На правах рукописи

Егорушкин Олег Игоревич

РАЗРЕШИМОСТЬ ПОЛИНОМИАЛЬНЫХ ГРАММАТИК И СИНТАКСИЧЕСКИЙ АНАЛИЗ КОНТЕКСТНО — СВОБОДНЫХ ЯЗЫКОВ НА ОСНОВЕ КОММУТАТИВНЫХ ОБРАЗОВ

Специальность 05.13.17 — Теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата физико–математических наук

Красноярск 2018

Работа выполнена в федеральном государственном бюджетном образовательном учреждении высшего образования «Сибирский государственный университет науки и технологий имени академика М. Ф. Решетнёва», г. Красноярск.

Научный руководитель: доктор физико–математических наук, профессор **Попов Алексей Михайлович**.

Официальные оппоненты: **Белов Алексей Яковлевич**, доктор физико–математических наук, доцент, Федеральное государственное автономное образовательное учреждение высшего образования «Московский физико–технический институт (государственный университет)», лаборатория продвинутой комбинаторики и сетевых приложений, главный научный сотрудник;

Кузнецов Степан Львович, кандидат физико–математических наук, Федеральное государственное бюджетное учреждение науки Математический институт им. В. А. Стеклова Российской академии наук, отдел математической логики, научный сотрудник.

Ведущая организация: Федеральное государственное бюджетное учреждение науки Институт программных систем им. А. К. Айламазяна Российской академии наук, г. Переславль–Залесский.

Защита состоится «20» декабря 2018 г. в 14.00 часов на заседании диссертационного совета Д 212.099.22 на базе Сибирского федерального университета по адресу: 660074, г. Красноярск, ул. Акад. Киренского, 26, аудитория УЛК 112.

С диссертацией можно ознакомиться в библиотеке и на сайте Сибирского федерального университета по адресу http://www.sfu-kras.ru/.

Автореферат разослан «___» ноября 2018 г.

Учёный секретарь
диссертационного совета Покидышева Людмила Ивановна

Общая характеристика работы

Актуальность темы и степень её разработанности. Основы теории формальных языков и грамматик были заложены в 50–х годах прошлого века выдающимся американским лингвистом Н. Хомским. В частности, он ввёл класс контекстно–свободных языков, которые адекватно описывают как естественные языки, так и языки программирования.

В теории формальных языков и грамматик исходным объектом является алфавит, т. е. множество символов $x_1, \ldots, x_m, z_1, \ldots, z_n$; над ними определены операции конкатенации, формальной суммы и умножения на числа. Символы $x_1, \ldots, x_m$ называются терминальными символами и интерпретируются как словарь формального языка, а символы $z_1, \ldots, z_n$ называются нетерминальными и нужны для задания совокупности грамматических правил (грамматики), порождающих язык. Грамматика задаёт «правильные» мономы от терминальных символов $x_1, \ldots, x_m$, которые интерпретируются как правильные предложения естественного языка либо правильные программы языка программирования, а язык рассматривается как формальный степенной ряд (ФСР), т. е. как сумма всех правильных мономов языка, порождённых его грамматикой.

Н. Хомский, используя совокупность правил подстановки, заменяющих один нетерминальный символ на целый моном от символов $x_1, \ldots, x_m, z_1, \ldots, z_n$, смоделировал возможность многократного включения в тексты на естественных языках или языках программирования новых текстовых фрагментов, что отражает возможность итерирования в языках сложноподчинённых и сложносочиненных предложений либо программ, содержащих подпрограммы.

Теория формальных языков и грамматик активно развивалась, и за прошедшие десятилетия опубликованы тысячи статей и сотни книг (N. Chomsky, M. P. Schützenberger, Sh. Greibach, E. Shamir, Y. Bar-Hillel, A. Salomaa, M. Soittola и мн. др.). Значительный вклад в эту теорию внесли советские и российские учёные (В. М. Глушков, А. В. Гладкий, Ан. А. Мучник, М. Р. Пентус, А. Л. Семёнов и др.). Изучались различные классы формальных языков и порождающих их грамматик: контекстно–свободные, линейные языки и грамматики, языки и грамматики непосредственно составляющих, языки в нормальной форме Грейбах, грамматики Ламбека и др. Исследования этих классов языков и грамматик продолжаются и в настоящее время (А. Я. Белов, С. Л. Кузнецов, А. Е. Пентус, В. М. Поляков, Н. С. Рыжкова, Ю. Д. Рязанов, Ю. В. Саватеев, К. В. Сафонов и др.).

Изучение языков и синтаксической структуры их мономов традиционно проводится методами, зависящими от класса порождающей грамматики.

Однако, по нашему мнению, грамматики, порождающие основные классы формальных языков, обладают общим свойством: они могут быть записаны в виде системы полиномиальных уравнений с некоммутативными переменными

$$P_j(z,x) = 0,\ P_j(0,0) = 0,\ j = 1,\dots,k, \tag{1}$$

которые решаются относительно символов $(z_1,\dots,z_n) = z$ в виде ФСР, зависящих от символов $(x_1,\dots,x_m) = x$.

Систему уравнений (1) назовём *полиномиальной* грамматикой, а первую компоненту её решения в виде ФСР — *полиномиальным* языком. Такое представление грамматики позволяет изучать полиномиальные языки для различных типов полиномов $P_j(z,x)$, входящих в систему, а также общие свойства таких систем.

Однако системы полиномиальных уравнений с некоммутативными переменными изучены мало. Известно лишь для частного случая системы уравнений Хомского — Щутценберже, определяющей контекстно–свободный язык, что она имеет единственное решение в виде ФСР. В общем случае условия разрешимости систем неизвестны — трудности связаны с тем, что исключению неизвестных препятствует как некоммутативность умножения, так и отсутствие операции деления, и потому применяемые ранее в теории языков и грамматик методы алгебры и дискретной математики недостаточны для исследования таких систем.

В связи с этим актуально расширение спектра методов исследования полиномиальных грамматик (систем полиномиальных уравнений с некоммутативными переменными) и порождаемых ими полиномиальных языков за счёт привлечения в эту теорию эффективных методов многомерного комплексного анализа и алгебраической геометрии, которые используются при изучении систем полиномиальных уравнений с коммутативными переменными, принимающими значения из поля комплексных чисел.

Для этого необходимо, во–первых, исследовать коммутативный образ полиномиальной грамматики, который получается в предположении, что все символы $x_1,\dots,x_m,z_1,\dots,z_n$, входящие в систему уравнений, являюся коммутативными комплексными переменными, во–вторых, установить, какие свойства исходной некоммутативной системы уравнений вытекают из свойств её коммутативного образа.

Решение указанных задач позволило бы получить новые методы исследования формальных языков и грамматик из различных классов, изучать с помощью этих методов языки и синтаксичскую структуру их мономов.

Объект исследования. Формальные языки и грамматики.

Предмет исследования. Полиномиальные языки и грамматики.

Целью диссертационной работы является установление условий разрешимости полиномиальных грамматик, а также разработка нового метода синтаксического анализа мономов контекстно–свободных языков, как подкласса полиномиальных языков, на основе исследования коммутативных образов языков и грамматик.

Поставленная цель достигается путем решения следующих **задач:**

а) установить связь между множеством решений полиномиальной грамматики и множеством решений её коммутативного образа;

б) найти условие существования и единственности полиномиального языка, порождённого полиномиальной грамматикой;

в) исследовать полиномиальные грамматики, порождающие бесконечное число языков, а также несовместные полиномиальные грамматики;

г) разработать новый метод синтаксического анализа мономов контекстно–свободных языков, образующих подкласс полиномиальных языков.

Соответствие диссертации паспорту специальности. Диссертация соответствует области исследований специальности 05.13.17 — Теоретические основы информатики по п. 10 «*Разработка основ математической теории языков и грамматик*, теории конечных автоматов и теории графов».

Методы исследования диссертационной работы. Основные результаты получены на основе методов алгебры, в том числе алгебры многочленов, дискретной математики, теории формальных языков и грамматик, а также комплексного анализа, включая методы теории интегральных представлений и теории вычетов.

Научная новизна:

а) впервые установлена связь между множествами решений полиномиальной грамматики и её коммутативного образа, которая позволяет исследовать полиномиальные языки и грамматики методами комплексного анализа и алгебраической геометрии;

б) найдено новое условие существования и единственности полиномиального языка, порождённого полиномиальной грамматикой в терминах её коммутативного образа. Условие легко проверяется при исследовании полиномиальных грамматик;

в) введено и исследовано понятие полиномиальной грамматики, порождающей бесконечное число языков, получены новые свойства несовместных полиномиальных грамматик, не имеющие аналогов для коммутативных переменных. Тем самым расширена совокупность исследуемых языков и грамматик;

г) разработан новый метод синтаксического полинома (который получен в виде интеграла по циклу от рациональной функции) для решения задачи синтаксического анализа мономов контекстно–свободных языков программирования. Данный метод может быть использован на практике.

Теоретическая и практическая значимость работы. Работа носит теоретический характер. Результаты могут быть использованы в теоретическом программировании и при разработке трансляторов. Предложенный метод синтаксического полинома может быть использован при разработке языков программирования.

Положения, выносимые на защиту диссертационной работы. На защиту выносятся следующие основные результаты:

а) описание связи между множествами полиномиальных языков, порождённых полиномиальной грамматикой, и решений её коммутативного образа;

б) условие существования и единственности полиномиального языка, порождённого полиномиальной грамматикой в терминах её коммутативного образа;

в) совокупность свойств полиномиальных грамматик, порождающих бесконечное число языков, а также несовместных полиномиальных грамматик;

г) метод синтаксического полинома для решения задачи синтаксического анализа языков программирования.

Достоверность результатов работы подтверждается математическими доказательствами основных положений.

Апробация результатов работы. Результаты диссертационной работы были доложены автором на следующих конференциях:

— Всероссийской конференции «Сибирская научная школа–семинар с международным участием "Компьютерная безопасность и криптография"» — Sibecrypt (Томск, 2006 г.; Новосибирск, 2016, 2017 гг.; Абакан 2018 г.);

— Международной научно–практической конференции «Решетнёвские чтения» (Красноярск, 2009, 2015 — 2018 гг.).

Результаты работы обсуждались на научно-исследовательских семинарах в Сибирском федеральном университете, Сибирском государственном университете науки и технологий имени академика М. Ф. Решетнёва, Национальном-исследовательском Томском государственном университете.

Публикации по теме диссертационной работы. По результатам диссертационного исследования опубликовано 16 работ, из которых 4 изданы в журналах, рекомендованных ВАК (из них 2 журнала входят также в базы цитирования Web of Science и Scopus), 12 – в материалах конференций и в других изданиях. Из статей, написанных в соавторстве, в диссертацию включены результаты, принадлежащие лично автору.

Структура и объём работы. Диссертационная работа изложена на 105 страницах и состоит из введения, двух глав, заключения и списка литературы.

Содержание диссертации

Во **введении** обоснована актуальность темы диссертационной работы, определены цель и задачи исследования, указаны применяемые в работе методы, представлены основные результаты.

В **первой главе** «Условия разрешимости полиномиальных грамматик» рассмотрены вопросы, связанные с существованием и единственностью полиномиальных языков, полиномиальными грамматиками, порождающими бесконечное множество языков, а также не порождающими никаких языков (решаются задачи а) — в) диссертационного исследования). Основные результаты первой главы опубликованы в работах [2—4].

Рассматривается коммутативный образ $ci(r)$ полинома или ФСР r от символов $x_1, \ldots, x_m, z_1, \ldots, z_n$ при условии, что все эти символы являются переменными, принимающими значения из поля комплексных чисел. Таким образом, коммутативный образ полиномиальной грамматики является системой полиномиальных уравнений в пространстве $\mathbb{C}^{n+m}_{z,x}$, а коммутативный образ ФСР становится кратным степенным рядом, представляющим алгебраическую функцию в $\mathbb{C}^m_x$, которая является решением этой системы. Впервые коммутативный образ ФСР был рассмотрен А. Л. Семёновым при решении алгоритмических проблем контекстно–свободных языков и грамматик.

Фундаментальная связь между множествами решений системы уравнений с некоммутативными переменными и её коммутативного образа описывает следующая теорема.

Теорема 1.1. *Если*

$$z = z(x) = (z_1(x), \ldots, z_n(x))$$

— решение некоммутативной системы уравнений (1) в виде символьных ФСР, то коммутативные кратные степенные ряды

$$z = ci(z(x)) = (ci(z_1(x)), \ldots, ci(z_n(x)))$$

над полем комплексных чисел сходятся в окрестности нуля, определяя ростки голоморфных алгебраических функций, и являются решением коммутативной системы уравнений

$$ci(P_j(z, x)) = 0, \ ci(P_j(0, 0)) = 0, \ j = 1, \ldots, k. \tag{2}$$

Согласно теореме 1.1, если некоммутативная система (1) совместна, то коммутативная система (2) имеет голоморфное в начале координат решение. Обратное, вообще говоря, неверно, что показывает следующий пример.

Система уравнений

$$z_1 - z_2 = x_1 x_2, \ z_1 - z_2 = x_2 x_1$$

является несовместной, а её коммутативный образ имеет решение:

$$z_1 = s + x_1 x_2,\ z_2 = s,$$

где s — произвольный коммутативный ФСР.

Таким образом, можно сделать следующие выводы:

1) множество решений системы уравнений (2), вообще говоря, шире, чем множество коммутативных образов решений системы уравнений (1), и из совместности системы уравнений (2) не следует совместность исходной некоммутативной системы (1);

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

Обозначим два множества ростков аналитических множеств в начале координат (множеств нулей аналитических функций в сколь угодно малых окрестностях нуля):

$$ci(S_P) = \{z = ci(z(x)) :\ P_j(z(x), x) = 0,\ j = 1, \dots, k\},$$

$$S_{ci(P)} = \{z = z(x) :\ ci(P_j(z(x), x)) = 0,\ j = 1, \dots, k\}.$$

Основываясь на последнем примере и сделанных обозначениях, сформулируем теорему следующим образом.

Теорема 1.2. *Имеет место включение*

$$ci(S_P) \subseteq S_{ci(P)}.$$

Рассмотрим простейший случай полиномиальной зависимости, какой является линейная зависимость.

В связи с этим мы называем рангом матрицы A число её строк, линейно независимых над полем комплексных чисел; обозначим его $rank(A)$.

Имеет место следующая теорема.

Теорема 1.3. *Пусть*

$$A = (a_{ij}(x))_{i=1, j=1}^{k,\, n}$$

— матрица, элементы которой являются многочленами от символьных некоммутативных переменных $x = (x_1, \dots, x_m)$, и матрица

$$ci(A) = (ci(a_{ij}(x)))_{i=1, j=1}^{k,\, n}$$

является её коммутативным образом. Тогда выполнено неравенство

$$rank(ci(A)) \leq rank(A).$$

Отметим также, что теореме 1.1 эквивалентна следующая теорема.

Теорема 1.4. *Если коммутативная система уравнений (2) не имеет решения, голоморфного в начале координат, то некоммутативная система (1) несовместна.*

Таким образом, условия несовместности коммутативной системы уравнений (2) также представляют интерес в теории формальных языков и грамматик.

Наибольший интерес для приложений представляют условия, которые обеспечивают совместность системы некоммутативных символьных уравнений (1), а также единственность её решения.

Как отмечено выше, таким условием не может быть существование голоморфного в начале координат решения системы коммутативных уравнений (2).

Это условие даёт такой инструмент, как якобиан системы функций.

Рассмотрим систему уравнений (1) в слRe: учае, когда $k = n$, и пусть

$$J(z,x) = \det\left(\frac{\partial ci(P_i(z,x))}{\partial z_j}\right) =$$

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

— якобиан системы уравнений (1) относительно переменных $z_1, \ldots, z_n$.

Символьным или дискретным аналогом известной теоремы о неявном отображении является следующая теорема.

Теорема 1.5. *Если для полиномиальной грамматики (некоммутативной символьной системы уравнений) (1) выполнено неравенство*

$$J(0,0) \neq 0,$$

то она имеет единственное решение в виде ФСР.

Как известно, неравенство нулю якобиана является условием теоремы о неявном отображении для коммутативной системы уравнений (2) с переменными в пространстве $\mathbb{C}^{n+m}_{z,x}$ и влечёт существование и единственность её голоморфного решения. Оказывается, что это неравенство влечёт также существование и единственность решения исходной некоммутативной символьной системы уравнений (1).

Неравенство $J(0,0) \neq 0$ обусловлено свойствами линейных частей некоммутативных многочленов из системы уравнений, а они совпадают со своими коммутативными образами. По этой причине неравенство влечёт совместность не только коммутативной, но и некоммутативной системы уравнений.

Далее рассмотрен вопрос о полиномиальных грамматиках, которые могут порождать более одного полиномиального языка.

Уравнения различной природы часто обладают тем свойством, что для них имеет место альтернатива: имеется единственное решение, либо бесконечно много решений. В связи с этим рассмотрен случай, когда полиномиальная грамматика, т. е. система некоммутативных уравнений (1) имеет бесконечно много решений (порождает бесконечно много полиномиальных языков).

Для уравнений с числовыми неизвестными наличие бесконечного числа решений обычно означает, что общее решение этого уравнения зависит от одного или нескольких числовых параметров. Для полиномиальных грамматик ситуация в определённом смысле аналогичная, однако смысл параметров иной.

Дадим следующее определение.

Определение 1.2. *Будем говорить, что полиномиальная грамматика (1) имеет бесконечно много решений (порождает бесконечно много языков), если множество решений системы (1) зависит хотя бы от одного произвольного ФСР от символов* $x_1, \ldots, x_m$.

Так, система из двух одинаковых уравнений

$$x_1 z_1 - z_2 x_2 = 0$$

имеет бесконечно много решений, поскольку её решения можно записать в виде

$$z_1 = s x_2, \; z_2 = x_1 s$$

, где s — произвольный ФСР от $x_1, \; x_2$.

Пусть, как выше,

$$J(z,x) = \det\left(\frac{\partial(ci(P_i(z,x)))}{\partial z_j}\right).$$

— якобиан системы уравнений (2) относительно переменных $z_1, ..., z_n$.

Для систем уравнений с комплексными переменными известна следующая теорема.

Теорема. *Пусть выполнено равенство*

$$J(z,x) \equiv 0,$$

тогда система уравнений (2) либо не имеет решения для каждого x *в пространстве* $\mathbb{C}^n_z$, *либо все её решения в этом пространстве — неизолированные.*

Рассмотрен вопрос об аналоге этой теоремы для случая полиномиальных грамматик и языков, представленных ФСР. Показано, что для некоммутативной системы уравнений (1) такая альтернатива (нет решений — бесконечно много решений) не имеет места.

Рассмотрена система, состоящая из двух одинаковых уравнений:

$$x_1z_1 + x_2z_2 - x_1x_2 - x_2x_1 = 0.$$

Её коммутативный образ имеет якобиан, тождественно равный нулю, тем не менее исходная некоммутативная система имеет единственное решение.

В самом деле, записывая систему уравнений в виде

$$x_1(z_1 - x_2) + x_2(z_2 - x_1) = 0,$$

получим, что первое слагаемое в этой сумме принадлежит левому идеалу, порождённому x_1, и второе слагаемое принадлежит левому идеалу, порождённому x_2. Очевидно, что сумма этих слагаемых может быть равной нулю только в случае, когда оба слагаемых равны нулю:

$$z_1 - x_2 = 0, z_2 - x_1 = 0,$$

следовательно, исходная система уравнений имеет единственное решение $z_1 = x_2, z_2 = x_1$.

Кроме того, пример системы из двух одинаковых уравнений

$$x_1(z_1 - x_1)(z_1 - x_2) + x_2(z_2 - x_1)(z_2 - x_2) = 0,$$

имеющей четыре решения

$$z_1 = x_1, z_2 = x_1;\ z_1 = x_1, z_2 = x_2;\ z_1 = x_2, z_2 = x_1;\ z_1 = x_2, z_2 = x_2,$$

показывает, что число рещений некоммутативной системы (коммутативный образ которой имеет якобиан, тождественно равный нулю) может быть любым.

Таким образом, в случае, когда якобиан коммутативного образа тождественно равен нулю, исходная некоммутативная система может:

1) не иметь решений;
2) иметь любое конечное число решений;
3) иметь бесконечно много решений.

Учитывая, что система

$$f = 0, \dots, f = 0,$$

из n одинаковых уравнений с n некоммутативными неизвестными $z_1, \dots, z_n$, имеющая тождественно равный нулю якобиан, эквивалентна одному уравнению $f = 0$, сформулируем следующий вывод.

Одно уравнение с некоммутативными неизвестными $P_1(z.x) = 0$ может: не иметь решений, а также иметь конечное и бесконечное число решений.

В этом состоит фундаментальное отличие от одного уравнения над полем комплексных чисел, которое всегда имеет решения в виде ростков аналитических функций.

Подтверждением того, что одно уравнение может не иметь решений, является следующий пример: уравнение

$$x_1(z_1 - x_1) + x_2(z_1 - x_2) = 0.$$

Первое слагаемое принадлежит левому идеалу, порождённому x_1, а второе — левому идеалу, порождённому x_2. Очевидно, что сумма этих слагаемых может быть равной нулю только в случае, когда оба слагаемых равны нулю:

$$z_1 - x_1 = 0, z_2 - x_2 = 0.$$

Значит, уравнение имеет решение

$$z_1 = x_1, z_1 = x_2,$$

что невозможно, т. е. уравнение не имеет решений.

Причина этого эффекта состоит в том, что некоммутативное полиномиальное уравнение может быть эквивалентно системе полиномиальных уравнений, а система уравнений может быть несовместной.

Так, рассмотренное уравнение вида

$$x_1 A_1 + x_2 A_2 = 0$$

равносильно системе полиномиальных уравнений $A_1 = A_2 = 0$.

В случае же комплексных переменных, как было отмечено выше, такого эффекта нет.

Во **второй главе** «Синтаксический анализ языков программирования методом интегрального представления синтаксических полиномов» предложен и обоснован новый метод решения проблемы синтаксического анализа программ — одной из важных проблем, связанных с языками программирования (решается задача г) диссертационного исследования). Основные результаты второй главы опубликованы в [1].

Отметим, что большинство языков программирования является контекстно–свободными языками (кс–языками), которые можно представить в виде ФСР, поэтому каждая программа, написанная на языке программирования, может рассматриваться как моном соответствующего ФСР.

Сформулируем проблему синтаксического анализа мономов кс–языка, грамматика которого является совокупностью правил подстановки (продукций):

$$z_j \to \; q_{jk}(z, x), \; j = 1, \ldots, n, \; k = 1, \ldots, p_j,$$

где $q_{jk}(z,x)$ — заданные мономы; правила подстановки можно применять к начальному символу z_1, а затем к другим символам в мономам неограниченное число раз в любом порядке, что позволяют выводить новые мономы, которые и образуют кс–язык. Известно также, что данный кс–язык $z_1(x)$ удовлетворяет системе полиномиальных уравнений Хомского — Щутценберже.

Проблема синтаксического анализа мономов состоит в том, чтобы решить обратную задачу: определить, принадлежит ли моном данному кс–языку, т. е. может ли быть получен из начального символа z_1 при помощи правил подстановки (этап синтаксического контроля), а также установить, какие правила подстановки и сколько раз использовались при выводе этого монома (этап синтаксического разбора), при этом порядок использования правил подстановки не имеет значения.

Для решения проблемы синтаксического анализа известен метод мономиальных меток, в котором рассматривается расширенная система уравнений Хомского — Щутценберже

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

где t_{jk} — символ из расширенного алфавита, играющий роль мономиальной метки соответствующего правила вывода: $z_j \to t_{jk}q_{jk}(z,x)$.

Решение этой системы можно получить методом последовательных приближений, о котором говорилось выше, в виде ФСР, в том числе первую компоненту (собственно расширенный кс–язык)

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

где w_i — мономы от символов $x_1, \ldots, x_m, t_{11}, t_{12}, \ldots, t_{np_n}$.

Синтаксический анализ монома v кс–языка $z_1(x)$ методом мономиальных меток проводится следующим образом. Считывая мономы степени, равной степени монома v относительно символов $x_1, \ \ldots \ , x_m$, и пропуская символы t_{jk}, можно установить, есть ли среди них моном v, а значит, можно ли вывести его с помощью системы продукций, а каждая мономиальная метка t_{jk}, содержащаяся в таком мономе, показывает, что использовалось правило $z_j \to t_{jk} \ q_{jk}(z,x)$. Таким образом, этот метод позволяет провести за конечное число шагов беступиковый синтаксический анализ любого монома (программы) кс–языка, заданного грамматикой, решая проблему синтаксического анализа: мономиальные метки показывают, какие правила вывода кс–языка и сколько раз использовались при выводе этого монома, с точностью до порядка их применения.

Однако, недостатком метода мономиальных меток является большое число громоздких итераций метода последовательных приближений, необходимых для получения начальных членов ФСР, причём это число возрастает вместе с ростом степени монома v. В связи с этим в диссертации метод мономиальных меток развивается: предложен и обоснован метод синтаксического полинома, позволяющий по–другому получить мономиальные метки некоммутативного монома.

А именно, рассмотрим коммутативный образ ФСР $z_1^*(x,t)$, сгруппированный в кратный ряд Гартогса по степеням x:

$$ci(z_1^*(x,t)) = \sum_{\alpha} s_\alpha(t)\ x^\alpha\ . \qquad (3)$$

Доказана следующая лемма.

Лемма 2.1. *При всех мультиндексах α голоморфные в нуле коэффициенты $s_\alpha(t)$ ряда Гартогса (3) являются полиномами.*

В связи с этим дано следующее определение.

Определение 2.1. *Синтаксическим полиномом монома v относительно кс–языка $z_1(x) = z_1^*(x,e)$ называется коэффициент $s_\alpha(t)$ ряда Гартогса (3), такой, что $x^\alpha = ci(v)$.*

В диссертации доказано, что мономиальные метки, содержащиеся в некоммутативных мономах кс–языка, не исчезают при переходе от ФСР $z_1^*(x,t)$ к его коммутативному образу $ci(z_1^*(x,t))$ и сохраняются в виде мономов синтаксических полиномов.

Доказано также, что если синтаксический полином монома относительно кс–языка равен нулю, то моном не принадлежит этому языку.

Метод синтаксического полинома состоит в следующем. Для монома v, такого, что $ci(v) = x^\alpha$, следует найти синтаксический полином $s_\alpha(t)$, затем остаётся проверить, можно ли получить исследуемый моном с помощью правил подстановки, соответствующих всем мономам синтаксического полинома $s_\alpha(t)$.

Доказано, что синтаксические полином $s_\alpha(t)$ монома, а значит, и содержащиеся в нём мономиальные метки этого монома можно получить в виде интеграла фиксированной кратности $n+m$ по циклу, где числа n и m не зависят от степени монома и равны соответственно числу нетерминальных и терминальных символов грамматики кс–языка.

А именно, доказана следующая теорема.

Теорема 2.2. *При всех t, достаточно близких к нулю, и всех мультииндексах α синтаксический полином $s_\alpha(t)$ задаётся равенством*

$$s_\alpha(t) = \frac{1}{(2\pi i)^{n+m}} \int\limits_{\gamma_z \times \Gamma_x} \frac{z_1 \det\bigl(\delta_{ij} - (ci(Q_i^*(z,x,t)))'_{z_j}\bigr)\ dz \wedge dx}{(z - ci(Q^*(z,x,t)))\ x^{\alpha+I}}, \qquad (4)$$

где $\gamma_z = \{|z_1| = \dots = |z_n| = \epsilon\}$ и $\Gamma_x = \{|x_1| = \dots = |x_m| = \delta\}$ — циклы интегрирования, $0 < \delta \ll \epsilon \ll 1$, $dz = dz_1 \wedge \dots \wedge dz_n$, $dx = dx_1 \wedge \dots \wedge dx_m$,

$x^{\alpha+I} = x_1^{\alpha_1+1} \cdot \ \cdots \ \cdot x_m^{\alpha_m+1}$, $(z - ci(Q^*(z,x,t))) = (z_1 - ci(Q_1^*(z,x,t))) \cdot \cdots$ $\cdots \cdot (z_n - ci(Q_n^*(z,x,t)))$, *и* δ_{ij} *— дельта Кронекера.*

Ещё одно интегральное представление синтаксического многочлена даёт следующая теорема.

Теорема 2.3. *При всех* t*, достаточно близких к нулю, и всех мультииндексах* α *синтаксический полином* $s_\alpha(t)$ *задаётся равенством*

$$s_\alpha(t) = \frac{1}{(2\pi i)^n \alpha!} \int\limits_{\gamma_z} \frac{\partial^{\|\alpha\|}}{\partial x^\alpha} \Big(\frac{z_1 \det\big(\delta_{ij} - (ci(Q_i^*(z,x,t)))'_{z_j}\big)}{z - ci(Q^*(z,x,t))} \Big)\Big|_{x=0} dz, \qquad (5)$$

где $\gamma_z = \{|z_1| = \cdots = |z_n| = \epsilon\}$ *— цикл интегрирования,* $0 < \epsilon \ll 1$, $dz = dz_1 \wedge \cdots \wedge dz_n$, $\alpha! = \alpha_1! \cdot \ \cdots \ \cdot \alpha_m!$, $z - ci(Q^*(z,x,t)) = (z_1 - ci(Q_1^*(z,x,t))) \cdot \cdots$ $\cdots \ \cdot (z_n - ci(Q_n^*(z,x,t)))$, $\frac{\partial^{\|\alpha\|}}{\partial x^\alpha} = \frac{\partial^{\alpha_1+\cdots+\alpha_m}}{\partial x_1^{\alpha_1} \cdot \ \ldots \ \cdot \partial x_m^{\alpha_m}}$, *и* δ_{ij} *— дельта Кронекера.*

Доказательства теорем 2.2 и 2.3 опираются на следующую лемму.

Лемма 2.4. *При всех* (x,t)*, достаточно близких к нулю, голоморфная в нуле алгебраическая функция* $ci(z_1^*(x,t))$*, представленная рядом Гартогса (3), задаётся формулой*

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

Формулы (4) и (5) из теорем 2.2. и 2.3 дают практическую возможность проводить синтаксический анализ монома (программы), что подтверждается разделом 2.4 диссертационной работы, в котором подробно рассмотрен синтаксический анализ программы, длина которой может быть сколь угодно большой, проведённый методом синтаксического полинома. Этот метод сравнивается с методами синтаксического анализа свёрткой и развёрткой, число операций в которых возрастает вместе с ростом длины программы.

Основные результаты и выводы

В заключении диссертации представлены следующие основные результаты и выводы.

Установлена связь между множествами решений полиномиальной грамматики и её коммутативного образа (**Определение 1.1**, **Теоремы 1.1 — 1.4**, **Пример 1.1**, **Замечания 1.1** и **1.2**). Тем самым в теорию формальных языков и грамматик привлечены новые для этой теории математические методы: комплексного анализа и алгебраической геометрии.

Найдено условие существования и единственности полиномиального языка, порождённого полиномиальной грамматикой (системой уравнений с некоммутативными символьными переменными) в терминах её коммутативного образа (**Теорема 1.5**, **Замечания 1.2** и **1.3**). Условие формулируется как теорема о неявном отображении, которая обычно используется в случае коммутативных переменных.

Введено и исследовано понятие полиномиальной грамматики, порождающей бесконечное число языков, а также получены новые свойства несовместных полиномиальных грамматик, не имеющие аналогов для коммутативного случая (**Определение 1.2**, **Примеры 1.2 — 1.4**, **Замечание 1.5** и **1.6**). Таким образом, расширяется совокупность полиномиальных грамматик, исследуемых в теории формальных языков и грамматик.

Разработан метод синтаксического полинома для решения задачи синтаксического анализа мономов контекстно-свободных языков программирования (**Определение 2.1**, **Леммы 2.1** и **2.4**, **Теоремы 2.2** и **2.3**, **Пример 2.1**). Синтаксический полином получен в виде интеграла по циклу от рациональной функции и может быть использован на практике в процессе синтаксического анализа программ.

Полученные результаты позволяют вести исследования в теории формальных языков и грамматик в направлениях:

– выявления новых условий существования полиномиальных языков и изучения их свойств, включая поиск критерия существования решения полиномиальных грамматик в виде ФСР;

– разработка новых методов синтаксического анализа формальных языков, которые могут применяться как в теоретическом программировании, так и при практическом осуществлении синтаксического анализа.

Публикации по теме диссертации

В изданиях, рекомендованных ВАК:

1. Егорушкин, О. И. О применении многомерного комплексного анализа в теории формальных языков и грамматик / О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов // Прикладная дискретная математика. — 2017. — Т. 37. — № 2. — С. 76—89.

2. Егорушкин, О. И. On solvability of systems of symbolic polynomial equations / О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов // Журнал Сибирского федерального университета. Математика и физика. — 2016. — Т. 9. — № 2. — С. 166—172.

3. Егорушкин, О. И. О разрешимости систем некоммутативных алгебраических уравнений, порождающих контекстно-свободные языки / О. И. Егорушкин, Д. А. Калугин-Балашов, К. В. Сафонов // Вестник Сибирского государственного аэрокосмического университета имени академика М. Ф. Решетнёва. — 2009. — № 2 (23). — С. 21—24.

4. Егорушкин, О. И. Контекстно-свободные языки в нормальной форме Грейбах: аналитический подход / О. И. Егорушкин // Вестник Си-

бирского государственного аэрокосмического университета имени академика М. Ф. Решетнёва. — 2008. — № 2 (19). — С. 24—25.

В материалах конференций и других изданиях:

5. Егорушкин, О. И. Синтаксический анализ программ методом интегральных представлений / О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов / Материалы 17 Всероссийской конференции «Сибирская научная школа–семинар с международным участием "Компьютерная безопасность и криптография"» — SIBECRYPT'18. Абакан. — 2018 // Прикладная дискретная математика. Приложение. — 2018. — № 11. — С. 128—130.

6. Егорушкин, О. И. Интегральное представление синтаксического многочлена О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов // Материалы XXI Международной научно–практической конференции «Решетнёвские чтения». Т. 2. Красноярск. — 2017. — С. 75—76.

7. Егорушкин, О. И. Аналог теоремы о неявном отображении для формальных грамматик / О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов / Материалы 16 Всероссийской конференции «Сибирская научная школа–семинар с международным участием "Компьютерная безопасность и криптография"» — SIBECRYPT'17. Новосибирск. — 2017 // Прикладная дискретная математика. Приложение. — 2017. — № 10. — С. 149—151.

8. Егорушкин, О. И. О совместности систем символьных полиномиальных уравнений и их приложении / О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов / Материалы 15 Всероссийской конференции «Сибирская научная школа–семинар с международным участием "Компьютерная безопасность и криптография"» — SIBECRYPT'16. Новосибирск. — 2016 // Прикладная дискретная математика. Приложение. — 2016. — № 9. — С. 119—121.

9. Егорушкин, О. И. Условия совместности систем символьных уравнений и приложения / О. И. Егорушкин, И. В. Колбасина, К. В. Сафонов // Материалы Международной научно–практической конференции «Актуальные проблемы авиации и космонавтики». Т. 1. Красноярск. — 2016. — С. 504—506.

10. Егорушкин, О. И. Об одном подходе к решению систем некоммутативных полиномиальных уравнений / О. И. Егорушкин , И. В. Колбасина, А. М. Попов, Н. А. Попов, К. В. Сафонов // Материалы XX Международной научно–практической конференции «Решетнёвские чтения». Т. 2. Красноярск. — 2016. — С. 100—102.

11. Егорушкин, О. И. Символьный аналог теоремы о неявном отображении и его приложения в теоретической информатике / О. И. Егорушкин, И. В. Колбасина, А. М. Попов, Н. А. Попов, К. В. Сафонов // Материалы XX Международной научно–практической конференции

«Решетнёвские чтения». Т. 2. Красноярск. — 2016. — С. 98—99.

12. Егорушкин, О. И. О решении систем некоммутативных полиномиальных уравнений и приложении результатов / О. И. Егорушкин, И. В. Колбасина, И. Р. Насыров, К. В. Сафонов // Материалы XIX Международной научно-практической конференции «Решетнёвские чтения». Т. 2. Красноярск. — 2015. — С. 124—125.

13. Егорушкин, О. И. Аналитический подход в теории контекстно-свободных языков в нормальной форме Грейбах / О. И. Егорушкин, К. В. Сафонов // Прикладная дискретная математика. — 2009. — № 3 (5). — С. 112—116.

14. Егорушкин, О. И. О решении систем алгебраических уравнений, ассоциированных с контекстно-свободными языками / О. И. Егорушкин, Д. А. Калугин-Балашов, К. В. Сафонов // Прикладная дискретная математика. — 2008. — № 2 (2). — С. 8—11.

15. Егорушкин, О. И. Некоторые свойства контекстно-свободных языков / О. И. Егорушкин // Материалы III Всероссийской научно-технической конференции студентов, аспирантов и молодых ученых «Молодежь и наука: начало XXI века». Красноярск. — 2007. — С. 65—67.

16. Егорушкин, О. И. О синтаксическом анализе и проблеме В. М. Глушкова распознавания контекстно-свободных языков Хомского / К. В. Сафонов, О. И. Егорушкин // Вестник Томского государственного университета. Приложение. — 2006. — № 17. — С. 63—66.

Автор выражает свою **благодарность** научному руководителю Попову Алексею Михайловичу за неоценимую помощь в работе, а также всему коллективу кафедры прикладной математики СибГУ им. М. Ф. Решетнёва за постоянную поддержку.