

Интеллектуальные системы на основе теории мультиопераций

С. И. Годиков¹, А. В. Показатская²

¹Санкт-Петербургский государственный электротехнический университет «ЛЭТИ» им. В.И. Ульянова (Ленина)

²Национальный исследовательский университет ИТМО

sergeytodikov@gmail.com

Аннотация. В работе рассматривается вопрос о практическом применении теории мультиопераций в области интеллектуальных систем. Приводятся недостатки классических решений и преимущества теории мультиопераций. Описана методика проектирования интеллектуальных систем на основе теории мультиопераций ранга 3. Приведено сравнение интеллектуальной системы на основе мультиопераций с интеллектуальной системой на основе упорядоченной нечеткой логики для задачи многокритериальной оценки поставщиков в условиях неопределенности.

Ключевые слова: мультиоперация, интеллектуальная система, искусственный интеллект, методика

I. ВВЕДЕНИЕ

Последние десятилетия искусственный интеллект развивается интенсивно по самым разным направлениям, а результаты в области искусственного интеллекта реализуются в виде интеллектуальных систем. С самого начала развития интеллектуальных систем предполагалось, что они должны уметь обрабатывать неточные и неопределенные данные.

Под неточными данными понимаются данные, полученные с некоторой погрешностью, шумом или искажением. Под неопределенными данными понимается отсутствие, неполнота или невалидность входных данных. Для представления знаний в интеллектуальных системах создаются модели представления знаний, которые используют различные математические модели.

Для обработки неточности во входных данных в основном используется аппарат нечеткой логики, где по входным данным определяется степень истинности лингвистических переменных. Одним из способов обработки неопределенности во входных данных является использование различных многозначных логик, где неопределенностью является одно логическое значение. Но для обработки неопределенности, которая соответствует множеству логических значений, аппарат классических многозначных логик не предназначен. Другим способом обработки неопределенности во входных данных является доопределение данных каждый раз при их появлении в процессе работы с интеллектуальной системой (дополнительный сбор данных, корректировка условий задачи, добавление новых условий задачи и т.д.). В случае, если время получения результата является важным фактором, то данный способ является неудовлетворительным, так как доопределение входных данных увеличивает время получения конечного результата.

Математическим аппаратом, обладающим встроенным инструментарием для работы с неточными и неопределенными данными, является теория мультиопераций, активно развивающаяся последние десятилетия. Теория мультиопераций может быть использована при обработке неопределенных входных данных, которые соответствуют множеству логических значений.

II. ОСНОВНЫЕ ПОНЯТИЯ

Пусть 2^A – множество всех подмножеств A . Отображение из A^n в 2^A называется n -местной мультиоперацией на A [1]. Под рангом мультиоперации понимается мощность множества A . Множества всех n -местных мультиопераций на A обозначается M_A^n , при $|A| = k$ используется обозначение M_k^n .

Мультиоперацию $f \in M_A^n$ для множества $A = \{a_0, \dots, a_{k-1}\}$ можно представить как отображение:

$$f: \{2^0, 2^1, \dots, 2^{k-1}\} \rightarrow \{0, 2^0, \dots, 2^k - 1\}$$

используя следующую кодировку:

$$a_i \rightarrow 2^i, \emptyset \rightarrow 0, \{a_{i_1}, \dots, a_{i_s}\} \rightarrow 2^{i_1} + \dots + 2^{i_s}$$

Также n -местную мультиоперацию f можно представить в виде вектора всех ее значений (a_1, \dots, a_{k^n}) , где $a_i \in \{0, 1, \dots, 2^{k-1}\}$, $a_i = f(2^{i_1}, \dots, 2^{i_n})$ и (i_1, \dots, i_n) есть представление $i - 1$ в системе счисления по основанию k n -разрядным числом.

Метаоперацию суперпозиция на множестве мультиопераций определяется следующим образом: если $f \in M_k^n$ и $f_1, \dots, f_n \in M_k^m$, то

$$s(f_0, f_1, \dots, f_n)(a_1, \dots, a_m) = \bigcup_{b_i \in f_i(a_{i_1}, \dots, a_{i_m})} f_0(b_1, \dots, b_n)$$

Терм с множеством неизвестных X и множеством констант K над множеством мультиопераций $F \subseteq M_k^n$ определяется следующим образом:

1. если $y \in K \cup X$, тогда $t \equiv y$ — это терм и $U(t) = \{y\}$;
2. если $f \in F^n \cup X$ и t_1, \dots, t_n являются термами, то $t \equiv f(t_1, \dots, t_n)$ — это терм и $U(t) = U(t_1) \cup \dots \cup U(t_n)$;
3. если t_0 — терм, $U(t_0) = \{y_1, \dots, y_n\}$ и t_1, \dots, t_n являются термами, то $t \equiv t_0(t_1, \dots, t_n)$ — это терм и $U(t) = U(t_1) \cup \dots \cup U(t_n)$.

Общая формула системы включений с неизвестными в теории мультиопераций имеет следующий вид: [2]

$$\begin{cases} t_1(\tilde{c}, \tilde{z}) \subseteq q_1(\tilde{c}, \tilde{z}) \\ \dots \dots \dots \\ t_m(\tilde{c}, \tilde{z}) \subseteq q_m(\tilde{c}, \tilde{z}) \end{cases}$$

где t_i, q_i термы с неизвестными \tilde{z} и константами \tilde{c} в $F_n \subseteq M_A^k$.

Решением системы включений является набор мультиопераций f_1, \dots, f_s таких, что все i выполняют включение при стандартном определении значения термов с помощью введенного оператора суперпозиции:

$$t_i(\tilde{c}, f_1(\tilde{c}), \dots, f_s(\tilde{c})) \subseteq q_i(\tilde{c}, f_1(\tilde{c}), \dots, f_s(\tilde{c})).$$

III. МЕТОДИКА ПРОЕКТИРОВАНИЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ НА ЛОГИКЕ МУЛЬТИОПЕРАЦИЙ

Идея интеллектуальной системы на основе теории мультиопераций заключается во взаимной работе нескольких источников данных, причем источники могут противоречить друг другу. Примером подобной работы может быть человек эксперт по практической области и ряд измерительных устройств или системы ИИ для определения значений отдельных входных данных. При неопределенности или неточности части входных данных, такая система формирует промежуточное решение, в котором указывает какие входные данные необходимо доопределить и сужает область поиска результата. Такой подход позволяет уменьшить время получения конечного.

В работах [3, 4] описывается концепция медицинской интеллектуальной системы на основе теории мультиопераций для определения области заболевания пациента, в работе [5] описывается техническая реализация системы. В текущей работе описан общий подход для построения интеллектуальной системы на основе теории мультиоперации.

Для создания базы знаний используются системы включений, где каждое включения является отдельным правилом интеллектуальной системы.

Методика для проектирования интеллектуальных систем включает следующие шаги:

1. Определить входные и выходные переменные.
2. Определить модели представления знаний для входных и выходных переменных.
3. Определить мультиоперации для построения правил на основе систем включений.
4. Задать базу знаний на основе систем включений.

Использование систем включений дает еще одно преимущество. После формирования базы знаний можно найти общее решение системы включений и провести анализ общего решения на различные условия и свойства верификации баз знаний, например на целостность, непротиворечивость, полноту и отсутствие избыточности. В работе [2] дается алгоритм для решения систем включений, в работе [6] описывается ПО для автоматизации решения систем включений. В работе [3] показан пример анализа общего решения. На текущий момент верификация путем решения систем включений является опциональным этапом, так как процесс поиска общего решения является достаточно сложным и требует большие вычислительные ресурсы.

После формирования базы знаний можно переходить к процессу работы с интеллектуальной системой.

IV. ПРИМЕНЕНИЕ МЕТОДИКИ

Применение методики рассмотрим на задаче многокритериальной оценки поставщиков в условиях неопределенности. Для сравнения выбрана система, описанная в статье [7], которая учитывает не только текущие значения критериев, но и направление их изменения во времени, а также вычисляет энтропию как меру неопределенности результата. Для сравнения двух интеллектуальных систем предметная область, набор критериев и содержательная база знаний приняты одинаковыми.

Сравниваемая система построена как нечеткая система вывода типа Мамдани на основе упорядоченных нечетких чисел (OFN). База знаний включает 216 правил вида IF-THEN [6], а входные и выходные лингвистические переменные задаются треугольными OFN с положительной и отрицательной ориентацией, что позволяет учитывать тренд изменения оценки. Процедура вывода включает фазификацию входов, вычисление силы правила, направленную импликацию DIMS, дефазификацию и вычисление энтропии по множеству активированных правил.

В статье [6] используются три входных критерия: доля локально нанятого персонала, полнота поставки и доля дефектных изделий. Для каждой выходной переменной задаются три уровня оценки — высокий, средний и низкий, а также два типа тренда — увеличивающийся и уменьшающийся. Таким образом, каждая переменная может принимать шесть лингвистических состояний, учитывающих уровень показателя и направление его изменения.

Для критерия «доля локально нанятого персонала» на диапазоне [50%; 100%] используются термы: низкая – [49,9, 50, 50, 75], средняя – [50, 75, 75, 100], высокая – [75, 100, 100, 100, 1] для увеличивающегося тренда; для уменьшающегося тренда используются те же опорные точки в обратном порядке. Для критерия «полнота поставки» на диапазоне [70%; 100%] используются наборы [69,9, 70, 70, 90], [70, 90, 90, 100], [90, 100, 100, 100, 1]. Для критерия «дефекты» как критерия затрат на диапазоне [0%; 30%] используются наборы [–0,1, 0, 0, 5], [0, 10, 10, 20], [20, 30, 30, 30, 1]. Для итоговой оценки поставщика на диапазоне [0; 100] заданы выходные термы [–25, 0, 0, 50], [25, 50, 50, 75], [50, 100, 100, 125] и их отрицательно ориентированные аналоги.

Система на основе теории мультиопераций использует базу правил из [6], формируя на их основе системы включений, а значения термов интерпретируются на языке мультиопераций. Это обеспечивает корректность дальнейшего сравнения результатов работы двух систем на одних и тех же входных данных.

В системе каждая переменная рассматривается как элемент конечного терм-множества, включающего комбинации уровня и направления изменения. Для переменной «Наём местных жителей» l вводится терм-множество $T_l = \{\text{высокий}\uparrow, \text{средний}\uparrow, \text{низкий}\uparrow, \text{высокий}\downarrow, \text{средний}\downarrow, \text{низкий}\downarrow\}$ или в символьном виде: $T_l = \{l_1, l_2, l_3, l_4, l_5, l_6\}$. Для переменной «Полнота заказа» w вводится $T_w = \{\text{высокий}\uparrow, \text{средний}\uparrow, \text{низкий}\uparrow,$

высокий↓, средний↓, низкий↓} или $T_w = \{w_1, w_2, w_3, w_4, w_5, w_6\}$. Для переменной «Дефекты» a вводится $T_a = \{\text{высокий}\uparrow, \text{средний}\uparrow, \text{низкий}\uparrow, \text{высокий}\downarrow, \text{средний}\downarrow, \text{низкий}\downarrow\}$ или $T_a = \{a_1, a_2, a_3, a_4, a_5, a_6\}$. Символ \uparrow означает увеличивающийся тренд, символ \downarrow — уменьшающийся.

Для построения системы на основе мультиопераций термы задаются отдельно, а неточность и неопределённость входных данных учитываются через модель представления знаний на языке мультиопераций.

Модель представления знаний для входных переменных

- 0 — \emptyset : ни один источник данных не квалифицирован для данного терма;
- 1 — {1}: терм ложен;
- 2 — {2}: терм истинен;
- 3 — {1,2}: терм не определён (противоречие между источниками);
- 4 — {4}: один из источников не дал значение / дал невалидное значение;
- 5 — {1,4}: часть источников недоступна/невалидна, но по доступным терм ложен;
- 6 — {2,4}: часть источников недоступна/невалидна, но по доступным терм истинен;
- 7 — {1,2,4}: необходимы уточнения.

Модель выходных переменных

- 0 — \emptyset : ни один источник данных не квалифицирован для терма;
- 1 — {1}: терм ложен;
- 2 — {2}: терм истинен;
- 3 — {1,2}: терм не определён из-за противоречия;
- 4 — {4}: невозможно определить терм, т.к. не по всем входам есть информация;
- 5 — {1,4}: по части источников делается вывод «ложно» (не по всем);
- 6 — {2,4}: по части источников делается вывод «истинно» (не по всем);
- 7 — {1,2,4}: необходимы уточнения.

Рассмотрим ситуацию, в которой значения всех входных параметров определены однозначно. Для анализа используются данные за шесть последовательных месяцев. Значения показателя «Наём местных жителей» составляют 72, 71, 80, 88, 88 и 90 процентов. Значения показателя «Полнота заказа» равны 80, 85, 90, 84, 90, 91. Значения показателя «Дефекты» равны 15, 10, 10, 12, 9, 4 процентам. Направление изменения показателя определяют по коэффициенту корреляции Пирсона между рядом значений критерия x и временем t : при значении выше -0.2 он считается

положительным, при значении -0.2 и ниже — отрицательным.

В рассматриваемом примере для показателя «Наём местных жителей» коэффициент корреляции составляет $r_{l,t} \approx 0.939 \Rightarrow r_{l,t} > -0.2 \Rightarrow l \uparrow$. Для показателя «Полнота заказа» коэффициент корреляции равен $r_{w,t} \approx 0.783 \Rightarrow r_{w,t} > -0.2 \Rightarrow w \uparrow$. Для показателя «Дефекты» коэффициент корреляции составляет $r_{a,t} \approx -0.824 \Rightarrow r_{a,t} \leq -0.2 \Rightarrow a \downarrow$. При этом критерий «дефекты» рассматривается как затратный, что означает ухудшение оценки при увеличении значения показателя. Это учитывается в системе через константу $\varepsilon = -1$.

После определения направления изменения параметров определяется уровень каждого показателя по значению последнего месяца. Для показателя «Наём местных жителей» последнее наблюдение равно 90%. Значение 90% находится в зоне пересечения среднего и высокого уровней, поэтому его можно отнести сразу к двум термам: $M\uparrow$ и $H\uparrow$. Аналогично показатель «Полнота заказа» со значением 91% также попадает в область перекрытия среднего и высокого уровней, что соответствует термам $M\uparrow$ и $H\uparrow$. Для показателя «Дефекты» последнее значение равно 4% и находится в зоне пересечения низкого и среднего уровней; с учётом убывающего тренда ему соответствуют термы $L\downarrow$ и $M\downarrow$.

По результатам фазификации каждой входной переменной ставятся в соответствие два допустимых терма. Таким образом, формируются восемь возможных комбинаций, представленные в табл. I, каждая из которых соотносится с отдельным правилом базы знаний.

ТАБЛИЦА I. РЕЗУЛЬТАТЫ ВЫВОДА ПРИ ПОЛНЫХ ВХОДНЫХ ДАННЫХ

Наём местных жителей	Полнота заказа	Дефекты	Результат
$L\downarrow$	$M\uparrow$	$M\uparrow$	$M\uparrow$
$M\downarrow$	$M\uparrow$	$M\uparrow$	$M\uparrow$
$L\downarrow$	$H\uparrow$	$M\uparrow$	$H\uparrow$
$M\downarrow$	$H\uparrow$	$M\uparrow$	$M\uparrow$
$L\downarrow$	$M\uparrow$	$H\uparrow$	$H\uparrow$
$M\downarrow$	$M\uparrow$	$H\uparrow$	$H\uparrow$
$L\downarrow$	$H\uparrow$	$H\uparrow$	$M\uparrow$
$M\downarrow$	$H\uparrow$	$H\uparrow$	$H\uparrow$

При полностью определённых входных данных активируется набор правил, приводящий к выходным термам $M\uparrow$ и $H\uparrow$; по составу активированных правил результаты системы на основе теории мультиопераций и системы на основе упорядоченных нечетких чисел эквивалентны.

Рассмотрим ситуацию, когда часть входной информации отсутствует. Показатель «Наём местных жителей» задан последовательностью значений 72, 71, 68, 73, 68 и 67. Показатель «Полнота заказа» характеризуется значениями 91, 85, 82, 84, 83 и 89. Для показателя «Дефекты» исторические данные представлены неполностью: первые три наблюдения отсутствуют, а для последующих моментов времени значения составляют 10, 8 и 12%.

По результатам фазификации для рассматриваемых параметров получены следующие термы: для показателя «Наём местных жителей» — $H\downarrow, M\downarrow, L\downarrow$; для показателя «Полнота заказа» — $M\downarrow, L\downarrow$; для показателя

«Дефекты» — М. Поскольку исторические данные по последнему показателю заданы не полностью, при дальнейшем выводе рассматриваются правила как с направлением ↑, так и с направлением ↓.

В системе упорядоченного нечёткого вывода построение результата в такой ситуации становится невозможным. Причина заключается в том, что процедура фазсификации требует точных входных значений, позволяющих определить степени принадлежности соответствующим лингвистическим термам. При отсутствии части наблюдений корректное построение входных OFN-значений становится невозможным, поэтому система не формирует итоговый вывод.

В системе, основанной на теории мультиопераций, отсутствие данных рассматривается как отдельное допустимое состояние переменной и кодируется значением {4}, означающим невозможность определить соответствующий терм. Вместо прекращения вычислений система формирует множество правил базы знаний, представленных в табл. II, которые не противоречат имеющейся информации.

ТАБЛИЦА II. РЕЗУЛЬТАТЫ ВЫВОДА ПРИ НЕПОЛНЫХ ВХОДНЫХ ДАННЫХ

<i>Наём местных жителей</i>	<i>Полнота заказа</i>	<i>Дефекты</i>	<i>Результат</i>
L↓	L↓	M↑↓	L↓
L↓	M↓	M↑↓	M↓
M↓	L↓	M↑↓	M↓
M↓	M↓	M↑↓	M↓
H↓	L↓	M↑↓	M↓
H↓	M↓	M↑↓	M↓

Отсутствие данных не даёт однозначно определить единственный итоговый вывод, однако позволяет сузить множество допустимых решений. В этом случае система формирует перечень всех возможных результатов для заданного поставщика. Анализ рассматриваемого примера показывает, что неопределённость направления одного из параметров не влияет на итоговый результат. Следовательно, при оценке поставщика допустима частичная неопределённость исходных данных на рассматриваемом временном интервале.

V. ЗАКЛЮЧЕНИЕ

Используя данные со сравнения двух систем, можно сделать вывод, что теория мультиопераций позволяет строить интеллектуальные системы, которые сохраняют работоспособность в условиях неточности и неопределенности во входных данных. Такой подход позволяет сократить количество входных данных, которые необходимо доопределить, на основе полученных промежуточных решений, а также сузить область поиска решения. Данные преимущества могут уменьшить время получения конечного результата в условиях неточности и неопределенности во входных данных.

СПИСОК ЛИТЕРАТУРЫ

- [1] Перязев Н.А. Тожества в алгебрах мультиопераций фиксированной размерности // Известия Иркутского государственного университета. Серия Математика, 2019, том 29, С. 86–97.
- [2] Peryazev N.A. Systems of Inclusions with Unknowns in Multioperations // Bulletin of Irkutsk State University – Series Mathematics. 2022. V. 38. P. 112-123.
- [3] Todikov S.I., Shichkina Y.A., Peryazev N.A. Applying the Theory of Multi-Operations to Building Decision-Making Systems with a Large Number of Uncertainties // Mathematics 2024. Vol. 12. P. 3694.
- [4] Тодиков С.И. Возможность применения теории мультиопераций для построения самообучающихся систем искусственного интеллекта со свойством объяснимости // Материалы IV международной конференции по нейронным сетям и нейротехнологиям (NeuroNT'23). 2023. С. 105-108.
- [5] Тодиков С.И., Показаккая А.В. Web-приложение для тестирования интеллектуальной системы на основе теории мультиопераций // Компьютерные инструменты в образовании. 2024. No 1. С. 58–70. doi:10.32603/2071-2340-2024-1-58-70.
- [6] Тодиков С.И. Автоматизация процесса решения систем включений с неизвестными в теории мультиопераций // Материалы XVII Международной конференции «Синтаксис и семантика логических систем». 2022. С. 53
- [7] Rudnik K., Chwastyk A., Pisz I. Approach Based on the Ordered Fuzzy Decision Making System Dedicated to Supplier Evaluation in Supply Chain Management // Entropy. 2024. Vol. 26, No. 10. Art. 860. DOI: 10.3390/e26100860.