

## ТЕОРЕТИКО-КАТЕГОРНЫЙ ПОДХОД К ОБУЧЕНИЮ НА ОСНОВЕ ОБОБЩЕНИЯ

В.Л. Стефанюк (*stefanuk@iitp.ru*)  
А.В. Жожикашвили (*zhozhik@iitp.ru*)

Институт проблем передачи информации РАН, Москва

В работе описывается новый подход к машинному обучению – обучение на основе обобщения. В работе излагается математический аппарат, основанный на языке теории категорий, который может быть использован в качестве математического формализма для такого обучения.

**Ключевые слова:** обучение, обобщение, образец, теория категорий.

### Введение

В настоящий момент в Искусственном интеллекте происходит бум, связанный с нейронными сетями. По мнению авторов настоящего доклада он связан не столько с самими нейронными сетями, сколько с тем, что утвердилась идея, состоящая в том, что компьютер может приобрести способность решать интеллектуальные задачи в результате обучения. Сама идея обучающегося компьютера достаточно стара, однако в последние годы многократно возрос доступный объем данных, которые можно использовать для обучения. Это позволило обучающимся системам, в частности, нейросетевым, успешно решать прикладные задачи.

Но нейронные сети являются лишь простым и удобным инструментом для обучения, не единственным, а возможно – и не лучшим. Авторы настоящего доклада много лет занимались продукционными интеллектуальными системами, в частности – экспертными системами. Такие системы имеют целый ряд привлекательных моментов. Упомянем, к примеру, их способность объяснить свои решения, указав правила, на основе которых эти решения были приняты. Нейронные сети таким свойством не обладают, человек не в состоянии понять, каким образом они получили результат. Это усложняет их применение в областях с высокой ценой ошибки, скажем – в медицине.

Однако возможность обучаться, свойственная нейронным сетям, пере-  
вешивает все эти недостатки. В классических продукционных сетях обу-  
чение не используется. Система продукций закладывается человеком-  
экспертом. Иногда этим даже занимается особый специалист – инженер  
по знаниям. Понимая, что этот подход тормозит использование продукци-  
онных систем, авторы доклада постоянно стремились создать механизмы  
автоматического формирования и развития базы знаний в процессе обу-  
чения, анализа правил, содержащихся в базе знаний, их модификации и  
создания новых правил.

К числу алгоритмов обучения, не связанных с нейронными сетями, ко-  
торые можно использовать и в продукционных системах, мы относим ал-  
горитм обучения на основе обобщения, многие годы развиваемый в на-  
шем коллективе. В свое время авторами доклада была предложена мате-  
матическая формализация ряда механизмов, связанных с представлением  
знаний [Стефанюк, 1999], основанная на аппарате теории категорий  
[Маклейн, 2004]. К их числу относится и механизм обучения на основе  
обобщения. Такая формализация позволяет формулировать алгоритмы  
обучения, не связанные буквально с тем, как устроены знания, которые  
система должна приобрести в результате этого обучения. Предложенная  
авторами теоретико-категорная технология создания интеллектуальных  
систем, основанных на знании [Zhozhikashvili, 2023] позволяет унифици-  
ровать и программирование таких систем.

## 1. Продукции, образцы, сопоставление и конкретизация

Авторы докладов много лет изучали интеллектуальные компьютерные  
системы, основанные на правилах, или продукциях – продукционные систе-  
мы. Именно такие системы были формализованы на теоретико-категорном  
языке. Все это было подробно описано в наших многочисленных статьях на  
эту тему. Приведем здесь возможно более короткое изложение.

Продукционная система, решая задачу, совершает целый ряд последо-  
вательных шагов. На каждом шаге она находит применимое в этот момент  
правило и применяет его. Затем она применяет следующее применимое  
правило и так далее, пока не достигнет поставленной цели. Применение  
каждой продукции увеличивает количество информации, известной сис-  
теме. Совокупность информации, известной в некоторый момент, мы на-  
зываем *ситуацией*. Таким образом, применение продукции меняет теку-  
щую ситуацию на новую, добавляя к ней информацию. Задача продукци-  
онной системы – получить ситуацию, являющуюся решением системы.  
Такая ситуация называется *целевой*. Поскольку под ситуацией мы пони-  
маем информацию, известную системе, под целевой ситуацией разумно  
понимать информацию, содержащуюся, среди прочего, то, что система  
должна определить в результате решения задачи. *Продукцией* мы называ-

ем описание двух ситуаций – ситуации, в которой продукция применима, и ситуации, которая возникает после ее применения. Эти описания ситуаций, составляющие правило, не являются абсолютно детализированными, некоторые моменты, не существенные для применимости данного правила, в них опущены. Такое описание, в котором опущен ряд деталей, мы называем образцом ситуации, или просто *образцом*. Если добавить к образцу эти опущенные детали, образец превратится в ситуацию. Такую операцию добавления опущенных деталей мы называем *конкретизацией* образца, а сами добавляемые детали – *конкретизатором*.

Пусть  $S$  – множество всех ситуаций, которые могут возникнуть при работе системы,  $O$  – некоторый образец,  $P$  – множество всех возможных конкретизаторов, которые могут быть использованы для конкретизации образца  $O$ . Выбирая любой конкретизатор  $p \in P$  и конкретизируя посредством этого конкретизатора образец  $O$ , мы получаем некоторую ситуацию  $s \in S$ . Таким образом, образец  $O$  можно рассматривать как отображение  $O \rightarrow P$ . Теперь можно уточнить описание продукции, именно, продукцией мы будем называть пару образов  $(O, p)$  с совпадающим множеством конкретизаторов, т.е. пару отображений  $O \rightarrow P$ ,  $p \in P$ . Если  $(O, p)$  – ситуация, то применению такой продукции к ситуации  $s$  дает ситуацию  $s'$  тогда и только тогда, когда  $s' = p(s)$ . Отметим, что результат применения продукции в общем случае определен неоднозначно, так как выбор элемента  $p \in P$  по условию  $s \in S$  может не быть однозначным.

В сложных задачах в качестве множеств ситуаций и множеств конкретизаторов могут использоваться различные множества, причем одно и то же множество может в одном случае выступать как множество ситуаций, в другом – как множество конкретизаторов, поэтому, если мы хотим использовать приведенный выше формальный подход в образцах, мы должны указать, какие множества могут быть множествами ситуаций/конкретизаторов и какие отображения между ними могут выступать в качестве образцов. При некоторых естественных условиях замкнутости это определяет подкатегорию категории множеств. Это соображение навело нас на мысль перевести всю теорию образов, сопоставления и конкретизации на язык теории категорий. Роль образов, в этом случае будут играть морфизмы категории. В этом случае пришлось отказаться от представления о том, что ситуация – элемент множества  $S$ , ибо объекты категории не обязаны быть множествами. Вместо этого ситуация понимается как частный случай образца. Это соответствует смыслу этих понятий: категория – это полностью конкретизированный образец, в котором не осталось опущенных деталей. Какие именно образцы считаются ситуациями – определяется спецификой конкретной задачи.

Все это приводит к следующему определению.

Системой образов называется категория <sup>1</sup>, в которой для любой пары объектов  $x, y$  и множества морфизмов  $M$  определено множество  $\{f(x), f(y)\}$ . Точное определение, приведенное в наших более ранних работах, содержит дополнительные требования к множествам  $M$ , которые мы здесь не приводим, чтобы не усложнять изложение. Образцом с областью конкретизаторов  $A$  и областью значений  $B$  считаем любой морфизм  $f: A \rightarrow B$ , ситуацией – морфизм  $g: A \rightarrow B$ . (Такой образец будем в дальнейшем называть  $f$ -образцом, а ситуацию –  $g$ -ситуацией.) Ситуация, таким образом, представляет собой частный случай образца. Ситуация сопоставима с образцом  $f$ , тогда и только тогда, когда существует морфизм  $h: A \rightarrow A$  такой, что  $h \circ f = g$ . Пусть  $x, y$  – объекты категории  $\mathcal{C}$ . Продукцией из  $A$  в  $B$  называется пара образов  $(f, g)$ , где  $f, g: A \rightarrow B$  – образцы категории  $\mathcal{C}$ . Будем говорить, что продукция применима к ситуации  $g$ , если эта ситуация сопоставима с образцом  $f$ , т.е. если существует морфизм  $h: A \rightarrow A$  для некоторого морфизма  $f: A \rightarrow B$ . В этом случае результатом применения продукции к ситуации является ситуация  $h \circ f$ .

## 2. Порядок на множестве образов

Образец можно рассматривать как способ единообразно описать множество ситуаций. Именно, образец может служить для описания тех ситуаций, которые сопоставимы с этим образцом. Определение сопоставимости ситуации с образцом легко обобщается на определение сопоставимости образца с образцом: говорят, образец  $f$  сопоставим с образцом  $g$ , если существует морфизм  $h: A \rightarrow A$  такой, что  $h \circ f = g$ . В этом случае будем также говорить, что образец  $f$  является частным случаем образца  $g$ . Логичность такой терминологии следует из легко доказуемого факта: если образец  $f$  сопоставим с образцом  $g$ , то всякая ситуация, сопоставимая с образцом  $f$ , сопоставима также и с образцом  $g$ , т.е. образец  $f$  описывает более широкое множество ситуаций.

Пользуясь понятием сопоставимости, можно для объекта категории  $\mathcal{C}$  ввести частичный порядок [Биркгоф, 1984] на множестве  $\mathcal{C}$ -образцов, полагая, что для  $f, g \in \mathcal{C}$  условие  $f \leq g$  означает, что образец  $f$  является частным случаем образца  $g$ .<sup>2</sup> Как будет понятно из дальнейшего,

---

<sup>1</sup> Будем считать, что все рассматриваемые категории являются малыми, чтобы избежать сложностей с теоретико-множественной аксиоматикой.

<sup>2</sup> Строго говоря, такое определение дает не порядок, а предпорядок. Для перехода к порядку надо провести некоторую факторизацию, как это всегда делается. Здесь мы не будем углубляться в эти тонкости, ибо дальнейший материал не содержит строгих математических доказательств и является достаточно понятным и без таких уточнений.

нас будет интересовать случай, когда получившееся упорядоченное множество является решеткой [Биркгоф, 1984]. Это бывает, естественно, не всегда, но во многих практически важных случаях это так.

Пусть  $\{a_1, \dots, a_n\}$  – образцы,  $a_i \leq a_j$  тогда и только тогда, когда  $i \leq j$ . Будем называть образец  $a_i$  обобщением образцов  $a_1, \dots, a_n$  и  $a_i$ , если существуют морфизмы  $f_1, \dots, f_n$  такие, что  $a_i = f_1(a_1) \vee \dots \vee f_n(a_n)$ . Будем записывать такую ситуацию как обобщение  $a_i$  образцов  $a_1, \dots, a_n$ . Очевидно, что в этом случае  $a_i \leq a_j$  тогда и только тогда, когда  $a_j$  обобщение образцов  $a_1, \dots, a_n$ . Будем называть обобщение  $a_i$  наименьшим обобщением образцов  $a_1, \dots, a_n$  и  $a_i$ , если для любого другого обобщения  $a_j$  образцов  $a_1, \dots, a_n$  существует морфизм  $f$  такой, что  $a_j = f(a_i)$ . Наименьшее обобщение образцов  $a_1, \dots, a_n$  и соответствует верхней грани  $a_1 \vee \dots \vee a_n$  в теории решеток.

### 3. Обучение на основе обобщения

Изложив некоторые детали необходимого нам математического аппарата мы переходим к изложению методов обучения, основанных на обобщении.

В общих чертах обучающаяся система работает следующим образом. Система должна научиться решать задачи определенного типа. Для этого используется процедура обучения с учителем. Учитель предлагает системе задачи, решения которых ему известны. Система пытается их решить. Если система не может решить задачу – учитель предоставляет ей решение, которое она использует для обучения, изменяя свою базу знаний. Говоря на языке продукционных систем решение, предоставляемое учителем – это просто описание целевой ситуации, которую система получила бы, если бы обладала достаточными знаниями. Если система решает задачу, предложенное ей решение проверяется учителем. В случае, если решение – правильное, никаких действий не предпринимается. В случае, если решение неверное, учитель информирует систему об этом, и система использует эту информацию для обучения. Как именно использовать такую «отрицательную» информацию – вопрос весьма непростой. Частично эта тема раскрывается в разделе 5, но на настоящий момент она проработана значительно слабее, чем использование «положительной» информации.

Встретив задачу, которую она не может решить, и получив от учителя решение, система первым делом создает и записывает в базу знаний правило, левой частью которого является условие задачи, правой – полученное от учителя решение. В дальнейшем она может использовать это правило только в одном случае: если на вход поступит точно такая же задача. Тогда она выдаст такое же решение. Ни для чего другого это правило не годится. Однако система сравнивает каждое новое правило с уже существующими. Если вновь созданное правило оказывается близким к одному из созданных ранее и эти два правила имеют много общего, система строит обобщение двух правил, помещает это обобщение в базу знаний, а исходные правила из базы удаляет.

Используем описанный выше формализм чтобы конкретизировать процесс обучения.

Проще всего изложить алгоритм обучения для задачи распознавания. Эта задача состоит в следующем. На вход системы распознавания поступают объекты определенного вида. Система должна научиться распознавать объекты, относящиеся к определенному классу, т.е. задача распознавания может рассматриваться также как задача классификации.

В нашей терминологии поступающие на вход объекты будут рассматриваться как ситуации и система должна определить, относится ли поступившая на вход ситуация к некоторому классу. Для этого формируется образец, описывающий ситуации этого класса. Система производит сопоставление поступившей ситуации с этим образцом. Если сопоставление прошло успешно, считается, что ситуация относится к нужному классу, если нет – не относится. В сложных случаях одного образца может не хватить, система формирует базу образцов и решение о том, что поступившая ситуация относится к классу, принимается в том случае, если ситуация сопоставима с одним из этих образцов из этой базы.

База образцов, используемая для распознавания, строится следующим образом. Если на вход поступает ситуация , относящаяся к исследуемому классу. Эта ситуация добавляется к базе как образец (выше говорилось, что ситуация – частный случай образца). Затем эта ситуация сравнивается со всеми образцами, содержащимися в базе. Если для ситуации и некоторого образца возможно построить нетривиальное наименьшее обобщение , то из базы удаляются и ситуация , и образец , вместо них в базу добавляется образец .

Тут возникает очень много вопросов.

Что такое нетривиальное наименьшее обобщение? Если упорядоченное множество образцов является решеткой, то в нем любая пара элементов имеет точную верхнюю грань. Следовательно, любая пара образцов имеет наименьшее обобщение. Какое из них считать нетривиальным? Этот вопрос может быть решен только в результате экспериментов. (Отметим, что в современной теории обучения многие решения принимаются в результате экспериментов. Иногда эти решения кажутся странными, но эксперименты показывают, что они работают.) В некоторых наших моделях обобщение считалось нетривиальным, если оно не совпадало с максимальным элементом решетки. В других случаях надо было вводить некоторые ограничения, требующие, чтобы обобщение было бы не слишком общим в каком-то смысле.

Более серьезный вопрос: откуда следует, что если ситуация относится к некоторому классу и всякая ситуация, сопоставимая с образцом , относится к этому классу, то и всякая ситуация, сопоставимая с образцом , тоже относится к этому классу? Ни откуда. Это утверждение нельзя

доказать, в общем случае оно неверно. Это – всего лишь разумная гипотеза, состоящая в том, что если ситуация относится к некоторому классу, ситуации, описываемые образцом , относится к этому классу, то условием принадлежности к классу может быть то общее, что содержится в описании ситуации и образца , а именно это общее содержит образец

. Но всякая гипотеза может оказаться ложной. В этом случае база должна быть подправлена, об этом говорится в следующем разделе.

Система распознавания проще, чем производственная система. В такой системе правило состоит из одной лишь левой части. При сопоставлении решается вопрос применимо – не применимо правило, а вопрос о том, что делать, если применимо, не рассматривается. Вопрос о том, как создавать правила в процессе обучения, в первую очередь – как определить операцию обобщения двух правил.

Пусть в результате обработки информации, поступившей от учителя, было создано правило , где , – образцы. Допустим, в базе знаний уже существует правило . Строим образец – наименьшее обобщение образцов и , образец – наименьшее обобщение образцов и , и заменяем пару правил и на правило .

Авторы отдают себе отчет, что столь краткое изложение не может дать сколько-нибудь полное изложение алгоритмов обучения. Но описать алгоритм детальнее пока не представляется возможным. Во-первых, мы находимся на стадии разработки основных идей обучения на основе обобщения. Очень многие вопросы пока не решены, многие детали не проработаны. Во-вторых, даже то, что проработано – слишком велико по объему для небольшого доклада.

#### **4. Неоправданно широкие обобщения**

Обучение на основе обобщения основано на следующей идее. Пусть и – описания двух множеств ситуаций (например, образцы). Пусть – обобщение этих описаний, т.е. описание, включающее в себя то общее, что есть в , и в . Допустим, мы знаем, что как ситуации, подходящие под описание , так и ситуации, подходящие под описание , удовлетворяют некоторому условию. Логично предположить, что выполнение этого условия является следствием той общей части, которая содержится в обоих описаниях. Из этого, вроде бы, следует, что ситуации, подходящие под описание , тоже удовлетворяют условию. Несмотря на некоторую логичность, высказанное утверждение отнюдь не всегда является верным. Да, под описание подходят все ситуации, подходящие под описание , и все ситуации, подходящие под описание , но под описание могут походить и многие другие ситуации, не подходящие ни под , ни под , и эти ситуации могут условию не удовлетворять. В этом случае мы говорим, что – неоправданно широкое обобщение и .

Совершая такие обобщения, обучающаяся система может прийти к неверным выводам, и нам не известен способ, позволяющий этого избежать. То же самое происходит при обучении человека: увидев, что нечто происходит часто, он может сделать вывод, что это происходит всегда, и пребывать в этом заблуждении то тех пор, пока не попадет в ситуацию, когда этого не произошло, или пока кто-то более образованный его не поправит.

Таким образом, в процессе обучения система формирует не безусловные знания, а гипотезы. Эти гипотезы надо еще проверять каким-либо способом. Отметим, что умение формулировать разумные гипотезы – важная способность интеллекта, в том числе и искусственного интеллекта.

Борьба с неоправданно широкими обобщениями – возможно, более сложная задача, чем само построение обобщений. Для того, чтобы уменьшить возможность появления таких обобщений и тем самым приблизить формируемые при обучении гипотезы к реальным знаниям используются два приема: первый – возможное уменьшение степени обобщения, второй – доработка базы знаний в процессе обучения. Об уменьшении степени обобщения мы напишем ниже, доработке базы посвящен следующий раздел.

Первым моментом, связанным с минимизацией степени обобщения, является использование наименьшего обобщения в терминах теории категорий или верхней грани – в терминах теории решеток. Но этого не достаточно. Полная решетка всегда имеет наибольший элемент. Это означает, что даже если два образца не имеют ничего общего, по ним все равно можно построить верхнюю грань – она будет совпадать с наибольшим элементом решетки и с ней будет сопоставима любая ситуация.

Чтобы этого избежать, надо несколько уточнить процедуру, содержащуюся в последнем абзаце предыдущего раздела. При описанном там построении наименьшего обобщения образцов  $A$  и  $B$  и наименьшего обобщения образцов  $C$  и  $D$ , проверить, не является ли это обобщение слишком общим, и, если является – отказаться от обобщения. Что значит – слишком общим? Можно просто считать слишком общим обобщение, совпадающее с наибольшим элементом решетки. Но возможно, следует выработать более тонкий критерий. На сегодняшний момент готовых решений у авторов доклада нет.

## **5. Доработка базы знаний в процессе обучения**

Если предыдущие разделы содержат научные результаты, то этот посвящен нерешенным пока вопросам, и их больше, чем вопросов решенных.

Одна из главных проблем – проблема неоправданно широких обобщений. Один из способов борьбы с ними – удаление из базы знаний правил, возникших в результате таких обобщений. Если предложенное системой решение будет отвергнуто учителем – правило, на основании которого было получено это решение, должно быть удалено.



Тут возникает целый ряд проблем. Система приходит к заключению, используя цепочку правил. Как понять, какое из правил цепочки, приведшей к неверному заключению, является ошибочным? Есть целый ряд соображений, посвященных этому вопросу, но ни одно из них не может считаться решением проблемы.

Вторая проблема состоит в том, что при создании по двум правилам их обобщения это обобщение замещает исходные правила, которые удаляются из базы. Это означает, что если обобщение будет впоследствии удалено из базы – там не останется ничего, ни исходных правил, которые были обобщены, ни результата. Как этого избежать? Один из способов – хранить вместе с каждым обобщением ссылку на те правила, из которых оно получено. Это позволит при удалении обобщения восстановить те правила, из которых оно было получено. Другой способ – ранжировать все правила по полезности и использовать сперва более полезные, и только в случае, когда они не подошли – менее полезные. Самый простой способ – расположить все правила базы знаний в виде последовательности и при поиске применимого правила перебирать правила в порядке, в котором они расположены в этой последовательности. Обобщение должно размещаться в этой последовательности раньше, чем правила, из которых оно было получено, так что после включения в базу знаний обобщения эти правила не будут использоваться, однако после удаления обобщения они снова начнут работать. Возможно, следует строить базу знаний не как последовательность правил, а как дерево, в котором правила расположены в вершинах и упорядочены по степени общности. Это значительно ускорит и выбор правила, применимого в некоторой ситуации.

Обучающаяся система должна постоянно анализировать базу знаний и, возможно, вносить в нее изменения. Рассмотрим, к примеру, правило, полученное обобщением двух других. Исходные правила при этом удаляются из базы или передвигаются в более далекие позиции в последовательности или в дереве правил. Может оказаться, что созданное обобщение является более общим, чем некоторые другие правила, не только те два, из которых обобщение было получено. В этом случае и эти правила следует переместить подальше или удалить, сохранив возможность восстановления.

Мы перечислили далеко не все нерешенные проблемы. Их решение – предмет дальнейшего исследования.

### **Заключение**

Обучение, основанное на обобщении – перспективный подход к обучению, который может быть использован при создании интеллектуальных компьютерных систем. Он может быть применен в традиционных для Искусственного интеллекта системах, таких, например, как экспертные

системы. Это позволяет объединить его с разработанными ранее методами построения интеллектуальных систем, включив в работу этих систем и обучение, и прямую передачу знаний от эксперта. Знания, полученные в результате обучения, представляются в форме, понятной человеку. Обучение, основанное на обобщении, является достаточно естественным, оно повторяет приемы обучения, свойственные человеку.

В настоящий момент у авторов нет сколько-нибудь завершенной системы, обучающейся на основе обобщения. Предложенные методы требуют апробации, которая и покажет, насколько они эффективны. Отсутствие законченной системы не позволяет провести сравнение такого алгоритма обучения с другими известными, прежде всего – с обучающимися нейронными сетями. Авторы могут высказывать лишь предположения, которые нуждаются в экспериментальной проверке.

Одно из таких важных на наш взгляд предположений мы выскажем прямо здесь. Знания, которые создаются при обучении на основе обобщения, много понятнее для человека, чем знания, создающиеся в результате обучения нейронной сети. Результатом обучения нейронной сети является набор весов. Понять его смысл про хото сколько-нибудь значительном размере сети для человека невозможно. Результатом обучения продукционной системы является набор правил, хорошо понятных для человека. Хотя в реальных задачах, когда количество правил исчисляется тысячами, разобраться в них тоже непросто, это все-таки много проще, чем оценивать веса, связанные с ребрами сети. Получив некоторый результат, продукционная система может объяснить свое решение, приведя те правила, на основании которых она пришла к этому результату. Ничего подобного нейронная сеть сделать не может.

Есть надежда, что такое обучение на основе обобщения будет идти быстрее, чем обучение нейронных сетей, и не потребует столь большого обучающего материала.

Отметим один существенный момент. Если нейронная сеть оперирует вещественными числами и обучении состоит в подборе этих чисел, в основе продукционной системы лежит бинарная логика да/нет. Достоинство это, или недостаток – покажут дальнейшие исследования.

### Список литературы

- [**Стефанюк, 1999**] Стефанюк В.Л., Жожикашвили А.В. Теоретико-категорные образцы для задач искусственного интеллекта, новые результаты // Известия РАН. Теория и системы управления. – 1999. – № 5. – С. 5-16.
- [**Маклейн, 2004**] Маклейн С. Категории для работающего математика. – М.: Физматлит, 2004.
- [**Zhozhikashvili, 2023**] Zhozhikashvili A.V., Stefanuk V.L. Category Technology to Design Intelligent Systems for Complicated Decisions, // Lecture Notes in Networks and Systems book series (LNNS, volume 566), Springer Nature.
- [**Биркгоф, 1984**] Биркгоф Г. Теория решеток. – М.: Наука, 1984.