

УДК 66.01-52

# МУЛЬТИМНОЖЕСТВЕННЫЕ ГРАММАТИКИ КАК БАЗОВАЯ МОДЕЛЬ ПРЕДСТАВЛЕНИЯ ЗНАНИЙ ДЛЯ ИНТЕЛЛЕКТУАЛЬНЫХ СИСТЕМ ИНЖИНИРИНГА ХИМИЧЕСКИХ РЕАКЦИЙ

© 2024 г. Академик РАН И. А. Шеремет<sup>1,\*</sup>

Поступило 15.06.2024 г.  
После доработки 09.07.2024 г.  
Принято к публикации 11.07.2024 г.

В работе предлагается использовать математический аппарат мультимножественных грамматик в качестве модели представления знаний для интеллектуальных систем анализа, оптимизации и проектирования сложных химических реакций. Это позволит интегрировать базы знаний об указанных реакциях с базами знаний интеллектуальных систем программно-целевого планирования в рамках базового стека цифровой экономики и на этой основе организовать целостные процессы управления функционированием промышленных кластеров, охватывающие как сборочные, так и химические производства.

*Ключевые слова:* мультимножество, мультимножественная грамматика, химическая реакция, маршрут сложной химической реакции

DOI: 10.31857/S2686953524040031, EDN: YEALJU

## 1. ВВЕДЕНИЕ

Целью создания интеллектуальных систем инжиниринга (ИСИ) сложных химических реакций (СХР) является решение как минимум трех основных задач:

- каким будет результат смешения заданных химических веществ (ХВ) в заданных условиях;
- как синтезировать заданное химическое вещество из доступных реагентов;
- как построить оптимальный по определенному критерию (системе критериев) маршрут СХР, посредством которого может быть реализован необходимый химико-технологический процесс.

Для решения перечисленных задач традиционно применяются фреймово-продукционные модели [1–3] и различные проблемно-ориентированные разновидности графов: потоковые (в том числе эксергетические потоковые), информационные, сигнальные, а также информационно-потоковые мультиграфы [4–7]. В частности, весьма распространенным средством решения указанных задач являются двудольные мультиграфы, обеспечивающие представление

систем уравнений химических реакций и именуемые графами Вольперта. Однако графовые модели имеют свои ограничения. Во-первых, графы охватывают только топологические аспекты химико-технологических процессов, тогда как для представления цепочек многопараметрических химических превращений, осуществляемых в рамках химико-технологического процесса, необходимо привлечение дополнительных языково-алгоритмических инструментов. Во-вторых, требуется выполнение значительного объема работы по предварительному построению графов, отображающих взаимосвязи между уравнениями различных химических реакций, и созданию алгоритмов и программ применения этих графов для решения различных задач. В этой связи возникает необходимость поиска некоторой альтернативной единой модели представления знаний о химических реакциях, которая минимизировала бы:

- трудоемкость создания и пополнения общей базы знаний о химических реакциях;
- временные затраты на освоение и использование ИСИ конечными пользователями (учеными и инженерами химической отрасли);
- вычислительную сложность решения задач применительно к очень большим объемам используемых баз знаний.

<sup>1</sup> Геофизический центр Российской академии наук, 119296 Москва, Россия

\*E-mail: sheremet@rfbr.ru

Для выполнения этих требований искомая модель представления знаний должна быть максимально совместима как с традиционным для химии представлением знаний в виде уравнений химических реакций, так и с общепринятым языком постановки многокритериальных оптимизационных задач.

Данная статья посвящена рассмотрению возможных способов применения мультимножественных грамматик (МГ), которые до настоящего времени использовались для решения теоретических и практических задач из областей системного анализа, исследования операций и цифровой экономики [8, 9], в качестве единой модели представления знаний для ИСИ СХР.

Во втором разделе описаны базовые математические конструкции, лежащие в основе мультиграмматического инструментария (МГИ) и предлагаемых решений упомянутых задач. Третий раздел посвящен способам использования различных классов мультимножественных грамматик для представления различных расширений химических уравнений и решению задачи построения оптимальных маршрутов сложных химических реакций.

## 2. МУЛЬТИГРАММАТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ХИМИЧЕСКИХ РЕАКЦИЙ

### 2.1. Исходные положения

Предлагаемый подход к мультиграмматическому представлению и моделированию химических реакций основывается на следующих предположениях:

– любое химическое вещество может быть представлено в виде мультимножества (ММ), объектами которого являются молекулы, а кратность каждого объекта в этом ММ есть количество идентичных молекул соответствующего вида в данном ХВ;

– уравнение любой химической реакции может быть представлено в виде соответствующего правила мультимножественной грамматики, в котором мультиобъекты из левой части представляют реагенты, а мультиобъекты из правой части – продукты реакции.

При этом применение некоторого конкретного правила к мультимножеству, получен-

ному на предшествующих шагах порождения, моделирует протекание соответствующей химической реакции, продуктами которой являются ХВ, представленные полученным мультимножеством. При таком подходе схема МГ представляет собой базу знаний о химических реакциях, имеющих отношение к рассматриваемому классу ХВ, а множество мультимножеств, порождаемых этой МГ из исходного ММ, представляющего исходное ХВ, есть не что иное, как множество результатов всех возможных химических реакций, которые в любых возможных последовательностях могут произойти, начиная от указанного исходного ХВ.

Перейдем к рассмотрению математической основы предлагаемого подхода.

Согласно работам [8, 9], мультимножество представляет собой совокупность мультиобъектов, так что запись

$$v = \{n_1 \cdot a_1, \dots, n_m \cdot a_m\} \quad (1)$$

означает, что ММ  $v$  содержит  $n_1$  объектов  $a_1, \dots, n_m$  объектов  $a_m$ ; целое число  $n_i$  именуется кратностью объекта  $a_i$ , а  $n_i \cdot a_i$  – мультиобъектом.

Как уже отмечалось, в рассматриваемом приложении объектами будут формулы молекул, а кратностями – количества этих молекул в ХВ, которое представляется в виде мультимножества. Синтаксическая структура объектов, то есть используемая запись формул химических соединений, может быть как общепринятой многоуровневой, так и эквивалентной ей линейной. В данной статье мы будем использовать линейную нотацию, смысл которой иллюстрируется примером 1.

*Пример 1.* Пусть имеется химическое вещество, состоящее из 8 молекул нитрата таллия  $\text{Pb}(\text{NO}_3)_3$  и 5 молекул сульфата ванадия  $\text{V}_2(\text{SO}_4)_3$ . Это ХВ представляется мультимножествами

$$v = \{8 \cdot \text{Pb}(\text{NO}_3)_3, 5 \cdot \text{V}_2(\text{SO}_4)_3\}$$

при использовании традиционной многоуровневой записи формул химических соединений и

$$v = \{8 \cdot \text{Pb}(\text{NO}_3)_3, 5 \cdot \text{V}_2(\text{SO}_4)_3\}$$

при использовании линейной нотации.

Уравнение химической реакции будет представляться в виде правила

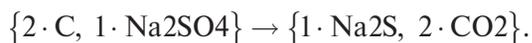


в котором  $v$  есть мультимножество, представляющее реагенты, а  $v'$  – продукты этой реакции.

*Пример 2.* В соответствии с (2) уравнение химической реакции



представляется в виде правила:



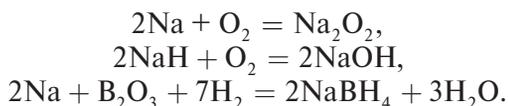
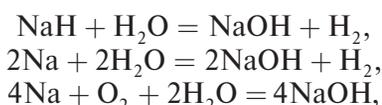
## 2.2. Прямая задача

Пусть имеется совокупность реагентов, представляемых мультимножествами  $v_0^1, \dots, v_0^k$ , и база знаний о химических реакциях в виде множества  $R$  представляющих их правил. Тогда множество возможных результатов цепочек химических превращений (сложных химических реакций), каждая из которых завершается получением своей совокупности продуктов, есть не что иное, как множество мультимножеств  $V_S$ , порождаемых мультимножественной грамматикой  $S = \langle v_0, R \rangle$ , где ядро

$$v_0 = v_0^1 + \dots + v_0^k. \quad (3)$$

При этом множество  $V_S$  включает мультимножественные представления всех химических веществ (в том числе промежуточных), получаемых в результате цепочек химических превращений, тогда как множество  $\bar{V}_S$  терминальных мультимножеств (ТММ), порождаемых этой МГ, включает только те мультимножества, которые представляют конечные результаты упомянутых цепочек (в общем случае  $\bar{V}_S$  включает более одного мультимножества, что соответствует возможности реализации различных цепочек).

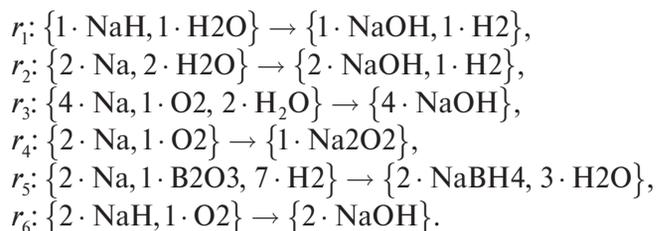
*Пример 3.* Пусть имеем три реагента, первый из которых включает восемь атомов натрия, второй – две молекулы  $B_2O_3$ , а третий – 15 молекул водорода. При этом база знаний о химических реакциях включает следующие реакции:



Мультимножественная грамматика  $S = \langle v_0, R \rangle$ , соответствующая этим исходным данным в указанном выше смысле, имеет следующие ядро и схему:

$$\begin{aligned} v_0 &= \{8 \cdot Na\} + \{2 \cdot B_2O_3\} + \{15 \cdot H_2\} = \\ &= \{8 \cdot Na, 2 \cdot B_2O_3, 15 \cdot H_2\}, \\ R &= \{r_1, r_2, r_3, r_4, r_5, r_6\}, \end{aligned}$$

где правила  $r_1, \dots, r_6$  имеют, в соответствии с выражением (2), следующий вид (уникальный идентификатор правила отделяется от правила двоеточием):



Порождение мультимножеств посредством применения правил осуществляется следующим образом:

$$\begin{aligned} v_0 &\stackrel{r_5}{\Rightarrow} \{8 \cdot Na, 2 \cdot B_2O_3, 15 \cdot H_2\} \\ &\quad - \{2 \cdot Na, 1 \cdot B_2O_3, 7 \cdot H_2\} \\ &\quad + \{2 \cdot NaBH_4, 3 \cdot H_2O\} = \\ &\{6 \cdot Na, 1 \cdot B_2O_3, 8 \cdot H_2, 2 \cdot NaBH_4, 3 \cdot H_2O\} \stackrel{r_5}{\Rightarrow} \\ &\{6 \cdot Na, 1 \cdot B_2O_3, 8 \cdot H_2, 2 \cdot NaBH_4, 3 \cdot H_2O\} \\ &\quad - \{2 \cdot Na, 1 \cdot B_2O_3, 7 \cdot H_2\} \\ &\quad + \{2 \cdot NaBH_4, 3 \cdot H_2O\} = \\ &\{4 \cdot Na, 1 \cdot H_2, 4 \cdot NaBH_4, 6 \cdot H_2O\} \stackrel{r_2}{\Rightarrow} \\ &\{4 \cdot Na, 1 \cdot H_2, 4 \cdot NaBH_4, 6 \cdot H_2O\} \\ &\quad - \{2 \cdot Na, 2 \cdot H_2O\} \\ &\quad + \{2 \cdot NaOH, 1 \cdot H_2\} = \\ &\{2 \cdot Na, 2 \cdot H_2, 4 \cdot NaBH_4, 4 \cdot H_2O, 2 \cdot NaOH\} \stackrel{r_2}{\Rightarrow} \\ &\{2 \cdot Na, 2 \cdot H_2, 4 \cdot NaBH_4, 4 \cdot H_2O, 2 \cdot NaOH\} \\ &\quad - \{2 \cdot Na, 2 \cdot H_2O\} \\ &\quad + \{2 \cdot NaOH, 1 \cdot H_2\} = \\ &\{3 \cdot H_2, 4 \cdot NaBH_4, 2 \cdot H_2O, 4 \cdot NaOH\}. \end{aligned}$$

Таким образом, множество терминальных мультимножеств, порождаемых мультимножествен-

ной грамматикой  $S$ , включает единственный элемент, представляющий химическое вещество, состоящее из трех молекул водорода, двух молекул воды, четырех молекул гидроксида натрия и четырех молекул  $\text{NaNH}_4$ :

$$\bar{V}_S = \{ \{ 3 \cdot \text{H}_2, 4 \cdot \text{NaNH}_4, 2 \cdot \text{H}_2\text{O}, 4 \cdot \text{NaOH} \} \}.$$

Такова базовая схема применения мультиграмматического инструментария для представления химических реакций и получения их продуктов. Эта схема без каких-либо дополнений обеспечивает решение первой из упомянутых во введении задач: **что получится в результате реакции, если смешать заданные химические вещества.**

### 2.3. Обратная задача

Рассмотрим логику применения МГИ для решения второй (обратной) задачи: **как синтезировать заданное химическое вещество из доступных реагентов.**

В этих целях построим мультимножественную грамматику  $S = \langle v_0, R \rangle$ , где ядро  $v_0$ , как и в прямой задаче, есть сумма мультимножеств, представляющих все доступные реагенты, а схема  $R$  есть база знаний, то есть множество правил, представляющих возможные химические реакции. Пусть химическое вещество, которое необходимо получить, представляется мультимножеством  $v = \{ n_1 \cdot a_1, \dots, n_m \cdot a_m \}$ , где, как и ранее,  $n_i$  есть количество молекул химического соединения  $a_i$  в ХВ  $v$ . Поскольку  $\bar{V}_S$  представляет множество ХВ, которые могут быть получены из  $v_0$ , то критерий для оценки возможности решения поставленной задачи очевиден: если  $v \in \bar{V}_S$ , то **задача разрешима**; в ином случае **получить требуемое вещество из имеющихся реагентов невозможно**. Разрешимость поставленной задачи означает, что существует последовательность непосредственных порождений мультимножеств

$$v_0 \xRightarrow{r_0} v_1 \xRightarrow{r_1} \dots \xRightarrow{r_i} v, \quad (4)$$

которая представляет последовательность соответствующих химических реакций, обеспечивающую получение требуемого вещества  $v$  из реагентов  $v_0^1, \dots, v_0^k$ , где  $v_0^1 + \dots + v_0^k = v_0$ . Таким образом,  $l$ -местный кортеж

$$\rho(v_0, v) = \langle r_{i_0}, r_{i_1}, \dots, r_{i_l} \rangle \quad (5)$$

есть не что иное, как представление маршрута сложной химической реакции, представлением совокупности реагентов которой является мультимножество  $v_0$ , а представлением совокупности ее продуктов — мультимножество  $v$ .

Данная задача может иметь и более сложную постановку, согласно которой **требуется получить ХВ**, представленное мультимножеством  $v$ , не из всей совокупности реагентов, представленной ММ  $v_0$ , а **из некоторого подмножества этих реагентов**, предварительно изъяв его и переместив в область реакции. В случае, если возможно несколько вариантов решения этой задачи, следует выбрать наилучший по какому-либо критерию.

Построим так называемую **зеркальную (mirror)** [9] мультимножественную грамматику  $S^{-1} = \langle v, R^{-1} \rangle$ , ядро которой есть мультимножественное представление ХВ, которое необходимо получить, а схема включает так называемые зеркальные (**инвертированные**) правила, полученные из правил, входящих во множество  $R$  переменной мест левой и правой частей:

$$R^{-1} = \{ v' \rightarrow v \mid v \rightarrow v' \in R \}. \quad (6)$$

Как видно, мультиграмматика  $S^{-1}$  обеспечивает порождение множеств  $V_{S^{-1}}$  мультимножеств, являющихся представлениями всех химических веществ, из которых в результате последовательностей химических реакций может быть получено вещество, представляемое мультимножеством  $v$ . Заметим, что в данном случае представлением множества **всех** ХВ, из которых может быть получено ХВ  $v$ , является именно множество **всех** ММ, порождаемых МГ  $S^{-1}$ , а не множество терминальных мультимножеств  $\bar{V}_{S^{-1}}$ .

Для получения множества искоемых ММ достаточно построить фильтрующую мультиграмматику  $S = \langle v, R^{-1}, \{ \subseteq v_0 \} \rangle$  с фильтром, включающим единственное граничное условие  $\subseteq v_0$ , которое обеспечивает отбор из множества  $V_{S^{-1}}$  таких мультимножеств  $v \in V_{S^{-1}}$ , которые являются подмультимножествами мультимножества  $v_0$ , то есть  $v \subseteq v_0$ , что означает принадлежность этих ММ множеству решений поставленной задачи. В случае, когда  $V_{S^{-1}}$  содержит более одного ММ, то есть  $|V_{S^{-1}}| > 1$ , достаточно использовать фильтрующую мультиграмматику  $S' = \langle v, R^{-1}, \{ \subseteq v_0 \} \cup F' \rangle$ , где  $F'$  есть множество граничных и оптимизирующих условий, обеспечивающих селекцию из множества  $V_S$  единственного решения, то есть  $|V_{S'}| = 1$ .

Таким образом в рамках базового представления обеспечивается решение первых двух задач, вербально сформулированных во введении. Для решения третьей из упомянутых задач, то есть построения оптимального маршрута сложной химической реакции, предварительно рассмотрим, каким образом дополнительная информация, необходимая для сравнения альтернативных маршрутов, может быть имплантирована в упомянутое базовое представление.

### 3. МУЛЬТИГРАММАТИЧЕСКОЕ ПРЕДСТАВЛЕНИЕ ХИМИЧЕСКИХ РЕАКЦИЙ С ДОПОЛНИТЕЛЬНОЙ ИНФОРМАЦИЕЙ И ПОСТРОЕНИЕ ОПТИМАЛЬНЫХ МАРШРУТОВ СЛОЖНЫХ ХИМИЧЕСКИХ РЕАКЦИЙ

#### 3.1. Стоимость синтеза вещества

В общем случае возможно существование нескольких последовательностей реакций, каждая из которых соответствует своему способу получения вещества  $v$ . При этом актуальной становится задача выбора наилучшего из возможных вариантов по некоторому заранее определенному показателю. Таким показателем в общем случае может быть **полная стоимость синтеза требуемого вещества из имеющихся реагентов**. В качестве исходных данных для вычисления значения этого показателя могут быть использованы стоимости химических реакций, мультимножественные представления уравнений которых имеются в базе знаний. Чтобы имплантировать эти стоимости в базу знаний, достаточно каждому правилу  $v \rightarrow v'$ , представляющему некоторое уравнение реакции, поставить в соответствие правило  $v \rightarrow v' + \{m \cdot c\}$ , где  $c$  есть обозначение денежной единицы, которая используется для исчисления стоимостей реакций и их цепочек, а  $m$  есть выраженная в единицах  $c$  стоимость реакции, представленной правилом  $v \rightarrow v'$ . Полученная таким образом мультиграмматика  $S_c = \langle v_0, R_c \rangle$ , где

$$R_c = \bigcup_{\langle r; v \rightarrow v' \rangle \in R} \{ \langle r; v \rightarrow v' + \{m_r \cdot c\} \rangle \}, \quad (7)$$

а  $m_r$  есть стоимость реакции, уравнение которой представляется правилом с именем  $r$ , обеспечивает порождение множества мультимножеств вида

$$\bar{v} + M \cdot c,$$

где, в свою очередь,  $M$  есть полная стоимость получения совокупности продуктов, представленной мультимножеством  $\bar{v}$ , из совокупности реагентов, представленных мультимножеством  $v_0$ , способом, представленным последовательностью правил  $r_{i_0}, r_{i_1}, \dots, r_{i_l}$ . Как видно,  $M$  есть сумма стоимостей всех реакций, последовательность которых обеспечивает получение  $\bar{v}$ :

$$v_0 \Rightarrow^{r_{i_0}} v_1 \Rightarrow^{r_{i_1}} \dots \Rightarrow^{r_{i_l}} \bar{v} + \{M \cdot c\}, \quad (8)$$

$$M = \sum_{j=1}^l m_{r_{i_j}}. \quad (9)$$

При этом для выбора способа с наименьшей стоимостью достаточно использовать фильтрующую мультиграмматику  $S_c = \langle v_0, R_c, F \rangle$  с терминальным фильтром  $F = \{c = \min\}$ .

Приведенная формализация не учитывает стоимости реагентов, которые необходимо приобрести для синтеза требуемого вещества. Указанная стоимость может быть включена в схему  $R_c$  следующим образом. Каждому реагенту  $k \cdot a$ , стоимость одной молекулы  $a$  которого составляет  $n_a$  единиц  $c$ , поставим в соответствие правило

$$\langle r_a; \{1 \cdot \#a\} \rightarrow \{1 \cdot a, n_a \cdot c\} \rangle, \quad (10)$$

где  $\#a$  – вспомогательный объект, полученный присоединением к лексеме  $a$  префикса “#”. Далее ядро  $v_0$  МГ  $S_c$  заменим на

$$v'_0 = \bigcup_{k \cdot a \in v_0} \{k \cdot \#a\}, \quad (11)$$

а схему  $R_c$  – на

$$R'_c = R_c \cup \left( \bigcup_{k \cdot a \in v_0} \{ \langle r_{\#a}; \{1 \cdot \#a\} \rightarrow \{1 \cdot a, n_a \cdot c\} \rangle \} \right), \quad (12)$$

в результате чего все элементы множества ТММ, порождаемых МГ

$$S'_c = v'_0, R'_c, \quad (13)$$

будут содержать мультиобъект  $M' \cdot c$ , где  $M' = M + M_0$ , а  $M_0$  есть стоимость реагентов.

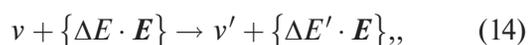
В некоторых случаях предпочтительными могут оказаться способы, которые характеризуются минимальными или ограниченными расходами тех или иных химических веществ, соединений и элементов. Для применения подобных крите-

риев на практике достаточно включить в фильтр соответствующие граничные и/или оптимизирующие условия.

### 3.2. Энергия и температура

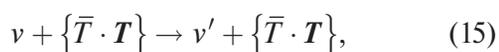
Как известно, для инициации химических реакций в общем случае могут понадобиться, наряду с реагентами, определенные **количества энергии**; с другой стороны, результатом реакции, помимо химических веществ, может быть выделение опять-таки некоторых количеств энергии. Кроме того, одним из триггеров химической реакции в общем случае является температура, которая применительно к конкретной реакции должна быть не ниже некоторого порогового значения либо находиться в некотором диапазоне. Наконец, существенно, что химические реакции протекают во времени и, соответственно, характеризуются продолжительностью. Рассмотрим, каким образом перечисленные факторы могут быть учтены в рамках мультиграмматического представления химических реакций.

Для представления химических реакций и процессов с учетом **потребляемой и выделяемой энергии** достаточно включить в левую и правую части правила  $v \rightarrow v'$ , отображающего химическое превращение, мультиобъекты, содержащие информацию о соответствующих количествах энергии. При этом правило  $v \rightarrow v'$  приобретает вид



где  $E$  – объект, представляющий единицу измерения энергии, а  $\Delta E$  и  $\Delta E'$  – количества энергии, необходимой для начала реакции и выделяемой в результате ее осуществления.

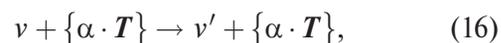
Химические реакции, для инициации которых необходима определенная температура, могут представляться правилами вида



где  $T$  есть объект, представляющий единицу измерения температуры, а  $\bar{T}$  – упомянутое пороговое значение температуры, при котором возможно химическое превращение совокупности реагентов, представленных мультимножеством  $v$ , в совокупность продуктов, представленных мультимножеством  $v'$ . Присутствие мультиобъекта  $\bar{T} \cdot T$  в левой и правой частях

правила отражает то обстоятельство, что температура в области пространства, где происходит реакция, по ее завершении не изменяется. Если изменение температуры в результате реакции имеет место, в правой части вместо мультиобъекта  $\bar{T} \cdot T$  будет использован мультиобъект  $\bar{T}' \cdot T$ , где  $\bar{T}'$  есть значение температуры по окончании реакции (это значение в общем случае может быть больше либо меньше  $\bar{T}$ ).

Более сложный с точки зрения мультиграмматического представления случай имеет место, когда реакция происходит в некотором диапазоне температур  $[\bar{T}, \bar{T}']$ . В этом случае требуется привлечение аппарата мультимножественных **метаграмматик (ММГ)** таким образом, что моделируемая система представляется в виде ММГ, схема которой, наряду с правилами вида (14) и (15), включает метаправила вида



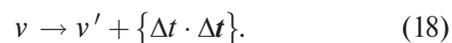
где  $\alpha$  – переменная, значение которой в процессе порождения мультимножеств есть температура в области пространства, где находятся реагенты. При этом фильтр ММГ включает описание данной переменной, имеющее вид

$$\bar{T} \leq \alpha \leq \bar{T}', \quad (17)$$

причем в каждом метаправиле используется своя уникальная переменная  $\alpha$ .

### 3.3. Продолжительность реакции

В качестве средства описания **протекающих во времени химических процессов** в МГИ могут быть использованы **темпоральные мультимножественные грамматики (ТМГ)**. Следуя семантике ТМГ, каждому уравнению реакции, совокупность реагентов которой представляется мультимножеством  $v$ , совокупность продуктов – мультимножеством  $v'$ , а продолжительность реакции составляет  $\Delta t$  единиц времени  $\Delta t$ , поставим в соответствие темпоральное правило



Исходной совокупности реагентов  $v_0$ , смешанных в момент  $t_0$ , поставим в соответствие темпоральное мультимножество  $v_0 = v_0 + \{t_0 \cdot t\}$ , где  $t$  есть объект, используемый для представления моментов времени. В результате множество терминальных темпоральных мультимножеств  $\bar{V}_s$ , порождаемое темпоральной

мультиграмматикой  $S = \langle v_0, R \rangle$ , где  $R$  есть множество темпоральных правил вида (18), представляющих уравнения химических реакций с их длительностями, есть не что иное, как мультимножественное представление совокупности результатов всех возможных цепочек химических реакций, инициируемых путем смешения реагентов в вещество  $v_0$  в момент времени  $t_0$ , причем каждый результат включает момент времени завершения этой цепочки (то есть момент получения продуктов СХР). С использованием подобного представления рассмотренные выше прямая и обратная задачи, в том числе с ограничениями на допустимую длительность последовательности реакций, решаются посредством применения фильтрующей ТМГ  $S' = \langle v_0, R, F \rangle$ , где фильтр  $F$  может включать граничное условие вида  $t \leq t_0 + \Delta \bar{t}$ , где  $\Delta \bar{t}$  есть допустимая длительность химического процесса, инициируемого совокупностью реагентов  $v_0$  (при  $t_0 = 0$  указанное условие трансформируется в  $t \leq \Delta \bar{t}$ ). В случае, если требуется цепочка химических реакций с минимальной продолжительностью, в фильтр достаточно включить оптимизирующее условие  $t = \min$ .

Обобщая рассмотренные способы применения элементов МГИ для представления уравнений химических реакций, можно утверждать, что универсальным средством этого класса, охватывающим все перечисленные особенности реакций (энергия, температура, продолжительность, стоимость), являются **темпоральные мультимножественные метаграмматики (ТММГ)**. В каждой такой ТММГ  $S = \langle v_0, R, F \rangle$

$$v_0 = v_0 + \{E \cdot E, T \cdot T, t_0 \cdot t\}, \quad (19)$$

то есть ядро этой метаграмматики в общем случае, помимо информации о реагентах в виде мультимножества  $v_0$ , включает сведения о количестве энергии  $E$ , доведенной до области реакции, о температуре  $T$ , имеющей место в данной области, а также о моменте времени  $t_0$  начала реакции. Схема  $R$  в общем случае наряду с правилами вида (6)–(15) и (18), представляющими уравнения реакций с учетом поглощаемой и выделяемой энергии, стоимости, продолжительности и пороговой температуры начала реакции, может содержать метаправила вида (16), представляющие диапазоны температуры, в которых могут протекать соответствующие реакции. Наконец, фильтр  $F$  этой ТММГ может содержать

граничные условия вида (17), фиксирующие упомянутые диапазоны температуры, а также граничные и/или оптимизирующие условия, определяющие допустимую продолжительность химического процесса, его стоимость и количества молекул тех или иных продуктов этого процесса.

### 3.4. Построение оптимальных маршрутов сложных химических реакций

Пусть имеем совокупность реагентов, представлением которой является мультимножество  $v_0$ , и базу знаний  $R$ . Множество терминальных мультимножеств  $\bar{V}_S$ , где  $S = \langle v_0, R \rangle$ , есть представление множества возможных химических веществ, которые могут быть получены из  $v_0$ . Целью соединения реагентов в  $v_0$  является получение некоторого ХВ, мультимножественное представление которого

$$\bar{v} = \{n_1 \cdot a_1, \dots, n_m \cdot a_m\} \quad (20)$$

в общем случае является подмультимножеством терминальных мультимножеств  $v \in \bar{V}_S$ , каждому из которых соответствует свой маршрут  $\rho(v_0, v)$  сложной химической реакции, обеспечивающей получение ХВ  $v$ . При этом мультимножество  $v - \bar{v}$  есть представление совокупности побочных продуктов СХР, маршрут которой есть  $\rho(v_0, v)$ .

Обозначим множество маршрутов СХР, начинающихся с совокупности реагентов  $v_0$ ,  $\rho_s$ :

$$\rho_s = \bigcup_{v \in \bar{V}_S} \{\rho(v_0, v)\}. \quad (21)$$

Критерий для селекции из множества  $\rho_s$  единственного оптимального маршрута может быть представлен в виде фильтра  $F$  фильтрующей темпоральной мультимножественной грамматики  $S = \langle v_0, R, F \rangle$ , включающего соответствующие граничные и/или оптимизирующие условия. При этом обязательным элементом фильтра должно быть мультимножественное граничное условие

$$\bar{v} \subseteq, \quad (22)$$

определяющее, что ХВ, полученное в результате СХР, должно содержать  $n_1$  молекул  $a_1, \dots, n_m$  молекул  $a_m$ . Наряду с граничным условием (22) фильтр может включать следующие условия:

$$c = \min, \quad (23)$$

которое обеспечивает селекцию маршрутов, имеющих минимальную стоимость;

$$E = \max, \quad (24)$$

которое обеспечивает селекцию маршрутов с минимальным энергопотреблением (поскольку минимальный расход энергии соответствует максимальному значению ее “остатка” по завершении СХР, условие (24) является максимизирующим);

$$a = \max, \quad (25)$$

которое обеспечивает селекцию маршрутов, характеризующихся минимальным расходом молекул вещества  $a$  (поскольку минимальный расход обеспечивает максимизацию числа молекул  $a$  в результате СХР  $v$ , условие (25) является максимизирующим);

$$\Delta t = \min, \quad (26)$$

которое обеспечивает селекцию маршрутов с минимальной продолжительностью;

$$b = \min, \quad (27)$$

где  $b$  есть обозначение ХВ, которое считается вредным побочным продуктом СХР, выброс которого негативно влияет на окружающую среду.

Как можно видеть, выполнение условия (23) гарантирует экономическую оптимальность маршрута; условия (24) – его максимальную энергоэффективность; условий вида (25) – его максимальную ресурсоэффективность (“бережливость” в отношении всех веществ, обозначения молекул которых имеют место в этих условиях); условия (26) – его минимально возможную продолжительность; тогда как условия вида (27) – его максимально возможную экологическую безопасность (опять-таки в отношении всех веществ, обозначения молекул которых имеют место в этих условиях).

Таким образом, применение фильтрующих темпоральных мультимножественных грамматик обеспечивает селекцию маршрутов СХР, оптимальных с точки зрения экономики, продолжительности реакции (то есть временных затрат

на получение ее продуктов), энерго- и ресурсопотребления, а также экологии. В случае, если ни один маршрут не удовлетворяет всем оптимизирующим условиям (23)–(27), некоторые из них могут быть исключены из фильтра  $F$  либо заменены граничными условиями. Этот процесс может продолжаться до получения хотя бы одного маршрута. Если, наоборот, маршрутов, удовлетворяющих условиям (23)–(27), больше одного, то фильтр  $F$  можно пополнить дополнительными условиями, которые позволяют исключить все маршруты, не удовлетворяющие этим условиям.

## 5. ЗАКЛЮЧЕНИЕ

Предложенная техника мультимножественного представления химических реакций позволяет формировать целостные модели производственно-экономических систем и процессов, на основе которых посредством того же мультимножественного инструментария можно решать задачи их оптимизации по различным критериям. Как можно видеть, предлагаемый подход отличается от известных подходов к использованию методов искусственного интеллекта в химической технологии двумя особенностями.

Во-первых, база знаний, используемая для решения указанных задач, содержит **не субъективные** знания, источниками которых являются эксперты и которые именно в силу их субъективности могут быть недостоверными, неполными и противоречивыми, а исключительно **объективные** знания – уравнения химических реакций в форме правил мультимножественных грамматик.

Во-вторых, **трудоемкость** решения рассмотренных задач со стороны конечных потребителей при использовании МГИ **минимальна** – по существу, достаточно указать цель (ХВ, которое необходимо получить) и ограничения, которым должен удовлетворять маршрут сложной химической реакции, обеспечивающей достижение этой цели. Предварительного построения графов, отображающих взаимосвязи между уравнениями различных химических реакций, и разработки алгоритмов и программ применения этих графов для решения задач не требуется – это делается системой логического вывода МГИ, обеспечивающей рекурсивное порождение и селекцию мультимножеств в соответствии с правилами, имеющимися в базе знаний, ис-

ходными данными (ядром мультиграмматики) и логическими условиями, которым должно удовлетворять решение поставленной задачи (фильтром МГ).

Как представляется, столь же естественным образом решаемы на основе МГИ задачи оценки устойчивости химико-технологических процессов и систем к разного рода деструктивным системоразрушающим воздействиям (стихийным бедствиям, техногенным катастрофам, экономическим санкциям, террористическим актам, кибератакам) [9–11].

### СПИСОК ЛИТЕРАТУРЫ

1. Мешалкин В.П. Экспертные системы в химической технологии. Основы теории, опыт разработки и применения. М: Химия, 1995. 368 с.
2. Мешалкин В.П. Основы интенсификации и ресурсоэнергоэффективности химико-технологических систем. Смоленск: Принт-экспресс, 2021. 442 с.
3. Мешалкин В.П., Бобков В.И., Дли М.И. Интеллектуальные методы и экспертные системы инжиниринга энергоресурсоэффективных химико-энерготехнологических процессов. Смоленск: Принт-Экспресс, 2021. 360 с.
4. Спивак С.И., Исмаилова А.С., Хамитова И.А. // Докл. АН. 2010. Т. 434. № 4. С. 499–501.
5. Баскин И.И. Формально-логический подход к органическим реакциям Н.С. Зефирова и С.С. Трача и его приложения для решения задач органической химии. В сб.: Органическая химия в работах Н.С. Зефирова. Уфа: Гилем, 2012. С. 219–238.
6. Спивак С.И., Исмаилова А.С., Ахмеров А.А. // Химия высоких энергий. 2015. Т. 49. № 4. С. 247–252. <http://doi.org/10.7868/S0023119315040166>
7. Баскин И.И., Маджидов Т.И., Антипин И.С., Варнек А.А. // Успехи химии. 2017. Т. 86. № 11. С. 1127–1156. <http://doi.org/10.1070/RCR4746>
8. Шеремет И.А. Рекурсивные мультимножества и их приложения. М.: Наука, 2010. 291 с.
9. Sheremet I.A. Multigrammatical Framework for Knowledge-Based Digital Economy. Cham: Springer Nature, 2022. 461 p.
10. Sheremet I. // Data Sci. J. V. 17. № 4. P. 1–17. <http://doi.org/10.5334/dsj-2018-004>
11. Kaper H., Roberts F., Sheremet I. Preparing for a Crisis: Improving the Resilience of Digitized Complex Systems. In: Mathematics for Action. Supporting Science-Based Decision-Making. Dhersin J.-S., Kaper H., Ndifon W., Roberts F., Rousseau C., Ziegler G.M. (Eds.). Paris: UNESCO, 2022. pp. 25–26.

## MULTISET GRAMMARS AS A BASIC KNOWLEDGE REPRESENTATION MODEL FOR INTELLIGENT SYSTEMS ENABLING CHEMICAL REACTIONS ENGINEERING

Academician of the RAS I. A. Sheremet<sup>a, #</sup>

<sup>a</sup>*Geophysical Center of the Russian Academy of Sciences, 119296 Moscow, Russian Federation*

<sup>#</sup>*E-mail: sheremet@rfbr.ru*

Application of multiset grammars as a knowledge representation model for intelligent systems, enabling analysis, optimization and design of complex chemical reactions, is proposed. This approach makes it possible to implant chemical knowledge bases into knowledge bases of intelligent systems, implementing smart planning and scheduling in cyberphysical industry, and thus to move to integrated end-to-end smart control of vast heterogeneous industrial clusters.

*Keywords:* multiset, multiset grammar, chemical reaction, route of a complex chemical reaction