МАТЕМАТИЧЕСКИЕ ОСНОВЫ ИНФОРМАТИКИ И ПРОГРАММИРОВАНИЯ

УДК 519.682

О ПРИМЕНЕНИИ МНОГОМЕРНОГО КОМПЛЕКСНОГО АНАЛИЗА В ТЕОРИИ ФОРМАЛЬНЫХ ЯЗЫКОВ И ГРАММАТИК¹

О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов

*Сибирский государственный университет науки и технологий им. акад. М. Ф. Решетнёва,
г. Красноярск, Россия*

Исследуются системы полиномиальных уравнений над полукольцом (относительно символов с некоммутативным умножением и коммутативным сложением). Такие системы уравнений интерпретируются как грамматики формальных языков и решаются относительно нетерминальных символов в виде формальных степенных рядов, зависящих от терминальных символов. Рассматривается коммутативный образ системы уравнений в предположении, что символы являются переменными, принимающими значения из поля комплексных чисел. Устанавливаются связи между решениями системы некоммутативных символьных уравнений и её коммутативного образа, тем самым методы многомерного комплексного анализа привлекаются в теорию формальных языков и грамматик. Доказывается дискретный аналог теоремы о неявном отображении для формальных грамматик: достаточным условием существования и единственности решения системы некоммутативных уравнений в виде формальных степенных рядов является отличие от нуля якобиана коммутативного образа этой системы. Предложен также новый метод синтаксического анализа мономов контекстно-свободного языка как модели языков программирования, основанный на интегральном представлении синтаксического многочлена программы. При этом показано, что интеграл фиксированной кратности по циклу позволяет найти синтаксический многочлен монома (программы) с неограниченным числом символов, что даёт новый подход к проблеме синтаксического анализа.

Ключевые слова: *формальный степенной ряд, коммутативный образ, синтаксический анализ, интегральное представление.*

DOI 10.17223/20710410/37/6

¹Исследование выполнено при финансовой поддержке РФФИ и Правительства Красноярского края в рамках научного проекта № 17-47-240318.

ON APPLICATION OF MULTIDIMENSIONAL COMPLEX ANALYSIS
IN FORMAL LANGUAGE AND GRAMMAR THEORY

O. I. Egorushkin, I. V. Kolbasina, K. V. Safonov

*Reshetnev State University of Science and Technology, Krasnoyarsk, Russia***E-mail:** safonovkv@rambler.ru

Systems of polynomial equations over a semiring (with respect to symbols with a non-commutative multiplication and a commutative addition) are investigated. These systems of equations are interpreted as the grammars of formal languages and are resolved with respect to the non-terminal symbols in the form of the formal power series (FPS) depending on the terminal symbols. The commutative image of the system of equations is determined under the assumption that the symbols are variables taking values from the field of complex numbers. The connections between solutions of the non-commutative symbolic system of equations and its commutative image are established, thus the methods for multidimensional complex analysis are involved in the theory of formal language grammars. A discrete analogue of implicit mapping theorem onto formal grammars is proved: a sufficient condition for the existence and the uniqueness of a solution of a non-commutative system of equations in the form of FPS is the inequality to zero of the Jacobian of the commutative image of this system. A new method for syntactic analysis of the monomials of a context-free language as a model of programming languages is also proposed. The method is based on the integral representation of the syntactic polynomial of a program. It is shown that the integral of a fixed multiplicity over a cycle allows finding the syntactic polynomial of a monomial (program) with the unlimited number of symbols, that gives a new approach to the problem of syntactic analysis.

Keywords: *formal power series, commutative image, syntactic analysis, integral representation.*

Введение

В теории формальных языков и грамматик исходным объектом является алфавит $\{z_1, \dots, z_n, x_1, \dots, x_m\}$, над которым определена некоммутативная операция умножения (конкатенации) и коммутативная операция формальной суммы; алфавит вместе с операциями образует полукольцо. Символы $x_1, \dots, x_m$ называются терминальными символами и составляют словарь формального языка, а символы $z_1, \dots, z_n$ — нетерминальными и нужны для задания совокупности грамматических правил (грамматики), которые порождают язык. По этим правилам определяются «правильные» мономы от терминальных символов $x_1, \dots, x_m$, которые являются правильными предложениями языка. Кроме того, определена коммутативная операция умножения символов алфавита на комплексные числа, что позволяет рассматривать символьные многочлены и формальные степенные ряды (ФСР) с числовыми коэффициентами. При этом формальным языком является такой ФСР, членами которого являются все правильные мономы, определённые данной грамматикой [1, 2].

Заметим, что порождающую грамматику для важных классов формальных языков можно записать в виде системы символьных полиномиальных уравнений

$$P_j(z, x) = 0, \quad P_j(0, 0) = 0, \quad j = 1, \dots, k, \quad (1)$$

которая решается относительно символов $(z_1, \dots, z_n) = z$ в виде ФСР, зависящих от символов $(x_1, \dots, x_m) = x$: $z = z(x) = (z_1(x), \dots, z_n(x))$; при этом ФСР $z_1(x)$ является формальным языком, который определяется этой грамматикой.

Так, для класса контекстно-свободных языков (кс-языков) система символьных уравнений (1) может быть представлена в виде системы уравнений Хомского — Шутценберже

$$z_j - Q_j(z, x) = 0, \quad j = 1, \dots, n, \quad (2)$$

где правые части удовлетворяют естественным условиям: $Q_j(0, 0) = 0$, многочлены $Q_j(z, 0)$ не содержат линейных членов [1]. Отметим, что классу кс-языков принадлежит большинство языков программирования [2].

Более широкий класс языков непосредственно составляющих задаётся системой символьных уравнений

$$m_j(z, x) - M_j(z, x) = 0, \quad j = 1, \dots, k,$$

где $m_j(z, x)$ — моном, а $M_j(z, x)$ — многочлен [2], которая также является частным случаем системы символьных уравнений (1).

Наконец, отметим языки в нормальной форме Грейбах, которые также можно задать системой символьных полиномиальных уравнений (1), где все многочлены $P_j(z, x)$ имеют степень не выше второй, вводя при необходимости новые терминальные и нетерминальные символы [2].

Таким образом, исследуя систему символьных полиномиальных уравнений (1), можно изучать свойства порождаемых ею формальных *полиномиальных* языков, в том числе языков из наиболее важных классов. Однако свойства некоммутативной системы уравнений (1) изучены мало. Так, в частном случае системы уравнений Хомского — Шутценберже (2) известно, что она имеет единственное решение $z = (z_1(x), \dots, z_n(x))$ в виде ФСР, но в общем случае условия разрешимости системы уравнений (1) неизвестны.

Для исследования нет подходящих методов: исключению неизвестных препятствует некоммутативность умножения и отсутствие операции деления. Например, методы дискретной математики, в частности теории графов, традиционно играющие важную роль в теории формальных языков и грамматик, не вполне приспособлены для исследования некоммутативной системы уравнений (1).

В случае коммутативных неизвестных системы полиномиальных уравнений изучаются в рамках алгебраической геометрии методами коммутативной алгебры. Если при этом рассматривать неизвестные над алгебраически замкнутым полем комплексных чисел, то можно использовать также методы многомерного комплексного анализа [3]. Однако методы коммутативной алгебры и комплексного анализа не выглядят удобными для изучения систем полиномиальных уравнений (1) с неизвестными, некоммутативными по умножению.

Таким образом, дальнейшее развитие теории формальных полиномиальных языков и грамматик (каждая грамматика может быть рассмотрена как система таких уравнений) затрудняется отсутствием подходящих методов, возникающих как в дискретной математике, так и в коммутативной алгебре и комплексном анализе. И всё же привлечь в теорию формальных языков и грамматик методы многомерного комплексного анализа можно на основе следующего подхода, предложенного в работах [4, 5].

Этот подход состоит в том, что наряду с некоммутативной системой (1) рассматривается её коммутативный образ, который получается в предположении, что все

переменные, входящие в систему, принимают значения из поля комплексных чисел. Коммутативный образ этой системы уравнений является системой полиномиальных уравнений в пространстве $\mathbb{C}_{z,x}^{n+m}$, а коммутативный образ формального языка становится кратным степенным рядом, представляющим алгебраическую функцию в $\mathbb{C}_x^m$, которая является решением этой системы. Далее можно исследовать коммутативный образ системы уравнений (1), применяя методы многомерного комплексного анализа. Наконец, остаётся установить, какие свойства некоммутативной системы уравнений (1) вытекают из свойств её коммутативного образа [4, 5].

В рамках этого подхода было отмечено, что из совместности некоммутативной системы (1) следует совместность её коммутативного образа, однако обратное утверждение неверно [4, 5]. В результате вопрос об исследовании некоммутативной системы (1) по свойствам её коммутативного образа оставался открытым.

Цель данной работы состоит в том, чтобы сделать следующий шаг по реализации указанного подхода, применяя в теории формальных языков и грамматик новые для неё методы многомерного комплексного анализа: во-первых, получить достаточное условие совместности (и единственности решения в виде ФСР) системы уравнений (1); во-вторых, решить важную прикладную задачу синтаксического анализа любой программы, написанной на кс-языке программирования.

Ниже доказано достаточное условие разрешимости некоммутативной системы (1) в терминах якобиана её коммутативного образа, а также предложен новый подход к проблеме синтаксического анализа, состоящий в том, что можно осуществить синтаксический анализ программы, переходя от «программирования к интегрированию» — вычислив кратный интеграл по циклу от некоторой рациональной формы.

1. Коммутативный образ ФСР

Упорядочим члены ФСР, следуя [4, 5]. Пусть все мономы от $x_1, \dots, x_m$ сгруппированы в однородные многочлены, расположенные по возрастанию степеней, затем пронумеруем мономы каждого из многочленов в лексикографическом порядке, переходя от меньшей степени к большей. При таком упорядочивании все мономы от символов $x_1, \dots, x_m$ единственным образом записываются в виде последовательности $u_0, u_1, \dots$, играющей роль базиса ФСР. Теперь каждый ряд s можно однозначно записать в виде разложения по этому базису:

$$s = \sum_{i=0}^{\infty} \langle s, u_i \rangle u_i, \quad (3)$$

где $\langle s, u_i \rangle$ — числовой коэффициент при мономе u_i .

Поставим в соответствие ФСР (3) его коммутативный образ $\text{ci}(s)$ — степенной ряд, который получается из s в предположении, что символы $x_1, \dots, x_m$ (равно как и $z_1, \dots, z_n$) обозначают коммутативные переменные, принимающие значения из поля комплексных чисел.

В этом предположении любой моном u_i от символов $x_1, \dots, x_m$ можно записать в виде $x_1^{\alpha_1} \cdot \dots \cdot x_m^{\alpha_m}$, где $\alpha_j = \deg_{x_j}(u_i)$ — число вхождений (степень) символа x_j в этот моном. Если обозначить мультииндекс $\alpha = (\alpha_1, \dots, \alpha_m)$, то можно записать равенство $\alpha = \deg_x(u_i)$, с учётом которого получаются следующие равенства:

$$\text{ci}(s) = \sum_{i=0}^{\infty} \langle s, u_i \rangle \text{ci}(u_i) = \sum_{\alpha} \left(\sum_{\alpha = \deg_x(u_i)} \langle s, u_i \rangle \right) x^{\alpha} \stackrel{\text{def}}{=} \sum_{\alpha} c_{\alpha} x^{\alpha}.$$

Впервые коммутативный образ ФСР рассмотрел А. Л. Семёнов [6], используя методы вещественного анализа для решения алгоритмических проблем, связанных с кс-

языками. Эта работа открыла дорогу для приложений как вещественного, так и многомерного комплексного анализа в теории формальных языков и грамматик.

Например, коммутативный образ ФСР может быть использован при решении проблемы академика В. М. Глушкова: установить критерии, с помощью которых можно выяснить, является данный ФСР кс-языком или нет [2, с. 196]. В самом деле, коммутативный образ кс-языка является кратным степенным рядом, представляющим алгебраическую функцию многих комплексных переменных, поэтому можно получить частичное решение этой проблемы, используя критерии алгебраичности для суммы степенного ряда [7–10].

Далее рассмотрим коммутативный образ системы символьных полиномиальных уравнений, представляющий собой систему полиномиальных уравнений в $\mathbb{C}_{z,x}^{n+m}$.

Теорема 1. Если $z = z(x) = (z_1(x), \dots, z_n(x))$ — решение некоммутиативной системы уравнений (1) в виде символьных ФСР, то коммутативные ФСР $z = \text{ci}(z(x)) = (\text{ci}(z_1(x)), \dots, \text{ci}(z_n(x)))$ над полем комплексных чисел сходятся в окрестности нуля, определяя ростки голоморфных алгебраических функций, и являются решением коммутативной системы уравнений

$$\text{ci}(P_j(z, x)) = 0, \quad \text{ci}(P_j(0, 0)) = 0, \quad j = 1, \dots, k. \quad (4)$$

Доказательство. Пусть решение $z(x) = (z_1(x), \dots, z_n(x))$ системы (1), представленное ФСР, имеет вид

$$z_1(x) = \sum_i \langle z_1, u_i \rangle u_i, \quad \dots, \quad z_n(x) = \sum_i \langle z_n, u_i \rangle u_i.$$

Отметим, что, подставляя это решение в многочлен $P_j(z, x)$, стоящий в левой части уравнений (1), получаем ФСР в виде разложения по универсальному базису с нулевыми коэффициентами:

$$l_j = \sum_i \langle l_j, u_i \rangle u_i = \sum_i 0 \cdot u_i, \quad j = 1, \dots, k.$$

Тогда имеют место равенства

$$\text{ci}(P_j(z, x))|_{z=\text{ci}(z(x))} = \text{ci}(P_j(z(x), x)) = \text{ci}(l_j) = \sum_i 0 \cdot \text{ci}(u_i), \quad j = 1, \dots, k,$$

т. е. совокупность коммутативных ФСР удовлетворяет системе уравнений (4).

Покажем, что все ФСР $\text{ci}(z(x))$ сходятся в некоторой окрестности нуля. В самом деле, если ФСР над полем комплексных чисел удовлетворяет уравнению с голоморфными коэффициентами (функциями, представленными абсолютно сходящимися в окрестности нуля степенными рядами), то он также абсолютно сходится в некоторой окрестности нуля, представляя голоморфную функцию [11]. Следовательно, коммутативные ФСР $\text{ci}(z(x))$ сходятся в некоторой окрестности нуля, порождая ростки алгебраической вектор-функции, а решения (z, x) системы (4), рассматриваемые как точки в комплексном пространстве $\mathbb{C}_{z,x}^{n+m}$, можно записать с помощью этой вектор-функции в виде $(z(x), x)$. ■

Частный случай теоремы 1, когда $k = n$, рассмотрен в [5].

Подчеркнём, что совместность понимается для некоммутиативной системы (1) и коммутативной системы (4) по-разному: в первом случае решением являются символьные ФСР от $x_1, \dots, x_m$, а во втором — точки в комплексном пространстве, параметризованные алгебраическим отображением $z = z(x)$; его голоморфные в нуле ветви

являются коммутативными образами символьных ФСР, представляющих компоненты решения системы (1). По теореме 1, если некоммутативная система (1) совместна, то коммутативная система (4) имеет голоморфное в начале координат решение. Обратное, вообще говоря, неверно. В самом деле, система уравнений

$$z_1 - z_2 = x_1x_2, \quad z_1 - z_2 = x_2x_1$$

является несовместной, тем не менее её коммутативный образ имеет решение: $z_1 = s + x_1x_2$, $z_2 = s$, где s — произвольный коммутативный ФСР. Таким образом, можно сделать следующие замечания.

Замечание 1. Множество решений системы уравнений (4), вообще говоря, шире, чем множество коммутативных образов решений системы уравнений (1), и из совместности системы уравнений (4) не следует совместность исходной некоммутативной системы (1).

Замечание 2. Естественная гипотеза о том, что система (1) совместна тогда и только тогда, когда система (4) имеет голоморфное в начале координат решение $z = z(x)$, неверна.

Рангом матрицы A называется наибольшее число её строк, линейно независимых над полем комплексных чисел; обозначается $\text{rank}(A)$.

Теорема 2. Пусть $A = (a_{ij}(x))_{k \times n}$ — матрица, элементы которой являются многочленами от символьных некоммутативных переменных $x = (x_1, \dots, x_m)$, и матрица $\text{ci}(A) = (\text{ci}(a_{ij}(x)))_{k \times n}$ является коммутативным образом матрицы A . Тогда выполнено неравенство

$$\text{rank}(\text{ci}(A)) \leq \text{rank}(A).$$

Доказательство. Пусть, для определённости, линейно независимыми являются строки с номерами $1, \dots, r$, следовательно, $\text{rank}(A) = r$. Приравнивая к нулю линейную комбинацию этих строк с коэффициентами $z_1, \dots, z_r$, получим систему полиномиальных уравнений

$$z_1 a_{1j}(x) + \dots + z_r a_{rj}(x) = 0, \quad j = 1, \dots, n,$$

которая имеет единственное решение в виде ФСР: $z_1 = \dots = z_r = 0$. Коммутативный образ этой системы уравнений

$$z_1 \text{ci}(a_{1j}(x)) + \dots + z_r \text{ci}(a_{rj}(x)) = 0, \quad j = 1, \dots, n,$$

в соответствии с замечанием 1, может иметь более широкое множество решений, включающее решение $z_1 = \dots = z_r = 0$. В этом случае коммутативные образы исходных строк являются линейно зависимыми, следовательно, $\text{rank}(\text{ci}(A)) \leq \text{rank}(A)$. ■

Отметим, что теореме 1 эквивалентна следующая теорема.

Теорема 3. Если коммутативная система уравнений (4) не имеет решения, голоморфного в начале координат, то некоммутативная система (1) несовместна.

Таким образом, условия несовместности системы уравнений (4) также представляют интерес в теории формальных языков и грамматик.

2. Дискретный аналог теоремы о неявном отображении

Конечно, наибольший интерес для приложений представляют условия, которые обеспечивают совместность системы некоммутативных символьных уравнений (1), а также единственность её решения. Согласно замечанию 2, таким условием не может

быть существование голоморфного в начале координат решения системы коммутативных уравнений (4). Тем не менее это условие можно получить с помощью такого инструмента, как якобиан системы функций.

Рассмотрим систему уравнений (1) в случае, когда $k = n$. Пусть

$$J(z, x) = \det((\text{ci}(P_i(z, x)))'_{z_j})$$

— якобиан системы уравнений (4) относительно переменных $z_1, \dots, z_n$.

Дискретным (символьным) аналогом теоремы о неявном отображении является следующая теорема.

Теорема 4. Если для некоммутиративной символьной системы уравнений (1) выполнено неравенство $J(0, 0) \neq 0$, то она имеет единственное решение в виде ФСР.

Замечание 3. Неравенство является условием теоремы о неявном отображении для коммутативной системы уравнений (4) с переменными в $\mathbb{C}_{z,x}^{n+m}$ и влечёт существование и единственность её голоморфного решения; тем не менее оказывается, что это неравенство влечёт также существование и единственность решения исходной некоммутиративной символьной системы уравнений (1).

Доказательство. Обозначим $L_1(z), \dots, L_n(z)$ линейные части многочленов $P_1(z, x), \dots, P_n(z, x)$ соответственно, зависящие только от $z_1, \dots, z_n$:

$$L_j(z) = a_{j1}z_1 + \dots + a_{jn}z_n, \quad j = 1, \dots, n.$$

Далее, обозначая многочлены $S_j(z, x) = L_j(z) - P_j(z, x)$, запишем исходную систему уравнений в виде

$$a_{j1}z_1 + \dots + a_{jn}z_n = S_j(z, x), \quad j = 1, \dots, n. \quad (5)$$

Рассмотрим число $J(0, 0)$; легко видеть, что имеют место равенства

$$J(0, 0) = \det((\text{ci}(L_i(z)))'_{z_j}) = \det(((L_i(z)))'_{z_j}) = \det(a_{ij}).$$

Поскольку $\det(a_{ij}) \neq 0$, матрицу (a_{ij}) можно привести к диагональному виду, умножая на числа уравнения системы (5) и складывая их [12]. При этом система (5) перейдёт в эквивалентную систему уравнений Хомского — Шутценберже (2), в которой правые части уравнений не содержат линейных членов.

Как отмечалось выше, эта система имеет единственное решение, которое можно получить методом последовательных приближений [1]:

$$z^{(k+1)}(x) = Q(z^{(k)}(x), x), \quad k = 0, 1, \dots; \quad z^{(0)} = 0;$$

$$Q(z, x) = (Q_1(z, x), \dots, Q_n(z, x)); \quad z(x) = \lim_{k \rightarrow \infty} z^{(k)}(x).$$

Теорема доказана. ■

Замечание 4. Неравенство $J(0, 0) \neq 0$ обусловлено свойствами линейных многочленов $L_1(z), \dots, L_n(z)$, а они совпадают со своими коммутативными образами. По этой причине неравенство влечёт совместность не только коммутативной, но и некоммутиративной системы уравнений.

Наконец, отметим, что система уравнений (1) имеет бесконечно много решений, если множество её решений зависит хотя бы от одного произвольного ФСР от символов $x_1, \dots, x_m$ [4].

Пример 1. Система из двух одинаковых уравнений $x_1z_1 - z_2x_2 = 0$ имеет бесконечно много решений, поскольку её решения можно записать в виде $z_1 = sx_2$, $z_2 = x_1s$, где s — произвольный ФСР от x_1, x_2 .

3. Синтаксический анализ мономов методом мономиальных меток

Одной из важных проблем, связанных с разработкой систем и языков программирования, является проблема синтаксического анализа программ [2]. Как отмечено выше, большинство языков программирования является кс-языками, которые можно представить в виде ФСР, поэтому каждая программа, написанная на языке программирования, может рассматриваться как моном соответствующего ФСР. В связи с этим рассмотрим проблему синтаксического анализа мономов кс-языка.

Для того чтобы сформулировать её, рассмотрим подробнее систему полиномиальных уравнений Хомского — Шутценберже (2), которая определяет кс-язык. Как известно [1, 2], грамматика кс-языка является множеством правил подстановки

$$z_j \rightarrow q_{j1}(z, x), \dots, z_j \rightarrow q_{jp_j}(z, x), \quad j = 1, \dots, n, \quad (6)$$

где $q_{jk}(z, x)$ является мономом от некоммутативных символьных переменных с числовым коэффициентом, равным единице. Правила подстановки можно применять к начальному символу z_1 , а затем к другим мономам в любом порядке неограниченное число раз, что позволяет выводить новые «правильные» мономы, образующие кс-язык. Известно также [1], что многочлены, стоящие в правой части системы уравнений Хомского — Шутценберже (2), определяются равенствами

$$Q_j(z, x) = q_{j1}(z, x) + \dots + q_{jp_j}(z, x), \quad j = 1, \dots, n.$$

Итак, проблема синтаксического анализа мономов состоит в том, чтобы определить, принадлежит ли моном данному кс-языку, т.е. может ли быть получен из начального символа z_1 при помощи правил подстановки (6), а также установить, какие правила подстановки и сколько раз использовались при выводе этого монома; при этом порядок использования правил подстановки не имеет значения [2].

В работе [8] предложен метод мономиальных меток, который позволяет провести беступиковый синтаксический анализ монома v от терминальных символов $x_1, \dots, x_m$. Метод состоит в следующем. Сначала каждое правило подстановки $z_j \rightarrow q_{jk}(z, x)$ заменяется правилом $z_j \rightarrow t_{jk}q_{jk}(z, x)$, имеющим мономиальную метку t_{jk} , которая является символом из расширенного алфавита, и для новых правил вывода рассматривается соответствующая система уравнений Хомского — Шутценберже:

$$z_j = Q_j^*(z, x, t) \stackrel{\text{def}}{=} t_{j1}q_{j1}(z, x) + \dots + t_{jp_j}q_{jp_j}(z, x), \quad j = 1, \dots, n. \quad (7)$$

Далее решение системы (7) можно получить методом последовательных приближений, о котором говорилось в доказательстве теоремы 4:

$$z^{(k+1)}(x, t) = Q^*(z^{(k)}(x, t), x, t), \quad k = 0, 1, \dots; \quad z^{(0)} = 0.$$

Как результат, решение получается в виде ФСР

$$z_j = z_j^*(x, t) = \sum_{i=0}^{\infty} \langle z_j^*, w_i \rangle w_i, \quad j = 1, \dots, n, \quad (8)$$

где w_i — мономы от символов $x_1, \dots, x_m, t_{11}, t_{12}, \dots, t_{np_n}$. Заметим, что если каждый символ t_{jk} заменить пустой цепочкой e , то решения систем уравнений (7) и (2) совпадают:

$$z_j^*(x, e) = z_j(x), \quad j = 1, \dots, n. \quad (9)$$

Теперь итерации метода последовательных приближений для системы уравнений (7) дают многочлены возрастающей степени относительно всех символов $x_1, \dots, x_m, t_{11}, t_{12}, \dots, t_{np_n}$, при этом мономы степени не выше $\deg_x(v)$ относительно символов $x_1, \dots, x_m$ после конечного числа итераций стабилизируются, не меняясь при последующих итерациях. Таким образом, можно получить начальные члены решения системы (7) в виде ФСР (8) до любой сколь угодно высокой степени, в том числе члены ФСР, представляющего первую компоненту этого решения:

$$z_1 = z_1^*(x, t) = \sum_{i=0}^{\infty} \langle z_j^*, w_i \rangle w_i. \quad (10)$$

Наконец, синтаксический анализ монома v кс-языка $z_1(x)$ можно провести следующим образом. Считывая мономы степени $\deg_x(v)$ относительно символов $x_1, \dots, x_m$ и пропуская символы t_{jk} , можно установить, есть ли среди них моном v , а значит, можно ли вывести его с помощью системы продукций (6). При этом каждая мономиальная метка t_{jk} , содержащаяся в таком мономе, показывает, что при его выводе использовалось правило $z_j \rightarrow t_{jk}q_{jk}(z, x)$. В самом деле, из системы уравнений (7) и метода последовательных приближений нетрудно видеть, что, применяя это правило вывода к моному, мы умножаем его слева на символ t_{jk} . Следовательно, мономиальные метки монома решают проблему его синтаксического анализа, показывая, какие правила вывода кс-языка и сколько раз использовались при выводе этого монома, с точностью до порядка их применения.

Доказано [8], что метод мономиальных меток позволяет провести за конечное число шагов беступиковый синтаксический анализ любого монома (программы) кс-языка, заданного грамматикой (6).

Очевидно, что недостатком метода мономиальных меток является большое число громоздких итераций метода последовательных приближений, необходимых для получения начальных членов ФСР (8), причём это число возрастает вместе с ростом степени монома v . В связи с этим ниже предлагается другой путь для получения мономиальных меток некоммутативного монома.

4. Интегральное представление синтаксического многочлена

Информацию о мономиальных метках монома можно получить в виде $(n+m)$ -кратного интеграла по циклу, где числа n и m не зависят от степени монома и равны числу нетерминальных и терминальных символов грамматики кс-языка соответственно.

Рассмотрим коммутативный образ ФСР (10)

$$\text{ci}(z_1^*(x, t)) = \sum_{\alpha} s_{\alpha}(t) x^{\alpha}, \quad (11)$$

сгруппированный по степеням x^{α} в кратный ряд Гартогса.

Лемма 1. При всех мультииндексах α голоморфные в нуле коэффициенты ряда Гартогса $s_{\alpha}(t)$ являются многочленами.

Доказательство. Как отмечено выше, числовые коэффициенты системы уравнений (7) равны единице, поэтому коэффициенты ФСР (10) являются целыми положительными числами. Отсюда следует, что коэффициенты каждого степенного ряда также являются целыми положительными числами, а число $s_{\alpha}(e)$ равно сумме этих коэффициентов. Из равенств (9) и (11) следует равенство

$$\text{ci}(z_1(x)) = \sum_{\alpha} s_{\alpha}(e) x^{\alpha},$$

которое влечёт неравенство $s_\alpha(e) < \infty$. Это значит, что сумма неотрицательных целых коэффициентов степенного ряда функции $s_\alpha(e)$ конечна, следовательно, эта функция является многочленом. ■

Определение 1. Синтаксическим многочленом монома v относительно кс-языка $z_1(x) = z_1^*(x, e)$ называется коэффициент $s_\alpha(t)$ ряда Гартогса (11), такой, что $x^\alpha = \text{ci}(v)$.

Замечание 5. Мономиальные метки, содержащиеся в некоммутативных мономах кс-языка, не исчезают при переходе от ФСР (10) к его коммутативному образу (11) и сохраняются в виде мономов синтаксических многочленов, поскольку все коэффициенты ФСР (10) являются целыми положительными числами. Следовательно, если синтаксический многочлен монома относительно кс-языка равен нулю, то моном не принадлежит этому языку.

Замечание 6. Для проведения синтаксического анализа монома v , такого, что $\text{ci}(v) = x^\alpha$, следует найти синтаксический многочлен $s_\alpha(t)$. Каждый моном многочлена $s_\alpha(t)$ является произведением мономиальных меток правил подстановки, которые позволяют получить некоторые мономы, имеющие тот же коммутативный образ x^α , поэтому для завершения синтаксического анализа монома v остаётся проверить, можно ли получить его с помощью правил подстановки, соответствующих всем мономам синтаксического многочлена $s_\alpha(t)$.

Следующая теорема даёт принципиальную возможность получить синтаксические многочлены $s_\alpha(t)$ в виде кратного интеграла по циклу, который может быть вычислен с помощью многомерных вычетов.

Теорема 5. При всех t , достаточно близких к нулю, и всех мультииндексах α синтаксический многочлен $s_\alpha(t)$ задаётся равенствами

$$s_\alpha(t) = \frac{1}{(2\pi i)^{n+m}} \int_{\gamma_z \times \Gamma_x} \frac{z_1 \det(\delta_{ij} - (\text{ci}(Q_i^*(z, x, t)))'_{z_j}) dz \wedge dx}{(z - \text{ci}(Q^*(z, x, t))) x^{\alpha+I}}; \quad (12)$$

$$s_\alpha(t) = \frac{1}{(2\pi i)^n \alpha!} \int_{\gamma_z} \frac{\partial^{\|\alpha\|}}{\partial x^\alpha} \left(\frac{z_1 \det(\delta_{ij} - (\text{ci}(Q_i^*(z, x, t)))'_{z_j})}{(z - \text{ci}(Q^*(z, x, t)))} \right) \Big|_{x=0} dz, \quad (13)$$

где $\gamma_z = \{|z_1| = \dots = |z_n|\}$ и $\Gamma_x = \{|x_1| = \dots = |x_m|\}$ — циклы интегрирования; $0 < \delta \ll \varepsilon \ll 1$; $dz = dz_1 \wedge \dots \wedge dz_n$; $dx = dx_1 \wedge \dots \wedge dx_m$; $x^{\alpha+I} = x_1^{\alpha_1+1} \dots x_m^{\alpha_m+1}$; $\alpha! = \alpha_1! \dots \alpha_m!$; $(z - \text{ci}(Q^*(z, x, t))) = (z_1 - \text{ci}(Q_1^*(z, x, t))) \cdot \dots \cdot (z_n - \text{ci}(Q_n^*(z, x, t)))$; $\frac{\partial^{\|\alpha\|}}{\partial x^\alpha} = \frac{\partial^{\alpha_1+\dots+\alpha_m}}{\partial x_1^{\alpha_1} \dots \partial x_m^{\alpha_m}}$; δ_{ij} — символ Кронекера.

Для доказательства теоремы 5 понадобится следующая лемма.

Лемма 2. При всех (x, t) , достаточно близких к нулю, голоморфная в нуле алгебраическая функция $\text{ci}(z_1^*(x, t))$, представленная рядом Гартогса (11), задаётся формулой

$$\text{ci}(z_1^*(x, t)) = \frac{1}{(2\pi i)^n} \int_{\gamma_z} \frac{z_1 \det(\delta_{ij} - (\text{ci}(Q_i^*(z, x, t)))'_{z_j})}{(z - \text{ci}(Q^*(z, x, t)))} dz. \quad (14)$$

Доказательство. Во-первых, отметим, что $\det(\delta_{ij} - (\text{ci}(Q_i^*(z, x, t)))'_{z_j}) = J_1(z, x)$ — якобиан системы функций

$$(z_1 - \text{ci}(Q_1^*(z, x, t)), \dots, z_n - \text{ci}(Q_n^*(z, x, t)))$$

по переменным $z_1, \dots, z_n$. Так как $\det((\text{ci}(Q_i^*(0, 0, 0)))'_{z_j}) = 0$, то $J_1(0, 0) = 1$, следовательно, по теореме о неявном отображении система уравнений

$$z_j - \text{ci}(Q_j^*(z, x, t)) = 0, \quad j = 1, \dots, n,$$

имеет в окрестности начала координат единственное и голоморфное решение:

$$z_j = \text{ci}(z_j^*(x, t)) = \text{ci}\left(\sum_{i=0}^{\infty} \langle z_j^*, w_j \rangle w_j\right), \quad j = 1, \dots, n.$$

Далее, если $z \in \gamma_z$, то при всех (x, t) , достаточно близких к нулю, выполнены неравенства

$$|z_j| \geq |\text{ci}(Q_j^*(z, x, t))|, \quad j = 1, \dots, n,$$

поэтому на цикле интегрирования знаменатель подынтегральной рациональной функции (14) не обращается в нуль. Наконец, из формулы многомерного логарифмического вычета [13] следует формула (14). ■

Доказательство теоремы 5. Коэффициенты ряда Гартогса, сходящегося в окрестности нуля, могут быть представлены формулой Коши

$$s_\alpha = \frac{1}{(2\pi i)^n} \int_{\Gamma_x} \frac{\text{ci}(z_1^*(x, t)) dx}{x^{\alpha+I}} \quad (15)$$

либо формулой Тейлора

$$s_\alpha = \frac{1}{\alpha!} \frac{\partial^{|\alpha|}}{\partial x^\alpha} (\text{ci}(z_1^*(0, t))).$$

Из этих формул и формулы (14) следует, что формулы (12) и (13) верны. ■

5. Пример вычисления синтаксического многочлена программы

Следуя [2], рассмотрим кс-язык арифметических выражений, порождённый грамматикой с системой правил подстановки

$$z_1 \rightarrow (z_1 * z_2), \quad z_1 \rightarrow (a + b), \quad z_2 \rightarrow b,$$

и выражение $((a + b) * b) * b$ как программу, написанную на этом языке. Известные методы синтаксического анализа показывают [2, с. 247], что для получения этой программы необходимо два раза использовать правило подстановки $z_1 \rightarrow (z_1 * z_2)$, один раз — правило $z_1 \rightarrow (a + b)$ и два раза — правило $z_2 \rightarrow b$.

Для того чтобы подчеркнуть, что интеграл (12), имеющий одну и ту же размерность $n + m$, позволяет анализировать мономы с неограниченными степенями, рассмотрим обобщённую программу, длина которой может быть сколь угодно большой.

Пример 2. Рассмотрим обобщённую программу $(\dots((a + b) * b) \dots * b)$, где символ «*» использован q раз и символ «(» — $q + 1$ раз. Легко проверить, что программа содержит $4q + 5$ символов. По аналогии с предыдущей программой можно видеть, что обобщённую программу можно получить, используя первое и третье правила подстановки q раз, а второе правило один раз.

Сначала переобозначим терминальные символы «+», «*», «(», «)», a , b символами $x_1, x_2, x_3, x_4, x_5, x_6$ соответственно, тогда система уравнений (7) для этого кс-языка примет вид

$$z_1 = t_{11}x_3z_1x_2z_2x_4 + t_{12}x_3x_5x_1x_6x_4, \quad z_2 = t_{21}x_6. \quad (16)$$

Перепишем также эту программу как моном

$$x_3 \cdots x_3 x_3 x_5 x_1 x_6 x_4 x_2 x_6 x_4 \cdots x_2 x_6 x_4,$$

имеющий степень $4q + 5$, которая может быть сколь угодно большой. Коммутативный образ этого монома равен

$$x_1 x_2^q x_3^{q+1} x_4^{q+1} x_5 x_6^{q+1}, \quad (17)$$

поэтому мультииндекс α в формуле (15) равен $(1, q, q + 1, q + 1, 1, q + 1)$.

Далее, якобиан системы уравнений (16) в этом случае равен

$$\det(\delta_{ij} - (\text{ci}(Q_i^*(z, x, t)))'_{z_j}) = 1 - t_{11} x_2 x_3 x_4 z_2.$$

Теперь вычислим кратный интеграл (12) как повторный, сначала вычисляя интеграл (14), а затем (15). Интеграл (14) в нашем случае равен

$$\frac{1}{(2\pi i)^2} \int_{|z_1|=|z_2|=\varepsilon} \frac{z_1(1 - t_{11} x_2 x_3 x_4 z_2) dz_1 \wedge dz_2}{(z_1 - t_{11} x_2 x_3 x_4 z_1 z_2 - t_{12} x_1 x_3 x_4 x_5 x_6)(z_2 - t_{21} x_6)}. \quad (18)$$

Вычисляя последний интеграл по формуле Коши сначала по переменной z_2 , а затем по z_1 , получим

$$\begin{aligned} \text{ci}(z_1^*) &= \frac{1}{2\pi i} \int_{|z_1|=\varepsilon} \frac{z_1(1 - t_{11} t_{21} x_2 x_3 x_4 x_6) dz_1}{z_1 - t_{11} t_{21} x_2 x_3 x_4 x_6 z_1 - t_{12} x_1 x_3 x_4 x_5 x_6} = \\ &= \frac{1}{2\pi i} \int_{|z_1|=\varepsilon} \frac{z_1(1 - t_{11} t_{21} x_2 x_3 x_4 x_6) dz_1}{z_1(1 - t_{11} t_{21} x_2 x_3 x_4 x_6) - t_{12} x_1 x_3 x_4 x_5 x_6} = \\ &= \frac{1}{2\pi i} \int_{|z_1|=\varepsilon} \frac{z_1(1 - t_{11} t_{21} x_2 x_3 x_4 x_6) dz_1}{\left(z_1 - \frac{t_{12} x_1 x_3 x_4 x_5 x_6}{1 - t_{11} t_{21} x_2 x_3 x_4 x_6}\right) (1 - t_{11} t_{21} x_2 x_3 x_4 x_6)} = \\ &= \frac{1}{2\pi i} \int_{|z_1|=\varepsilon} \frac{z_1 dz_1}{z_1 - \frac{t_{12} x_1 x_3 x_4 x_5 x_6}{1 - t_{11} t_{21} x_2 x_3 x_4 x_6}} = \frac{t_{12} x_1 x_3 x_4 x_5 x_6}{1 - t_{11} t_{21} x_2 x_3 x_4 x_6}. \end{aligned}$$

Теперь вычислим интеграл (15), где $\alpha = (1, q, q + 1, q + 1, 1, q + 1)$, $\Gamma_x = \{|x_1| = \dots = |x_6|\}$, $t = (t_{11}, t_{12}, t_{21})$, используя разложение подынтегральной рациональной функции в степенной ряд для геометрической прогрессии:

$$\begin{aligned} s_\alpha(t) &= \frac{1}{(2\pi i)^6} \int_{\Gamma_x} \frac{t_{12} x_1 x_3 x_4 x_5 x_6 dx_1 \wedge \cdots \wedge dx_6}{(1 - t_{11} t_{21} x_2 x_3 x_4 x_6) x_1^2 x_2^{q+1} x_3^{q+2} x_4^{q+2} x_5^2 x_6^{q+2}} = \\ &= \frac{1}{(2\pi i)^6} \int_{\Gamma_x} \left(\sum_{k=0}^{\infty} t_{11}^k t_{12} t_{21}^k x_1^k x_2^{k+1} x_3^{k+1} x_4^{k+1} x_5 x_6^{k+1} \right) \frac{dx_1 \wedge \cdots \wedge dx_6}{x_1^2 x_2^{q+1} x_3^{q+2} x_4^{q+2} x_5^2 x_6^{q+2}} = \\ &= \sum_{k=0}^{\infty} \frac{t_{11}^k t_{12} t_{21}^k}{(2\pi i)^6} \int_{\Gamma_x} x_2^{k-q} x_3^{k-q} x_4^{k-q} x_6^{k-q} \frac{dx_1 \wedge \cdots \wedge dx_6}{x_1 \dots x_6} = t_{11}^q t_{12} t_{21}^q. \end{aligned}$$

Очевидно, что интегралы от всех слагаемых равны нулю, кроме случая $k = q$, когда интеграл равен $(2\pi i)^6$. Итак, синтаксический многочлен мономов, имеющих коммутативный образ (17), равен $t_{11}^q t_{12} t_{21}^q$; согласно замечанию 6, мы должны проверить единственный путь для получения нашей программы: q раз использовать правило подстановки $z_1 \rightarrow (z_1 * z_2)$, один раз — правило $z_1 \rightarrow (a + b)$ и q раз — правило $z_2 \rightarrow b$, начиная

с исходного символа z_1 . Этот путь в самом деле позволяет получить программу, что и завершает её синтаксический анализ.

Таким образом, пример 2 подтверждает, что интегральное представление фиксированной кратности даёт возможность осуществлять синтаксический анализ мономов (программ), степени (длины) которых неограничены.

ЛИТЕРАТУРА

1. *Salomaa A. and Soittola M.* Automata-Theoretic Aspects of Formal Power Series. N.Y.: Springer Verlag, 1978. 167 p.
2. *Глушков В. М., Цейтлин Г. Е., Ющенко Е. Л.* Алгебра. Языки. Программирование. Киев: Наукова думка, 1973. 319 с.
3. *Гриффитс Ф., Харрис Дж.* Принципы алгебраической геометрии. Т. 1. М.: Мир, 1982. 259 с.
4. *Егорушкин О. И., Колбасина И. В., Сафонов К. В.* О совместности систем символьных полиномиальных уравнений и их приложении // Прикладная дискретная математика. Приложение. 2016. №9. С. 119–121.
5. *Egorushkin O. I., Kolbasina I. V., and Safonov K. V.* On Solvability of Systems of Symbolic Polynomial Equations // Журнал Сибирского федерального университета. Математика и физика. 2016. Т. 9. №2. С. 166–172.
6. *Семёнов А. Л.* Алгоритмические проблемы для степенных рядов и контекстно-свободных грамматик // Доклады Академии наук СССР. 1973. Т. 212. С. 50–52.
7. *Safonov K. V.* On power series of algebraic and rational functions in C^n // J. Math. Analysis and Appl. 2000. V. 243. P. 261–277.
8. *Сафонов К. В., Егорушкин О. И.* О синтаксическом анализе и проблеме В. М. Глушкова распознавания контекстно-свободных языков Хомского // Вестник Томского государственного университета. 2006. Приложение №17. С. 63–67.
9. *Safonov K. V.* An algebraicity criterion for the sum of a power series (a generalization of Kronecker's criterion) and its application // Doklady Mathematics. 2009. V. 79. No. 1. P. 13–15.
10. *Pemantle R. and Wilson M. C.* Analytic Combinatorics in Several Variables. Cambridge: Cambridge University Press, 2013. 414 p.
11. *Herve M.* Several Complex Variables. Local Theory. Oxford: Oxford University, 1963. 158 p.
12. *Глухов М. М., Елизаров В. П., Нечаев А. А.* Алгебра. СПб.; М.; Краснодар: Лань, 2015. 608 с.
13. *Aizenberg L. A. and Yuzhakov A. P.* Integral Representations and Residues in Multidimensional Complex Analysis. Translations of Mathematical Monographs. V. 58. Providence: AMS, 1983. 283 p.

REFERENCES

1. *Salomaa A. and Soittola M.* Automata-Theoretic Aspects of Formal Power Series. N.Y., Springer Verlag, 1978. 167 p.
2. *Glushkov V. M., Tseytlin G. E., and Yushchenko E. L.* Algebra. Yazyki. Programirovanie [Algebra. Languages. Programming]. Kiev, Naukova dumka, 1973. 319 p. (in Russian)
3. *Griffiths P. and Harris J.* Principles of Algebraic Geometry. Vol. 1. N.Y., John Wiley and Sons, 1978. 348 p.
4. *Egorushkin O. I., Kolbasina I. V., and Safonov K. V.* O sovместnosti sistem simvol'nykh polinomial'nykh uravneniy i ikh prilozhenii [On consistency of systems of symbolic polynomial equations and its application]. Prikladnaya Discretnaya Matematika. Prilozhenie, 2016, no. 9, pp. 119–121. (in Russian)

5. *Egorushkin O. I., Kolbasina I. V., and Safonov K. V.* On solvability of systems of symbolic polynomial equations. *J. Siberian Federal University. Mathematics and Physics*, 2016, vol. 9, no. 2, pp. 166–172.
6. *Semenov A. L.* Algoritmicheskie problemy dlya stepennykh ryadov i kontekstno-svobodnykh grammatik [Algorithmic problems for power series and context-free grammars]. *Doklady Akademii nauk SSSR*, 1973, vol. 212, pp. 50–52. (in Russian)
7. *Safonov K. V.* On power series of algebraic and rational functions in C^n . *J. Math. Analysis and Appl.*, 2000, vol. 243, pp. 261–277.
8. *Safonov K. V. and Egorushkin O. I.* O sintaksicheskoy analize i probleme V. M. Glushkova raspoznavaniya kontekstno-svobodnykh yazykov Khomskogo [Syntactical analysis and the V. M. Glushkov problem of recognition of Chomsky Context-Free Languages]. *Vestnik TSU. Prilozhenie*, 2006, no. 17, pp. 63–67. (in Russian)
9. *Safonov K. V.* An algebraicity criterion for the sum of a power series (a generalization of Kronecker's criterion) and its application. *Doklady Mathematics*, 2009, vol. 79, no. 1, pp. 13–15.
10. *Pemantle R. and Wilson M. C.* *Analytic Combinatorics in Several Variables*. Cambridge, Cambridge University Press, 2013. 414 p.
11. *Herve M.* *Several Complex Variables. Local Theory*. Oxford, Oxford University, 1963. 158 p.
12. *Glukhov M. M., Elizarov V. P., and Nechaev A. A.* *Algebra [Algebra]*. SPb., Moscow, Krasnodar, LAN' Publ., 2015. 608 p. (in Russian)
13. *Aizenberg L. A. and Yuzhakov A. P.* *Integral Representations and Residues in Multidimensional Complex Analysis*. *Translations of Mathematical Monographs*. Vol. 58. Providence, AMS, 1983. 283 p.