На правах рукописи

Кишкан Владимир Владимирович

РАЗРАБОТКА И ОЦЕНКА СЛОЖНОСТИ АЛГОРИТМОВ, НАХОДЯЩИХ ПРИМЕНЕНИЕ В АППАРАТНОМ И ПРОГРАММНОМ ОБЕСПЕЧЕНИИ МНОГОПРОЦЕССОРНЫХ СИСТЕМ

Специальность 05.13.17 – Теоретические основы информатики

АВТОРЕФЕРАТ
диссертации на соискание учёной степени
кандидата физико-математических наук

Красноярск 2020

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

Научный руководитель:	доктор физико-математических наук, профессор **Кузнецов Александр Алексеевич**.
Официальные оппоненты:	**Винокуров Сергей Федорович**, доктор физико-математических наук, профессор, Федеральное государственное бюджетное образовательное учреждение высшего образования «Иркутский государственный университет», кафедра алгебраических и информационных систем, профессор;
	Овчаренко Алёна Юрьевна, кандидат технических наук, Федеральное государственное бюджетное образовательное учреждение высшего образования «Сибирский государственный университет телекоммуникаций и информатики» (г. Новосибирск), кафедра высшей математики, доцент.
Ведущая организация:	Федеральный исследовательский центр «Информатика и управление» Российской академии наук, г. Москва.

Защита состоится «22» декабря 2020 г. в 14.30 часов на заседании диссертационного совета Д 212.099.22, созданного 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 2020 г.

Учёный секретарь
диссертационного совета Покидышева Людмила Ивановна

Общая характеристика работы

Актуальность темы и степень её разработанности. В настоящее время постоянно увеличивающийся спрос на облачные вычисления приводит к росту крупномасштабных центров обработки данных (ЦОД). Современные ЦОД содержат сотни тысяч узлов, соединенных между собой сетью. Топология такой сети, т.е. способ соединения узлов, является ключевым звеном, от которого зависит быстродействие, отказоустойчивость, надежность и другие характеристики ЦОД. По этой причине проектирование сети является очень важной задачей, включающей в себя поиск моделей графов, которые имеют хорошие топологические свойства и позволяют использовать эффективные алгоритмы маршрутизации. Этими качествами обладают графы Кэли, имеющие такие привлекательные топологические свойства как высокая симметрия, иерархическая структура, рекурсивная конструкция, высокая связность и отказоустойчивость. Определение графа Кэли подразумевает, что вершины графа являются элементами некоторой алгебраической группы. Выбор группы и ее порождающих элементов позволяет получить граф, отвечающий необходимым требованиям по диаметру, степени вершин, количеству узлов и т. д.

К сожалению, в общем случае задача вычисления кратчайшего пути на графе Кэли, согласно классическому результату С. Ивена и О. Голдрейха 1981 года, является NP–трудной. Отметим, что существующие на сегодняшний день алгоритмы маршрутизации на графах Кэли могут быть отнесены в одну из следующих категорий: а) те, которые предназначены для конкретных графов Кэли; б) универсальные с высокой пространственно-временной сложностью и в) с низкой сложностью, которые не обеспечивают кратчайших путей.

Поэтому, разработка эффективных алгоритмов маршрутизации на графах Кэли является актуальной задачей.

Для многопроцессорных вычислительных систем (МВС) актуален ряд направлений исследований. Одним из них являются исследования, связанные с перспективными языками программирования, которые могут разрабатываться в дальнейшем для обеспечения работы с МВС.

Практически все известные в настоящее время языки программирования являются контекстно–свободными языками (кс–языками), порождёнными контекстно–свободными грамматиками (кс–грамматиками).

Однако, с точки зрения разработки перспективных языков программирования, в том числе для МВС, известные алгоритмы синтаксического анализа не применимы. Как правило, известные алгоритмы реализованы в виде специальных программ (парсеров), предназначенных для анализа выражений, написанных на определённом языке программирования.

Заметим, что в ситуации, когда разрабатывается новый язык программирования, никаких парсеров, естественно, нет. В случае, когда необходимо провести синтаксический анализ (разбор) некоторого выражения относительно совокупности грамматических правил, находящихся в стадии разработки, могут быть полезными различные алгоритмы, в том числе имеющие высокую сложность. Как правило, тестируются выражения ограниченной длины, и потому высокая сложность алгоритма в такой ситуации не играет роли – важно, чтобы он был простым в программной реализации и вполне конструктивным. Сложность такого алгоритма может быть даже выше экспоненциальной, что вполне допустимо для случаев, когда длина программы N не слишком велика.

Таким образом, разработка беступиковых конструктивных алгоритмов для решения расширенной проблемы синтаксического анализа является достаточно актуальной задачей.

Объект исследования. Графы Кэли, порождённые конечными группами подстановок, а также контекстно–свободные языки и грамматики (связанные с аппаратно–программным обеспечением МВС).

Предмет исследования. Алгоритмы маршрутизации на графах Кэли, а также алгоритмы решения расширенной проблемы синтаксического анализа мономов контекстно–свободных языков программирования.

Целью диссертационной работы является создание универсального алгоритма маршрутизации на графах Кэли групп подстановок, имеющего лучшие пространственно–временные характеристики по сравнению с известными алгоритмами (теоретические основы аппаратного обеспечения МВС), а также разработка беступикового алгоритма решения расширенной проблемы синтаксического анализа мономов контекстно–свободных языков (теоретические основы программного обеспечения МВС).

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

1. разработать эффективный алгоритм маршрутизации на графах Кэли, заданных конечными группами подстановок;
2. исследовать ранее неизвестные характеристики для некоторых графов Кэли;
3. разработать беступиковый алгоритм решения расширенной проблемы синтаксического анализа мономов контекстно–свободных языков;
4. оценить сложность беступикового алгоритма решения расширенной проблемы синтаксического анализа, позволяющего установить все возможные выводы монома контекстно–свободного языка.

Соответствие диссертации паспорту специальности. Диссертационная работа 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 и теории групп, дискретной математики, теории графов, теории формальных языков и грамматик.

Научная новизна:

1. Разработан новый эффективный алгоритм, позволяющие вычислять кратчайшие маршруты между вершинами графа Кэли $Cay(G, X)$, заданного произвольной конечной группой подстановок $G = \langle X \rangle$. Обоснована корректность представленного алгоритма, даны оценки его сложности. В этом случае для произвольного слова $w \in X^*$ проблема поиска минимального слова разрешается за время $T = O(|w|)$.

2. При помощи МВС вычислены ранее неизвестные характеристики графов Кэли модифицированной пузырьковой сортировки MBS(n) для случаев $n = 14$ и 15. Получены новые результаты о некоторых графах Кэли, заданных конечными двупорождёнными группами периода 7.

3. Впервые разработан беступиковый алгоритм синтаксического анализа, использующий иерархию маркированных скобок и позволяющий решать расширенную проблему синтаксического анализа, включая вычисление всех выводов монома (программы), который находит применение при разработке языков программирования для МВС.

4. Получена оценка сложности алгоритма синтаксического анализа, использующего иерархию маркированных скобок, особенно эффективного при разработке новых языков программирования.

Теоретическая и практическая значимость работы. Работа носит теоретический характер. Новый эффективный алгоритм для вычисления кратчайших маршрутов между вершинами графа Кэли $Cay(G, X)$, порожденного произвольной конечной группой подстановок $G = \langle X \rangle$ может быть использован при разработке аппаратно–программного обеспечения МВС. Предложенный алгоритм решения расширенной проблемы синтаксического анализа может быть эффективно использован при разработке языков программирования, в том числе для обеспечения МВС, в условиях отсутствия парсеров.

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

1. Новый эффективный алгоритм для вычисления кратчайших маршрутов между вершинами графа Кэли $Cay(G, X)$, порожденного произвольной конечной группой подстановок $G = \langle X \rangle$.

2. Ранее неизвестные характеристики некоторых графов Кэли, полученных при помощи МВС.

3. Новый беступиковый алгоритм синтаксического анализа решения расширенной проблемы синтаксического анализа на основе иерархии вложенных скобок.

4. Оценка сложности разработанного беступикового алгоритма синтаксического анализа мономов контекстно–свободных языков на основе иерархии вложенных скобок.

Достоверность результатов работы подтверждается математическими доказательствами основных положений.

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

– Международной научно-практической конференции «Решетнёвские чтения» (Красноярск, 2017–2019 гг.);

– Всероссийской конференции «Сибирская научная школа–семинар с международным участием «Компьютерная безопасность и криптография» – Sibecrypt (Томск, 2019 –2020 гг.);

– Всероссийской с международным участием научно-методической конференции, посвященной 80-летию профессора Ларина С. В. (Красноярск, 2019 г.).

Результаты работы обсуждались на научно-исследовательских семинарах:

– Сибирском государственном университете науки и технологий имени академика М. Ф. Решетнёва;

Публикации по теме диссертационной работы. По результатам диссертационного исследования опубликовано 16 работ, из которых:

– 4 статьи в журналах, рекомендованных ВАК;

– 2 свидетельства на программное обеспечение;

– 2 статьи в изданиях, индексируемых Web of Science и Scopus.

Из статей, написанных в соавторстве, в диссертацию включены результаты, принадлежащие лично автору.

Структура и объём работы. Диссертационная работа изложена на 101 странице и состоит из введения, двух глав, заключения, списка литературы из 104 источников и приложения.

Содержание диссертации

Глава 1 **«Алгоритмы на графах Кэли групп подстановок»**, посвящена разработке новых алгоритмов на группах подстановок, а также алгоритмов на графах Кэли, порождёнными группами указанного вида.

В разделе 1.1 «Алгоритмы на группах подстановок» представлены основные определения, а также новые вспомогательные алгоритмы.

Пусть $G := \langle X \rangle$ – конечная группа, порожденная упорядоченным множеством $X := \{x_1 \prec x_2 \prec \ldots \prec x_m\}$, которое также называют алфавитом. Множество всех слов (строк) над алфавитом X будем обозначать X^*. Пусть $w := x_1 x_2 \ldots x_l$ – слово над X и $|w| := l$ – его длина. На множестве X^* также определим отношение порядка. Пусть v и w – два произвольных слова в алфавите X. Тогда $v \prec w$, если $|v| < |w|$, а в случае равенства длин слов, меньшее слово будет определяться согласно введенному лексикографическому порядку на порождающих. Если необходимо подчеркнуть, что строка $v \in X^*$ соответствует элементу $g \in G$, то мы будем писать v_g. Строку v будем называть минимальным словом элемента g, если для всех других $w \in X^*$, таких что $v_g = w_g$, будет выполняться $v \prec w$. Очевидно, что каждому $g \in G$ соответствует уникальное минимальное слово. Длиной элемента группы g будем называть длину его минимального слова v, т. е. $|g| := \min\{|v_g| : v_g \in X^*\}$.

Графом Кэли $\Gamma := Cay(G, X)$ группы G относительно X называют ориентированный невзвешенный помеченный граф с множеством вершин $V(\Gamma) := \{g \mid g \in G\}$ и множеством ребер $E(\Gamma) := \{(g, gx) \mid x \in X, g \in G\}$. Пусть $(g, gx) \in E(\Gamma)$, тогда генератор x называют меткой данного ребра. Если $X = X \cup X^{-1}$, то граф Γ будет неориентированным. Будем считать, что единичный элемент $e \notin X$, т. е. в Γ нет петель. Как известно, кратчайшее расстояние между двумя произвольными вершинами графа g и h, которое мы обозначим $d(g, h)$, равно длине минимального слова элемента $g^{-1}h$, т. е. $d(g, h) := |g^{-1}h|$.

Пусть G – конечная группа подстановок, заданная на множестве точек $\Omega = \{1, 2, \ldots, n\}$. Обозначим $\alpha^g := g[\alpha]$ образ элемента $\alpha \in \Omega$ под действием $g \in G$. Орбитой точки $\alpha \in \Omega$ называется множество $\alpha^G := \{\alpha^g \mid g \in G\}$. Стабилизатором точки $\alpha \in \Omega$ будем называть множество $G_\alpha := \{g \in G \mid \alpha^g = \alpha\}$. Для заданных $\beta_1, \beta_2, \ldots, \beta_i \in \Omega$ индуктивно определим

$$G_{\beta_1, \beta_2, \ldots, \beta_i} := (G_{\beta_1, \beta_2, \ldots, \beta_{i-1}})_{\beta_i} = \{g \in G \mid \beta_j^g = \beta_j,\ j = 0, 1, \ldots, i\}.$$

Последовательность различных элементов $B := (\beta_1, \beta_2, \ldots, \beta_m)$ будем называть базой группы G, если $G_{\beta_1, \beta_2, \ldots, \beta_m} = e$. Таким, образом только единичный элемент группы оставляет неподвижными все точки базы.

Пусть $G^{(i)} := G_{\beta_1, \beta_2, \ldots, \beta_{i-1}}$. Далее определим цепь стабилизаторов

$$G := G^{(1)} \geq G^{(2)} \geq \cdots \geq G^{(m)} \geq G^{(m+1)} = e.$$

Если $G^{(i+1)}$ является собственной подгруппой $G^{(i)}$ для $i \in [1, m]$, то базу B называют несократимой.

Смежные классы группы $G^{(i)}$ по подгруппе $G^{(i+1)}$ имеют взаимно однозначное соответствие с элементами орбиты $\Delta^{(i)} := \beta_i^{G^{(i)}}$. Действительно, если a, $b \in G^{(i)}$ и $G^{(i+1)}a = G^{(i+1)}b$, то для некоторого $h \in G^{(i+1)}$ будет выполняться $a = hb$. Поэтому, $\beta_i^a = \beta_i^{hb} = \beta_i^b$.

Указанный выше факт дает возможность вычислить семейство представителей смежных классов (трансверсаль) $U^{(i)}$ группы $G^{(i)} \mod G^{(i+1)}$. Для каждого $\gamma \in \Delta^{(i)}$ определим $u_i(\gamma) \in G^{(i)}$, который отображает β_i в γ, т. е. $\beta_i^{u_i(\gamma)} = \gamma$. В частном случае, если $\beta_i \to \beta_i$, то $u_i(\beta_i) := e$.

Согласно введенным обозначениям мы получим упорядоченную последовательность $U^{(i)} := (u_i(\gamma) \mid \gamma \in \Delta^{(i)})$. Очевидно, что $|U^{(i)}| = |\Delta^{(i)}|$.

Объединив все $U^{(i)}$, мы получим полное семейство представителей смежных классов группы:

$$U = \bigcup_{i=1}^{m} U^{(i)}.$$

По теореме Лагранжа $|G| = |G^{(1)} : G^{(2)}| \cdot |G^{(2)}| = |U^{(1)}| \cdot |G^{(2)}|$. Аналогично, $|G^{(2)}| = |G^{(2)} : G^{(3)}| \cdot |G^{(3)}| = |U^{(2)}| \cdot |G^{(3)}|$. Продолжив данный процесс, мы получим

$$|G| = |U^{(1)}| \cdot |U^{(2)}| \cdots |U^{(m)}|.$$

Пусть $B = (\beta_1, \beta_2, \ldots, \beta_m)$ – база G, тогда для любого элемента группы мы можем определить его базовый образ $B^g := (\beta_1^g, \beta_2^g, \ldots, \beta_m^g)$. Нетрудно показать, что $\forall g \in G$ базовый образ B^g имеет уникальное представление.

Далее в разделе представлен алгоритм A–1, который получив на входе слово, состоящее из произведения элементов $g_1 g_2 \ldots g_l$ группы G, ее базу B, полное семейство представителей смежных классов U и его инверсию $\hat{U}$, а также вспомогательный массив A, возвращает каноническое представление $u_m u_{m-1} \ldots u_1$ данного произведения.

Пусть $T_i = O(f)$ и $M_i = O(f)$ – верхние асимптотические оценки вычислительной и пространственной сложности i–го алгоритма соответственно. Доказана следующая

Лемма 1.2. *Алгоритм A–1 корректен и* $T_1 = O(l \cdot m + m^2)$.

Разложение элементов в каноническом виде дает возможность эффективно нумеровать все $g \in G$, используя метод перечисления элементов кортежа в смешанной системе счисления. Пусть $c := (c_m, c_{m-1}, \ldots, c_1)$ – основание смешанной системы счисления, в которой $c_1 := 1$ и $c_i = c_{i-1} \cdot |U^{(i-1)}|$ для $i \geq 2$.

Пусть $g := u_m u_{m-1} \ldots u_1$. Определим биективное отображение

$$\mathcal{N}\colon u_m u_{m-1} \ldots u_1 \longrightarrow (a_m, a_{m-1}, \ldots, a_1),$$

где $a_i \in [0, |U^{(i)}| - 1]$ – номер элемента u_i в $U^{(i)}$.

Заметим, что вектор $(a_m, \dots, a_1)$ представляет собой число $\mathcal{N}(g) \in [0, |G|-1]$ в системе счисления со смешанным основанием $(c_m, \dots, c_1)$. Пусть $k := \mathcal{N}(g)$ – номер элемента g. В следующей лемме даны оценки сложности указанных операций.

Лемма 1.3. *Вычислительная сложность $\mathcal{N}(g)$ и $\mathcal{N}^{-1}(k)$ не превышает $O(m)$.*

Отметим, что для нумерации элементов группы нам необходимо знать ее базу B и полное семейство представителей смежных классов U. Для их вычисления мы будем использовать известный алгоритм Шрайера-Симса, предложенный Ч. Симсом в 1970 году. В настоящее время существует множество его модификаций. Наиболее эффективные версии алгоритма имеют низкую вычислительную сложность и реализованы в таких системах компьютерной алгебры, как GAP, Magma и Mathematica, а также в библиотеке SymPy для языка Python.

Рассмотрим краткое содержание раздела 1.2 «Алгоритмы маршрутизации на графах Кэли».

В начале раздела дано определение таблицы маршрутизации $P_{2\times|G|}$, которую также называют родительским деревом. Затем представлен новый алгоритм А–2, который, получив на входе порождающее множество группы X, ее базу B, полное семейство представителей смежных классов U и его инверсию $\hat{U}$, а также вспомогательный массив A, возвращает указанную таблицу маршрутизации P. Доказана следующая

Теорема 1.1. *Алгоритм А–2 корректен и $T_2 = O(m^2 \cdot |X| \cdot |G|)$.*

Продолжает изложение данного раздела алгоритм А–3, который при помощи таблицы маршрутизации P вычисляет кратчайший путь $w := x_1x_2\dots x_s$ между вершинами $a, b \in Cay(G, X)$, где $x_i \in X$.

Пусть D – диаметр графа. В сформулированной ниже теореме обоснована корректность алгоритма А–3, а также получена оценка его вычислительной сложности.

Теорема 1.2. *Алгоритм А–3 корректен, $T_3 = O(m^2 + D)$ и $M_3 = O(m \cdot n + |G|)$.*

В разделе 1.3 «Проблема минимального слова» представлен алгоритм А–4, который вычисляет минимальное слово w из произвольной строки $v := x_1x_2\dots x_r$ в алфавите порождающих X. Данный алгоритм является модифицированной версией алгоритма А–3. Доказана

Теорема 1.3. *$T_4 = O(|v|)$ и $M_4 = O(m \cdot n + |G|)$.*

Заметим, что во многих прикладных задачах при исследовании графа Кэли $\Gamma := Cay(G, X)$ порядок порождающей группы G значительно превышает ее степень, т. е. $m \leq n \ll |G|$. В этом случае $M_3 = O(|G|)$ и $M_4 = O(|G|)$.

В разделе 1.4 «Сравнительный анализ алгоритмов» показано, что алгоритм А–4 имеет значительное преимущество перед широко известном алгоритмом А–0, основанном на использовании конечных автоматов. Напомним, что $T_0 = O(|v|^2)$ и $M_0 = O(|G|)$.

Представленный в данной главе алгоритм А–3 послужит отправной точкой для создания новых ресурсно-эффективных алгоритмов маршрутизации. Здесь можно выделить два направления. Во-первых, создание алгоритмов, учитывающих топологию сети. В этом случае алгоритмы будут проектироваться для конкретных классов графов Кэли. Во-вторых, разработка гибридных алгоритмов, включающих в себя как статическую, так и динамическую таблицы маршрутизации. Это позволит рассчитывать оптимальные маршруты в зависимости от текущего состояния сети.

Раздел 1.5 «Исследования графов Кэли некоторых групп» содержит результаты вычислительных экспериментов по определению ранее неизвестных характеристик некоторых больших графов Кэли.

Пусть $G = \langle X \rangle$. *Шаром* K_s радиуса s группы G будем называть множество всех её элементов, которые могут быть представлены в алфавите X в виде несократимых групповых слов длины не больше s. Все элементы одинаковой длины i образуют *сферу* P_i радиуса i. Единица группы e является пустым словом, длина которого равна нулю. Согласно данным определениям, $K_s = \bigcup\limits_{i=0}^{s} P_i$.

Для каждого целого неотрицательного i можно определить (сферическую) функцию роста группы $F(G)$, которую мы будем записывать в виде вектора: $F(G) = (F_0, F_1, \ldots, F_i, \ldots)$, где $F_i = |P_i|$. Пусть $F_{s_0} > 0$, но $F_{s_0+1} = 0$, тогда s_0 является диаметром графа Кэли группы G в алфавите порождающих X, который будем обозначать $D_X(G)$. Средний диаметр $\overline{D}_X(G)$ равен $\frac{1}{|G|} \sum\limits_{s=0}^{s_0} s \cdot F_s$.

Рассмотрим семейство графов MBS(n), каждого представителя которого называют «Modified bubble-sort graph» или в переводе на русский «Модифицированный граф пузырьковой сортировки». Указанные графы, задаются симметрическими группами S_n, порождаемыми транспозициями следующего вида:

$$S_n = \langle X_n \rangle, \text{ где } X_n = \{(i\; i+1) \mid i = 1, 2, \ldots, n-1\} \cup \{(1\; n)\}.$$

Используя перечисленные выше алгоритмы, при помощи МВС, удалось исследовать графы Кэли MBS(14) и MBS(15). Доказана

Теорема 1.4. *Пусть* $S_n = \langle X_n \rangle$. *В этом случае, при* $n = 14, 15$, *верны следующие утверждения:*

(i) $D_{X_n}(S_n)$ *и* $\overline{D}_{X_n}(S_n)$ *такие как в таблице* 1.1 *диссертации,*

(ii) Функции роста S_n *заданы таблицами* 1.2–1.3 *диссертации.*

Полученные результаты подтверждают известную гипотезу, в которой утверждается, что $D = \left\lfloor \frac{n^2}{4} \right\rfloor$ и $\overline{D} = \frac{n^2-n+1}{6}$.

Пусть $B_k = \langle a_1, a_2 \rangle$ – некоторая конечная двупорожденная группа периода 7, где a_1 и a_2 – порождающие элементы группы и $|B_k| = 7^k$. Для каждой из B_k при помощи системы компьютерной алгебры GAP легко получить pc-представление элементов группы (power commutator presentation). В этом случае:

$$\forall g \in g \Longrightarrow g = a_1^{x_1} \dots a_k^{x_k},\ x_i \in \mathbb{Z}_7.$$

Пусть $a_1^{x_1} \dots a_n^{x_n}$ и $a_1^{y_1} \dots a_n^{y_n}$ – два произвольных элемента в группе B_k, записанные в коммутаторном виде. Тогда их произведение равно

$$a_1^{x_1} \dots a_n^{x_n} \cdot a_1^{y_1} \dots a_n^{y_n} = a_1^{z_1} \dots a_n^{z_n}.$$

Обозначим $X := \{a_1, a_2\}$ – минимальное порождающее множество B_k и $Y := X \cup X^{-1} = \{a_1, a_2, a_1^{-1}, a_2^{-1}\}$ – симметричное порождающее множество.

Для исследования графов Кэли групп B_k относительно порождающих множеств X и Y в алгоритм A–4 был добавлен указанный выше быстрый способ нумерации элементов. В итоге, используя вычисления на суперкомпьютере, удалось получить функции роста групп B_k при $k \leq 14$, а также характеристики соответствующих графов Кэли. Результаты вычислительных экспериментов представлены в виде теорем 1.5 и 1.6:

Теорема 1.5. *Пусть $B_k = \langle X \rangle$ при $k \leq 14$. В этом случае значения диаметров $D_X(B_k)$ и средних диаметров $\overline{D}_X(B_k)$ такие как в таблице* 1.4 *диссертации, а соответствующие функции роста заданы в таблицах* 1.6–1.18.

Теорема 1.6. *Пусть $B_k = \langle Y \rangle$ при $k \leq 14$. В этом случае значения диаметров $D_Y(B_k)$ и средних диаметров $\overline{D}_Y(B_k)$ такие как в таблице* 1.5 *диссертации, а соответствующие функции роста заданы в таблицах* 1.19–1.31.

Перейдём к изложению основного содержания Главы 2 **«Алгоритм расширенного синтаксического анализа языков программирования методом иерархии маркированных скобок».**

В разделе 2.1 «Постановка расширенной проблемы синтаксического анализа с учётом порядка применения продукций» уточняется формулировка проблемы синтаксического анализа, как результат, формулируется расширенная проблема синтаксического анализа мономов кс–языков: *разработать беступиковый алгоритм, который позволяет установить, может ли быть выведен данный моном при помощи системы продукций заданного кс–языка, а также найти сразу все выводы этого монома, указав, какие продукции, сколько раз и в каком порядке применяются для вывода этого монома.*

В разделах 2.2 «Алгоритм решения расширенной проблемы синтаксического анализа с использованием маркированных скобок» и 2.3 «Оценка сложности алгоритма синтаксического анализа на основе иерархии маркированных скобок» рассматривается два способа маркировки скобок, которые можно использовать в алгоритме синтаксического анализа, основанного на иерархии маркированных таким образом скобок. Рассмотрим подробнее один из способов маркировки.

Рассматривается грамматика кс–языка, заданная в виде совокупности правил подстановки (продукций):

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

где $q_{jk}(z,x)$ – мономом от некоммутативных символов алфавита с числовым коэффициентом, который равен 1.

Кс–язык, порождённый данной грамматикой, является решением системы уравнений Хомского – Шютценберже:

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

и представляется в виде ФСР, который выражает первую компоненту решения $(z_1(x), \dots, z_n(x))$ этой системы уравнений:

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

где w_i – мономы от терминальных символов $x_1, \dots, x_m$.

Для решения расширенной проблемы синтаксического анализа рассматривается соответствующая расширенная грамматика

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

где метка t_{jk} – символ из расширенного алфавита, причём она «привязана» к стоящей справа скобке, помечая правило вывода $z_j \to t_{jk} q_{jk}(z,x)$. Такая расширенная грамматика позволяет определить порядок применения продукций.

Итак, можно маркировать левую (открывающуюся) скобку мономиальной меткой, привязанной к ней слева, считая, что сразу за мономиальной меткой продукции t_{jk} должна следовать привязанная к ней открывающаяся скобка. Показано (лемма 2.2), что для открывающейся скобки всегда можно найти соответствующую закрывающуюся скобку, а значит, по метке, маркирующей стоящую справа от неё скобку, можно устанавливать иерархию (вложение) скобок.

Иерархию скобок позволяет определить соответствующая расширенная система уравнений Хомского – Шютценберже, которая имеет вид:

$$z_j = Q^*_j(z,x,t) \stackrel{\text{def}}{=} t_{j1}[\ q_{j1}(z,x)\ ] + \ \cdots \ + t_{jp_j}[\ q_{jp_j}(z,x)\ ], \ j = 1, \ \dots \ , n.$$

Решение этой системы можно получать методом последовательных приближений. Его итерации, генерирующие начальные члены ФСР решения системы Хомского – Шютценберже, осуществляются по формулам

$$z_j^{(k+1)}(x,t) = Q_j^*(z^{(k)}(x,t),x,t);\ k=0,1,\ldots;\ z^{(0)}(x,t)=0,\ j=1,\ldots,n.$$

Далее начальные члены этого ряда анализируются в соответствии с алгоритмом, описанным в конце раздела 2.2. При этом, «внутренние» скобки соответствуют продукциям, которые использованы позже.

Таким образом, алгоритм синтаксического анализа методом иерархии маркированных скобок можно использовать применительно к расширенной системе уравнений Хомского – Шютценберже.

Для применения метода последовательных приближений необходимо иметь оценку числа итераций, обеспечивающих стабилизацию начальных членов ФСР, которые представляют решение системы уравнений Хомского – Шютценберже.

Такую оценку даёт следующая лемма.

Лемма 2.1. *Пусть $k_2 > k_1$, тогда мономы степени не выше k_1 по переменным x многочленов*

$$z_j^{(k_2)}(x,t),\ j=1,\ldots,n,$$

не зависят от k_2.

Оценка числа итераций, таким образом, состоит в том, что мономы многочленов $z_j^{(k)}(x,t)$, $j=1,\ldots,n$, стабилизированы до степени k включительно и уже не изменятся при дальнейшем росте числа итераций.

Для решения расширенной проблемы синтаксического анализа предложен алгоритм метода иерархии маркированных скобок.

Алгоритм синтаксического анализа, основанный на методе иерархии маркированных скобок, состоит в следующем.

Пусть дан моном (программа) w степени (длины) N.

1. Проводим N итераций метода последовательных приближений для решения соотвествующей расширенной системы уравнений Хомского – Шютценберже.

2. Перебираем все полученные в п. 1 мономы, степень которых равна N, определяя те мономы, которые, с точностью до множителей t_{jk}, совпадают с мономом w, множители t_{jk} при этом пропускаются.

Если таких мономов нет, то моном w вывести невозможно, если такие мономы есть, то они дают решение расширенной проблемы синтаксического анализа в соответствии со следующим п. 3.

3. Считываем все найденные в п. 2 мономы слева направо, устанавливая иерархию маркированных скобок (по признаку сравнения – внутренняя или внешняя скобка) и определяя тем самым порядок применения продукций при выводе монома w, т. е. все возможные способы его вывода.

На вопрос о работоспособности предложенного алгоритма отвечает следующая теорема.

Теорема 2.1. *Алгоритм синтаксического анализа методом иерархии маркированных скобок, основанный на расширенной системе уравнений Хомского – Шютценберже, позволяет за конечное число шагов осуществить беступиковый синтаксический анализ, с учётом порядка применения продукций, любого монома кс–языка, порождённого кс–грамматикой.*

Далее в разделе 2.3 рассмотрен вопрос о сложности разработанного конструктивного беступикового алгоритма синтаксического анализа, позволяющего провести синтаксический анализ монома w, который имеет степень (длину) N.

Этот вопрос решается леммами 2.3 – 2.6, а также следующей теоремой.

Теорема 2.2. *Сложность алгоритма, основанного на методе иерархии маркированных скобок, использующем расширенную систему уравнений Хомского – Шютценберже, равна*

$$O(Ng^{d^N}).$$

Как видно, сложность данного алгоритма достаточно высока. Следует отметить, что она равна сложности метода мономиальных меток, предложенного К.В. Сафоновым для решения более ограниченной задачи синтаксического анализа.

С другой стороны, алгоритм, основанный на иерархии маркированных скобок, является конструктивным, который достаточно прост.

Кроме того, данный алгоритм является беступиковым (безвозвратным, безостановочным) в отличие от известных ранее.

Наконец, отметим, что предложенный алгоритм особенно эффективен в ситуации, когда степень (длина) исследуемого монома (программы) не слишком большая, а других возможностей провести синтаксический анализ, например, в виде парсеров, нет. Такая ситуация, как уже отмечалось, может возникать при разработке языков программирования.

В **Заключении** работы представлены основные результаты и выводы диссертационного исследования.

В **Списке литературы** приведены используемые в диссертации литературные источники.

В **Приложении** содержатся копии свидетельств о государственной регистрации программ для ЭВМ.

Основные результаты и выводы

В результате проделанной работы получены следующие результаты:

1. Разработан новый эффективный алгоритм, позволяющий вычислять кратчайшие маршруты между вершинами графа Кэли $Cay(G, X)$, заданного произвольной конечной группой подстановок $G = \langle X \rangle$. Обоснована корректность представленного алгоритма, даны оценки его сложности. В этом случае для произвольного слова $w \in X^*$ проблема поиска минимального слова разрешается за время $T = O(|w|)$.

2. При помощи МВС вычислены ранее неизвестные характеристики графов Кэли модифицированной пузырьковой сортировки MBS(n) для случаев $n = 14$ и 15. Получены новые результаты о некоторых графах Кэли, заданных конечными двупорождёнными группами периода 7.

3. Впервые разработан беступиковый алгоритм синтаксического анализа, использующий иерархию маркированных скобок и позволяющий решать расширенную проблему синтаксического анализа, включая вычисление всех выводов монома (программы), который находит применение при разработке языков программирования для МВС.

4. Получена оценка сложности алгоритма синтаксического анализа, использующего иерархию маркированных скобок, особенно эффективного при разработке новых языков программирования.

Таким образом, решены все поставленные в диссертации задачи, цель работы достигнута.

Список публикаций автора

Свидетельства о регистрации программ для ЭВМ

1. Кишкан В.В., Кузнецов А.А., Кузнецова А.С. Произведение элементов в двупорожденных группах периода 7, основанное на полиномах Холла // Свидетельство о регистрации программы для ЭВМ RU 2019611975, 07.02.2019. Заявка № 2018661656 от 16.10.2018.

2. Кишкан В.В., Кузнецов А.А., Кузнецова А.С. Произведение элементов в двупорожденных группах периода 5, основанное на полиномах Холла // Свидетельство о регистрации программы для ЭВМ RU 2015663307, 15.12.2015. Заявка № 2015619899 от 20.10.2015.

Публикации, входящие в международные базы цитирования

3. Kuznetsov A.A., Kuznetsova A.S., Kishkan V.V. The Cayley graphs of a centralizer of the Burnside group $B_0(2,5)$ // IOP Conference Series: Materials Science and Engineering. – 2020. Vol. 822, 012043 (1–5). **(Scopus)** DOI:10.1088/1757-899X/822/1/012043.

4. Kishkan V.V., Safonov K.V., Tsarev R.Yu. Syntactical analysis of context-free languages taking into account order of application of productions // Journal of Physics: Conference Series. – 2019. Vol. 1333, no. 3, 032072 (1–6). **(Web of Science, Scopus)** DOI:10.1088/1742-6596/1333/3/032072.

Публикации в журналах из перечня ВАК

5. Кишкан В.В., Сафонов К.В. Беступиковый алгоритм расширенного синтаксического анализа и его приложение к языкам программирования для квантовых компьютеров // Computational Nanotechnology. 2020 Т. 20. № 2. С. 42–48.

6. Кишкан В. В. , Кузнецов А. А. Об одном алгоритме маршрутизации на графах Кэли, порожденных группами подстановок // Сибирский журнал науки и технологий. 2020 Т. 21. № 2. С. 187–194.

7. Kuznetsov A.A., Kishkan V.V. The Cayley graphs of finite two-generator burnside groups of exponent 7 // Сибирский журнал науки и технологий. 2018 Т. 19. № 2. С. 217–222.

8. Кишкан В. В. , Кузнецов А. А. Исследование графов модифицированной пузырьковой сортировки на основе высокопроизводительных вычислений // Сибирский журнал науки и технологий. 2017 Т. 18. № 3. С. 525–529.

Публикации в материалах конференций

9. Кишкан В.В., Сафонов К.В. Синтаксический анализ мономов контекстно-свободных языков с учетом порядка применения продукций // Материалы XXIII Международной научно–практической конференции «Решетнёвские чтения». Красноярск. – 2019. С. 19–20.

10. Кузнецов А.А., Кишкан В.В. Алгоритм эффективной нумерации элементов произвольной группы подстановок // Материалы XXIII Международной научно–практической конференции «Решетнёвские чтения». Красноярск. – 2019. С. 23–24.

11. Кузнецов А.А., Кишкан В.В. Об одном алгоритме исследования роста в конечной группе подстановок // Материалы VIII Всероссийской с международным участием научно–методической конференции, посвященной 80-летию профессора Ларина С.В. – 2019. С. 43–47.

12. Кишкан В.В., Сафонов К.В. О новом подходе в синтаксическом анализе мономов контекстно-свободных языков // Материалы VIII Всероссийской с международным участием научно–методической конференции, посвященной 80-летию профессора Ларина С.В. – 2019. С. 54–56.

13. Кишкан В.В., Сафонов К.В. Синтаксический анализ программ методом интегральных представлений / Материалы 18 Всероссийской конференции «Сибирская научная школа-семинар с международным участием «Компьютерная безопасность и криптография» – SIBECRYPT'19. // Прикладная дискретная математика. Приложение. – 2019. № 12. С. 194–196.

14. Кишкан В.В., Колбасина И.В., Сафонов К.В. Синтаксический анализ контекстно-свободных языков с учетом порядка применения продукции // Материалы XXII Международной научно–практической конференции «Решетнёвские чтения». Красноярск. – 2018. Т. 2. С. 14–15.

15. Кишкан В.В., Колбасина И.В., Сафонов К.В. О синтаксическом анализе контекстно-свободных языков // Материалы XXI Международной научно–практической конференции «Решетнёвские чтения». Красноярск. – 2017. Т. 2. С. 77–78.

16. Кузнецов А.А., Кишкан В.В. Исследование графов модифицированной пузырьковой сортировки на основе высокопроизводительных вычислений // Материалы XXI Международной научно–практической конференции «Решетнёвские чтения». Красноярск. – 2017. С. 82–83.

Автор выражает свою благодарность научному руководителю профессору Кузнецову Александру Алексеевичу за всестороннюю поддержку, полученную во время работы над диссертацией.

Кроме того автор благодарит заведующего кафедрой прикладной математики СибГУ профессора Сафонова Константина Владимировича за ценные научные консультации, полученные при написании второй главы диссертации.