

О РАСШИРЕНИИ ПОНЯТИЯ ИНФОРМАЦИИ СОВРЕМЕННЫМИ СРЕДСТВАМИ ИНФОРМАТИКИ

Проведен анализ понятия информация с учетом последних достижений в теоретической информатике. Показано, что сложившееся для технических наук определение, базирующееся на отображении интерпретации сообщений в информацию, не укладывается в классические представления математики, формализованно рассматривающих только статические множества. В информатике действующие множества сообщений и множества образов интерпретации оказываются динамическими совокупностями состояний объектов. Информационный аспект связей элементов оказывается принципиально динамическим, и сам характер изменения множества интерпретируемых сообщений и образов интерпретации составляет существо получения и обработки информации.

Ключевые слова: информация, сообщение, интерпретация, отображение, динамические множества.

Введение. Общетеchnическое понятие информации сложилось на основе работ К. Шеннона по математической теории связи [1]. В последующие годы, в связи с бурным развитием теории и практики технического управления, оно стало широко употребляться для описания процессов взаимодействия в технических системах. В связи с созданием и развитием компьютеров его стали применять в области компьютерных наук и их ближайших приложений, оно было использовано также в междисциплинарной науке кибернетике. Сама кибернетика была определена Н. Винером как наука «об общих закономерностях получения, хранения, передачи и преобразования информации в сложных управляющих системах, будь то машины, живые организмы или общество» [2]. Кроме этого, Н. Винер вводил определение, что «*Информация — это обозначение содержания, полученное нами из внешнего мира в процессе приспособления к нему нас и наших чувств*».

В математической теории информации ее описание базируется на статистическом подходе, что дает цельное и глубокое представление о количественной стороне передачи информации. Но он оказывается содержательно значимым только для закономерностей и процессов передачи и получения информации, поскольку хранение информации оказывается, по существу, средством ее передачи от одного момента времени (запоминания) к другому моменту (извлечения из памяти). Это понимание может легко переноситься и на задачи хранения информации в технических или естественных системах. Что же касается преобразования информации в широком ее применении, то количественный подход, основанный на вероятностных методах, оказывается, по существу, малоприменимым. В рассмотрении технических устройств, работающих по отдельным фиксированным алгоритмам, фактическое взаимодействие этих устройств с окружающим миром, в частности, целевое функционирование таких устройств оказывается существенным

именно со стороны взаимодействия, то есть опять же с позиций случайного в общем случае поступления внешних воздействий и потока вводимых данных [3]. Но существо процесса преобразования информации в более сложных современных информационных системах в подавляющем большинстве ситуаций определяется не случайными факторами, а функционально детерминированными внутренними конструкциями и функциональной конструкцией таких систем.

Более того, на смену жестким алгоритмам обработки всего поступающего потока данных, имевшим место в первых применениях компьютеров, стали широко применяться программные методы обработки, зависящие от данных. Под жесткий алгоритм в данном контексте понимается алгоритм, не использующий динамических данных, а именно данных с изменяемым размещением в памяти как внутренних состояний системы и не использующий указатели на размещение данных. В чисто техническом аспекте эти методы, зависящие от данных, называются объектно-ориентированными и для них характерно, что вместо постоянного алгоритма преобразования входных данных используются для общей цели различные процедуры преобразования, причем выбор конкретного варианта осуществляется в зависимости от конкретного типа данных.

К настоящему времени понятие «информация» определяется международным стандартом ISO/IEC/IEEE 24765: 2010 Systems and software engineering. Согласно пункту 3.1396 этого стандарта, информация есть сведения или знания, которыми их пользователи обмениваются о вещах, фактах, понятиях и т. п. В процессах обработки информации, информация — это знание, связанное с объектами, такими как факты, события, вещи или идеи, включая понятия, которые имеют частное значение внутри определенного контекста.

При этом особенно отмечается, что хотя информация должна иметь форму представления, чтобы ею можно было обмениваться, но она, в первую

очередь, есть интерпретация этого представления в виде данных. (Although information will necessarily have a representation form to make it communicable, it is the interpretation of this representation (the meaning) that is relevant in the first place). Практически этим подчеркивается, что информация в строгом смысле отличается от данных, но в неформальном контексте эти два понятия используются как синонимы.

С точки зрения теоретико-формализованных оснований, такое определение опирается на формализацию Ф. Бауэра, который определял информацию на основе более простого понятия сообщения, описываемого как цепочка символов из некоторого алфавита [4]. Его подход состоял в указании связи между сообщениями и информацией, которая записывается в виде отображения $N \rightarrow J$, где N — сообщение, а J — информация, причем правило соответствия называют *интерпретацией* сообщения. Используя сложившийся математический подход, Бауэр рассматривал множество \mathfrak{R} сообщений и множество сведений (информаций), между которыми определено математическое понятие отображения.

Указанный подход позволил математически формализовать представление об обработке информации. Для этого служит коммутативная диаграмма

$$\begin{array}{ccc} \mathfrak{R} & \xrightarrow{\alpha} & \mathfrak{Z} \\ \downarrow \gamma & & \downarrow \sigma \\ \mathfrak{R}' & \xrightarrow{\alpha'} & \mathfrak{Z}' \end{array}$$

Здесь \mathfrak{R} , \mathfrak{R}' — множества сообщений, \mathfrak{Z} , \mathfrak{Z}' — множества сведений (информаций), γ — отображение преобразования сообщения из множества \mathfrak{R} исходных сообщений в множество \mathfrak{R}' обработанных сообщений, α , α' — отображения интерпретации. В общем случае для произвольных отображений α , α' и γ отношение σ между элементами множеств \mathfrak{Z} и \mathfrak{Z}' может не быть отображением. Для того, чтобы вместо более общего неоднозначного отношения имело место именно отображение каждого элемента из \mathfrak{R} в один-единственный элемент множества \mathfrak{Z}' , необходимо выполнение математического условия

$$\forall n_1 \in \mathfrak{R}, n_2 \in \mathfrak{R}, n_1 \neq n_2, \alpha(n_1) = \alpha(n_2) \text{ должно быть} \\ \alpha'(\gamma(n_1)) = \alpha'(\gamma(n_2)).$$

Иначе говоря, прообраз каждого элемента информации $j \in \mathfrak{Z}$ после отображения γ обработки информации и последующего отображения α' должен состоять из единственного элемента множества \mathfrak{Z}' . Поскольку конструктивное строение и непосредственное использование множества информации обычно принципиально не детализируется и практически недоступно, то такое требование легко декларируемо, но далеко не всегда доступно для проверки. Именно отсюда в конечном случае вытекает множество проблем достоверности результатов обработки, единственности их понимания и доступности отладки с устранением алгоритмических причин ошибок.

Для текущего рассмотрения принципиально важным является тот факт, что совокупность элементов, в которую происходит отображение интерпретации, принципиально имеет не постоянный состав.

Например, английское слово *file* еще по данным англо-русского словаря 1990 г. [5] интерпретируется как досье, подшитые бумаги и еще множество

других вариантов интерпретации. Интерпретация этого слова как обозначение именованного объекта хранения информации в компьютерной системе для массового пользователя появилось относительно недавно, а в первых поколениях компьютеров никаких файлов и представления о них не было. Практически большинство слов естественного языка имеет устойчивую интерпретацию только на не очень большом промежутке времени порядка многих десятков или сотен лет.

В современных компьютерных системах интерпретация ряда стандартных подпрограмм заметно отличается в зависимости от версии соответствующих библиотек (в них со временем появляются дополнительные возможности или удаляются устаревшие). Даже стандартизованный язык программирования по существу своих возможностей и действий зависит от стандарта, в частности, язык C стандарта ISO C89 заметно отличается от него же стандарта ISO C11 [6].

Чисто теоретически можно вместо обычных математических множеств для прообраза и образа отображения рассматривать семейство таких множеств, проиндексированных параметром времени. Какие-то утверждения относительно такой системы отображений будут математически точными только для фиксированного значения параметра, а возможность общего строгого обсуждения всего отображения без постоянного учета момента времени исчезает. В то же время большинство обсуждений понятия информации и их использование никак не готово прямо учитывать указанный параметр.

Важно понимать, что указанная особенность со спецификой обработки информации имела место всегда, но из-за внутренности таких процессов внутри человеческого разума не попадала во внимание. Компьютерные же особенности не просто бросаются в глаза, они широко используются, но уже не для разовой реакции на поступление сообщения, а в широкой временной панораме длительных и преемственных процессов обработки информации в современных информационных системах, включающих как технические средства, так и человеческую деятельность по изменению и усовершенствованию систем обработки.

Обычно подразумевается, что хранение и обработка информации в компьютерах — это специализированное техническое использование ее, а аналогичные действия с информацией у людей гораздо более общие и охватывают человеческое знание. Это убеждение складывается просто по рассуждению, что у людей есть разум, а у машин его нет. На самом деле искусственные средства действий с информацией просто в более детализированной и четко различающейся форме реализуют методы и подходы принципиально общих подходов к информации, но сами общие подходы могут быть одними и теми же. Ситуация здесь подобна химии, используемой в человеческих организмах, особенности имеют место только путем использования количественно более сложных конструкций молекул, но все реальные механизмы и законы одни и те же что в организме, что в технической химии.

Отсюда видна принципиальная возможность рассмотреть и подвергнуть дальнейшему анализу внутренние механизмы обработки информации в любых системах, в том числе в человеке, оттачиваясь от изучения и осмысления сложившихся и работающих подходов в современных технических средствах. Прежде всего можно взглянуть

на складывающееся к современному моменту существо содержания и основного использования информации в технических, в первую очередь программных системах, и обобщить там, где это явно не вызывает противоречия с наблюдаемыми реально фактами, на общечеловеческую область. Тогда можно утверждать следующее.

Информация — распознанное или отраженное (зафиксированное во внутренних состояниях системы) внешнее воздействие, являющееся основанием для последующих действий этой системы.

Под это определение тут же подпадают команды, вызовы программ, в частности, команды интерпретатора, фрагмент или элемент директив управления. Таким фрагментом является командный код программы в любых внешних или внутренних обозначениях, в частности, в виде команд естественного языка. Например, «пойти туда-то», «сделать сначала то, потом другое, посмотрев на промежуточном шаге на состояния, запомненные ранее в результате предыдущих воздействий». Примером явной информации оказывается условная фраза, по которой исполняющая сторона немедленно приступает к выполнению заранее запланированной последовательности действий.

Хорошей иллюстрацией является малоизвестный, но исторически очень существенный пример. Во время Второй мировой войны начало действий на Тихом океане с нападением Японии на базу военно-морского флота США на Гавайях началось с выдачи в «радиоэфир» сообщения, замаскированного под обычные сообщения о погоде. Были ретранслированы несколько раз подряд слова «западный ветер — ясная погода», по которому в тех посольствах Японии за рубежом, которые относились к союзникам США, и в самих США персонал стал срочно уничтожать все секретные документы. Эта процедура была запланирована заранее по предыдущим секретным контактам, а запущена одной условной фразой в открытом радиодоступе [7].

Непостоянность множеств \mathfrak{R} , \mathfrak{J} , в теоретически формализованном соответствии $\mathfrak{R} \rightarrow \mathfrak{J}$, оказывается не просто досадным неудобством разрабатываемой теории. Оно является принципиальной особенностью рассматриваемого предмета изучения и отражает тот важнейший факт, что в практических областях применения информации используются *динамические множества*. Данный термин порожден по аналогии с динамическими массивами современных языков и систем объектно-ориентированного программирования [8] и соответствует удивительно для многих особенностей не детерминированно изменяющихся совокупностей. Именно с такими совокупностями объектов или некоторых элементов приходится иметь дело во множестве практических областей и, в частности, при функционировании современных информационных систем. Эта ситуация совершенно нехарактерна для математики более чем столетнего интервала времени, после утверждения теории множеств.

Принципиальное свойство непостоянства множеств лишает человека возможности использовать (для их изучения и эффективного в приложениях формализованного описания) классические приемы доказательств по принадлежности элементов, на которых основаны алгебра множеств и основанные на ней более сложные разделы математики [9]. Представляется, что именно поэтому формализация Бауэра не нашла серьезного распространения и прямых теоретических приложений. Но этот

недостаток не помешал активному практическому развитию систематизации и изучению области преобразований и технического использования информации, в том числе в прикладной дисциплине информатики.

Информация, как зафиксированное во внутренних состояниях системы внешнее воздействие, являющееся основанием последующих действий этой системы, широко применяется при построении языков программирования, программных систем программирования со всеми вспомогательными средствами и в операционных системах. Но при этом вместо привычного для математики языка теории множеств и ее доказательств, имеющих всеобщую и постоянную значимость, приходится выполнять описание, создание технических систем и изучение функционирования последних исключительно на программно-описательном уровне и декларативных средствах описания данных, например, языках запросов данных к базам данных.

С учетом выявленной специфики понятия информации и принципиальных особенностей ее обработки представляет несомненный интерес уточнение с возможно более общих позиций той связи сообщений с информацией, которая для отдельных элементов дается отображением $N \rightarrow J$, но не укладывается в формализм отображения классических множеств.

Результаты отображения интерпретации сообщений практически всегда, за исключением искусственных жестких схем и устройств, является динамическим множеством. Оно практически постоянно или систематически пополняется за счет эффектов «накопления опыта», фиксации новых существенных для функционирования системы новых сообщений и попыток их интерпретаций как на основе ранее имеющейся информации, так и аналогий или случайного выбора поведения. Это утверждение достаточно очевидно для системы человеческого разума, для социальных и биологических систем. Для технических систем его действие во многих случаях поддерживается человеческим вмешательством по явному или неявному запросу таких систем. Так, при поступлении исходного кода программы на вход транслятора в том случае, когда последний не может однозначно разрешить ситуацию, пользователю выдается сообщение о такой неоднозначности, обычным результатом которой является уточнение исходного кода с последующим перезапуском транслятора. Не столь частным результатом является переделка — усовершенствование транслятора для распознавания им указанной конкретной ситуации с исходным сообщением, которым в данном случае является сам исходный код.

Очень близкой по существу с этим техническим приемом служит ситуация с профессиональным обучением, когда экзаменатор отмечает неполноценно информативное содержание ответа или действия обучаемого, и последний вынужден выдавать дополнительные тексты «продолжения сообщения», позволяющие преподавателю однозначно интерпретировать сообщения от него.

Динамическим множеством оказывается и собственно отображение интерпретации как совокупность упорядоченных пар (сообщение, информация), причем эта динамичность обусловлена не только изменяемым составом динамического «множества информации» в системе, но и тем, что одним и тем же сообщениям в последующем оказываются сопоставляемыми другие элементы результирующей

щего множества информации. Эта особенность хорошо видна в естественных системах человеческого общества, когда одно и то же слово естественного языка теряет прежнее значение и становится обозначающим другое понятие или объект.

Область прообраза отображения интерпретации также является динамическим множеством (за исключением простейших технически реализованных жестких алгоритмов). В частности, современный «компьютер» (практически текущее состояние его операционной системы с множеством программных приложений) с течением времени получает возможность запускать новые по наименованию программы (а некоторые старые могут далее не вызываться — не интерпретироваться им). Для человеческой информационной системы подобное замечание тем более очевидно, человек со временем научается интерпретировать новые слова и новые понятия.

Динамические множества есть характерная особенность современных (но не первых) компьютерных систем. Так, даже традиционные базы данных постоянно пополняются и модифицируются. Это явление стало очевидным уже с тех компьютеров, которые стали хранить в постоянной памяти множество файлов (а не только запускать вводимую извне в них отдельную программу без хранения между такими запусками промежуточной или накопленной информации).

Классическое определение алгоритма никак не включает динамическое управление данными. Данные в алгоритме неявно предполагаются связанными фиксированными связями с локальным местом хранения данных (фиксированной ссылкой на именованное данное как переменную). Эта содержательная система связывания значений данных с именами переменных составляла важнейшую особенность первых языков программирования.

Расширение программирования можно рассматривать как расширение интерпретации сообщений от фиксированных в системе на динамическое связывание с элементами области интерпретаций [10]. Такое расширение потребовало не только формальных средств соответствующих языков программирования, но и создания и использования функционирующих технических средств такого связывания. Эти средства оказалось возможным построить только за счет расширенного управления памятью, выполняемого в многопрограммных системах только с помощью операционной системы. Эти средства, видимые со стороны формального программирования как системные вызовы операторов new в объектно-ориентированном программировании [8] или как системные функции malloc обеспечивают динамическое связывание данных с управляемой по распределению памятью системой. С учетом указанного динамического характера формируемых при этом динамических множеств, в классической математике не оказалось прямых аналогов этих конструкций и операций. Средства управления памятью касаются не отображений, не изменений значений в переменных, а размещения значений где-то в реальном мире, кажущихся для традиционного математического подхода далеко не абстрактной частностью. Поэтому сложилась привычка не замечать те частности, которые связаны с размещением изменяемых значений по техническим состояниям управляющей системы. Хотя, как показал проведенный выше анализ сложившейся ситуации, эти далеко не частности, а принципиальные особенности составляют в зна-

чительной степени существо проблематики информации. Именно они в первую очередь определяют существенное различие абстрактной математики, прямо никак не связанной своими конструкциями с элементами реального мира, и информатики, как дисциплины изучающей и закладывающей основы создания технических систем преобразования информации и интерпретации сообщений. Причем не только в собственно искусственных технических системах, но и в любых реально существующих информационных системах, в том числе внутри разумных организмов и социумов.

Жесткий алгоритм — не использующий динамических данных (данных с изменяемым размещением в памяти (внутренних состояниях системы) и указателей на размещение данных).

С учетом проведенного анализа представление «о получении информации» имеет не один, а несколько семантических значений (смыслов). Во-первых, получение информации обозначает *вызов информации по сообщению*, (который действует в соответствие с отображением $\mathfrak{R} \rightarrow \mathfrak{S}$). Для технических приложений, в частности, жестких алгоритмов поведения управляющих систем это основное значение. Для специалистов, знакомых с системами управления базами данных, это «получение информации» соответствует получению данных из базы по указанному значению главного индекса данных. Практически этот вариант использования обозначает не получение, а текущее использование информации, а получается собственно сообщение, информация, соответствующая ему, в системе уже хранилась ранее и только используется по назначению управления в таком вызове.

Во-вторых, «получение информации» приводит к расширению области определения отображения $\mathfrak{R} \rightarrow \mathfrak{S}$, когда в систему хранения поступает воздействие, формирующее новое, относительно ранее используемых, значение для динамического множества \mathfrak{R} . Для системы управления базами данных это соответствует созданию новой записи в базе данных с новым значением главного индекса.

Более редкое, но используемое применение «получения информации» ведет к уменьшению области определения рассматриваемого отображения, когда использовавшееся ранее значение сообщения удаляется из динамического множества \mathfrak{R} . Этой ситуации соответствует удаление записи базы данных с указанным главным индексом.

Наиболее часто используемое в сложных изменяемых системах управления, в частности в компьютерных системах и человеческом разуме, получение информации заключается в изменении области значения отображения $\mathfrak{R} \rightarrow \mathfrak{S}$, которое приводит к новому элементу области отображения. В частном случае такое изменение может заключаться в сокращении числа элементов хранения состояния (как информации внутри системы). Например, сокращение текста программы, которая вызывается по имени, служащему сообщением ее вызова. Получение информации, а фактически изменение элемента области значений отображения, может и часто заключается в изменениях ссылок на элементы хранения информации — состояния управляющей системы. В частности, в программе могут изменяться имена вызываемых ее процедур. Может просто усложняться структура или длина последовательности действий, закодированные в элементе интерпретации, что дает более сложную и предположительно более общую и для большего

числа реальных воздействий информацию [10]. Для систем управления базами данных эта разновидность получения информации заключается в изменении полей уже существующих записей данных.

Таким образом, под получением информации понимаются два основных существенно различных варианта поведения управляющих систем. Первое скорее следовало бы называть более точно именно *получением сообщений*, поскольку сообщению детерминированно для текущего состояния системы сопоставляется значение хранимой в ней информации как последовательности действий управления. Новая информация при этом не получается, хотя в ряде ситуаций динамически значимым является момент получения такого сообщения или конкретный вид последовательности сообщений. Последний вариант очень важен и должен изучаться более тщательно, так как реальное управление в системе из множества управляющих систем задается не отдельными не связанными друг с другом сообщениями, а функционально значимой последовательностью их. Ярким примером этого служит почти любая программа.

Второй вариант поведения управляющих систем состоит действительно в получении информации как перестройки или, скорее, реконфигурации внутренней подсистемы состояний, рассматриваемой на практике как система хранения данных. С формализованной стороны, в первом варианте множества прообраза и образа отображения $\mathcal{X} \rightarrow \mathcal{Y}$ сохраняются постоянными, как это имеет место в традиционной математике. Во втором варианте происходит динамическое изменение указанных множеств, чему в классической основе математики теории множеств нет прямых аналогов. По существу, оказывается, что полноценное получение информации выводит научную тематику за традиционные границы математики, что делает понятным, почему хорошее понимание информации и действий с ней до сих пор оказывались мало формализованными и находились вне внимания большинства ученых.

Заключение. Показано, что современные информационные технические системы используют отображение интерпретации информации с переменным составом прообраза и образа отображения. Информационный аспект сложных систем управления состоит в непостоянстве базовых множеств интерпретации, а получение и обработка информации включает как основополагающий момент расшире-

ние или сужение этих множеств. Информационные связи и соответствия, в отличие от математических отображений и связей, принципиально отличаются от вторых динамическим управляемым характером изменения таких множеств.

Библиографический список

1. Шеннон К. Математическая теория связи. Работы по теории информации и кибернетике. М.: Изд-во ИЛ, 1963. С. 243–322.
2. Винер Н. Кибернетика, или Управление и связь в животном и машине. М.: Советское радио, 1968. 326 с.
3. Флоренсов А. Н. Асинхронные процессы взаимодействия как основа понятия информации // Омский научный вестник. 2017. № 1 (151). С. 131–134.
4. Бауэр Ф., Гооз Г. Информатика. М.: Мир, 1976. 484 с.
5. Мюллер В. К. Англо-русский словарь. Изд. 23-е, стер. М.: Русский язык. 1990. 848 с. ISBN 5-200-01309-7.
6. Клеменс Б. Язык C в XXI веке. М.: ДМК Пресс, 2015. 376 с. ISBN 978-5-97060-101-3, 978-1-491-90389-6.
7. Можейко И. В. Западный ветер — ясная погода. М.: АСТ, 2001. 541 с. ISBN 5-17-005-862-4.
8. Баад Т. Объектно-ориентированное программирование в действии. СПб.: Питер, 1997. 464 с. ISBN 5-88782-270-8, 0-201-82419-1.
9. Куратовский К., Мостовский А. Теория множеств. М.: Мир, 1970. 416 с.
10. Буч Г. Объектно-ориентированный анализ и проектирование с примерами приложений. 3-е изд. М.: Вильямс, 2008. 720 с.

ФЛОРЕНЦОВ Александр Николаевич, кандидат технических наук, доцент (Россия), доцент кафедры «Информатика и вычислительная техника».
SPIN-код: 1116-8357
AuthorID (РИНЦ): 661799
ORCID: 0000-0001-6466-5202
Адрес для переписки: florensov@yandex.ru

Для цитирования

Флоренсов А. Н. О расширении понятия информации современными средствами информатики // Омский научный вестник. 2018. № 6 (162). С. 235–239. DOI: 10.25206/1813-8225-2018-162-235-239.

Статья поступила в редакцию 18.09.2018 г.

© А. Н. Флоренцов