

Наследование свойств категории при переходе от статических систем, использующих знания, к динамическим¹

Аннотация. Ранее авторами была показана целесообразность привлечения аппарата теории категорий для абстрактного описания знаний в системах продукционного типа. Первоначально этот аппарат предназначался для описания обычных, т.е. статических экспертных систем, получивших самое широкое распространение в искусственном интеллекте и смежных областях. При этом несколько в стороне оказались так называемые динамические экспертные системы, допускающие изменение данных и знаний в процессе своего функционирования. Поскольку в предложенных нами ранее схемах динамическая экспертная система строилась на основе статической, то естественно возник вопрос о распространении теоретико-категорного аппарата и на такие системы. Для этого потребовалось определенное обобщение развитого прежде теоретико-категорного аппарата, связанное с введением так называемых расслоенных категорий и соответствующих контекстных решеток. Совокупность возникающих при этом проблем подробно рассматривается в настоящей работе.

Ключевые слова: динамическая экспертная система, теоретико-категорная продукция, ТК-продукция, расслоенная категория, контекстная решетка

Введение

Теоретическое исследование в области систем, основанных на знаниях, или экспертных систем, является затруднительным, поскольку на практике сталкиваются как с большим разнообразием областей их применения, так и разнообразием используемых средств их программной реализации.

Для описания и исследования систем авторы настоящей работы предложили достаточно абстрактный формальный язык [1-3], позволяющий с единых позиций рассматривать некоторые моменты, являющиеся общими для большинства подобных систем. Этот язык основан на формализме теории категорий [4]. В частности, в рамках этого языка были описаны интеллектуальные компьютерные системы, основанные на правилах, или продукционные системы. Характерно, что продукция при этом понималась весьма широко, она рассматрива-

лась как форма записи базовой элементарной операции, которую использует в своей работе интеллектуальная система, основанная на знаниях. Эта операция состоит в том, чтобы узнать ситуацию, с которой столкнулась система в своей работе, найти в памяти описание этой ситуации и преобразовать ситуацию в соответствии с найденными в памяти инструкциями, в результате чего возникает новая ситуация. Подобные продукции были представлены авторами на теоретико-категорном языке и получили название ТК-продукций [5-8].

В основу теоретических построений в этих работах были положены *ситуации*, *образцы* (обобщенные описания ситуаций) и *операция сопоставления*. Следует заметить, что для каждой предметной области эти три понятия приобретают, вообще говоря, специфический конкретный вид. Введенные понятия позволяют построить категорию, которая описывает все возможные действия со знаниями.

¹ Работа частично финансировалась РФФИ по проектам 07-07-0039, 05-01-00382 и по программе №14 Президиума РАН.

Вторым направлением исследования, послужившим отправной точкой для настоящей статьи, являются проводимые авторами в течение многих лет исследования так называемых динамических экспертных систем, впервые введенных в работе [9]. В процессе работы таких систем допускается пересмотр данных, полученных ранее, что может приводить к отказу от сделанных ранее логических выводов и замене их другими. В частности, в [9] была предложена квазистатическая реализация динамической экспертной системы, при которой производился «перезапуск» статической экспертной системы, позволяя пересматривать введенные ранее данные и вовлекаемые при этом знания.

Несколько позже авторы [10] предложили свою архитектуру для реализации динамики в одной конкретной прикладной задаче, впрочем, также отталкиваясь от некой статической экспертной системы.

С другой стороны, соображения оптимизации работы динамической системы привели к тому, что используемые в [9] данные были разбиты на две группы – статические и динамические данные.

В течение некоторого времени эти две линии нашего исследования – разработка достаточно универсального языка описания интеллектуальных систем и описание работы динамических экспертных систем – не пересекались. Это было связано с концентрацией исследования на том, как именно предложенный язык ТК-продукций использовался для описания самой продукции (правила) и того, как она преобразует ситуацию. Однако выяснилось, что многие свойства динамических систем вполне могут быть описаны на языке ТК-продукций, поскольку предложенная в [9] динамическая экспертная система фактически отличается от обычной (статической) лишь определенными особенностями использования алгоритма логического вывода на продукциях [11].

Дело в том, что как бы ни подходить к обеспечению динамики, т.е. изменению данных и знаний в экспертных системах в процессе их функционирования, общим во всех подходах будет то, что с фактами, установленными системой, должна быть каким-то образом связана информация о том, при каких условиях истинность полученных фактов может быть подверг-

нута сомнению. Сходная мысль была высказана еще М. Минским, сформулировавшим так называемую «проблему фрейма» [12]. (Проблема фрейма в случае динамической экспертной системы обсуждается в работе [13].)

Поскольку на языке ТК-продукций факты задаются определенными морфизмами, возникла идея связать подобную информацию с морфизмами категории. В результате было показано, что каждый морфизм категории [4] в динамических системах «расслаивается», т.е. заменяется на конечное множество морфизмов, причем на каждом этапе логического вывода используется один и только один элемент этого множества, а именно тот морфизм, для которого определенные условия оказываются выполненными. Сама категория при этом оказывается состоящей из отдельных слоев, с каждым из которых связываются некоторые условия. А слой представляет собой подкатегорию исходной категории, в которую включаются морфизмы, существование которых допускается этими условиями. При изменении условий мы переходим от слоя к слою, при этом некоторые морфизмы исключаются из рассмотрения, следовательно, описываемые ими факты перестают быть истинными и нуждаются в проверке, предложенной в [9].

Такой расслоенной категории и посвящена настоящая работа. Основное внимание в ней уделено вопросу о том, как те или иные свойства категории, построенной для описания некоторой статической продукционной системы, наследуются расслоенной категорией, построенной для описания соответствующей динамической системы.

1. Язык ТК-продукций

Язык ТК-продукций неоднократно был описан нами в цитированных выше работах, поэтому здесь мы дадим лишь краткое изложение его основных понятий, призванное облегчить понимание дальнейшего материала.

В основе наших построений лежит понятие *ситуации*, описывающее те знания о решаемой задаче, которые известны системе к некоторому моменту в процессе рассуждений. Ситуация меняется с получением новой информации. Для описания множества сходных ситуаций, в которых следует предпринимать одни и те же дей-

ствия, невзирая на то, что они различаются в деталях, служит понятие *образца* – обобщенного описания ситуации. Основной операцией в нашей теории является *операция сопоставления* ситуации с образцом, дающая ответ на вопрос, подходит ли ситуация под образец, а также выдающая некоторую побочную информацию, соответствующую особенностям данной ситуации. Продукция понимается как пара образцов, один из которых описывает ситуацию, в которой продукцию можно применять, другой – ситуацию, возникающую после ее применения (точнее, некоторое множество ситуаций, определяемых образцом).

На языке ТК-продукций задание системы образцов состоит в задании категории, среди морфизмов которой выделяется некоторый класс морфизмов, называемых ситуациями. При этом, если морфизмы α , β и φ связаны соотношением $\alpha = \beta\varphi$, морфизм α является ситуацией в том и только в том случае, когда ситуацией является морфизм β . Под образцом понимается произвольный морфизм $\varphi: X \rightarrow S$ категории. Ситуация $\alpha: I \rightarrow S$ считается сопоставимой с образцом $\varphi: X \rightarrow S$, если существует морфизм $\beta: I \rightarrow X$, удовлетворяющий условию $\alpha = \beta\varphi$. Продукцией из S в T , где S и T – объекты категории, называется пара образцов $\varphi: X \rightarrow S$, $\psi: X \rightarrow T$, т.е. продукция задается парой морфизмов. Продукция считается применимой к ситуации $\alpha: I \rightarrow S$, если для некоторого морфизма $\beta: I \rightarrow X$ имеем $\alpha = \beta\varphi$. Результатом применения продукции к ситуации α будет морфизм $\beta\psi: I \rightarrow T$.

Для того, чтобы включить конкретную задачу, решаемую интеллектуальной системой, в эту схему, следует определить, какого рода ситуации могут возникать при решении этой задачи, какие образцы разумно выбрать для описания множеств сходных ситуаций и как должна быть устроена операция сопоставления. После этого может быть построена категория, в рамках которой будут описаны такие ситуации и образцы. Теперь база знаний может быть представлена в виде некоторого набора морфизмов, задающих продукции, а текущее состояние вывода определяется морфизмом-ситуацией.

2. Расслоенная категория

Как уже отмечалось, в динамической системе истинность факта не является абсолютной, она зависит от контекста и при определенных условиях может быть пересмотрена. Для формализации понятия «контекст» введём в рассмотрение алгебраическую решетку P , обладающую наименьшим элементом 0 , которую в дальнейшем будем называть контекстной решеткой. С каждым фактом f мы связываем некоторый контекст $p \in P$, в котором этот факт верен. Если текущее состояние системы описывается контекстом $q \in P$, и $p \leq q$, будем считать, что факт f является истинным, если же неравенство $p \leq q$ не выполнено, истинность факта f следует перепроверить заново. В этом смысл частичного порядка на множестве P .

Пусть теперь C – произвольная категория, а P – контекстная решетка.

Определение 1. Категорией C_P , расслоенной с помощью контекстной решетки P , будем называть категорию, построенную следующим образом. Объектами категории C_P объявим объекты категории C . Пусть X, Y – объекты категории C (следовательно, и C_P). Определим множество морфизмов из X в Y в категории C_P формулой $C_P(X, Y) = C(X, Y) \times P$. Пусть $(\varphi, p) \in C_P(X, Y)$, $(\psi, q) \in C_P(Y, Z)$. Определим морфизм $(\varphi, p) \circ (\psi, q) \in C_P(X, Z)$ формулой $(\varphi, p) \circ (\psi, q) = (\varphi \circ \psi, p \vee q)$, где символ $\vee$ обозначает верхнюю грань.

Таким образом, каждый морфизм категории C_P включает ссылку на контекст, в котором этот морфизм можно использовать. Композиция двух морфизмов определена в таком контексте, в котором определен каждый из них.

Корректность введенных определений мы здесь проверять не будем, ибо она станет очевидной в конце параграфа, где будет дано эквивалентное определение расслоенной категории. Отметим лишь, что роль единичного морфизма объекта X категории C_P играет морфизм $(1_X, 0)$, где символом 1_X обозначен единичный морфизм объекта X категории C ,

а символом 0 – наименьший элемент контекстной решетки.

Приведенная процедура позволяет по произвольной категории $\mathbf{C}$ и контекстной решетке P строить новую категорию, расслоенную с помощью контекстной решетки P . Каждому контексту соответствует слой категории со своими морфизмами. Слои упорядочены (частично), и то, что присутствует в более низком слое, сохраняется в более высоком.

Известно, что каждую полугруппу с 1 можно рассматривать как категорию. Делается это следующим образом. Пусть M – полугруппа с 1 . Построим категорию с единственным объектом, взяв в качестве множества эндоморфизмов этого объекта саму полугруппу M . Такая конструкция удовлетворяет всем требованиям определения категории.

Поскольку решетка с наименьшим элементом образует коммутативную полугруппу с 1 относительно операции $\vee$ (роль единицы играет наименьший элемент), контекстную решетку P можно рассматривать как категорию P с единственным объектом, множеством эндоморфизмов которого является само множество P . Роль композиции играет операция $\vee$, а роль единичного морфизма – наименьший элемент 0 множества P . Теперь расслоенной категории можно дать другое определение, эквивалентное приведенному ранее.

Определение 2. Пусть $\mathbf{C}$ – категория, P – решетка, обладающая наименьшим элементом. Расслоением категории $\mathbf{C}$ с помощью решетки P назовем категорию $\mathbf{C}_P = \mathbf{C} \times P$ – произведение категории $\mathbf{C}$ и построенной выше однообъектной категории P .

Определение 2 является более удобным в теоретическом плане. Оно не требует проверки того, что построенная структура является категорией, ибо она является произведением категорий. Кроме того, оно облегчает применение теоретико-категорной техники для определения свойств $\mathbf{C}_P$.

Для получения содержательных результатов, относящихся к некоторой категории, обычно накладывают на нее ряд дополнительных условий, для построения в ней определенных универсальных конструкций. Естественным в этом плане является следующий вопрос: при каких условиях, наложенных на контекстную решет-

ку, те или иные свойства переносятся с исходной категорией на расслоенную. Будем говорить, что контекстная решетка P сохраняет некоторое свойство категории, если всякий раз, когда категория $\mathbf{C}$ обладает этим свойством, свойство наследует и категория $\mathbf{C}_P$. В последующих разделах мы рассмотрим вопрос о сохранении контекстной решеткой некоторых свойств категории.

3. Пределы в расслоенной категории

Использование языка теории категорий в той или иной задаче будет плодотворным, если конструкции, построенные в этой теории, будут соответствовать определенным реалиям решаемой задачи. Проведенные авторами исследования показывают, что многие построения, возникающие в теории ТК-продукций, естественно переводятся на теоретико-категорный язык. Так, в работе [8] при построении продукционных деревьев активно использовалось понятие прямого произведения. В [3] при исследовании свойств систем продукций были использованы такие понятия, как уравниватели, образующие, расщепляемость морфизмов и некоторые другие. Стало быть, для того, чтобы описанные в этих работах теоретические построения были бы осуществимы, необходимо, чтобы категория допускала построение соответствующих конструкций. Отсюда естественный интерес к вопросу о том, осуществимы ли те или иные построения в расслоенной категории, если они осуществимы в исходной.

Начнем с изучения вопроса, можно ли из существования некоторых пределов в категории $\mathbf{C}$ вывести существование этих пределов и в категории $\mathbf{C}_P$. В теории категорий существует большое разнообразие в определении понятия «предел». Мы будем пользоваться определением предела функтора, данное в [4].

Пусть $\mathbf{C}$ – произвольная категория, $\mathbf{J}$ – малая категория, F – функтор из $\mathbf{J}$ в $\mathbf{C}$. То, что задан предел функтора F , означает, что задан объект S категории $\mathbf{C}$ и для каждого объекта X категории $\mathbf{J}$ задан морфизм $\varphi_X : C \rightarrow F(X)$, причем для любого морфизма $\xi : X \rightarrow Y$ категории $\mathbf{J}$ имеет место равенство $\varphi_Y = \varphi_X F(\xi)$. Кроме того, предел обладает свойством универсально-

сти. Это означает, что для любого объекта C' категории $\mathbf{C}$ и для любой системы морфизмов $\varphi'_X : C' \rightarrow F(X)$ такой, что $\varphi'_Y = \varphi'_X F(\xi)$, где $\xi : X \rightarrow Y$ - произвольный морфизм категории $\mathbf{J}$, существует единственный морфизм $\varphi : C' \rightarrow C$, удовлетворяющий условиям $\varphi'_X = \varphi \varphi_X$ для всех объектов X категории $\mathbf{J}$.

Пусть заданы категории $\mathbf{C}$ и $\mathbf{D}$. Каждый функтор F из категории $\mathbf{J}$ в категорию $\mathbf{C} \times \mathbf{D}$ имеет вид $F = F_C \times F_D$, $F_C : \mathbf{J} \rightarrow \mathbf{C}$ и $F_D : \mathbf{J} \rightarrow \mathbf{D}$. Это означает, что для любого объекта X категории $\mathbf{J}$ выполнено равенство $F(X) = (F_C(X), F_D(X))$ и для любого морфизма $\varphi : X \rightarrow Y$ имеем $F(\varphi) = (F_C(\varphi), F_D(\varphi))$. Следующую лемму авторам не удалось найти в литературе (видимо, из-за ее элементарности), потому она приводится здесь вместе с доказательством.

Лемма 1. *Предел функтора $F = F_C \times F_D : \mathbf{J} \rightarrow \mathbf{C} \times \mathbf{D}$ существует, если существуют пределы функторов $F_C : \mathbf{J} \rightarrow \mathbf{C}$ и $F_D : \mathbf{J} \rightarrow \mathbf{D}$.*

Доказательство. Пусть определен предел функтора F_C . Это означает, что задан объект C категории $\mathbf{C}$ и для каждого объекта X категории $\mathbf{J}$ задан морфизм $\varphi_X : C \rightarrow F_C(X)$, причем для любого морфизма $\xi : X \rightarrow Y$ категории $\mathbf{J}$ имеет место равенство $\varphi_Y = \varphi_X F_C(\xi)$. Кроме того, для любого объекта C' категории $\mathbf{C}$ и для любой системы морфизмов $\varphi'_X : C' \rightarrow F_C(X)$ такой, что $\varphi'_Y = \varphi'_X F_C(\xi)$, где $\xi : X \rightarrow Y$ - произвольный морфизм категории $\mathbf{J}$, существует единственный морфизм $\varphi : C' \rightarrow C$, удовлетворяющий условиям $\varphi'_X = \varphi \varphi_X$ для всех объектов X категории $\mathbf{J}$. Аналогично существование предела функтора F_D означает существование объекта D категории $\mathbf{D}$ семейства морфизмов $\psi_X : D \rightarrow F_D(X)$ таких, что для любого морфизма $\xi : X \rightarrow Y$ категории $\mathbf{J}$ выполнено равенство $\psi_Y = \psi_X F_D(\xi)$. При этом, если D' - объект категории $\mathbf{D}$ и $\psi'_X : D' \rightarrow F_D(X)$ - морфизмы, удовлетворяющие равенствам $\psi'_Y = \psi'_X F_D(\xi)$ для любого морфизма $\xi : X \rightarrow Y$ категории $\mathbf{J}$, то существует единственный морфизм

$\psi : D' \rightarrow D$, удовлетворяющий условиям $\psi'_X = \psi \psi_X$ для всех объектов X категории $\mathbf{J}$.

Тогда объект (C, D) категории $\mathbf{C} \times \mathbf{D}$ и семейство морфизмов $(\varphi_X, \psi_X) : C \rightarrow (F_C(X), F_D(X))$ той же категории определяют предел функтора F . В самом деле, если $\xi : X \rightarrow Y$ - морфизм категории $\mathbf{J}$, то

$$\begin{aligned} (\varphi_Y, \psi_Y) &= (\varphi_X F_C(\xi), \psi_X F_D(\xi)) = \\ &= (\varphi_X, \psi_X) (F_C(\xi), F_D(\xi)) = (\varphi_X, \psi_X) F(\xi). \end{aligned}$$

Пусть, далее (C', D') - объект категории $\mathbf{C} \times \mathbf{D}$ и $(\varphi'_X, \psi'_X) : C' \rightarrow (F_C(X), F_D(X))$ - морфизмы этой категории, причем $(\varphi'_Y, \psi'_Y) = (\varphi'_X, \psi'_X) F(\xi)$ для любого морфизма $\xi : X \rightarrow Y$ категории $\mathbf{J}$. Последнее равенство означает, что $\varphi'_Y = \varphi'_X F_C(\xi)$ и $\psi'_Y = \psi'_X F_D(\xi)$. В силу универсального свойства предела функтора F_C существует морфизм $\varphi : C' \rightarrow C$, удовлетворяющий условиям $\varphi'_X = \varphi \varphi_X$ для всех объектов X категории $\mathbf{J}$. Аналогично, существует морфизм $\psi : D' \rightarrow D$, удовлетворяющий условиям $\psi'_X = \psi \psi_X$. Это означает, что для морфизма $(\varphi, \psi) : (C', D') \rightarrow (C, D)$ категории $\mathbf{C} \times \mathbf{D}$ имеет место равенство

$$(\varphi'_X, \psi'_X) = (\varphi \varphi_X, \psi \psi_X) = (\varphi, \psi) (\varphi_X, \psi_X).$$

Морфизм (φ, ψ) определяется однозначно, ибо однозначно определяются его компоненты φ и ψ . Доказательство леммы закончено.

Для полноты изложения сформулируем дуальную лемму.

Лемма 2. *Копредел функтора $F = F_C \times F_D : \mathbf{J} \rightarrow \mathbf{C} \times \mathbf{D}$ существует, если существуют копределы функторов $F_C : \mathbf{J} \rightarrow \mathbf{C}$ и $F_D : \mathbf{J} \rightarrow \mathbf{D}$.*

Доказательство ее аналогично, мы его не приводим и в остальной части статьи к копределам обращаться не будем.

Теперь мы можем исследовать следующий вопрос. Допустим, категория $\mathbf{C}$ обладает некоторыми свойствами, связанными с существованием в ней некоторых универсальных объектов. При каких условиях, наложенных на контекстную решетку P , это свойство переносится на категорию $\mathbf{C}_P$. Доказанные леммы позволяют утверждать, что существование некоторого

предела или копредела переносится из исходной категории в расслоенную, если такой предел (копредел) существует в контекстной решетке, рассматриваемой как категория.

Одним из наиболее простых и, в то же время, наиболее важных частных случаев предела является прямое произведение. Это понятие, являясь одним из наиболее фундаментальных понятий теории категорий, играет важную роль и в наших исследованиях. Попробуем применить лемму 1 для исследования вопроса о наследовании расслоенной категорией свойства замкнутости относительно прямых произведений. Именно, если категория C обладает конечными прямыми произведениями, то для того, чтобы таковыми обладала бы и категория C_P , достаточно, чтобы прямые произведения существовали бы в решетке P , рассматриваемой как категория. Поскольку P - категория с одним объектом, речь может идти только о произведении этого объекта самого на себя, при этом результатом будет тот же объект и пара проекций $p \in P$ и $q \in P$. Эти проекции должны удовлетворять следующему условию: если $x \in P$ и $y \in P$, то существует единственный элемент $z \in P$ такой, что $x = p \vee z$ и $y = q \vee z$. Из первого равенства следует, что $x \geq p$. Поскольку x - произвольный элемент множества P , можно сделать вывод, что p - наименьший элемент множества P . Аналогично заключаем, что и q - наименьший элемент множества P , откуда следует, что $p = q$. Из этого, в свою очередь, получаем $x = p \vee z = q \vee z = y$, но поскольку x и y - произвольные элементы множества P , мы видим, что это множество должно состоять из одного элемента.

Вывод – категория P не обладает прямыми произведениями за исключением, может быть, случая, когда решетка P одноэлементная. Следовательно, за исключением этого тривиального случая лемма 1 не применима. Легко убедиться, что дело не только в невозможности применить лемму, категория C_P не обладает прямыми произведениями, за исключением случая одноэлементной решетки P . Это означает, что для динамических продукционных систем придется переделывать определенные фрагменты теории ТК-продукций, заменяя

прямые произведения чем-то иным, более подходящим к динамическому случаю.

В других случаях, однако, вопрос о переносе свойств с категории C на категорию C_P может привести к более содержательным результатам. Рассмотрим, к примеру, вопрос о существовании в категории C_P уравнителей. Уравнители возникали при исследовании категорий, описывающих продукционные системы, в [3]. Уравнитель – частный случай предела, потому для существования уравнителей в категории C_P необходимо и достаточно их существование в категориях C и P .

Несложно видеть, что уравнителем пары элементов $p \in P$ и $q \in P$ является элемент $k \in P$ такой, что

- i. $k \vee p = k \vee q$;
- ii. если $m \in P$ и $m \vee p = m \vee q$, то существует единственный элемент $n \in P$ такой, что $m = k \vee n$.

Заметим, что для пары элементов k, m множества P условие существования элемента $n \in P$ такого, что $m = k \vee n$, эквивалентно условию $m \geq k$. Действительно, очевидно, что $k \vee n \geq k$, с другой стороны, если $m \geq k$, то $m = m \vee n$, так что в качестве k можно взять m . С другой стороны, условие существования единственного элемента $n \in P$ такого, что $m = k \vee n$, выглядит для решетки довольно экзотично. Так, для довольно «хорошей» решетки – решетки всех подмножеств некоторого множества, такое условие выполняется только, если $k = \emptyset$. Это говорит о том, что утверждать существование уравнителей нельзя для многих контекстных решеток, т.е. для этих решеток не применима лемма 1. И действительно, для многих решеток P существование уравнителей в категории C не дает существования уравнителей в категории C_P .

Прежде, чем двигаться дальше, отметим следующее важное обстоятельство. Конструкции теории категорий не всегда точно соответствуют построениям, полезным для исследования ТК-продукций. Вот наиболее характерный пример.

В наших работах мы называем образец $\varphi: X \rightarrow S$ частным случаем образца $\psi: Y \rightarrow S$, если существует морфизм $\sigma: X \rightarrow Y$, удовлетворяющий условию $\varphi = \sigma\psi$. В этом случае

всякая ситуация, сопоставимая с образцом $\varphi: X \rightarrow S$, сопоставима также с образцом $\psi: Y \rightarrow S$. Пусть $\varphi: X \rightarrow S$ и $\psi: Y \rightarrow S$ - два образца. Тот факт, что образец $\chi: Z \rightarrow S$ является частным случаем каждого из них, может быть выражен следующей коммутативной диаграммой:

$$\begin{array}{ccc} Z & \xrightarrow{\sigma} & X \\ \downarrow \tau & \searrow \chi & \downarrow \varphi \\ Y & \xrightarrow{\psi} & S \end{array}$$

В теории ТК-продукций определенную роль играло понятие наиболее общего частного случая. Определить его в терминах теории категорий можно так: образец $\chi: Z \rightarrow S$ является наиболее общим частным случаем образцов $\varphi: X \rightarrow S$ и $\psi: Y \rightarrow S$, если он является их общим частным случаем, т.е. имеет место выписанная выше коммутативная диаграмма, и для любого другого общего частного случая, т.е. образца $\chi': Z' \rightarrow S$, для которого имеет место коммутативная диаграмма

$$\begin{array}{ccc} Z' & \xrightarrow{\sigma'} & X \\ \downarrow \tau' & \searrow \chi' & \downarrow \varphi \\ Y & \xrightarrow{\psi} & S \end{array}$$

найдется морфизм $\rho: Z' \rightarrow Z$, для которого $\sigma' = \rho\sigma$ и $\tau' = \rho\tau$. Это определение очень близко к теоретико-категорному определению расслоенного произведения или декартова (коуниверсального) квадрата. Единственным, но очень существенным отличием является отсутствие требования единственности морфизма ρ . Это отличие не позволяет применить некоторые полученные в теории категорий результаты. В частности, невозможно доказать единственность наиболее общего частного случая с точностью до изоморфизма. Однако, отсутствие требования единственности является принципиальным моментом, ибо в большинстве интересных систем образцов она места не имеет.

Подобная ситуация и с другой важной в нашей теории конструкции – понятии наименьшего обобщения пары образцов. Обобщение двух образцов определяется определенной диаграммой, наименьшее же обобщение – это такое, из которого существует определенный морфизм в любое другое обобщение. Однако,

если потребовать единственность такого морфизма, мы получим определение, не приложимое ко многим интересным частным случаям.

В ряде других случаев стандартное определение предела, принятое в теории категорий, оказывается неприменимым. Для того, чтобы оно было бы применимо, нужно несколько ослабить формулировку определения предела, убрав из него единственность.

Поступим следующим образом. Будем говорить, что объект C категории $\mathbf{C}$ и семейство морфизмов $\varphi_X: C \rightarrow F(X)$ задает предел функтора F в ослабленной формулировке, если выполнены все сформулированные выше условия, кроме, может быть, требования единственности морфизма $\varphi: C' \rightarrow C$. Как уже говорилось, такое ослабление формулировки приводит к потере ряда свойств предела, в частности, невозможно доказать его единственность с точностью до изоморфизма. Тем не менее во многих категориях, возникающих при формализации систем образцов, связанных с конкретными интересными для практики задачами, можно утверждать существование лишь таких, ослабленных пределов.

Аналогичные рассуждения можно провести и для копределов. Мы не будем останавливаться на этом подробно, ибо в последующей части статьи копределы не рассматриваются.

Сформулируем теперь аналог леммы 1 для ослабленной формулировки определения предела.

Лемма 3. *Предел функтора $F = F_C \times F_D: \mathbf{J} \rightarrow \mathbf{C} \times \mathbf{D}$, понимаемый в описанном выше ослабленном смысле, существует, если существуют такие же ослабленные пределы функторов $F_C: \mathbf{J} \rightarrow \mathbf{C}$ и $F_D: \mathbf{J} \rightarrow \mathbf{D}$.*

Доказательство леммы представляет собой в точности доказательство леммы 1, если удалить из последнего завершающее его рассуждение о единственности.

Перейдем теперь от уравнивателя пары элементов к уравнителю, понимаемому в ослабленном смысле. Для этого случая из выписанного выше условия ii) следует убрать слово «единственный»:

ii') если $m \in P$ и $m \vee p = m \vee q$, то существует элемент $n \in P$ такой, что $m = k \vee n$.

Лемма 4. *Пара элементов $p \in P$ и $q \in P$ обладает уравнителем в ослабленном смысле тогда и только тогда, когда множество*

$\{m \in P : m \vee p = m \vee q\}$ имеет наименьший элемент.

Доказательство. Пусть $M = \{m \in P : m \vee p = m \vee q\}$. Если $k \in P$ - уравнитель p и q , то, в силу i, $k \in M$. Далее, если $m \in M$, т.е. $m \vee p = m \vee q$, то, в силу ii', $m = k \vee n$ для некоторого $n \in P$, откуда $m \geq k$. Мы видим, что k - наименьший элемент множества M .

Пусть, напротив, известно, что множество M обладает наименьшим элементом k . Поскольку $k \in M$, свойство i выполнено. Если $m \in P$ и $m \vee p = m \vee q$, то $m \in M$, потому $m \geq k$ и, следовательно, $m = k \vee m$, т.е. и свойство ii' выполнено. Лемма доказана.

Теорема 1. Если контекстная решетка является булевой алгеброй, то она сохраняет свойство существования уравнителей, понимаемых в ослабленном смысле.

Доказательство. В силу лемм 3 и 4 нам достаточно доказать, что для любой пары элементов $p \in P$ и $q \in P$ множество $M = \{m \in P : m \vee p = m \vee q\}$ имеет наименьший элемент. Покажем, что искомым наименьшим элементом является элемент $k = (p \wedge \bar{q}) \vee (\bar{p} \wedge q)$. Во-первых, $k \in M$. Действительно,

$$k \vee p = (p \wedge \bar{q}) \vee (\bar{p} \wedge q) \vee p = (\bar{p} \wedge q) \vee p = (\bar{p} \vee p) \wedge (q \vee p) = (q \vee p),$$

аналогично,

$$k \vee q = (q \vee p).$$

Пусть теперь $m \in M$, т.е. $m \vee p = m \vee q$. Тогда

$$m \vee (p \wedge \bar{q}) = (m \vee p) \wedge (m \vee \bar{q}) = (m \vee q) \wedge (m \vee \bar{q}) = m \vee (q \wedge \bar{q}) = m,$$

и, аналогично,

$$m \vee (\bar{p} \wedge q) = m,$$

поэтому

$$m \vee k = m \vee (p \wedge \bar{q}) \vee (\bar{p} \wedge q) = m,$$

т.е. $m \geq k$. Мы видим, что k , действительно, является наименьшим элементом множества M , что и доказывает лемму.

Булевы алгебры – весьма ограниченный класс решеток. Тем не менее, ряд практически важных контекстных решеток относится к этому классу. Типичным примером контекстной решетки является решетка, в которой каждый

контекст представляет собой некоторое множество условий. Контекстная решетка в этом случае представляет собой решетку всех подмножеств множества допустимых условий, являясь, следовательно, булевой алгеброй. Булевы алгебры часто возникают и в тех случаях, когда контекст представляет собой некоторое логическое выражение, а в роли верхней и нижней грани выступает конъюнкция и дизъюнкция.

От требования существования дополнения в решетке можно отказаться, если заменить дистрибутивную решетку вполне дистрибутивной. Напомним, что полная решетка называется вполне дистрибутивной, если в ней выполнены бесконечные тождества дистрибутивности. Нам, в частности, понадобится условие

Теорема 2. Вполне дистрибутивная решетка сохраняет свойство существования уравнителей, понимаемых в ослабленном смысле.

Доказательство. Нам снова достаточно доказать, что для любой пары элементов $p \in P$ и $q \in P$ множество $M = \{m \in P : m \vee p = m \vee q\}$ имеет наименьший элемент. Положим $k = \wedge \{m : m \in M\}$. Докажем, что $k \in M$. Имеем $k \vee p = (\wedge \{m : m \in M\}) \vee p = \wedge \{m \vee p : m \in M\} = \wedge \{m \vee q : m \in M\} = (\wedge \{m : m \in M\}) \vee q = k \vee q$.

То, что k - наибольший элемент этого множества, очевидно.

Отметим, что во многих задачах, важных с практической точки зрения, решетка P является конечной, так что полная дистрибутивность сводится к дистрибутивности.

4. Наиболее общий частный случай в расслоенной категории

Как уже отмечалось выше, в наших исследованиях важную роль играли наименьшее обобщение и наибольший частный случай пары образцов. Посмотрим, как переносятся эти понятия на расслоенную категорию. Ограничимся понятием наиболее общего частного случая, которое, как уже говорилось, представляет собой ослабленный вариант декартова квадрата. Это означает, что мы можем попытаться применить лемму 3. Для этого нужно посмотреть, существуют ли наиболее общие частные случаи в решетках, рассматриваемых как категории.

Рассмотрим категорию, полученную из контекстной решетки P . В этой категории построить элементы $p \in P$ и $q \in P$ до декартова квадрата, понимаемого в ослабленном смысле – это значит найти элементы $u \in P$ и $v \in P$ такие, что

$$i. \quad p \vee u = q \vee v,$$

ii. если $u \in P$, $v \in P$ таковы, что $p \vee u' = q \vee v'$, то для некоторого $w \in P$ имеем $u' = u \vee w$ и $v' = v \vee w$.

Изучим простой и, в то же время, важный пример – решетку подмножеств некоторого множества. В этом примере элементы p , q , u , v и другие суть подмножества, а знак $\vee$ означает теоретико-множественное объединение. Первое условие тогда записывается в виде $p \cup u = q \cup v$, откуда

$$\begin{aligned} u \cap (q - p) &= \\ &= (u \cap (q - p)) \cup \emptyset = (u \cap (q - p)) \cup (p \cap (q - p)) = \\ &= (u \cup p) \cap (q - p) = (v \cup q) \cap (q - p) = q - p, \end{aligned}$$

т.е. $u \supset q - p$. Аналогично доказывается, что $v \supset p - q$. С другой стороны, если положить $u' = q - p$, $v' = p - q$, то

$p \cup u' = p \cup (q - p) = p \cup q = q \cup (p - q) = q \cup v'$, откуда, согласно ii, $u' = u \vee w$ и $v' = v \vee w$, поэтому $u \subset u' = q - p$ и $v \subset v' = p - q$. Сопоставляя это с включениями, полученными ранее, находим, что $u = q - p$ и $v = p - q$. Условие ii теперь означает, что если $p \cup u' = q \cup v'$, то $u' = u \cup w$ и $v' = v \cup w$ для некоторого $w \in P$, поэтому

$$\begin{aligned} u' \cap p \cap q &= \\ &= (u \cup w) \cap p \cap q = (u \cap p \cap q) \cup (w \cap p \cap q) = \\ &= ((q - p) \cap p \cap q) \cup (w \cap p \cap q) = w \cap p \cap q \end{aligned}$$

и, аналогично,

$$v' \cap p \cap q = w \cap p \cap q,$$

откуда

$$u' \cap p \cap q = v' \cap p \cap q.$$

Между тем очевидно, что значение $p \cup u'$ не зависит от того, какие элементы множества $p \cap q$ входят в u' , и то же самое касается множества $q \cap v'$, поэтому равенство $p \cup u' = q \cup v'$ не может накладывать никаких ограничений на $u' \cap p \cap q$ и $v' \cap p \cap q$.

Вывод: в категории, полученной из решетки подмножеств, наиболее общие частные случаи могут не существовать. То же самое можно сказать о ряде других, важных для исследования динамических систем, контекстных решеток. Следовательно, к этим решеткам не применима лемма 3 и мы не можем утверждать, что они сохраняют свойство существования наиболее общих частных случаев, поэтому в соответствующих расслоенных категориях это свойство может не выполняться.

Было бы нежелательно отказаться в случае расслоенных категорий от понятия наибольшего частного случая и искать ему какую-то замену, подобно тому, как мы ранее предполагали поступить с понятием прямого произведения. Действительно, можно отказаться от той или иной теоретико-категорной конструкции, увидев, что она не находит применения в теории ТК-продукций, но не хотелось бы отказываться от конструкции, рожденной в рамках самой этой теории.

Дело в том, что приведенное выше определение наибольшего обобщения является слишком «сильным». Мы иногда пользовались им, ибо оно более точно соответствует конструкциям, используемым в теории категорий, являясь ослабленным вариантом декартова квадрата. Это означает, что оно является более удобным в тех случаях, когда оно применимо, а оно применимо в целом ряде практически полезных задач. Однако в предыдущих работах мы пользовались и другой формой определения наибольшего частного случая, являющейся более слабой и имеющей поэтому более широкую область применимости. Для того, чтобы перейти к этому определению, заметим, что если рассматривать множество образцов как множество, упорядоченное отношением «быть частным случаем», то наибольший общий частный случай есть ни что иное, как точная нижняя грань этих образцов. Иначе говоря, образец $\chi: Z \rightarrow S$ является наибольшим частным случаем пары образцов $\varphi: X \rightarrow S$ и $\psi: Y \rightarrow S$, если

- i. он является частным случаем этих образцов;
- ii. если образец $\chi': Z' \rightarrow S$ тоже является частным случаем образцов $\varphi: X \rightarrow S$ и $\psi: Y \rightarrow S$, то он является частным случаем и образца $\chi: Z \rightarrow S$.

На теоретико-категорном языке первое условие записывается так:

i. существуют морфизмы $\sigma: Z \rightarrow X$ и $\tau: Z \rightarrow Y$, для которых имеют место равенства $\sigma\varphi = \chi$ и $\tau\psi = \chi$, что эквивалентно условию i прежнего определения наибольшего частного случая.

Второе условие записывается так:

ii. если для некоторого образца $\chi': Z' \rightarrow S$ существуют морфизмы $\sigma': Z' \rightarrow X$ и $\tau': Z' \rightarrow Y$ такие, что выполнены равенства $\sigma'\varphi = \chi'$ и $\tau'\psi = \chi'$, то для некоторого $\rho: Z' \rightarrow Z$ имеем $\chi' = \rho\chi$.

Это условие отличается от условия ii предыдущего определения, ибо не требует выполнения равенств $\sigma' = \rho\sigma$ и $\tau' = \rho\tau$. Точнее, оно является более слабым, чем условие предыдущего определения, так как из равенств $\sigma' = \rho\sigma$ и $\tau' = \rho\tau$ (или хотя бы одного из них) следует

$$\chi' = \sigma'\varphi = \rho\sigma\varphi = \rho\chi$$

обратное же в общем случае не верно.

Заметим, что определенный таким образом наибольший частный случай пары образцов не является пределом (даже в ослабленном смысле), поэтому для его исследования мы не можем применить леммы 1 и 3. Сформулируем и докажем аналогичную лемму, относящуюся к случаю наибольшего частного случая.

Пусть заданы две категории $\mathbf{C}$ и $\mathbf{D}$, объекты S_C, X_C, Y_C категории $\mathbf{C}$ и S_D, X_D, Y_D категории $\mathbf{D}$, и морфизмы $\varphi_C \in \mathbf{C}(X_C, S_C)$, $\psi_C \in \mathbf{C}(Y_C, S_C)$, $\varphi_D \in \mathbf{D}(X_D, S_D)$ и $\psi_D \in \mathbf{D}(Y_D, S_D)$ этих категорий.

Лемма 5. *Наибольший общий частный случай пары образцов $(\varphi_C, \varphi_D): (X_C, X_D) \rightarrow (S_C, S_D)$ и $(\psi_C, \psi_D): (Y_C, Y_D) \rightarrow (S_C, S_D)$ в категории $\mathbf{C} \times \mathbf{D}$ существует тогда и только тогда, когда существуют наибольшие частные случаи пары образцов $\varphi_C: X_C \rightarrow S_C$ и $\psi_C: Y_C \rightarrow S_C$ в категории $\mathbf{C}$ и пары образцов $\varphi_D: X_D \rightarrow S_D$ и $\psi_D: Y_D \rightarrow S_D$ в категории $\mathbf{D}$.*

Доказательство. Допустим, объект Z_C категории $\mathbf{C}$ и пара морфизмов $\sigma_C: Z_C \rightarrow X_C$, $\tau_C: Z_C \rightarrow Y_C$ определяет наибольший частный случай пары образцов $\varphi_C: X_C \rightarrow S_C$ и $\psi_C: Y_C \rightarrow S_C$ в категории $\mathbf{C}$, а объект Z_D и

морфизмы $\sigma_D: Z_D \rightarrow X_D$, $\tau_D: Z_D \rightarrow Y_D$ - наибольший частный случай образцов $\varphi_D: X_D \rightarrow S_D$ и $\psi_D: Y_D \rightarrow S_D$ в категории $\mathbf{D}$. Тогда объект (Z_C, Z_D) категории $\mathbf{C} \times \mathbf{D}$ и пара морфизмов $(\sigma_C, \sigma_D): (Z_C, Z_D) \rightarrow (X_C, X_D)$ и $(\tau_C, \tau_D): (Z_C, Z_D) \rightarrow (Y_C, Y_D)$ той же категории определяют наибольший частный случай пары образцов $(\varphi_C, \varphi_D): (X_C, X_D) \rightarrow (S_C, S_D)$ и $(\psi_C, \psi_D): (Y_C, Y_D) \rightarrow (S_C, S_D)$. Действительно, равенства i для этих морфизмов очевидно выполнены, ибо они выполнены покомпонентно. Пусть теперь (Z'_C, Z'_D) - объект категории $\mathbf{C} \times \mathbf{D}$, а $(\sigma'_C, \sigma'_D): (Z'_C, Z'_D) \rightarrow (X_C, X_D)$ и $(\tau'_C, \tau'_D): (Z'_C, Z'_D) \rightarrow (Y_C, Y_D)$ - морфизмы этой категории, причем $(\sigma'_C, \sigma'_D)(\varphi_C, \varphi_D) = (\chi'_C, \chi'_D)$ и $(\tau'_C, \tau'_D)(\psi_C, \psi_D) = (\chi'_C, \chi'_D)$. Выписанная пара равенств эквивалентна двум парам равенств в категориях $\mathbf{C}$ и $\mathbf{D}$: $\sigma'_C\varphi_C = \chi'_C$, $\tau'_C\psi_C = \chi'_C$ и $\sigma'_D\varphi_D = \chi'_D$, $\tau'_D\psi_D = \chi'_D$. Применяя свойство ii к наибольшему частному случаю Z_C , находим морфизм $\rho_C: Z'_C \rightarrow Z_C$ такой, что $\chi'_C = \rho_C\chi_C$, а применяя его к наибольшему частному случаю Z_D - морфизм $\rho_D: Z'_D \rightarrow Z_D$ такой, что $\chi'_D = \rho_D\chi_D$. Тогда для морфизма $(\rho_C, \rho_D): (Z'_C, Z'_D) \rightarrow (Z_C, Z_D)$ имеем $(\chi'_C, \chi'_D) = (\rho_C, \rho_D)(\chi_C, \chi_D)$, чем доказано свойство ii.

Теорема 3. *Любая контекстная решетка сохраняет свойство существования наибольших частных случаев.*

Доказательство. Согласно предыдущей лемме достаточно доказать, что во всякой контекстной решетке, рассматриваемой как категория, для каждой пары образцов существует наибольший частный случай. Посмотрим, во что превращается определение наибольшего частного случая для решетки. Пусть P - контекстная решетка. Поскольку порожденная ею категория имеет лишь один объект, будем говорить только о морфизмах. Образец, стало быть, задается элементом решетки P . Пусть $p \in P$, $q \in P$ - два образца. Понятно, что наибольший частный случай - это такой элемент $w \in P$, что

i. для некоторых $u \in P$, $v \in P$ имеем $u \vee p = v \vee q = w$;

ii. если образец $w' \in P$ таков, что для некоторых элементов $u' \in P$, $v' \in P$ имеем $u' \vee p = v' \vee q = w'$, то существует элемент $z \in P$ такой, что $w' = zw$.

Положим $w = p \vee q$ и покажем, что такой элемент w удовлетворяет условиям i и ii. Очевидно, что $w \vee p = p \vee q \vee p = p \vee q = w$ и, аналогично, $w \vee q = w$, так что в качестве u и v можно взять w . Далее, если $u' \vee p = v' \vee q = w'$, то $w' \geq p$ и $w' \geq q$, откуда $w' \geq p \vee q = w$, поэтому $w' = w'w$, т.е. в качестве z из условия ii можно взять w' . Доказательство теоремы закончено.

Мы видим, что перенос свойств с категории на расслоенную категорию является весьма нетривиальной задачей. Одни свойства не наследуются расслоенной категорией, другие – наследуются, третьи наследуются при дополнительных ограничениях на контекстную решетку. Тут есть довольно обширное поле для дальнейших исследований.

Литература

1. Жожикашвили А.В., Стефанюк В.Л. Теория категорий в задачах представления знаний и обучения // Известия АН СССР. Техническая кибернетика. 1986, N.2, с.17-31.
2. Жожикашвили А.В., Стефанюк В.Л. Представление знаний и обучение в терминах теории категорий // Алгоритмы обработки экспериментальных данных. М.: Наука, 1986, с.130-155.
3. В.Л.Стефанюк, А.В. Жожикашвили. Теоретико-категорные образцы для задач искусственного интеллекта, новые результаты. Известия академии наук. «Теория и системы управления», №5, 1999, с. 5-16.
4. С.Маклейн, Категории для работающего математика, М., Физматлит, 2004.
5. А.В.Жожикашвили, Стефанюк В.Л. Алгебраическая теория продукционных систем. Восьмая национальная конференция по искусственному интеллекту. КИИ'2002, 7-12 октября, Коломна, Россия. Труды конференции. М., Физматлит, т. 1, с. 428-436.
6. V.L. Stefanuk, A.V. Zhzhikashvili. Productions and rules in artificial intelligence, KYBERNETES, The International Journal of Systems & Cybernetics. MCB University Press: 2002, pp. 817-826.
7. А.В.Жожикашвили, В.Л.Стефанюк. О понятии продукции в искусственном интеллекте. Известия академии наук, серия «Теория и системы управления», N4, 2002, с.76-81.
8. А.В.Жожикашвили, В.Л.Стефанюк. Продукционные сети: развитие теории ТК-продукций. Вестник Российского университета дружбы народов. Серия «Прикладная и компьютерная математика», т.2, 2003, с. 118-126.
9. Стефанюк В.Л. Поведение квазистатической оболочки в изменяющей нечеткой среде. Труды IV национальной конференции с международным участием «Искусственный интеллект - 94», Рыбинск, сентябрь 16-24, 1994, т.1, с.199-203
10. Лебедева Т.Г., Осипов Г.С. Архитектура и управляемость дискретных динамических систем, основанных на знаниях // Изв. РАН. Теория и системы управления, 2000, №5, с. 703-709
11. V.L. Stefanuk, Dynamic expert systems. KYBERNETES, The International Journal of Systems & Cybernetics. V.29, Number 5/6, 2000, pp. 702-709.
12. Минский М. Фреймы для представления знаний. В кн. Уинстона П. Психология машинного зрения, М., Мир, 1978.
13. Стефанюк В.Л. Динамическая экспертная система и логическая проблема фрейма. Международная конференция по искусственному интеллекту (AIS'07/CAD-2007), т.2, с. 107-113, М, Физматлит, 2007.

Стефанюк Вадим Львович. Ведущий научный сотрудник Института проблем передачи информации им. А.А. Харкевича РАН. В 1962 г. окончил физический факультет Московского государственного университета. Доктор технических наук, профессор РУДН. Свыше 190 печатных работ и монографий. Область научных интересов: математическое и компьютерное моделирование мышления, искусственный интеллект, коллективное поведение, локально-организованные системы.

Жожикашвили Александр Владимирович. Старший научный сотрудник Института проблем передачи информации им. А.А. Харкевича РАН. В 1977 году окончил механико-математический факультет Московского государственного университета. Кандидат технических наук, старший научный сотрудник. Свыше 30 печатных работ и монографий. Область научных интересов: искусственный интеллект, представление знаний, экспертные системы.