Б. А. КУЛИК
Институт проблем машиноведения РАН, Санкт-Петербург

СЕМАНТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИСЧИСЛЕНИЯ ПРЕДИКАТОВ ПЕРВОГО ПОРЯДКА НА ОСНОВЕ АЛГЕБРЫ МНОЖЕСТВ

Семантическое моделирование исчисления предикатов первого порядка основано на предложенной Э. Мендельсоном интерпретации, в которой предикатам и формулам со свободными переменными соответствуют n-местные отношения. В эту интерпретацию внесены следующие изменения: разным областям изменения переменных (доменам) соответствуют разные имена (атрибуты); n-местные отношения представлены как объединения декартовых произведений. Установлено, что математической моделью измененной интерпретации является алгебра кортежей, которая изоморфна алгебре множеств. Показаны некоторые позитивные результаты такой интерпретации, перечислены нерешенные проблемы.

Введение. В докладе представлены результаты исследований, с помощью которых намечается подход к решению некоторых задач, сформулированных в Детализированных сведениях Приоритетных направлений фундаментальных и поисковых научных исследований (ПФНИ 2021–2030) в разделе 1.1.1.6. «Математическая логика», а именно, «развитие теории семантического моделирования», «дальнейшая разработка логических формализмов и подходов для работы с естественными языками», «исследования по взаимосвязи синтаксических и семантических свойств в логике».

В качестве математической основы семантического моделирования в классической логике предлагается выбрать алгебру множеств, поскольку в естественных рассуждениях речь идет об отношениях типа «часть – целое», «объекты – свойства» и т. д., которые можно представить как отношения и операции с множествами. Некоторые преимущества такого выбора показаны для исчисления предикатов первого порядка.

В докладе рассмотрены отличия алгебры множеств от аксиоматической теории множеств.

Алгебра множеств и теория множеств. В основе современной логики лежит сформировавшийся на рубеже XIX и XX столетий ***аксиоматический подход***, в котором аксиомы и правила вывода сформулированы в символьных выражениях так, что их трудно интерпретировать. В рамках этого подхода развивалась современная ***теория множеств***, основы которой были заложены исследованиями Г. Кантора, Р. Дедекинда и др. в последней четверти XIX века. Через некоторое время были открыты парадоксы теории множеств (Г. Кантор, Ч. Бурали–Форти и др.), а на рубеже XIX и XX столетий стали завоевывать популярность публикации математиков и философов, заложивших основы современного аксиоматического подхода (Г. Фреге, Ч. С. Пирс, Дж. Пеано, Б. Рассел и др.) [1].

На фоне этих событий начала свое развитие современная логика, в которой термин «множество» оказался под запретом, в силу обнаруженных парадоксов теории множеств. Хотя даже не специалисту в логике понятно, что термин сам по себе не может быть противоречивым – его противоречивость может зависеть от того, как этот термин определили и как его связали с другими терминами теории.

Впоследствии из теории множеств был выделен более простой ее вариант – ***алгебра множеств***, законы которой могут быть обоснованы без аксиом на основе лишь определений операций (дополнение, пересечение, объединение) и отношений (включения и равенства) [2].

Законы алгебры множеств соответствуют законам классической логики, поэтому возможность доказать законы алгебры множеств без аксиом означает, что для классической логики нет необходимости в аксиомах.

Источником противоречий термина «множества» в теории множеств является то, что в этой теории разрешается множеству быть элементом множества. Такое допущение в некоторых разделах математики присутствует и в настоящее время. Споры на эту тему до сих пор не утихли. Но одно, несомненно: в алгебре множеств это допущение можно убрать – законы алгебры множеств от этого не изменятся. Обусловлено это тем, что в алгебре множеств в отличие от теории множеств системообразующим является не отношение принадлежности элемента и множества ($\in$), а отношение включения множеств ($\subseteq$). Поэтому запрет термина «множество» в логике нельзя считать обоснованным.

Интерпретация исчисления предикатов первого порядка. Современная математическая логика также развивается в рамках аксиоматического подхода и содержит два основных раздела, предназначенных для дедуктивного анализа: относительно простое ***исчисление высказываний*** и весьма сложное ***исчисление предикатов***.

Хотя имеется немало широко известных публикаций по математической логике, в докладе в основном цитируется наиболее популярная и достаточно ясно изложенная книга на эту тему, а именно 6-е издание книги Э. Мендельсона [3]. На русском языке в 1971 г. было опубликовано 3-е издание этой книги.

Обучение математической логике начинается с двух моделей. Первая модель – это ***таблицы истинности*** для логических связок $\neg$ (не), $\wedge$ (и), $\vee$ (или), $\supset$ (если, то). С помощью таких таблиц можно доказывать теоремы исчисления высказываний, но для обоснования теорем исчисления предикатов они малопригодны. Таблицы истинности по сути тоже являются аксиомами, так как в основе многих неклассических логик лежат измененные таблицы истинности логических связок.

Вторая (основная) модель начинается с определения ***языка первого порядка*** ($\mathcal{L}$), где предусматривается использование определенного алфавита для обозначения ***переменных***, ***констант*** (значений переменных), ***функций*** и ***предикатов***. В языке $\mathcal{L}$ также используются ***логические связки***, в состав которых, помимо $\neg$, $\wedge$, $\vee$ и $\supset$, входят ***кванторы*** $\forall$ (для всех) и $\exists$ (существует). Излагаются правила, с помощью которых формируются правильно построенные формулы (***ппф***).

Язык $\mathcal{L}$, в свою очередь, используется для построения ***теории первого порядка*** K, в которой используются ***ппф*** языка $\mathcal{L}$, а также ***аксиомы*** и ***правила вывода***, причем аксиомы делятся на два класса: ***логические*** и ***собственные*** (или нелогические). Собственные аксиомы служат для построения различных теорий (например, теории групп или формальной арифметики). Если собственных аксиом нет, то теория K называется ***исчислением предикатов первого порядка***.

В [3] предлагается следующая интерпретация языка первого порядка $\mathcal{L}$. В качестве области интерпретации (*domain*) переменной используется множество D элементов (констант), а для n-местных предикатов и формул со свободными переменными областью интерпретации является n-местное отношение, т. е. подмножество n-местных кортежей элементов из декартова произведения множеств D^n.

Рассмотрим более подробно, что это означает. Если переменная не находится в области действия какого-либо квантора, то она считается ***свободной***. Например, в формуле $\forall y(P(x)\wedge Q(x, y))$ переменная x свободная, в то время как переменная y – ***связанная***. То, что в этой формуле оказывается свободной единственная переменная x, означает, что интерпретацией этой формулы является какое-то подмножество (возможно, пустое) области D. Если бы не было квантора $\forall y$, то свободными были бы переменные x и y, а интерпретацией этой формулы было бы некоторое двухместное отношение, т. е. подмножество D^2.

В рассмотренную выше интерпретацию $\mathcal{L}$ были внесены следующие изменения.

Изменение 1. Для разных переменных языка $\mathcal{L}$ предложено использовать не одну какую-то область интерпретации D, а разные области интерпретации. Поэтому, во избежание возможных несогласованностей, было предложено по аналогии с базами данных приписывать к именам интерпретаций формул языка $\mathcal{L}$ ***схему отношения***, т.е. последовательность имен областей интерпретации переменных, формирующих это отношение. С учетом этого, имена областей интерпретации переменных названы ***атрибутами***, а области интерпретации атрибутов – ***доменами***.

Изменение 2. Для многих задач логического анализа более удобно рассматривать n-местное отношение не как множество кортежей элементов, а как объединение декартовых произведений. Поскольку ДП формируется из множеств, то в качестве значений атрибута используются не элементы его домена, а имена или обозначения (например, A_2 или $\{b, d\}$) всех подмножеств домена. Множества с этими именами или обозначениями названы ***компонентами*** атрибута. Короче: *компоненты – это произвольные подмножества домена атрибута*.

Алгебра кортежей. Оказалось, что измененную интерпретацию языка $\mathcal{L}$ можно выразить с помощью алгебры множеств. Но для этого потребовалось разработать и обосновать новую математическую структуру, получившую название ***алгебра кортежей*** [4]. С алгеброй множеств ее связывает то, что в ней используются те же операции (дополнение, пересечение, объединение), те же отношения (равенства и включения) и те же законы (де Моргана, транзитивности, непротиворечия и т. д.). Отличие только в том, что в ней используются не обычные множества, а структуры, которые можно с помощью вычислений представить множествами n-местных кортежей элементов (т.е. n-местными отношениями). Эти структуры – ***декартовы произведения множеств*** (ДП) ***и их объединения***. Как выяснилось в процессе исследований, они являются интерпретациями основных типов формул математической логики.

Объединение ДП, рассматриваемое как отдельная структура, ранее не исследовалось. Для этой структуры не были известны алгоритмы операций (дополнение, пересечение, объединение), алгоритмы проверок включения одной структуры в другую и т. д. В публикациях содержались только отдельные операции для ДП (их пересечение и разность), а также алгоритм проверки включения одного ДП в другое. Оказалось, что алгоритмы для новой структуры и их обоснования можно существенно упростить, если отказаться от общепринятых обозначений ДП (D^n, $A \times B \times C$, $\prod_{i=1}^{n} D_i$ и т. д.). Вместо этого предложено представлять ДП как кортежи компонент, при этом каждая компонента с помощью схемы отношения привязывается к определенному атрибуту. Одна из четырех структур алгебры кортежей – ***C- кортеж*** – как раз и является подобной записью ДП. Тогда более сложные структуры записываются в виде матриц [4–6].

В широком смысле ***интерпретация*** – это толкование, объяснение, раскрытие смысла чего-нибудь. Предполагается, что интерпретация близка к семантике. В книге [3] во Введении сказано, что в логике «семантические понятия носят теоретико-множественный характер». Доказано, что алгебра кортежей изоморфна алгебре множеств [4]. Поэтому вполне правомерно предположить, что интерпретация исчисления предикатов на основе алгебры множеств является ***семантическим моделированием***.

Нерешенные проблемы.

1. Задача Steamroller (№ 47 в [10]), которая выражается на языке исчисления предикатов, является иллюстрацией сложности логического вывода. Предполагается, что ее формулировка на языке алгебры кортежей позволит упростить ее решение. Однако до настоящего времени такая формулировка не найдена. Также отсутствует обоснование того, что этого нельзя сделать.
2. Не рассмотрена интерпретация и область ее применения для функциональных символов.
3. Не исследована возможность замены универсума Эрбрана [11] более простым вариантом на основе алгебры кортежей.
4. Не исследована возможность интерпретации исчисления предикатов высших порядков.
5. Не исследована возможность интерпретации теоремы Геделя о неполноте [3].

Список можно продолжить.

Заключение. Исследования показали, что, помимо логического анализа, алгебру кортежей можно использовать в следующих областях дискретной математики и информационных технологий: 1) реляционные модели; 2) графы и сети; 3) системы искусственного интеллекта (экспертные системы, семантические сети, фреймы, онтологии); 4) логико-вероятностные методы, включая вероятностную логику; 5) дискретные автоматы; 6) задачи удовлетворения ограничений (Constraint Satisfaction Problem – CSP); 7) модели вопросно-ответных систем; 8) при машинной реализации – сокращение трудоемкости алгоритмов решения сложных задач логического анализа за счет специфических свойств АК, а также за счет возможности эффективного распараллеливания алгоритмов [5–7].

С помощью средств математической логики трудно, а порой невозможно применять многие необходимые в естественных рассуждениях методы логического анализа, такие как проверка гипотез, анализ неопределенностей, распознавание и анализ ошибок и некорректностей в рассуждениях, вывод абдуктивных заключений, анализ пресуппозиций и т. д. В то же время эти задачи решаются с помощью алгебры кортежей [5, 6, 8].

С помощью исчисления предикатов можно решать лишь часть задач дедуктивного анализа, т. е. поиск доказательств теорем, в случае если их формулировки известны. Однако задача поиска следствий с заранее заданными свойствами в этой системе не решается. В то же время решение этой задачи стало возможным с помощью алгебры кортежей [6, 9].

ЛИТЕРАТУРА

1. **Бурбаки Н.** Теория множеств. Книга 1. Основные структуры анализа. М.: Мир, 1965. 456 с.
2. **Курант Р., Роббинс Г.** Что такое математика? 3-е изд., испр. и доп. М.: МЦНМО, 2001. 568 с.
3. Mendelson, E. Introduction to Mathematical Logic. Boca Raton, London, New York: Taylor & Francis Group, 2015 (6th ed.). 499 pp.
4. **Кулик Б.А.** Логический анализ систем на основе алгебраического подхода: дис. докт. физ.-мат.наук / Санкт-Петербург, 2007. 291 с.
5. **Кулик Б.А., Зуенко А.А., Фридман А.Я.** Алгебраический подход к интеллектуальной обработке данных и знаний. СПб.: Изд-во Политехн. ун-та, 2010. 235 с.
6. **Кулик Б.А.** Логика и математика: просто о сложных методах логического анализа. СПб.: Политехника, 2020. 141 с.
7. Kulik B., Fridman A. Complicated Methods of Logical Analysis Based on Simple Mathematics. Cambridge Scholars Publishing, 2022. 195 pp.
8. **Кулик Б.А.** Исследование противоречий в естественных рассуждениях на примерах метафор и пресуппозиций. *Труды Семнадцатой Национальной конференции по искусственному интеллекту с международным участием*. КИИ-2019 (21–25 октября 2019 г., г. Ульяновск, Россия). Ульяновск: УлГТУ, 2019. Т. 2. С. 192-200.
9. **Кулик Б.А.** Вывод следствий с предварительно заданными свойствами // Системный анализ в проектировании и управлении. В 3 ч. Ч.2: сб. научных трудов XXV Международной научной и учебно-практической конференции, 13-14 октября 2021 г. СПб.: ПОЛИТЕХ-ПРЕСС, 2021. Часть 2. С. 89-97.
10. Pelletier F.J. Seventy-Five Problems for Testing Automatic Theorem Provers. *Journal of Automated Reasoning*, 1986, Vol. 2. P. 191-216.
11. **Чень Ч., Ли Р.** Математическая логика и автоматическое доказательство теорем. М., Наука. 1983. 360 с.

B.A.Kulik (Institute for Problems in Mechanical Engineering of the Russian Academy of Sciences, Saint Petersburg)

Semantic modeling of first-order predicate calculus based on set algebra

Semantic modeling of first-order predicate calculus is based on an interpretation in which predicates and formulas with free variables correspond to n-place relations. The following changes have been made to this interpretation: i) different names (attributes) correspond to different domains, ii) n-plase relations are represented as unions of Cartesian products. The resulting mathematical structure – the n-tuple algebra – is isomorphic to the algebra of sets. Some positive results of such interpretation are shown, unsolved problems are listed.

Авторы готовы представить текст на английском языке для сборника материалов мультиконференции, который будет подан для индексирования в Scopus.