

УДК 512.567, 004.825
AMS MSC2020: 08A72

Основанные на решетках алгебраические системы и их приложения в задачах управления знаниями¹

Махортов С. Д.

Воронежский государственный университет

Аннотация. Алгебраические методы предоставляют эффективные средства формального построения и исследования моделей в различных областях информатики, включая интеллектуальные системы. Это положение в полной мере относится к логическим системам продукционного типа, которые образуют теоретическую основу важного направления в дисциплине искусственного интеллекта. В докладе исследуется алгебраическая структура, моделирующая распределенную интеллектуальную систему продукционного типа с нечеткими правилами (FDLP-структура). Вводится класс уравнений, служащий теоретической основой нового метода обратного логического вывода. Метод направлен на минимизацию числа обращений к внешним источникам информации (FDLP-вывод). Обсуждаются стратегии обратного вывода, использующие несколько параметров релевантности. Рассматриваются вопросы компьютерной реализации FDLP-структур.

Ключевые слова: алгебраическая система, FDLP-структура, продукционно-логическое уравнение, нечеткая распределенная продукционная система, релевантный обратный вывод, компьютерная реализация.

Введение

В последнее десятилетие автор разрабатывает алгебраическую теорию LP-структур (lattice production structures) [1], которая эффективно решает задачи, связанные с интеллектуальными системами

¹Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 19-07-00037.

продукционного типа. Один из важных разделов теории составляет метод релевантного обратного логического вывода (LP-вывод), направленный на снижение числа обращений к внешним источникам информации. Его теоретической основой служит аппарат продукционно-логических уравнений.

Существенной особенностью современных интеллектуальных систем является нечеткий характер знаний и рассуждений. Отсюда возникает задача распространения теории LP-структур на нечеткие продукционные системы. В то же время прогресс информационных технологий приводит к значительному увеличению масштабов и сложности решаемых задач. Компьютерные системы становятся большими и распределенными, что также требует продвижений в методах их построения. В ряде работ (в частности, [4]) было представлено соответствующее обобщение концепции LP-структур — определена алгебраическая модель, охватывающая нечеткие распределенные продукционно-логические системы.

Настоящая работа развивает алгебраическую модель исследованием класса продукционно-логических уравнений в FDLP-структуре. Процесс решения уравнения моделирует нечеткий обратный логический вывод в распределенной интеллектуальной системе. Обсуждаются новые стратегии обратного вывода с учетом нескольких параметров релевантности. Затрагиваются вопросы компьютерной реализации FDLP-структур.

Следует отметить, что известен ряд других интересных алгебраических подходов к формальному построению и исследованию интеллектуальных систем. К таковым, в частности, относится *formal concept analysis* (FCA) [2]. Он так же основан на решетках и рассматривает бинарные отношения между множествами. Однако предмет исследований настоящей работы и представленные здесь результаты не имеют пересечений с направлением FCA и другими известными альтернативами.

1. Основные понятия теории LP-структур

Вводится определение нечеткой LP-структуры. Это атомно-рожденная решетка, на которой задано (вторичное) нечеткое бинарное отношение R . В целях моделирования логического вывода

используется композиция отношений в классической семантике — $(\max\text{-}\min)$ -композиция.

Отношение на решетке называется логическим, если обладает рядом дополнительных свойств: рефлексивность, транзитивность, дистрибутивность. Для произвольного отношения R может быть рассмотрено его логическое замыкание \bar{R} . Алгебраическая система, образованная решеткой и заданным на ней нечетким логическим отношением \bar{R} , называется FLP-структурой (fuzzy lattice production structure).

Рассматриваемый подход к моделированию и исследованию продукционных систем основан на представлении множеств фактов и правил FLP-структурой. Каждый элементарный факт отображается атомом решетки, предпосылка и заключение правила — соответствующими элементами решетки, а правила представляются нечетким бинарным отношением R , то есть его функцией принадлежности. Свойства отношения \bar{R} отражают возможности логического вывода в моделируемой базе знаний.

Далее на FLP-структуре задается отображение в булеан, порождаемый множеством узлов некоторой вычислительной сети. Таким путем формируется расширенная алгебраическая система (FDLP-структура), моделирующая нечеткую распределенную базу знаний. На основе этой системы рассматриваются формальные признаки «хорошего» распределения информации по узлам сети.

2. Продукционно-логические уравнения в FDLP-структуре

Вводится связанный с FDLP-структурами класс уравнений. Правая часть уравнения — это заданный элемент решетки («гипотеза»). Решением называется любой минимальный прообраз гипотезы при отношении \bar{R} , принадлежащий так называемому начальному множеству решетки. Процесс нахождения решения уравнения моделирует обратный логический вывод в распределенной нечеткой базе знаний.

Основной источник трудностей для процесса решения состоит в том, что обычно задано лишь отношение R . Оно в моделируемой продукционной системе соответствует известному множеству продукций — базе знаний. Решение же требуется найти как прообраз гипотезы при отношении \bar{R} — логическом замыкании R . При

этом полное построение логического замыкания нецелесообразно, поскольку на практике требует чрезмерный объем ресурсов. Другим фактором, усложняющим не только метод решения, но и саму постановку задачи, является нечеткость отношения R .

В представленных исследованиях получены результаты о разрешимости уравнений в нечеткой FDLP-структуре, а также обоснованы способы их решения. Приводятся формулировки основных теорем.

Аппарат уравнений в FDLP-структуре создает основу для продвижений в области оптимизации нечеткого распределенного логического вывода и верификации соответствующих баз знаний.

3. О применениях уравнений в FDLP-структурах

Излагаются идеи практического применения аппарата продукционно-логических уравнений для оптимизации нечеткого распределенного обратного вывода (FDLP-вывода) в интеллектуальных системах.

Стратегия направлена на минимизацию количества медленно выполняемых запросов (к базе данных или пользователю). Запросы по возможности должны соответствовать лишь тем фактам, которые действительно необходимы при выводе. Кроме того, при LP-выводе предпочтение отдается тестированию множеств фактов минимальной мощности.

Первая стадия обратного вывода на основе уравнения состоит в его решении — построении всех минимальных начальных прообразов гипотезы в FDLP-структуре. Далее в построенном множестве достаточно найти любой прообраз, атомы которого отображают лишь истинные факты, после чего сразу можно сделать заключение об истинности гипотезы.

Эффективный способ состоит в приоритетном просмотре прообразов, содержащих наиболее «релевантные» атомы решетки. Таковы в первую очередь атомы, присутствующие в максимальном количестве построенных прообразов. Тогда единственный отрицательный ответ на запрос исключает из рассмотрения сразу большое количество прообразов, что соответственно ускоряет исследование. Вторым показателем релевантности атома — его присутствие в прообразах минимальной мощности. Таким образом, предпочтение отдается

тем прообразам, проверка истинности которых потребует меньшего количества вопросов пользователю (или обращений к базе данных).

Нечеткий и распределенный характер базы знаний создает дополнительные возможности (и трудности) для повышения эффективности логического вывода.

К задачам моделирования нечеткой продукционной системы относится достижение для доказываемой гипотезы более высокого значения коэффициента уверенности за приемлемое время. Таким образом, при формировании стратегий нечеткого LP-вывода необходимо учитывать еще один показатель релевантности атомов решетки. Он представляет значение функции принадлежности, вычисляемое для каждого найденного решения уравнения.

Еще одна очевидная цель моделирования интеллектуальной системы рассматриваемого типа — снижение трафика между узлами вычислительной сети в процессе выполнения логического вывода. Стратегия FDLP-вывода, наряду с упомянутыми выше тремя параметрами релевантности, должна использовать дополнительные показатели, связанные с атрибутами хранения фактов и правил на узлах сети.

Простейший вариант алгоритма определения релевантных атомов решетки в процессе FDLP-вывода состоит в элементарном суммировании значений используемых показателей. При этом имеется широкое поле для практических экспериментов, комбинирующих параметры релевантности с «весами», выбираемыми в зависимости от задач в конкретной предметной области. В будущем возможны более глубокие теоретические обоснования стратегий подсчета релевантности атомов, например, на основе методов многокритериальной оптимизации.

Компьютерная реализация FDLP-структур может основываться на представлении решеток битовыми векторами [3]. Разработана программная библиотека `FDLPStructure`, реализующая функциональные возможности новой теории. Эксперименты показывают, что при применении FDLP-вывода снижение числа выполняемых внешних запросов достигает в среднем 20–25%.

Заключение

В представленной работе определен класс уравнений в FDLP-структуре, расширяющей границы применения алгебраической теории до нечетких распределенных интеллектуальных систем продукционного типа.

Установлены теоретические результаты, обосновывающие разрешимость таких уравнений и методы их решения. Процесс решения уравнения моделирует обратный нечеткий логический вывод в распределенной интеллектуальной системе. Таким образом, сформирована теоретическая база для оптимизации логического вывода.

На основе теории предложена концепция FDLP-вывода. Кратко описаны стратегии выбора параметров релевантности.

Определены принципы компьютерной реализации FDLP-структур. Проведены эксперименты, демонстрирующие эффективность теории.

Список литературы

- [1] *Махортов, С. Д.* Математические основы искусственного интеллекта: теория LP-структур для построения и исследования моделей знаний продукционного типа / Под ред. В. А. Васенина. — М. : МЦНМО, 2009. — 304 с.
- [2] *Ferré, S.* Formal Concept Analysis: From Knowledge Discovery to Knowledge Processing / S. Ferré, M. Huchard, M. Kaytoue [et al.] // A Guided Tour of Artificial Intelligence Research / Eds. Marquis P., Papini O., Prade H. — Cham : Springer, 2020. — Vol. 2. — P. 411–435.
- [3] *Halib M.* Bit-vector encoding for partially ordered sets / M. Halib, L. Nourine // Lecture Notes in Computer Science. — 1994. — Vol. 831. — P. 1–12.
- [4] *Makhortov, S. D.* Algebraic Models for Big Data and Knowledge Management // 22nd International Conference on Data Analytics and Management in Data Intensive Domains (DAMDID/RCDL 2020), Selected Proceedings. CCIS. — Cham : Springer, 2021. — Vol. 1427. — P. 19–26.

Библиографическая ссылка

Махортов, С. Д. Основанные на решетках алгебраические системы и их приложения в задачах управления знаниями // Всероссийская научная конференция «Математические основы информатики и информационно-коммуникационных систем». Сборник трудов. — Тверь : ТвГУ, 2021. — С. 50–56.

<https://doi.org/10.26456/mfcsics-21-6>

Сведения об авторах

СЕРГЕЙ ДМИТРИЕВИЧ МАХОРТОВ

Воронежский государственный университет. Заведующий кафедрой

Россия, 394018, Воронеж, Университетская пл., 1, ВГУ

E-mail: msd_exp@outlook.com