

На правах рукописи



Шабля Юрий Васильевич

**АЛГОРИТМИЧЕСКОЕ ОБЕСПЕЧЕНИЕ
КОМБИНАТОРНОЙ ГЕНЕРАЦИИ НА ОСНОВЕ
ПРИМЕНЕНИЯ ТЕОРИИ ПРОИЗВОДЯЩИХ
ФУНКЦИЙ**

Специальность 05.13.17 —
«Теоретические основы информатики»

Автореферат
диссертации на соискание учёной степени
кандидата технических наук

Томск — 2019

Работа выполнена в Федеральном государственном бюджетном образовательном учреждении высшего образования «Томский государственный университет систем управления и радиоэлектроники» (ТУСУР).

Научный руководитель: доктор технических наук, профессор
Шелупанов Александр Александрович

Официальные оппоненты: **Рябко Борис Яковлевич**,
доктор технических наук, профессор,
главный научный сотрудник, и.о. заведующего лабораторией информационных систем и защиты информации Федерального государственного бюджетного учреждения науки «Институт вычислительных технологий Сибирского отделения Российской академии наук»

Мальчуков Андрей Николаевич,
кандидат технических наук, доцент,
доцент отделения информационных технологий Федерального государственного автономного образовательного учреждения высшего образования «Национальный исследовательский Томский политехнический университет»

Ведущая организация: Федеральное государственное бюджетное образовательное учреждение высшего образования «Новосибирский государственный технический университет»

Защита состоится 12 декабря 2019 г. в **15** часов **15** минут на заседании диссертационного совета Д 212.268.05 на базе Томского государственного университета систем управления и радиоэлектроники по адресу: 634050, г. Томск, пр. Ленина, 40, ауд. 201.

С диссертацией можно ознакомиться в библиотеке ТУСУРа по адресу: 634045, г. Томск, ул. Красноармейская, 146 и на сайте ТУСУРа по адресу: <https://postgraduate.tusur.ru/urls/mvi6k984>.

Автореферат разослан «___» _____ 2019 года.

Ученый секретарь
диссертационного совета



Костюченко Евгений Юрьевич

ОБЩАЯ ХАРАКТЕРИСТИКА РАБОТЫ

Актуальность темы. Многие информационные объекты, в том числе с большим объемом данных, носят иерархическую или рекурсивную природу описания. В таком случае древовидная структура является естественной формой представления информационного объекта, что позволяет описать данный информационный объект формальным комбинаторным множеством и применить в дальнейшем алгоритмы комбинаторной генерации.

Комбинаторная генерация — раздел на стыке информатики и комбинаторики, развивающий методы и алгоритмы ранжирования и генерации комбинаторных множеств, таких как классы перестановок, разбиений, графов, деревьев, таблиц и многие другие. Под ранжированием и генерацией комбинаторных множеств понимается нумерация объектов комбинаторных множеств, то есть их кодирование в виде чисел для удобства хранения информации о них, а также последующее восстановление из чисел самих объектов. Комбинаторная генерация тесно связана с процедурой индексации информационных объектов, а разработка эффективных алгоритмов комбинаторной генерации является актуальной задачей. Разработкой методов построения алгоритмов комбинаторной генерации занимались такие ученые, как Э. Рейнгольд, Д. Крехер, Е. Баркуччи, С. Баччелли, А. Дель Лунго, В. Вайновски, Ф. Флажолле, К. Мартинес и К. Мулинеро, Б.Я. Рябко, Ю.С. Медведева, В.В. Кручинин.

Существующие универсальные методы построения алгоритмов комбинаторной генерации требуют представления рассматриваемых комбинаторных объектов в специальном виде за счет применения некоторых эвристик, что является трудной задачей ввиду отсутствия соответствующих методик. Предлагаемое научное исследование направлено на разработку и формализацию новых методов построения алгоритмов комбинаторной генерации. Для этого планируется применить для получения выражений функций мощности комбинаторных множеств математический аппарат теории производящих функций, который ранее не использовался в рамках рассматриваемой научной задачи. Производящие функции являются мощным инструментом решения задач из самых разных областей математических наук, таких как комбинаторика, теория чисел, теория вероятностей и др.

Цели и задачи исследования. Целью диссертационной работы является развитие методов построения алгоритмов комбинаторной генерации за счет применения теории производящих функций.

Для достижения поставленной цели были решены следующие задачи:

1. Провести аналитический обзор современного состояния исследований в области разработки алгоритмов комбинаторной генерации.
2. Исследовать и формализовать метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ.
3. Исследовать метод получения явных выражений коэффициентов производящих функций.

4. Модифицировать метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ за счет применения теории производящих функций.

5. Провести апробацию модифицированного метода построения алгоритмов комбинаторной генерации, разработав и исследовав новые алгоритмы комбинаторной генерации.

6. Создать программное обеспечение для ранжирования и генерации по рангу элементов комбинаторных множеств на основе разработанных алгоритмов комбинаторной генерации.

Объект исследования. Объектом исследования являются методы и алгоритмы комбинаторной генерации.

Предмет исследования. Предметом исследования является построение новых эффективных алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ с использованием теории производящих функций.

Методы исследования. В диссертационной работе применялись методы построения алгоритмов комбинаторной генерации, получения явных выражений коэффициентов производящих функций, анализа алгоритмов, объектно-ориентированного программирования.

Научная новизна полученных результатов:

1. Предложен модифицированный метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ, который отличается применением метода получения явных выражений коэффициентов производящих функций для нахождения выражения функции мощности комбинаторного множества, для которого не известно выражение функции мощности, принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$.

2. Разработаны новые алгоритмы ранжирования и генерации по рангу для множества комбинаторных объектов, отражающих вторичную структуру РНК, отличающиеся от аналогов меньшей вычислительной сложностью, оценка которой равна $O(m^2(n - m))$.

3. Впервые созданы алгоритмы ранжирования и генерации по рангу для множества комбинаторных объектов, определяемых числовым треугольником Эйлера-Каталана.

Теоретическая значимость работы. Теоретическая значимость результатов диссертационной работы заключается в развитии методов построения алгоритмов комбинаторной генерации. Модифицированный метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ за счет применения теории производящих функций позволяет использовать данный метод для получения новых алгоритмов ранжирования и генерации по рангу как для известных, так и для новых комбинаторных множеств.

Практическая значимость работы. Предложенный модифицированный метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ расширяет перечень инструментов проведения

исследований в области разработки алгоритмов комбинаторной генерации. Полученные в рамках апробации модифицированного метода алгоритмы комбинаторной генерации подтверждают эффективность и универсальность его применения для широкого разнообразия комбинаторных множеств.

Разработанное программное обеспечение для системы компьютерной алгебры «Matha» позволяет в автоматизированном режиме решать задачи комбинаторной генерации по ранжированию и генерации по рангу элементов комбинаторных множеств. Это обеспечивает ускорение выполнения вычислений в рамках работы с алгоритмами комбинаторной генерации.

Практическая значимость результатов диссертационной работы подтверждается их внедрением в ООО «ПлантаПлюс», ООО «Удостоверяющий центр Сибири» и учебный процесс ФГБОУ ВО «ТУСУР», а также использованием в ходе выполнения научно-исследовательских работ. Например, в процессе создания программного продукта для работы с алгоритмами получения простых чисел в ООО «Удостоверяющий центр Сибири», комбинирование полученных результатов проверки генерируемых чисел с помощью теста простоты числа с результатами проверки теста, основанного на полученных с помощью разработанного программного обеспечения критериев простоты числа, позволило сократить от 5 до 7% количество необнаруженных псевдопростых чисел, что является значимым результатом при обработке больших выборок чисел.

Положения, выносимые на защиту:

1. Модифицированный метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ позволяет строить алгоритмы ранжирования и генерации по рангу для таких комбинаторных множеств, для которых не известно выражение функции мощности, принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$, но известно выражение производящей функции для последовательности значений функции мощности.

Соответствует п. 3 паспорта специальности 05.13.17: Исследование методов и разработка средств кодирования информации в виде данных. Принципы создания языков описания данных, языков манипулирования данными, языков запросов. Разработка и исследование моделей данных и новых принципов их проектирования.

2. Оценка вычислительной сложности разработанных алгоритмов ранжирования и генерации по рангу для множества комбинаторных объектов, отражающих вторичную структуру РНК длины n с m пар нуклеотидов, соединенных водородной связью, равна $O(m^2(n-m))$, при этом не требуется вспомогательных предварительных вычислений.

Соответствует п. 3 паспорта специальности 05.13.17: Исследование методов и разработка средств кодирования информации в виде данных. Принципы создания языков описания данных, языков манипулирования данными, языков запросов. Разработка и исследование моделей данных и новых принципов их проектирования.

3. Оценка вычислительной сложности разработанных алгоритмов ранжирования и генерации по рангу для множества комбинаторных объектов, определяемых числовым треугольником Эйлера-Каталана EC_n^m , равна $O(nm(n - m))$, при этом не требуется вспомогательных предварительных вычислений.

Соответствует п. 3 паспорта специальности 05.13.17: Исследование методов и разработка средств кодирования информации в виде данных. Принципы создания языков описания данных, языков манипулирования данными, языков запросов. Разработка и исследование моделей данных и новых принципов их проектирования.

4. Алгоритмы комбинаторной генерации для создания программного обеспечения, автоматизирующего процесс ранжирования и генерации по рангу для комбинаторных множеств.

Соответствует п. 14 паспорта специальности 05.13.17: Разработка теоретических основ создания программных систем для новых информационных технологий.

Достоверность результатов. Достоверность результатов диссертационной работы обеспечивается сравнением разработанных алгоритмов с аналогичными алгоритмами других авторов, проверкой теоретических положений вычислительными экспериментами, положительным эффектом от внедрения полученных результатов.

Внедрение результатов работы. Результаты диссертационной работы использованы в ходе выполнения следующих научно-исследовательских работ:

– базовая часть государственного задания Министерства науки и высшего образования РФ (проект № 2.8172.2017/8.9 «Метод и модели определения уровня защищенности информационных систем» на 2017–2019 гг.);

– грант «Российского научного фонда» (проект № 18-71-00059 «Разработка алгоритмов и программного обеспечения индексирования больших объемов данных на основе новых методов комбинаторной генерации» на 2018–2020 гг.);

– грант «Российского фонда фундаментальных исследований» (проект № 18-31-00201 «Методы комбинаторной генерации на основе деревьев И/ИЛИ с применением теории производящих функций» на 2018–2019 гг.).

Результаты диссертационной работы использованы при систематизации архива проведения экспериментальных исследований в ООО «ПлантаПлюс», а также в процессе создания программного продукта для работы с алгоритмами получения простых чисел в ООО «Удостоверяющий центр Сибири».

Результаты диссертационной работы внедрены в учебный процесс ФГБОУ ВО «ТУСУР» на факультете безопасности и используются при изучении дисциплин «Дискретная математика» и «Теория игр и исследование операций».

Личный вклад автора. Постановка цели и задач научного исследования осуществлялась совместно с научным руководителем. Автором разработан модифицированный метод построения алгоритмов комбинаторной генерации на основе применения теории производящих функций; разработаны алгоритмы комбинаторной генерации для множества комбинаторных объектов, отражающих выражения обобщенного языка Дика, для множества комбинаторных объектов, отражающих вторичную структуру РНК, и для множества комбинаторных объектов, определяемых треугольником Эйлера-Каталана; создано программное обеспечение для ранжирования и генерации по рангу элементов комбинаторных множеств.

Апробация работы. Основные результаты диссертационной работы докладывались на следующих конференциях:

1. Международная научно-практическая конференция «Электронные средства и системы управления» (2015–2018 гг., г. Томск, ТУСУР).

2. Международная научно-техническая конференция студентов, аспирантов и молодых ученых «Научная сессия ТУСУР» (2016–2019 гг., г. Томск, ТУСУР).

3. The 3rd Algorithmic and Enumerative Combinatorics Summer School (1–5 августа 2016 г., Австрия, г. Хагенберг, Университет им. И. Кеплера).

4. The Mediterranean International Conference of Pure & Applied Mathematics and Related Areas (26–29 октября 2018 г., Турция, г. Анталья, Университет Акдениз).

5. XII Всероссийская научная конференция «Наука. Технологии. Инновации» (3–7 декабря 2018 г., г. Новосибирск, НГТУ).

6. V Международная научно-практическая конференция молодых ученых «Прикладная математика и информатика: современные исследования в области естественных и технических наук» (22–24 апреля 2019 г., г. Тольятти, ТГУ).

7. The 32th International Conference of the Jangjeon Mathematical Society (17–19 июля 2019 г., г. Владивосток, ДВФУ).

8. The 2nd Mediterranean International Conference of Pure & Applied Mathematics and Related Areas (28–31 августа 2019 г., Франция, г. Париж, Университет Эври).

9. Томский IEEE семинар «Интеллектуальные системы моделирования, проектирования и управления» (2015–2019 гг., г. Томск, ТУСУР).

Публикации по теме диссертации. Основные результаты диссертационного исследования опубликованы в 21 работе, в том числе 4 публикации в рецензируемых журналах из перечня ВАК, 6 публикаций в научных изданиях, индексируемых Web of Science и Scopus, 9 публикаций в тезисах и материалах научных конференций, получены 3 свидетельства о государственной регистрации программы для ЭВМ.

Объем и структура работы. Диссертация состоит из введения, четырех глав основной части, заключения, списка литературы и двух приложений. Полный объем диссертации составляет 123 страницы, включая 29 рисунков и 8 таблиц. Список литературы содержит 102 наименования.

ОСНОВНОЕ СОДЕРЖАНИЕ РАБОТЫ

Во **введении** обоснована актуальность исследований, проводимых в рамках диссертационной работы, сформулирована цель, поставлены задачи работы, изложены научная новизна, теоретическая и практическая значимость представляемой работы, а также положения, выносимые на защиту. Указаны сведения об апробации и внедрении результатов.

В **первой главе** представлен анализ современного состояния исследований в области разработки алгоритмов комбинаторной генерации. Среди основных подходов к разработке алгоритмов комбинаторной генерации выделены: метод поиска с возвратом (backtracking), ЕСО-метод, метод Ф. Флажоле, метод Б.Я. Рябко, метод В.В. Кручинина. Для каждого из приведенных подходов рассмотрены подробно принципы их работы с указанием имеющихся недостатков. Также рассмотрен подробно процесс разработки алгоритмов комбинаторной генерации с помощью каждого метода для таких классических комбинаторных множеств, как перестановки и сочетания.

Проведенный анализ универсальных методов построения алгоритмов комбинаторной генерации показал, что:

- часть методов (метод поиска с возвратом и ЕСО-метод) направлены только на разработку алгоритмов последовательной генерации комбинаторных объектов;
- существуют ограничения на возможность применения части методов (ЕСО-метод и метод Ф. Флажоле) для комбинаторных множеств, описываемых более чем одним параметром;
- большинство методов требуют представления комбинаторного объекта в специальном виде (слово, последовательность, спецификация, дерево И/ИЛИ), что не всегда является тривиальной задачей и требует дополнительного исследования;
- существуют требования к наличию дополнительной информации, описывающей комбинаторное множество.

В таблице 1 представлены результаты сравнения выявленных основных характеристик методов построения алгоритмов комбинаторной генерации:

- характеристика «Listing»: имеется возможность построения алгоритмов последовательной генерации комбинаторных объектов;
- характеристика «Ranking/Unranking»: имеется возможность построения алгоритмов ранжирования и генерации комбинаторных объектов в соответствии с заданным рангом;
- характеристика «Более одного параметра»: имеется возможность применения метода для комбинаторных множеств, описываемых более чем одним параметром;
- характеристика «Биекция»: имеется требование представления комбинаторного объекта в специальном виде;

– характеристика «Дополнительные требования»: имеются требования к наличию дополнительной информации, описывающей комбинаторное множество.

Таблица 1 — Сравнение характеристик общих методов построения алгоритмов комбинаторной генерации

Метод	Характеристика				
	Listing	Ranking/ Unranking	Более одного параметра	Биекция	Дополнительные требования
Метод поиска с возвратом	+	–	+	–	
ЕСО-метод	+	–	–	+	ЕСО-правило
Метод Ф. Флажоле	+	+	–	+	Функция мощности, производящая функция, спецификация
Метод Б.Я. Рябко	+	+	+	+	Функция мощности, вспомогательные функции
Метод В.В. Кручинина	+	+	+	+	Функция мощности из алгебры $\{\mathbb{N}, +, \times, R\}$

Для дальнейшего исследования выбран метод комбинаторной генерации В.В. Кручинина, основанный на применении деревьев И/ИЛИ, так как данный метод:

- позволяет разрабатывать все типы алгоритмов комбинаторной генерации (listing, ranking и unranking);
- не имеет ограничений на количество параметров, которыми описываются комбинаторные множества, что позволяет рассматривать более сложные дискретные структуры;
- требует в качестве дополнительной информации, описывающей комбинаторное множество, только выражение функции мощности, на основе которой строится структура дерева И/ИЛИ.

Однако при этом необходимо знать выражение функции мощности комбинаторного множества, принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$. Это является одним из недостатков применения данного метода для разработки алгоритмов комбинаторной генерации, который в рамках диссертационного исследования предлагается устранить за счет применения математического аппарата производящих функций.

Во **второй главе** представлены полученные результаты разработки модификации метода построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ за счет применения теории производящих функций.

Метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ базируется на представлении комбинаторных множеств в виде структуры дерева И/ИЛИ, число вариантов которого должно совпадать со значением функции мощности комбинаторного множества. С помощью такого дерева И/ИЛИ можно строить алгоритмы последовательной генерации комбинаторных объектов, их ранжирования и генерации в соответствии с их рангами.

Деревом И/ИЛИ называется дерево, которое содержит узлы двух типов: И-узел и ИЛИ-узел. Вариантом дерева И/ИЛИ называется дерево, которое получается из данного путем отсечения у всех ИЛИ-узлов всех дуг кроме одной. Если для комбинаторного множества A известно выражение функции мощности $f = |A|$, которое принадлежит алгебре $\{\mathbb{N}, +, \times, R\}$, где R — оператор примитивной рекурсии, тогда существует возможность построения дерева И/ИЛИ, число вариантов которого будет совпадать со значением функции мощности комбинаторного множества. В обратном случае структура дерева И/ИЛИ не будет построена и, соответственно, применение метода построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ становится невозможным. Предлагается решить эту проблему за счет применения математического аппарата производящих функций, так как он является основополагающим в современной комбинаторике и для многих комбинаторных множеств уже известны выражения производящих функций или существует возможность их получения.

Производящие функции являются мощным инструментом решения задач из самых разных областей математических наук, таких как комбинаторика, теория чисел, теория вероятностей и др. Обыкновенной производящей функцией (или производящей функцией) произвольной числовой последовательности $(a_n)_{n \geq 0}$ называется формальный степенной ряд вида

$$A(t) = a_0 + a_1 t + a_2 t^2 + \dots = \sum_{n \geq 0} a_n t^n.$$

Композитой обыкновенной производящей функции $G(t)$ называется функция $G^\Delta(n, k)$, которая является функцией коэффициентов k -й степени производящей функции $G(t)$

$$(G(t))^k = \sum_{n \geq k} G^\Delta(n, k) t^n.$$

Математический аппарат степеней производящих функций обеспечивает такие операции над композициями, как сдвиг, сложение, умножение, композиция, а также определение взаимных и обратных композит. Такой набор операций над композициями позволяет получать явные выражения для коэффициентов производящих функций.

Если для некоторого комбинаторного множества A рассмотреть его подмножество $A_n \subset A$, которое содержит только комбинаторные объекты размерности n , тогда функция мощности $f(n) = |A_n|$ такого комбинаторного множества A_n может быть описана некоторой производящей функцией

$$F(t) = \sum_{n \geq 0} f_n t^n = \sum_{n \geq 0} f(n) t^n = \sum_{n \geq 0} |A_n| t^n.$$

Следовательно, для получения выражения функции мощности $f(n)$ комбинаторного множества, для которого известна производящая функция, можно использовать метод получения явных выражений коэффициентов производящих функций. Запишем модифицированный метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ в виде последовательности шагов:

- Шаг 1.** Если известно выражение функции мощности f комбинаторного множества A , принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$, то переход на шаг 4.
- Шаг 2.** Если известно выражение производящей функции F для последовательности значений функции мощности f комбинаторного множества A , то применить метод получения явных выражений коэффициентов производящих функций, иначе дальнейшее применение метода невозможно.
- Шаг 3.** Если получено выражение функции мощности f комбинаторного множества A , принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$, то переход на шаг 4, иначе дальнейшее применение метода невозможно.
- Шаг 4.** На основе выражения функции мощности f комбинаторного множества A построить структуру дерева И/ИЛИ D .
- Шаг 5.** Определить биекцию $A \leftrightarrow W(D)$ между элементами комбинаторного множества A и множества вариантов дерева И/ИЛИ D в виде некоторых алгоритмов $\text{ObjectToVariant}(a, D) : A \rightarrow W(D)$, где $a \in A$, и $\text{VariantToObject}(v, D) : W(D) \rightarrow A$, где $v \in W(D)$.
- Шаг 6.** Определить биекцию $W(D) \leftrightarrow \mathbb{N}_{|W(D)|}$ между элементами множества вариантов дерева И/ИЛИ D и конечного множества натуральных чисел $\mathbb{N}_{|W(D)|} = \{0, 1, \dots, |W(D)| - 1\}$ с помощью алгоритмов $\text{RankVariant}(\text{root}, v, D) : W(D) \rightarrow \mathbb{N}_{|W(D)|}$, где root — корень дерева И/ИЛИ D , $v \in W(D)$, и $\text{UnrankVariant}(r, D) : \mathbb{N}_{|W(D)|} \rightarrow W(D)$, где $r \in \mathbb{N}_{|W(D)|}$.

Совокупность алгоритмов, определенных на последних двух шагах модифицированного метода (Шаг 5 и Шаг 6), формируют биекцию $A \leftrightarrow \mathbb{N}$ и представляют собой алгоритмы комбинаторной генерации $\text{Rank}(a) : A \rightarrow \mathbb{N}$ и $\text{Unrank}(r) : \mathbb{N} \rightarrow A$. Изменение оригинального метода заключается в наличии дополнительных шагов (Шаг 2 и Шаг 3), которые в случае успешного их выполнения позволяют воспользоваться методом построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ.

Третья глава посвящена апробации предложенной модификации метода построения алгоритмов комбинаторной генерации путем разработки алгоритмического обеспечения.

В разделе 3.1 разработаны алгоритмы ранжирования и генерации по рангу для множества выражений обобщенного языка Дика на примере множества последовательностей правильно вложенных n пар скобок m типов. Оценка вычислительной сложности полученных алгоритмов комбинаторной генерации равна $O(n^3)$, что является хуже по сравнению с существующими аналогами. Однако это показывает возможность применения разработанной модификации для получения новых алгоритмов комбинаторной генерации.

В разделе 3.2 разработаны алгоритмы ранжирования и генерации по рангу для множества комбинаторных объектов, отражающих вторичную структуру РНК длины n с m пар нуклеотидов, соединенных водородной связью. Рибонуклеиновая кислота (РНК) — это один из основных и наиболее важных элементов клеточной структуры любого живого организма. Молекула РНК представляет собой цепочку нуклеотидов, состоящих из азотистых оснований четырех видов: А (Аденин), С (Цитозин), G (Гуанин) и U (Урацил). Вид последовательности нуклеотидов в цепочке РНК образует так называемую первичную структуру РНК. С точки зрения математической записи первичная структура РНК является последовательностью элементов множества $\{A, C, G, U\}$, то есть $a = (a_1, a_2, \dots)$, где $a_i \in \{A, C, G, U\}$. Также в цепочке РНК могут образовываться дополнительные водородные связи между двумя основаниями, за счет образования которых в цепочке РНК появляются всевозможные изгибы и петли, что, в свою очередь, усложняет вид структуры молекулы РНК. С точки зрения математической записи вторичная структура РНК может быть представлена в виде структуры графа, в котором помеченные вершины являются нуклеотидами с указанием типа азотистого основания, а ребра показывают наличие связи между нуклеотидами. На рисунке 1 приведен пример представления вторичной структуры цепочки РНК, состоящей из 27 нуклеотидов и 9 водородных связей.

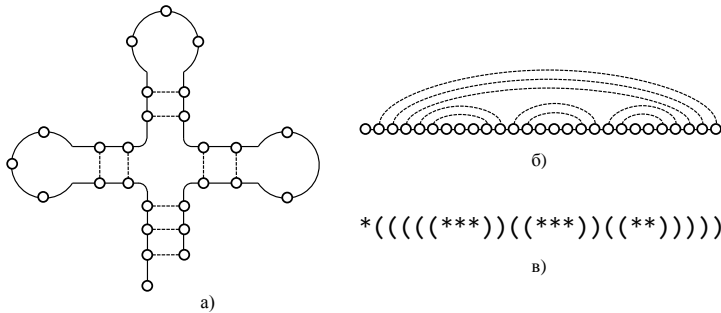


Рис. 1 — Пример представления вторичной структуры цепочки РНК:
 а) в виде структуры «листа клевера»; б) в виде прямой цепочки;
 в) в виде последовательности скобок

На рисунке 2 представлено полученное дерево И/ИЛИ на основе известного выражения функции мощности S_n^m для комбинаторного множества вторичных структур РНК.

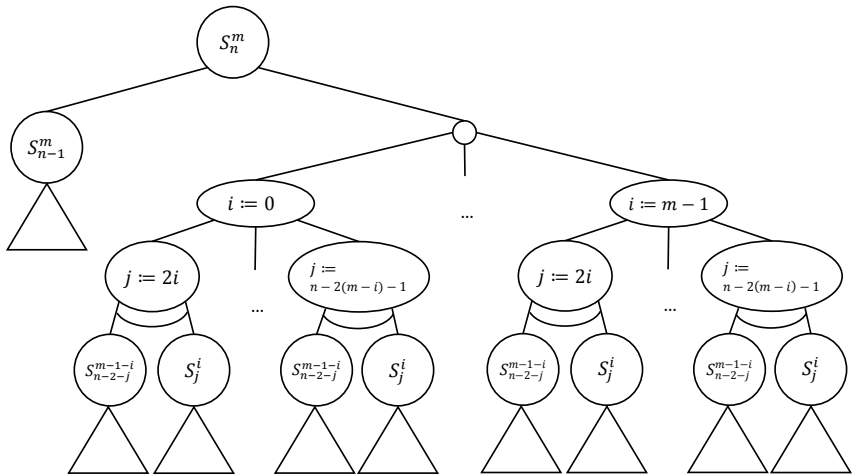


Рис. 2 — Дерево И/ИЛИ для множества вторичных структур РНК

Оценка вычислительной сложности разработанных алгоритмов $\text{RNASS_Rank}(a, n, m)$ для ранжирования и $\text{RNASS_Unrank}(r, n, m)$ для генерации по рангу для множества вторичных структур РНК длины n с m пар нуклеотидов, соединенных водородной связью, равна $O(m^2(n - m))$. По сравнению с существующими аналогами данная оценка является лучшей, так как не требуется никаких предварительных вычислений.

В разделе 3.3 были разработаны алгоритмы ранжирования и генерации по рангу для множества комбинаторных объектов, определяемых треугольником Эйлера-Каталана, на примере множества помеченных путей Дика длины $2n$ с m подъемами на возвратных шагах. Оценка вычислительной сложности разработанных алгоритмов равна $O(m^2(n - m))$. Для данного комбинаторного множества не было известно выражение функции мощности, принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$, но было известно выражение производящей функции $EC(x, y)$ для последовательности значений функции мощности. В результате с помощью метода получения явных выражений коэффициентов производящих функций (Шаг 2 модифицированного метода) было получено выражение функции мощности EC_n^m , принадлежащее алгебре $\{\mathbb{N}, +, \times, R\}$.

Таким образом, по итогам апробации представленной модификации метода построения алгоритмов комбинаторной генерации можно сделать вывод о том, что применение данного метода является эффективным с точки зрения построения новых алгоритмов комбинаторной генерации как для известных комбинаторных множеств, так и для новых.

В **четвертой** главе представлено созданное на основе системы компьютерной алгебры «Mathia» программное обеспечение для ранжирования и генерации по рангу элементов комбинаторных множеств, которая автоматизирует процессы вычислений в рамках разработанных алгоритмов комбинаторной генерации. На рисунке 3 представлена структура разработанного программного обеспечения.

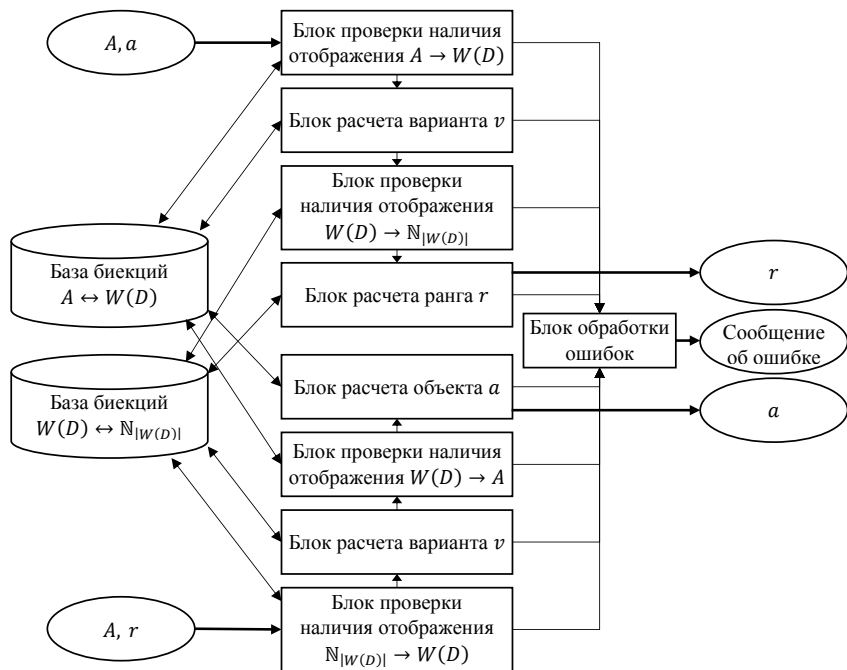


Рис. 3 — Структура программного обеспечения для ранжирования и генерации по рангу элементов комбинаторных множеств

Используя программную реализацию разработанных алгоритмов комбинаторной генерации, был проведен ряд вычислительных экспериментов, направленных на выявление зависимостей времени вычислений от значений параметров исследуемого комбинаторного множества.

Например, для алгоритмов комбинаторной генерации для множества вторичных структур РНК длины n с m пар нуклеотидов, соединенных водородной связью, был проведен вычислительный эксперимент для фиксированного значения $m = 2$, в рамках которого была рассмотрена работа алгоритмов для комбинаторных объектов с максимальным значением ранга $r = S_n^m - 1$. На рисунке 4 представлены полученные результаты средних значений времени вычислений по каждому алгоритму, которые подтверждают наличие линейной зависимости от параметра n .

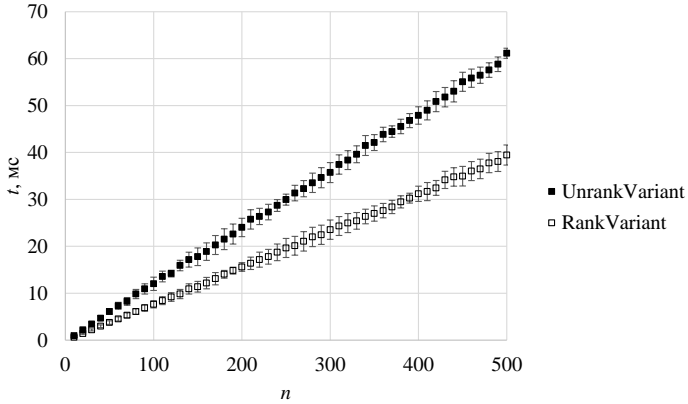


Рис. 4 — Среднее время вычислений в зависимости от параметра n (на основе максимального значения ранга r при $m = 2$)

Для проверки зависимости алгоритмов `UnrankVariant_RNASS`(r, n, m) и `RankVariant_RNASS`(v, n, m) от значений параметра m был проведен вычислительный эксперимент, в рамках которого был зафиксирован второй параметр $n = 100 + 2m$. На рисунке 5 представлены полученные результаты, которые подтверждают наличие квадратичной зависимости от параметра m (при $m < n - 2m$ и постоянном значении разницы $n - 2m$) и линейной зависимости от параметра m (при $m > n - 2m$ и постоянном значении разницы $n - 2m$).

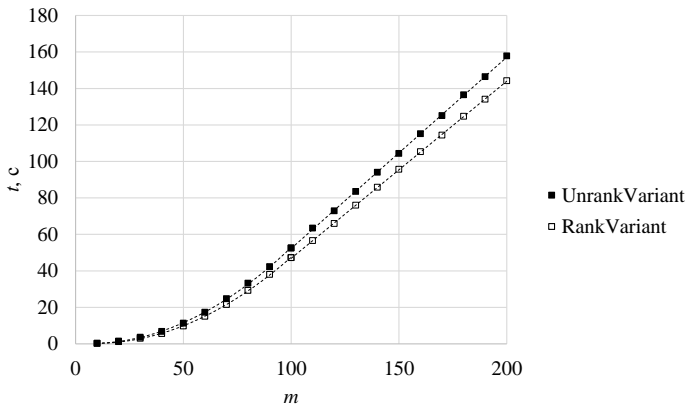


Рис. 5 — Среднее время вычислений в зависимости от параметра m (на основе максимального значения ранга r при $n = 100 + 2m$)

В **заклучении** приведены основные результаты и выводы по проделанной работе.

ЗАКЛЮЧЕНИЕ

В диссертационной работе решена задача развития методов построения алгоритмов комбинаторной генерации на основе применения теории производящих функций за счет использования метода получения явных выражений коэффициентов производящих функций.

Основные результаты диссертационной работы:

1. Проведен аналитический обзор современного состояния исследований в области разработки алгоритмов комбинаторной генерации, который показал, что в настоящее время существует несколько универсальных методов построения алгоритмов комбинаторной генерации: метод поиска с возвратом, ЕСО-метод, метод Ф. Флажолле, метод Б.Я. Рябко, метод В.В. Кручинина. При этом: метод поиска с возвратом и ЕСО-метод направлены только на разработку алгоритмов последовательной генерации комбинаторных объектов; существуют ограничения на возможность применения ЕСО-метода и метода Ф. Флажолле для комбинаторных множеств, описываемых более чем одним параметром; большинство методов требует представления комбинаторного объекта в специальном виде; существуют требования к наличию дополнительной информации, описывающей комбинаторное множество.

2. Проведено исследование метода построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ и метода получения явных выражений коэффициентов производящих функций. Используя данные методы предложен модифицированный метод построения алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ. Предложенный модифицированный метод отличается применением метода получения явных выражений коэффициентов производящих функций для нахождения выражения функции мощности комбинаторного множества и позволяет построить алгоритмы ранжирования и генерации по рангу для таких комбинаторных множеств, для которых не известно требуемое выражение функции мощности, но известно выражение производящей функции для последовательности значений функции мощности.

3. Проведена апробация модифицированного метода построения алгоритмов комбинаторной генерации, в ходе которой разработаны новые алгоритмы комбинаторной генерации:

– для множества выражений обобщенного языка Дика на примере множества последовательностей правильно вложенных n пар скобок t типов (оценка вычислительной сложности полученных алгоритмов равна $O(n^3)$, что является хуже по сравнению с существующими аналогами, однако это показывает возможность применения разработанного модифицированного метода для получения новых алгоритмов комбинаторной генерации);

– для множества комбинаторных объектов, отражающих вторичную структуру РНК длины n с m пар нуклеотидов, соединенных водородной связью (оценка вычислительной сложности равна $O(m^2(n-m))$), что по сравнению с существующими аналогами является лучшей, так как не требуется никаких предварительных вычислений);

– для множества комбинаторных объектов, определяемых треугольником Эйлера-Каталана, на примере множества помеченных путей Дика длины $2n$ с m подъемами на возвратных шагах (оценка вычислительной сложности разработанных алгоритмов равна $O(m^2(n-m))$).

4. На основе разработанных алгоритмов комбинаторной генерации создано программное обеспечение для ранжирования и генерации по рангу элементов комбинаторных множеств. Наличие общего интерфейса взаимодействия, а также возможность подключения дополнительных файлов с программной реализацией алгоритмов комбинаторной генерации позволяют в дальнейшем расширять перечень комбинаторных множеств, работу с которыми будет поддерживать разработанное программное обеспечение. Результаты проведенных вычислительных экспериментов полностью подтверждают полученные теоретические оценки вычислительной сложности для разработанных алгоритмов комбинаторной генерации.

СПИСОК РАБОТ, ОПУБЛИКОВАННЫХ АВТОРОМ ПО ТЕМЕ ДИССЕРТАЦИИ

Публикации в ведущих рецензируемых журналах, рекомендованных ВАК для публикации результатов кандидатских и докторских диссертационных работ:

1. Кручинин Д. В., **Шабля Ю. В.** Программное обеспечение для анализа тестов простоты натурального числа // Доклады ТУСУРа. — 2014. — № 4 (34). — С. 95–99.
2. **Шабля Ю. В.**, Кручинин Д. В., Шелупанов А. А. Генератор критериев простоты натурального числа на основе свойств композиции производящих функций // Доклады ТУСУРа. — 2015. — № 4 (38). — С. 97–101.
3. Сравнительный анализ вычислительных способов нахождения коэффициентов ряда Тейлора в математических пакетах / В. С. Мельман, **Ю. В. Шабля**, Д. В. Кручинин, В. В. Кручинин // Доклады ТУСУРа. — 2017. — Т. 20, № 4. — С. 71–74.
4. **Шабля Ю. В.**, Кручинин Д. В. Модификация метода построения алгоритмов комбинаторной генерации на основе применения теории производящих функций // Доклады ТУСУРа. — 2019. — Т. 22, № 3. — 6 с.

Публикации в научных изданиях, индексируемых Web of Science и Scopus:

5. Kruchinin D. V., **Shablya Y. V.** Explicit formulas for Meixner polynomials // International Journal of Mathematics and Mathematical Sciences. — 2015. — Vol. 2015. — P. 1–5.

6. A method for obtaining coefficients of compositional inverse generating functions / D. V. Kruchinin, **Y. V. Shablya**, V. V. Kruchinin, A. A. Shelupanov // Proceeding of International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2015). — Vol. 1738. — 2016. — P. 1–4.
 7. Integer properties of a composition of exponential generating functions / D. V. Kruchinin, **Y. V. Shablya**, O. O. Evsutin, A. A. Shelupanov // Proceeding of International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2016). — Vol. 1863. — 2017. — P. 1–4.
 8. Properties of a composition of exponential and ordinary generating functions / D. V. Kruchinin, **Y. V. Shablya**, V. V. Kruchinin, A. A. Shelupanov // Communications in Mathematics and Applications. — 2018. — Vol. 9, no. 4. — P. 705–711.
 9. A library for calculating polynomials based on composatae of generating functions / D. V. Kruchinin, V. S. Melman, **Y. V. Shablya**, A. A. Shelupanov // Proceeding of International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2018). — Vol. 2116. — 2019. — P. 1–4.
 10. Explicit formulas for the Eulerian numbers of the second kind / D. V. Kruchinin, V. V. Kruchinin, **Y. V. Shablya**, A. A. Shelupanov // Proceeding of International Conference of Numerical Analysis and Applied Mathematics (ICNAAM 2018). — Vol. 2116. — 2019. — P. 1–4.
- Публикации в тезисах и материалах научных конференций:
11. **Shablya Y. V.** The Catalan and Euler number triangles and their application // Материалы международной научно-технической конференции студентов, аспирантов и молодых ученых «Научная сессия ТУСУР-2018» (16-18 мая 2018 г.). — Т. 4. — Томск: 2018. — С. 216–219.
 12. **Шабля Ю. В.**, Мельман В. С., Репкин А. С. Исследование метода построения алгоритмов генерации комбинаторных объектов на основе деревьев И/ИЛИ // Материалы докладов XIV международной научно-практической конференции «Электронные средства и системы управления» (28-30 ноября 2018 г.), в 2 частях. — Т. 2. — Томск: В-Спектр, 2018. — С. 20–22.
 13. **Шабля Ю. В.**, Репкин А. С., Кручинин Д. В. Представление комбинаторных объектов в виде структуры дерева И/ИЛИ // Материалы докладов XIV международной научно-практической конференции «Электронные средства и системы управления» (28-30 ноября 2018 г.), в 2 частях. — Т. 2. — Томск: В-Спектр, 2018. — С. 12–15.
 14. **Shablya Y. V.**, Kruchinin D. V. Euler-Catalan's number triangle and its application // Proceedings book of the Mediterranean International Conference of Pure & Applied Mathematics and Related Areas (26-29 October 2018). — Antalya: 2018. — P. 212–216.
 15. **Шабля Ю. В.**, Репкин А. С., Кручинин Д. В. Анализ метода разработки алгоритмов комбинаторной генерации на основе деревьев И/ИЛИ // Сборник научных трудов XII всероссийской научной конференции «Наука. Технологии. Инновации» (3-7 декабря 2018 г.), в 9 частях. — Т. 2. — Новосибирск: Изд-во НГТУ, 2018. — С. 78–82.

16. **Шабля Ю. В.**, Кручинин Д. В., Репкин А. С. Сравнение подходов к разработке алгоритмов комбинаторной генерации на примере множеств перестановок и сочетаний // Материалы V международной научно-практической конференции молодых ученых «Прикладная математика и информатика: современные исследования в области естественных и технических наук» (22-24 апреля 2019 г.). — Тольятти: 2019. — С. 75–80.
17. Репкин А. С., Филиппов Г. А., **Шабля Ю. В.** Анализ современного состояния исследований в области разработки алгоритмов комбинаторной генерации // Сборник избранных статей научной сессии ТУСУРа (22-24 мая 2019 г.), в 2 частях. — Т. 2. — Томск: В-Спектр, 2019. — С. 35–37.
18. Kruchinin V. V., Kruchinin D. V., **Shablya Y. V.** On some properties of generalized Narayana numbers // Abstracts book of the 32th International Conference of the Jangjeon Mathematical Society (17-19 July 2019). — 2019.
19. **Shablya Y. V.**, Kruchinin D. V. Application of the method of compositae in combinatorial generation // Proceedings book of the 2nd Mediterranean International Conference of Pure & Applied Mathematics and Related Areas (28-31 August 2019). — Paris: 2019.

Публикации в других изданиях:

20. Realization of a method for calculating Bell polynomials based on compositae of generating functions / V. S. Melman, **Y. V. Shablya**, D. V. Kruchinin, A. A. Shelupanov // Journal of Informatics and Mathematical Sciences. — 2018. — Vol. 10, no. 4. — P. 659–672.
21. Kruchinin D. V., Kruchinin V. V., **Shablya Y. V.** Obtaining explicit formulas and identities for polynomials defined by generating functions of the form $F(t)^x G(t)^a$ // Polynomials — Theory and application. — IntechOpen, 2019. — P. 1–20.

Свидетельства о государственной регистрации программы для ЭВМ:

22. **Шабля Ю. В.**, Кручинин Д. В., Мельман В. С. Программа «PCG: Primality Criterion Generator» для генерации критериев простоты числа // Свидетельство о государственной регистрации программы для ЭВМ №2018611242 от 26.01.2018 г.
23. **Шабля Ю. В.**, Кручинин Д. В., Мельман В. С. Программа «PTA: Primality Test Analyser» для анализа тестов простоты числа // Свидетельство о государственной регистрации программы для ЭВМ №2018611241 от 26.01.2018 г.
24. **Шабля Ю. В.**, Кручинин Д. В. Программа для ранжирования и генерации по рангу элементов комбинаторных множеств // Свидетельство о государственной регистрации программы для ЭВМ №2019660020 от 29.07.2019 г.

Шабля Юрий Васильевич

Алгоритмическое обеспечение комбинаторной генерации на основе применения
теории производящих функций

Автореф. дис. на соискание ученой степени канд. техн. наук

Тираж 100 экз. Заказ ____.

Томский государственный университет
систем управления и радиоэлектроники
634050, г. Томск, пр. Ленина, 40.

Тел. (3822) 53-30-18.