воспользоваться сайтом <http://kalkulyatoronline.ru/>.



### 1.3.3.3. Алгоритмический метод

В теории информации существует алгоритмический метод оценки количества информации в сообщении. В. А. Острейковский предлагает этот метод охарактеризовать следующими рассуждениями.

Под алгоритмом всегда понималась процедура, которая позволяла путём выполнения последовательных элементарных шагов (действий) получать однозначный результат (независимо от того, кто выполнял эти шаги) или за конечное число шагов прийти к выводу о том, что решение не существует.

Каждый согласится, что слово 0101010101...01 сложнее слова 0000000000...00, а слово, где 0 и 1 выбираются из эксперимента – бросания монеты (где 0 – герб, 1 – аверс<sup>19</sup>), сложнее обоих предыдущих.

Компьютерная программа, производящая слово из одних нулей, крайне проста: печатать один и тот же символ. Для получения 010101 ... 01 нужна чуть более сложная программа, печатающая символ, противоположный только что напечатанному. Случайная, не обладающая никакими закономерностями последовательность не может быть произведена никакой «короткой» программой. Длина программы, производящей хаотическую последовательность, должна быть близка к длине последней.

Приведённые рассуждения позволяют предположить, что любому сообщению можно приписать количественную характеристику, отражающую сложность программы, которая позволяет её произвести.

Так как имеется много разных вычислительных машин и разных языков программирования (разных способов задания алгоритма), то для определённости задаются некоторой конкретной вычислительной машиной, например машиной Тьюринга<sup>20</sup>.

### 1.3.4. Семантическая (смысловая) мера информации

정보의 의미론적 척도는 전술한 정보측정방법에 포함시킬 수 없는데, 의미론적 수준에서의 정보량 추정의 결과는 정보를 받는 사람, 예를 들어 객관적이지 않은 상황을 제시하는 특정한 사람의 특성에 의존하기 때문이다.

Если вы пропустили предыдущий абзац и начали сразу читать данную строчку, то вы попались на уловку придуманную авторами, ещё раз подтвердив истинность утверждения, что:

Семантическая мера информации не может быть отнесена к вышеописанным методам измерения информации, поскольку результат оценки количества информации на семантическом уровне зависит от свойств её получателя, например конкретного человека, что даёт не объективную картину.

Если среди читателей, их родственников или знакомых случайно найдутся знатоки корейского языка, то они подтвердят, что первый и третий абзацы данного параграфа идентичны по смыслу, то есть несут одинаковое количество информации<sup>21</sup> для чита-

<sup>19</sup> Аверс (фр. *avers*, лат. *adversus* – «обращённый лицом») – лицевая, главная сторона монеты.

<sup>20</sup> Машина Тьюринга – абстрактный исполнитель (абстрактная вычислительная машина). Была предложена Аланом Тьюрингом в 1936 году для формализации понятия алгоритма.

<sup>21</sup> Заметим, что из равенства количества информации на семантическом уровне (в нашем случае у двух предложений одинаковый смысл) не следует, что объёмы памяти ЭВМ занимаемые этими же предложениями будут равны. Подробнее о кодировании текстовой информации см. параграф 2.2. Представление текстовой информации в ЭВМ.

теля, однако воспринять эту информацию получится не у всех. Подумайте на досуге, что бы могла означать используемая в обиходе фраза «читайте между строк» с научной точки зрения.

По мнению Акулова О.А., Н.В.Медведева [5] и Макаровой Н.В., Волкова В.Б. [6], для измерения смыслового содержания информации, то есть её количества на семантическом уровне, наибольшее признание получила тезаурусная мера, которая связывает семантические свойства информации со способностью пользователя понимать поступившее сообщение. Для этого используется понятие тезауруса пользователя.

Тезаурус<sup>22</sup> – множество смысловыражающих единиц некоторого языка с заданной на нём системой семантических отношений. Тезаурус фактически определяет семантику языка (национального языка, языка конкретной науки или формализованного языка для автоматизированной системы управления) [7]. В отличие от толкового словаря, тезаурус позволяет выявить смысл не только с помощью определения, но и посредством соотнесения слова с другими понятиями и их группами. Более коротко:

**Тезаурус** – это совокупность сведений, которыми располагает пользователь или система [6].

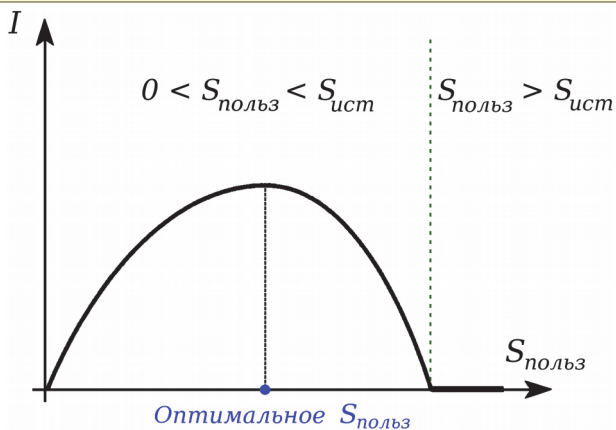


Рисунок 1.2. Зависимость количества воспринимаемой пользователем информации  $I$  от его тезауруса  $S_{польз}$  ( $S_{ист} = const$ )

Отметим, что когда тезаурус пользователя стремится к нулю, то и воспринимаемая им информация также стремится к нулю. Согласитесь, что попытка объяснить высшую математику только научившемуся говорить двухлетнему ребёнку практически бесполезное занятие.

Также, когда  $S_{польз} \geq S_{ист}$ , источник не несёт новой информации пользователю, как если рассказывать про количество букв в алфавите или про таблицу умножения выпускникам школы.

Максимальное количество семантической информации потребитель приобретает при согласовании её смыслового содержания со своим тезаурусом (См. значение «Оптимальное  $S_{польз}$ » на Рисунке 1.2 соответствующее экстремуму графика), когда поступающая информация понятна пользователю и несёт ему ранее неизвестные (отсутствующие в его тезаурусе) сведения [5, 6].

В зависимости от соотношения между смысловым содержанием информации источника  $S_{ист}$  и тезаурусом пользователя  $S_{польз}$  изменяется количество семантической информации  $I$ , воспринимаемой пользователем (и включаемой им в дальнейшем в свой тезаурус) [5, 6].

Если зафиксировать  $S_{ист}$ , то характер зависимости  $I(S_{польз})$  представлен на Рисунке 1.2.

<sup>22</sup>От греч.  $\theta\eta\sigma\alpha\upsilon\rho\acute{o}\varsigma$  (thesaurós) – сокровище, сокровищница

**Вывод:** Народное шутливое утверждение, что «Гениальному писателю нужен гениальный читатель.» возможно подвергнуть сомнению, поскольку последний, если окажется умнее первого, может заскучать при чтении, так как не откроет для себя ничего нового.

### 1.3.5. Прагматическая мера информации (мера полезности)

*После прочтения раздела подумайте, почему в народе говорят: «Гусь свиные не товарищи» или «Дорога ложка к обеду»?*



Александр Александрович Харкевич (1904–1965), советский учёный в области электродинамики и электроакустики. Его «Очерки общей теории связи» (1955) стали одной из основополагающих работ в теории информации.

Довольно интересный подход к измерению меры полезности информации был предложен отечественным учёным А. А. Харкевичем, который предложил принять за меру ценности информации приращение вероятности достижения некоторой цели. Зачастую это более понятно обывателю, нежели какая-то там вероятность сообщения или энтропия.

Например, вы школьник средних классов и мечтаете лечить людей от болезней. Это ваша цель. Для её достижения у вас может быть несколько путей, заняться фундаментальными исследованиями в медицине и в итоге создавать новые лекарства, придумывать новые методы лечения болезней, либо стать врачом и использовать уже имеющиеся препараты и методы. В любом случае, в старших классах вам предстоит сделать выбор своего жизненного пути и, возможно, по окончании школы поступить в какой-нибудь вуз. Ваш друг, как ему кажется, напротив, далёк от медицины, интересуется компьютерами и планирует поступать в какой-нибудь технический университет. Он сообщает вам, что ездил на выходных в интересующий его вуз и рассказывает подробную информацию о работе в нём с абитуриентами.

Если не брать в рассмотрение социальные и психологические аспекты общения людей, как то желание двух друзей поговорить, умение преподнести информацию, вопросы вежливости и т.п., то сообщённая вам информация скорее всего будет бесполезна, поскольку она не приблизит вас к вашей цели, хотя она будет вам полностью понятна (воспринята), полна, достоверна, актуальна и т.д. Объективно ситуация видится следующим образом: Если до получения сообщения вероятность достижения вами цели равнялась  $p_0$ , а после её получения  $p_1$ , то ценность информации определяется как:

$$I = \log p_1 - \log p_0 = \log \frac{p_1}{p_0}$$

Причём эта формула допускает и отрицательную ценность информации, в случае снижения вероятности достижения цели после приёма сообщения. В конкретном случае это может произойти по той причине, что ваш друг убедит вас пойти с ним учиться на программиста в его вуз за компанию, а курсе на третьем, например, он женится, ста-

нет семейным человеком и будет общаться с вами реже, а вы вспомните, что первоначальная жизненная цель у вас была совсем другая.

В общем случае, отрицательная ценность может говорить о том, что сообщение содержит дезинформацию. Создана такая ситуация кем-то специально или так получилось случайно, – это уже другой вопрос.

Дальнейшее развитие данного подхода базируется на статистической теории информации и теории решений. Сущность метода состоит в том, что, кроме вероятностных характеристик неопределённости объекта, после получения информации вводятся функции штрафов или потерь, и оценка информации производится с учётом возможной минимизации потерь. Максимальной ценностью обладает информация такого сообщения, которое уменьшает потери до нуля при достижении поставленной цели.

## 1.4. Заключение

Всё в жизни крутится вокруг нас, людей, даже информация измеряется относительно нас и для нас. Иначе откуда взяться таким расплывчатым понятиям как «обилие информации», «избыток информации», «информационная зависимость», «информационная диета» или «гиперинформированность»?

Обсуждение различных информационных вопросов с гуманитарной точки зрения не менее интересно, но к сожалению выходит за возможности данного учебника, поэтому лишь порекомендуем читателям самостоятельно изучить интересное на наш взгляд интервью с Алексеем Козыревым<sup>23</sup> по теме «Изобилие информации: это благо или бремя?» [11].

Что же вопросов **измерения информации**, то в жизни мы чаще имеем дело не с абстрактной информацией, а с конкретными физическими объектами и явлениями вокруг нас, поэтому все наши интересы сводятся к измерению тех или иных физических величин. Прямо или косвенно. Если не углубляться в вопросы квантовой физики и метрологии, то измерения получаются двух видов:

- либо счётные (точно измеримые): например 2 карандаша, один ластик;
- либо непрерывные – измеряемые примерно.

С первыми обычно всё понятно, а вот со вторыми возникают трудности. Информация как характеристика какого-либо физического объекта, например о цвете яблока созревающего на дереве, сильно зависит от времени получения этой информации. Поэтому нам важно понять (договориться) когда и как измерять. Сместем предположить, что в июне и в июле результаты будут разными. А ёлка, как мы знаем с детства, наоборот, зимой и летом окажется одним цветом. По сути, в примере с яблоком, как цвет есть непрерывная величина плавно переходящая от зелёного к красному<sup>24</sup>, так и момент его измерения может быть растянут во времени. Вот тут и возникает трудность: как «бинарному» по своей сути компьютеру, используя «нули» и «единицы», записать в себя что-то непрерывное (бесконечное) и желателно абсолютно точно? Это невозможно, но, поскольку очень хочется, можно приблизиться и значение физической величины (способной принимать бесконечное множество значений на некотором диапа-

<sup>23</sup> Деканом философского факультета МГУ на момент интервьюирования.

<sup>24</sup> Будем считать, что при созревании яблоко окрашивается равномерно по всем направлениям.

Конец ознакомительного фрагмента.

Приобрести книгу можно

в интернет-магазине

«Электронный универс»

[e-Univers.ru](http://e-Univers.ru)