

Информатика



Радченко Надежда Петровна
Учитель информатики и ИКТ
ГБОУ СОШ №311 г. Москвы.

Структурирование данных – путь от информации к модели

Весь когда-то существовавший мир и мир, окружающий нас сейчас, и мы в нём, а также мир, который сложится в будущем – всё это система, и всё – объект нашего познания. Как облегчить путь познания и как распорядиться полученной информацией – проблема, решаемая с помощью моделирования, всё чаще информационного. Модель в определённых условиях заменяет объект-оригинал, так как обладает необходимым набором характеристик и некоторыми преимуществами по сравнению с объектом-оригиналом: наглядностью и доступностью для исследования. При этом любая модель является только приблизительной копией объекта-оригинала. Модель всегда проще, так как при её разработке учитываются свойства объекта, существенные только для конкретных целей моделирования.

Информационное моделирование – один из методов познания. В процессе моделирования описываются внешний вид, связи, поведение объекта, для установления причинно-следственных связей рисуется схема, составляется упорядоченный перечень планируемых действий. Разработка, изучение и использование информационных моделей характерны для всех отраслей знаний.

В отличие от материальных моделей информационные модели потрогать нельзя, хотя они имеют свой материальный носитель и представлены

в различных биологических, физических или химических формах. Это обусловлено тем, что основу нематериальных моделей составляет информация, характеризующая объект-оригинал, а форма представления информации может быть различной, как различна природа символов, кодирующих информацию.

Информационная модель – это информация об объекте, которая описывает его с помощью выбранного способа кодирования, отражая существенные для целей моделирования сведения об объекте.

Объекты и их свойства

Выясним, какую информацию об объекте нужно считать свойством, которое следует учитывать при построении модели.

1) Любой объект, состоящий из какого-либо количества упорядоченных взаимосвязанных частей, существующих как единое целое, можно рассматривать как систему. Составные части системы называют элементами системы. Каждый элемент имеет своё назначение, своё функциональное место в структуре системы. Именно это означает упорядоченность элементов в системе в отличие от хаоса беспорядочного множества (сравните груды кирпичей и кирпичную кладку; на части разобранный велосипед и велосипед, находящийся в рабочем состоянии).

В свою очередь, системы могут вступать в различные отношения с другими системами и взаимодействовать. Приведём примеры.

- Семья – это, с одной стороны, самостоятельно функционирующая система. С другой стороны, параметры её жизнедеятельности взаимосвязаны с параметрами другой общественной системы – государственной.

- Завод – сложная техническая система со своими параметрами и функциями, состоящая из множества других технических подсистем. В то же время завод является элементом социотехнической системы, поскольку его функционирование взаимосвязано с уровнем доходов государства в целом и семей его работников в частности. Кроме того, любой завод включён в некоторую биотехническую систему, так как непосредственно влияет на экологическую обстановку как в региональных масштабах, так и планетарных.

2) Любой объект характеризует набором *параметров* (свойств) и

действий (деятельных проявлений), которые позволяют выделить этот объект среди других. Объекту присваивают характеризующее его имя, отличающее его среди других объектов.

3) В результате сравнения объектов друг с другом можно установить, в каких *отношениях* друг к другу они находятся:

- пространственных (река Волга дальше от экватора, чем река Нил);
- временных (война с войсками Наполеона случилась раньше, чем первая мировая);
- можно в целом выделить часть (компьютер – это целое, а процессор – его часть);
- при рассмотрении взаимоотношений между людьми, чтобы их охарактеризовать, можно выделить общественные отношения.

Таким образом, отношения между объектами характеризуют эти объекты. Следовательно, отношения между объектами – это их свойства.

4) Каждый объект может быть охарактеризован тем *действием*, проявление которого изменяет состояние объекта или объектов, с которыми он находится в тех или иных обозначенных выше отношениях. Тогда говорят об установленных *связях* между объектами, то есть таких отношениях между ними, когда изменение параметров одного объекта приводит к изменению параметров другого. Связи между объектами различны по типу, по способу описания – их можно описать словами, представить формулами, отобразить схемой. Рассмотрим примеры.

- Если социологические опросы, проведённые после мер, направленных на понижение уровня корруп-

ции в стране, указывают на повышение авторитета властных структур, то это означает наличие обратной связи между объектами «властные структуры» и «электорат».

• Если за рулём разбитого автомобиля сидит пьяный человек, то, анализируя связь между ними, можно утверждать, что этот человек был управляющей системой, а разбитый автомобиль – управляемой. Заметим, что, анализируя произошедшую аварию, следовательно выяснит все связи и отношения в системе водитель – автомобиль: скорость движения автомобиля, состояние дорожного покрытия, наличие помех на пути следования, техническое состояние

автомобиля до аварии. Но существенным в процессе следствия при моделировании состояния системы водитель – автомобиль, приведшего к происшествию, может оказаться указанная выше связь между водителем и автомобилем (действительно, не автомобиль же управлял человеком) и физиологическое состояние водителя как параметр объекта.

Модель является только приближенной копией объекта-оригинала: не все реальные свойства объекта могут быть выявлены, и тем более не все свойства объекта могут быть известны субъекту моделирования к моменту создания модели.

Формы представления информационных моделей

Информационные модели могут быть представлены разными способами. В таблице 1 приведена одна из

возможных структуризаций форм представления информационных моделей.

Таблица 1

Формы представления информационных моделей	Образно-знаковые			Знаковые		
	Графические	Структурные	Алгоритмические	Описание на естественном языке	Описание на специальных языках отраслей знаний	
	Чертеж, фотография, голограмма	Диаграмма, схема, таблица	Блок-схема, нумерованный список	Последовательность предложений в виде устной или письменной речи	Математические формулы, уравнения	Химические формулы, нотные записи, компьютерные программы, ...
Отражаемые в модели характеристики объекта	Внешний вид	Структура и связи параметров	Последовательность действий	Любые характеристики (но с разной степенью успешности отражения)	Количественные	Специфические для отрасли знаний

Из таблицы понятно, что форма представления информацион-

ной модели зависит от того, какие характеристики объекта являются

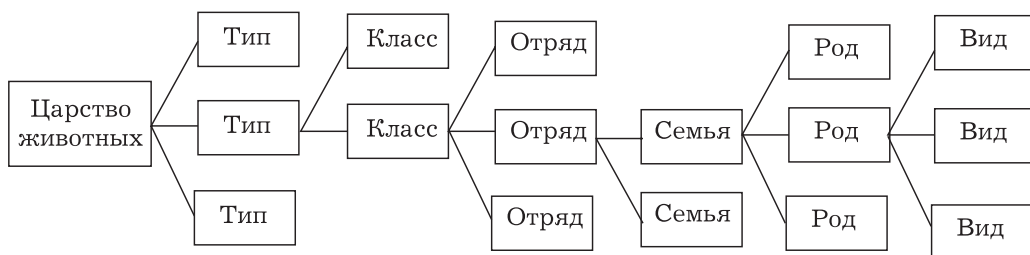
существенными для целей моделирования.

Информационные модели, представленные в образно-знаковых формах, доступны зрительному восприятию и могут быть зафиксированы на таких материальных носителях, как бумага, холст, фотоплёнка и т. п. Информационные модели при знаковой форме представления допускают использование языков, алфавиты которых могут состоять, например, из символов-звуков, электрических сигналов, то есть символов, не написанных на бумаге. Поэтому про знаковые информационные модели нельзя утверждать, что они всегда доступны зрительному восприятию.

Отметим также, что для описания некоторых объектов могут использоваться одновременно несколько различных языков, например, языки математических соотношений и химических (физических) формул. Это позволяет создавать информаци-

онные модели, отображающие характеристики и связи различных параметров объекта-оригинала в виде математических формул, описывать химические (физические) процессы и представлять изменения характеристик объекта-оригинала в течение заданного интервала времени под воздействием меняющихся условий. Последующее описание таких математических моделей на языках программирования позволяет проводить компьютерные исследования объектов-оригиналов – компьютерные вычислительные эксперименты, имитационное моделирование.

Некоторые модели создаются только для того, чтобы фиксировать состояние объекта-оригинала в заданный момент времени. Такие модели называют *статическими* информационными моделями. Именно статическую информационную модель животного мира образует его классификация по родам, видам, классам:

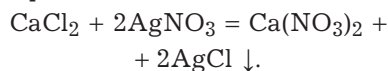


Статической моделью вещества является его структурная формула, например, аминифенола



В том случае, когда информационная модель отражает изменения в состоянии объекта с течением времени, её называют *динамической*.

Уравнение, описывающее химическую реакцию, – динамическая модель, так как содержит информацию о процессе изменения во времени некоторых веществ под воздействием определённых факторов, в частности, образования осадка хлорида серебра:



К числу динамических относится модель прогнозирования изменений популяции животных в заданных условиях.

Информационное моделирование в табличной форме

Информацию, целенаправленно отобранную для построения модели, принято называть «данными». Одним из путей удобного пользования данными является их структуризация. Действительно, вот пример неструктурированных данных:

Иванов Иван Иванович проживает в 1-ой квартире; в квартире №2 живёт Пётр Петрович Петров, а в кв. 3 – Сидоров Сидор Сидорович.

При неструктурированной форме хранения сложно организовать поиск нужных данных, а упорядочить подобную информацию практически нереально.

Структурирование данных осуществляется в соответствии со связями, которые были выявлены между элементами системы. В структурированном виде та же информация может быть упорядочена по какому-то признаку, и тогда среди всей информации можно быстро отыскать необходимую:

Таблица 2

№ кв.	Фамилия	Имя	Отчество
1	Иванов	Иван	Иванович
2	Петров	Пётр	Петрович
3	Сидоров	Сидор	Сидорович

В информационном моделировании широкое применение получили три способа организации данных: *табличный, иерархический и сетевой*. В соответствии с таблицей 1 они позволяют представить информационные модели в образно-знаковой или знаковой форме. В данной статье мы остановимся на особенностях табличного представления данных.

Это наиболее привычный из перечисленных выше способ формализованного описания данных о разно-

го рода объектах. Например, табличную форму имеет одна из информационных моделей организации школьного учебного процесса – всем знакомое расписание уроков.

Данные, размещённые в ячейках таблицы, не только наглядны для восприятия. Табличная форма представления данных облегчает их анализ и обработку. Таблице можно дать название, приписать ей номер. Очевидно, что в каждой из таблиц может быть разное количество столбцов и строк, поименованных в соответствии с целями моделирования. В качестве основных структурных элементов рассматривают *строки* таблицы (обычно называемые «записями»), *столбцы* («поля») и *реквизиты* («атрибуты») – конкретные значения, находящиеся в ячейках таблицы.

Чтобы структурировать данные в виде таблицы, нужно чётко представлять цель такого моделирования. Поле (столбец) таблицы должно содержать данные, одинаковые по типу, но в разных полях типы данных, естественно, могут не совпадать. Таким образом, запись (строка) состоит из ячеек одного уровня, но разных полей. Следовательно, она может содержать реквизиты, отличающиеся по типу, но характеризующие обычно один объект.

Проводя *системный анализ*, нужно выделить элементы системы и проанализировать применительно к ним (как объектам), какие из известных свойств и взаимоотношений существенны для целей моделирования. Затем выяснить, какие объекты можно объединить в одно множество (или несколько множеств). На основании этого становится ясным, необходимы ли в таблице различные уровни заголовков в строках (ступени). При определении общего количества столбцов, поряд-

ка их расположения, наименования и типов данных, помещаемых в них, необходимо учесть, что некоторые данные одного типа и наименования могут подразделяться на подмножества одинакового назначения для

каждой записи. Если такое случается, то при формировании столбцов появятся ярусы – разные уровни заголовков. Ниже в общем виде представлена многоярусная и многоступенчатая таблица 3.

Таблица 3

		Заголовок 1 столбца 1-го яруса				Заголовок 2 столбца 1-го яруса
		Подзаголовок_1 Ярус_2	Подзаголовок_2 Ярус_2	Подзаголовок_3 Ярус_2	Подзаголовок_4 Ярус_2	
Заголовок строки 1, ступень 1	Заголовок_1 Ступень_2					
	Заголовок_2 Ступень_2					
	Заголовок_3 Ступень_2					

После того, как структура таблицы определена, можно переходить к заполнению её ячеек данными (реквизитами). Отметим, что данные, структурированные в таблицу, могут образовывать как статические, так и динамические информационные модели. Так, например, табель успеваемости ученика с полугодовыми итоговыми оценками представляет собой статическую информационную модель. С динамической информационной моделью мы сталкиваемся всякий раз, когда читаем школьный дневник любого ученика.

Помимо различий в учёте временного фактора, таблицы могут иметь ещё ряд различающих их характеристик.

1) **Таблицы типа «объект-свойство»** могут отражать информацию об объектах и их свойствах. Такая таблица содержит данные о свойствах объектов, входящих в одно множество, причём объектов и свойств, их характеризующих, должно быть хотя бы по одному.

Примером такой таблицы является таблица 4 «Мои друзья», кото-

рую каждый может составить, чтобы вовремя поздравить своих близких с днём рождения:

Таблица 4. Мои друзья

№	Имя	День рождения	Телефон
1	Аня	05.01	111-11-11
2	Света	06.02	222-22-22
3	Петя	07.03	333-33-33
4	Игорь	08.04	444-44-44

Эта таблица образует статическую информационную модель множества друзей составителя таблицы. В самом деле, как любая модель, таблица отражает не всю информацию о друзьях, а только целенаправленно отобранную в соответствии с целями моделирования. А поскольку содержащиеся в таблице данные не зависят от временного фактора, то эта модель статическая. Проверим, является ли таблица 4 таблицей типа «объект-свойство». Она содержит информацию о четырёх объектах, при этом для каждого из них приводятся три свойства,

характеризующие объекты. Следовательно, определяющая такие типы таблиц характеристика (хотя бы один объект и хотя бы одно его свойство) соблюдена.

2) **Таблица типа «объект-объект».** В такой таблице можно отобразить *взаимосвязи между объектами*, причём объектов должно быть не меньше двух.

Таблица 5. Результаты соревнований (время в секундах)

Вид Участник	Бег 60 м	Бег 100 м	Бег 400 м с барьерами	Бег 800 м
Иванов	12	15	60	120
Петров	30	50	0	0
Сидоров	10	12	55	110

Здесь первое множество объектов – участники соревнований, второе множество – виды спорта, в соревнованиях по которым приняли участие наши спортсмены. Взаимосвязь между каждым из объектов 1-го множества и каждым из объектов 2-го множества охарактеризована всего одним свойством, значения которого вписаны в соответствующие ячейки таблицы. Это время (измеренное в секундах), которое показал каждый из участников соревнования.

3) Таблицы более сложного вида, например, **«объекты-свойства-объекты»**, получаются путём комбинирования двух описанных выше ти-

пов таблиц.

В таблицах всех указанных выше типов присутствуют свойства объектов, устанавливающие отношения между реквизитами и объектами. Поэтому информационные модели с табличным структурированием данных ещё называют реляционными (от англ. *relation* – «отношение»).

4) **Двоичные матрицы.** Это одна из разновидностей таблиц типа «объект-объект». В отличие от примера таблицы 5, таблица 6 хотя и является таблицей именно такого типа, но представленная в ней информация образует динамическую информационную модель.

Таблица 6. Наличие рыбы в Медвежьих озёрах

Вид рыбы \ День обследования	01.07.2000	01.07.2005	01.07.2011
Щука	1	1	1
Сом	0	0	0
Карп	1	0	0
Ёрш	1	1	1
Окунь	1	0	0

Кроме того, таблица представляет модель, разработанную с целью установить наличие или отсутствие связей между объектами. Поэтому в качестве значения реквизита, фик-

сирующего наличие или отсутствие связи, в соответствующих ячейках таблицы записаны «1» или «0». При этом каждый из объектов одного множества может иметь связь с



более чем одним объектом другого множества. Главное, что характер связи между объектами не рассматривается, а получаемые числовые таблицы (матрицы, более того, двоичные матрицы, поскольку состоят из символов двоичного алфавита) поддаются компьютерной обработке

Программные средства создания и обработки табличных структур данных

В качестве инструмента для создания таблиц любого вида могут использоваться такие прикладные компьютерные программы, как текстовые редакторы (процессоры), допустим, Microsoft Word, или электронные таблицы, например, Microsoft Excel. Только надо понимать, что задача табличного структурирования данных под силу человеку, а не компьютеру, поскольку это – задача создания информационной модели.

Для табличной организации данных применяют системы управления базами данных (СУБД) – прикладные компьютерные программы, позволяющие создавать табличные (реляционные) базы данных. Примерами СУБД реляционного типа являются Microsoft Access и СУБД из интегрированного пакета Microsoft Works.

Надо понимать, что даже в простых случаях построение информационных моделей в табличной фор-

и анализу.

Обобщая всё вышесказанное, сформулируем последовательность этапов приведения информационной модели к табличной форме:

- 1) анализ имеющейся информации для выделения объектов;
- 2) выявление и анализ свойств объектов и/или отношений между ними;
- 3) определение возможности объединения объектов и их характеристик в подмножества с целью выявления количества уровней и ступеней будущей таблицы;
- 4) определение общего количества столбцов и порядка их следования;
- 5) определение наименований столбцов и типов данных в них;
- 6) определение порядка строк;
- 7) занесение реквизитов в ячейки таблицы.

ме – вовсе не обычное заполнение реквизитами ячеек готовой табличной формы, которое является только последним этапом при табличном структурировании данных.

Отметим, что в случае исследования сложных биологических, технических, социальных систем, модели которых могут быть представлены в табличных формах, в процессе компьютерного моделирования зачастую приходится прибегать к разработке уникальных программных комплексов, реализующих обработку матриц с использованием методов высшей математики.

Следует также иметь в виду, что отражение в таблице всех существенных для целей моделирования свойств, получение информационной модели, адекватной моделируемому объекту и позволяющей решить относительно объекта оригинала поставленную задачу, чаще всего оказывается трудноразрешимой задачей.