МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ
имени М. В. ЛОМОНОСОВА

Механико-математический факультет

На правах рукописи

Митрофанов Иван Викторович

УДК 512.5+512.64+519.1

Алгоритмические проблемы, связанные с морфическими последовательностями

01.01.06 — математическая логика, алгебра и теория чисел

Диссертация на соискание ученой степени
кандидата физико-математических наук

Научные руководители:
доктор физико-математических наук,
профессор А. Я. Белов
доктор физико-математических наук,
профессор А. В. Михалёв

Москва — 2016

Оглавление

Глава 1

Введение

1.1 Основные теоремы диссертации.

Теорема 1.1.1. *Факторный язык непериодичного почти периодичного сверхслова является языком некоторой морфической последовательности тогда и только тогда, когда протокол детерминированной эволюции его схем Рози периодичен, возможно, с предпериодом.*

Теорема 1.1.2. *Задача проверки заключительной периодичности морфических последовательностей алгоритмически разрешима.*

Теорема 1.1.3. *Задача проверки почти периодичности морфических последовательностей алгоритмически разрешима.*

1.2 Актуальность темы

Комбинаторика слов имеет важное значение в самых разных областях математики и computer science. Ряд проблем, относящихся к комбинаторике слов, находится на стыке алгебры (при изучении базисов и нормальных форм) и символической динамики. В известной монографии французских авторов под коллективным псевдонимом М.Лотэйр (M.Lothaire [100, 101]) значительная часть посвящена приложениям в алгебре, см. также монографию М. В. Сапира [119]. Комбинаторика слов широко используется в задачах комбинаторной теории групп ([34, 42–45]), в теории алгебр Ли, вопросах бернсайдовского типа и задачах, связанных с мономиальными алгебрами [13, 52, 65, 75, 119].

Многие алгебраисты использовали в работах достаточно тонкие теоремы комбинаторики слов.

Многие проблемы комбинаторики слов представляют самостоятельный интерес. История комбинаторики бесконечных слов (сверхслов) начинается с работ Туэ, а также М. Морса и Г. Хедлунда [104]. Работы Туэ и М. Морса, а именно построение примера бесконечного бесквадратного слова (осуществленное с помощью подстановочных систем, подстановочные системы исследуются в данной диссертации), послужили отправной точкой для решения проблемы Бернсайда П. С. Новиковым и С. И. Адяном, явившимися одними из основоположников применения комбинаторики слов в теории групп и полугрупп. П. С. Новиков и С. И. Адян [42–45] провели детальное исследование свойств периодичности, находящее свое применение в самых разных разделах математики [100, 101]. Е. С. Голод и И. Р. Шафаревич [16, 17] построили конечно порожденную бесконечную периодическую группу (с неограниченной экспонентой) на основе рассмотрения нормальных форм алгебр и оценки функкий роста.

Комбинаторные соображения, возникшие в символической динамике (автоматные группы), нашли свое применение в работах С. В. Алешина [3–5] (Изложение примера С. В. Алешина – см. в книге [25]). В работах Р. И. Григорчука [18–21] идеи символической динамики и автоматные конструкции были использованы при решении проблемы Милнора – посторении групп промежуточного роста (при этом группы Григорчука периодичны). Автоматные конструкции активно используются в самых разных ситуациях (см. [13, 31, 52, 117]). Возникают они и у нас (графы Рози).

Задачи комбинаторики, и комбинаторики слов в частности, позволили решить большое число задач в теории колец [13, 52]. Комбинаторика слов активно используется в алгебрах Ли, особенно в проблемах бернсайдовского типа [28]. В теории алгебр Ли описание базиса дается в терминах так называемых “правильных слов” (базис Холла-Линдона-Ширшова).

Применив методы символической динамики (равномерно рекуррентные слова и соображения компактности), Д. Бэкелин установил [52], что любое сверхслово содержит подслово вида uvu, где u и v – правильные слова, получив, тем самым, короткое доказательство локальной нильпотентности подалгебры алгебры Ли, порожденной сэндвичами, упростив соответствующие работы А. И. Кострикина.

Одним из основоположников применения комбинаторики в теории колец является А. И. Ширшов [97], ряд глубоких комбинаторно-алгебраических результатов был получен представителями школы А. И. Ширшова [23,97]. Применив гомологическое соображение, связанное с невозможностью зацепления правильного слова за самого себя, А. И. Ширшов показал алгоритмическую разрешимость проблемы равенства в алгебрах Ли с одним определяющим соотношением. С помощью комбинаторики слов А. И. Ширшов [97] доказал теорему о свободе подалгебры свободной алгебры Ли. Для супералгебр это обобщил А. А. Михалев [38,39]. Он показал алгоритмическую разрешимость проблемы равенства для алгебр ли с одним определяющим соотношением.

Комбинаторика слов активно используется в теории PI-алгебр, достаточно упомянуть знаменитую теорему Ширшова о высоте [97] (ей посвящен раздел в монографии М. Лотэйр [101]), утверждающую возможность приведения слов к кусочно периодическому виду.

Теорема А.И.Ширшова о высоте. *Пусть A – конечно порожденная PI-алгебра степени m. Тогда существует конечный набор элементов Y и число $h \in \mathbb{N}$ такие, что A линейно представима (то есть порождается линейными комбинациями) множеством элементов вида:*

$$w = {u_1}^{k_1}{u_2}^{k_2}\cdots{u_r}^{k_r},\ \text{где } u_i \in Y \text{ и } r \leq h.$$

При этом в основе оригинальных доказательств А. И. Ширшова (как теоремы о свободе так и теоремы о высоте) лежала техника, связанная с преобразованием алфавита путем подстановок. Эта же техника используется при работе с равномерно–рекуррентными словами и в символической динамике.

Последующие доказательства теоремы о высоте [54,63] и ее обобщение для алгебр Ли [40] использовали анализ свойств периодичности. Комбинаторный анализ слов, связанный с теоремой А. И. Ширшова о высоте, явился тематикой исследования ряда авторов (см. [10,13,54]).

Понятие *роста* в алгебре является важным комбинаторным инвариантом, ему посвящена монография [96] (см. также [13,52]). Если размерность пространства, порожденного словами степени не выше n от образующих A растет как n^{λ}, то величина λ называется *размерностью Гельфанда–Кириллова алгебры A*. Размерность Гельфанда–Кириллова может быть равной 0, 1, а также любому числу $\geqslant 2$, ∞ или не существовать. То обстоятельство, что она не может принимать промежуточные значения на интервале $(1, 2)$ составляет содержание знаменитой *Bergman gap theorem.* Ассоциативная алгебра размерности Гельфанда–Кириллова 0 конечномерна. Л. Смолл показал, что ассоциативная алгебра размерности Гельфанда–Кириллова 1 является PI-алгеброй. Базисы ассоциативных алгебр размерности Гельфанда–Кириллова больше 1 с минимальной функцией роста исследовались в работах А. Т. Колотова [26, 27]. Их описание дается в терминах так называемых *последовательностей Штурма*, находящихся в центре внимания данной работы.

Обобщение понятия роста на бесконечномерный случай является понятие *ряда коразмерностей.* Первоначальное доказательство А. Регева об экспоненциальной оценке ряда коразмерности относительно свободных алгебр было улучшено В. Н. Латышевым с помощью оценки числа полилинейных n-разбиваемых слов на основе теоремы Дилуорса [32]. Использование идеи В.Н.Латышева позволило в дальнейшем получить субэкспоненциальную оценку [14,54]. Само же понятие *n-разбиваемого слова* возникло у А. И. Ширшова в его теореме о высоте. Ряды коразмерности в различных структурах исследовались в работах В. Н. Латышева, С. П. Мищенко, М. В. Зайцева, A. Giambruno. Понятию роста посвящена монография [96].

Комбинаторика слов работает в теории полугрупп. Следует указать работы Екатеринбургской школы Л. Н. Шеврина, в часности, работы М. В. Сапира, О. Г. Харлампович. Они активно применяли методы символической динамики к теории полугрупп (см. обзор [119]).

Теория колец оказалась связана с символической динамикой. В терминах почти периодических слов удалось построить теорию радикала мономиальных алгебр и решить известный вопрос о совпадении ниль-радикала и радикала Джекобсона [13, 64], кроме того, получается описание слабо нетеровых мономиальных алгебр. Каждое ненулевое слово слабо нетеровой мономиальной алгебры есть подслово из набора (сверх)слов, удовлетворяющего следующему условию: каждое слово из этого набора либо конечное, либо является бесконечным (односторонними или двухсторонними) словом, которое при выбрасывании некоторого конечного куска распадается на почи-периодические (равномерно-рекуррентные) части [12, 13].

В настоящей диссертации исследуются бесконечные слова или *сверхслова*, то есть бесконечные вправо последовательности элементов конечного множества (алфавита) и подстановочные системы.

1.2.1 Символическая динамика и комбинаторика слов.

Символическая динамика рассматривает слова как коды траектории точки в динамической системе: пусть M — компакт, $U \subset M$ — его открытое подмножество, $f : M \to M$ — гомеоморфизм компакта в себя и $x_0 \in M$ — начальная точка. *Эволюцией* точки x_0 называется бесконечное слово W над бинарным алфавитом $\{a, b\}$:

$$W_n = \begin{cases} a, & \text{если } f^n(x_0) \in U; \\ b, & \text{если } f^n(x_0) \notin U, \end{cases}$$

Если рассматривать несколько подмножеств $U_1, U_2, \ldots, U_n$, то получаются слова над алфавитом с большим числом символов.

Слова Штурма допускают определение через символическую динамику: они получаются из систем, в которых M – окружность единичной длины, f – поворот окружности на дугу α иррациональной длины, U – дуга длины α, а x_0 – произвольная точка M.

Для любой последовательности можно построить порождающую ее динамическую систему, хотя далеко не всегда динамическая система выглядит так наглядно, как для слов Штурма. Рассмотрим A^ω – множество всех сверхслов над данным алфавитом, снабженное тихоновской топологией, и оператор сдвига $s : A^\omega \to A^\omega$, действующий по правилу $(s(w))_n = w_{n+1}$. Тогда для произвольного сверхслова W это слово является кодом траектории точки, соответствующей W. Если в качестве компакта взять подмножество A^ω, являющееся замыканием орбиты W под действием s, получится топологическая динамическая система, *порожденная* сверхсловом W. Замкнутые инвариантные относительно s подмножества A^ω называются *субшифтами (subshifts)*.

Прямые и *обратные* задачи символической динамики изучают связь свойств динамической системы и комбинаторных свойств порождающего его сверхслова.

Так, динамическим системам, в которых компакт M – конечное множество, отвечают периодичные сверхслова, а *минимальным* динамическим системам, то есть системам, не содержащим нетривиальных замкнутых подсистем, соответствует свойство *почти периодичности* или *равномерно рекуррентности.* Это свойство было введено Х. Фюрстенбергом [86]. Почти периодическим словам посвящен обзор [41]. Сверхслово W называется *почти периодичным*, если для любого конечного подслова u существует такая константа $C(u)$, что u является подсловом в любом подслове W длины $C(u)$. Также почти периодичные сверхслова называют *равномерно рекуррентными.* Имеет место следующая

Теорема 1.2.1. *Если W – сверхслово, то существует почти периодичное сверхслово W' такое, что любое конечное подслово W' является конечным подсловом W.*

Эта теорема исключительно важна в комбинаторике слов, поскольку очень часто позволяет свести изучение произвольных слов к изучению почти периодических (или равномерно-реккурентных) слов (см. алгебраические приложения [12, 13, 52, 64]).

Минимальным динамическим системам с единственной мерой Хаара соответствуют сверхслова, у которых для любого конечного подслова его нижняя и верхняя плотности вхождения совпадают.

В терминах *функции рассогласования* можно сформулировать комбинаторное условие на то, что динамическая система сопряжена сдвигу тора, что дачт подход к проблематике, связанной с гипотезой Пизо. Более подробно см. в разделе 1.2.4, также в [15, 68].

1.2.2 Слова Штурма

Примером своего рода “комбинаторного рая”, источником обобщений, образцом того, как свойства динамической системы отражаются на комбинаторике слов явилась классическая красивая теорема М. Морса и Г. Хедлунда:

Теорема 1.2.2 (Теорема эквивалентности [104])**.** *Пусть W — бесконечное рекуррентное сверхслово над бинарным алфавитом $A = \{0; 1\}$.*

Следующие условия «почти» эквивалентны:

1. *Количество различных подслов длины n слова W равно $p_n(W) = n + 1$ для любого $n \geq 1$.*

2. *Слово W не периодично и является* сбалансированным*, то есть для любых двух подслов u, $v \sqsubset W$ одинаковой длины выполняется неравенство $||v|_a - |u|_a| \leq 1$, где $|w|_a$ обозначает количество вхождений символа a в слово w.*

3. *Слово $W = (w_n)$ является механическим словом с иррациональным α, то есть существуют такое иррациональное α, $x_0 \in [0, 1]$ и интервал $U \subset \mathbb{S}^1$, $U = \alpha$ такие, что выполняется условие:*
$$w_n = \begin{cases} a, & \text{если } f^n(x_0) \in U; \\ b, & \text{если } f^n(x_0) \notin U \end{cases}$$

4. *Слово W получается путем предельного перехода последовательности слов, каждое из которых получается из предыдущего путем подстановки вида*
$$a^k b \to b, a^{k+1} b \to a$$
либо подстановки вида
$$b^k a \to a, b^{k+1} a \to b.$$
Показатель k зависит от шага. Если эти показатели k_i периодически повторяются, то α есть квадратичная иррациональность.

Слово «почти» значит то, что симметрические разности этих классов сверхслов не более чем счетны, и все исключения описаны. Например, при $\alpha \in \mathbb{Q}$ механическое слово не принадлежит первому классу. Пересечение всех этих классов называется множеством *слов Штурма.*

Условие 2 говорит о *комбинаторной сложности*, или *функции* сверхслова. М. Морс и Г. Хедлунд показали, что если у сверхслова W для какого-то n $p_w(n) < n + 1$, то p_w ограниченна и слово w является периодичным. Поэтому *слова Штурма* – это непериодичные слова с минимально возможной функцией роста.

Условие 4 связано с понятиями *подстановки* и *морфизма.* Так, сверхслово $01001010\ldots$, переходящее само в себя при одновременной замене каждого нуля на 01, а единицы – на 0, называется *словом Фибоначчи*, а предел отношения числа единиц и нулей в нем равен золотому сечению.

Для наших целей важно и иное задание слов Штурма, а именно *одноразвилковыми графами Рози* (см. раздел 1.2.3). Если при этом слово Штурма связано с квадратичной иррациональностью, то последовательность событий в этих графах оказывается заключительно периодической. Исследование многоразвилковых схем приводит к теоремам типа теоремы Вершика-Лившица.

Понятие *последовательностей Штурма* и их обобщения послужило отправной точкой многих исследований. Естественными обобщениями слов Штурма являются слова с минимальным ростом, то есть слова с функцией роста $T(n) = n + k$, начиная с некоторого n. Для двубуквенных алфавитов они носят название *квазиштурмовых слов.* Сверхслова с функцией роста, удовлетворяющей соотношению

$$\lim_{n\to\infty} T(n)/n = 1$$

изучены в работе A. Aberkane [59]. Другим обобщением слов Штурма является обобщение, связанное с понятием *сбалансированности*, а также *m-сбалансированности.* Сбалансированные непериодические слова над n-буквенным алфавитом изучены в работе R. L. Graham [87]. В работе А. Л. Чернятьева [72] построена порождающая динамическая система для сбалансированных непериодических слов.

Обощением понятия поворота окружности служат *перекладывания отрезков* (поворот окружности реализуется перекладыванием двух отрезков). В работах А. Я. Белова и А. Л. Чернятьева [95] (см. также диссертацию А. Л. Чернятьева [56]), дан критерий того, что сверхслово получается с помощью перекладывания отрезков (отдельно рассмотрен случай, когда некоторые отрезки переворачиваются), в терминах *размеченных графов Рози.* Позднее S. Ferenczi, L. Zamboni [85] был получен критерий для перекладываний (без переворотов), удовлетворяющих условию т.н. *IDOC-condition.*

Итак, с одной стороны, в терминах размеченных графов Рози мы получаем перевод топологических свойств динамической системы, а с другой стороны, важнейшим комбинаторным свойством является морфичность последовательности. Одним из основных результатов автора является перевод свойства морфичности на язык периодичности схем Рози. Утверждение о периодичности схем Рози можно рассматривать как обобщение теоремы Вершика-Лившица, оно позволяет решить ряд алгоритмических проблем, ряд из которых был исследован и поставлен А. А. Мучником, А. Л. Семеновым с учениками [41, 47] (см. разделы 1.2.4, 1.2.5).

Перекладываниям отрезков посвящена обширная библиография, они связаны с рядом важных и интересных задач в математике, в частности, в теории дифференциальных уравнений, динамических систем. Рассмотрим векторное поле на двумерном многообразии, разделенном разрезами на клетки. Движение частиц, летящих вдоль векторного поля, доставляет кусочно-непрерывное преобразование множества разрезов. Инвариантная мера задачт метрику, а тем самым и перекладывания отрезков, которые управляют динамическими системами в двумерном случае. В частности, рассмотрим биллиард с рациональными углами наклона. Тогда если рассмотреть динамическую систему «сторона + угол траектории», то, проведя аналогичные рассуждения, можно заключить, что перекладывания отрезков управляют поведением траектории.

Возвращаясь к словам Штурма, можно сказать, что слова Штурма, задающиеся подстановочными системами, связаны с квадратичными иррациональностями. Если мы рассматриваем общие перекладывания отрезков, то сочетание теоремы о периодичности схем Рози [125] с критерием того, когда последовательность получается из перекладывания отрезков [56, 95], предоставляет возможность строить подстановочные системы, связанные с высшими иррациональностями, которые задают динамическую систему, связанную с перекладыванием отрезков. Есть надежда получить глобальный анализ поведения для некоторых биллиардных траекторий.

1.2.3 Графы Рози

Факторный язык – такое множество конечных слов, что вместе с любым словом языка ему принадлежат и все его подслова. Примером могут служить факторные языки всех конечных подслов данного сверхслова, такой язык будет обозначаться $L(W)$. *Граф Рози* (или граф подслов) $G_W(n)$ порядка n сверхслова W – это орграф, вершинами которого являются слова $L(W)$ длины n, и из слова $a_1 \ldots a_n$ ведет ребро в $a_2 \ldots a_{n+1}$ если и только если $a_1 a_2 \ldots a_{n+1} \in L(W)$.

Графы Рози были введены в работе Ж. Рози, и они часто используются для изучения сверхслов низкой комбинаторной сложности. Например, так как сложность слов Штурма равна $n+1$, то любой граф Рози для любого слова Штурма содержит ребер на одно больше, чем вершин, следовательно, является объединением двух циклов. пересекающихся по вершине или цепочке ребер.

Комбинаторный анализ графов Рози позволил П. А. Лаврову [98] и независимо И. И. Богданову и Г. Р. Челнокову [69] получить верхнюю оценку на длину минимального периода периодичного слова, задаваемого n запретами. Ж. Кассинь [70] использовал технику графы Рози при доказательстве того, что у не более чем линейной функции комбинаторной сложности сверхслова не могут быть неограниченные первые разности, это эквивалентно тому, что графы Рози сверхслов с не более чем линейной функцией сложности имеют ограниченное число *развилок*, т.е. вершин степени > 1.

Рассмотрение слов с линейной функцией сложности приводит к изучению слов, порождаемых *перекладыванием отрезков.* Эти преобразования были введены В. И. Оселедецом, следовавшим идее В. И. Арнольда. Известно, что если перекладывание k отрезков *регулярно*, то есть траектория любого из концов отрезка перекладывания не попадает на конец другого отрезка, то слово, порождаемое данным перекладыванием, имеет комбинаторную сложность $p_n = n(k-1)+1$. Ж. Рози впервые показал, что связь между вращениями круга и последовательностями Штурма обобщается, если рассматривать перекладывания отрезков, и задал вопрос описания последовательностей, связанных с перекладываниями отрезков.

А. Я. Беловым и А. Л. Чернятьевым был получен комбинаторный критерий в терминах графов Рози [95] (в том числе для случая, когда отрезки переворачивались), в работе S.Ferenczi, L. Zamboni [85] был получен другой критерий (для частного случая наличия IDOC condition). В этой связи естественно возник вопрос о переводе свойств подстановочности на язык графов Рози. Если подстановочность означает своего рода периодичность форм графов Рози, а с другой стороны, свойство быть перекладыванием отрезков переводится на язык графов Рози, то возникают конструкции перекладываний, связанные с подстановочными системами. При этом собственные числа или плотности букв связаны с алгебраическими иррациональностями произвольных степеней, а отнюдь не только с квадратичными иррациональностями, как в случае последовательностей Штурма.

Таким образом, исследование графов (схем) Рози оказалась тесно связанной с теоремами типа теоремы Вершика-Лившица.

1.2.4 Морфические последовательности и теоремы типа Вершика-Лившица

Чисто подстановочные (чисто морфические) последовательности имеют вид $\varphi^\infty(a)$, подстановочные (ими морфические) имеют вид $h(\varphi^\infty(a))$. Это – основной объект изучения в данной диссертации.

Слово Туэ – бесквадратная чисто подстановочная последовательность.

$$abcabacaac\cdots = \varphi^\infty(a),$$

где

$$\varphi(a) = abcab, \varphi(b) = acabcb, \varphi(c) = acbcacb.$$

Легко понять, что над алфавитом из двух букв не может быть бесквадратного сверхслова, а слово *Туэ-Морса* $0110100\ldots$, полученное из буквы 0 подстановкой $\varphi(0) = 01$, $\varphi(1) = 10$, является примером *бескубного* бинарного слова.

Избегание квадратов – это избегание паттерна AA, кубов – паттерна AAA. Для большинства *избегаемых* паттернов избегающие их слова впервые построены с помощью подстановочных систем [101].

Матрица подстановки φ – это такая матрица M_φ, у которой в i-ом столбце в j-ой строке стоит число вхождений буквы a_j в $\varphi(a_i)$. Если некоторая степень матрицы не содержит нулей, то эта матрица (а также соответствующая подстановка и подстановочное слово) называется *примитивной*. У примитивной матрицы по теореме Перрона-Фробениуса есть вектор, соответствующий наибольшему положительному собственному значению.

Если все остальные собственные значения по модулю меньше единицы, то подстановка называется *подстановкой Пизо.*

Гипотеза 1.2.1 (Гипотеза Пизо)**.** *У сверхслова, полученного с помощью неприводимой подстановки Пизо, динамическая система имеет число дискретный спектр и измеримо сопряжена сдвигу тора.*

На данный момент гипотеза доказана для двухбуквенного алфавита M.Hollander, B.Solomyak [89], есть продвижения в общем случае.

Как известно, у примитивных морфических последовательностей комбинаторная сложность – не более чем линейная функция, это делает графы Рози перспективным инструментом для их изучения.

А. Э. Фрид [57] описала последовательность графов Рози для некоторого подкласса равноблочных подстановочных слов, что позволило, например, точно вычислить для таких сверхслов функцию комбинаторной сложности, так как графы Рози очень хорошо описывают факторный язык сверхслова. В последовательности графов Рози оказывается конечное число *топологических типов* (топологический тип – то, что получится из графа, если заменить все длинные простые пути ребрами), и эти типы сменяют друг друга периодическим образом.

Последовательность событий (называемых «переключениями»), наблюдаемых при изучении последовательности графов Рози сверхслова Штурма, заключительно периодична тогда и только тогда, когда это сверхслово является подстановочным. Эта заключительная периодичность соответствует заключительная периодичности разложения в цепные дроби квадратичных иррациональностей. Это наблюдение, а также работа [95], привели А. Я. Белова к гипотезе, что аналогичный результат верен для намного более широкого класса морфических последовательностей, а именно – всех почти периодичных. В дальнейшем оказалось, что конструкция графов Рози нуждается в видоизменении. Соответствующий объект (последовательность *схем Рози*) описан в разделе 3.

Отметим другие результаты, связывающие подстановочность сверхслова с комбинаторно-динамическими свойствами.

Ф. Дюранд [78] изучал дискретный аналог отображения первого возвращения Пуанкаре – *производные последовательности (derived sequences)*. Пусть x – рекуррентное сверхслово, а U – какой-то его начальный кусок. Отметим все первые буквы вхождений U в x, они разбивают x на конечные слова. Если различных типов этих слов конечное число, закодируем их последовательность всеми первыми несколькими членами натурального ряда так, чтобы одинаковые слова кодировались одинаковыми числами, а различные – разными. Такая последовательность чисел называется *производной последовательностью* x относительно U.

Теорема 1.2.3. *Пусть x – почти периодичное сверхслово. Оно является морфическим тогда и только тогда, когда у x конечное число производных последовательностей относительно всевозможных начальных кусков.*

Теорема Вершика-Лившица говорит о периодичности диаграмм Брателли для марковских компактов, порожденных подстановочными системами [121, 122]. Ее доказательство основано на явной конструкции.

Рассмотрим образ $V_n = \varphi^{(n)}(a) = (\varphi^{(n-2)})(\varphi^{(2)}(a))$ где a есть буква алфавита, φ есть некоторый примитивный морфизм полугруппы слов. V_n состоит из блоков отвечающих применению $\varphi^{(n-2)}$ к буквам слова $\varphi^{(2)}(a)$. Кроме того, V_n можно представить в виде произведения блоков отвечающих применению $\varphi^{(n-1)}$ к буквам слова $\varphi(a)$. Конечные множества, образующие диаграммы Брателли состоят из последовательностей пар первого рода (блок, его позиция в блоке на единицу большего размера), которые отвечают собственным подсловам блоков на единицу большего размера (все блоки одного уровня) а также последовательности второго рода – последовательность первого рода для размера n начинающаяся или заканчивающаяся последовательностью первого рода для предыдущего размера. При этом естественная замена этой дополнительной последовательности на пару (соответствующий блок чьим концом или началом она является, его позиция) естественным образом приводит к появлению последовательности первого рода. (Особо надо рассмотреть случай когда при заполнении возникает блок на единицу большего рода чем все присутствующие). Естественным образом на этих множествах вводятся стрелки, означающие что соответствующий объект есть начальное или концевое подслово другого объекта. Детали конструкции – см. [121, 122] а также [99].

1.2.5 Алгоритмические проблемы

Весьма плодотворным оказывается взгляд на комбинаторику слов с точки зрения математической логики, во многом инициированный обзором А. А. Мучника, А. Л. Семенова, Ю. Л. Притыкина [41] (см. также диссертации [47, 49]). Этот обзор позволил авторам посмотреть на почти периодические последовательности, адические компакты и теоремы типа теоремы Вершика-Лившица с другой стороны. Нам представляется это интересным как в плане самих алгоритмических задач, так и в для осознания интересных связей, для понимания роли техники графов Рози и адических компактов.

В этоих работах были поставлены вопросы, относящиеся, в частности, к алгоритмическим проблемам проверки периодичности, почти периодичности $HDOL$-систем. Для случая примитивной подстановки алгоритмическая разрешимость этих вопросов была установлена Ю. Л. Притыкином [47]. Благодаря применению техники связанной с теоремами типа теоремы Вершика-Лившица нам удалось доказать алгоритмическую разрешимость в общем случае.

Морфическая последовательность $h(\varphi(a))$ задается конченым набором информации – подстановочной системой (A, B, φ, h, a), где A и B – алфавиты, $\varphi : A^* \to A^*$ – подстановка, $h : A^* \to B^*$ – кодирование, $a \in A$ – начальная буква.

Естественным образом встают вопросы об алгоритмическом описании различных свойств морфической последовательности и об алгоритмическом сравнении различных последовательностей [41]. Везде далее ставится вопрос о существовании алгоритма, на вход которому подается одна или две подстановочные системы.

Проблема 1.2.1 (Периодичность). *Является ли слово $h(\varphi^\infty(a))$ заключитльно периодичным?*

Нам не известно, когда эта задача была поставлена впервые. Были построены алгоритмы в ряде частных случаев. Для чисто морфических последовательностей алгоритм был найден в 1986 году в работах T. Harju и M. Linna [88] и J.-J. Pansiot [109]. В 1986 году, J. Honkala [90] описан алгоритм для *автоматных последовательностей* (это последовательности вида $h(\varphi^\infty(a))$ такие, что образы φ от всех букв имеют одинаковую длину).

В работе J. Honkala и M. Rigo [91] дается эквивалентная формулировка проблемы в терминах распознаваемых подмножеств натуральных чисел в абстрактных системах счисления. Описание распознаваемых подмножеств натуральных чисел в абстрактных системах счисления (A. Maes и M. Rigo [103]), вкупе с теоремой 4.0.1 дает алгоритм проверки того, представляется ли распознаваемое подмножество натуральных чисел в произвольной системе счисления в виде объединения конечного количества арифметических прогрессий.

Общий случай рассматривается в разделе 4.

Проблема 1.2.2 (Совпадение слов)**.** *Пусть есть две морфические последовательности w_1 и w_2, можно ли алгоритмически понять, равны ли эти сверхслова?*

Для число подстановочных последовательностей, эта задача (называемая задачей ω-эквивалентности для D0L систем) была решена в 1986 году M. Linna и T. Harju [88], а также J.J. Pansiot [109]. В 2012 году F. Duran [77] показал алгоритмическую разрешимость этой задачи для примитивных морфических слов.

Следующий вопрос был поставлен А. Л. Семеновым [41]:

Проблема 1.2.3 (Почти периодичность)**.** *Является ли данное морфическое слово почти периодичным?*

В работе F.Nicolas, Ю. Л. Притыкина [107] также ставится алгоритмическая задача проверки почти периодичности морфических слов. В обзоре Ю. Л. Притыкина, Ан. А. Мучник и А. Л. Семенова [41] была выдвинута гипотеза о том, что эта задача алгоритмически разрешима.

Для автоматных последовательностей, в работе F. Nicolas и Ю. Л. Притыкина [107] строится полиномиальный алгоритм в классах чисто морфических и автоматных слов.

В разделе 5 доказывается алгоритмическая разрешимость почти периодичности морфических сверхслов.

Проблема 1.2.4 (Совпадение факторных языков). *Пусть есть две морфические последовательности. Существует ли алгоритм, проверяющий, совпадают ли их факторные языки?*

В 1997 I. Fagnot строит алгоритм для некоторого класса число морфических последовательностей [83] и для автоматных последовательностей [84]. Случай, в котором одна из последовательностей примитивна, разбирается как основная лемма в разделе 5.

Проблема 1.2.5 (Фактор и сопряженность динамических систем). *Пусть есть два субшифта, порожденных двумя морфическими последовательностями. Можно ли алгоритмически определить, сопряжены ли эти системы? Можно ли определить, является одна из них фактором другой?*

В работе E. Coven, M. Keane и M. LeMasurier [74] были найдены некоторые ответы на этот вопрос для случая, когда один из субшифтов порожден словом Морса.

1.3 Цель работы

Целью настоящей работы является изучение морфических последовательностей. Перед автором возникли следующие задачи:

- Обобщить соответствие между подстановочностью и периодичным поведением графов Рози на произвольные почти периодичное слова, рассматривая при этом, возможно, конструкцию, несколько отличающуюся от графов Рози. Создать теорию связанную с теоремами типа теоремы Вершика-Лившица. При этом должен получиться новый способ описания структуры морфических последовательностей.

- Получить алгоритмически проверяемые критерии для почти периодичности произвольного морфического сверхслова; периодичности и заключительной периодичности произволного морфического серхслова.

Все эти задачи успешно решены автором.

1.4 Научная новизна

Все результаты диссертации являются новыми. Основные результаты диссертации заключаются в следующем:

1. Создана теория связанная со схемами Рози и теоремами типа Вершика-Лившица. Дано определение схем Рози сверхслова, эволюции схем Рози, протокола эволюции схем Рози. Доказано, что протокол эволюции схем Рози для почти периодичного сверхслова периодичен тогда и только тогда, когда его факторный язык языком некоторой морфической последовательности.

2. Доказана алгоритмическая разрешимость проверки периодичности и заключительной периодичности морфических последовательностей. Тем самым был получен положительный ответ на вопрос, известный с 70-х годов [110].

3. Доказана алгоритмическая разрешимость проверки почти периодичности морфических последовательностей. Тем самым был получен положительный ответ на вопрос, поставленный А. А. Мучником и А. Л. Семеновым в 2003 году.

4. Доказана алгоритмическая разрешимость проверки совпадения факторных языков двух морфических последовательностей в случае, когда одна из них примитивна.

Результат 1 получен в соавторстве с А. Я. Беловым, остальные результаты получены самостоятельно. Результаты 2 и 3 при этом получены одновременно и независимо с Ф. Дюраном [80, 81].

1.5 Краткое содержание диссертации

Глава 1 является введением диссертации. Она содержит описание актуальности темы, цели работы, список основных результатов и сведения об апробации работы.

В **главе 2** приведен обзор основных понятий, используемых в диссертации.

В **главе 3** показана периодичность схем Рози, тем самым сделан перевод языка подстановочных систем на язык схем Рози.

Определение 1.5.1. *Графом со словами* будем называть связный ориентированный граф, у которого на каждом ребре которого написано по два слова – *переднее* и *заднее*, а кроме того, каждая вершина либо имеет входящую степень 1, а исходящую больше 1 *(раздающая вершина)*, либо входящую степень больше 1 и исходящую степень 1 *(собирающая вершина)*.

Симметричный путь в графе со словами – это путь, первое ребро которого начинается в собирающей вершине, а последнее ребро кончается в раздающей. Каждый путь можно записать словом над алфавитом, являющимся множеством ребер графа. Оно называется *реберной записью пути.* Ребро пути s, идущее i-тым по счету, будем обозначать $s[i]$. Естественно определены отношения *подпути* (пишем $s_1 \sqsubseteq s_2$), *начала* и *конца.* $s_1 \sqsubseteq_k s_2$, если для соответствующих слов u_1 и u_2 – реберных записей путей s_1 и s_2 – выполнено $u_1 \sqsubseteq_k u_2$. Если последнее ребро пути s_1 идет в ту же вершину, из которой выходит первое ребро пути s_2, путь, реберная запись которого является конкатенацией реберных записей путей s_1 и s_2, будем обозначать $s_1 s_2$.

Введем понятие *переднего слова $F(s)$, соответствующего пути s в графе со словами.* Пусть $v_1 v_2 \dots v_n$ – реберная запись пути s. В $v_1 v_2 \dots v_n$ возьмем подпоследовательность: включим в нее v_1, а также те и только те ребра, которые выходят из раздающих вершин графа. Эти ребра назовем *передними образующими для пути s.* Возьмем передние слова этих ребер и запишем их последовательную конкатенацию, так получаем $F(s)$. Аналогично определяется понятие *заднего слова $B(s)$*. В пути $v_1 v_2 \dots v_n$ возьмем ребра, входящие в собирающие вершины и ребро v_n в порядке следования – это *задние образующие для пути s.* Тогда последовательной конкатенацией задних слов этих ребер получается *заднее слово $B(s)$ пути s.*

Определение 1.5.2. Граф со словами будет являться *схемой Рози* для рекуррентного непериодичного сверхслова W, если он удовлетворяет следующим свойствам:

1. Граф сильносвязен и состоит более чем из одного ребра.

2. Все ребра, исходящие из одной раздающей вершины графа, имеют передние слова с попарно разными первыми буквами. Все ребра, входящие в одну собирающую вершину графа, имеют задние слова с попарно разными последними буквами.

3. Для любого симметричного пути, его переднее и заднее слова совпадают.

4. Если есть два симметричных пути s_1 и s_2 и выполнено $F(s_1) \sqsubseteq_k F(s_2)$, то $s_1 \sqsubseteq_k s_2$.

5. Все слова, написанные на ребрах графа, являются подсловами W.

6. Для любого $u \sqsubset W$ существует симметричный путь, слово которого содержит u.

7. Для любого ребра s существует такое слово u_s, принадлежащее W, что любой симметричный путь, слово которого содержит u_s, проходит по ребру s.

Симммегричный путь s называется *допустимым*, если $F(s) \sqsubseteq W$.

Если S – сильносвязный граф, не являющийся циклом, и s – путь в графе, то *естественное продолжение* пути *s вправо* – это минимальный путь, началом которого является s и который оканчивается в раздающей вершине. *Естественное продолжение* пути *s влево* – это минимальный путь, концом которого является s и который начинается в собирающей вершине.

Пусть W – рекуррентное непериодичное сверхслово, S – схема Рози для этого сверхслова. Из сильносвязности S следует, что существует ребро v, идущее из собирающей вершины в раздающую. Такие ребра будем называть *опорными*.

Определяется S'' – *эволюция* (W, S, v) *схемы Рози S по опорному ребру v*: само ребро v уничтожается и заменяется на минимальные пути через v строго содержащие v, отвечающие подсловам W. При этом естественным образом определяются передние и задние слова для нового графа со словами. Построенный таким образом граф со словами S'' назовем *элементарной эволюцией* (W, S, v).

Лемма 1.5.1. *Пусть S – схема Рози. Тогда элементарная эволюция (W, S, v) является схемой Рози для сверхслова W (т.е. удовлетворяет свойствам (1)–(7) определения 1.5.2).*

Далее показывается связь между эволюцией схем Рози и графами Рози. Описывается процесс, ставящий в соответствие каждому графу Рози схему Рози и доказывается, что схемы, полученные их двух последоватеьных графов Рози, либо совпадают, либо получаются одна из другой одной или несколькими элементарными эволюциями.

Эволюция схем Рози – это последовательность применений элементарных эволюций. На каждом шаге могут быть несколько опорных ребер для применения элементарной эволюции, и следующее определение введено для снятия неоднозначности выбора.

Пронумеруем ребра схемы Рози и зафиксируем *метод эволюции* – алгоритм, который по пронумерованной схеме определяет, какое опорное ребро использовать, а после перенумеровывает ребра в проэволюционировавшей схеме. Тогда последовательность схем Рози определяется однозначно. Потребуем также, чтобы метод эволюции не зависел от слов, написанных на ребрах, то есть работал с *облегченной нумерованной схемой.*

Облегченная нумерованная схема для схемы Рози – граф с теми же вершинами и ребрами, ребра графа пронумерованы, а слов на них нет.

Протокол детерминированной эволюции – последовательность облегченных нумерованных схем Рози, номеров соответствующих опорных ребер, а также информация о минимальных путях, содержащих опорное ребро и не являющихся допустимыми.

Теорема 1.5.1. *Если W – почти периодичное сверхслово и выбран некоторый метод эволюции, то протокол эволюции периодичен тогда и только тогда, когда у W такой же факторный язык, как у некоторого морфического слова.*

Автор считает этот результат диссертации самым интересным.

Сложной частью является доказательство периодичности протокола эволюции для морфических слов. Сначала задача решается в предположении, что подстановка примитивна, а общий случай сводится к примитивному после с помощью леммы из работы Ю. Л. Притыкина, а также дополнительного соображения:

Лемма 1.5.2. *Следующие три условия эквивалентны: а) Слово* $\varphi^\infty(a)$ *не является почти периодическим. б) В* $\varphi^\infty(a)$ *есть бесконечно много* φ*-ограниченных подслов. в) Существует непустое* $w \in A^*$ *такое, что* w^n *является подсловом* $\varphi^\infty(a)$ *для любого* n*.*

Доказательство того, что из периодичности протокола следует эквивалентность морфическому слову, проще.

Глава 4 посвящена задаче проверки периодичности морфических последовательностей. Первые продвижения по этой задаче (для чисто морфических последовательностей) были получены в работе T. Harju and M. Linna [88], также задача была поставлена в обзоре А. Л. Семеновым [41], где была высказана гипотеза об ее алгоритмической разрешимости.

Иными словами, строится алгоритм для решения следующей задачи:

Дано: *два алфавита* A *и* B*, подстановка* $\varphi : A^* \to A^*$*, продолжающаяся над буквой* a *и морфизм* $h : A^* \to B^*$*.*

Определить: *существуют ли такие конечные слова* u *и* v*, что* $h(\varphi^\infty(a)) = uv^\infty$ *(т.е. является ли сверхслово заключительно периодичным.)*

Сначала задача решается в предположении, что подстановка φ примитивна, в этом случае достаточно найти период протокола эволюции схем Рози. После общий случай сводится к примитивному.

В **главе 5** решается задача о почти периодичности морфических последовательностей, поставленная в обзоре Ан. А. Мучник, Ю. Л. Притыкина и А. Л. Семенова. [41].

Она сводится к следующей алгоритмической задаче: **Дано:**

1. Три конечных алфавита A*,* B*,* C*;*

2. Числа K_1*,* K_2*,* $1 < \Theta_1 < \Theta_2$*;*

3. Четыре морфизма $\varphi : A^* \to A^*$*,* $\psi : A^* \to C^*$*,* $g : B^* \to B^*$*,* $h : B^* \to C^*$ *такие, что*

а) все морфизмы нестирающие;

b) *морфизм g продолжается над b_1, морфизм φ примитивен и продолжается над $a_1 \in A$;*

c) *для некоторого $\lambda \in [\Theta_1; \Theta_2]$ и всех $a_i \in A$, $b_j \in B$, $k \in \mathbb{N}$ выполнено $K_1\lambda^k < |\psi(\varphi^k(a_i))| < K_2\lambda^k$ и $K_1\lambda^k < |h(g^k(b_j))| < K_2\lambda^k$.*

d) *сверхслово $\psi(\varphi^\infty(a_1))$ непериодично;*

Определить: *верно ли, что любое подслово сверхслова $h(g^\infty(b_1))$ является подсловом сверхслова $\psi(\varphi^\infty(a_1))$?*

А эта задача, в свою очередь, исследуется с помощью анализа схем Рози.

В **главе 6** предлагаются альтернативные решения задач периодичности и почти периодичности морфических последовательностей.

1.6 Основные методы исследования

В работе используются методы комбинаторики слов, теории графов Рози и схем Рози, символической динамики, теории неотрицательных матриц.

1.7 Аппробация работы

Результаты диссертации докладывались автором на следующих научно-исследовательских семинарах:

1. Семинар «Кольца и модули» под руководством профессора В. Н. Латышева, профессора А. В. Михалева неоднократно в 2011-2014 гг.;
2. Научно-исследовательский семинар кафедры алгебры, 2015.
3. Семинар «Алгебра и теория моделей» под руководством профессора Е. И. Буниной в 2016г.
4. Научно-исследовательский семинар А. М. Райгородского в 2011–2012 гг.
5. Bar-Ilan Algebra Seminar (Bar-Ilan University) (January, 2014)
6. PI-Seminar (Technion (Israel Institute of Technology))(January, 2014)

7. Петербургский семинар по теории представлений и динамическим системам (2015, ПОМИ)

Результаты диссертации докладывались на всероссийских и международных конференциях:

1. Международная конференция «CANT-2012» г. Люмини в 2012 г
2. Международный воркшоп Decidability problems for substitutive sequences, tilings and numerations, организованный F.Durand (2012, Амьен)
3. Международная конференция «SubTite» г. Люмини в 2013 г
4. Franco-Russian workshop on Algorithms, complexity and applications 2013, Москва
5. Number Theory and Dynamics + Journees Arithmetiques, Institut Fourier, 2013, Гренобль
6. Growth, Symbolic Dynamics and Combinatorics of Words in Groups (2015, Париж)

1.8 Теоретическая и практическая ценность работы

Диссертация имеет теоретический характер. Результаты диссертации могут представляют интерес для специалистов в комбинаторике слов, логике, алгебре, информатике, динамических системах, специалистов МГУ, МПГУ, НГУ, СПбГУ, LAMFA, l'Institut de Mathematiques de Luminy.

1.9 Публикации

Результаты автора по теме диссертации опубликованы в 3 работах, список которых приведен в конце библиографии.

1.10 Структура и объем диссертации

Диссертация состоит из шести глав, заключения и списка литературы.

1.11 Благодарности

Автор глубоко благодарен своем научным руководителям — доктору физико-математических наук профессору Алексею Яковлевичу Белову и доктору физико-математических наук Александру Васильевичу Михалеву за постановку задач, обсуждение результатов и постоянное внимание к работе.

Также автор хотел бы поблагодарить за внимание и обсуждения работы доктора физико-математических наук, профессора Виктора Николаевича Латышева и всех участников семинара “Теория колец”.

Автор выражает свою отдельную благодарность за полезные обсуждения А. Л. Семенову, А. М. Вершику и всем участникам его семинара, А. М. Райгородскому и всем участникам его семинара за полезные обсуждения.

Работа выполнена при частичной финансовой поддержке фонда Дмитрия Зимина «Династия» и гранта РФФИ No14-01-00548.

Глава 2

Основные определения, обозначения и технические утверждения.

Слово u над конечным алфавитом $\{a_i\}$ — это последовательность букв: $a_{i_1}a_{1_2}\dots a_{i_n}$. Слова бывают *конечными* и *бесконечными*, бесконечные слова бывают *бесконечными вправо*, *бесконечными влево* или *бесконечными в обе стороны*. Бесконечные слово также будет называться *сверхсловом*. На буквах слова можно ввести нумерацию, будем требовать, чтобы в словах, не бесконечных слева, нумерация совпадала с естественной нумерацией от левого конца. Тот факт, что на i-том месте в слове u стоит буква a, будем обозначать $u[i] = a$.

Для конечного слова u определена *длина* $|u|$ — количество букв в нем. Если слово u_1 не бесконечно справа, а u_1 — не бесконечно слева, то определена их *конкатенация* u_1u_2 – слово, получающееся приписыванием второго к первому справа.

Слово v является *подсловом* слова u, если $u = v_1vv_2$ для некоторых слов v_1, v_2. В случае, когда v_1 или v_2 – пустое слово, v называется *началом* или соответственно *концом* слова u. На словах существует естественная структура частично упорядоченного множества: $u_1 \sqsubseteq u_2$, если u_1 является подсловом u_2. Будем обозначать $u_1 \sqsubseteq_k u_2$, если слово u_1 входит в u_2 хотя бы k раз.

Сверхслово W называется *рекуррентным*, если любое его подслово встречается в W бесконечно много раз, иначе говоря, $v \sqsubseteq W \Rightarrow v \sqsubseteq_\infty W$. Сверхслово W называют *почти периодичным* (*равномерно рекуррентным*), если для любого его подслова v существует такое число $k(v, W)$, что если $u \sqsubseteq W$ и $|u| \geq k(v, W)$, то $v \sqsubseteq u$.

Сверхслово W называется *почти периодичным*, если для любого его подслова u конечной длины существует такое число $N(u)$, что для любого подслова сверхслова W длины $N(u)$ в этом подслове есть хотя бы одно вхождение u (следовательно, W является рекуррентным).

Сверхслово W называется *периодичным*, если $W = uuuuu\dots$ для некоторого непустого u. Само слово u, равно как и его длина, называются *периодом* сверхслова W. Сверхслово называется *заключительно периодичным*, если оно представляется в виде конкатенации $W = uW'$, где u — конечное слово, а W' – периодичное сверхслово.

Предложение 2.0.1. *Если сверхслово заключительно периодично и рекуррентно, то оно периодично.*

Сверхслово W называется *периодичным* с периодом k, если для любого i выполнено $W[i+k] = W[i]$, и *заключительно периодичным* с периодом k, если для любого i начиная с некоторого N выполнено $W[i+k] = W[i]$.

Множество слов A^* над алфавитом A можно считать свободным моноидом с операцией конкатенации и единицей — пустым словом. Отображение $\varphi\colon A^* \to B^*$ называется *морфизмом*, если оно сохраняет операцию моноида. Очевидно, морфизм достаточно задать на буквах алфавита A. Морфизм называется *нестирающим*, если образом никакой буквы не является пустое слово.

Если алфавиты A и B совпадают и существует такая буква a_1, что $\varphi(a_1) = a_1 u$ для некоторого слова u и $\varphi^k(u)$ не является пустым словом ни для какого k, то бесконечное слово

$$a_1 u \varphi(u) \varphi^2(u) \varphi^3(u) \varphi^4(u) \dots$$

называется *чисто морфическим*, пишут $W = \varphi^\infty(a_1)$.

Если существует такая степень морфизма φ^k, что для любых двух букв a_i содержится в $\varphi^k(a_j)$, морфизм называют *примитивным*.

Для бесконечного вправо слова W определены *графы Рози*. Граф Рози *порядка* k обозначается $G_k(W)$, если же понятно, о каком слове идет речь, то будем писать просто G_k. Вершины его соответствуют всевозможным различным подсловам длины k сверхслова W. Две вершины графа u_1 и u_2 соединяются направленным ребром, если в W есть такое подслово v, что $|v| = k + 1$,

$v[1]v[2]\dots v[k] = u_1$ и $v[2]v[3]\dots v[k+1] = u_2$. Если w — подслово W длины $k+l$, ему соответствует в G_k путь длины l, проходящий по ребрам, соответствующим подсловам слова w длины $k+1$.

Глава 3

Схемы Рози

3.1 Определение схем Рози.

В графах Рози, определчнных в конце части 2, естественно заменить блинные цепочки рчбер единичными рчбрами. На этом пути получаются *схемы Рози*.

Графом со словами будем называть связный ориентированный граф, у которого на каждом ребре которого написано по два слова – *переднее* и *заднее*, а кроме того, каждая вершина либо имеет входящую степень 1, а исходящую больше 1, либо входящую степень больше 1 и исходящую степень 1. Вершины первого типа назовем *раздающими*, а второго — *собирающими*.

Путь в графе со словами — это последовательность ребер, каждое следующее из которых выходит из той вершины, в которую входит предыдущая. *Симметричный путь* — это путь, первое ребро которого начинается в собирающей вершине, а последнее ребро кончается в раздающей.

Каждый путь можно записать словом над алфавитом — множеством ребер графа, которое называется *реберной записью пути*. Иногда мы будем отождествлять путь и его реберную запись. Ребро пути s, идущее i-тым по счету, будем обозначать $s[i]$. Для двух путей так же, как и для слов, определены отношения *подпути* (пишем $s_1 \sqsubseteq s_2$), *начала* и *конца*. Кроме того, пишем $s_1 \sqsubseteq_k s_2$, если для соответствующих слов u_1 и u_2 — реберных записей путей s_1 и s_2 – выполнено $u_1 \sqsubseteq_k u_2$. Если последнее ребро пути s_1 идет в ту же вершину, из которой выходит первое ребро пути s_2, путь, реберная запись которого является конкатенацией реберных записей путей s_1 и s_2, будем обозначать $s_1 s_2$.

Замечание 3.1.1. *Определения пути, подпути, реберной записи, начала и конца имеют смысл для любых графов ориентированных графов.*

Введем понятие *переднего слова $F(s)$, соответствующего пути s в графе со словами.* Пусть $v_1v_2\dots v_n$ — реберная запись пути s. В $v_1v_2\dots v_n$ возьмем подпоследовательность: включим в нее v_1, а также те и только те ребра, которые выходят из раздающих вершин графа. Эти ребра назовем *передними образующими для пути s*. Возьмем передние слова этих ребер и запишем их последовательную конкатенацию, там получаем $F(s)$.

Аналогично определяется понятие *заднего слова $B(s)$*. В пути $v_1v_2\dots v_n$ возьмем ребра, входящие в собирающие вершины и ребро v_n в порядке следования — это *задние образующие для пути s*. Тогда последовательной конкатенацией задних слов этих ребер получается *заднее слово $B(s)$ пути s*.

Определение 3.1.1. Граф со словами будет являться *схемой Рози* для рекуррентного непериодичного сверхслова W, если он удовлетворяет следующим свойствам, которые в дальнейшем будут называться *свойствами схем Рози*:

1. Граф сильно связен и состоит более чем из одного ребра.
2. Все ребра, исходящие из одной раздающей вершины графа, имеют передние слова с попарно разными первыми буквами. Все ребра, входящие в одну собирающую вершину графа, имеют задние слова с попарно разными последними буквами.
3. Для любого симметричного пути, его переднее и заднее слова совпадают. То есть можно говорить просто о слове симметричного пути.
4. Если есть два симметричных пути s_1 и s_2 и выполнено $F(s_1) \sqsubseteq_k F(s_2)$, то $s_1 \sqsubseteq_k s_2$.
5. Все слова, написанные на ребрах графа, являются подсловами W.
6. Для любого $u \sqsubset W$ существует симметричный путь, слово которого содержит u.

7. Для любого ребра s существует такое слово u_s, принадлежащее W, что любой симметричный путь, слово которого содержит u_s, проходит по ребру s.

3.2 Свойства схем Рози.

Лемма 3.2.1. *Пусть S — схема Рози для сверхслова W. Тогда верно следующее:*

1. *Если путь s_1 оканчивается в раздающей вершине и в этой же вершине начинается путь s_2, то $F(s_1 s_2) = F(s_1)F(s_2)$.*

2. *Если путь s_1 оканчивается в собирающей вершине и в этой же вершине начинается путь s_2, то $B(s_1 s_2) = B(s_1)B(s_2)$.*

1. Множество образующих передних ребер пути $s_1 s_2$ — это в точности образующие передние ребра пути s_1 и образующие передние ребра пути s_2, записанные последовательно.

2. Множество образующих задних ребер пути $s_1 s_2$ — это образующие задние ребра пути s_1 и образующие задние ребра пути s_2, записанные последовательно.

Замечание 3.2.1. *При доказательстве этой леммы никак не использовалось, что S – схема Рози. То есть утверждение верно для любого графа со словами.*

Следствие 3.2.1. *Если в графе со словами S выполнено свойство 3 схем Рози, s – симметричный путь в S, s_1 — произвольный путь в S и $s_1 \sqsubseteq s$, то $F(s_1) \sqsubseteq F(s)$.*

Лемма 3.2.2. *Если s_1 и s_2 — два симметричных пути таких, что $s_1 \sqsubseteq_k s_2$, то $F(s_1) \sqsubseteq_k F(s_2)$.*

Доказательство. Пусть $v_1 v_2 \dots v_n$ — реберная запись пути s_2, а $w_1 w_2 \dots w_m$ — реберная запись пути s_1. Для каждого вхождения слова $w_1 w_2 \dots w_m$ в $v_1 v_2 \dots v_n$ ту часть $v_1 v_2 \dots v_n$, которая идет до последнего ребра вхождения (включительно), назовем *путевым началом*, а все, что после — *путевым* *концом.* Так как ребро w_m входит в раздающую вершину, то любое путевое начало является симметричным путем, а первое ребро любого путевого конца является в пути s_2 передним образующим ребром. Если s_b и s_e — путевые начало и конец соответственно, то $F(s_2) = F(s_b)F(s_e)$, при этом $F(s_b)$ оканчивается на $F(s_1)$. Для доказательства леммы достаточно показать, что передние слова всех путевых концов различные. Для любого путевого конца s_e множество его передних образующих ребер — это подмножество передних образующих ребер пути s_2, пересеченное с s_e. Но для различных путевых концов такие подмножества вложены одно в другое, а так как для любого путевого конца его первое ребро является образующим, то они вложены строго. $\square$

Замечание 3.2.2. *Доказательство леммы проходит для любого графа со словами, для которого выполнено свойство (3) схем Рози.*

Определение 3.2.1. Симмметричный путь s называется *допустимым*, если $F(s) \sqsubseteq W$.

Лемма 3.2.3. *Если $u \sqsubseteq W$, то в схеме S есть допустимый путь s такой, что $u \sqsubseteq F(s)$.*

Пусть $l_{\max}$ — максимальная из длин слов на ребрах S. Так как сверхслово W рекуррентно, то в нем есть вхождение слова u такое, что первая буква этого вхождения имеет в W номер более $l_{\max}$. Рассмотрим w — подслово W, имеющее вид u_1uu_2, где $|u_1| > l_{\max}$, $|u_2| > l_{\max}$. Согласно свойству (6), в схеме S существует симметричный путь s_1 такой, что $w \sqsubseteq F(s_1)$. Можно считать, что s_1 – минимальный (относительно отношения $\sqsubseteq$) симметричный путь с таким свойством.

Пусть $v_1v_2\dots v_n$ — реберная запись этого пути. Пусть v_{n_1} — последнее ребро, являющееся передним образующим для пути s_1. Это либо v_1, либо ребро, выходящее из раздающей вершины. Если верно первое, то в пути s_1 только одно переднее образующее ребро и $|F(s_1)| \leq l_{\max}$. Значит, можно рассмотреть симметричный путь с реберной записью $v_1v_2\dots v_{n_1-1}$. Из минимальности s_1 следует, что $w \not\sqsubseteq F(v_1v_2\dots v_{n_1-1})$.

Рассуждая аналогично, получим, что если v_{n_2} — первое ребро, являющееся задним образующим, то $w \not\sqsubseteq F(v_{n_2}v_{n_2+1}\dots v_n)$.

Докажем следующий факт: $n_2 \leq n_1$.

В самом деле, иначе, в силу минимальности n_2, среди ребер $v_1, v_2, \dots v_{n_1}$ ни одно не входит в собирающую вершину, то есть в симметричном пути $v_1 v_2 \dots v_{n_1-1}$ ровно одно заднее образующее ребро (а именно v_{n_1-1}) и, стало быть, $|B(v_1 v_2 \dots v_{n_1-1})| \leq l_{\max}$. Но

$$F(s_1) = F(v_1 v_2 \dots v_{n_1-1}) F(v_{n_1}) = B(v_1 v_2 \dots v_{n_1-1}) F(v_{n_1}) \leq 2 l_{\max}.$$

Противоречие с тем, что $|F(s_1)| > 2 l_{\max}$.

Значит, мы можем рассматривать симметричный путь

$$s_2 = v_{n_2} v_{n_2+1} \dots v_{n_1-1}$$

. Пусть $w' = F(s_2)$. Мы можем записать $F(s_1) = B(v_{n_2-1}) w' F(v_{n_1})$. Так как слово $F(s_1)$ содержит w, а слова $B(v_{n_2-1}) w'$ и $w' F(v_{n_1})$ не содержат, то $w' \sqsubseteq w$. Стало быть, $w' \sqsubseteq W$.

С другой стороны, $F(s_1) = u_1' u_1 u u_2 u_2'$, где $\min\{|u_1' u_1|, |u_2 u_2'|\} \geq l_{\max}$. А так как $\max\{B(v_{n_2-1}), F(v_{n_1})\} \leq l_{\max}$, то $u \sqsubseteq w'$. Следовательно, путь s_2 — искомый.

Замечание 3.2.3. *В доказательстве леммы 3.2.3 не использовалось свойство (7) схем Рози.*

Лемма 3.2.4. *Пусть имеется допустимый путь l со словом u. Если $u u_1 \sqsubseteq W$ для некоторого u_1, то в схеме Рози S есть такой допустимый путь l', что его началом является l, а его слово начинается с $u u_1$.*

Согласно лемме **3.2.3** в схеме S существует допустимый путь l_2 такой, что $u u_1 \sqsubseteq F(l_2)$. Рассмотрим слово $F(l_2)$. В нем может быть несколько вхождений слова u, причем среди них есть такое (допустим, k—тое, если считать с конца), после которого сразу идет u_1.

Тогда $l \sqsubseteq_k l_2$. Пусть $l_2 = s_1 l s_2$ (здесь рассматривается k—тое с конца вхождение пути l). Рассмотрим k-тое с конца вхождение пути l в путь l_2. Первое ребро пути l выходит из собирающей вершины, а последнее ребро пути s_2 входит в раздающую вершину. Следовательно, $l s_2$ — симметричный путь. Этот путь является допустимым, так как $B(l_2) = B(s_1) B(l s_2)$ и $B(l_2) \sqsubseteq W$.

Так как l имеет ровно k вхождений в ls_2, то, согласно свойству 4 определения схем Рози 3.1.1 и лемме **3.2.4**, $u = F(l)$ имеет ровно k вхождений в $F(ls_2)$. Кроме того, $F(ls_2)$ начинается с $F(l) = u$. Этому свойству удовлетворяет ровно одно окончание слова $F(l_2)$. Значит, $F(ls_2)$ начинается со слова uu_1. Стало быть, путь ls_2 – искомый.

Замечание 3.2.4. *Аналогично доказывается следующий факт: пусть имеется допустимый путь l со словом u. Если $u_1u \sqsubseteq W$ для некоторого u_1, то в схеме Рози S есть такой допустимый путь l', что его концом является l, а его слово кончается на u_1u.*

Следствие 3.2.2. *Если u — биспециальное подслово W такое, что оно содержит слово некоторого симметричного пути l_1 схемы S, то в схеме S существует такой симметричный путь l, что $F(l) = u$.*

Пусть $u = u_1F(l_1)u_2$. Из биспециальности u следует, что существуют буквы a_1 и a_2 такие, что $F(l_1)u_2a_1 \sqsubseteq W$, $F(l_1)u_2a_2 \sqsubseteq W$. Тогда существуют два пути l_1s_1 и l_1s_2 такие, что слово первого начинается с $F(l_1)u_2a_1$, а второго — с $F(l_1)u_2a_2$.

Пусть $v_1v_2 \dots v_{k-1}v_k \dots$ — реберная запись пути l_1s_1, а $v_1v_2 \dots v_{k-1}v'_k \dots$ — реберная запись пути l_1s_2, различие происходит в ребре с номером k. Очевидно, ребра v_k и v'_k выходят из одной и той же вершины. Значит, эта вершина раздающая и путь $v_1v_2 \dots v_{k-1}$ — симметричный. По свойству 2 схем Рози (см. определение 3.1.1), $F(v_k)$ и $F(v'_k)$ начинаются с разных букв. Так как $F(l_1s_1)$ начинается с $F(v_1v_2 \dots v_{k-1})F(v_k)$, а $F(l_1s_2)$ — с $F(v_1v_2 \dots v_{k-1})F(v'_k)$, то $F(v_1v_2 \dots v_{k-1}) = F(l_1)u_2$. При этом путь $v_1v_2 \dots v_{k-1}$ начинается с l_1.

Аналогично рассуждая, найдем симметричный путь $w_1w_2 \dots w_m$, концом которого является путь l_1 и словом которого является $u_1F(l_1)$. Пусть $w_1w_2 \dots w_m = l'l_1$. Тогда путь $l'v_1v_2 \dots v_{k-1}$ является искомым.

3.3 Получение схем Рози из графов Рози.

Пусть W — рекуррентное непериодичное сверхслово. Фиксируем натуральное число k. Для упрощения дальнейшего считаем, что в W нет биспециальных слов длины ровно k.

Рассмотрим G_k — граф Рози порядка k для сверхслова W.

Предложение 3.3.1. *G_k есть сильносвязный орграф, не являющийся циклом.*

Если u_1 и u_2 — подслова W длины k, то, в силу рекуррентности сверхслова W, в W найдется подслово вида $a_{i_1}a_{i_2}\dots a_{i_n}$, где $a_{i_1}a_{i_2}\dots a_{i_k} = u_1$, $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n} = u_2$.

Тогда слова вида $a_{i_l}a_{i_{l+1}}\dots a_{i_{l+k}}$ имеют длину $k+1$ и соответствуют ребрам графа G_k, образующим путь, соединяющий вершины, соответствующие словам u_1 и u_2. Кроме того, если бы G_k был циклом длины n, то сверхслово W было бы периодичным с периодом n.

В силу выбора k, в графе G_k нет вершин с входящей и одновременно исходящей степенью более 1. Построим граф со словами S, вершинами которого будут специальные вершины графа G_k, а ребра — простыми цепями, соединяющие специальные вершины G_k.

Предложение 3.3.2. *Пути в графе S соответствуют путям в G_k, начинающимся и кончающимся в специальных вершинах.*

Пусть простой путь проходит в графе G_k по вершинам

$$a_{i_1}a_{i_2}\dots a_k,\ a_{i_2}a_{i_3}\dots a_{k+1},\ \dots, a_{i_l}a_{i_{l+1}}\dots a_{i_{l+n-1}},\ \dots,\ a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$$

и соединяет специальные вершины $a_{i_1}a_{i_2}\dots a_k$ и $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$. Сопоставим этому пути *переднее* и *заднее* слова по следующему правилу:

1. Если вершина, соответствующая $a_{i_1}a_{i_2}\dots a_{i_k}$, явлется раздающей, то *переднее* слово пути — это $a_{i_{k+1}}a_{i_{k+2}}\dots a_{i_n}$.

2. Если вершина, соответствующая $a_{i_1}a_{i_2}\dots a_{i_k}$, явлется в собирающей, то *переднее* слово пути — это $a_{i_1}a_{i_2}\dots a_{i_n}$.

3. Если вершина, соответствующая $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$, является собирающей, то *заднее* слово пути — это $a_{i_1}a_{i_2}\dots a_{i_n-k}$.

4. Если вершина, соответствующая $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$, является раздающей, то *заднее* слово пути — это $a_{i_1}a_{i_2}\dots a_{i_n}$.

Определение 3.3.1. Если S — сильно связный граф, не являющийся циклом, и s – путь в графе, то *естественное продолжение* пути s *вправо* — это минимальный путь, началом которого является s и который оканчивается в раздающей вершине. *Естественное продолжение* пути s *влево* — это минимальный путь, концом которого является s и который начинается в собирающей вершине.

Очевидно, для сильносвязных нецикличных графом естественное продолжение существует всегда и единственно. Теперь мы готовы написать слова на ребрах S: для каждого ребра в качестве переднего слова берется переднее слово того пути в G_k, который соответствует естественному расширению вправо этого ребра.

Заднее же слово — это заднее слово пути в G_k, соответствующего естественному продолжению влево рассматриваемого ребра.

Предложение 3.3.3. *В полученном графе со словами S для любого пути s слово $F(s)$ — это переднее слово того пути графа G_k, который соответствует естественному продолжению вправо пути s. Аналогично, $B(s)$ — это заднее слово того пути в G_k, который соответствует естественному продолжению пути s влево (в графе S).*

Действительно, в графе S для любого пути s его естественное продолжение вправо разбивается на естественные продолжения вправо его передних образующих ребер, причем все естественные продолжения, начиная со второго, выходят из раздающих вершин и, стало быть, слова на них пишутся по правилу для первого из четырех типов. Значит, слова для этих ребер в объединении и дадут слово для длинного пути в G_k.

Лемма 3.3.1. *Пусть S — определенный выше граф со словами. Тогда он является схемой Рози для сверхслова W.*

Докажем выполнение всех свойств схемы Рози.

Свойство 1. Так как G_k является сильносвязным и не является циклом, то же самое верно и для S.

Свойство 2. Два пути, выходящие из одной раздающей вершины графа S с различными первыми ребрами, соответствуют путям в графе G_k, которые выходят из одной раздающей вершины $a_{i_1}a_{i_2}\dots a_{i_k}$, причем первые ребра у путей разные. Пусть вторые вершины путей — это $a_{i_2}a_{i_3}\dots a_{i_k}b$ и $a_{i_2}a_{i_3}\dots a_{i_k}c$ соответственно. Тогда слова этих путей начинаются на буквы b и c соответственно, и на эти буквы начинаются слова соответствующих путей в S.

Свойство 3 следует из того, что для симметричного пути его естественным продолжением вправо и естественным продолжением влево является он сам. Согласно 3.3.3, передним и задним словом такого пути будут соответственно переднее и заднее слово пути

$$a_{i_1}a_{i_2}\dots a_k,\ a_{i_2}a_{i_3}\dots a_{k+1},\ \dots, a_{i_l}a_{i_{l+1}}\dots a_{i_{l+n-1}},\ \dots,\ a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$$

графа G_k. Переднее слово этого пути определяется по типу 2, а заднее — по типу 4, то есть они совпадают.

Докажем **свойство 4**. Пусть в S нашлись два симметричных пути s_1 и s_2 такие, что $F(s_1) \sqsubseteq_k F(s_2)$. Соответственные пути в G_k обозначим s_1' и s_2'. Очевидно, достаточно показать, что $s_1' \sqsubseteq_k s_2'$. пути s_1' и s_2' выходят из собирающих вершин графа G_k и входят в раздающие вершины. Передние слова этих путей — это $F(s_1)$ и $F(s_2)$. Последовательности их $(k+1)$-буквенных подслов — это последовательности ребер s_1' и s_2'. Значит, реберная запись пути s_1' встречается в реберной записи s_2' хотя бы k раз.

Свойство 5 достаточно доказать для передних слов. Пусть v — ребро схемы S. Рассмотрим в S естественное продолжение ребра v вправо. Оно соответствует пути s в G_k. Пусть s содержит L ребер. Первая вершина этого пути — слово $a_{i_1}a_{i_2}\dots a_{i_k}$. Первое ребро пути s соответствует подслову u вида $a_{i_1}a_{i_2}\dots a_{i_k}b$, при этом $a_{i_1}a_{i_2}\dots a_{i_k}b \sqsubseteq W$. Очевидно, существует u_1 — подслово сверхслова W длины $L+k$, начинающееся с $a_{i_1}a_{i_2}\dots a_{i_k}b$. Ему соответствует путь s' длины L по ребрам G_k с первым ребром таким же, как у пути s.

Так как среди промежуточных вершин пути s нет раздающих вершин (иначе s не соответствует естественному продолжению ребра v вправо), то s' должен совпадать с путем s. Стало быль, слово пути s является подсловом u_1 и, следовательно, подсловом W.

Свойство 6. Существует такое число M, что любом пути длины M по ребрам G_k будет хотя бы одна раздающая и хотя бы одна собирающая вершина. В самом деле, в качестве M подойдет количество вершин в G_k, иначе существовал бы цикл, в котором не было бы либо раздающих, либо собирающих вершин и G_k был бы либо не сильносвязным, либо циклом.

Пусть $w \sqsubseteq W$. Тогда существует $w' \sqsubseteq W$ вида w_1ww_2, где $|w_1| = 2M$, $|w_2| = 2M$. Рассмотрим в G_k путь, соответствующий слову w'. Среди его последних M вершин есть раздающая, а среди первых M — собирающая. Значит, можно взять путь, который короче не более, чем на $2M$ и соответствует симметричному пути в S. Этот путь соответствует подслову w', которое имеет длину не менее $|w'| - 2M$ и, следовательно, содержит w. А это значит, что слово некоторого симметричного пути в S содержит w.

Докажем **свойство 7**. Пусть v — ребро в графе S. В графе Рози G_k ребру v соответствует путь v', проходящий через вершины

$$a_{i_1}a_{i_2}\ldots a_{i_k},\ a_{i_2}a_{i_3}\ldots a_{i_{k+1}},\ \ldots, a_{i_l}a_{i_{l+1}}\ldots a_{i_{l+n-1}},\ \ldots,\ a_{i_{n-k+1}}a_{i_{n-k}}\ldots a_{i_n}.$$

Ни одна из вершин пути кроме первой и последней не является раздающей или собирающей.

Для S уже доказаны свойства (1)–(6). Согласно замечанию 3.2.3, для S справедлива лемма 3.2.3. Для слова $a_{i_1}a_{i_2}\ldots a_{i_k}a_{i_{k+1}}$, являющегося подсловом сверхслова W, найдем в S допустимый путь w такой, что $a_{i_1}a_{i_2}\ldots a_{i_k}a_{i_{k+1}} \sqsubseteq F(w)$. Соответственный путь в графе G_k содержит первое ребро пути v', а следовательно, и весь v'. Стало быть, w содержит v, а так как для S выполнено свойство (4), то любой симметричный путь, содержащий $F(w)$, содержит ребро v.

3.4 Эволюция схем Рози.

Далее W — рекуррентное непериодичное сверхслово, S — схема Рози для этого сверхслова. Из сильной связности S следует, что существует ребро v, идущее из собирающей вершины в раздающую. Такие ребра будем называть *опорными*.

Цель этого раздела — определить *эволюцию* (W, S, v) схемы Рози S по опорному ребру v. Опорное ребро является симметричным путем, стало быть, переднее и заднее слова этого ребра совпадают.

Пусть $\{x_i\}$ — множество ребер. входящих в начало v, а $\{y_i\}$ — множество ребер, идущих из конца v. Эти два множества могут пересекаться. Обозначим $F(y_i) = Y_i$, $B(x_i) = X_i$, $F(v) = V$.

Рассмотрим все слова вида X_iVY_j. Если такое слово не входит в W, то пару (x_i, y_j) назовем *плохой*, в противном случае — *хорошей*.

Построим граф S'. Он получается из S заменой ребра v на $K_{\#\{x_i\},\#\{y_j\}}$. Более подробно: ребро v удаляется, его начало заменяется на множество из $\{A_i\}$ из $\#\{x_i\}$ вершин так, что для любого i ребро x_i идет в A_i; конец ребра v заменяется на множество $\{B_j\}$ из $\#\{y_j\}$ вершин так, что для любого j ребро y_j ведет в B_j; вводятся ребра $\{v_{i,j}\}$, соединяющие вершины множества $\{A_i\}$ с вершинами множества $\{B_j\}$.

По сравнению с S, у графа S' нет ребра v, но есть новые ребра $\{v_{ij}\}$. Остальные ребра графа S имеют соответственные ребра в S', ребра в первом и во втором графе зачастую будут обозначаться одними и теми же буквами.

На ребрах S' расставим слова следующим образом. На всех ребрах S', кроме ребер из $\{y_i\}$ и $\{v_{i,j}\}$, передние слова пишутся те же, что и передние слова соответственных ребер в S. Для каждого i и j, переднее слово ребра y_j в S' — это VY_j. Переднее слово ребра $v_{i,j}$ — это Y_j. Аналогично, на всех ребрах, кроме ребер из $\{x_i\}$ и $\{v_{i,j}\}$, задние слова переносятся с соответствующих ребер S; для всех i и j в качестве заднего слова ребра x_i возьмем X_iV, а в качестве заднего слова ребра $v_{i,j}$ — X_i.

Теперь построим граф S''. Ребра $v_{i,j}$, соответствующие плохим парам (x_i, y_j), назовем *плохими*, а все остальные ребра графа S' — хорошими. Граф S'' получается, грубо говоря, удалением плохих ребер из S'. Более точно, в графе S'' раздающие вершины — подмножество раздающих вершин S', а собирающие вершины — подмножество собирающих вершин S'. В графе S' эти подмножества — вершины, из которых выходит более одного хорошего ребра, и вершины, в которые входит более одного хорошего ребра соответственно. Назовем эти вершины S' неисчезающими. Ребра в графе S'' соответствуют таким путям в S', которые идут лишь по хорошим ребрам, начинаются в неисчезающих вершинах, заканчиваются в неисчезающих, а все промежуточные вершины которых не являются неисчезающими.

Согласно лемме 3.4.1, которая будет доказана ниже в этом же разделе, S'' – сильно связный граф и не цикл. Следовательно, у каждого ребра в S'' есть естественное продолжение вперед и назад. Естественное продолжение вперед соответствует пути в S'; переднее слово этого пути в S' возьмем в качестве переднего слова для соответствующего ребра S''. Аналогично для задних слов: у ребра в S'' есть естественное продолжение влево, этому пути в S'' соответствует путь в S'. Заднее слово этого пути и будет задним словом ребра в S''.

Определение 3.4.1. Построенный таким образом граф со словами S'' назовем *элементарной эволюцией* (W, S, v).

Доказательству следующей теоремы будет посвящена основная часть раздела.

Теорема 3.4.1. *Пусть W – схема Рози. Тогда элементарная эволюция (W, S, v) является схемой Рози для сверхслова W (то есть удовлетворяет свойствам (1)—(7) определения 3.1.1).*

Определение 3.4.2. Пусть S — схема Рози для сверхслова W. На ее ребрах можно написать различные натуральные числа (или пары чисел). Такую схему мы назовем *нумерованной.* Если с ребер пронумерованной схемы стереть слова, получится *облегченная нумерованная схема.*

Определение 3.4.3. *Метод эволюции* — это функция, которая каждой облегченной пронумерованной схеме (с нумерацией, допускающей двойные индексы) дает этой же схеме новую нумерацию, такую, что в ней используются числа от 1 до n для некоторого n по одному разу каждое.

Зафиксируем какой-либо метод эволюции и далее не будем его менять.

Среди опорных ребер пронумерованной схемы S возьмем ребро v с наименьшим номером, и совершим эволюцию (W, S, v). Укажем естественную нумерацию новой схемы.

Напомним, что сначала строится схема S', а потом — S''. В схеме S' все ребра можно пронумеровать по следующему правилу: ребра кроме v сохраняют номера, а ребра вида v_{ij} нумеруют соответствующим двойным индексом.

Каждое ребро в схеме S'' — это некоторый путь по ребрам схемы S', различным ребрам из S'' соответствуют в S' пути с попарно различными первыми ребрами. Таким образом, ребра схемы S'' можно пронумеровать номерами первых ребер соответствующих путей S'. Теперь применим к облегченной нумерованной схеме S'' метод эволюции. Получится новая облегченная нумерованная схема, и в нумерованной (не облегченной) схеме S'' перенумеруем ребра соответственным образом.

Определение 3.4.4. Описанное выше соответствие, ставящее нумерованной схеме Рози другую нумерованную схему Рози, назовем *детерминированной эволюцией.* Будем обозначать это соответствие $S'' = \mathrm{Evol}(S)$.

Определение 3.4.5. Применяя детерменированную эволюцию к нумерованной схеме S много раз, получаем последовательность нумерованных схем Рози. Кроме того, на каждом шаге получаем множество пар чисел, задающие плохие пары ребер. *Протокол детерминированной эволюции* — это последовательность таких множеств пар чисел и облегченных нумерованных схем.

Предложение 3.4.1. *Облегченная нумерованная схема для* $\mathrm{Evol}(S)$ *однозначно определяется по облегченной нумерованной схеме* S *и множеству пар чисел, задающие плохие пары ребер.*

Сформулируем основной результаты работы:

Теорема 3.4.2. *Если сверхслово* W *является морфическим с примитивным порождающим морфизмам, то протокол детерминированной эволюции с некоторого места периодичен.*

В разделе 3.8 теорема обобщается на случай произвольного почти периодичного морфического слова.

Дальнейшая часть раздела посвящена доказательству теоремы 3.4.1.

Рассмотрим следующее соответствие f между симметричными путями в S за исключением опорного ребра v и симметричными путями в S'. Пусть s — симметричный путь в S имеет реберную запись $v_1v_2\dots v_n$. Часть этих ребер (может быть, ни одно) являются ребром v. Очевидно, если ребро v входит в путь, то оно находится в между ребрами x_i и y_j для некоторых i и j. Если v является первым или последним ребром s, то соответствующее x_i или

y_j находится только с одной стороны. Соответствие f преобразует $v_1v_2\dots v_n$ (реберную запись s) следующим образом: для каждого вхождения v определяется, находится ли оно на конце реберной записи или в середине, если на конце, то "v" просто отбрасывается, а если в середине, то "v" заменяется на "v_{ij}". Индексы i и j берутся те же, что и у ребер x_i и y_j, стоящих рядом с v в реберной записи s. Путь с соответствующей реберной записью является образом $f(s)$ при соответствии.

Обратное соответствие f^{-1} таково: если S' есть путь l, и его реберная запись — $v_1v_2\dots v_n$, то каждое вхождение "v_{ij}" заменяется на "v". Кроме того, если v_1 — это одно из $\{y_i\}$, то в начало реберной записи приписывается "v", если же v_n — одно из ребер $\{x_i\}$, то "v" приписывается в конец реберной записи.

Очевидно, f и f^{-1} сохраняют свойство последовательности ребер *начало каждого следующего ребра является концом предыдущего*, то есть пути переводят в пути. Кроме того, f и f^{-1} на самом деле являются взаимно обратными. Далее, в графе S' ребра вида y_j выходят из собирающих вершин, ребра вида x_i входят в раздающие, ребра v_{ij} выходят из раздающих вершин и входят в собирающие; в графе S ребро v, как опорное, выходит из собирающей и входит в раздающую, ребра x_i входят в собирающую вершину, а y_j выходят из раздающей, для всех остальных ребер графа S свойство выходить из собирающей вершины или входить в раздающую сохраняется в графе S'. Таким образом, при соответствиях f и f^{-1} сохраняется симметричность путей.

Предложение 3.4.2. *При этом соответствии сохраняются слова симметричных путей.*

Докажем этот факт для передних слов, так как для задних все аналогично. А следовательно, мы докажем, что передние и задние слова симметричных путей в S' совпадают.

Пусть в графе S имеется симметричный путь s с реберной записью $v_1v_2\dots v_n$. Разберем два случая.

1. Пусть первое ребро $v_1 = v$, а последнее $v_n \neq v$. Рассмотрим все вхожде-

ния ребра v:

$$s = vy_{j_1} \dots vy_{j_2} \dots vy_{j_3} \dots v_n.$$

Пусть $f(s) = s'$.

$$s' = y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots v_{i_3 j_3} y_{j_3} \dots v_n.$$

Рассмотрим передние образующие ребра в s и s''. В пути s в парах вида vy_j окажется выбранным ребро y_j, кроме первой пары, где выбраны будут оба ребра. В s' в парах $v_{ij}y_j$ окажутся выбраны ребра v_{ij}, при этом первое ребро y_{j_1} тоже будет передним образующим. Остальные передние образующие ребра в обоих путях одни и те же. Таким образом,

$$F(s) = F(v)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$$F(s') = F(y_{j_1}) \dots F(v_{i_2 j_2}) \dots F(v_{i_3 j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

Фрагменты, отмеченные многоточиями, одинаковы в обоих наборах.

2. Пусть $v_1 = v$, $v_n = v$. Аналогично первому случаю, рассмотрим вхождения v в реберную запись пути s.

$$s = vy_{j_1} \dots vy_{j_2} \dots vy_{j_3} \dots x_{i_k} v.$$

$$s' = f(s) = y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots v_{i_3 j_3} y_{j_3} \dots x_{i_k}.$$

Тогда

$$F(s) = F(v)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$$F(s') = F(y_{j_1}) \dots F(v_{i_2 j_2}) \dots F(v_{i_3 j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

Как мы видим, слова совпадают.

3. Пусть $v_1 \neq v$, $v_n = v$. Аналогично, рассмотрим вхождения v в реберную запись пути s.

$$s = v_1 \dots vy_{j_1} \dots vy_{j_2} \dots x_{i_k} v.$$

$$s' = f(s) = v_1 \dots v_{i_1 j_1} y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots x_{i_k}.$$

Тогда

$$F(s) = F(v_1)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = F(v_1)Y_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$$F(s') = F(v_1)F(v_{i_1j_1})\dots F(v_{i_2j_2})\dots F(v_{i_3j_3})\dots =$$
$$= F(v_1)Y_{j_1}\dots Y_{j_2}\dots Y_{j_3}\dots$$

4. Пусть $v_1 \neq v$, $v_n \neq v$. Рассмотрим реберную запись пути s.

$$s = v_1 \dots v y_{j_1} \dots v y_{j_2} \dots v_n.$$

$$s' = f(s) = v_1 \dots v_{i_1j_1} y_{j_1} \dots v_{i_2j_2} y_{j_2} \dots v_n.$$

Тогда

$$F(s) = F(v_1)F(y_{j_1})\dots F(y_{j_2})\dots F(y_{j_3})\dots = F(v_1)Y_{j_1}\dots Y_{j_2}\dots Y_{j_3}\dots$$

$$F(s') = F(v_1)F(v_{i_1j_1})\dots F(v_{i_2j_2})\dots F(v_{i_3j_3})\dots =$$
$$= F(v_1)Y_{j_1}\dots Y_{j_2}\dots Y_{j_3}\dots$$

Предложение 3.4.3. *Если в графе S два симметричных пути $s_1 \neq v$, $s_2 \neq v$ и $s_2 \sqsubseteq_n s_1$, то $f(s_1) \sqsubseteq_n f(s_2)$.*

Напомним, что v — это выделенное опорное ребро. Рассмотрим реберные записи путей s_1 и s_2. Вторая реберная запись входит в первую не менее n раз.

Соответствие f делает с этими записями следующее: каждое вхождение “v” либо отбрасываются, если оно находилось с краю реберной записи, либо меняется на некоторою ребро, зависящее только от соседей “v” в записи справа и слева.

Рассмотрим некое вхождение второй реберной записи в первую. Все буквы-ребра, не являющиеся “v”, при соответствии f не меняются. Если буква “v” находится не на конце реберной записи s_2, то в обеих реберных записях она меняется на одну и ту же букву. Если “v” находится в начале или конце второго слова, при соответствии f во втором слове она исчезает.

Следовательно, при отображении s подпоследовательность реберной записи пути s_1, являющаяся реберной записью s_2, переходит в подпоследовательность реберной записи s', содержащую реберную запись s_2. Кроме того, так как $s_2 \neq v$, в реберной записи s_2 есть первое с начала ребро, не меняющееся при отображении s (то есть не являющееся v). Для каждого вхождения второй реберной записи в первую отметим это ребро. Рассмотрим образ первого слова при соответствии f. Отмеченные ребра перейдут в отмеченные ребра, стоящие в различным местах реберной записи пути $f(s)$. При этом образ каждого из отмеченных ребер будет началом реберной записи $f(s_2)$. Таким образом, в реберной записи первого слова будет не менее n подслов, являющихся реберной записью пути $f(s_2)$.

Доказанное утверждение вкупе с предложением 3.4.2 моментально влечет тот факт, что для графа S' выполняется свойство 4 схем Рози.

Далее, из построения S' очевидно, что для любой раздающей вершины передние слова ребер, выходящих из нее, начинаются с различных букв, а задние слова ребер, входящих в одну собирающую вершину, кончаются на различнные буквы.

Предложение 3.4.4. *Для графа со словами S' выполнены свойства (1)–(4) схем Рози.*

Осталось доказать свойство **1**. В силу сильной связности S, в S существует цикл, проходящий по всем ребрам S, а также по всем тройкам ребер вида x_ivy_j. Если применить к нему отображение f, то получится цикл в S', проходящий по всем ребрам. То, что S' — не цикл, очевидно, так как у него есть раздающие ребра.

Составим $\{u_i\}$ — набор подслов сверхслова W. Для каждого ребра $v \in S$ включим в этот набор слово u_s — слово, соответствующее этому ребру по свойству 7 для схем Рози. Кроме того, включим в набор те слова вида X_iVY_j, которые являются подсловами W. В сверхслове W существует подслово U_1, содержащее все слова набора $\{u_i\}$. Пусть $l_{\max}$ – максимальная из длин слов на ребре S. В W найдется слово U_2 вида U_1wU_1, где $|w| > 10l_{\max}$.

Применим к U_2 лемму 3.2.3. Получим допустимый путь s с реберной записью $v_1v_2\dots v_n$. Слово $F(s)$ имеет вид $w_1U_1wU_1w_2$. Множество симметричных путей, являющихся началом s, можно упорядочить по включению:

$$s_1 \sqsubseteq s_2 \sqsubseteq \dots \sqsubseteq s_k = s.$$

Если $v_{i_1}, v_{i_2}, \dots, v_{i_k}$ — передние образующие для пути s, то

$$F(s_j) = F(v_{i_1})F(v_{i_2})\dots F(v_{i_j}).$$

Значит, $F(s_{i+1}) - F(s_i) \sqsubseteq l_{\max}$ для любого i.

Рассмотрим минимальное i такое, что $w_1U_1 \sqsubseteq F(s_i)$. Реберная запись пути s_i имеет вид $s_i = v_1v_2\dots v_{k_1}$. Очевидно, $F(s_i) = w_1U_1w_1'$ для некоторого w_1'. При этом $|w_1'| \le l_{\max}$, иначе слово $F(s_{i-1})$, являющееся началом $F(s_i)$ длиной не менее $|F(s_i)| - l_{\max}$, содержало бы w_1U_1.

Аналогично, существует такой симметричный путь s', что его реберная запись имеет вид $v_{k_2}v_{k_2+1},\dots v_n$, а его слово $F(s') = w_2'U_1w_2$, где $|w_2'| \le l_{\max}$.

Итак, у нас есть пути $s_i = v_1v_2\dots v_{k_1}$ и $s' = v_{k_2}v_{k_2+1},\dots v_n$, являющиеся соответственно началом и концом пути $s = v_1v_2\dots v_n$. Докажем, что $k_1 < k_2$. Предположим противное: $k_1 \ge k_2$.

Введем **обозначения:** $A = B(v_1v_2\dots v_{k_2-1})$, $B = F(v_{k_2}\dots v_{k_1})$ (этот путь симметричен), $C = F(v_{k_1+1}v_{k_1+2},\dots v_n)$.

Тогда слова путей $v_1v_2\dots v_{k_1}$, $v_{k_2}v_{k_2+1},\dots v_n$ и $v_1v_2\dots v_n$ — это AB, BC и ABC соответственно. Следовательно, $F(s) < F(s_i)+F(s')$. С другой стороны,

$$F(s) = |w_1U_1wU_1w_2| = |w_1U_1w_1'| + |w_2'U_1w_2| + (|w| - |w_1'| - |w_2'|) \ge$$

$$\ge F(s_i) + F(s') + 8l_{\max}.$$

Мы получили противоречие. Стало быть, верно неравенство $k_2 > k_1$.

Так как $U \sqsubseteq F(s_i)$, то путь s_i проходит по всем ребрам S. Докажем, он проходит и по всем тройкам ребер вида $x_{i_1}vy_{i_2}$ для хороших пар (x_{i_1}, y_{i_2}). В самом деле: если (x_{i_1}, y_{i_2}) — хорошая пара ребер, то слово $X_{i_1}VY_{i_2}$ является подсловом U_1. С другой стороны, $X_{i_1}VY_{i_2}$ является словом симметричного пути, образованного конкатенацией естественного продолжения влево ребра x_{i_1}, ребра v и естественного продолжения вправо ребра y_{i_2}. Осталось воспользоваться свойством 4 схем Рози, примененным к схеме S.

То же самое верно и для пути s'.

Если симметричный путь проходит по тройке ребер $x_{i_1}vy_{i_2}$ для плохой пары (x_{i_1}, y_{i_2}), то слово этого пути содержит $X_{i_1}VY_{i_2}$ в качестве подслова и путь не является допустимым. Следовательно, путь s, как допустимый, не проходит по “плохим” тройкам ребер.

Замечание 3.4.1. *Конструкция пути $f(s)$ в графе S' понадобится далее.*

Лемма 3.4.1. *Граф S'' сильно связен и не цикл.*

Рассмотрим в графе S' путь $f(s)$. Этот симметричный путь проходит только по хорошим ребрам графа S'. В этом пути можно выделить начало $f(s_i)$, конец $f(s')$ и среднюю часть, причем начало и конец проходят по всем хорошим ребрам графа S'. Пусть d и e — ребра схемы S''. Им соответствуют пути в S'. Ребру d — путь $d_1d_2\dots d_{n_1}$, ребру e — путь $e_1e_2\dots e_{n_2}$. Путь $f(s_i)$ проходит по ребру d_1. Из концов каждого из ребер $d_1, d_2, \dots, d_{n_1-1}$ идет ровно одно хорошее ребро. Следовательно, в пути s после ребра d_1 обязаны идти ребра d_2,d_3,$\dots$, d_{n_1}. Аналогично, в s' есть вхождение ребра e_{n_2}, а перед вхождением e_{n_2} в $f(s)$ обязана идти последовательность ребер $e_1e_2\dots e_{n_2-1}$. Часть пути $f(s)$, начинающуюся с $d_1d_2\dots d_{n_1}$ и кончающуюся на $e_1e_2\dots e_{n_2}$, можно разбить на последовательные группы ребер, образующие ребра графа S''.

Таким образом, по ребрам графа S'' можно добраться от любого ребра до любого другого. Следовательно, граф S'' сильно связен.

Предположим, что S'' — цикл. Тогда из каждой вершины графа S' выходит ровно одно хорошее ребро. Пусть всего хороших ребер n. Рассмотрим какое-нибудь хорошее ребро, которое в схеме S' выходит из раздающей вершины, и назовем его d_1. Такие ребра существуют, например, сойдет ребро вида v_{ij}. В схеме S' хорошие ребра образуют цикл $d_1d_2\dots d_n$. Возьмем w_0 – произвольное подслово W. В S существует симметричный путь $v_1v_2\dots v_n$ такой, что $w_0 \sqsubseteq F(v_1v_2\dots v_n) \sqsubseteq W$. Соответствующий путь в S' должен проходит только по хорошим ребрам — следовательно, этот путь имеет вид $d_{\mathrm{begin}}(d_1d_2\dots d_e)^k d_{\mathrm{end}}$, где участки d_{begin} и d_{end} состоят менее чем из n ребер. Пусть переднее слово пути $d_1d_2\dots d_n$ — это D. Тогда w_0 является подсловом слова $D_{\mathrm{begin}}D^kD_{\mathrm{end}}$, где длина слов D_{begin} и D_{end} не превосходит $nl_{\max}$. Пусть N — длина слова D. Тогда любое подслово сверхслова W длины N является циклическим сдвигом D (если D' — подслово W длины N, выберем w_0, содержащее D в середине и имеющее длину $N + 10\cdot nl_{\max}$). Таким образом, W периодично. Противоречие, значит, S'' — не цикл.

Предложение 3.4.5. *Для S'' выполнено второе свойство схем Рози.*

В самом деле, ребра, выходящие из одной раздающей вершины в графе S'', соответствуют путям, выходящим из одной раздающей вершины графа S' и имеющим различные первые ребра. Стало быть, передние слова ребер в S'' —

передние слова некоторых путей, выходящих из одной раздающей вершины графа S' и имеющих различные первые ребра. Передние слова этих путей начинаются с передних слов соответствующих первых ребер, первые буквы которых, согласно предложению 3.4.4, различны.

Так как симметричные пути в S'' являются симметричными путями в S' с теми же словами, а для графа S' выполнены свойства 3 и 4, то эти же свойства выполнены и для графа S''.

Предложение 3.4.6. *Для графа со словами S'' выполнено свойство (5).*

Рассмотрим в графе S' симметричный путь $f(s)$, построенный ранее (см. замечание 3.4.1). Слово этого пути — подслово сверхслова W. Как доказывалось ранее, если ребру d графа S'' в S' соответствует путь $d_1d_2\dots d_k$, то $d_1d_2\dots d_k \sqsubseteq f(s)$. Из следствия 3.2.1 имеем $F(d_1d_2\dots d_k) \sqsubseteq F(f(s))$ а также $B(d_1d_2\dots d_k) \sqsubseteq F(f(s))$. Осталось заметить, что $F(f(s)) \sqsubseteq W$.

Предложение 3.4.7. *Для S'' выполнено свойство (7).*

Рассмотрим путь $f(s)$ (см. замечание 3.4.1). Докажем, что если l — симметричный путь в s и $F(f(s)) \sqsubseteq F(l)$, то для любого ребра $v \in S''$ выполняется $v \sqsubseteq l$. В самом деле: путь l является симметричным и в графе S'. По доказанному ранее свойству 4 для схемы S', путь l содержит путь $f(s)$. А стало быть, в графе S'' он проходит по всем ребрам.

Для доказательства теоремы **3.4.1** осталось проверить свойство 6. Пусть u — интересующее нас подслово сверхслова W. Каждому ребру графа S'' в S' соответствует путь, пусть все такие пути имеют длину не более N ребер, а слова всех ребер имеют в S длину не более $l_{\max}$. Рассмотрим w_0 – подслово в W вида u_1uu_2, где $|u_1| = |u_2| = 10 \cdot Nl_{\max}$. В S существует допустимый путь l, слово которого содержит w_0. Рассмотрим в S' путь $f(l)$. В этом пути содержится более $10 \cdot N$ ребер. Среди последних N ребер пути есть ребро, из конца которого выходит более одного хорошего ребра. Также среди первых N ребер пути есть ребро, в начало которого входят хотя бы два хороших ребра. Отрезок пути между этими двумя ребрами обозначим r. Путь r является симметричным, а его слово содержит слово u в качестве подслова. Кроме того, r является симметричным путем в схеме S''.

Итак, **теорема 3.4.1 доказана**.

3.5 Свойства слов, порожденных примитивными морфизмами.

Замечание 3.5.1. *Мы можем считать, что примитивный морфизм φ таков, что в $a_j \sqsubseteq \varphi(a_i)$ для любых букв a_i и a_j (иначе возьмем соответствующую степень этого морфизма).*

Определение 3.5.1. *Пословная сложность* $P(N)$ сверхслова W — количество различных подслов W длины N.

Определение 3.5.2. *Показатель рекуррентности* $P_2(N)$ для почти периодичного сверхслова W — минимальное число $P_2(N)$ такое, что в любом подслове сверхслова W длины $P_2(N)$ встретятся все подслова W длины N.

Лемма 3.5.1. *Для примитивного морфизма φ существуют такие $C_1 > 0$, $C_2 > 0$, $\lambda_0 > 1$, что $C_1\lambda_0^k < |\varphi^k(a_j)| < C_2\lambda_0^k$ для любых k и буквы a_j.*

Образы каждой буквы алфавита под действием морфизма φ являются словами, содержащими весь алфавит (см. замечание 3.5.1). Пусть $\varphi(a_i) = A_i$. Рассмотрим матрицу A размера $n \times n$ (где n - мощность алфавита $\{a_i\}$) такую, что A_j^i — количество букв a_j в A_i. Тогда k-й столбец матрицы A^N является вектором количеств вхождений букв алфавита в $\varphi^N(a_k)$.

Все элементы матрицы A положительны. По теореме Перрона-Фробениуса, у матрицы A есть λ_0 – положительное собственное значение, которое является собственным значением с строго максимальным модулем, причем у этого значения есть собственный вектор v_0 с положительными координатами. Тогда для любого вектора v существует константа $c(v)$ такая, что $A^k(v) = c(v)\lambda_0^k v + \bar{o}(k)$. Константа $c(v)$ – коэффициент при v_0 в разложении v по жорданову базису. При этом для любого ненулевого вектора с неотрицательными координатами $c(v) > 0$. Таким образом, существуют такие положительные константы C_1 и C_2, что для любого k и любой буквы a_j верно двойное неравенство: $C_1\lambda_0^k < |\varphi^k(a_j)| < C_2\lambda_0^k$.

Лемма 3.5.2. *Если W — чисто морфическое слово, порожденное примитивным морфизмом φ, то его показатель рекуррентности $P_2(N)$ ограничен сверху некой линейной функцией: $P_2(N) \leq C_3N$.*

Для любого k сверхслово W можно разбить на $\varphi^k(a_i)$ для различных a_i. Если u_1 — подслово сверхслова W — имеет длину не менее $2C_2\lambda_0^k$, то u содержит образ φ^k от некоторой буквы. Если u_2 – подслово W — имеет длину не более $C_1\lambda_0^k$, то u_2 содержится в образе $\varphi^k(a_ia_j)$ для некоторых букв a_i и a_j, встречающихся подряд в W.

Существует такое натуральное число m, что любая пара букв, встречающаяся подряд в W, встретится подряд в $\varphi^m(a_1)$. Следовательно, если для некоторого k длина слова u_1 не более $C_1\lambda_0^k$, а длина u_2 не менее $C_2\lambda_0^{k+m+1}$, то u_2 содержит u_1, так как u_2 содержит $\varphi^{m+k}(a_1)$ и, следовательно, содержит образ φ^k от любой встречающейся подряд в W пары букв.

Если отношение длин слов u_2 и u_1 более $\lambda_0^{m+2}\frac{C_2}{C_1}$, то такое k найдется.

Лемма 3.5.3. *Пусть W — почти периодичное сверхслово над алфавитом A с не более, чем линейным показателем рекуррентности $P_2(k) \leq C_3k$, а $\psi\colon A^* \to A^*$ — такой морфизм, что образом каждой буквы a_i является либо эта же буква a_i, либо пустое слово. Тогда, если слово $\psi(W)$ не является пустым, оно имеет не более, чем линейный показатель рекуррентности.*

В W должна быть буква, образ которой — не пустое слово. Пусть $\psi(a_1) = a_1$. Существует такое M, что среди любых M подряд идущих букв слова W встречается a_1. Тогда всякое подслово w сверхслова $\psi(W)$ является подсловом $\psi(u)$, где u — некоторое подслово W длины не более $M|w|$. Следовательно, любое подслово W длины $P_2(M|w|) \leq C_3M|w|$ содержит w.

Лемма 3.5.4. *Пусть W — почти периодичное сверхслово над алфавитом A с не более, чем линейным показателем рекуррентности $P_2(k)$, также путь B — конечный алфавит, а $\psi\colon A^* \to B^*$ – нестирающий морфизм, продолжающийся на W. Тогда слово $\psi(W)$ является почти периодичным с не более, чем линейным показателем рекуррентности.*

Пусть $|\psi(a_i)| \leq M$ для любой буквы a_i. Разобьем $\psi(W)$ на образы $\psi(a_i)$. Любое подслово u_1 длины k содержится в образе идущих подряд k букв. Если длина u_2 – подслова $\psi(W)$ — не менее, чем $M(F_2(k)+1)$, то u_2 содержит образ ψ от $F_2(k)$ идущих подряд букв сверхслова W, и, следовательно, содержит u_1.

Следствие 3.5.1. *Пусть W — почти периодичное сверхслово над алфавитом A с не более, чем линейным показателем рекуррентности, B – конечный алфавит, а $\psi\colon A^* \to B^*$ — произвольный морфизм, продолжающийся на W. Тогда $\psi(W)$ имеет не более чем линейный показатель рекуррентности.*

Лемма 3.5.5. *Если W — сверхслово с не более, чем линейным показателем рекуррентности, то пословная сложность W также не более чем линейна.*

Пусть $F_2(k) \leq C_3k$. Рассмотрим u — подслово W длиной C_3k. В нем должны присутствовать все подслова W длины k, но в u таких подслов не более, чем $C_3k - k + 1$.

Лемма 3.5.6. *Если для сверхслова W существует такая константа C, что для всех N выполнено $P(N) \leq CN$, то функция $P(N+1) - P(N)$ ограничена.*

Доказательство приведено в работе [70]

Лемма 3.5.7. *Если сверхслово $W = \psi(\varphi^\infty(a_1))$, где морфизм φ примитивный, то существует такая константа C_4, что во всех графах Рози $G_k(W)$ количество вершин суммарной степени более 2 не превосходит C_4.*

К сверхслову W применимы леммы 3.5.1 — 3.5.6.

Для каждого подслова W длины n в W существует хотя бы одно подслово длины $n+1$ с тем же началом. Если же слову в графе G_n соответствует вершина, у которой исходящая степень более одного, то в W есть хотя бы два подслова с тем же началом. Следовательно, для каждого n число $P(n+1) - P(n)$ не менее числа вершин G_n с исходящей степенью более единицы. Аналогично, $P(n+1) - P(n)$ не менее количества вершин входящей степени более единицы. Из того, что $P(n+1) - P(n)$ ограничено, следует утверждение леммы.

Лемма 3.5.8. *Если для почти периодичного непериодичного W показатель рекуррентности $P_2(k) \leq Ck$, то существует такое $C_5 > 0$, что если $u \sqsubseteq W$, то для любых двух различных вхождений u в W их левые концы находятся на расстоянии не меньшем, чем $C_5|u|$.*

Пусть в W есть два вхождения u длины k, а их левые концы находятся на расстоянии $l < \frac{k}{2C}$.

$$\rlap{$\overbrace{\phantom{a_1\dots a_l\, a_{l+1}a_{l+2}\dots a_{k'-1}a_{k'}}}^{u}$}a_1\dots a_l\,\underbrace{a_{l+1}a_{l+2}\dots a_{k'-1}a_{k'}\,a_{k'+1}\dots a_{k'+l}}_{u}$$

По очереди рассматривая оба вхождения слова u, приходим к выводу, что фрагмент $a_1 \dots a_l$ повторится подряд как минимум $[\frac{k_1}{l}]$ раз, что превышает C. Любое подслово W длины l является подсловом слова $(a_1 a_2 \dots a_l)^C$, что немедленно влечет периодичность W и противоречие.

3.6 Свойства схем Рози для слов, порожденных примитивными морфизмами.

В этом разделе $W = \psi(\varphi^\infty(a_1))$, где ψ – произвольный морфизм, φ — примитивный морфизм, W не является заключительно периодичным, а S — схема Рози для сверхслова W.

Определение 3.6.1. *Масштаб схемы* есть наименьшая из длин слов опорных ребер.

Возьмем собирающие вершины этой схемы. Для каждой из них возьмем ребро, выходящее из вершины, и его естественное продолжение вперед. Получится набор симметричных путей $\{s_i\}$ схемы S. Рассмотрим раздающие вершины схемы S. Для каждой из них возьмем естественное продолжение назад ребра, входящего в вершину. Получим набор симметричных путей $\{t_i\}$.

Лемма 3.6.1. *Существует такое C_6, что для любых двух путей s_1, s_2 из $\{s_i\}$, длины слов $F(s_1)$ и $F(s_2)$ отличаются не более, чем в C_6 раз.*

Так как $F(s_1)$ является передним словом первого ребра пути s_1, то $F(s_1) \sqsubseteq W$. Аналогично, $F(s_2) \sqsubseteq W$. Ровно одно из ребер пути s_1 входит в раздающую вершину. То же самое верно и для s_2. Следовательно, случаи $s_1 \sqsubseteq_2 s_2$ или $s_2 \sqsubseteq s_2 s_1$ невозможны. Значит, по свойству 4 схем Рози, в $F(s_1)$ может быть не более одного вхождения $F(s_2)$ и наоборот. Для слова W выполняется $P_2(n) \leq C_3 n$ для некоторого C_3 (см. лемму 3.5.2). Если бы длины слов $F(s_1)$ и $F(s_2)$ различались более, чем в $2C_3$ раз, то одно из слов (например, $F(s_1)$) можно было бы разбить на два слова, каждое из которых содержало бы $F(s_2)$.

Замечание 3.6.1. *Доказательство следующего утверждения полностью аналогично доказательству леммы 3.6.1: для любых путей из $\{t_i\}$ отношение длин их слов не превосходит C_6.*

Следствие 3.6.1. *Если M — масштаб схемы S, то для любого пути из $\{s_i\}$ или $\{t_i\}$ длина его слова не превосходит $C_6 M$.*

Лемма 3.6.2. *Для любого пути из $\{s_i\}$, его слово является в W специальным слева и оканчивается на некоторое биспециальное слово длины не менее M, где M — масштаб схемы.*

Пусть рассматриваемый путь имеет реберную запись $v_1 v_2 \dots v_n$. Тогда ребро v_n выходит из собирающей вершины, иначе естественное продолжение ребра v_1 было бы короче хотя бы на одно ребро. Ребро v_n является опорным.

Докажем, что слово любого опорного ребра v является биспециальным в W. Докажем, например, что $F(v)$ является специальным справа. Пусть из правого конца v выходят y_1 и y_2. По свойству 7 схем Рози для ребра y_1 есть u_{y_1} – подслово слова W такое, что любой симметричный путь, слово которого содержит u_{y_1}, проходит по ребру y_1. По лемме 3.2.3, в схеме S есть допустимый путь, слово которого содержит u_{y_1}. Реберная запись этого допустимого пути обязана содержать vy_1. Следовательно, в слово этого пути входит переднее слово пути vy_1, стало быть, $F(vy_1) \sqsubseteq W$.

Аналогично доказывается, что $F(vy_2) \sqsubseteq W$. Так как $F(vy_i) = F(v)F(y_i)$ и слова $F(y_i)$ имеют различные первые буквы по свойству 2 схем Рози, v является специальным справа. Специальность слева доказывается абсолютно так же.

Таким образом, $F(v_n)$ — биспециальное. То, что $|F(v_n)| \geq M$, очевидно.

Теперь докажем, что слово пути $v_1 v_2 \dots v_n$ является специальным слева. Пусть x_1 и x_2 входят в начало ребра v_1. Для ребра x_1 есть u_{x_1} — такое подслово W, что любой симметричный путь, слово которого содержит u_{x_1}, содержит x_1. В S существует допустимый путь l_{x_1}, слово которого содержит u_{x_1}. Этот путь проходит по ребру x_1. Следующие n ребер этого пути могут быть только ребра

$v_1, v_2, \ldots v_n$. Рассмотрим слово $F(l_{x_1})$: $B(x_1)F(v_1v_2\ldots v_n) \sqsubseteq B(l_{x_1})$. Значит, $B(x_1)F(s_1s_2\ldots s_n) \sqsubseteq W$.

Аналогично, $B(x_2)F(v_1v_2\ldots v_n) \sqsubseteq W$. Пользуясь свойством 2 схем Рози для $B(x_1)$ и $B(x_2)$, получаем, что $F(v_1v_2\ldots v_n)$ является специальным слева.

Аналогично доказывается соответствующий факт для путей $\{t_i\}$.

Лемма 3.6.3. *Существует константа C_7 такая, что в любой схеме Рози для W количество вершин не превосходит C_7.*

Достаточно показать, что для любой схемы S количество элементов в $\{s_i\}$ и $\{t_i\}$ ограничено константой, зависящей только от сверхслова W. Докажем для $\{s_i\}$, так как для $\{t_i\}$ аналогично.

Пусть M — масштаб схемы S. Рассмотрим для W граф Рози G_k, где $k = [\frac{M}{2}]$. Пусть $s_1 \in \{s_i\}$. Из 3.6.1, его слово $F(s_1)$ имеет длину, не превосходящую C_6M. Из леммы 3.6.2 следует, что последние k букв этого слова являются словом, специальным справа, а первые k букв – словом, специальным слева. При этом $F(s_1)$ является подсловом W. Все подслова в $F(s_1)$ длины $k+1$ образуют путь в G_k, ведущий из собирающей вершины в раздающую. Пусть в этом пути обнаружился цикл длины l.

Это значит, что в W есть подслово длины $k+l$, первые k букв которого образуют то же слово, что и последние k букв. Применяя лемму 3.5.8, получим $l \geq C_5k$.

Слову $f(s_1)$ соответствует путь по ребрам графа G_k из одной специальной вершины в другую, этот путь не может попадать в одну и ту же вершину G_k менее, чем через C_5k шагов, а его длина не превосходит C_6M. Значит, одну и ту же вершину графа G_k этот путь посетит не более $\frac{C_6M}{C_5k}$ раз.

Из леммы 3.5.7, в графе G_k может быть не более $C_4(W)$ специальных вершин. Тогда посещений специальных вершин у пути будет не более $\frac{C_6MC_4(W)}{C_5k}$. Если $|B|$ - мощность нашего алфавита, то всего различных путей, содержащих не более $\frac{C_6M}{C_5k}$ (с учетом количества вхождений) специальных вершин, выходящих из специальной вершины и входящих в специальную вершину, будет не более, чем

$$\sum_{i=1}^{\frac{C_6MC_4(W)}{C_5k}} C_4(W)B^i.$$

Для различных путей из $\{s_i\}$ их слова также различны, иначе, по свойству 4 схем Рози, пути бы совпадали. Значит, этим путям мы поставили в соответствие различные пути в G_k. Поэтому верна оценка

$$\#\{s_i\} \leq \sum_{i=1}^{\frac{C_6 M C_4(W)}{C_5 k}} C_4(W) B^i.$$

Следствие 3.6.2. *Количество различных облегченных нумерованных схем, возникающих при детерменированной эволюции, конечно для W.*

Лемма 3.6.4. *Существует такая константа C_8, что если S_1 и S_2 — нумерованные схемы Рози для W и $S_2 = \mathrm{Evol}(S_1)$, то масштаб схем S_1 и S_2 относятся не более, чем в C_8 раз.*

Для некоторого C_3 выполняется $P_2(k) \leq C_3 k$. Если v – опорное ребро схемы S_2, то ему соответствует симметричный путь s по ребрам схемы S_1, причем $F(s) = F(v_1)$. Путь s задается в S_1 номерами его ребер.

Номера ребер пути s однозначно определяются по следующему набору:

$\big\{$*Облегченная нумерованная схема S_1, множество пар плохих ребер в S_1(задаваемых номерами этих ребер), облегченная нумерованная схема S_2, номер ребра v в схеме S_2*$\big\}$.

Из следствия 3.6.2 вытекает, что таких наборов для слова W конечное число. Следовательно, для W существует такая константа C_9, что путь s не длиннее C_9 ребер (для всех S_1, S_2 и v). А значит, $F(s) \leq C_9 C_3 M$, где M — масштаб схемы S_1, иначе для некоторого опорного ребра v_0 графа S_1 будет $F(v_0) \sqsubseteq_{C_9+1} F(s)$ и, по свойству 4 схем Рози, $v_0 \sqsubseteq_{C_9+1} s$.

Неравенство $|F(v)| = |F(s)| \geq M$ очевидно.

Лемма 3.6.5. *Существует такая константа C_{10}, что для любой схемы S длины всех слов на ребрах этой схемы не превосходят $C_{10}M$, где M – масштаб схемы.*

Достаточно доказать существования такой константы для передних слов, для задних доказательство не будет отличаться. Для некоторого C_3 выполняется $P_2(k) \leq C_3 k$.

Для ребер, выходящих из собирающих вершин, существование такой константы следует из леммы 3.6.1. Рассмотрим произвольное ребро v, выходящее из раздающей вершины. По свойствам схем Рози, в S есть допустимый путь l_v, проходящий через v.

Рассмотрим l'_v — минимальный симметричный подпуть этого пути, проходящий через вершину v. Его реберная запись имеет вид

$$v_1 v_2 \dots v \dots v_n.$$

Так как l'_v — минимальный, то среди ребер $v_1 v_2 \dots v$ ровно одно выходит из собирающей вершины (а именно v_1). Также из минимальности следует, что среди ребер $v \dots v_n$ ровно одно входит в раздающую вершину (а именно v_n). Тогда l'_v проходит не более, чем по двум опорным ребрам. А так как его слово является подсловом W, то из свойства 4 для схем его длина $|F(l'_v)|$ не более, чем $2C_3M$. Осталось заметить, что $F(v) \sqsubseteq F(l'_v)$.

Лемма 3.6.6. *Существует такая константа C_{11}, что если S — схема Рози для W, а s_0 — допустимый путь, то $|F(s_0)| \geq C_{11}MN$, где M — масштаб схемы, а N — количество ребер пути.*

Пусть v_1 — первое ребро произвольного допустимого пути s. Слово $F(v_1)$ является началом $F(s)$. Естественное продолжение вправо ребра v_1 — симметричный путь, принадлежащий $\{s_i\}$. По лемме 3.6.2 $F(s_1)$ является специальным слева, а его длина не меньше, чем M. Аналогично доказывается, что $F(s)$ оканчивается на специальное справа слово длины не менее M.

Следовательно, слово $F(s)$ соответствует пути l по ребрам графа $G_{[\frac{M}{2}]}$, при этом l начинается в левой специальной вершине, а кончается в правой специальной вершине.

Рассмотрим путь s_0. Среди его $N-1$ промежуточных вершин найдутся либо половина раздающих, либо половина собирающих. Не умаляя общности, будем считать, что выполнено первое. Рассмотрим множество допустимых путей, являющихся началами s_0. Их хотя бы $\frac{N-1}{2}$. Их можно упорядочить так, чтобы слово каждого следующего пути являлось началом слова предыдущего.

$$F(s_0) \succneqq F(s_1) \succneqq \dots \succneqq F(s_k), k \geq \frac{N-1}{2}.$$

Словам $F(s_i)$ соответствуют пути l_i, начинающиеся в одной собирающей вершине, кончающиеся в раздающих вершинах и такие, что каждый следующий путь — начало предыдущего. В $G_{[\frac{M}{2}]}$ не более C_7 специальных вершин (см. лемму 3.5.7). Следовательно, через одну из раздающих вершин путь l_0 проходит не менее $\frac{N-1}{C_7}$ раз.

Из леммы 3.5.8 следует, что между последовательными посещениями одной и той же вершины графа $G_{[\frac{M}{2}]}$ в пути l_0 должно быть не менее $[\frac{M}{2}]C_5$ ребер. Значит, l_0 проходит не менее, чем по $[\frac{M}{2}]C_5\frac{N-1}{C_7}$ ребрам. Осталось сказать, что длина l_0 не более $|F(s_0)|$.

Лемма 3.6.7. *Существует такая константа C_{12}, что для любой схемы S и любого пути s длина слова этого пути $F(s)$ не превосходит $C_{12}NM$, где M — масштаб схемы, а N — количество ребер пути.*

Если симметричный путь проходит по N ребрам, то длина его слова не превосходит $N \cdot l_{\max}$, где $l_{\max}$ — максимальная из длин слов на ребрах. В силу леммы 3.6.5, $l_{\max} \leq C_{10}M$. Лемма доказана.

3.7 Построение оснасток.

В этом разделе $W = \psi(\varphi^\infty(a_1))$, где φ – примитивный морфизм, а W не является заключительно периодичным.

Определение 3.7.1. Пусть $A \sqsubseteq W$, S — схема Рози для W. Множество симметричных путей в S, слова которых являются подсловами A, обозначим $S(A)$. Если $S(A)$ не пусто, в этом множестве есть максимальный элемент относительно сравнения $\sqsubseteq$, который будем называть *нерасширяемым путем.* Очевидно, для каждого пути из $S(A)$ есть путь, который содержит его и является нерасширяемым. Назовем этот путь *максимальным расширением.*

Лемма 3.7.1. *Если s — нерасширяемый путь в $S(A)$, то*

$$|A| - 2l_{\max} \leq |F(s)| \leq |A|$$

, где $l_{\max}$ — максимальная из длин слов на ребрах S.

Неравенство $|F(s)| \leq |A|$ очевидно, так как $F(s) \sqsubseteq A$.

Пусть A имеет вид $u_1F(s)u_2$. Если $|F(s)| < |A| - 2l_{\max}$, то либо $|u_1| > l_{\max}$, либо $|u_2| > l_{\max}$. Не умаляя общности, предположим второе. По лемме 3.2.4 существует такой допустимый путь s_1, что началом его является путь s, а его слово имеет вид $F(s_1) = F(s)u_2u_3$. Пусть s имеет реберную запись $v_1v_2\dots v_n$, а s_2 — реберную запись $v_1v_2\dots v_nv_{n+1}\dots v_{n+k}$. Рассмотрим путь, образованный s и естественным продолжением вправо ребра v_{n+1}. Слово этого пути является подсловом A, следовательно, s не является нерасширяемым в $S(A)$.

Лемма 3.7.2. *Существует такая константа C_{13}, что для любой схемы Рози S слова W, слов X, Y, Z таких, что $XYZ \sqsubseteq W$ а также $\min\{|X|, |Y|, |Z|\} > C_{13}M$, где M — масштаб схемы S, и симметричного пути l в схеме S следующие два условия эквивалентны:*

1. *$l \in S(XYZ)$*
2. *Существует такой симметричный путь l', что $l \sqsubseteq l'$, l' разбивается на три части $l' = xyz$, где xy — нерасширяемый путь для $S(XY)$, y – нерасширяемый путь для $S(Y)$, yz — нерасширяемый путь для $S(YZ)$.*

Согласно лемме 3.5.8 существует такая константа C_5, что любое подслово сверхслова W длины n не может иметь два различных вхождения в подслово W длины, не превосходящей $n + C_5n$.

Сначала докажем, что при достаточно больших C_{13} 1 влечет 2.

Будем доказывать, что в качестве l' можно взять максимальное расширение l в $S(XYZ)$. По лемме 3.7.1, $F(l') \geq |XYZ| - 2l_{\max}$, а из леммы 3.6.5 следует $l_{\max} \leq C_{10}M$.

Рассмотрим в l' такое максимальное начало l_1, что $l_1 \in S(XY)$. Докажем, что l_1 — нерасширяемый в $S(XY)$. Пусть $F(l') = F(l_1)u_1$. Рассмотрим его вхождение в XYZ, как подслова u.

Так как l' — нерасширяемый в XYZ, то

$$XYZ = d_1F(l')d_2,$$

где $|d_1| \leq l_{\max}$, $|d_1| \leq l_{\max}$.

$$XYZ = d_1F(l_1)u_1d_2.$$

Есть два случая:

1. Пусть $|d_1F(l_1)| \leq |XY|$. В этом случае получаем $|d_1F(l_1)| \geq |XY| - l_{\max}$, иначе у l' есть начало l_1' такое, что $l_1 \sqsubseteq l_1'$ и $|F(l_1')| \leq |F(l_1)| + l_{\max}$, а стало быть, $F(l_1') \sqsubseteq XY$.

   Тогда $|F(l_1)| \geq |XY| - 2l_{\max}$. Предположим, что l_1 расширяемый в XY. Тогда для некоторого l_1'' выполняется $l_1 \sqsubseteq l_1''$ и $F(l_1'') \sqsubseteq XY$. Можно считать, что l_1 — конец или начало l_1''.

   (a) $l_1'' = l_b l_1$. Тогда $XY = e_1 B(l_b) F(l_1) e_2$. Если выполнено

   $$2C_{10}M < C_5(2C_{13}M - 2C_{10}M),$$

   то у $F(l_1)$ только одно вхождение в XY, то есть $d_1 = e_1 B(l_b)$ и

   $$XYZ = e_1 B(l_b) F(l') d_2.$$

   Тогда симметричный путь $l_b l' \in S(XYZ)$ и путь l' расширяемый.

   (b) $l_1'' = l_1 l_e$. Тогда $XY = e_1 F(l_1) F(l_e) e_2$. Если выполнено

   $$2C_{10}M < C_5(2C_{13}M - 2C_{10}M),$$

   то у $F(l_1)$ только одно вхождение в XY, то есть $e_1 = d_1$ и

   $$XY = d_1 F(l_1) F(l_e) e_2.$$

   Следовательно, $F(l_1 l_e) \sqsubseteq XY$. Кроме того, слово $F(l_1 l_e)$ является началом $F(l')$. Тогда, по второму свойству схем Рози, путь $l_1 l_e$ является началом пути l'. Противоречие.

2. Пусть $|d_1F(l_1)| > |XY|$. В этом случае $d_1F(l_1) = XYd_3$, где $|d_3| < |d_1| \leq l_{\max}$. Так как

   $$F(l_1) \sqsubseteq_2 XYd_3 \sqsubseteq W,$$

   $|XYd_3| < |XY| + C_{10}M$, $|F(l_1)| > |XY| - C_{10}M$, то при

   $$2C_{10}M < C_5(2C_{13}M - C_{10}M)$$

   получаем противоречие.

Аналогично, если l_2 — такой максимальный конец пути l', что $l_2 \in S(YZ)$, то l_2 — нерасширяемый в $S(YZ)$.

Докажем следующий факт: *пути l_1 и l_2 внутри l' перекрываются, то есть в реберных записях l_1 и l_2 суммарно больше ребер, чем в реберной записи пути l'.*

В самом деле, иначе возьмем s — естественное продолжение влево последнего ребра l_1. Тогда $|F(l')| + |F(s)| \geq F(l_1) + F(l_2)$, то есть

$$|XYZ| \geq (|XY| - 2C_{10}M) + (|YZ| - 2C_{10}M) - C_{10}M,$$

что не выполняется при достаточно больших C_{13}.

Итак, можно считать, что $l_1 = xy$, $l_2 = yz$, $l' = xyz$. Для доказательства импликации осталось показать, что путь y – нерасширяемый путь в $S(Y)$.

Имеем $XYZ = d_1B(x)F(y)F(z)d_2$, $XY = d_1B(x)F(y)e_1$,

и $YZ = e_2F(y)F(z)d_2$, где

$$\max\{|d_1|, |d_2|, |e_1|, |e_2|\} \leq l_{\max}.$$

Отсюда

$$|Y| = |e_1| + |e_2| + |F(y)|,$$

то есть $Y = e_2F(y)e_1$.

Если y расширяемый в $S(Y)$, то, не умаляя общности, y – начало пути yl_e такого, что $F(y)F(l_e) \sqsubseteq Y$, то есть $Y = f_1F(y)F(l_e)f_2$. Если выполняется неравенство

$$2MC_{10} < C_5(C_{13}M - 2MC_{10}),$$

то $F(y)$ имеет ровно одно вхождение в Y, стало быть, $e_2 = f_1$ и $e_1 = F(l_e)f_2$.

Тогда $XY = d_1B(x)F(y)F(l_e)f_2$, то есть $l_1l_e \sqsubseteq S(XY)$. Кроме того, $F(l_1l_e)$ является началом слова $F(l_1)e_1 = B(x)F(y)e_1$, которое, в свою очередь, является началом $B(x)F(Y)F(z) = F(l')$. Следовательно, l_1l_e — начало l'. Противоречие.

Теперь докажем, что при достаточно больших C_{13} свойство (2) влечет свойство (1).

Так как xy — нерасширяемый путь в $S(XY)$, а y — нерасширяемый путь в Y, XY имеет вид $XY = d_1B(x)F(y)d_2$, а Y имеет вид $Y = e_1F(y)e_2$, причем $|d_i|, |e_i| \leq l_{\max}$. Если $C_{10}M < C_5(C_{13}M - C_{10}M)$, то $d_2 = e_2$.

Аналогично, $YZ = f_1 F(y) F(z) f_2$, где $f_1 = e_1$. Тогда

$$XYZ = d_1 B(x) F(y) F(z) d_2,$$

а значит,

$$F(l) \sqsubseteq F(l') = F(xyz) \sqsubseteq XYZ$$

.

Определение 3.7.2. *Набор проверочных слов порядка k* — это набор слов $\{q_i\}$, в который входят $\psi(\varphi^k(a_i))$ для всех букв алфавита $\{a_i\}$, а также $\psi(\varphi^k(a_i a_j))$ для всевозможных пар последовательных букв слова $\varphi^\infty(a_1)$.

Кроме того, для любого k в наборе проверочных слов порядка k одно и то же число подслов.

Предложение 3.7.1. *Существуют такие положительные C_{14}, C_{15} и $\lambda_0 > 1$, что в наборе проверочных слов с номером k длина любого слова q удовлетворяет двойному неравенству $C_{14}\lambda_0^k < |q| < C_{15}\lambda_0^k$.*

Существуют константы k_1 и k_2 такие, что, с одной стороны, в $\varphi^k(a_i)$ среди любых k_1 идущих подряд букв есть буква, которая не удаляется при действии ψ, а с другой стороны, образ ψ от каждой буквы — это слово не длиннее k_2. Осталось применить лемму 3.5.1.

Определение 3.7.3. *Оснастка* (S, k) порядка k определяется для пронумерованной схемы Рози S. Для получения оснастки берется набор проверочных слов порядка k, для каждого проверочного слова v_i рассматривается набор $S(q_i)$. Каждый путь в схеме задается упорядоченным набором чисел — номеров ребер схемы. Таким образом, $S(q_i)$ задается множеством таких упорядоченных наборов, а оснастка получается, если взять такие наборы для всех проверочных слов и облегченную нумерованную схему S.

Определение 3.7.4. *Размер оснастки* — максимальная длина (в ребрах) по всем путям из $S(q_i)$ для всех q_i — проверочных слов порядка k.

Лемма 3.7.3. *Существуют такие положительные C_{16}, C_{17} и C_{18}, что для любой схемы S размер оснастки (S, k) заключен между $C_{16}\frac{\lambda_0^k}{M} - C_{17}$ и $C_{18}\frac{\lambda_0^k}{M}$.*

Длина каждого проверочного слова, согласно 3.7.1, хотя бы $C_{14}\lambda_0^k$. Если q_i — проверочное, то длина слова нерасширяемого в $S(q_i)$ пути не менее $C_{14}\lambda_0^k - -2C_{10}M$. А длина этого пути в ребрах составляет, согласно лемме 3.7.1, не менее $\frac{C_{14}\lambda_0^k-2C_{10}M}{C_{12}M}$.

С другой стороны, длина слова нерасширяемого пути не может быть больше $C_{15}\lambda_0^k$, а его длина в ребрах, как следует из леммы 3.6.6, не может быть более $\frac{C_{15}\lambda_0^k}{C_{11}M}$.

Следствие 3.7.1. *Существуют такие C_{19}, C_{20} и C_{21}, что если размер оснастки (S, k) больше C_{19}, то*

1. *Длины всех проверочных слов составляют не менее $C_{13}M$.*
2. *Для любой хорошей тройки ребер существует проходящий через нее путь, слово которого содержится во всех проверочных словах.*
3. *Размер оснастки (S, k) относится к размеру оснастки $(\mathrm{Evol}(S), k)$ не более, чем в C_{20} раз.*
4. *Размеры оснасток (S, k) и $(S, k + 1)$ отличаются не более, чем в C_{20} раз.*
5. *Размер оснастки $(S, k + C_{21})$ больше размера оснастки (S, k) хотя бы в два раза.*

1. Из лемм 3.7.3 и 3.7.1 следует, что если x – размер оснастки, а $|q|$ — размер проверочного слова q, то $|q| > C_{14}\lambda_0^k > C_{14}(C_{18}xM)$.
2. Симметричный путь, получающийся естественным расширением хорошей тройки ребер сначала вправо, а потом влево, является допустимым и его слово имеет длину не более $3l_{\max} \leq 3C_{10}M$. Если проверочное слово q имеет длину хотя бы $3C_{10}MC_5$, где C_5 — показатель рекуррентности, то подсловом q является и слово рассматриваемого пути. (Как видно из предыдущего пункта, выбором C_{19} этого легко добиться.)
3. Пусть M и M_1 — масштабы схемы S и $\mathrm{Evol}(S)$, а x и x_1 — размеры оснасток. Согласно лемме 3.6.4, $M_1 \leq C_8M$. Тогда

$$x_1 > C_{16}\frac{\lambda_0^k}{C_8M} - C_{17} > C_{16}\frac{C_{18}x}{C_8} - C_{17}$$

, что при достаточно большом x больше, чем $C_{16}\frac{C_{18}x}{2C_8}$.

4. Пусть x и x_1 — размеры оснасток (S,k) и $(S,k+1)$.
   Тогда $C_{16}\frac{\lambda_0^{k+1}}{M} - C_{17} < x_1 < C_{18}\frac{\lambda_0^{k+1}}{M}$.
   С другой стороны, $\frac{x+C_{17}}{C_{16}} > \frac{\lambda_0^k}{M} > \frac{x}{C_{18}}$.
   Таким образом, $C_{16}\frac{\lambda_0 x}{C_{18}M} - C_{17} < x_1 < C_{18}\frac{\lambda_0(x+C_{17})}{C_{16}M}$.

5. Пусть x и x_1 — размеры оснасток (S,k) и $(S,k+C_{21})$.
   Тогда $C_{16}\frac{\lambda_0^{k+C_{21}}}{M} - C_{17} < x_1$. С другой стороны, $\frac{\lambda_0^k}{M} > \frac{x}{C_{18}}$.
   Таким образом, $x_1 > C_{16}\frac{\lambda_0^{C_{21}}x}{C_{18}M} - C_{17}$, что не меньше $2x$ при достаточно большом C_{21}.

Лемма 3.7.4. *Если размер оснастки (S,k) хотя бы C_{19}, то оснастка $(S,k+1)$ является функцией оснастки (S,k).*

Рассмотрим проверочное слово порядка $k+1$.

Оно имеет вид $\psi(\varphi^k(\varphi(a)))$, где a — либо одна, либо две буквы. В любом случае, $\varphi(a)$ - подслово φ^∞. Таким образом, проверочное слово порядка $k+1$ разбивается на блоки — проверочные слова порядка k, причем любые два последовательных блока образуют проверочное слово порядка k. (заметим, что то, на какие типы слов порядка k разбивается $\psi(\varphi^{k+1}(a))$, не зависит от номера k. *Типом слова* $\psi(\varphi^k(a_i))$ называется буква a_i.)

Пусть $A_1A_2\dots A_m$ — проверочное слово порядка $k+1$, A_i – проверочные слова порядка k. Согласно 3.7.1, длины слов A_i удовлетворяют условию леммы 3.7.2. Применим эту лемму к словам A_1,A_2,A_3. Так как свойство пути быть *нерасширяемым* определяется по оснастке, то по оснастке можно определить, какие пути принадлежат $S(A_1A_2A_3)$ (а точнее, какие пути в соответствующей облегченной пронумерованной схеме). После этого применим лемму 3.7.2 к словам A_1A_2, A_3, A_4 и определим, какие пути принадлежат $S(A_1A_2A_3A_4)$. Потом определим, какие пути принадлежат $S(A_1A_2A_3A_4A_5)$, и так далее, наконец определим, какие пути схемы S принадлежат $S(A_1A_2\dots A_m)$. Делая так для всех проверочных слов порядка $k+1$, получим что хотели.

Лемма 3.7.5. *Если размер оснастки (S,k) не менее C_{19}, то по оснастке (S,k) однозначно определяются плохие пары ребер и оснастка* $(\mathrm{Evol}(S),k)$.

Из следствия 3.7.1 очевидно, что по оснастке определяется множество хороших и плохих тройки ребер: по хорошим тройкам ребер пути из оснастки проходят, а по плохим — нет (так как пути в оснастках допустимые). Следовательно, определяется и нумерованная облегченная схема $\mathrm{Evol}(S)$. Симметричные пути в $\mathrm{Evol}(S)$ соответствуют некоторым симметричным путям в S с теми же словами. При этом мы можем указать соответствие, глядя лишь на облегченные схемы. Таким образом, для каждого симметричного пути в облегченной нумерованной схеме $\mathrm{Evol}(S)$ по оснастке (S, k) можно определить, каким из $\mathrm{Evol}(S)(A_i)$ этот путь принадлежит. (Здесь A_i — проверочные слова порядка k).

Симметричным путям в $\mathrm{Evol}(S)$ соответствуют симметричные пути в S с теми же словами, при этом в $\mathrm{Evol}(S)$ пути не длиннее, чем соответственные пути в S. Стало быть, размер оснастки $(\mathrm{Evol}(S), k)$ не более, чем размер оснастки (S, k).

Завершение доказательства теоремы 3.4.2. Построим последовательность оснасток по следующему правилу. Возьмем $T = 2C_{19}C_{20}^{C_{21}}$. Построим первую оснастку для такого порядка проверочных слов, чтобы размер оснастки был хотя бы T. Каждая следующая оснастка определяется по предыдущей. Если ее размер менее T, то по оснастке вида (S, k) строится оснастка вида $(S, k+1)$. Иначе делается операция перехода от оснастки вида (S, k) к оснастке вида $((\mathrm{Evol}(S)), k)$. Таким образом, размер оснастки никогда не упадет ниже C_{19}, значит, каждый шаг определен однозначно. С другой стороны, размер оснастки не поднимается выше $C_{20}T$. Значит, различных оснасток в последовательности конечное число (так как различных нумерованных облегченных схем тоже конечно.) То есть с некоторого момента оснастки повторяются периодично. При этом не может быть так, что на каждом шаге повышается порядок проверочных слов. Значит, время от времени в последовательности появляются оснастки с новыми схемами. А значит, и протокол эволюции с некоторого места периодичен.

3.8 Переход к произвольным почти периодичным подстановочным словам

Пусть φ — произвольный морфизм из A^* в A^*, h — морфизм из A^* в B^*.

Лемма 3.8.1. *Если непериодичное слово $W = h(\varphi^\infty(a))$ является почти периодичным, то для некоторого алфавита D существуют такие морфизмы $\rho : D^* \to D^*$ и $g : D^* \to B^*$, где ψ – примитивный морфизм, что множества подслов W и $g(\rho^\infty(d))$ для некоторой буквы $d \in D$ совпадают.*

Для доказательства этой леммы нам потребуются следующие определения. Слово $w \in A^*$ будем называть *φ—ограниченным*, если последовательность

$$w, \varphi(w), \varphi^2(w), \varphi^3(w), \dots$$

периодична начиная с некоторого момента. В противном случае, $|\varphi^n(w)| \to \infty$ при $n \to \infty$ и слово w называется *φ-растущим.* Очевидно, слово является φ-ограниченным тогда и только тогда, когда оно состоит из φ-ограниченных букв.

Назовем слово w *$\varphi - h$-уничтожаемым*, если $h(\varphi^n(w)) = \Lambda$ для любого $n \geq 0$. Легко видеть, что слово является $\varphi{-}h{-}$уничтожаемым тогда и только тогда, когда оно состоит из $\varphi{-}h{-}$уничтожаемых букв. Далее, можно считать, что в алфавите A нет $\varphi - h-$уничтожаемых букв:

Предложение 3.8.1. *Пусть A' — множество не $\varphi - h-$уничтожаемых букв. Рассмотрим морфизм $\varphi' : A'^* \to A'^*$, определяемый как $\varphi'(a_i) =$ "$\varphi(a_i)$ с вычеркнутыми $\varphi - h-$уничтожаемыми буквами". Тогда слова $h(\varphi^\infty(a))$ и $h(\varphi'^\infty(a))$ совпадают.*

Доказательство очевидно.

Лемма 3.8.2 (A. Ehrenfeucht, G. Rozenberg)**.** *Следующие три условия эквивалентны:*

1. *Слово $\varphi^\infty(a)$ не является почти периодичным.*

2. *В $\varphi^\infty(a)$ есть бесконечно много $\varphi-$ограниченных подслов.*

3. *Существует непустое* $w \in A^*$ *такое, что* w^n *является подсловом* $\varphi^\infty(a)$ *для любого* n.

Доказательство импликации $1 \to 2$ можно найти в работе [107], импликации $2 \to 1$ — в работе [82], а импликация $3 \to 1$ очевидна.

Предположим, что слово $\varphi^\infty(a)$ не является почти периодичным. Тогда некоторое w встречается в нем сколь угодно много раз подряд. Так как w не является $\varphi - h$-уничтожаемым (см. 3.8.1), то для некоторого k слово $h(\varphi^k(w))$ не является пустым словом. Тогда слово $h(\varphi^k(w))$ встречается в W сколь угодно много раз подряд, и из почти периодичности W следует его периодичность.

Таким образом, мы можем считать, что $\varphi^\infty(a)$ почти периодично и, следовательно, в нем лишь конечное число φ—ограниченных слов.

Пусть I_φ — множество всех φ-растущих букв, B_φ — множество φ-ограниченных слов (включая пустое). В силу 3.8.2, B_φ конечно. Рассмотрим (конечный) алфавит C, состоящий из символов $[twt']$, где t и t' буквы из I_φ, а w — слово из B_φ и слово twt' является подсловом $\varphi^\infty(a)$.

Определим морфизм $\psi : C^* \to C^*$ следующим образом:

$$\psi([twt']) = [t_1 w t_2][t_2 w t_3] \dots [t_k w_k t_{k+1}],$$

где $\varphi(tw) = w_0 t_1 w_1 t_2 \dots t_k w'_k$, слово $\varphi(t')$ начинается с $w''_k t_{k+1}$ и $w_k = w'_k w''_k$ (слова w_i, w'_k и w''_k принадлежат B_φ).

Также определим $f : C^* \to C^*$ по правилу

$$f([twt']) = h(tw).$$

Предложение 3.8.2. *Все буквы алфавита* C *являются* ψ*-растущими и ни одна не является* $\psi - f-$*уничтожаемой.*

Очевидно, в образе $\varphi(t)$ от произвольной буквы $t \in I_\varphi$ содержится хотя бы одна буква из I_φ. Более того, в слове $\varphi^n(t)$ для некоторого n содержатся хотя бы две буквы из I_φ, иначе $\varphi^n(t) = w_n t_{i_n} v_n$, где w_n и v_n принадлежат B_φ. Таких слов, согласно 3.8.2, конечное число. Стало быть, буква t не является φ-растущей.

Далее, если в $\varphi^n(t)$ встречается k букв из I_φ, то длина $|\psi^n([twt'])|$ не меньше, чем k. Следовательно, буква $[twt']$ алфавита C является ψ-растущей.

Предположим, что буква $[twt']$ является $\psi - h-$уничтожаемой. Тогда в некоторая степень морфизма ψ от этой буквы является словом из C^* длины не менее 2. Значит, для некоторых t_i, w_i слово вида $[t_1w_1t_2][t_2w_2t_3]$ является $\psi - f-$уничтожаемым. Буква t_2 не является $\varphi - h-$уничтожаемой. Значит, существует последовательность букв $\{a_i\}$ такая, что $t_2 = a_0$, a_{i+1} содержится в $\varphi(a_i)$ и $h(a_k) \neq \Lambda$.

Докажем по индукции следующий факт: $\psi^n([t_1w_1t_2][t_2w_2t_3])$ содержит подслово вида $[t_{i_n}w_{i_n}t_{j_n}][t_{j_n}w_{j_n}t_{k_n}]$ такое, что в w_{n_1}, t_{n_2} или в w_{n_2} содержится a_n. пусть это верно для n, докажем для $n+1$. Возможны следующие случаи:

1. Пусть a_n содержится в w_{i_n}. Тогда $\varphi(t_{i_n}w_{i_n})$ содержит подслово вида $t_{i_{n+1}}w_{i_{n+1}}$, где a_{n+1} содержится в $w_{i_{n+1}}$. Тогда $\psi([t_{i_n}w_{i_n}t_{j_n}])$ содержит букву вида $[t_{i_{n+1}}w_{i_{n+1}}t_{j_{n+1}}]$, где $w_{i_{n+1}}$ содержит a_{n+1}. Так как у буквы $[t_{i_{n+1}}w_{i_{n+1}}t_{j_{n+1}}]$ в $\psi^{n+1}([t_1w_1t_2][t_2w_2t_3])$ есть левый либо правый сосед, искомое подслово встретится.

2. Если a_n содержится в w_{j_n}, то доказательство полностью аналогично.

3. Пусть $a_n = t_{j_n}$, $\varphi(t_{j_n}) = w_0t_1w_1t_2\ldots$ Если a_{n+1} встречается в w_0, то a_{n+1} содержится в “средней части” последней буквы слова $\psi([t_{i_n}w_{i_n}t_{j_n}])$. Иначе a_{n+1} содержится как “первая буква” или в “средней части” одной из букв слова $[t_{j_n}w_{j_n}t_{k_n}]$.

Таким образом, в $\psi^k([t_1w_1t_2][t_2w_2t_3])$ найдется буква, которая не обнуляется морфизмом f.

Пусть $\varphi^\infty(a)$ имеет вид $a_1w_1a_2\ldots$, где $a_1, a_2 \in I_\varphi$, $w_1 \in B_\varphi$. Тогда, как легко заметить, для любого n слово $h(\varphi^n(a))$ является началом слова $f(\psi^n([a_1w_1a_2]))$, следовательно, слова W и $f(\psi^\infty([a_1w_1a_2]))$ совпадают.

Замечание 3.8.1. *Конструкция морфизма ψ была рассмотрена в работе [107].*

Рассмотрим ориентированный граф G_φ, вершинами которого являются буквы алфавита C, и из c_i ведет стрелка в c_j тогда и только тогда, когда c_j содержится в c_i. Пусть D – сильносвязная компонента этого графа, до которой можно дойти по стрелочкам из $[a_1w_1a_2]$. Тогда ψ можно рассматривать как морфизм из D^* в D^*. Существует буква $d \in D$ и натуральное k такое, что

$\psi^k(d)$ начинается с d. Обозначим $\rho = \psi(k)$. Слово $\rho^\infty(d)$ является словом, все конечные подслова которого являются подсловами $\psi^\infty([a_1 w_1 a_2])$. Стало быть, все конечные подслова $f(\rho^\infty(d))$ являются подсловами W. Очевидно, что ρ является примитивным.

При этом образ при f хотя бы одной буквы из D не равен пустому слову. Стало быть, $f(\rho^\infty(d))$ бесконечное. А так как W почти периодичное, то все его конечные подслова являются конечными подсловами $f(\rho^\infty(d))$. В самом деле, пусть w — подслово W. Тогда для некоторого n любое подслово W длины n содержит w. Но в W есть такое подслово, являющееся подсловом $f(\rho^\infty(d))$.

Это рассуждение завершает доказательство леммы 3.8.1.

Следствие 3.8.1. *Пусть W — морфическое почти периодичное бесконечное слово. Тогда эволюция схем Рози для этого слова периодична с предпериодом.*

Это утверждение вытекает из теоремы 3.4.2, леммы 3.8.1 и того соображения, что эволюция схем Рози однозначно задается множеством подслов сверхслова.

Глава 4

Периодичность морфических последовательностей

В этой главе доказывается разрешимость следующей задачи:

Дано: *два алфавита A и B, подстановка $\varphi : A^* \to A^*$, продолжающаяся над буквой a и морфизм $h : A^* \to B^*$.*

Определить: *существуют ли такие конечные слова u и v, что $h(\varphi^\infty(a)) = uv^\infty$ (т.е. является ли сверхслово заключительно периодичным.)*

Далее мы будем называть эту задачу *проблемой заключительной периодичности.*

Теорема 4.0.1. *Проблема заключительной периодичности разрешима.*

В разделе 4.1 производится сведение общей задачи к случаю примитивного сверхслова, дальнейшие разделы (4.2 — 4.6) посвящены как раз примитивному случаю.

В разделе существенно используются определения и результаты предыдущего раздела, связанные со *схемами Рози.* Схемы Рози, используемые здесь, немного отличаются, поэтому в разделах 4.2 — 4.6 все определения, связанные со схемами Рози, даются заново в уточненном виде.

4.1 Сведение общего случая к примитивному.

Сперва воспользуемся следующим классическим результатом:

Теорема 4.1.1 (см. [60], глава 7). *Если $f : A^* \to B^*$ и $g : A^* \to A^*$ — произвольные морфизмы и $f(g^\infty(a_1))$ – бесконечное слово, то можно найти такие алфавит A', букву $a'_1 \in A'$, нестирающую подстановку φ, действующую на алфавите A и кодирование $\tau : A' \to B$ такие, что $f(g^\infty(a_1)) = \tau(\varphi^\infty(a'_1))$.*

Далее считаем φ нестирающей подстановкой, а h — кодированием.

Пусть имеется оракул, который для примитивной подстановки g, буквы a и морфизма f (то есть для тройки $\{g, a, f\}$) говорит, является ли $f(g^\infty(a))$ периодичным и если является, какой у него период.

Цель данного раздела — для произвольных нестирающего морфизма φ, кодирования h и буквы a_1 научиться отвечать на вопрос, является ли сверхслово $h(\varphi^\infty(a_1))$ периодичным, используя этот оракул.

Теорема 4.1.2. *Пусть φ — подстановка, действующая на алфавите A, h — морфизм из A^* в B^*, u — конечное слово из B^*. Тогда существует алгоритм, проверяющий, встречается ли u в слове $W = h(\varphi^\infty(a_1))$ и, если встречается, конечное ли число раз.*

Доказательство. Пусть $|u| = n$. Очевидно, можно считать, что все буквы алфавита A встретятся в $\varphi^\infty(a_1)$. Также можно считать, что все буквы алфавита встречаются в $\varphi(a_1)$ (иначе возьмем нужную степень морфизма). Если $a \in A$, то $\chi_n(a)$ — это число вхождений слова u в $h(\varphi^k(a))$, $l_k(a)$ и $r_k(a)$ — это два слова длины n, являющиеся, соответственно, началом и коцом $h(\varphi^k(a))$. Если $|h(\varphi^k(a))| < n$, то $l_k(a) = r_k(a) = h(\varphi^k(a))$.

Обозначим Ω_k набор из $2|A|$ слов: $r_k(a_i)$ и $l_k(a_i)$ для всех $a_i \in A$. Числа $\chi_k(a_i)$ образуют вектор χ_k.

Если $\varphi(a_i) = a_{i_1} a_{i_2} \ldots a_{i_m}$, то

$$h(\varphi^{k+1}(a_i)) = h(\varphi^k(a_{i_1}))h(\varphi^k(a_{i_2})) \ldots h(\varphi^k(a_{i_m})).$$

Слово l_{k+1} получается, если взять первые n букв слова $l_k(a_{i_1})l_k(a_{i_2}) \ldots l_k(a_{i_k})$. Слово r_{k+1} получается, если взять последние n букв слова

$$r_k(a_{i_1})r_k(a_{i_2}) \ldots r_k(a_{i_k}).$$

Стало быть, набор Ω_{k+1} однозначно определяется по набору Ω_k. Так как различных наборов не больше, чем $|B|^{2|A|n}$, то последовательность Ω_k зациклится.

Слово $h(\varphi^{k+1}(a_i))$ делится на блоки $h(\varphi^k(a_{i_k}))$, и вхождения u в $h(\varphi^{k+1}(a_i))$ бывают двух типов: те, которые целиком лежат в каком-либо блоке и те, которые принадлежат хотя бы двум блокам. Количество первых равно $\chi_k(a_{i_1}) + \chi_k(a_{i_2}) + \cdots + \chi_k(a_{i_m})$. Это число равно $i-$й компоненте вектора $M\chi_k$, где M — это матрица подстановки φ.

Количество вхождений, принадлежащих хотя бы двум блокам, можно вычислить по Ω_k. Таким образом, $\chi_{k+1} = A\chi_k + f(\Omega_k)$.

Пусть последовательность наборов Ω_k периодична начиная с $k = k_0$, и длина периода равна T. Тогда получим

$$\chi_{k_0+(k+1)T} = M(M(\ldots(M\chi_{k_0+kT} + f(\Omega_{k_0+kT})) + \ldots))+$$

$$+f(\Omega_{k_0+(k+1)T-2})) + f(\Omega_{k_0+(k+1)T-1}).$$

Раскроем все скобки: $\chi_{k_0+(k+1)T} = M^T\chi_{k_0+kT} + C$ для некоторого вектора C с неотрицательными компонентами.

Вспомним, что нас интересует поведение первой компоненты вектора χ_k при k, стремящемся к бесконечности. Очевидно, $\chi_k(a_i) = 0$ тогда и только тогда, когда $|\chi_k| = 0$, то есть u не является подсоловом W тогда и только тогда, когда $C = 0$ и $\chi_{t_0} = 0$.

Далее надо узнать, стремится ли к бесконечности $\chi_k(a_1)$. Пусть $C \neq 0$. Тогда $\chi_{k_0+kT} \geq C + MC + M^2C + \ldots M^kC$ (здесь сравнение покоординатное.) Но у матриц M^k первые столбцы состоит из целых положительных чисел, следовательно, $\chi_{k_0+kT}(a_1) \geq k$ и u встречается в W бесконечно много раз.

Пусть $C = 0$. Слово u встречается в W бесконечное число раз тогда и только тогда, когда существует такое $i \leq |A|$, что a_i встречается в $\varphi^\infty(a_1)$ бесконечно много раз и $\chi_{k_0}(a_i) > 0$.

Если a_1 встречается в $\varphi(a_1)$ хотя бы два раза, то все буквы встречаются в $\varphi^\infty(a_1)$ бесконечно много раз. Пусть a_1 встречается в $\varphi(a_1)$ один раз. Рассмотрим ориентированный граф G_φ, у которого $|A|$ вершин — буквы алфавита A и из вершины, соответствующей a_i, в a_j ведет ребро тогда и только тогда, когда $a_j \in \varphi(a_i)$. В G_φ есть петля на вершине a_1, удалим ее. Нетрудно понять, что a_i встречается в W бесконечное число раз тогда и только тогда, когда в полученном графе существует самопересекающийся путь из a_1 в a_i. А это свойство графа легко проверяется.

$\square$

Следствие 4.1.1. *Таким образом, можно определить, верно ли, что сверхслово заключительно периодично с данным пириодом u, то есть верно ли, что все слова длины $|u|$, которые бесконечно много раз встречаются в W, являются циклическими сдвигами u.*

Далее φ — нестирающий морфизм из A^* в A^*, h — кодирование из A в B, $W = h(\varphi^\infty(a_1))$.

Слово $w \in A^*$ будем называть *φ—ограниченным*, если последовательность

$$w, \varphi(w), \varphi^2(w), \varphi^3(w), \dots$$

периодична начиная с некоторого момента. В противном случае, $|\varphi^n(w)| \to \infty$ при $n \to \infty$ и слово w называется *φ-растущим.* Очевидно, слово является φ-ограниченным тогда и только тогда, когда оно состоит из φ-ограниченных букв.

Теорема 4.1.3. *Существует алгоритм, который определяет, конечно ли в $\varphi^\infty(a)$ число различных φ—ограниченных подслов. Если это число бесконечно, то можно проверить, является ли W периодическим. Если оно конечно, то все эти слова алгоритмически находятся.*

Доказательство. Прежде всего отметим, что все φ—возрастающие буквы алгоритмически находятся (см., например, [107]). Далее φ-растущие буквы будем писать, как a_1, a_2 и т.д.

Построим ориентированный граф Q, на ребрах которого будут записаны упорядоченные пары слов. Вершинами этого графа будут служить φ-растущие буквы из A а также всевозможные упорядоченные пары φ-растущих букв. Введем фиктивную букву t, также к вершинам Q добавим всевозможные пары вида $a_i t$, где a_i — φ-растущая буква.

Из вершины a_i в a_j идет ребро, если $a_j \in a_i$. На таких ребрах пара слов — $\{\varepsilon, \varepsilon\}$ (ε– пустое слово). Из вершины a_i в $a_j a_k$ ведет ребро со словами $\{\omega, \varepsilon\}$, если для некоторого φ—ограниченного слова ω слово $a_j \omega a_k$ является подсловом $\varphi(a_i)$ (из a_i в $a_j a_k$ могут вести несколько ребер.) Из a_i и $a_i t$ ведет по ребру с парой $\{\omega, \varepsilon\}$ в $a_j t$, если ω — φ—ограниченное и $\varphi(a_i)$ оканчивается на $a_j \omega$.

Из $a_i a_j$ ведет ребро в $a_k a_l$ с парой φ—ограниченных слов $\{\omega_1; \omega_2\}$, если $\varphi(a_1)$ кончается на $a_k\omega_1$, а $\varphi(a_2)$ начинается на $\omega_2 a_l$.

Предложение 4.1.1. *Пусть $k \in \mathbb{N}$. Рассмотрим какой-нибудь путь длины k по ребрам графа Q, выходящий из a_1. Последовательность пар слов на ребрах этого пути*

$$\{u_1, v_1\}, \{u_2, v_2\}, \ldots, \{u_k, v_k\}.$$

Тогда в $\varphi^k(a_1)$ есть φ—нерасширяемое слово

$$u_k\varphi(u_{k-1}) \ldots \varphi^{k-1}(u_1)\varphi^{k-1}(v_1)\varphi^{k-2}(v_2) \ldots v_k.$$

При этом, если путь оканчивается на $a_i a_j$, то есть вхождение этого слова, обрамленное буквами a_i и a_j.

Наоборот, любое φ—ограниченное подслово $\varphi^k(a_1)$ можнно получить, получив слово по указанному правилу слово и взяв его подслово.

Это несложно показывается индукцией по k.

Первый случай: в любом ориентированном цикле графа Q, до которого можно добраться из a_1, на ребрах цикла написаны пары пустых слов. Тогда в любом пути, выходящем из вершины a_1, число ребер, на которых написаны не пустые слова, не превосходит количества вершин в Q. Следовательно, число различных φ—ограниченных слов конечно, а по графу Q можно их всех найти.

Второй случай: есть цикл, и в этом цикле не все слова пустые. Пусть, например, в цикле есть пары слов, в которых первое слово не пустое. Тогда для некоторой буквы a_i и непустого φ—ограниченного u, слово $\varphi(a_i)$ оканчивается на $a_i u$. Образ u при подстановке φ зациклится. Следовательно, для некоторого непустого слова U для любого $k \in \mathbb{N}$ слово $h(U^k)$ является подсловом W. Тогда сверхслово W заключительно периодично если и только если оно заключительно периодично с периодом $h(U)$. А проверять, есть ли заключительная периодичность с данным периодом умеем (см. следствие 4.1.1).

□

Пусть I_φ — множество всех φ-растущих букв, B_φ — множество φ-ограниченных подслов сверхслова $\varphi^\infty(a_1)$(включая пустое). Можно считать, что B_φ конечно и что мы знаем все слова в B_φ. Рассмотрим (конечный) алфавит C, состоящий из символов $[twt']$, где t и t' буквы из I_φ, а w — слово из B_φ и слово twt' является подсловом $\varphi^\infty(a)$.

Определим морфизм $\psi : C^* \to C^*$ следующим образом:

$$\psi([twt']) = [t_1wt_2][t_2wt_3]\dots[t_kw_kt_{k+1}],$$

где $\varphi(tw) = w_0t_1w_1t_2\dots t_kw'_k$, слово $\varphi(t')$ начинается с w''_kt_{k+1} и $w_k = w'_kw''_k$ (слова w_i, w'_k и w''_k принадлежат B_φ).

Также определим $f : C^* \to A^*$ по правилу

$$f([twt']) = tw.$$

Предложение 4.1.2. *Все буквы алфавита C являются ψ-растущими.*

Доказательство. Заметим, что в $\psi^n([twt'])$ столько же букв, сколько в слове $\varphi^n(t)$ φ-растущих букв. Очевидно, в образе $\varphi(t)$ от произвольной буквы $t \in I_\varphi$ содержится хотя бы одна буква из I_φ. Более того, в слове $\varphi^n(t)$ для некоторого n содержатся хотя бы две буквы из I_φ, иначе $\varphi^n(t) = w_nt_{i_n}v_n$ (где w_n и v_n принадлежат B_φ) и $|\varphi^nt|$ ограниченно. $\square$

Пусть $\varphi^\infty(a)$ имеет вид $a_1w_1a_2\dots$, где $a_1, a_2 \in I_\varphi$, $w_1 \in B_\varphi$. Тогда, несложно убедиться, что для любого n слово $\varphi^n(a)$ является началом слова $f(\psi^n([a_1w_1a_2]))$, следовательно, слова W и $(h \circ f)(\psi^\infty([a_1w_1a_2]))$ совпадают. Заметим, что морфизм $h' = h \circ f$ является нестирающим.

Замечание 4.1.1. *Конструкция морфизма ψ встречалась в работах [107, 108].*

Букву $c_i \in C$ назовем *рекуррентной*, если для некоторого $k \in \mathbb{N}$ выполнено $c_i \sqsubseteq \psi^k(c_i)$. Для каждой рекуррентной буквы $c_i \in C$ существует такое число $k(c_i)$, что если $k(c_i)|n$, то $c_i \sqsubseteq \psi^n(c_i)$. Следовательно, существует такое $n \in \mathbb{N}$, что если c_i — произвольная рекуррентная буква алфавита C, то $c_i \sqsubseteq \psi^n([a_1w_1a_2])$. Положим $\rho = \psi^n$.

Рассмотрим ориентированный граф G_ρ, вершинами которого являются буквы алфавита C, и из c_i ведет стрелка в c_j тогда и только тогда, когда c_j содержится в $\rho(c_i)$.

Пусть D — сильносвязная компонента этого графа, до которой можно дойти по стрелочкам из $[a_1wa_2]$. Рассмотрим ограничение ρ на D^*. Все буквы из

D являются рекуррентными. Следовательно, если $d \in D$, то $d \sqsubseteq \rho(d)$. Также для любого k выполнено $\rho^k(d) \sqsubseteq \rho^{k+1}(d)$. Следовательно, существует такое m, что для любых букв d_1 и $d_2 \in D$ $d_2 \sqsubseteq \rho^m(d_1)$. Поэтому морфизм ρ в ограничении на D^* является примитивным.

Найдется такая буква $d \in D$, что $\rho^l(d)$ начинается на d для некоторого l. Обозначим $\rho_2 = \rho^l$. Так как все буквы из D являются ρ_2-растущими, то $\rho_2^\infty(d)$ является бесконечным сверхсловом, при этом все его конечные подслова являются подсловами $\varphi^\infty(a_1)$.

Спросим оракула о периодичности сверхслова $h'(\psi^\infty(d))$. Если оно непериодично, то в W бесконечно много специальных справа подслов и само W не является заключительно периодичным.

Если же оракул сказал, что сверхслово $h'(\psi^\infty(d))$ периодично и его период — слово u, то, если W заключительно периодично, его периодом является то же самое слово u. Согласно следствию 4.1.1, мы можем проверить, верно ли, что $h'(\psi^\infty(d))$ заключительно периодично с периодом u. Таким образом, утверждение теоремы 4.0.1 верно, если разшешима проблема периодичности для примитивных морфических слов.

4.2 Графы и схемы Рози.

Пусть W — бесконечное слово. *Граф Рози порядка k* этого сверхслова обозначается $G_k(W)$, если же понятно, о каком слове идет речь, то будем писать просто G_k. Вершинами этого графа являются всевозможные различные подслова сверхслова W, имеющие длину k. Две вершины графа u_1 и u_2 соединяются направленным ребром, если в W есть такое подслово v, что $|v| = k+1$, $v[1]v[2]\ldots v[k] = u_1$ и $v[2]v[3]\ldots v[k+1] = u_2$. (Запись $v[i]$ обозначает $i-$тую букву слова v.)

Если w — подслово W длины $k+l$, ему соответствует в G_k путь длины l. Этот путь проходит по ребрам, соответствующим подсловам слова w длины $k+1$.

Свойства W, которые определяются множеством его конечных подслов, можно определить, зная последовательность его графов Рози. Так, рекуррентность W означает, что все графы G_k являются сильносвязными, а периодичность — то, что при некотором k граф G_k является циклом. Из 4.1.2 следует, что для морфического слова и любого k по морфизму можно алгоритмически найти граф G_k. Поведение же последовательности графов G_k определить затруднительно; для некоторого подкласса чисто морфических слов это было сделано в работе [57].

Схемы Рози также определяют множество его конечных подслов, но при этом их поведение проще описывается.

Графом со словами будем называть сильносвязный ориентированный граф, у которого на каждом ребре написано по два слова — *переднее* и *заднее*. Также потребуем, чтобы каждая вершина либо имеет входящую степень 1, а исходящую больше 1, либо входящую степень больше 1 и исходящую степень 1. Вершины первого типа назовем *раздающими*, а второго — *собирающими*.

Путь в графе со словами — это последовательность ребер, каждое следующее из которых выходит из той вершины, в которую входит предыдущая. *Симметричный путь* — это путь, первое ребро которого начинается в собирающей вершине, а последнее ребро кончается в раздающей.

Каждый путь можно записать словом над алфавитом — множеством ребер графа, и это слово называется *реберной записью пути*. Иногда мы будем отождествлять путь и его реберную запись. Ребро пути s, идущее i-тым по счету, будем обозначать $s[i]$. Для двух путей так же, как и для слов, определены отношения *подпути* (пишем $s_1 \sqsubseteq s_2$), *начала* и *конца*. Кроме того, пишем $s_1 \sqsubseteq_k s_2$, если для соответствующих слов u_1 и u_2 — реберных записей путей s_1 и s_2 — выполнено $u_1 \sqsubseteq_k u_2$. Если последнее ребро пути s_1 идет в ту же вершину, из которой выходит первое ребро пути s_2, путь, реберная запись которого является конкатенацией реберных записей путей s_1 и s_2, будем обозначать $s_1 s_2$.

Замечание 4.2.1. *Определения пути, подпути, реберной записи, начала и конца имеют смысл для любых ориентированных графов.*

Введем понятие *переднего слова $F(s)$, соответствующего пути s в графе со словами*. Пусть $v_1 v_2 \dots v_n$ — реберная запись пути s. В $v_1 v_2 \dots v_n$ возьмем подпоследовательность: включим в нее v_1, а также те и только те ребра, которые выходят из раздающих вершин графа. Эти ребра назовем *передними* *образующими для пути* s. Возьмем передние слова этих ребер и запишем их последовательную конкатенацию, там получаем $F(s)$.

Аналогично определяется $B(s)$. В $v_1 v_2 \ldots v_n$ возьмем ребра, входящие в собирающие вершины и ребро v_n в порядке следования — это *задние образующие для пути* s. Тогда последовательной конкатенацией задних слов этих ребер получается *заднее слово* $B(s)$ *пути* s.

Определение 4.2.1. Граф со словами будет являться *схемой Рози* для рекуррентного сверхслова W, если он удовлетворяет следующим свойствам, которые в дальнейшем будут называться *свойствами схем Рози*:

1. Граф сильносвязен и состоит более чем из одного ребра.
2. Все ребра, исходящие из одной раздающей вершины графа, имеют передние слова с попарно разными первыми буквами. Все ребра, входящие в одну собирающую вершину графа, имеют задние слова с попарно разными последними буквами.
3. Для любого симметричного пути, его переднее и заднее слова совпадают. То есть можно говорить просто о слове симметричного пути.
4. Если есть два симметричных пути s_1 и s_2 и выполнено $F(s_1) \sqsubseteq_k F(s_2)$, то $s_1 \sqsubseteq_k s_2$.
5. Все слова, написанные на ребрах графа, являются подсловами W.
6. Для любого u – подслова W существует симметричный путь, слово которого содержит u.
7. Для любого ребра s существует такое слово u_s, принадлежащее W, что любой симметричный путь, слово которого содержит u_s, проходит по ребру s.

В работе [125] понятие схем Рози вводилось только для непериодических слов.

Лемма 4.2.1. *Пусть* S *— произвольный граф со словами.*

1. *Если путь* s_1 *оканчивается в раздающей вершине и в этой же вершине начинается путь* s_2*, то* $F(s_1 s_2) = F(s_1)F(s_2)$.

2. *Если путь s_1 оканчивается в собирающей вершине и в этой же вершине начинается путь s_2, то $B(s_1 s_2) = B(s_1)B(s_2)$.*

Доказательство. 1.Множество образующих передних ребер пути $s_1 s_2$ — это в точности образующие передние ребра пути s_1 и образующие передние ребра пути s_2, записанные последовательно.

2.Множество образующих задних ребер пути $s_1 s_2$ — это образующие задние ребра пути s_1 и образующие задние ребра пути s_2, записанные последовательно. □

Следствие 4.2.1. *Если для графа со словами S выполнено свойство 3 схем Рози, s — симметричный путь в S, s_1 — произвольный путь в S и $s_1 \sqsubseteq s$, то $F(s_1) \sqsubseteq F(s)$.*

Лемма 4.2.2. *Пусть S — граф со словами, для которого выполнено свойство 3 схем Рози. Если s_1 и s_2 — два симметричных пути таких, что $s_1 \sqsubseteq_k s_2$, то $F(s_1) \sqsubseteq_k F(s_2)$.*

Доказательство. Пусть $v_1 v_2 \dots v_n$ — реберная запись пути s_2, а $w_1 w_2 \dots w_m$ — реберная запись s_1. Для каждого вхождения слова $w_1 w_2 \dots w_m$ в слово $v_1 v_2 \dots v_n$ разобьем слово $v_1 v_2 \dots v_n$ на две части. Ту часть, которая находится правее соответствующего вхождения буквы w_m, назовем *путевым концом*, а остальное — *путевым началом.* Так как ребро w_m входит в раздающую вершину, то любое путевое начало является симметричным путем, а первое ребро любого путевого конца является в пути s_2 передним образующим ребром. Если s_b и s_e — путевые начало и конец соответственно, то $F(s_2) = F(s_b)F(s_e)$, при этом $F(s_b)$ оканчивается на $F(s_1)$. Для доказательства леммы достаточно показать, что передние слова любых двух путевых концов различные. Для любого путевого конца s_e множество его передних образующих ребер — это те передние образующие ребра пути s_2, которые находятся в s_e. Но для различных путевых концов такие подмножества вложены одно в другое, а так как для любого путевого конца его первое ребро является образующим, то они вложены строго. □

Определение 4.2.2. Симмметричный путь s называется *допустимым*, если $F(s) \sqsubseteq W$.

Лемма 4.2.3. *Пусть S — граф со словами, для которого в применении к сверхслову W выполнены свойства* $1-6$ *графов Рози. Если $u \sqsubseteq W$, то в схеме S есть допустимый путь s такой, что $u \sqsubseteq F(s)$.*

Доказательство. Пусть $l_{\max}$ — максимальная из длин слов на ребрах S. Так как сверхслово W рекуррентно, то в нем есть вхождение слова u такое, что первая буква этого вхождения имеет в W номер более $l_{\max}$. Рассмотрим w — подслово W, имеющее вид u_1uu_2, где $|u_1| > l_{\max}$, $|u_2| > l_{\max}$. Согласно свойству 6, в схеме S существует симметричный путь s_1 такой, что $w \sqsubseteq F(s_1)$. Можно считать, что s_1 — минимальный (относительно $\sqsubseteq$) симметричный путь с таким свойством.

Пусть $v_1v_2\ldots v_n$ — реберная запись этого пути. Пусть v_{n_1} — последнее ребро, являющееся передним образующим для пути s_1. Это либо v_1, либо ребро, выходящее из раздающей вершины. Если верно первое, то в пути s_1 только одно переднее образующее ребро и $|F(s_1)| \leq l_{\max}$, противоречие с длиной слова w.

Значит, можно рассмотреть симметричный путь с реберной записью $v_1v_2\ldots v_{n_1-1}$. Из минимальности s_1 следует, что $w \not\sqsubseteq F(v_1v_2\ldots v_{n_1-1})$.

Рассуждая аналогично, получим, что если v_{n_2} — первое ребро, являющееся задним образующим, то $w \not\sqsubseteq F(v_{n_2}v_{n_2+1}\ldots v_n)$. Докажем следующий факт: $n_2 \leq n_1$. В самом деле, иначе, в силу минимальности n_2, среди ребер $v_1, v_2, \ldots v_{n_1}$ ни одно не входит в собирающую вершину, то есть в симметричном пути $v_1v_2\ldots v_{n_1-1}$ ровно одно заднее образующее ребро (а именно v_{n_1-1}) и, стало быть, $|B(v_1v_2\ldots v_{n_1-1})| \leq l_{\max}$. Но

$$F(s_1) = F(v_1v_2\ldots v_{n_1-1})F(v_{n_1}) = B(v_1v_2\ldots v_{n_1-1})F(v_{n_1}) \leq 2l_{\max}.$$

Противоречие с тем, что $|F(s_1)| > 2l_{\max}$.

Значит, мы можем рассматривать симметричный путь

$$s_2 = v_{n_2}v_{n_2+1}\ldots v_{n_1-1}.$$

Пусть $w' = F(s_2)$. Мы можем записать $F(s_1) = B(v_{n_2-1})w'F(v_{n_1})$. Так как слово $F(s_1)$ содержит w, а слова $B(v_{n_2-1})w'$ и $w'F(v_{n_1})$ не содержат, то $w' \sqsubseteq w$. Стало быть, $w' \sqsubseteq W$.

С другой стороны, $F(s_1) = u_1'u_1uu_2u_2'$, где $\min\{|u_1'u_1|, |u_2u_2'|\} \geq l_{\max}$. А так как $\max\{B(v_{n_2-1}), F(v_{n_1})\} \leq l_{\max}$, то $u \sqsubseteq w'$. Следовательно, путь s_2 — искомый.

□

Лемма 4.2.4. *Пусть имеется допустимый путь l со словом u. Если $uu_1 \sqsubseteq W$ для некоторого u_1, то в схеме Рози S есть такой допустимый путь l', что его началом является l, а его слово начинается с uu_1.*

Доказательство. Согласно лемме 4.2.3 в схеме S существует допустимый путь l_2 такой, что $uu_1 \sqsubseteq F(l_2)$. Рассмотрим слово $F(l_2)$. В нем может быть несколько вхождений слова u, причем среди них есть такое (допустим, k—тое, если считать с конца), после которого сразу идет u_1.

Тогда $l \sqsubseteq_k l_2$. Пусть $l_2 = s_1ls_2$ (здесь рассматривается k—тое с конца вхождение пути l). Рассмотрим k-тое с конца вхождение пути l в путь l_2. Первое ребро пути l выходит из собирающей вершины, а последнее ребро пути s_2 входит в раздающую вершину. Следовательно, ls_2 — симметричный путь. Этот путь является допустимым, так как $B(l_2) = B(s_1)B(ls_2)$ и $B(l_2) \sqsubseteq W$.

Так как l имеет ровно k вхождений в ls_2, то, согласно свойству 4 определения схем Рози 4.2.1 и лемме 4.2.4, $u = F(l)$ имеет ровно k вхождений в $F(ls_2)$. Кроме того, $F(ls_2)$ начинается с $F(l) = u$. Двум этим свойствам удовлетворяет ровно одно окончание слова $F(l_2)$. Значит, $F(ls_2)$ начинается со слова uu_1. Стало быть, путь ls_2 — искомый. □

Замечание 4.2.2. *Аналогично доказывается следующий факт: пусть имеется допустимый путь l со словом u. Если $u_1u \sqsubseteq W$ для некоторого u_1, то в схеме Рози S есть такой допустимый путь l', что его концом является l, а его слово кончается на u_1u.*

Следствие 4.2.2. *Если u — биспециальное подслово W такое, что оно содержит слово некоторого симметричного пути l_1 схемы S, то в схеме S существует такой симметричный путь l, что $F(l) = u$.*

Доказательство. Пусть $u = u_1F(l_1)u_2$. Из биспециальности u следует, что существуют буквы a_1 и a_2 такие, что $F(l_1)u_2a_1 \sqsubseteq W$, $F(l_1)u_2a_2 \sqsubseteq W$. Тогда существуют два пути l_1s_1 и l_1s_2 такие, что слово первого начинается с $F(l_1)u_2a_1$, а второго — с $F(l_1)u_2a_2$.

Пусть $v_1v_2\dots v_{k-1}v_k\dots$ — реберная запись l_1s_1, а $v_1v_2\dots v_{k-1}v'_k\dots$ — реберная запись l_1s_2, различие происходит в ребре с номером k. Очевидно, ребра v_k и v'_k выходят из одной и той же вершины. Значит, эта вершина раздающая и путь $v_1v_2\dots v_{k-1}$ — симметричный. По свойству 2 схем Рози (см. определение 4.2.1), $F(v_k)$ и $F(v'_k)$ начинаются с разных букв. Так как $F(l_1s_1)$ начинается с $F(v_1v_2\dots v_{k-1})F(v_k)$, а $F(l_1s_2)$ — с $F(v_1v_2\dots v_{k-1})F(v'_k)$, то $F(v_1v_2\dots v_{k-1}) = F(l_1)u_2$. При этом путь $v_1v_2\dots v_{k-1}$ начинается с l_1.

Аналогично рассуждая, найдем симметричный путь $w_1w_2\dots w_m$, концом которого является путь l_1 и словом которого является $u_1F(l_1)$. Пусть $w_1w_2\dots w_m = l'l_1$. Тогда путь $l'v_1v_2\dots v_{k-1}$ является искомым. □

Напомним, как получать схемы Рози из графов Рози.

Фиксируем натуральное число k. Рассмотрим G_k — граф Рози порядка k для сверхслова W.

Предложение 4.2.1. *G_k есть сильносвязный орграф. Если слово W не является периодичным, то этот граф не является циклом.*

Доказательство. Если u_1 и u_2 — подслова W длины k, то, в силу рекуррентности сверхслова W, в W найдется подслово вида $a_{i_1}a_{i_2}\dots a_{i_n}$, где $a_{i_1}a_{i_2}\dots a_{i_k} = u_1$, $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n} = u_2$.

Тогда слова вида $a_{i_l}a_{i_{l+1}}\dots a_{i_{l+k}}$ имеют длину $k+1$ и соответствуют ребрам графа G_k, образующим путь, соединяющий вершины, соответствующие словам u_1 и u_2. Кроме того, если бы G_k был циклом длины n, то сверхслово W было бы периодичным с периодом n. □

Если слово W периодично, то для всех достаточно больших n граф G_n является циклом, графы же с малыми номерами циклами могут и не являться.

Замечание 4.2.3. *Дальнейшие построения проводятся для графа Рози, не являющегося циклом. Про периодичность самого сверхслова ничего не утверждается.*

Граф S получается из G_k следующими операциями:

1. Все простые цепи, соединяющие вершины, заменяются на единичные ребра.

2. Каждая биспециальная вершина a заменяется на ребро v_a. При этом все ребра, которые раньше вели в a, будут вести в начало v_a, а ребра, которые раньше шли из a, будут выходить из конца ребра v_a.

В графе S все вершины являются либо раздающими, либо собирающими.

Рассмотрим те пути в графе G_k, которые соединяют две специальные вершины. Пусть путь ведет из вершины $a_{i_1}a_{i_2}\dots a_{i_k}$ в вершину $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$. Сопоставим ему слово по следующему правилу:

1. Если вершина, соответствующая $a_{i_1}a_{i_2}\dots a_{i_k}$, явлется раздающей, то переднее слово пути — это $a_{i_{k+1}}a_{i_{k+2}}\dots a_{i_n}$.

2. Если вершина, соответствующая $a_{i_1}a_{i_2}\dots a_{i_k}$, явлется в собирающей, то переднее слово пути — это $a_{i_1}a_{i_2}\dots a_{i_n}$.

3. Если вершина, соответствующая $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$, является собирающей, то заднее слово пути — это $a_{i_1}a_{i_2}\dots a_{i_n-k}$.

4. Если вершина, соответствующая $a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$, является раздающей, то заднее слово пути — это $a_{i_1}a_{i_2}\dots a_{i_n}$.

Определение 4.2.3. Если S — сильносвязный граф, не являющийся циклом, и s — путь в графе, то *естественное продолжение* пути s *вправо* — это минимальный путь, началом которого является s и который оканчивается в раздающей вершине. *Естественное продолжение* пути s *влево* — это минимальный путь, концом которого является s и который начинается в собирающей вершине.

Очевидно, для сильносвязных нецикличных графов естественное продолжение существует всегда и единственно. Теперь мы готовы написать слова на ребрах S: если v — ребро в S, то в графе G_k ему соответствует либо биспециальная вершина, либо простая цепь. Если это биспециальная вершина, то есть слово длины k, то в качестве $F(v)$ и $B(v)$ возьмем это слово. Если ребру v соответствует простая цепь, то в качестве $F(v)$ возьмем то переднее слово, которое соответствует естественному продолжению вправо этой цепи. Заднее же слово ребра v — это заднее слово пути в G_k, соответствующего естественному продолжению влево рассматриваемой цепи.

В G_k биспециальные вершины можно считать путями длины 0.

Предложение 4.2.2. *В полученном графе со словами S для любого пути s слово $F(s)$ — это переднее слово того пути графа G_k, который соответствует естественному продолжению вправо пути s. Аналогично, $B(s)$ — это заднее слово того пути в G_k, который соответствует естественному продолжнию пути s влево (в графе S).*

Доказательство. Проведем доказательство для передних слов. В графе S для любого пути s его естественное продолжение вправо разбивается на естественные продолжения вправо его передних образующих ребер, причем все естественные продолжения, начиная со второго, выходят из раздающих вершин и, стало быть, слова на них пишутся по правилу для первого из четырех типов. Значит, слова для этих ребер в объединении и дадут слово для длинного пути в G_k. Для задних слов все аналогично. $\square$

Лемма 4.2.5. *Пусть S — определенный выше граф со словами. Тогда он является схемой Рози для сверхслова W.*

Доказательство. **Свойство 1.** Так как G_k является сильносвязным и не является циклом, то же самое верно и для S.

Свойство 2. Два пути, выходящие из одной раздающей вершины графа S с различными первыми ребрами, соответствуют путям в графе G_k, которые выходят из одной раздающей вершины $a_{i_1}a_{i_2}\dots a_{i_k}$, причем первые ребра у путей разные. Пусть вторые вершины путей — это $a_{i_2}a_{i_3}\dots a_{i_k}b$ и $a_{i_2}a_{i_3}\dots a_{i_k}c$ соответственно. Тогда слова этих путей начинаются на буквы b и c соответственно, и на эти буквы начинаются слова соответствующих путей в S.

Свойство 3 следует из того, что для симметричного пути его естественным продолжением вправо и естественным продолжением влево является он сам. Согласно 4.2.2, передним и задним словом такого пути будут соответственно переднее и заднее слово пути

$$a_{i_1}a_{i_2}\dots a_k,\ a_{i_2}a_{i_3}\dots a_{k+1},\ \dots, a_{i_l}a_{i_{l+1}}\dots a_{i_{l+n-1}},\ \dots,\ a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}$$

графа G_k. Переднее слово этого пути определяется по типу 2, а заднее — по типу 4, следовательно, они совпадают.

Докажем **свойство 4**. Пусть в S нашлись два симметричных пути s_1 и s_2 такие, что $F(s_1) \sqsubseteq_l F(s_2)$. Соответственные пути в G_k обозначим s_1' и s_2'. Очевидно, достаточно показать, что $s_1' \sqsubseteq_l s_2'$. пути s_1' и s_2' выходят из собирающих вершин графа G_k и входят в раздающие вершины. Передние слова этих путей — это $F(s_1)$ и $F(s_2)$. Последовательности их $(k+1)$-буквенных подслов — это последовательности ребер s'_1 и s'_2. Значит, реберная запись пути s'_1 встречается в реберной записи s'_2 хотя бы l раз. Если же $F(s_1)$–биспециальное слово длины k, то, очевидно, путь s'_2 хотя бы l раз проходит по вершине $F(s_1)$.

Свойство 5 достаточно доказать для передних слов. Пусть v — ребро схемы S. Рассмотрим в S естественное продолжение ребра v вправо. Оно соответствует пути s в G_k. Пусть s содержит L ребер. Первая вершина этого пути — слово $a_{i_1}a_{i_2}\dots a_{i_k}$. Первое ребро пути s соответствует подслову u вида $a_{i_1}a_{i_2}\dots a_{i_k}b$, при этом $a_{i_1}a_{i_2}\dots a_{i_k}b \sqsubseteq W$. Очевидно, существует u_1 — подслово сверхслова W длины $L+k$, начинающееся с $a_{i_1}a_{i_2}\dots a_{i_k}b$. Ему соответствует путь s' длины L по ребрам G_k с первым ребром таким же, как у пути s.

Так как среди промежуточных вершин пути s нет раздающих вершин (иначе s не соответствует естественному продолжению ребра v вправо), то s' должен совпадать с путем s. Стало быль, слово пути s является подсловом u_1 и, следовательно, подсловом W.

Свойство 6. Существует такое число M, что любом пути длины M по ребрам G_k будет хотя бы одна раздающая и хотя бы одна собирающая вершина. В самом деле, в качестве M подойдет количество вершин в G_k, иначе существовал бы цикл, в котором не было бы либо раздающих, либо собирающих вершин и G_k был бы, либо не сильносвязным, либо циклом.

Пусть $w \sqsubseteq W$. Тогда существует $w' \sqsubseteq W$ вида w_1ww_2, где $|w_1| = 2M$, $|w_2| = 2M$. Рассмотрим в G_k путь, соответствующий слову w'. Среди его последних M вершин есть раздающая, а среди первых M — собирающая. Значит, можно взять путь, который короче не более, чем на $2M$ и соответствует симметричному пути в S. Этот путь соответствует подслову w', которое имеет длину не менее $|w'|-2M$ и, следовательно, содержит w. А это значит, что слово некоторого симметричного пути в S содержит w.

Докажем **свойство 7**. Пусть v — ребро в графе S. В графе Рози G_k ребру v соответствует путь v', проходящий через вершины

$$a_{i_1}a_{i_2}\dots a_{i_k},\ a_{i_2}a_{i_3}\dots a_{i_{k+1}},\ \dots, a_{i_l}a_{i_{l+1}}\dots a_{i_{l+n-1}},\ \dots,\ a_{i_{n-k+1}}a_{i_{n-k}}\dots a_{i_n}.$$

Ни одна из вершин пути кроме первой и последней не является раздающей или собирающей.

Для S уже доказаны свойства 1–6. Для S справедлива лемма 4.2.3. Для слова $a_{i_1}a_{i_2}\dots a_{i_k}a_{i_{k+1}}$, являющегося подсловом сверхслова W, найдем в S допустимый путь w такой, что $a_{i_1}a_{i_2}\dots a_{i_k}a_{i_{k+1}} \sqsubseteq F(w)$. Соответственный путь в графе G_k содержит первое ребро пути v', а, следовательно, и весь v'. Стало быть, w содержит v, а так как для S выполнено свойство 4, то любой симметричный путь, содержащий $F(w)$, содержит ребро v. $\square$

4.3 Элементарная эволюция схем Рози.

В этом разделе мы распространим определенную в [125] эволюцию на схемы для периодических слов. Грубо говоря, сначала схемы для периодических слов выглядят так же, как и для периодических, но на некотором шаге эволюция останавливается.

Далее W — рекуррентное сверхслово, S — схема Рози для этого сверхслова. Из сильносвязности S следует, что существует ребро v, идущее из собирающей вершины в раздающую. Такие ребра будем называть *опорными*.

Цель этого раздела — определить *эволюцию* (W, S, v) схемы Рози S по опорному ребру v. Опорное ребро является само по себе симметричным путем, стало быть, переднее и заднее слова этого ребра совпадают.

Пусть $\{x_i\}$ — множество ребер. входящих в начало v, а $\{y_i\}$ — множество ребер, идущих из конца v. Эти два множества могут пересекаться. Обозначим $F(y_i) = Y_i$, $B(x_i) = X_i$, $F(v) = V$.

Рассмотрим все слова вида X_iVY_j. Если такое слово не входит в W, то пару (x_i, y_j) назовем *плохой*, в противном случае — *хорошей*. Также *хорошей* или *плохой* будем называть соответствующую тройку ребер (x_i, v, y_j).

Построим граф S'. Он получается из S заменой ребра v на $K_{\#\{x_i\},\#\{y_j\}}$, где $K_{m,n}$ — полный двудольный граф. Более подробно: ребро v удаляется, его начало заменяется на множество $\{A_i\}$ из $\#\{x_i\}$ вершин так, что для любого i ребро x_i идет в A_i; конец ребра v заменяется на множество $\{B_j\}$ из $\#\{y_j\}$ вершин так, что для любого j ребро y_j выходит из B_j; вводятся ребра $\{v_{i,j}\}$, соединяющие вершины множества $\{A_i\}$ с вершинами множества $\{B_j\}$.

По сравнению с S, у графа S' нет ребра v, но есть новые ребра $\{v_{ij}\}$. Остальные ребра граф S взаимно однозначно соответствуют ребрам графа S'. Соответственные ребра в первом и втором графе зачастую будут обозначаться одними и теми же буквами.

На ребрах S' расставим слова следующим образом. На всех ребрах S', кроме ребер из $\{y_i\}$ и $\{v_{i,j}\}$, передние слова пишутся те же, что и передние слова соответственных ребер в S. Для каждого i и j, переднее слово ребра y_j в S' — это VY_j. Переднее слово ребра $v_{i,j}$ — это Y_j. Аналогично, на всех ребрах, кроме ребер из $\{x_i\}$ и $\{v_{i,j}\}$, задние слова переносятся с соответствующих ребер S; для всех i и j в качестве заднего слова ребра x_i возьмем X_iV, а в качестве заднего слова ребра $v_{i,j}$ — X_i.

Теперь построим граф S''. Ребра $v_{i,j}$, соответствующие плохим парам (x_i, y_j), назовем *плохими*, а все остальные ребра графа S' — *хорошими.* Граф S'' получается, грубо говоря, удалением плохих ребер из S'. Более точно, в графе S'' раздающие вершины — подмножество раздающих вершин S', а собирающие вершины — подмножество собирающих вершин S'. В графе S' эти подмножества — вершины, из которых выходит более одного хорошего ребра, и вершины, в которые входит более одного хорошего ребра соответственно. Назовем в S' вершины этих двух подмножеств S' неисчезающими. Ребра в графе S'' соответствуют таким путям в S', которые идут лишь по хорошим ребрам, начинаются в неисчезающих вершинах, заканчиваются в неисчезающих, а все промежуточные вершины которых не являются неисчезающими.

Согласно лемме 4.3.2, которая будет доказана ниже, S'' — сильносвязный граф. Если он является циклом, то, как видно из этой же леммы, сверхслово W периодично, при этом период сверхслова определяется по схеме S и множеству плохих пар ребер. В таких случаях будем говорить, что *элементарная эволюция (W,S,v) выявила периодичность сверхслова.*

Если же S'' не является циклом, то у каждого ребра в S'' есть естественное продолжение вперед и назад. Естественное продолжение вперед соответствует пути в S'; переднее слово этого пути в S' возьмем в качестве переднего слова для соответствующего ребра S''. Аналогично для задних слов: у ребра в S'' есть естественное продолжение влево, этому пути в S'' соответствует путь в S'. Заднее слово этого пути и будет задним словом ребра в S''.

Определение 4.3.1. Построенный таким образом граф со словами S'' назовем *элементарной эволюцией* (W, S, v).

Теорема 4.3.1. *Пусть W — схема Рози. Тогда элементарная эволюция (W, S, v) является схемой Рози для сверхслова W (то есть удовлетворяет свойствам* 1—7 *определения 4.2.1).*

Доказательству этого и будет посвящена оставшаяся часть раздела.

Опишем соответствие f между симметричными путями (за исключением опорного ребра v) в схеме S и симметричными путями в S'. Пусть s — симметричный путь в S — имеет реберную запись $v_1 v_2 \ldots v_n$. Некоторые (может быть, ни одно) из этих ребер являются ребром v. Очевидно, что если ребро v встречается в реберной записи, то оно находится между ребрами x_i и y_j для некоторых i и j. Если v является первым или последним ребром пути s, то соответствующее x_i или y_j находится только с одной стороны. Соответствие f преобразует $v_1 v_2 \ldots v_n$ (реберную запись s) следующим образом: для каждого вхождения v определяется, находится ли оно на краю реберной записи или в середине, если на конце, то "v" просто отбрасывается, а если в середине, то "v" заменяется на "v_{ij}". Индексы i и j берутся те же, что и у ребер x_i и y_j, стоящих рядом с v в реберной записи s. Путь с соответствующей реберной записью является образом $f(s)$ при соответствии.

Обратное соответствие f^{-1} таково: если в графе S' есть путь l и его реберная запись — $v_1 v_2 \ldots v_n$, то каждое вхождение ребра "v_{ij}" заменяется на "v". Кроме того, если v_1 (первое ребро пути) — это одно из $\{y_i\}$, то в начало реберной записи приписывается "v", если же v_n (последнее ребро пути) — одно из ребер $\{x_i\}$, то "v" приписывается в конец реберной записи.

Очевидно, f и f^{-1} сохраняют свойство последовательности ребер "начало каждого следующего ребра является концом предыдущего", то есть пути переводят в пути. Кроме того, f и f^{-1} на самом деле являются взаимно обратными. Далее, в графе S' ребра вида y_j выходят из собирающих вершин, ребра вида x_i входят в раздающие, ребра v_{ij} выходят из раздающих вершин и входят в собирающие; в графе S ребро v, как опорное, выходит из собирающей и входит в раздающую, ребра x_i входят в собирающую вершину, а y_j выходят из раздающей, для всех остальных ребер графа S свойство выходить из собирающей вершины или входить в раздающую сохраняется в графе

S'. Таким образом, при соответствиях f и f^{-1} сохраняется симметричность путей.

Предложение 4.3.1. *При этом соответствии сохраняются слова симметричных путей.*

Доказательство. Докажем этот факт для передних слов, так как для задних все аналогично. Автоматически будет доказано, что передние и задние слова симметричных путей в S' совпадают.

Пусть в графе S имеется симметричный путь s с реберной записью $v_1 v_2 \dots v_n$. Разберем четыре случая.

1. Пусть первое ребро $v_1 = v$, а последнее $v_n \neq v$. Рассмотрим все вхождения ребра v:
$$s = vy_{j_1} \dots vy_{j_2} \dots vy_{j_3} \dots v_n.$$
Пусть $f(s) = s'$.
$$s' = y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots v_{i_3 j_3} y_{j_3} \dots v_n.$$
Рассмотрим передние образующие ребра в s и s''. В пути s в парах вида vy_j окажется выбранным ребро y_j, кроме первой пары, где выбраны будут оба ребра. В s' в парах $v_{ij} y_j$ окажутся выбраны ребра v_{ij}, при этом первое ребро y_{j_1} тоже будет передним образующим. Остальные передние образующие ребра в обоих путях одни и те же. Таким образом,
$$F(s) = F(v)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$
$$F(s') = F(y_{j_1}) \dots F(v_{i_2 j_2}) \dots F(v_{i_3 j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$
Фрагменты, отмеченные многоточиями, одинаковы в обоих наборах.

2. Пусть $v_1 = v$, $v_n = v$. Аналогично первому случаю, рассмотрим вхождения v в реберную запись пути s.
$$s = vy_{j_1} \dots vy_{j_2} \dots vy_{j_3} \dots x_{i_k} v.$$
$$s' = f(s) = y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots v_{i_3 j_3} y_{j_3} \dots x_{i_k}.$$
Тогда
$$F(s) = F(v)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$$F(s') = F(y_{j_1}) \dots F(v_{i_2 j_2}) \dots F(v_{i_3 j_3}) \dots = VY_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

Как мы видим, слова совпадают.

3. Пусть $v_1 \neq v$, $v_n = v$. Аналогично, рассмотрим вхождения v в реберную запись пути s.

$$s = v_1 \dots vy_{j_1} \dots vy_{j_2} \dots x_{i_k} v.$$

$$s' = f(s) = v_1 \dots v_{i_1 j_1} y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots x_{i_k}.$$

Тогда

$$F(s) = F(v_1)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = F(v_1)Y_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$$F(s') = F(v_1)F(v_{i_1 j_1}) \dots F(v_{i_2 j_2}) \dots F(v_{i_3 j_3}) \dots = F(v_1)Y_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

4. Пусть $v_1 \neq v$, $v_n \neq v$. Рассмотрим реберную запись пути s.

$$s = v_1 \dots vy_{j_1} \dots vy_{j_2} \dots v_n.$$

$$s' = f(s) = v_1 \dots v_{i_1 j_1} y_{j_1} \dots v_{i_2 j_2} y_{j_2} \dots v_n.$$

Тогда

$$F(s) = F(v_1)F(y_{j_1}) \dots F(y_{j_2}) \dots F(y_{j_3}) \dots = F(v_1)Y_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$$F(s') = F(v_1)F(v_{i_1 j_1}) \dots F(v_{i_2 j_2}) \dots F(v_{i_3 j_3}) \dots = F(v_1)Y_{j_1} \dots Y_{j_2} \dots Y_{j_3} \dots$$

$\square$

Предложение 4.3.2. *Если в графе S два симметричных пути $s_1 \neq v$, $s_2 \neq v$ и $s_2 \sqsubseteq_n s_1$, то $f(s_1) \sqsubseteq_n f(s_2)$.*

Доказательство. Напомним, что v — это выделенное опорное ребро. Рассмотрим реберные записи путей s_1 и s_2. Вторая реберная запись входит в первую не менее n раз.

Соответствие f делает с этими записями следующее: каждое вхождение “v” либо отбрасываются, если оно находилось с краю реберной записи, либо меняется на некоторою ребро, зависящее только от соседей “v” в записи справа и слева.

Рассмотрим некое вхождение второй реберной записи в первую. Все буквы-ребра, не являющиеся “v”, при соответствии f не меняются. Если буква “v”

находится не на конце реберной записи s_2, то в обеих реберных записях она меняется на одну и ту же букву. Если "v" находится в начале или конце второго слова, при соответствии f во втором слове она исчезает.

Следовательно, при отображении s подпоследовательность реберной записи пути s_1, являющаяся реберной записью s_2, переходит в подпоследовательность реберной записи s', содержащую реберную запись s_2. Кроме того, так как $s_2 \neq v$, в реберной записи s_2 есть первое с начала ребро, не меняющееся при отображении s (то есть не являющееся v). Для каждого вхождения второй реберной записи в первую отметим это ребро. Рассмотрим образ первого слова при соответствии f. Отмеченные ребра передут в отмеченные ребра, стоящие в различным местах реберной записи пути $f(s)$. При этом образ каждого из отмеченных ребер будет началом реберной записи $f(s_2)$. Таким образом, в реберной записи первого слова будет не менее n подслов, являющихся реберной записью пути $f(s_2)$. $\square$

Доказанное утверждение вкупе с предложением 4.3.1 моментально влечет тот факт, что для графа S' выполняется свойство 4 схем Рози.

Далее, из построения S' очевидно, что для любой раздающей вершины передние слова ребер, выходящих из нее, начинаются с различных букв, а задние слова ребер, входящих в одну собирающую вершину, кончаются на различные буквы.

Предложение 4.3.3. *Для графа со словами S' выполнены свойства 1 — 4 схем Рози.*

Доказательство. Осталось доказать свойство 1. В силу сильной связности S, в S существует цикл, проходящий по всем ребрам S, а также по всем тройкам ребер вида $x_i v y_j$. Если применить к нему отображение f, то получится цикл в S', проходящий по всем ребрам. То, что S' — не цикл, очевидно, так как у него есть раздающие ребра. $\square$

Лемма 4.3.1. *В графе S существует допустимый путь, который можно разбить на три куска так, что первый и третий проходят по всем ребрам и всем "хорошим" тройкам ребер.*

Доказательство. Составим $\{u_i\}$ — набор подслов сверхслова W. Для каждого ребра $v \in S$ включим в этот набор слово u_s — слово, соответствующее этому ребру по свойству 7 для схем Рози. Кроме того, включим в набор те слова вида X_iVY_j, которые являются подсловами W. В сверхслове W существует подслово U_1, содержащее все слова набора $\{u_i\}$. Пусть $l_{\max}$ — максимальная из длин слов на ребрах S. В W найдется подслово U_2 вида U_1wU_1, где $|w| > 10l_{\max}$.

Применим к U_2 лемму 4.2.3. Получим допустимый путь s с реберной записью $v_1v_2\dots v_n$. Слово $F(s)$ имеет вид $w_1U_1wU_1w_2$. Множество симметричных путей, являющихся началами пути s, можно упорядочить по включению:

$$s_1 \sqsubseteq s_2 \sqsubseteq \dots \sqsubseteq s_k = s.$$

Несложно понять, что у пути s столько же передних образующих ребер, сколько у него начал-симметричных путей. Если $v_{i_1}, v_{i_2}, \dots, v_{i_k}$ — передние образующие для пути s, то

$$F(s_j) = F(v_{i_1})F(v_{i_2})\dots F(v_{i_j}).$$

Значит, $F(s_{i+1}) - F(s_i) \leq l_{\max}$ для любого i.

Рассмотрим минимальное i такое, что $w_1U_1 \sqsubseteq F(s_i)$. Реберная запись $p_1 = s_i$ имеет вид $p_1 = v_1v_2\dots v_{k_1}$. Очевидно, $F(p_1) = w_1U_1w_1'$ для некоторого w_1'. При этом $|w_1'| \leq l_{\max}$, иначе слово $F(s_{i-1})$, являющееся началом $F(s_i)$ длиной не менее $|F(s_i)| - l_{\max}$, содержало бы w_1U_1.

Аналогично, существует такой симметричный путь p_2, что его реберная запись имеет вид $v_{k_2}v_{k_2+1}, \dots v_n$, а его слово $F(s') = w_2'U_1w_2$, где $|w_2'| \leq l_{\max}$.

Итак, у нас есть пути $p_1 = v_1v_2\dots v_{k_1}$ и $p_2 = v_{k_2}v_{k_2+1}, \dots v_n$, являющиеся соответственно началом и концом пути $s = v_1v_2\dots v_n$. Докажем, что $k_1 < k_2$. Предположим противное: $k_1 \geq k_2$. Введем обозначения: $A = B(v_1v_2\dots v_{k_2-1})$, $B = F(v_{k_2}\dots v_{k_1})$ (этот путь симметричен), $C = F(v_{k_1+1}v_{k_1+2}, \dots v_n)$. Тогда слова путей $v_1v_2\dots v_{k_1}$, $v_{k_2}v_{k_2+1}, \dots v_n$ и $v_1v_2\dots v_n$ — это AB, BC и ABC соответственно. Следовательно, $F(s) < F(p_1) + F(p_2)$. С другой стороны,

$$F(s) = |w_1U_1wU_1w_2| = |w_1U_1w_1'| + |w_2'U_1w_2| + (|w| - |w_1'| - |w_2'|) \geq$$

$$\geq F(p_1) + F(p_2) + 8l_{\max}.$$

Мы получили противоречие. Стало быть, верно неравенство $k_2 > k_1$.

Так как $U \sqsubseteq F(p_1)$, то путь p_1 проходит по всем ребрам S. Докажем, он проходит и по всем хорошим тройкам ребер. В самом деле: если (x_{i_1}, y_{i_2}) — хорошая пара ребер, то слово $X_{i_1}VY_{i_2}$ является подсловом U_1. С другой стороны, $X_{i_1}VY_{i_2}$ является словом симметричного пути, образованного конкатенацией естественного продолжения влево ребра x_{i_1}, ребра v и естественного продолжения вправо ребра y_{i_2}. Осталось воспользоваться свойством 4 схем Рози, примененным к схеме S.

То же самое верно и для пути p_2.

Если симметричный путь проходит по тройке ребер $x_{i_1}vy_{i_2}$ для плохой пары (x_{i_1}, y_{i_2}), то слово этого пути содержит $X_{i_1}VY_{i_2}$ в качестве подслова и путь не является допустимым. Следовательно, путь s, как допустимый, не проходит по плохим тройкам ребер.

□

Лемма 4.3.2. *Граф S'' сильносвязен. Если он является циклом, то сверхслово W периодично и его период определяется по схеме S, опорному ребру и множеству плохих пар ребер.*

Доказательство. Рассмотрим в графе S' путь $f(s)$, где s — путь, построенный в лемме 4.3.1. Этот симметричный путь проходит только по хорошим ребрам графа S'. В нем можно выделить начало $f(p_1)$, конец $f(p_2)$ и среднюю часть, причем начало и конец проходят по всем хорошим ребрам графа S'. Пусть d и e — ребра схемы S''. Им соответствуют пути в S'. Ребру d — путь $d_1d_2\dots d_{n_1}$, ребру e — путь $e_1e_2\dots e_{n_2}$. Путь $f(p_1)$ проходит по ребру d_1. Из концов каждого из ребер $d_1, d_2, \dots, d_{n_1-1}$ идет ровно одно хорошее ребро. Следовательно, в пути s после ребра d_1 обязаны идти ребра d_2,d_3,$\dots$, d_{n_1}. Аналогично, в p_2 есть вхождение ребра e_{n_2}, а перед вхождением e_{n_2} в $f(s)$ обязана идти последовательность ребер $e_1e_2\dots e_{n_2-1}$. Часть пути $f(s)$, начинающуюся с $d_1d_2\dots d_{n_1}$ и кончающуюся на $e_1e_2\dots e_{n_2}$, можно разбить на последовательные группы ребер, образующие ребра графа S''.

Таким образом, по ребрам графа S'' можно добраться от любого ребра до любого другого. Следовательно, граф S'' сильносвязен.

Предположим, что S'' — цикл. Тогда из каждой вершины графа S' выходит ровно одно хорошее ребро. Пусть всего хороших ребер n. Рассмотрим какое-нибудь хорошее ребро, которое в схеме S' выходит из раздающей вершины, и назовем его d_1. Такие ребра существуют, например, сойдет ребро вида v_{ij}. В схеме S' хорошие ребра образуют цикл $d_1d_2\dots d_n$. Возьмем w_0 — произвольное подслово W. В S существует симметричный путь $v_1v_2\dots v_n$ такой, что $w_0 \sqsubseteq F(v_1v_2\dots v_n) \sqsubseteq W$. Соответствующий путь в S' должен проходит только по хорошим ребрам — следовательно, этот путь имеет вид $d_{\text{begin}}(d_1d_2\dots d_e)^k d_{\text{end}}$, где участки d_{begin} и d_{end} состоят менее чем из n ребер. Пусть переднее слово пути $d_1d_2\dots d_n$ — это D. Тогда w_0 является подсловом слова $D_{\text{begin}}D^kD_{\text{end}}$, где длина слов D_{begin} и D_{end} не превосходит $nl_{\max}$. Пусть N — длина слова D. Тогда любое подслово сверхслова W длины N является циклическим сдвигом D (если D' — подслово W длины N, выберем w_0, содержащее D в середине и имеющее длину $N + 10 \cdot nl_{\max}$). Таким образом, W периодично с периодом D. □

Предложение 4.3.4. *Для S'' выполнено свойство 2 схем Рози.*

Доказательство. В самом деле, ребра, выходящие из одной раздающей вершины в графе S'', соответствуют путям, выходящим из одной раздающей вершины графа S' и имеющим различные первые ребра. Стало быть, передние слова ребер в S'' — передние слова некоторых путей, выходящих из одной раздающей вершины графа S' и имеющих различные первые ребра. Передние слова этих путей начинаются с передних слов соответствующих первых ребер, первые буквы которых, согласно предложению 3.4.4, различны. □

Так как симметричные пути в S'' являются симметричными путями в S' с теми же словами, а для графа S' выполнены свойства 3 и 4, то эти же свойства выполнены и для графа S''.

Предложение 4.3.5. *Для S'' выполнено свойство 5.*

Доказательство. Рассмотрим в графе S' симметричный путь $f(s)$, где s — путь, построенный в лемме 4.3.1. Слово этого пути — подслово сверхслова W. Если ребру d графа S'' в S' соответствует путь $d_1d_2\dots d_k$, то $d_1d_2\dots d_k \sqsubseteq f(s)$. Из следствия 4.2.1 имеем $F(d_1d_2\dots d_k) \sqsubseteq F(f(s))$ а также $B(d_1d_2\dots d_k) \sqsubseteq F(f(s))$. Осталось заметить, что $F(f(s)) \sqsubseteq W$. □

Предложение 4.3.6. *Для S'' выполнено свойство 7.*

Доказательство. Рассмотрим путь $f(s)$, где s — путь из леммы 4.3.1. Докажем, что если l — симметричный путь в s и $F(f(s)) \sqsubseteq F(l)$, то для любого ребра $v \in S''$ выполняется $v \sqsubseteq l$. В самом деле: путь l является симметричным и в графе S'. По доказанному ранее свойству 4 для схемы S', путь l содержит путь $f(s)$. А стало быть, в графе S'' он проходит по всем ребрам. $\square$

Для доказательства теоремы 4.3.1 осталось проверить свойство 6. Пусть u — интересующее нас подслово сверхслова W. Каждому ребру графа S'' в S' соответствует путь, пусть все такие пути имеют длину не более N ребер, а слова всех ребер имеют в S длину не более $l_{\max}$. Рассмотрим w_0 — подслово в W вида u_1uu_2, где $|u_1| = |u_2| = 10 \cdot Nl_{\max}$. В S существует допустимый путь l, слово которого содержит w_0. Рассмотрим в S' путь $f(l)$. В этом пути содержится более $10 \cdot N$ ребер. Среди последних N ребер пути есть ребро, из конца которого выходит более одного хорошего ребра. Также среди первых N ребер пути есть ребро, в начало которого входят хотя бы два хороших ребра. Отрезок пути между этими двумя ребрами обозначим r. Путь r является симметричным, а его слово содержит слово u в качестве подслова. Кроме того, r является симметричным путем в схеме S''.

Итак, **теорема 4.3.1 доказана**.

4.4 Эволюция схем Рози.

Определение 4.4.1. Пусть S — схема Рози для сверхслова W. На ее ребрах можно написать различные натуральные числа (или пары чисел). Такую схему мы назовем *нумерованной.* Если с ребер пронумерованной схемы стереть слова, получится *облегченная нумерованная схема.*

Определение 4.4.2. *Метод эволюции* — это функция, которая каждой облегченной пронумерованной схеме (с нумерацией, допускающей двойные индексы) дает этой же схеме новую нумерацию, такую, что в ней используются числа от 1 до n для некоторого n по одному разу каждое.

Зафиксируем какой-либо метод эволюции и далее не будем его менять.

Среди опорных ребер пронумерованной схемы S возьмем ребро v с наименьшим номером, и совершим элементарную эволюцию (W, S, v). Эта эволюция либо выявит периодичность сверхслова, либо даст новую схему. Укажем естественную нумерацию новой схемы.

Напомним, что сначала строится схема S', а потом — S''. В схеме S' все ребра можно пронумеровать по следующему правилу: ребра кроме v сохраняют номера, а ребра вида v_{ij} нумеруют соответствующим двойным индексом.

Каждое ребро в схеме S'' — это некоторый путь по ребрам схемы S', различным ребрам из S'' соответствуют в S' пути с попарно различными первыми ребрами. Таким образом, ребра схемы S'' можно пронумеровать номерами первых ребер соответствующих путей S'. Теперь применим к облегченной нумерованной схеме S'' метод эволюции. Получится новая облегченная нумерованная схема, и в нумерованной (не облегченной) схеме S'' перенумеруем ребра соответственным образом.

Определение 4.4.3. Описанное выше соответствие, ставящее нумерованной схеме Рози другую нумерованную схему Рози (если не выявилась периодичность) назовем *детерменированной эволюцией.* Будем обозначать это соответствие $S'' = \mathrm{Evol}(S)$

Определение 4.4.4. Применяя детерменированную эволюцию к нумерованной схеме S много раз, получаем последовательность нумерованных схем Рози. Кроме того, на каждом шаге получаем множество пар чисел, задающие плохие пары ребер. *Протокол детермеминованной эволюции* — это последовательность таких множеств пар чисел и облегченных нумерованных схем. Если на каком-то шаге выявилась периодичность, то протокол эволюции внезапно обрывается.

Из определения элементарной эволюции следует

Предложение 4.4.1. *Облегченная нумерованная схема для* $\mathrm{Evol}(S)$ *однозначно определяется по облегченной нумерованной схеме* S *и множеству пар чисел, задающие плохие пары ребер.*

4.5 Схемы Рози слов с не более чем линейным показателем рекуррентности.

Замечание 4.5.1. *Мощность алфавита далее предполагается фиксированной.*

Определение 4.5.1. *Пословная сложность* $P(N)$ сверхслова W — количество различных подслов W длины N.

Определение 4.5.2. *Показатель рекуррентности* $P_2(N)$ для почти периодичного сверхслова W — минимальное число $P_2(N)$ такое, что в любом подслове сверхслова W длины $P_2(N)$ встретятся все подслова W длины N.

Определение 4.5.3. Число C называется ограничителем рекуррентности сверхслова W, если $P_2(N) \leq CN$.

Замечание 4.5.2. *Мы не требуем какой-либо минимальности для ограничителя рекуррентности.*

Лемма 4.5.1. *Существует такая вычислимая функция* $C_{\text{comp}} : \mathbb{N} \to \mathbb{N}$, *что если* W *— сверхслово с ограничителем рекуррентности* C_{rec}, *то* $P(N) \leq C_{\text{comp}}(C_{\text{rec}})N$.

Доказательство. Рассмотрим конечное подслово длины nC_{rec}. С одной стороны, в нем содержатся все n−буквенные подслова слова W. С другой стороны, различных подслов длины n в нем не больше, чем его длина. □

Лемма 4.5.2. *Существует такая вычислимая функция* $C_{\text{cass}} : \mathbb{N} \to \mathbb{N}$, *что если для сверхслова* W *и некоторого* C *выполнено* $P(N) \leq CN$, *то* $P(N+1) - P(N) \leq C_{\text{cass}}$.

Доказательство этого нетривиального утверждения приведено в работе [70].

Следствие 4.5.1. *[125] Существует такая вычислимая функция* $C_4 : \mathbb{N} \to \mathbb{N}$, *что если у сверхслова* W C_{rec} *— ограничитель рекуррентности, то во всех графах Рози* $G_k(W)$ *количество вершин суммарной степени более* 2 *не превосходит* $C_4(C_{\text{rec}})$.

Лемма 4.5.3. *Существует такая вычислимая функция* $C_5 : \mathbb{N} \to (0; \frac{1}{2})$, *что если для сверхслова* W *выполнены свойства*

1. $C_{\rm rec}$ *является ограничителем рекуррентности,*

2. *Некоторое подслово* u *встречается в* W *два раза со сдвигом меньшим, чем* $C_5(C_{\rm rec})|u|$,

то W *периодично с длиной периода, равной величине этого сдвига.*

Доказательство. Пусть в W есть два вхождения u длины k, а их левые концы находятся на расстоянии $l < \frac{k}{2C_{\rm rec}}$.

$$\overbrace{a_1 \dots a_l\, a_{l+1} a_{l+2} \dots a_{k'-1} a_{k'}}^{u} a_{k'+1} \dots a_{k'+l}$$
$$a_1 \dots a_l\, \underbrace{a_{l+1} a_{l+2} \dots a_{k'-1} a_{k'} a_{k'+1} \dots a_{k'+l}}_{u}$$

По очереди рассматривая оба вхождения слова u, приходим к выводу, что фрагмент $a_1 \dots a_l$ повторится подряд как минимум $[\frac{k}{l}]$ раз, что превышает $C_{\rm rec}$. Любое подслово W длины l является подсловом слова $(a_1 a_2 \dots a_l)^{C_{\rm rec}}$, что немедленно влечет периодичность W с периодом $a_1 a_2 \dots a_l$. □

Следствие 4.5.2. *Если сверхслово* W *непериодично и* $C_{\rm rec}$ *— его ограничитель рекуррентности, то, если* $u \sqsubseteq W$, *то для любых двух различных вхождений* u *в* W *их левые концы находятся на расстоянии не меньшем, чем* $C_5(C_{\rm rec})|u|$.

Определение 4.5.4. *Масштаб схемы* есть наименьшая из длин слов, написанных на опорных ребрах этой схемы.

Пусть S — схема Рози для сверхслова W. Возьмем собирающие вершины этой схемы. Из каждой собирающей вершины выходит единственное ребро, возьмем его естественное продолжение вперед. В схеме S получится набор симметричных путей $\{s_i\}$. Рассмотрим раздающие вершины схемы S. Для каждой из них возьмем естественное продолжение назад ребра, входящего в вершину. Получим набор симметричных путей $\{t_i\}$.

Лемма 4.5.4. *Существует такая вычислимая функция* $C_6 : \mathbb{N} \to \mathbb{N}$, *что если* W *— сверхслово с ограничителем рекурреентности* $C_{\rm rec}$, S *— его схема, а* $\{s_i\}$ *— построенный выше набор, то для любых двух путей* s_1, s_2 *из* $\{s_i\}$ *длины слов* $F(s_1)$ *и* $F(s_2)$ *отличаются не более, чем в* $C_6(C_{\rm rec})$ *раз.*

Доказательство. Так как $F(s_1)$ является передним словом первого ребра пути s_1, то $F(s_1) \sqsubseteq W$. Аналогично, $F(s_2) \sqsubseteq W$. Ровно одно из ребер пути s_1 входит в раздающую вершину. То же самое верно и для s_2. Следовательно, случаи $s_1 \sqsubseteq_2 s_2$ или $s_2 \sqsubseteq_2 s_1$ невозможны. Значит, по свойству 4 схем Рози, в $F(s_1)$ может быть не более одного вхождения $F(s_2)$ и наоборот. Если бы длины слов $F(s_1)$ и $F(s_2)$ различались более, чем в $2C_3$ раз, то одно из слов (например, $F(s_1)$) можно было бы разбить на два слова, каждое из которых содержало бы $F(s_2)$. $\square$

Замечание 4.5.3. *В условиях предыдущей леммы, для любых путей из $\{t_i\}$ отношение длин их слов не превосходит $C_6(C_{\text{rec}})$. Доказательство полностью аналогично.*

Следствие 4.5.3. *Пусть S — схема Рози для W, а C_{rec} — ограничитель рекуррентности. Если M — масштаб схемы S, то для любого пути из $\{s_i\}$ или $\{t_i\}$ длина его слова не превосходит $C_6(C_{\text{rec}})M$.*

Лемма 4.5.5. *Для любого пути из $\{s_i\}$, его слово является в W специальным слева и оканчивается на некоторое биспециальное слово длины не менее M, где M — масштаб схемы.*

Доказательство. Пусть рассматриваемый путь имеет реберную запись $v_1v_2\ldots v_n$. Тогда ребро v_n выходит из собирающей вершины, иначе естественное продолжение ребра v_1 было бы короче хотя бы на одно ребро. Ребро v_n является опорным.

Докажем, что слово любого опорного ребра v является биспециальным в W. Докажем, например, что $F(v)$ является специальным справа. Пусть из правого конца v выходят y_1 и y_2. По свойству 7 схем Рози для ребра y_1 есть u_{y_1} — подслово слова W такое, что любой симметричный путь, слово которого содержит u_{y_1}, проходит по ребру y_1. По лемме 4.2.3, в схеме S есть допустимый путь, слово которого содержит u_{y_1}. Реберная запись этого допустимого пути обязана содержать vy_1. Следовательно, в слово этого пути входит переднее слово пути vy_1, стало быть, $F(vy_1) \sqsubseteq W$.

Аналогично доказвается, что $F(vy_2) \sqsubseteq W$. Так как $F(vy_i) = F(v)F(y_i)$ и слова $F(y_i)$ имеют различные первые буквы по свойству 2 схем Рози, v

является специальным справа. Специальность слева доказывается абсолютно так же.

Таким образом, $F(v_n)$ — биспециальное. То, что $|F(v_n)| \geq M$, очевидно.

Теперь докажем, что слово пути $v_1 v_2 \ldots v_n$ является специальным слева. Пусть x_1 и x_2 входят в начало ребра v_1. Для ребра x_1 есть u_{x_1} — такое подслово W, что любой симметричный путь, слово которого содержит u_{x_1}, содержит x_1. В S существует допустимый путь l_{x_1}, слово которого содержит u_{x_1}. Этот путь проходит по ребру x_1. Следующие n ребер этого пути могут быть только ребра $v_1, v_2, \ldots v_n$. Рассмотрим слово $F(l_{x_1})$: $B(x_1)F(v_1 v_2 \ldots v_n) \sqsubseteq B(l_{x_1})$. Значит, $B(x_1)F(s_1 s_2 \ldots s_n) \sqsubseteq W$.

Аналогично, $B(x_2)F(v_1 v_2 \ldots v_n) \sqsubseteq W$. Пользуясь свойством 2 схем Рози для $B(x_1)$ и $B(x_2)$, получаем, что $F(v_1 v_2 \ldots v_n)$ является специальным слева. $\square$

Аналогично доказывается соответствующий факт для путей $\{t_i\}$.

Следствие 4.5.4. *Если сверхслово периодично с длиной периода N, то масштаб любой схемы Рози для этого слова меньше, чем N.*

Доказательство. Если масштаб схемы Рози равен $M \geq N$, то в сверхслове есть специальное справа подслово длины $N - 1$, чего для периодичного с периодом N слова не иожет быть. $\square$

Лемма 4.5.6. *Существует такая вычислимая функция $C_7 : \mathbb{N} \to \mathbb{N}$, что если C_{rec} — ограничитель рекуррентности сверслова W, а S — схема Рози этого сверхслова, то количество вершин в S не превосходит $C_7(C_{\text{rec}})$.*

Доказательство. Очевидно, в схеме S количество собирающих вершин равно числу путей в наборе $\{s_i\}$, количество раздающих вершин — числу путей в наборе $\{t_i\}$. Оценим число путей в $\{s_i\}$. Пусть M — масштаб схемы S. Рассмотрим для W граф Рози G_k, где $k = [\frac{M}{2}]$. Рассмотрим слово $s_1 \in \{s_i\}$. Из 4.5.3, его слово $F(s_1)$ имеет длину, не превосходящую $C_6(C_{\text{rec}})M$. Из леммы 4.5.5 следует, что последние k букв этого слова являются словом, специальным справа, а первые k букв — словом, специальным слева. При этом $F(s_1)$ является подсловом W. Все подслова в $F(s_1)$ длины $k + 1$ образуют путь в G_k, ведущий из собирающей вершины в раздающую. Пусть в этом пути обнаружился цикл длины l.

Это значит, что в W есть подслово длины $k+l$, первые k букв которого образуют то же слово, что и последние k букв. Применяя лемму 4.5.3, получим, что либо $l \geq C_5(C_{\text{rec}})k$, либо слово W периодично с периодом l. Но, так как $l < k < M$, второй случай противоречит утверждению леммы 4.5.4.

Слову $F(s_1)$ соответствует путь по ребрам графа G_k из одной специальной вершины в другую, этот путь не может попадать в одну и ту же вершину G_k менее, чем через $C_5(C_{\text{rec}})k$ шагов, а его длина не превосходит $C_6(C_{\text{rec}})M$. Значит, одну и ту же вершину графа G_k этот путь посетит не более

$$\frac{C_6(C_{\text{rec}})M}{C_5(C_{\text{rec}})k}$$

раз.

Из леммы 4.5.1, в графе G_k может быть не более $C_4(C_{\text{rec}})$ специальных вершин. Тогда посещений специальных вершин у пути будет не более

$$\frac{C_6(C_{\text{rec}})MC_4(C_{\text{rec}})}{C_5(C_{\text{rec}})k}$$

.

Если $|B|$ - мощность нашего алфавита, то всего различных путей, выходящих из специальной вершины, входящих в специальную вершину и содержащих не более $\frac{C_6(C_{\text{rec}})M}{C_5(C_{\text{rec}})k}$ (с учетом количества вхождений) специальных вершин, будет не более, чем

$$\sum_{i=1}^{\frac{C_6(C_{\text{rec}})MC_4(C_{\text{rec}})}{C_5(C_{\text{rec}})k}} C_4(C_{\text{rec}})B^i.$$

Для различных путей из $\{s_i\}$ их слова также различны, иначе, по свойству 4 схем Рози, пути бы совпадали. Значит, этим путям мы поставили в соответствие различные пути в G_k. Поэтому верна оценка

$$\#\{s_i\} \leq \sum_{i=1}^{\frac{C_6(C_{\text{rec}})MC_4(C_{\text{rec}})}{C_5(C_{\text{rec}})k}} C_4(C_{\text{rec}})B^i.$$

Аналогичная оценка выполняется для числа путей в $\{t_i\}$. $\square$

Следствие 4.5.5. *Существует такая вычислимая функция* $C_{\text{sch}} : \mathbb{N} \to \mathbb{N}$, *что если взять всевозможные сверхслова с ограничителем рекуррентности* C_{rec} *и рассмотреть все облегченные нумерованные схемы, возникающие при* *детерменированных эволюциях этих слов, то количество различных схем будет конечно и не будет превосходить $C_{\mathrm{sch}}(C_{\mathrm{rec}})$.*

Доказательство. Количество таких схем не превосходит числа различных ориентированных графов, у которых не более, чем $C_7(C_{\mathrm{rec}})$ вершинам, степени вершин не превосходят мощности алфавита (который у нас фиксирован), а ребра пронумерованы начальным отрезком натуральных чисел. $\square$

Лемма 4.5.7. *Существует такая вычислимая функция $C_8 : \mathbb{N} \to \mathbb{N}$, что если C_{rec} — ограничитель рекуррентности для сверхслова W, а S_1 и S_2 — нумерованные схемы Рози для и $S_2 = \mathrm{Evol}(S_1)$, то масштабы схем S_1 и S_2 относятся не более, чем в $C_9(C_{\mathrm{rec}})$ раз.*

Доказательство. Если v — опорное ребро схемы S_2, то ему соответствует некий симметричный путь s по ребрам схемы S_1, причем $F(s) = F(v_1)$. Путь s задается в S_1 номерами его ребер в соответствующей облегченной схеме.

Путь s (то есть номера его ребер в облегченной пронумерованной схеме для S_1) однозначно определяются по следующему набору: {Облегченная нумерованная схема S_1, множество пар плохих ребер в S_1(задаваемых номерами этих ребер), облегченная нумерованная схема S_2, номер ребра v в схеме S_2}.

Из 4.5.5 следует, что таких наборов для слова W конечное число. Более того, все такие наборы для заданного C_{rec} можно явно указать и для каждого набора можно найти длину пути s. Следовательно, можно найти такое число $C_8(C_{\mathrm{rec}})$, что для для всех S_1, S_2 и v длина соответствующего пути s не превосходит $C_8(C_{\mathrm{rec}})$.

А значит, $F(s) \leq C_8 C M$, где M — масштаб схемы S_1, иначе для некоторого опорного ребра v_0 графа S_1 будет $F(v_0) \sqsubseteq_{C_8(C_{\mathrm{rec}})+1} F(s)$ и, по свойству 4 схем Рози, $v_0 \sqsubseteq_{C_8(C_{\mathrm{rec}})+1} s$.

Неравенство в другую сторону ($|F(v)| = |F(s)| \geq M$) очевидно. $\square$

Лемма 4.5.8. *Существует такая вычислимая функция $C_{10} : \mathbb{N} \to \mathbb{N}$, что если C_{rec} — ограничитель рекуррентности для сверхслова W, то для любой схемы S этого сверхслова длины всех слов на ребрах этой схемы не превосходят $C_{10}(C_{\mathrm{rec}})M$, где M — масштаб схемы.*

Доказательство. Достаточно доказать для передних слов, для задних доказательство не будет отличаться.

Для ребер, выходящих из собирающих вершин, существование такой константы следует из леммы 4.5.4. Рассмотрим произвольное ребро v, выходящее из раздающей вершины. По свойствам схем Рози, в S есть допустимый путь l_v, проходящий через v.

Рассмотрим l'_v — минимальный симметричный подпуть этого пути, проходящий через v. Его реберная запись имеет вид

$$v_1 v_2 \dots v \dots v_n.$$

Так как l'_v — минимальный, то среди ребер $v_1 v_2 \dots v$ ровно одно выходит из собирающей вершины (а именно v_1). Также из минимальности следует, что среди ребер $v \dots v_n$ ровно одно входит в раздающую вершину (а именно v_n). Тогда l'_v проходит не более, чем по двум опорным ребрам. А так как его слово является подсловом W, то из свойства 4 для схем его длина $|F(l'_v)|$ менее, чем $3C_{\mathrm{rec}}M$. Осталось заметить, что $F(v) \sqsubseteq F(l'_v)$. □

Лемма 4.5.9. *Существует такая вычислимая функция $C_{11} : \mathbb{N} \to \mathbb{Q}_+$, что если C_{rec} — ограничитель рекуррентности для сверхслова W, S — схема Рози для W, а p — допустимый путь, то $|F(p)| \geq C_{11}(C_{\mathrm{rec}})MN$. Здесь M — масштаб схемы, а N — количество ребер пути.*

Доказательство. Пусть v_1 есть первое ребро произвольного допустимого пути s. Тогда слово $F(v_1)$ является началом $F(s)$. Естественное продолжение вправо ребра v_1 – симметричный путь, принадлежащий множеству путей $\{s_i\}$ (оно определялось перед леммой 4.5.4). Согласно лемме 4.5.5, $F(v_1)$ является специальным слева, а его длина не меньше, чем M. Аналогично доказывается, что $F(s)$ оканчивается на специальное справа слово длины не менее M.

Следовательно, слово любого допустимого пути соответствует некоторому пути по ребрам графа $G_{[\frac{M}{2}]}$, ведущему из какой-то левоспециальной вершины в какую-то правоспециальную.

Рассмотрим путь p. Среди его $N-1$ промежуточных вершин найдутся либо половина раздающих, либо половина собирающих. Не умаляя общности, будем считать, что выполнено первое. Рассмотрим множество допустимых путей, являющихся началами пути p. Их хотя бы $\frac{N-1}{2}$. Упорядочим эти пути так, чтобы слово каждого следующего пути являлось началом слова предыдущего.

$$F(p) = F(p_0) \succcurlyeq F(p_1) \succcurlyeq \ldots \succcurlyeq F(p_k), \ \ k \geq \frac{N-1}{2}.$$

Словам $F(p_i)$ в графе $G_{[\frac{M}{2}]}$ соответствуют пути l_i, начинающиеся в одной собирающей вершине, кончающиеся в раздающих вершинах и такие, что каждый следующий путь является подпутем-началом предыдущего. В $G_{[\frac{M}{2}]}$ не более $C_7(C_{\mathrm{rec}})$ специальных вершин (см. лемму 4.5.1). Следовательно, через одну из раздающих вершин путь l_0 проходит не менее $\frac{N-1}{2C_7(C_{\mathrm{rec}})}$ раз.

Из леммы 4.5.3 следует, что либо между любыми двумя последовательными посещениями одной и той же вершины графа $G_{[\frac{M}{2}]}$ в пути l_0 не менее $[\frac{M}{2}]C_5(C_{\mathrm{rec}})$ ребер, либо W периодично с периодом, меньшим $C_5(C_{\mathrm{rec}})\frac{M}{2}$. Согласно утверждению 4.5.4, второй случай невозможен (так как C_5 меньше единицы). Значит, l_0 проходит не менее, чем по $[\frac{M}{2}]C_5(C_{\mathrm{rec}})\frac{N-1}{C_7(C_{\mathrm{rec}})}$ ребрам. Осталось сказать, что длина l_0 не более $|F(s_0)|$. $\square$

Лемма 4.5.10. *Существует такая вычислимая функция* $C_{12} : \mathbb{N} \to \mathbb{N}$, *что если* C_{rec} *— ограничитель рекуррентности для сверхслова* W, S *— схема Рози масштаба* M, *а* s *— путь, проходящий по* N *ребрам, то длина слова этого пути* $F(s)$ *не превосходит* $C_{12}(C_{\mathrm{rec}})NM$.

Доказательство. Если симметричный путь проходит по N ребрам, то длина его слова не превосходит $Nl_{\max}$, где $l_{\max}$ — максимальная из длин слов на ребрах. В силу леммы 4.5.8, $l_{\max} \leq C_{10}(C_{\mathrm{rec}})M$. Лемма доказана. $\square$

4.6 Оснастки и построение алгоритма для морфического случая.

Определение 4.6.1. Пусть A является конечным подсловом W, S — схема Рози для W. Множество симметричных путей в S, слова которых являются подсловами A, обозначим $S(A)$. Если $S(A)$ непусто, в этом множестве есть максимальный элемент относительно сравнения $\sqsubseteq$, который будем называть *нерасширяемым путем.* Очевидно, для каждого пути из $S(A)$ есть путь, который содержит его и является нерасширяемым. Назовем этот путь *максимальным расширением.*

Замечание 4.6.1. *Если бы схема называлась не S, а, допустим, S', то множество путей мы бы обозначали $S'(A)$.*

Лемма 4.6.1. *Если s — нерасширяемый путь в $S(A)$, то*

$$|A| - 2l_{\max} \leq |F(s)| \leq |A|$$

, где $l_{\max}$ — максимальная из длин слов на ребрах S.

Доказательство. Неравенство $|F(s)| \leq |A|$ очевидно, так как $F(s) \sqsubseteq A$.

Пусть A имеет вид $u_1F(s)u_2$. Если $|F(s)| < |A| - 2l_{\max}$, то либо $|u_1| > l_{\max}$, либо $|u_2| > l_{\max}$. Не умаляя общности, предположим второе. По лемме 4.2.4 существует такой допустимый путь s_1, что началом его является путь s, а его слово имеет вид $F(s_1) = F(s)u_2u_3$. Пусть s имеет реберную запись $v_1v_2\dots v_n$, а s_2 — реберную запись $v_1v_2\dots v_nv_{n+1}\dots v_{n+k}$. Рассмотрим путь, образованный s и естественным продолжением вправо ребра v_{n+1}. Слово этого пути является подсловом A, следовательно, s не является нерасширяемым в $S(A)$. □

Лемма 4.6.2. *Существует такая вычислимая функция $C_{13} : \mathbb{N} \to \mathbb{N}$, что если C_{rec} — ограничитель рекуррентности для сверхслова W, то для любой схемы Рози S, симметричного пути l и слов X, Y, Z таких, что выполнены следующие условия:*

1. *Слово XYZ является подсловом W;*
2. $\min\{|X|, |Y|, |Z|\} > C_{13}(C_{\text{rec}})M$, *где M — масштаб схемы S;*
3. *Либо W не имеет периода, либо его минимальный период не менее $3MC_{10}(C_{\text{rec}})$,*

Следующие условия эквивалентны:

1. $l \in S(XYZ)$
2. *Существует такой симметричный путь l', что $l \sqsubseteq l'$, l' разбивается на три части $l' = xyz$, где xy — нерасширяемый путь для $S(XY)$, y — нерасширяемый путь для $S(Y)$, yz — нерасширяемый путь для $S(YZ)$.*

Доказательство. Выберем C_{13} таким, чтобы при всех n выполнялись неравенства:

$$C_{13}(n) \geq 6C_{10}(n);$$

$$2C_{10}(n)M < C_5(n)\big(C_{13}(n)M - 2C_{10}(n)M\big).$$

Далее, если для функции не указан аргумент, подразумевается, что он равен C_{rec}.

Сначала докажем, что 1 влечет 2.

Будем доказывать, что в качестве l' можно взять максимальное расширение l в $S(XYZ)$. По лемме 4.6.1, $F(l') \geq |XYZ| - 2l_{\max}$, а из леммы 4.5.8 следует $l_{\max} \leq C_{10}M$.

Рассмотрим в l' такое максимальное начало l_1, что $l_1 \in S(XY)$. Докажем, что l_1 — нерасширяемый в $S(XY)$. Пусть $F(l') = F(l_1)u_1$. Рассмотрим вхождение слова $F(l')$ в XYZ, как подслова u.

Так как l' — нерасширяемый в XYZ, то

$$XYZ = d_1F(l')d_2,$$

где $|d_1| \leq l_{\max}$, $|d_2| \leq l_{\max}$.

$$XYZ = d_1F(l_1)u_1d_2.$$

Есть два случая:

1. $|d_1F(l_1)| \leq |XY|$. В этом случае получаем $|d_1F(l_1)| \geq |XY| - l_{\max}$, иначе у l' есть начало l'_1 такое, что $l_1 \sqsubseteq l'_1$ и $|F(l'_1)| \leq |F(l_1)| + l_{\max}$, а стало быть, $F(l'_1) \sqsubseteq XY$.

   Тогда $|F(l_1)| \geq |XY| - 2l_{\max} \geq 2C_{13}M - 2C_{10}M$.

   Если есть два различных вхождения $F(l_1)$ в XY, то сдвиг между ними не превосходит $2l_{\max} \leq 2C_{10}M$. Так как выполнено

   $$2C_{10}M < C_5(2C_{13}M - 2C_{10}M),$$

   то сверхслово W, согласно лемме 4.5.3, периодично с периодом не превосходящим $2C_{10}M$. Противоречие. Следовательно, есть ровно одно вхождение $F(l_1)$ в XY.

Предположим, что l_1 расширяемый в XY. Тогда для некоторого симметричного пути l_1'' выполняется $l_1 \sqsubseteq l_1''$ и $F(l_1'') \sqsubseteq XY$. Можно считать, что l_1 — конец или начало l_1''.

(a) $l_1'' = l_b l_1$. Тогда $XY = e_1 B(l_b) F(l_1) e_2$.

Так как есть ровно одно вхождение $F(l_1)$ в XY, то $d_1 = e_1 B(l_b)$ и

$$XYZ = e_1 B(l_b) F(l') d_2.$$

Тогда симметричный путь $l_b l' \in S(XYZ)$ и путь l' расширяемый.

(b) $l_1'' = l_1 l_e$. Тогда $XY = e_1 F(l_1) F(l_e) e_2$. Так как у $F(l_1)$ только одно вхождение в XY, то $e_1 = d_1$ и

$$XY = d_1 F(l_1) F(l_e) e_2.$$

Следовательно, $F(l_1 l_e) \sqsubseteq XY$. Кроме того, $F(l_1 l_e)$ является началом $F(l')$. Тогда, по 2-му свойству схем Рози, $l_1 l_e$ является началом l' (иначе рассмотрим первое ребро пути $l_1 l_e$, не совпадающее с ребром пути l'). Противоречие.

2. $|d_1 F(l_1)| > |XY|$. В этом случае $d_1 F(l_1) = XY d_3$, где $|d_3| \leq |d_1| \leq l_{\max}$. Так как $F(l_1) \sqsubseteq_2 XY d_3 \sqsubseteq W$,

$$|XY d_3| \leq |XY| + C_{10} M, \ \ |F(l_1)| > |XY| - C_{10} M,$$

то слово $F(l_1)$, имеющее длину не менее $2C_{13}M - C_{10}M$, встречается со сдвигом, меньшим $2C_{10}M$. Из леммы 4.5.3 выводим противоречие.

Рассмотрим l_2 — такой максимальный конец пути l', что $l_2 \in S(YZ)$. Аналогичными рассуждениями получим, что l_2 — нерасширяемый в $S(YZ)$.

Докажем следующий **факт**: *пути l_1 и l_2 внутри l' перекрываются, то есть в путях l_1 и l_2 суммарно больше ребер, чем в пути l'.*

В самом деле, иначе возьмем s — естественное продолжение влево последнего ребра l_1. Путь s является симметричным, $F(S) \leq C_{10}M$. Тогда $|F(l')| + |F(s)| \geq F(l_1) + F(l_2)$, то есть

$$|XYZ| \geq (|XY| - 2C_{10}M) + (|YZ| - 2C_{10}M) - C_{10}M,$$

или $5C_{10}M \geq C_{13}M$, что неверно в силу выбора C_{13}.

Итак, можно считать, что $l_1 = xy$, $l_2 = yz$, $l' = xyz$. Для доказательства импликации $1 \to 2$ осталось показать, что путь y — нерасширяемый путь в $S(Y)$.

Имеем

$$XYZ = d_1B(x)F(y)F(z)d_2, \;\; XY = d_1B(x)F(y)e_1$$

и $YZ = e_2F(y)F(z)d_2$, где $\max\{|d_1|, |d_2|, |e_1|, |e_2|\} \leq l_{\max}$. Пользуясь очевидным равенством $|XYZ| + |Y| = |XY| + |YZ|$, получаем

$$|Y| = |e_1| + |e_2| + |F(y)|.$$

Так как Y оканчивается на $F(y)e_1$ и начинается на $e_2F(y)$, можно сделать вывод, что $Y = e_2F(y)e_1$.

Слово $F(y)$ имеет длину не меньшую, чем $C_{13}M - 2MC_{10}$, и, если бы оно встречалось в Y два раза, то встречалось бы со сдвигом, не превосходящим $2C_{10}M$. В силу леммы 4.5.3 и неравенства

$$C_{13}M - 2MC_{10} > 2MC_{10},$$

это влекло бы периодичность W с периодом, меньшим $2C_{10}M$. Поэтому $F(y)$ встречается в Y один раз.

Если путь y — расширяемый для $S(Y)$, то, не умаляя общности, y — начало симметричного пути yl_e такого, что $F(y)F(l_e) \sqsubseteq Y$. В таком случае, $Y = f_1F(y)F(l_e)f_2$, $e_2 = f_1$ и $e_1 = F(l_e)f_2$.

Тогда $XY = d_1B(x)F(y)F(l_e)f_2$, то есть $l_1l_e \sqsubseteq S(XY)$. Кроме того, $F(l_1l_e)$ является началом слова $F(l_1)e_1 = B(x)F(y)e_1$, которое, в свою очередь, является началом $B(x)F(y)F(z) = F(l')$. Следовательно, l_1l_e — начало l'. Противоречие.

Теперь докажем, что 2 влечет 1.

Так как xy — нерасширяемый путь в $S(XY)$, а y — нерасширяемый путь в Y, XY имеет вид $XY = d_1B(x)F(y)d_2$, а Y имеет вид $Y = e_1F(y)e_2$, причем $|d_i|, |e_i| \leq l_{\max}$. Так как $C_{10}M < C_5(C_{13}M - C_{10}M)$, то $d_2 = e_2$.

Аналогично, $YZ = f_1F(y)F(z)f_2$, где $f_1 = e_1$. Тогда

$$XYZ = d_1B(x)F(y)F(z)d_2,$$

а значит, $F(l) \sqsubseteq F(l') = F(xyz) \sqsubseteq XYZ$. $\square$

Лемма 4.6.3. *Существует такая вычислимая функция $C_{\mathrm{per}} : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$, что если C_{rec} — ограничитель рекуррентности для сверхслова W, S — схема Рози для сверхслова W, а W периодично с длиной периода, не превосходящей KM, где M — масштаб схемы, то в схеме S для любого допустимого пути его реберная запись периодична с длиной периода, не превосходящей $C_{\mathrm{per}}(C_{\mathrm{rec}}, K)$.*

Замечание 4.6.2. *Считается, что конечное слово aabaabaabaaba периодично с длиной периода* 3.

Доказательство. Пусть s — допустимый путь, длина которого достаточно велика. Найдется s' — допустимый путь, являющийся началом этого пути и такой, что длина его слова $F(s')$ лежит в промежутке $[KM; KM+C_{10}(C_{\mathrm{rec}})M]$. Если два подслова слова W начинаются на $F(s')$, то одно из них является началом другого. Следовательно, если два допустимых пути имеют общим началом s', то один из них является началом другого. Если длина некоторого подслова W не меньше $2C_{\mathrm{rec}}(K + C_{10}(C_{\mathrm{rec}}))M$, то это подслово содержит хотя бы два вхождения слова $F(s')$. Возьмем такое минимальное начало s'' пути s, что $|F(s'')| \geq 2C_{\mathrm{rec}}(K + C_{10}(C_{\mathrm{rec}}))M$. Из минимальности s'' следует, что $|F(s'')| \leq 2C_{\mathrm{rec}}(K + C_{10}(C_{\mathrm{rec}}))M + C_{10}(C_{\mathrm{rec}})M$. В таком случае, согласно лемме 4.5.9, в пути s'' не более

$$C_{\mathrm{per}} := \frac{2C_{\mathrm{rec}}(K + C_{10}(C_{\mathrm{rec}})) + C_{10}(C_{\mathrm{rec}})}{C_{11}(C_{\mathrm{rec}})}$$

ребер.

Очевидно, $s' \sqsubseteq_2 s''$. Рассмотрим реберную запись пути s. В ней на расстоянии, меньшем, чем C_{per}, встречаются две реберных записи пути s'. А так как среди любых двух допустимых путей, начинающихся на s', один является началом другого, то реберная запись s периодична с периодом, меньшим C_{per}. □

Замечание 4.6.3. *В дальнейшем запись $C_{\mathrm{per}}(C_{\mathrm{rec}})$ будет обозначать $C_{\mathrm{per}}(C_{\mathrm{rec}}, 3C_{10}(C_{\mathrm{rec}}))$.*

Определение 4.6.2. *Набор проверочных слов порядка k* — это упорядоченный набор слов $\{q_i\}$, в который входят $\psi(\varphi^k(a_i))$ для всех букв алфавита

$\{a_i\}$, а также $\psi(\varphi^k(a_i a_j))$ для всевозможных пар последовательных букв слова $\varphi^\infty(a_1)$. Каждая буква алфавита, а также все двубуквенные слова, являющиеся подсловами $\varphi^\infty(a_1)$, назовем *источниками.*

Для любого k проверочных слов порядка k столько же, сколько и источников.

Пусть $W = \psi(\varphi^\infty(a_1))$, где подстановка φ примитивная. Начнем строить алгоритм, определяющий его периодичность.

Прежде всего отметим, что все источники, согласно 4.1.2, находятся алгоритмически.

Лемма 4.6.4. *[125] По примитивному морфизму φ алгоритмически находятся такие числа $C_1 > 0$, $C_2 > 0$, что для некоторого $\lambda_0 > 1$ $C_1\lambda_0^k < |\varphi^k(a_j)| < C_2\lambda_0^k$ для любых k и буквы a_j.*

Следствие 4.6.1. *Можно найти такие положительные C_{14}, C_{15}, что в наборе проверочных слов с номером k длина любого слова q удовлетворяет двойному неравенству $C_{14}\lambda_0^k < |q| < C_{15}\lambda_0^k$.*

Лемма 4.6.5. *[60, 125] По примитивному φ и произвольному ψ алгоритмически находится число C_{rec}, являющееся ограничителем рекуррентности сверхслова $\psi(\varphi^\infty(a_1))$.*

Замечание 4.6.4. *Для всех введенных ранее функций (например, C_5) мы будем писать C_5, подразумевая $C_5(C_{\mathrm{rec}})$, где C_{rec} — вычисленный для W показатель рекуррентности.*

Определение 4.6.3. Определим *оснастку* (S, k). Здесь $k \in \mathbb{N}$ — порядок оснастки, S — пронумерованная схема Рози. Для получения оснастки берется набор проверочных слов порядка k и для каждого проверочного слова q_i рассматривается набор путей $S(q_i)$. Каждый путь в схеме задается можно задать упорядоченным набором чисел — номерами ребер.

Сама оснастка — это следующий набор информации:

1. облегченная нумерованная схема S (то есть без слов, но с циферками на ребрах),

2. Для каждого источника p_i указывается, какие пути находятся в множестве $S(\psi(\varphi^k(p_i)))$. Пути указываются на облегченной нумерованной схеме S упорядоченными наборами чисел.

Определение 4.6.4. *Размер оснастки* — максимальная длина (в ребрах) по всем путям из $S(q_i)$ для всех q_i — проверочных слов порядка k.

Замечание 4.6.5. *Глядя на оснастку, ее размер мы определить можем. А k (то есть порядок проверочных слов) — не можем.*

Лемма 4.6.6. *Можно найти такие положительные C_{16}, C_{17} и C_{18}, что для любой схемы S размер оснастки (S,k) заключен между $C_{16}\frac{\lambda_0^k}{M}-C_{17}$ и $C_{18}\frac{\lambda_0^k}{M}$.*

Доказательство. Длина каждого проверочного слова, согласно 4.6.1, хотя бы $C_{14}\lambda_0^k$. Если q_i — проверочное, то длина слова нерасширяемого в $S(q_i)$ пути не менее $C_{14}\lambda_0^k - 2C_{10}M$. А длина этого пути в ребрах составляет, согласно лемме 4.6.1, не менее $\frac{C_{14}\lambda_0^k-2C_{10}M}{C_{12}M}$.

С другой стороны, длина слова нерасширяемого пути не может быть больше $C_{15}\lambda_0^k$, а его длина в ребрах, как следует из леммы 4.5.9, не может быть более $\frac{C_{15}\lambda_0^k}{C_{11}M}$. □

Следствие 4.6.2. *Можно указать такие C_{19}, C_{20} и C_{21}, что если размер оснастки (S,k) больше C_{19}, то*

1. *Размер оснастки не меньше, чем $2C_{\mathrm{per}}C$.*

2. *Длины всех проверочных слов составляют не менее $C_{13}M$.*

3. *Для любой хорошей тройки ребер (то есть хорошей пары ребер, соединенной опорным ребром) существует проходящий через нее путь, слово которого содержится во всех проверочных словах.*

4. *Размер оснастки (S,k) относится к размеру оснастки $(\mathrm{Evol}(S),k)$ не более, чем в C_{20} раз (если, конечно, $\mathrm{Evol}(S)$ существует.)*

5. *Размеры оснасток (S,k) и $(S,k+1)$ отличаются не более, чем в C_{20} раз.*

6. *Размер оснастки $(S, k+C_{21})$ больше размера оснастки (S,k) хотя бы в два раза.*

Доказательство. 1. Достаточно взять большой размер оснастки.

2. Из лемм 4.6.6 и 4.6.1 следует, что если x — размер оснастки, а $|q|$ — размер проверочного слова q, то $|q| > C_{14}\lambda_0^k > C_{14}(C_{18}xM)$.

3. Симметричный путь, получающийся естественным расширением хорошей тройки ребер сначала вправо, а потом влево, является допустимым и его слово имеет длину не более $3l_{\max} \leq 3C_{10}M$. Если проверочное слово q имеет длину хотя бы $3C_{10}MC$, где C — показатель рекуррентности, то подсловом q является и слово рассматриваемого пути. Как видно из предыдущего пункта, выбором C_{19} этого легко добиться.

4. Пусть M и M_1 — масштабы схемы S и $\mathrm{Evol}(S)$, а x и x_1 — размеры оснасток. Согласно лемме 4.5.7, $M_1 \leq C_8M$. Тогда $x_1 > C_{16}\frac{\lambda_0^k}{C_8M} - C_{17} > C_{16}\frac{C_{18}x}{C_8} - C_{17}$, что при достаточно большом x больше, чем $C_{16}\frac{C_{18}x}{2C_8}$.

5. Пусть x и x_1 — размеры оснасток (S,k) и $(S,k+1)$.
Тогда $C_{16}\frac{\lambda_0^{k+1}}{M} - C_{17} < x_1 < C_{18}\frac{\lambda_0^{k+1}}{M}$. С другой стороны, $\frac{x+C_{17}}{C_{16}} > \frac{\lambda_0^k}{M} > \frac{x}{C_{18}}$.
Таким образом, $C_{16}\frac{\lambda_0 x}{C_{18}M} - C_{17} < x_1 < C_{18}\frac{\lambda_0(x+C_{17})}{C_{16}M}$.

6. Пусть x и x_1 — размеры оснасток (S,k) и $(S,k+C_{21})$.
Тогда $C_{16}\frac{\lambda_0^{k+C_{21}}}{M} - C_{17} < x_1$. С другой стороны, $\frac{\lambda_0^k}{M} > \frac{x}{C_{18}}$.

   Таким образом, $x_1 > C_{16}\frac{\lambda_0^{C_{21}}x}{C_{18}M} - C_{17}$, что не меньше $2x$ при достаточно большом C_{21}.

   Заметим, что точное значение λ_0 мы не знаем, но из леммы 4.6.4 мы можем найти какую-то оценку снизу, отделяющую его от 1, и какую-то оценку сверху.

□

Лемма 4.6.7. *Если размер оснастки (S,k) не меньше, чем C_{19}, то, зная оснастку (S,k), можно сделать хотя бы один из двух выводов:*

1. *Сверхслово W является периодичным и длина периода не превосходит $MC_{\text{per}}C_{10}$, где M — масштаб схемы S. (Заметим, что по оснастке оценить M мы не можем, так как в оснастке используется облегченная схема).*

2. *Вывод о том, какова оснастка $(S, k+1)$.*

Доказательство. Пусть R — размер оснастки (S, k). Среди путей оснастки есть путь, у которого $R \geq 2C_{\text{per}}C$ ребер. Посмотрим на его реберную запись и определим, является ли она периодичной с длиной периода, не превосходящей C_{per}. Если является, то в W встречается u^C для некоторого u, при этом длина $|u| \leq MC_{\text{per}}C_{10}$. Можно сделать первый вывод (о переодичности), так как C любое слово длины $|u|$ встречается в u^C.

Далее считаем, что реберная запись этого пути не является периодичной. В таком случае из леммы 4.6.3 следует, что либо W непериодично, либо его минимальный период больше, чем $3C_{10}M$, то есть выполнено третье условие леммы 4.6.2.

Для того, чтобы узнать оснастку $(S, k+1)$, нужно для каждого источника p_i определить множество путей в $S(\psi(\varphi^{k+1}(p_i)))$. Согласно лемме 4.6.2, размер оснастки $(S, k+1)$ не будет больше $C_{20}R$. Поэтому достаточно проверить принадлежность к $S(\psi(\varphi^{k+1}(p_i)))$ для каждого симметричного пути с не более, чем $C_{20}R$ ребрами.

Пусть $\varphi(p_i) = a_{i_1}a_{i_2}\ldots a_{i_k}$.

$$\psi(\varphi^{k+1}(p_i)) = \psi(\varphi^k(a_{i_1}))\psi(\varphi^k(a_{i_2}))\ldots\psi(\varphi^k(a_{i_k})).$$

Каждая из букв a_{i_1}, a_{i_2}, …, a_{i_k} является источником. Также источниками являются $a_{i_1}a_{i_2}$, $a_{i_2}a_{i_3}$,…, $a_{i_{k-1}}a_{i_k}$.

Согласно лемме 4.6.2, длины слов $\psi(\varphi^k(a_{i_1}))$, $\psi(\varphi^k(a_{i_2}))$, $\psi(\varphi^k(a_{i_3}))$ не менее $C_{13}M$, поэтому для схемы S, слов $\psi(\varphi^k(a_{i_1}))$, $\psi(\varphi^k(a_{i_2}))$, $\psi(\varphi^k(a_{i_3}))$ и произвольного симметричного пути l выполняются условия леммы 4.6.2.

Из оснастки (S, k) нам известны следующие множества путей:

$$S(\psi(\varphi^k(a_{i_1}))\psi(\varphi^k(a_{i_2}))),\ S(\psi(\varphi^k(a_{i_2}))\psi(\varphi^k(a_{i_3}))),\ S(\psi(\varphi^k(a_{i_2}))).$$

В каждом из этих множеств находим максимальные по включению элементы — нерасширяемые пути.

Симметричный путь l принадлежит множеству $S(\psi(\varphi^k(a_{i_1}a_{i_2}a_{i_3})))$ тогда и только тогда, когда *существует такой симметричный путь* l', *что* $l \sqsubseteq l'$, *а* l' *разбивается на три части* x,y *и* z *такие, что* xy *— нерасширяемый путь для* $S(\psi(\varphi^k(a_{i_1}a_{i_2})))$, yz *— для* $S(\psi(\varphi^k(a_{i_2}a_{i_3})))$, *а* y *— для* $S(\psi(\varphi^k(a_{i_2})))$. А это свойство легко проверяется по оснастке (S,k).

После определения множества $S(\psi(\varphi^k(a_{i_1}a_{i_2}a_{i_3})))$ воспользуемся леммой 4.6.2, примененной к словам $\psi(\varphi^k(a_{i_1}a_{i_2}))$, $\psi(\varphi^k(a_{i_3}))$, $\psi(\varphi^k(a_{i_4}))$ и аналогичным способом определим множество $S(\psi(\varphi^k(a_{i_1}a_{i_2}a_{i_3}a_{i_4})))$.

Потом определим множество $S(\psi(\varphi^k(a_{i_1}a_{i_2}a_{i_3}a_{i_4}a_{i_5})))$ и, действуя подобным образом, доберемся до $S(\psi(\varphi^k(a_{i_1}a_{i_2}\ldots a_{i_k})))$.

$\square$

Лемма 4.6.8. *Если размер оснастки* (S,k) *не менее* C_{19}, *то по оснастке* (S,k) *однозначно определяются плохие пары ребер, а также определяется, выявит ли элементарная эволюция периодичность слова* W. *Если не выявит, то определяется оснастка* $(\mathrm{Evol}(S),k)$.

Доказательство. Из леммы 4.6.2 следует, что по оснастке определяется множество хороших и плохих пар ребер: по хорошим тройкам ребер пути из оснастки проходят, а по плохим — нет (так как все пути в оснастках допустимые). Следовательно, определяется, существует ли $\mathrm{Evol}(S)$ и какая у $\mathrm{Evol}(S)$ (в случае существования) облегченная нумерованная схема. Симметричные пути в $\mathrm{Evol}(S)$ соответствуют некоторым симметричным путям в S с теми же словами. При этом мы можем указать соответствие, глядя лишь на облегченные схемы. Таким образом, для каждого симметричного пути в облегченной нумерованной схеме $\mathrm{Evol}(S)$ по оснастке (S,k) можно определить, каким из $\mathrm{Evol}(S)(A_i)$ этот путь принадлежит. (Здесь A_i — проверочные слова порядка k). $\square$

Симметричным путям в $\mathrm{Evol}(S)$ соответствуют симметричные пути в S с теми же словами, при этом в $\mathrm{Evol}(S)$ пути не длиннее, чем соответственные пути в S. Стало быть, размер оснастки $(\mathrm{Evol}(S),k)$ не более, чем размер оснастки (S,k).

Запустим основную часть алгоритма. Возьмем $T = 2C_{19}C_{20}^{C_{21}}$. Построим какую-нибудь схему Рози и возьмем такой порядок проверочных слов, чтобы размер оснастки был не менее T. Далее будем работать только с оснастками. Первая оснастка у нас есть, а каждая следующая оснастка определяется по предыдущей. Если размер оснастки менее T, то по оснастке вида (S, k) строится оснастка вида $(S, k+1)$ (операция перехода описана в лемме 4.6.7). Иначе делается операция перехода от оснастки вида (S, k) к оснастке вида $(\mathrm{Evol}(S), k)$ (операция описана в лемме 4.6.8). Таким образом, размер оснастки никогда не упадет ниже C_{19}. С другой стороны, размер оснастки не поднимается выше $C_{20}T$.

Значит, различных оснасток в последовательности не может быть больше, чем некоторое наперед заданное число H (так как различных нумерованных облегченных схем тоже конечно.) Следовательно, есть альтернатива: либо не более чем за H шагов последовательность оснасток зациклится, либо она оборвется не позднее шага с номером H.

Если она обрывается, значит, либо в одном из переходов согласно лемме 4.6.7 был сделан вывод о периодичности, либо в одном из переходов согласно лемме 4.6.8 было обнаружено, что элементарная эволюция выявляет периодичность. И в том и в другом случае, проследив переходы на необлегченных схемах, легко обнаружить период сверхслова W.

Если же последовательность зацикливается, то последовательность оснасток периодична (так как каждая следующая определяется по предыдущей). Если с некоторого момента совершаются только переходы типа $(S, k) \to (S, k+1)$, то размер оснасток неограниченно увеличивается, чего не может быть. Следовательно, совершается бесконечно много переходов типа $(S, k) \to (\mathrm{Evol}(S), k)$. То есть у сверхслова W бесконечно много схем Рози. Докажем, что оно непериодично.

В самом деле, пусть оно периодично. Тогда длины всех слов ра ребрах всех схем одновременно ограничены. Так как облегченных схем конечное число, то какая-то схема (как граф со словами) должна повториться дважды. Но у каждой следующей схемы множество слов всех симметричных путей согласно 4.3.1 строго содержится в аналогичном множестве для предыдущей схемы.

Таким образом, проблема периодичности для примитивных морфических слов разрешима и теорема 4.0.1 доказана.

Глава 5

Почти периодичность морфических последовательностей

Сформулируем задачу проверки почти периодичности для морфических слов, поставленную в [107] и [41]:

Дано: *два конечных алфавита A и B, буква $a_1 \in A$, подстановка $\varphi : A^* \to A^*$, продолжающаяся над a_1 и морфизм $\psi : A^* \to B^*$.*

Определить: *является ли слово $\psi(\varphi^\infty(a_1))$ почти периодичным?*

В этой главе строится алгоритм для произвольного морфического слова:

Теорема 5.0.1. *Проблема почти периодичности для морфических слов разрешима.*

5.1 Приведение морфизмов к удобному виду.

Напомним *проблему заключительной периодичности для морфических слов.*

Дано: *два конечных алфавита A и B, буква $a_1 \in A$, подстановка $\varphi : A^* \to A^*$, продолжающаяся над a_1 и морфизм $\psi : A^* \to B^*$.*

Определить: *является ли слово $\psi(\varphi^\infty(a_1))$ заключительно периодическим. Если является, явно указать его период.*

Теорема 5.1.1. *Проблема заключительной периодичности для морфических слов разрешима.*

Эта теорема была доказана в разделе 4.

Теорема 5.1.2 (доказательство см., например, в предыдущем разделе). *Пусть φ — подстановка, действующая на алфавите A, ψ — морфизм из A^* в B^*, u — конечное слово из B^*. Тогда существует алгоритм, проверяющий, встречается ли u в слове $W = \psi(\varphi^\infty(a_1))$ и, если встречается, конечное ли число раз.*

Предложение 5.1.1. *Если почти периодичное сверхслово W для некоторого непустого U и любого натурального k содержит подслово U^k, то оно является число периодическим с периодом u.*

Доказательство. Это следует из того, что если $|V| \sqsubseteq W$ и $|V| = |U|$, то V является циклическим сдвигом слова U. □

Мы хотим проверить на почти периодичность сверхслово $W = \psi(\varphi^\infty(a_1))$. Морфизм φ можно считать нестирающим, а h — кодированием (то есть для любой буквы $a \in A$ $|\varphi(a)| > 0$, $|\psi(a)| = 1$). Это следует из результата

Теорема 5.1.3 (см. [60], глава 7). *Если $f : A^* \to B^*$ и $g : A^* \to A^*$ — произвольные морфизмы и $f(g^\infty(a_1))$ – бесконечное слово, то можно найти такие алфавит A', букву $a_1' \in A'$, нестирающую подстановку φ, действующую на алфавите A и кодирование $\tau : A' \to B$, что $f(g^\infty(a_1)) = \tau(\varphi^\infty(a_1'))$.*

Слово $w \in A^*$ будем называть *φ—ограниченным*, если последовательность

$$w, \varphi(w), \varphi^2(w), \varphi^3(w), \ldots$$

периодична начиная с некоторого момента. В противном случае,

$$|\varphi^n(w)| \to \infty$$

при $n \to \infty$ и слово w называется *φ-растущим.* Очевидно, слово является φ-ограниченным тогда и только тогда, когда оно состоит из φ-ограниченных букв.

Теорема 5.1.4. *Существует алгоритм, который определяет, конечно ли в $\varphi^\infty(a)$ число различных φ—ограниченных подслов. Если это число бесконечно, то можно указать такое непустое U, что U^k является подсловом $\varphi^\infty(a)$ для любого $k \in \mathbb{N}$. Если оно конечно, то все φ—ограниченные слова алгоритмически находятся.*

Доказательство. Прежде всего отметим, что все $\varphi-$возрастающие буквы алгоритмически находятся (см., например, [107]). Далее φ-растущие буквы будем писать как a_1, a_2 и т.д.

Построим ориентированный граф Q, на ребрах которого будут записаны упорядоченные пары слов. Вершинами этого графа будут служить φ-растущие буквы из A а также всевозможные упорядоченные пары φ-растущих букв. Введем фиктивную букву t, также к вершинам Q добавим всевозможные пары вида a_it, где a_i — φ-растущая буква.

Из вершины a_i в a_j идет ребро, если $a_j \in a_i$. На таких ребрах пара слов — $\{\varepsilon, \varepsilon\}$ ($\varepsilon-$ пустое слово). Из вершины a_i в a_ja_k ведет ребро со словами $\{\omega, \varepsilon\}$, если для некоторого $\varphi-$ограниченного слова ω слово $a_j\omega a_k$ является подсловом $\varphi(a_i)$ (из a_i в a_ja_k могут вести несколько ребер.) Из a_i и a_it ведет по ребру с парой $\{\omega, \varepsilon\}$ в a_jt, если ω — $\varphi-$ограниченное и $\varphi(a_i)$ оканчивается на $a_j\omega$.

Из a_ia_j ведет ребро в a_ka_l с парой $\varphi-$ограниченных слов $\{\omega_1; \omega_2\}$, если $\varphi(a_1)$ кончается на $a_k\omega_1$, а $\varphi(a_2)$ начинается на ω_2a_l.

Предложение 5.1.2. *Пусть $k \in \mathbb{N}$. Рассмотрим какой-нибудь путь длины k по ребрам графа Q, выходящий из a_1. Последовательность пар слов на ребрах этого пути*

$$\{u_1, v_1\}, \{u_2, v_2\}, \dots, \{u_k, v_k\}.$$

Тогда в $\varphi^k(a_1)$ есть $\varphi-$нерасширяемое слово

$$u_k\varphi(u_{k-1})\dots\varphi^{k-1}(u_1)\varphi^{k-1}(v_1)\varphi^{k-2}(v_2)\dots v_k.$$

При этом, если путь оканчивается на a_ia_j, то есть вхождение этого слова, обрамленное буквами a_i и a_j.

Наоборот, любое $\varphi-$ограниченное подслово $\varphi^k(a_1)$ можнно получить, получив слово по указанному правилу слово и взяв его подслово.

Это утверждение несложно показывается индукцией по k.

Первый случай: в любом ориентированном цикле графа Q, до которого можно добраться из a_1, на ребрах цикла написаны пары пустых слов. Тогда в любом пути, выходящем из вершины a_1, число ребер, на которых написаны не пустые слова, не превосходит количества вершин в Q. Следовательно, число различных φ—ограниченных слов конечно, а по графу Q можно их всех найти.

Второй случай: есть цикл, и в этом цикле не все слова пустые. Пусть, например, в цикле есть пары слов, в которых первое слово не пустое. Тогда для некоторой буквы a_i и непустого φ—ограниченного u, слово $\varphi(a_i)$ оканчивается на $a_i u$. Образ u при подстановке φ зациклится. Следовательно, для некоторого непустого слова U для любого $k \in \mathbb{N}$ слово U^k является подсловом $\varphi^\infty(a)$. $\square$

Применим этот алгоритм к φ и a_1. Если некоторое слово U повторяется в $\varphi(a_1)$ сколь угодно много раз подряд, то $\psi(U)$ повторяется сколь угодно много раз подряд в W. Согласно 5.1.1, для установления почти периодичности W достаточно проверить, является ли оно чисто периодическим с периодом $\psi(U)$.

Предложение 5.1.3. *Бесконечное слово является чисто периодическим с заданным периодом A тогда и только когда, когда все его конечные подслова длины $|A|$ являются циклическими сдвигами A.*

Согласно 4.1.2, все конечные подслова W длины $|\psi(U)| = |U|$ можно найти, стало быть, в этом случае определить, является ли W почти периодичным, мы можем.

В дальнейшем рассматриваем случай, когда в $\varphi^\infty(a_1)$ конечное число φ—ограниченных подслов.

Пусть I_φ — множество всех φ-растущих букв, B_φ — множество φ-ограниченных подслов сверхслова $\varphi^\infty(a_1)$(включая пустое слово). Можно считать, что B_φ конечно и что мы знаем все слова в B_φ. Рассмотрим (конечный) алфавит C, состоящий из символов $[twt']$, где t и t' буквы из I_φ, а w — слово из B_φ и слово twt' является подсловом $\varphi^\infty(a)$.

Определим морфизм $\varphi' : C^* \to C^*$ следующим образом:

$$\varphi'([twt']) = [t_1 w t_2][t_2 w t_3] \ldots [t_k w_k t_{k+1}],$$

где $\varphi(tw) = w_0 t_1 w_1 t_2 \ldots t_k w'_k$, слово $\varphi(t')$ начинается с $w''_k t_{k+1}$ и $w_k = w'_k w''_k$ (слова w_i, w'_k и w''_k принадлежат B_φ).

Также определим $f : C^* \to A^*$ по правилу

$$f([twt']) = tw.$$

Предложение 5.1.4. *Все буквы алфавита C являются φ'-растущими.*

Доказательство. Заметим, что в $\varphi'^n([twt'])$ столько же букв, сколько в слове $\varphi^n(t)$ φ-растущих букв. Очевидно, в образе $\varphi(t)$ от произвольной буквы $t \in I_\varphi$ содержится хотя бы одна буква из I_φ. Более того, в слове $\varphi^n(t)$ для некоторого n содержатся хотя бы две буквы из I_φ, иначе $\varphi^n(t) = w_n t_{i_n} v_n$ (где w_n и v_n принадлежат B_φ) и $|\varphi^n t|$ ограниченно. $\square$

Пусть $\varphi^\infty(a)$ имеет вид $a_1 w_1 a_2 \dots$, где $a_1, a_2 \in I_\varphi$, $w_1 \in B_\varphi$. Тогда, несложно убедиться, что для любого n слово $\varphi^n(a)$ является началом слова $f(\varphi'^n([a_1 w_1 a_2]))$, следовательно, слова W и $(\psi \circ f)(\varphi'^\infty([a_1 w_1 a_2]))$ совпадают. Заметим, что морфизм $h' = h \circ f$ является нестирающим.

Замечание 5.1.1. *Конструкция морфизма ψ встречалась в работах [107, 108].*

Таким образом, можно считать, что все буквы алфавита A являются $\varphi-$ растущими, а h - произвольный нестирающий морфизм.

Букву $a_i \in A$ назовем *рекуррентной*, если для некоторого $k \in \mathbb{N}$ выполнено $a_i \sqsubseteq \varphi^k(a_i)$. Для каждой рекуррентной буквы алфавита существует такое число $k(a_i)$, что если $k|n$, то $a_i \sqsubseteq \varphi^n(a_i)$. Следовательно, существует такое $n \in \mathbb{N}$, что если a_i — произвольная рекуррентная буква алфавита, то $a_i \sqsubseteq \varphi^n(a_i)$. Положим $\rho = \varphi^n$.

Рассмотрим ориентированный граф G_ρ, вершинами которого являются буквы алфавита A, и из a_i ведет стрелка в a_j тогда и только тогда, когда a_j содержится в $\rho(a_i)$.

Пусть D — сильносвязная компонента этого графа, до которой можно дойти по стрелочкам из a_1. Рассмотрим ограничение ρ на D^*. Все буквы из D являются рекуррентными. Следовательно, если $d \in D$, то $d \sqsubseteq \rho(d)$. Также для любого k выполнено $\rho^k(d) \sqsubseteq \rho^{k+1}(d)$. Следовательно, существует такое m, что для любых букв d_1 и $d_2 \in D$ $d_2 \sqsubseteq \rho^m(d_1)$. Поэтому морфизм ρ в ограничении на D^* является примитивным.

Найдется такая буква $d \in D$, что $\rho^l(d)$ начинается на d для некоторого l. Обозначим $\rho_2 = \rho^l$. Так как все буквы из D являются ρ_2-растущими, то $\rho_2^\infty(d)$ является бесконечным сверхсловом, все конечные подслова которого являются подсловами $\varphi^\infty(a_1)$.

Слово $H = \psi(\rho_2^\infty(d))$ является почти периодичным как примитивное (см., например, в [60].)

Предложение 5.1.5. *Пусть W и H — сверхслова, H является почти периодичным, и все конечные подслова H являются подсловами сверхслова W. Тогда W является почти периодичным тогда и только тогда, когда любое его конечное подслово является подсловом сверхслова H.*

Доказательство можно найти в [107].

Проверим, является ли H периодичным. Если является, то достаточно проверить периодичность слова W, что мы делать умеем. Поэтому далее считаем, что H — непериодичное слово.

5.2 Порядок роста букв.

Напомним: φ — продолжающаяся над a_1 подстановка на алфавите A, для которой все буквы являются φ—возрастающими, $\psi : A \to B^+$ — нестирающий морфизм. Можно считать, что все буквы из A принадлежат $\varphi^\infty(a_1)$.

Известно следующее утверждение (см. [118]):

Предложение 5.2.1. *Пусть $\sigma : A^* \to A^*$ — подстановка. Для любой буквы $a \in A$ выполнено одно из двух условий:*

1. $\exists k \in \mathbb{N} : \sigma^k(a) = \varepsilon$.

2. $\exists d(a) \in \mathbb{N}_0, c(a) \in \mathbb{R}_+, C(a) \in \mathbb{R}, \theta(a) \in \mathbb{R}$ *такие, что для всех k выполнено*
$$c(a) < \frac{\sigma^n(a)}{c(a)n^{d(a)}\theta(a)^n} < C(a)$$

Очевидно, для φ выполнено второе условие. Для каждой буквы a_i пара $(d(a_i), \theta(a_i))$ называется *порядком роста буквы.* Также (см. [79,118]) известно, что если в A есть буква порядка роста (n, θ), где $n > 0$, то в A есть буква с порядком роста $(n-1,\theta)$. Таким образом, если для некоторой буквы $a_i \in A$ $\theta(a_i)=1$, то в A есть буква роста $(0,1)$, то есть $\varphi-$ограниченная. Так как у нас все буквы φ-растущие, то для любой буквы a_i $\theta(a_i)>1$.

На порядках роста можно ввести операцию сравнения: $(d_1,\theta_1)<(d_2,\theta_2)$, если $\theta_1<\theta_2$ или $\theta_1<\theta_2$ и $d_1<d_2$. Если у буквы a_1 порядок роста меньше, чем у a_2, то $\lim_{k\to\infty}\dfrac{|\varphi^k(a_1)|}{|\varphi^k(a_2)|}=0$. Аналогично порядок роста буквы определяется *порядок роста конечного слова*, который для слова u будем обозначать $r(u)$.

Предложение 5.2.2. *Очевидно,* $r(a_1a_2\dots a_n)=\max_{i\in[1;n]} r(a_i)$. *Кроме того,* $r(u)=r(\varphi(u))$.

Лемма 5.2.1. *Пусть* $r(u)=(d,\theta)$, *при этом* $\theta>1$. *Тогда для некоторого* $C(u)$ *для всех* k *выполнено*

$$|u|+|\varphi(u)|+|\varphi^2(u)|+\dots+|\varphi^k(u)|<C(u)k^d\theta^k.$$

Доказательство. В самом деле, для некоторого C для любого k выполнено $|\varphi^k(u)|<Ck^d\theta^k$. Тогда

$$\sum_{i=0}^{k}|\varphi^i(u)|<C\sum_{i=0}^{k}i^d\theta^i<C\sum_{i=0}^{k}k^d\theta^i<\frac{\theta C}{\theta-1}k^d\theta^k.$$

□

Положим D,Θ — самый большой порядок роста среди порядков роста для всех букв из алфавита. Буквы с таким порядком роста будем называть *быстрорастущими.* Очевидно, a_1 является быстрорастущей буквой.

Лемма 5.2.2. *В слове* $\varphi^\infty(a_1)$ *бесконечно много быстрорастущих букв.*

Доказательство. Пусть $\varphi(a_1)=a_1u$. Тогда

$$\varphi^\infty(a_1)=a_1u\varphi(u)\varphi^2(u)\varphi^3(u)\dots$$

Если в слове u есть хотя бы одна быстрорастущая буква, то хотя бы одна быстрорастущая буква есть в $\varphi^k(u)$ для любого k.

Предположим, что в u быстрорастущих букв нет.

Пусть $r(u)=(d,\theta)<(D,\Theta)$. Как мы знаем, $\theta>1$. Тогда $|\varphi^k(a)|=1+\sum_{i=1}^{k}|\varphi^{i-1}(u)|<C(u)k^d\theta^k$ согласно лемме 5.2.1. То есть $(D,\Theta)=(d,\theta)$. Противоречие. □

Лемма 5.2.3. *Если W почти периодично, то все буквы из A являются быстрорастущими.*

Доказательство. Если в A не все буквы являются быстрорастущими, то в $\varphi^\infty(a_1)$ найдется подслово вида $e_0C_0f_0$, где e_0 и f_0 — быстрорастущие буквы, а $r(C_0) < (D,\Theta)$. Для каждого $k \in \mathbb{N}$ представим $\varphi^k(e_0) = V'_k e_k V_k$, где $r(e_k) = (D,\Theta)$ и $r(V_k) < (D,\Theta)$. Последовательность $\{e_k\}$ заключительно периодична с неким периодом T_1.

Аналогично $\varphi^k(f_0) = U_k f_k U'_k$, $r(U_k) < (D,\Theta)$, $r(f_k) = (D,\Theta)$, посдедовательность $\{f_k\}$ заключительно периодична с периодом T_2.

Пусть $\sigma = \varphi^{T_1T_2}$. Тогда существуют такие буквы e, f и слово C, что $\sigma(e) = V'eV$, $\sigma(f) = UfU'$, $r(e) = r(f) = (D,\Theta)$, $r(U) < (D,\Theta)$, $r(V) < (D,\Theta)$, $r(C) < (D,\Theta)$, eCf — подслово сверхслова $\varphi^\infty(a_1)$.

Введем обозначения:

$$E_k = \varphi^{k-1}(V')\varphi^{k-2}(V')\dots V'e;$$

$$F_k = fU'\varphi(U)\varphi^2(U)\dots\varphi^{k-1}(U);$$

$C_k = V\varphi(V)\dots\varphi^{k-1}(V)\varphi^k(C)\varphi^{k-1}(U)\varphi^{k-2}(U)\dots U.$

Для любого k, $E_kC_kF_k$ является подсловом $\varphi^\infty(a_1)$. Каждое следующее E_k оканчивается на предыдущее, каждое следующее F_k начинается на предыдущее. Из леммы 5.2.1 следует, что

$$\lim_{k\to\infty}\frac{|E_k|}{|C_k|} = \lim_{k\to\infty}\frac{|F_k|}{|C_k|} = \infty$$

. Так как морфизм ψ является нестирающим, то

$$\lim_{k\to\infty}\frac{|\psi(E_k)|}{|\psi(C_k)|} = \lim_{k\to\infty}\frac{|\psi(F_k)|}{|\psi(C_k)|} = \infty$$

.

Предположим, что $W = \psi(\varphi^\infty(a_1))$ является почти периодичным. Тогда, как было показано в конце предыдущего раздела, множество его конечных подслов совпадает с множеством подслов — непериодичного сверхслова H, порожденного примитивной подстановкой. Следовательно, для W выполнены следующие свойства:

Предложение 5.2.3. *Существуют такие положительные действительные P и K, что*

1. *Если u_1 и u_2 — подслова W и $|u_2| \geq P|u_1|$, то $u_1 \sqsubseteq u_2$.*

2. *Если $u \sqsubseteq W$, то то для любых двух различных вхождений u в W их левые концы находятся на расстоянии не меньшем, чем $K|u|$ (подразумевается, что $K < 1$.)*

Возьмем такое натуральное m, что $\dfrac{3P}{m} < \dfrac{K}{2}$. Найдется такое k, что $|\psi(E_k)| > m|\psi(C_{k+m})|$ и $|\psi(F_k)| > m|\psi(C_{k+m})|$.

В W есть подслова $\psi(E_k)\psi(C_{k+i})\psi(F_k)$ для всех $i = 0, 1, \ldots, m$.

Пусть $N = \min\{|\psi(E_k)|, |\psi(F_k)|\}$. Таким образом, существует такой набор слов A, B, C_i для $i = 0, \ldots, m$, что

1. $|A| = |B| = N$;

2. $|C_0| < |C_1| < \cdots < |C_m| < \dfrac{N}{m}$;

3. $AC_iB \sqsubseteq W$ для любого i.

Все слова AC_iB, согласно 5.2.3, можно поместить в U — подслово сверхслова W длины $3PN$. По принципу Дирихле, для каких-то i и j левые концы слов AC_iB и AC_jB будут находиться в U на расстоянии, не большем $\dfrac{3PN}{m}$. Тогда правые концы этих слов будут находиться на расстоянии, не большем, чем $\dfrac{N}{m} + \dfrac{3PN}{m} < KN$. Слово B встретилось со сдвигом, меньшим, чем KN. Противоречие. □

Замечание 5.2.1. *В работе [79] была сформулирована и доказана аналогичная лемма.*

Если у всех букв алфавита одинаковый порядок роста, то $D = 0$. Существуют константы C_1 и C_2 такие, что для любого $k \in \mathbb{N}$ и любой буквы a_i выполнено $C_1\Theta^k < |\psi(\varphi^k(a_i))| < C_2\Theta^k$.

Предложение 5.2.4. *Алгоритмически можно определить, одинаковые ли порядки роста у букв алфавита. Если одинаковые, то можно явно указать C_1 и C_2, а также оценить сверху и снизу число Θ.*

Доказательство. Сначала проверим, одинаковые ли порядки роста у букв. Напомним, что буква называется рекуррентной, если она принадлежит некоторой итерации φ от самой себя.

Лемма 5.2.4. *Порядки роста у всех букв одинаковые если и только если у всех рекуррентных букв порядки роста одинаковые.*

Доказательство. В одну сторону это очевидно, докажем в другую, показав, что для любой буквы a есть рекуррентная буква с таким же порядком роста.

Рассмотрим последовательность слов $\{u_i\}$, где $u_1 = a$, а слово u_{i+1} получается из $\varphi(u_i)$ вычеркиванием всех рекуррентных букв. Если какое-то слово $u_k = \varepsilon$, то, пользуясь 5.2.2, несложно показать, что порядок роста каждого из u_i является порядком роста некоторой рекуррентной буквы. С другой стороны, для каждой буквы в u_{i+1} в u_i есть та буква, из которой она непосредственно возникла, поэтому, если $|u_{|A|+1}| > 0$, то среди букв слов u_i по принципу Дирихле есть рекуррентная буква. □

Очевидно, что при замене φ на φ^k одинаковые порядки роста останутся одинаковыми, а различные — различными. Рассмотрим ρ — такую степень φ, что если a_i — рекуррентная буква, то $a_i \sqsubseteq \rho(a_i)$. Для каждой рекуррентной буквы a_i рассмотрим ограничение φ на то множество букв, которое можно получить из a_i итерациями подстановки ρ. Для a_i рассмотрим на этом множестве букв матрицу подстановки M_i, пусть букве a_i в матрице отвечают первая строка и столбец.

Порядок роста $r(a_i)$ — это порядок роста вектора $M_i^k(a_i)$ в l_1-норме. Рассмотрим жорданов базис для оператора M_i (возможно, придется комплексифицировать пространство). Рассмотрим те жордановы клетки, у которых модуль числа на диагонали наибольший. Пусть этот модуль — это λ, из природы оператора M_i следует, что $\lambda > 1$. Среди этих клеток возьмем клетку самого большого размера d. Тогда скорость роста любого вектора в эрмитовой норме не превосходит $Cn^{d-1}|\lambda|^n$, причем вектора, у которых порядок роста меньше, лежат в подпространстве коразмерности 1. Так как все нормы в конечномерном пространстве эквивалентны, то в исходном (некомпактифицированном) пространстве есть вектор с положительными координатами и таким же порядком роста нормы. Заметим, что у вектора, соответствующего букве a_i, порядок роста не меньше (ибо он не меньше, чем порядки роста всех остальных букв, соответствующих строкам матрицы). Итак, мы выяснили, что если $r(a_i) = (d, \theta)$, то θ — это максимальный из модулей собственных значений оператора M_i, а d — размер жордановой клетки.

Так как комплексные числа можно представить в виде матриц 2×2 из действительных чисел, то в сигнатуре $\{=, <, 0, 1, +, \times\}$ можно выразить свойство "жордановы клетки оператора M_i имеют такой-то вид; клетки оператора M_j имеют такой-то вид; наборы модулей собственных значений, соответствующих клеткам, упорядочены так-то".

Из теоремы Тарского-Зайденберга следует, что равенство $r(a_i) = r(a_j)$ алгоритмически проверяемо.

Пусть уже установлено, что все буквы имеют одинаковый порядок роста. В таком случае, этот порядок — $(0, \theta)$ для некоторого θ.

Из доказательства леммы 5.2.4 следует, что достаточно уметь указывать ограничивающие константы для каждой рекуррентной буквы. Кроме того, можно указать константы для подстановки ρ.

Итак, мы снова работаем с матрицей M_i. Ограничивающие константы C_1 и C_2 можно вывести, если известны:

1. Вид жордановой формы (то есть упорядочивание жордановых клеток по модулям коэффициентов и их размеры);

2. Информация про коэффициенты разложения вектора $(1, 0, \ldots, 0)$ по некоторому жорданову базису (коэффициент при одном из базисных векторов, отвечающий наибольшему собственному значению, ненулевой. Нам достаточно оценки снизу на его отношение к остальным коэффициентам).

3. Оценки сверху на число обусловленности матрицы перехода к данному жорданову базису.

Первый пункт делается применением теоремы Тарского-Зайденберга. Покажем, как можно получить оценки во втором и третьем пунктах. Из алгоритма поиска жорданова базиса следует, что существует матрица перехода к жорданову базису, коэффициенты которой являются рациональными функциями от корней характеристического многочлена матрицы M_i. Значит, для второго и третьего пунктов существуют оценки, являющиеся отношениями двух многочленов от корней характеристического многочлена, при этом эти оценки не равны нулю. В таком случае можно как угодно оценить сверху значение знаменателя, а значение числителя оценить снизу с помощью обобщенной теоремы Лиувилля. $\square$

Далее можно считать, что порядки роста всех букв равны. Тогда, согласно 5.1.5, задачу можно свести к следующей:

Дано: алфавиты A, B, C; морфизмы $\varphi : C^* \to C^*$, $\psi : C^* \to B^*$, $g : A^* \to A^*$, $h : A^* \to B^*$, а также числа C_1', C_2', $\Theta_1 > 1$, Θ_2. При этом все морфизмы нестирающие, морфизм φ примитивен и продолжается над c_1 и известно, что $C_1'\lambda^k < |h(g^k(c_i))| < C_2'\lambda^k$ для всех k и некоторого $\lambda \in [\Theta_1; \Theta_2]$.

Определить: верно ли что для всех натуральных k и букв $a_i \in A$ каждое из слов $h(g^k(a_i))$ является подсловом сверхслова $W = \psi(\varphi^\infty(c_1))$?

Теорема 5.2.1. *Эта задача алгоритмически разрешима.*

Для доказательства потребуется язык схем Рози.

5.3 Схемы Рози.

Напомним основные определения и факты, связанные со схемами Рози и описанные в разделах 3 и 4.

Графом со словами будем называть сильносвязный ориентированный граф, у которого на каждом ребре написано по два слова — *переднее* и *заднее*. Также потребуем, чтобы каждая вершина либо имеет входящую степень 1, а исходящую больше 1, либо входящую степень больше 1 и исходящую степень 1. Вершины первого типа назовем *раздающими*, а второго — *собирающими*.

Путь в графе со словами — это последовательность ребер, каждое следующее из которых выходит из той вершины, в которую входит предыдущая. *Симметричный путь* — это путь, первое ребро которого начинается в собирающей вершине, а последнее ребро кончается в раздающей.

Каждый путь можно записать словом над алфавитом — множеством ребер графа, и это слово называется *реберной записью пути.* Для двух путей выполняются отношения *подпути* (пишем $s_1 \sqsubseteq s_2$), *начала* или *конца*, если для их реберных записей выполняются соответственно отношения подслова, начала или конца. Кроме того, пишем $s_1 \sqsubseteq_k s_2$, если для соответствующих слов u_1 и u_2 — реберных записей путей s_1 и s_2 — выполнено $u_1 \sqsubseteq_k u_2$. Если последнее ребро пути s_1 идет в ту же вершину, из которой выходит первое ребро пути s_2, путь, реберная запись которого является конкатенацией реберных записей путей s_1 и s_2, будем обозначать s_1s_2.

Для каждого пути s в графе со словами определим *переднее слово* $F(s)$. Пусть $v_1v_2\dots v_n$ — реберная запись пути s. В $v_1v_2\dots v_n$ возьмем подпоследовательность: включим в нее v_1, а также те и только те ребра, которые выходят из раздающих вершин графа. Эти ребра назовем *передними образующими для пути* s. Возьмем передние слова этих ребер и запишем их последовательную конкатенацию, это и будет $F(s)$.

Замечание 5.3.1. *Запись* $|s|$ *обозначает длину пути, которая измеряется в ребрах.* $|F(s)|$ *— это длина слова.*

Аналогично определяется $B(s)$. В $v_1v_2\dots v_n$ возьмем ребра, входящие в собирающие вершины и ребро v_n в порядке следования — это *задние образующие для пути* s. Тогда последовательной конкатенацией задних слов этих ребер получается *заднее слово* $B(s)$ *пути* s.

Определение 5.3.1. Если S Ц сильносвязный граф, не являющийся циклом, и s Ц путь в графе, то естественное продолжение пути s вправо Ц это минимальный путь, началом которого является s и который оканчивается в раздающей вершине. Естественное продолжение пути s влево Ц это минимальный путь, концом которого является s и который начинается в собирающей вершине.

Определение 5.3.2. Граф со словами будет являться *схемой Рози* для рекуррентного сверхслова W, если он удовлетворяет следующим свойствам, которые в дальнейшем будут называться *свойствами схем Рози*:

1. Граф сильносвязен и состоит более чем из одного ребра.

2. Все ребра, исходящие из одной раздающей вершины графа, имеют передние слова с попарно разными первыми буквами. Все ребра, входящие в одну собирающую вершину графа, имеют задние слова с попарно разными последними буквами.

3. Для любого симметричного пути, его переднее и заднее слова совпадают. То есть можно говорить просто о слове симметричного пути.

4. Если есть два симметричных пути s_1 и s_2 и выполнено $F(s_1) \sqsubseteq_k F(s_2)$, то $s_1 \sqsubseteq_k s_2$.

5. Все слова, написанные на ребрах графа, являются подсловами W.

6. Для любого u – подслова W существует симметричный путь, слово которого содержит u.

7. Для любого ребра s существует такое слово u_s, принадлежащее W, что любой симметричный путь, слово которого содержит u_s, проходит по ребру s.

Опорным ребром в схеме называется любое ребро, входящее в раздающую вершину и выходящее из собирающей. В любой схеме Рози присутствует хотя бы одно опорное ребро, и у каждого опорного ребра переднее слово равно заднему. *Масштабом* схемы Рози называется наименьшая из длин слов на опорных ребрах. В разделе 3 показано, как для непериодичного рекуррентного сверхслова получать с помощью *графов Рози* схему Рози сколь угодно большого масштаба.

Определим *эволюцию* (W, S, v) схемы Рози S по опорному ребру v: пусть $\{x_i\}$ — множество ребер. входящих в начало v, а $\{y_i\}$ — множество ребер, идущих из конца v (эти два множества могут пересекаться). Обозначим $F(y_i) = Y_i$, $B(x_i) = X_i$, $F(v) = V$. Рассмотрим все слова вида X_iVY_j. Если такое слово не входит в W, то пару (x_i, y_j) назовем *плохой*, в противном случае — *хорошей*. Также *хорошей* или *плохой* будем называть соответствующую тройку ребер (x_i, v, y_j).

Построим граф S'. Он получается из S заменой ребра v на $K_{\#\{x_i\},\#\{y_j\}}$, где $K_{m,n}$ — полный двудольный граф. Более подробно: ребро v удаляется, его начало заменяется на множество $\{A_i\}$ из $\#\{x_i\}$ вершин так, что для любого i ребро x_i идет в A_i; конец ребра v заменяется на множество $\{B_j\}$ из $\#\{y_j\}$ вершин так, что для любого j ребро y_j выходит из B_j; вводятся ребра $\{v_{i,j}\}$, соединяющие вершины множества $\{A_i\}$ с вершинами множества $\{B_j\}$.

По сравнению с S, у графа S' нет ребра v, но есть новые ребра $\{v_{ij}\}$. Остальные ребра граф S взаимно однозначно соответствуют ребрам графа S'. Соответственные ребра в первом и втором графе зачастую будут обозначаться одними и теми же буквами.

На ребрах S' расставим слова следующим образом. На всех ребрах S', кроме ребер из $\{y_i\}$ и $\{v_{i,j}\}$, передние слова пишутся те же, что и передние слова соответственных ребер в S. Для каждого i и j, переднее слово ребра y_j в S' — это VY_j. Переднее слово ребра $v_{i,j}$ — это Y_j. Аналогично, на всех ребрах, кроме ребер из $\{x_i\}$ и $\{v_{i,j}\}$, задние слова переносятся с соответствующих ребер S; для всех i и j в качестве заднего слова ребра x_i возьмем X_iV, а в качестве заднего слова ребра $v_{i,j}$ — X_i.

Теперь построим граф S''. Ребра $v_{i,j}$, соответствующие плохим парам (x_i, y_j), назовем *плохими*, а все остальные ребра графа S' — *хорошими*. Граф S'' получается, грубо говоря, удалением плохих ребер из S'. Более точно, в графе S'' раздающие вершины — подмножество раздающих вершин S', а собирающие вершины — подмножество собирающих вершин S'. В графе S' эти подмножества — вершины, из которых выходит более одного хорошего ребра, и вершины, в которые входит более одного хорошего ребра соответственно. Назовем в S' вершины этих двух подмножеств S' неисчезающими. Ребра в графе S'' соответствуют таким путям в S', которые идут лишь по хорошим ребрам, начинаются в неисчезающих вершинах, заканчиваются в неисчезающих, а все промежуточные вершины которых не являются неисчезающими.

Несложно показать, что S'' — сильносвязный граф, не являющийся циклом. У каждого ребра в S'' есть естественное продолжение вперед и назад. Естественное продолжение вперед соответствует пути в S'; переднее слово этого пути в S' возьмем в качестве переднего слова для соответствующего ребра S''. Аналогично для задних слов: у ребра в S'' есть естественное продолжение влево, этому пути в S'' соответствует путь в S'. Заднее слово этого пути и будет задним словом ребра в S''.

Определение 5.3.3. Построенный таким образом граф со словами S'' назовем *элементарной эволюцией* (W, S, v).

Теорема 5.3.1. *Элементарная эволюция* (W, S, v) *является схемой Рози для сверхслова* W *(то есть удовлетворяет свойствам* 1—7 *определения 5.3.2).*

Определение 5.3.4. Пусть S — схема Рози для сверхслова W. На ее ребрах можно написать различные натуральные числа (или пары чисел). Такую схему мы назовем *нумерованной.* Если с ребер пронумерованной схемы стереть слова, получится *облегченная нумерованная схема.*

Определение 5.3.5. *Метод эволюции* — это функция, которая каждой облегченной пронумерованной схеме (с нумерацией, допускающей двойные индексы) дает этой же схеме новую нумерацию, такую, что в ней используются числа от 1 до n для некоторого n по одному разу каждое.

Зафиксируем какой-либо метод эволюции и далее не будем его менять.

Среди опорных ребер пронумерованной схемы S возьмем ребро v с наименьшим номером, и совершим элементарную эволюцию (W, S, v). Укажем естественную нумерацию новой схемы.

Напомним, что сначала строится схема S', а потом — S''. В схеме S' все ребра можно пронумеровать по следующему правилу: ребра кроме v сохраняют номера, а ребра вида v_{ij} нумеруют соответствующим двойным индексом.

Каждое ребро в схеме S'' — это некоторый путь по ребрам схемы S', различным ребрам из S'' соответствуют в S' пути с попарно различными первыми ребрами. Таким образом, ребра схемы S'' можно пронумеровать номерами первых ребер соответствующих путей S'. Теперь применим к облегченной нумерованной схеме S'' метод эволюции. Получится новая облегченная нумерованная схема, и в нумерованной (не облегченной) схеме S'' перенумеруем ребра соответственным образом.

Определение 5.3.6. Описанное выше соответствие, ставящее нумерованной схеме Рози другую нумерованную схему Рози, назовем *детерменированной эволюцией.* Будем обозначать это соответствие $S'' = \mathrm{Evol}(S)$.

Определение 5.3.7. Применяя детерменированную эволюцию к нумерованной схеме S много раз, получаем последовательность нумерованных схем Рози. Кроме того, на каждом шаге получаем множество пар чисел, задающие плохие пары ребер. *Протокол детермеминованной эволюции* — это последовательность таких множеств пар чисел и облегченных нумерованных схем.

Из определения элементарной эволюции следует

Предложение 5.3.1. *Облегченная нумерованная схема для* $\mathrm{Evol}(S)$ *однозначно определяется по облегченной нумерованной схеме* S *и множеству пар чисел, задающие плохие пары ребер.*

Предложение 5.3.2. *Каждому симметричному пути в* $\mathrm{Evol}(S)$ *соответствует симметричный путь в* S *с таким же словом; длина соответствующего пути в* S *не меньше, чем в* $\mathrm{Evol}(S)$ *(строится соответствие так:* $\mathrm{Evol}(S) = S''$ *естественным образом вкладывается в* S'*, а после применяется отображение путей* f^{-1}*, описанное в [125]). При этом соответствии сохраняется отношение* $\sqsubseteq_k$*, а также отношения "начала пути" и "конца пути". Опорное ребро, по которому делалась эволюция, не соответствует в* $\mathrm{Evol}(S)$ *никакому пути. Для любого симметричного пути в* S*, не являющегося этим опорным ребром, в* S *существует минимальный симметричный путь, соответствующий некоторому симметричному пути в* $\mathrm{Evol}(S)$*. Все соответствующие пути можно определить по облегченным нумерованным схемам.*

В работе [125] доказана следующая теорема:

Теорема 5.3.2. *Если* W *— примитивное подстановочное непериодичное сверхслово, то протокол его детерменированной эволюции периодичен с предпериодом.*

При этом предпериод и период алгоритмически находятся по морфизмам, порождающим слово W.

Определение 5.3.8. Если S — схема Рози для сверхслова W, а T — натуральное число, то в соответствующей облегченной нумерованной схеме можно указать, какие из симметричных путей длины не более T являются допустимыми. Этот набор (облегченная нумерованная схема + набор допустимых путей не длиннее T) будем называть *T−разруленной схемой.*

Из доказательства теоремы 5.3.2 (см. [125]) вытекает следующее предложение:

Предложение 5.3.3. *Если W — примитивное подстановочное сверхслово, а T — произвольное число, то для протокола детерменированной эволюции последовательность соответствующих T-разруленных схем периодична с предпериодом. Опять же, предпериод и период алгоритмически находятся.*

5.4 Построение алгоритма для теоремы 5.2.1.

Предложение 5.4.1. *Пусть $W = \psi(\varphi^\infty(c_1))$ — непериодичное примитивное сверхслово. Тогда:*

1. *Можно явно найти такое число P, что если u_1 и u_2 — подслова W и $|u_2| \geq P|u_1|$, то $u_1 \sqsubseteq u_2$.*
2. *Можно явно найти такое число C, что если $u \sqsubseteq W$, то то для любых двух различных вхождений u в W их левые концы находятся на расстоянии не меньшем, чем $C|u|$.*
3. *Можно явно найти такое число $C_{\max}$, что в любой схеме Рози S длина слова любого допустимого пути не менее $C_{\min}Mn$, где n — количество ребер в пути, а M — масштаб схемы S.*
4. *Можно явно найти такое число $C_{\min}$, что в любой схеме Рози S длина любого переднего и заднего слова на ребрах схемы не превосходит $C_{\max}M$, где M — масштаб схемы S. Следовательно, если путь содержит n ребер, то длина переднего и заднего слов этого пути не превосходит $C_{\max}Mn$.*
5. *Масштаб схемы* $\mathrm{Evol}(S)$ *не меньше масштаба схемы S. Можно явно найти такое C_m, что масштабы схем* $\mathrm{Evol}(S)$ *и S относятся не более, чем в C_m раз для любой схемы S.*

Доказательства всех утверждений предложения 5.4.1 можно найти в [125] или в разделе 3 диссертации.

Далее слово W, а также морфизмы φ и ψ считаем неизменными.

Пусть s — симметричный путь по ребрам схемы S. Пусть последовательность его ребер — это $v_1, v_2, \ldots, v_n$. Предположим, что $v_{i_1}, v_{i_2}, \ldots, v_{i_m}$ — его передние образующие. Если длина слова $F(s)$ превосходит $C_{\max}M$, то передних образующих ребер более одного. Путь $v_1 v_2 \ldots v_{i_m - 1}$ также является симметричным. Будем обозначать его s_{right} и по отношению к s называть *урезанным справа.* Так как $F(s) = F(s_{\text{right}})F(v_{i_m})$, то $F(s_{\text{right}}) \geq F(s) - C_{\max}M$.

Аналогично определяется путь s_{left}, который является по отношению к s *урезанным слева.*

Если $F(s) \geq 2C_{\max}M$, то существует путь $(s_{\text{left}})_{\text{right}} = (s_{\text{right}})_{\text{left}}$, который будем обозначать $\operatorname{cut}(s)$. Очевидно, $F(s) = B(v_1)\operatorname{cut}(s)F(v_2)$ для некоторых ребер v_1 и v_2.

Пусть S — схема Рози для сверхслова W, A — слово. Если $A \sqsubseteq W$, то $A \sqsubseteq F(s)$ для некоторого допустимого пути s.

Определение 5.4.1. Любой минимальный по включению путь s будем обозначать $l(S, A)$. Здесь S — схема, A — слово. Далее будет показано, что во многих случаях такой путь единственный.

Лемма 5.4.1. *Длина слова пути $l(S, A)$ не превышает $|A| + 2C_{\max}M$.*

Доказательство. Пусть $s = l(S, A)$. Слово $F(s)$ можно представить в виде конкатенации $B(v_1)F(\operatorname{cut}(s))F(v_2)$. В этом слове содержится A, но при этом A из-за минимальности s не содержится в $B(v_1)F(\operatorname{cut}(s)) = F(s_{\text{right}})$ и в $F(\operatorname{cut}(s))F(v_2) = F(s_{\text{left}})$. Следовательно, $A = u_1 F(\operatorname{cut}(s)) u_2$ для некоторых непустых u_1 и u_2. А как мы знаем, $F(\operatorname{cut}(s)) \geq F(s) - 2C_{\max}M$. $\square$

Лемма 5.4.2. *Существует такое число K, что если M — масштаб схемы S, а длина $|A|$ не меньше KM, то $l(S, A)$ единственен.*

Доказательство. Зафиксируем один минимальный путь s. Также рассмотрим произвольный минимальный путь s'. Из доказательства леммы 5.4.1 следует, что $A = u_1 F(\operatorname{cut}(s)) u_2$ для некоторых непустых u_1 и u_2. Так как $A \sqsubseteq F(s')$, то $F(\operatorname{cut}(s)) \sqsubseteq F(s')$. По свойству 4 схем Рози, $\operatorname{cut}(s) \sqsubseteq s'$. Длины слов $F(\mathrm{cut}(s))$ и $F(s')$ отличаются не более, чем на $4MC_{\max}$, следовательно, если

$$\frac{4MC_{\max}}{|A| - 2MC\max} < C,$$

то у слова $F(\mathrm{cut}(s))$ ровно одно вхождение в $F(s')$. Выполнения этого неравенства легко достичь, выбирая достаточно большое K.

Пусть $v_1v_2\dots v_n$ — ребра пути s', при этом ребра $v_k, v_{k+1}, \dots v_l$ образуют путь $\mathrm{cut}(s)$. Тогда $B(v_{k-1})F(\mathrm{cut}(s))F(v_{l+1}) \sqsubseteq F(s')$. Так как у $F(\mathrm{cut}(s))$ ровно одно вхождение в $F(s')$, то $F(v_{l+1})$ и u_2 начинаются на одну и ту же букву, а $B(v_{k-1})$ и u_1 кончаются на одну и ту же букву. Следовательно, у путей s и s' первое ребро после вхождения $\mathrm{cut}(s)$ одно и то же (а именно то, переднее слово которого начинается с первой буквы слова u_1.) Аналогично, ребро непосредственно перед вхождением $\mathrm{cut}(s)$ одно и то же. Следовательно, s' содержит путь s. Из минимальности следует, что $s = s'$. $\square$

Предложение 5.4.2. *Если S — схема с масштабом M, s — допустимый путь, $|A| \geq KM$, $A \sqsubseteq s$, $A = u_1F(\mathrm{cut}(s))u_2$ для непустых u_1, u_2. Тогда $l(S, A) = s$.*

Доказательство. Пусть $A \sqsubseteq F(s_{\mathrm{right}})$. Тогда у $F(\mathrm{cut}(s))$ есть вхождение в $A \sqsubseteq F(\mathrm{right})$, не являющееся его концом. Но у слова $F(\mathrm{cut}(s))$ есть ровно одно вхождение в $F(s)$ (см. доказательство леммы 5.4.2). Противоречие.

Аналогично доказывается, что $F(s_{\mathrm{left}})$ не содержит A. Следовательно, $s = l(S, A)$. $\square$

Лемма 5.4.3. *Пусть S — схема масштаба M, а длины слов A и B не менее KM. Тогда путь $l(S, A)$ — начало пути $l(S, AB)$.*

Доказательство. Пусть $v_1v_2\dots v_n$ — последовательность ребер $l(S, AB)$. Для некоторого k, $v_{k+1}v_{k+2}\dots v_n$ — последовательность ребер пути $l(S, AB)_{\mathrm{left}}$. Тогда $AB = u_1u_2$, $B(v_k)$ кончается на u_1 и $F(v_{k+1}v_{k+2}\dots v_n)$ начинается на u_2. Так как $|u_1| \leq C_{\max}M < |A|$, то $A = u_1u_3$. Слово $F(l(S, AB)_{\mathrm{left}})$ начинается на u_3, следовательно, существует такое минимальное k', что $v_{k+1}v_{k+2}\dots v_{k'}$ — симметричный путь, слово которого начинается с u_3. Докажем, что $l(S, A) = v_1v_2\dots v_{k'}$. В самом деле,

$$A = u_1u_3 \sqsubseteq B(v_k)F(v_{k+1}v_{k+2}\dots v_{k'}) = F(v_1v_2\dots v_{k'}).$$

Из минимальности k' следует, что $u_3 = F((v_{k+1}v_{k+2}\dots v_{k'})_{\text{right}})u_4$ для некоторого непустого u_4. Следовательно, $A = u_1F(\text{cut}(v_1v_2\dots v_{k'}))u_4$. Утверждение леммы следует из предложения 5.4.2. □

Из доказательства этой леммы следует следующее предложение:

Предложение 5.4.3. *Пусть S — схема масштаба M, а длины слов A и B не менее KM. Если $F(l(S,A)) = u_1Au_2$, а $F(l(S,AB)) = u_3ABu_4$. Тогда $u_1 = u_3$.*

Лемма 5.4.4. *Пусть S — схема Рози сасштаба M, слова A,B,C имеют длину более KM, слова AB и BC являются подсловами W. Тогда следующие условия эквивалентны:*

1. *Слово ABC является подсловом W.*
2. *Склейка путей $l(S,AB)$ и $l(S,BC)$ по пути $l(S,B)$ является допустимым путем (согласно лемме 5.4.3, путь $l(S,B)$ является концом $l(S,AB)$ и началом $l(S,BC)$). Иначе говоря, путь l, имеющий длину $|l(S,AB)| + |l(S,BC)| - |l(S,B)|$ и такой, что $l(S,AB)$ является началом пути l, а $l(S,BC)$ — концом l, является допустимым.*

При этом путем $l(S,ABC)$ будет являться построенный путь l.

Доказательство. Предположим, что $ABC \sqsubseteq W$. Рассмотрим путь $s = l(S,ABC)$. Согласно лемме 5.4.3, $l(S,AB)$ является началом s, а $l(S,BC)$ — концом s. Пусть $F(ABC) = u_1ABCu_2$. Тогда, согласно 5.4.3, $F(l(S,AB)) = u_1ABu_3$, а $F(l(S,BC)) = u_4BCu_2$. Также путь $l(S,B)$ является концом пути $l(S,AB)$ и началом пути $l(S,BC)$, а слово $F(l(S,B))$ представляется в виде u_4Bu_3.

Пусть $v_1v_2\dots v_n$ — последовательность ребер s. Тогда найдутся такие k_1 и k_2, что $l(S,AB) = v_1v_2\dots v_{k_1}$ и $l(S,BC) = v_{k_2}v_{k_2+1}\dots v_n$. Предположим, что $k_1 < k_2$. Тогда рассмотрим минимальный симметричный путь s', последнее ребро которого — v_{k_1}. Так как $B(s') = B(v_{k_1})$, то $F(s') \le C_{\max}M$. С другой стороны,

$$F(s') + F(s) = F(v_1v_2\dots v_{k_1}) + (F(s') + F(v_{k_1}v_{k_1+1}\dots v_n)) >$$

$$> F(l(S,AB)) + F(l(S,BC)).$$

Но правая часть превосходит левую хотя бы на $|B| - C_{\max}M$. Противоречие. Таким образом, $k_1 \le k_2$ и мы можем рассмотреть симметричный путь $v_{k_2}v_{k_2+1}\dots v_{k_1}$. Длина слова этого пути

$$|F(v_{k_2}v_{k_2+1}\dots v_{k_1})| = |F(l(S,AB))| + |F(l(S,BC))| - |F(s)| = |u_3| + |B| + |u_4|.$$

У слова $F(l(S,B))$ такая же длина и оно также является окончанием слова $F(l(S,AB))$. Следовательно, $F(v_{k_2}v_{k_2+1}\dots v_{k_1}) = F(l(S,B))$ и $v_{k_2}v_{k_2+1}\dots v_{k_1} = l(S,B)$ по 4-му свойству схем Рози. Значит, допустимый путь s — это путь l, фигурирующий во втором условии. Доказана импликация $1 \to 2$.

Предположим теперь, что выполнено второе условие. Докажем, что $F(l)$ содержит ABC. Воспользуемся предложением 5.4.3 и представим слова в следующем виде: $F(l(S,AB)) = u_1ABu_2$, $F(l(S,BC)) = u_3BCu_4$, $F(l(S,B)) = u_3Bu_2$. Пусть $v_1v_2\dots v_n$ — последовательность ребер пути l, при этом рассматриваемый подпуть $l(S,B)$ — это $v_{k_1}v_{k_1+1}\dots v_{k_2}$. Обозначим $D = F(v_{k_2+1}v_{k_2+2}\dots v_n)$. Тогда $u_3BCu_4 = u_3Bu_2D$, следовательно, $Cu_4 = u_2D$. Но

$$F(l) = u_1ABu_2D = u_1ABCu_4.$$

Доказана импликация $2 \to 1$. □

Определение 5.4.2. Все подслова сверхслова $g^\infty(a)$, имеющие длину 1 или 2, назовем $g-$*источниками.*

Все $g-$источники, согласно 4.1.2, находятся алгоритмически. Пусть их количество равно m.

Определение 5.4.3. Слова $h(g^k(q_1)), h(g^k(q_2)), \dots, h(g^k(q_m))$, где q_i- это $g-$источники, назовем *рабочими словами порядка k.*

Предложение 5.4.4. *Можно найти такие положительные числа C_1 и C_2, что для любых $k \in \mathbb{N}$ и g-источника q_i выполнено $C_1\lambda^k < |h(g^k(q_i))| < C_2\lambda^k$.*

Далее фиксируем число T. Оно предполагается достаточно большим; явно укажем его позднее.

Определение 5.4.4. Пусть S — нумерованная схема Рози, k — натуральное число, все рабочие слова порядка k являются подсловами W. Тогда множество, состоящее из

1. Соответствующей облегченной нумерованной схемы для S;

2. Множества всех допустимых путей не длиннее T, рассматриваемых как пути в облегченной схеме (эти пути назовем *проверочными*);

3. Номера $T-$разруленной схемы S в периоде или предпериоде (согласно 5.3.3, эти схемы периодичны с предпериодом);

4. Множества путей $l(S, h(g^k(q_i)))$ для всех источников q_i (пути рассматриваются в облегченной схеме, назовем их *основными путями*).

назовем *антиоснасткой* (S, k).

Определение 5.4.5. *Размер антиоснастки* (S, k) — это максимальная длина (в ребрах) по путям из $l(S, h(g^k(q_i)))$ для всех $g-$источников q_i.

Лемма 5.4.5. *Можно найти такие положительные C_3, C_4 и C_5, что для любой схемы S и любого k размер антиоснастки (S, k) заключен между $C_3\frac{\lambda^k}{M}$ и $C_4\frac{\lambda^k}{M} + C_5$, где M — масштаб схемы.*

Доказательство. Пусть p_i — рабочее слово порядка k. Согласно 5.4.4, $|p_i| > C_1\lambda^k$. Если $s = l(S, p_i)$, то $|F(s)| \geq |p_i|$. С другой стороны, $F(s) \leq C_{\max}M|s|$. Стало быть, $|s| \geq \frac{C_1\lambda^k}{C_{\max}M}$.

С другой стороны,

$$C_{\min}M|s| \leq F(s) \leq |p_i| + 2MC_{\max} < C_2\lambda^k + 2MC_{\max}$$

.

Следовательно, $|s| < \frac{C_2\lambda^k}{C_{\min}M} + 2\frac{C_{\max}}{C_{\min}}$. □

Следствие 5.4.1. *Можно указать явно такие C_6, C_7, C_8 и C_9, что если размер антиоснастки (S, k) больше C_6, а M — масштаб схемы S, то*

1. *Длины всех рабочих слов составляют не менее KM.*

2. *Размеры антиоснасток (S, k) и $(\mathrm{Evol}(S), k)$ относятся не более, чем в C_7 раз.*

3. *Если схема S_1 получена из S применением некоторого количества операций* Evol, *то размер антиоснастки (S_1, k) не превосходит размера антиоснастки (S, k), умноженного на C_7.*

4. *Размеры антиоснасток (S, k) и $(S, k+1)$ отличаются не более, чем в C_8 раз.*

5. *Размер оснастки $(S, k+C_9)$ больше размера оснастки (S, k) хотя бы в два раза.*

Доказательство. Пусть x — размер антиоснастки (S, k). В различных пунктах будут различные требования вида ограничения на x снизу. Выберем C_6 исходя из самого сильного требования.

1. Из 5.4.5 и 5.4.4 следует, что если x — размер антиоснастки, а q — рабочее слово порядка k, то $\lambda^k > \frac{(x-C_5)M}{C_4}$ и $|q| > C_1\lambda^k$. Стало быть, $|q| > \frac{C_1(x-C_5)M}{C_4}$. Достаточно выполнения неравенства $KC_4 < C_1(x - C_5)$.

2. Пусть M_1 — масштаб схемы $\mathrm{Evol}(S)$, x и x_1 — размеры антиоснасток. Тогда выполнены двойные неравенства $C_4\frac{\lambda^k}{M_1} + C_5 > x_1 > C_3\frac{\lambda^k}{M_1}$ и $C_4\frac{\lambda^k}{M} + C_5 > x > C_3\frac{\lambda^k}{M}$.

   Мы знаем, что $M \le M_1 \le C_m M$ (см. 5.4.1). Получаем $x_1 > \frac{C_3 M(x-C_5)}{M_1 C_4}$. Если $x > 2C_5$, то $x_1 > \frac{C_3}{2C_4C_m}x$.

   С другой стороны, $x_1 < \frac{C_4 x M}{M_1 C_3} + C_5 \le 2\frac{C_4}{C_3}x$ при $x > \frac{C_5C_3}{C_4}$.

3. Согласно 5.4.1, масштаб схемы S_1 не меньше, чем M. Далее см. доказательство предыдущего пункта.

4. Пусть x и x_1 — размеры антиоснасток (S, k) и $(S, k+1)$.

   Тогда $C_3\frac{\lambda^{k+1}}{M} < x_1 < C_4\frac{\lambda^{k+1}}{M} + C_5$. С другой стороны, $\frac{x-C_5}{C_4} < \frac{\lambda^k}{M} < \frac{x}{C_3}$.

   Таким образом, $\frac{C_3\lambda(x-C_5)}{C_4} < x_1 < \frac{C_4\lambda x}{C_3} + C_5$. При $x > 2C_5$ и $x > \frac{C_5C_3}{C_4}$ получаем, что
   $$\frac{C_3\lambda}{2C_4} < \frac{x_1}{x} < \frac{2C_4\lambda}{C_3}.$$

   Заметим, что нам не обязательно знать λ точно. Достаточно знать какую-нибудь оценку сверху.

5. Пусть x и x_1 — размеры антиоснасток (S, k) и $(S, k + C_9)$.

   Тогда $C_3 \frac{\lambda^{k+C_9}}{M} < x_1$. С другой стороны, $\frac{\lambda^k}{M} > \frac{x - C_5}{C_4}$.

   Таким образом, при $x > 2C_5$ выполняется $x_1 > \frac{C_3 \lambda^{C_9}}{2C_4} x$, что не меньше $2x$ при достаточно большом C_9.

   Заметим, что нам не обязательно знать λ точно. Достаточно знать какую-нибудь оценку снизу, отделяющую λ от 1.

$\square$

Лемма 5.4.6. *Если размер антиоснастки* (S, k) *не меньше* C_6*, то по антиоснастке* (S, k) *можно алгоритмически определить, верно ли, что все рабочие слова порядка* $k + 1$ *являются подсловами* W*, и, если верно, то алгоритмически найти антиоснастку* $(S, k + 1)$*.*

Доказательство. Пусть x — размер антиоснастки (S, k). Необходимо для каждого источника p_i определить, является ли $h(g^{k+1}(p_i))$ подсловом W, и если является, то найти путь $l(S, h(g^{k+1}(p_i)))$.

Пусть $g(p_i) = a_{i_1} a_{i_2} \dots a_{i_m}$.

$$h(g^{k+1}(p_i)) = h(g^k(a_{i_1}))h(g^k(a_{i_2})) \dots h(g^k(a_{i_m})).$$

Каждая из букв a_{i_1}, a_{i_2}, $\dots$, a_{i_m} является источником. Также источниками являются $a_{i_1}a_{i_2}$, $a_{i_2}a_{i_3}$,$\dots$, $a_{i_{m-1}}a_{i_m}$.

Согласно лемме 5.4.1, длины слов $h(g^k(a_{i_1}))$, $h(g^k(a_{i_2}))$, $h(g^k(a_{i_3}))$ не менее KM, поэтому для схемы S, слов $A = h(g^k(a_{i_1}))$, $B = h(g^k(a_{i_2}))$ и $C = h(g^k(a_{i_3}))$ выполняются условия леммы 5.4.4.

Из антиоснастки (S, k) нам известны пути $l(S, AB)$, $l(S, BC)$ и $l(B)$. Чтобы определить, является ли ABC подсловом сверхслова W, нужно узнать, является ли некоторый путь допустимым. Длина этого пути равна

$$|l(S, AB)| + |l(S, BC)| - |l(S, B)|$$

и не превосходит $2x$. Если $T > 2x$, то определить допустимость пути мы можем, проверив, лежит ли он среди проверочных путей. Если он допустимый, то он и является путем $l(ABC)$.

Далее воспользуемся леммой 5.4.4, примененной к словам $h(g^k(a_{i_1}a_{i_2}))$, $h(g^k(a_{i_3}))$, $h(g^k(a_{i_4}))$ и аналогичным способом определим путь

$$l(S, h(g^k(a_{i_1}a_{i_2}a_{i_3}a_{i_4}))),$$

если он существует. На этом шаге потребуется неравенство $T > 3x$.

Потом определим, существует ли путь $l(S, h(g^k(a_{i_1}a_{i_2}a_{i_3}a_{i_4}a_{i_5})))$ и, действуя подобным образом, доберемся до $l(S, h(g^k(a_{i_1}a_{i_2}\dots a_{i_k})))$. На последнем шаге нам хватит оценки $T > x(m-1)$. □

Лемма 5.4.7. *Если размер антиоснастки* S, k *хотя бы* C_8C_6*, то по* (S, k) *можно найти антиоснастку* $(\mathrm{Evol}(S), k)$.

Доказательство. По номеру $T-$разруленной схемы S в периоде или предпериоде можно узнать следующую разруленную схему. Для определения антиоснастки $(\mathrm{Evol}(S), k)$ осталось найти все основные пути.

Согласно 5.4.1, размер $(S, k+1)$ хотя бы C_6. Следовательно, в антиоснастке $(S, k+1)$ для каждого источника ровно один основной путь. Пусть s — основной путь антиоснастки $(\mathrm{Evol}(S), k)$, соответствующий источнику p. Согласно 5.3.2, в облегченной схеме S пути s соответствует путь s', у которого (в необлегченной схеме) будет такое же слово, как и у s. Из минимальности $s_1 = l(S, h(g^k(p_i)))$ следует, что $s_1 \sqsubseteq s'$.

Так как $|s_1| > 1$, то, согласно 5.3.2, в S можно найти минимальный путь, содержащий s_1, который соответствует пути с таким же словом в $\mathrm{Evol}(S)$. Этим путем, очевидно, и будет являться s. □

Теперь перейдем к построению алгоритма. Возьмем нумерованную схему Рози S_0 и подберем такое k_0, чтобы размер антиоснастки (S_0, k_0) был больше, чем $\Gamma = 2C_6C_7^2C_8^{C_9}$ (если такого k не существует, то алгоритм выдает ответ "нет".) Далее рассмотрим следующую последовательность антиоснасток (S_i, k_i): если размер антиоснастки (S_{i-1}, k_{i-1}) меньше, чем Γ, то $k_i = k_{i+1} + 1$, $S_i = S_{i-1}$. Если же размер антиоснастки (S_i, k_i) больше либо равен Γ, то $k_i = k_{i-1}$, $S_i = \mathrm{Evol}(S_{i-1})$. Размер первой оснастки обозначим x_0.

Из утверждений леммы леммы 5.4.1, следует, что размер антиоснасток этой последовательности не опускается ниже C_6C_7 и не поднимается выше $X = \max\{C_7x_0, C_7C_8\Gamma\}$. (Теперь можно указать T, возьмем его равным

$2X \max_{a\in A} |g(a)|$.) Следовательно, различных антиоснасток в этой последовательности не может быть больше, чем некоторое алгоритмически определяемое число R. Попытаемся построить первые $R+1$ член последовательности антиоснасток. Если для некоторых k_i и S_i соответствующая антиоснастка не существует, то некоторое рабочее слово не является подсловом W и алгоритм выдает ответ “нет”.

Если первые $R+1$ антиоснастки удалось построить, то среди них есть две одинаковых. Пусть алгоритм выдаст ответ “да”. Докажем, что этот ответ правильный. Согласно леммам 5.4.6 и 5.4.7, существование каждого следующего члена последовательности антиоснасток, а также сам следующий член определяются по предыдущему члену последовательности (то есть без информации о словах на ребрах схемы и о числе k.) Значит, последовательность антиоснасток не оборвется ни на каком члене. Для доказательства корректности работы алгоритма осталось показать, что в последовательности (S_i, k_i) числа k_i неограниченно возрастают.

Предположим противное. Тогда из 5.4.5 следует, что в бесконечной последовательности схем Рози, полученной итерированием детерменированной эволюции из некоторой схемы, масштабы схем в ограничены в совокупности. Следовательно, различных слов на ребрах всех схем последовательности тоже конечное число. А так как облегченных схем в последовательности конечно, то и число самих схем Рози в последовательности конечно.

Для каждой схемы Рози рассмотрим множество слов ее симметричных путей. Согласно 5.3.2, такое множество, построенное для $\mathrm{Evol}(S)$, является собственным подмножеством множества слов симметричных путей схемы S. Отсюда следует, что никакая схема не может быть получена из себя же итерированием операции Evol. Получаем противоречие с конечностью числа схем в последовательности.

Глава 6

Более короткие доказательства алгоритмической разрешимости.

6.1 Проверка периодичности в примитивном случае.

В этом разделе мы дадим более простой алгоритм для следующей задачи:

Дано: *два конечных алфавита A и B, примитивная подстановка $\varphi : A^* \to A^*$, продолжающаяся над буквой $a_1 \in A$ и нестирающий морфизм $h : A^* \to B^*$.*

Определить: *является ли слово $h(\varphi^\infty(a)) = u^\infty$ периодичным и, если является, слово u, являющееся наименьшим периодом.*

В дальнейшем мы считаем, что все входные данные зафиксированы. Это значит, что величину, зависящую только от h и φ, мы будем рассматривать как константу и писать не $C(\varphi, h)$, а просто C.

Напомним два полезных свойства примитивной подстановочной системы [60].

Предложение 6.1.1. *Существуют положительные $K_1 < K_2$ и $\lambda > 1$ такие, что для любых k и a_i выполнено*

$$K_1\lambda^k < |h(\varphi^k(a_i))| < K_2\lambda^k.$$

Предложение 6.1.2. *Существует такое K, что если u_1 и u_2 — подслова слова $h(\varphi^\infty(a_1))$ такие, что $|u_2| > K|u_1|$, то u_1 является подсловом u_2.*

6.1.1 Схемы расположений подслов.

Здесь мы дадим определение *схемы расположений подслов* — объекта, с которым работает описываемый в дальнейшем алгоритм.

Неформальное описание: хочется сказать, что вхождения слов $u_1 = abba$ и $u_2 = baab$ расположены в слове $u_3 = abbabaabaab$ так же, как вхождения слов $v_1 = aa$ и $v_2 = bb$ расположены внутри слова $v_3 = aabbb$. Слово u_1 является началом u_3, а слово u_2 имеет в u_3 два перекрывающихся вхождения: первое начинается там, где кончается единственное вхождение слова u_1, а второе является концом слова u_3. Та же самая фраза остается верной, если вместо u_1, u_2, u_3 написать соответственно v_1, v_2, v_3. Хочется сказать, что у $(\{u_3\}, \{u_1, u_2\})$ такая же схема вхождения, как у $(\{v_3\}, \{v_1, v_2\})$

Определение 6.1.1. *Узлом* конечного слова u длины n будем называть одну из $n+1$ позиций: начало слова (*начальный* узел), конец слова (*конечный* узел) или один из $n-1$ промежутков между его буквами (*обычный* узел).

Каждые пара узлов определяет некоторое подслово в u, возможно, пустое.

Определение 6.1.2. Пусть $U = (U_1, U_2, \dots, U_n)$ и $u = (u_1, u_2, \dots, u_m)$ — два упорядоченных множества конечных слов над одним и тем же алфавитом. Для каждого $U_i \in U$ определим множество *интересных* узлов: начальный, конечный, а также все те узлы, которые являются концами или началами каких-либо вхождений слов из u.

Назовем *схемой расположения подслов* для U и u следующую пару:

1. Упорядоченное множество из n чисел $t_1, t_2, \dots, t_n$, где t_i — это число интересных узлов в слове U_i.

2. Таблицу размером $n \times m$, в клетках которой находятся множества упорядоченных пар чисел, полученные по следующему правилу: в клетке, стоящей на строке i в столбце j парами чисел, не превосходящих t_i, описываются все вхождения слова u_j в U_i (каждая пара чисел задает начало и конец некоторого вхождения).

Такую схему будем обозначать $S(U, u)$.

Пример 6.1.1. $n = 3$, $m = 2$, $U_1 = abcabc$, $U_2 = bcabca$, $U_3 = bbbbbc$, $u_1 = abc$, $u_2 = bc$.

В слове U_1 пять интересных узлов: ${}_1a_2bc_3a_4bc_5$.

В слове U_2 также пять интересных узлов: ${}_1bc_2a_3bc_4a_5$.

А в слове U_3 три интересных узла: ${}_1bbbb_2bc_3$.

Тогда $S((U_1, U_2, U_3), (u_1, u_2))$ состоит из вектора $(5, 5, 3)$ и следующей таблицы:

(1,3), (3,5)	(2,3), (4,5)
(2,4)	(1,2), (3,4)
∅	(2,3)

6.1.2 Схемы вхождений подслов, связанные с итерациями подстановки.

Определение 6.1.3. Упорядоченное множество подслов слова $\varphi^\infty(a_1)$, состоящих из одной или двух букв, назовем *порождающими словами* и будем обозначать G. Согласно, например, 4.1.2, все слова из G алгоритмически находятся, порядок на G выбирается произвольный.

Введем обозначение: $U^k := \{h(\varphi^k(g))|g \in G\}$. Мы считаем, что элементы множества U^k упорядочены так же, как соответствующие элементы G.

Пример 6.1.2. Найдем $S(U^5, U^4)$ для подстановочной системы Фибоначчи.

$A = B = \{a, b\}$, $\varphi(a) = ab$, $\varphi(b) = a$, $h = Id$.

$h(\varphi^\infty(a) = abaababaab\dots$

$G = (a, b, aa, ab, ba)$

$U^4 = (abaababa, abaab, abaababaabaababa, abaababaabaab, abaababaababa)$

U^5 с отмеченными интересными узлами: $({}_1abaab_2aba_3ab_4aab_5,$
${}_1abaab_2aba_3, {}_1abaab_2aba_3ab_4aab_5aba_6ab_7aba_8ab_9aab_{10},$
${}_1abaab_2aba_3ab_4aab_5aba_6ab_7aba_8, {}_1abaab_2aba_3ab_4aab_5aba_6ab_7aab_8)$

Таким образом, $S(U^5, U^4)$ состоит из вектора $(5, 3, 10, 8, 8)$ и следующей таблицы:

(1,3)	(1,2),(2,4),(3,5)	∅	(1,5)	∅
(1,3)	(1,2)	∅	∅	∅
(1,3),(3,6),(5,8)	(1,2),(2,4),(3,5),(5,7),(7,9),(8,10)	(1,6)	(1,5),(5,10)	(3,8)
(1,3),(3,6), (5,8)	(1,2),(2,4),(3,5),(5,7)	(1,6)	(1,5)	(3,8)
(1,3),(3,6)	(1,2),(2,4),(3,5),(5,7),(6,8)	(1,6)	(1,5),(3,8)	∅

Теорема 6.1.1. *Можно алгоритмически (по подстановочной системе) найти такое число d, что из условий $k_1 > l_1 + d$, $k_2 > l_2 + d$ и $S(U^{k_1}, U^{l_1}) = S(U^{k_2}, U^{l_2})$ следует $S(U^{k_1+1}, U^{l_1+}1) = S(U^{k_2+1}, U^{l_2+1})$. Равенство схем вхождения подслов следует понимать как совпадение таблиц.*

Это утверждение можно интерпретировать следующим образом: $S(U^{k+1}, U^{l+1})$ можно однозначно определить по $S(U^k, U^l)$.

Доказательство. Нам потребуется следующее

Предложение 6.1.3. *$S(U^k, U^{l+1})$ можно однозначно определить по $S(U^k, U^l)$.*

Доказательство. Пусть $\varphi(g_i) = a_{i_1}a_{i_2}\dots a_{i_t}$.

Тогда

$$U_i^{l+1} = h(\varphi^{l+1}(g_i)) = h(\varphi^l(a_{i_1}))h(\varphi^l(a_{i_2}))\dots h(\varphi^l(a_{i_t})).$$

Каждое из $h(\varphi^l(a_i))$ принадлежит U^l, поэтому у нас есть проверяемый критерий: пара узлов (p, p') слова $u \in U^k$ задает некоторое вхождение слова U_i^{l+1} тогда и только тогда, когда существует последовательность узлов $p_0 = p, p_1, \dots, p_t = p'$ такая, что (p_0, p_1) — вхождение слова $h(\varphi^l(a_{i_1}))$, пара (p_1, p_2) — вхождение слова $h(\varphi^l(a_{i_2}))$, и так далее, пара (p_{t-1}, p_t) — вхождение слова $h(\varphi^l(a_{i_t}))$. □

Если мы знаем, как именно внутри слова v расположены все вхождения слов A и B, то мы можем определить и все вхождения слова AB. Но если слово v не входит ни в A, ни в B, то в AB оно может как входить, так и не входить. Поэтому переход от $S(U^k, U^l)$ к $S(U^{k+1}, U^l)$ менее тривиальный.

Лемма 6.1.1. *Пусть про конечные слова $A, B, C, u_1, u_2, \dots, u_t$ известно, что $|B| \geqslant 2\max_i(|u_i|)$, а также известна схема $S(AB, B, BC), (u_1 \dots u_t))$. Тогда можно найти $S(ABC, (u_1 \dots u_t))$.*

Доказательство. В слове ABC вхождение B, идущее после A и перед C, назовем *центральной* частью. *Центральными левыми* узлами назовем узлы, являющиеся левыми концами каких-либо вхождений слов из (u_i), целиком лежащих в центральной части. Аналогично введем определение *центральных правых* узлов. Объединение этих множеств назовем *центральными* узлами.

Пусть W — конечное слово и фиксировано какое-то вхождение слова W в ABC. Тогда числам в $S(W,(u_i))$ соответствуют узлы слова ABC, а парам чисел — подслова слова ABC. Среди вхождений B в ABC выберем центральную часть, также естественным образом фиксируются вхождения AB (начало) и BC (конец).

Центральным левым узлам взаимно однозначно соответствуют числа i, встречающееся в таблице $S(B,(u_i))$ в парах первыми членами, то есть как $(i,\cdot)$. Так по схеме $S(B,(u_i))$ можно найти количество центральных левых узлов. Также центральным левым узлам соответствуют некоторые из чисел в таблице $S(AB,(u_i))$, встречающиеся в парах первыми членами.

Пусть какому-то центральному левому узлу в $S(AB,(u_i))$ соответствует число i. Тогда, если в таблице $S(AB,(u_i))$ встречается пара вида $(i',\cdot)$, где $i'>i$, то числу i' также соответствует некоторый центральный левый узел. Таким образом, по схемам вхождения подслов можно найти все числа, которым в $S(AB,(u_i))$ соответствуют центральные левые узлы (ибо известно общее количество таких чисел). Далее можно найти все числа, соответствующие в $S(AB,(u_i))$ центральным правым узлам. Числа из $S(AB,(u_i))$, которым не соответствуют центральные узлы, назовем *левыми периферийными*. Каждому левому периферийному числу соответствует интересный узел в ABC, исключение: число r, кодирующее конец слова AB, может не соответствовать интересному узлу в ABC.

Каждый узел слова ABC, соответствующий левому периферийному числу, являются концом или началом некоторого вхождения u_i, пересекающегося со начальным вхождением слова A. Аналогично можно выяснить, каким числам из $S(BC,(u_i))$ соответствуют центральные узлы и ввести определение *правых периферийных* чисел.

Также по схеме выясняется, каким парам чисел, где одно число из $S(AB,(u_i))$, а второе — из $S(BC,(u_i))$, соответствует один и то же центральный узел.

Так как $|B|\geqslant 2\max(|u_i|)$, то все узлы, соответствующие левым периферийным числам, расположены строго левее, чем узлы, соответствующие правым периферийным числам.

Так как $|B|\geqslant 2\max(|u_i|)$, то каждое вхождение слова из (u_i) в ABC

соответствует хотя какой-то паре чисел из схемы вхождения подслов, а все интересные узлы в ABC либо центральные, либо соответствуют правым периферийным числам, либо соответствуют левым периферийным числам.

Таким образом, восстанавливается схема $ABC, (u_i)$. □

Пример 6.1.3. Пусть дана $S((AB, B, BC), (u_1))$.

Вектор: $(5, 4, 6)$, таблица:

(1,2), (2,3), (3,4)
(2,3)
(2,3), (3,4), (4,5)

Из второй строчки делаем вывод, что есть один центральный левый узел и один центральный правый узел. В первой строчке центральному левому узлу соответствует цифра 2, а центральному правому — цифра 3. В третьей строчке этим узлам соответствуют 1 и 2 соответственно.

Значит, в ABC интересных узлов 7 (два центральных, два, соответствующих левым периферийным числам и три, соответствующих правым периферийным числам).

А таблица будет выглядеть так:

(1,2), (2,3), (3,4), (4,5), (5,6)

Следствие 6.1.1. *Если известно, что любое из слов U^k хотя бы в два раза длиннее любого из слов U^l, то $S(U^{k+1}, U^l)$ однозначно определяется по $S(U^k, U^l)$.*

Доказательство. Нас интересует схема расположения подслов для слова $u \in U^{k+1}$ и множества U^l. Слово u представляется в виде конкатенации $u = u_1u_2 \dots u_t$, где все $u_i \in U^k$ и все $u_iu_{i+1} \in U^k$. Согласно правилу, описанному в лемме, найдем все схемы вхождения подслов вида $S(u_iu_{i+1}u_{i+2}, U^l)$, потом $S(u_iu_{i+1}u_{i+2}u_{i+3}, U^l)$, действуя так, доберемся до $S(u_1u_2 \dots u_t, U^l)$. □

Согласно лемме 6.1.1, можно найти (алгоритмически) такое d, что из условия $k > l + d$ следует, что $\min_{u \in U^k} |u| > 2 \max_{u \in U^l} |u|$.

Пусть $k - l > d$ и $S(U^{k_1}, U^{l_1}) = S(U^{k_2}, U^{l_2})$. Из утверждения 6.1.3 следует, что $S(U^{k_1+1}, U^{l_1}) = S(U^{k_2+1}, U^{l_2})$, а из 6.1.1 следует, что $S(U^{k_1+1}, U^{l_1+1}) = S(U^{k_2+1}, U^{l_2+1})$.

$\square$

6.1.3 Алгоритм.

В дальнейшем полагаем, что d — это число d, возникающее в 6.1.1. Оно может быть эффективно найдено как функция от (φ, h).

Размером схемы вхождения подслов будем называть наибольшее количество элементов в клетке соответствующей таблицы. Так, размер схемы из примера 6.1.2 равен 5.

Найдем такое число N, что $N > \frac{2K_2K\lambda^d}{K_1}$, где K — константа равномерной рекуррентности, а K_1, K_2, λ — числа из утверждения 6.1.1 (заметим, что точно эти числа мы найти не можем, но можем получить на них какую-то алгоритмическую оценку.)

Вот порядок действий для определения периодичности:

1. Выбрать произвольное k_0
2. Последовательно для всех натуральных i строить $S(U^{k_0+i+d}, U^{k_0+i})$, пока не будут выполнены условия пункта 3 или 4.
3. Если какие-то две из построенных схем совпали, то слово $h(\varphi^\infty(a_1))$ непериодично.
4. Если размер некоторой схемы $S(U^{k_0+i+d}, U^{k_0+i})$ больше, чем N, то $h(\varphi^\infty(a_1))$ периодично и длина периода не превосходит $\max_{u \in U^{k_0+i}} |u|$.

Доказательство. Несложно понять, что различных схем размера, не превосходящего N, может встретится лишь конечное число. Значит, либо условия пункта 3, либо условия пункта 4 на каком-то шаге окажутся выполнены.

3. Пусть $S(U^{k_0+i_1+d}, U^{k_0+i_1}) = S(U^{k_0+i_2+d}, U^{k_0+i_2})$. Тогда последовательность схем будет, согласно 6.1.1, периодична с длиной периода $i_2 - i_1$ и, в частности, размеры всех схем в последовательности будут ограничены. Но если $|u_1| > 2|u_2|$, $u_1, u_2 \in (\varphi^\infty(a_1))$, а $(\varphi^\infty(a_1))$ периодично с длиной периода p, то у слова u_2 не менее, чем $\frac{|u_2|}{p}$ вхождений в u_1. Выбирая $u_1 \in U^{i+d}$, $u_2 \in U^i$ для достаточно большого i, получаем противоречние с периодичностью.

4. Пусть нашлись два слова $u_1 \in U^{i+d}$, $u_2 \in U^i$ такие, что у u_2 не менее, чем N вхождений в u_1. По принципу Дирихле, какие-то два вхождения расположены друг относительно друга на расстоянии x, не превосходящем $\frac{|u_1|}{N}$. Согласно 6.1.1, $\frac{|u_1|}{|u_2|} \leqslant \frac{K_2\lambda^d}{K_1}$.

   Из выбора N следует, что $Kx < |u_2|$. Рассмотрим эти два вхождения слова u_2:

$$\overbrace{b_1 \dots b_x \, b_{x+1} b_{x+2} \dots b_{k'-1} b_{k'}}^{u_2} b_{k'+1} \dots b_{k'+x}.$$
$$b_1 \dots b_x \, \underbrace{b_{x+1} b_{x+2} \dots b_{k'-1} b_{k'} b_{k'+1} \dots b_{k'+x}}_{u_2}.$$

   Несложно видеть, что любые две буквы, находящиеся в слове u_2 на расстоянии x, одинаковы, то есть любое подслово длины x слова u_2 является циклическим сдвигом слова $b_1 b_2 \dots b_x$. Из определения константы равномерной рекуррентности, любое подслово слова $h(\varphi^\infty(a_1))$ является подсловом слова u_2 и, следовательно, циклическим сдвигом слова $b_1 b_2 \dots b_x$. Отсюда следует, что $h(\varphi^\infty(a_1))$ периодично с длиной периода x.

$\square$

Сведение общего случая к примитивному выполняется так же, как в главе 4.

6.2 Альтернативный алгоритм для проверки почти периодичности.

Напомним, что что задача о почти периодичности сводится к следующей алгоритмической задаче:

Дано: *1. Три конечных алфавита A, B, C;*
2. Числа K_1, K_2, $1 < \Theta_1 < \Theta_2$;
3. Четыре морфизма $\varphi : A^* \to A^*$, $\psi : A^* \to C^*$, $g : B^* \to B^*$, $h : B^* \to C^*$ такие, что

a) все морфизмы нестирающие;

b) морфизм g продолжается над b_1, морфизм φ примитивен и продолжается над $a_1 \in A$;

c) для некоторого $\lambda \in [\Theta_1; \Theta_2]$ *и всех* $a_i \in A$, $b_j \in B$, $k \in \mathbb{N}$ *выполнено* $K_1\lambda^k < |\psi(\varphi^k(a_i))| < K_2\lambda^k$ *и* $K_1\lambda^k < |h(g^k(b_j))| < K_2\lambda^k$.

d) сверхслово $\psi(\varphi^\infty(a_1))$ *непериодично;*

Определить: *верно ли, что любое подслово сверхслова* $h(g^\infty(b_1))$ *является подсловом сверхслова* $\psi(\varphi^\infty(a_1))$*?*

Теорема 6.2.1. *Эта задача алгоритмически разрешима.*

Упорядоченное множество подслов слова $\varphi^\infty(a_1)$, состоящих из одной или двух букв, назовем *порождающими словами* и будем обозначать A_G. Все слова из A_G алгоритмически находятся, порядок на G выбирается произвольный.

Введем обозначения: $U^k := \{\psi(\varphi^k(x))|x \in A_G\}$, $V^k := \{h(g^k(x))|x \in B\}$. Мы считаем, что элементы множества U^k упорядочены так же, как соответствующие элементы A_G, а элементы V^k упорядочены так же, как буквы в алфавите B.

Предложение 6.2.1. *[60] Существует такое* $K(I)$*, что если* u_1 *и* u_2 *— два подслова сверхслова* $\psi(\varphi^\infty(a_1))$ *такие, что* $|u_2| > K(I)|u_1|$*, то* u_1 *является подсловом* u_2.

Запись $d(I)$ означает, что d — это число, алгоритмически определяемое по входным данным.

Лемма 6.2.1. *Можно алгоритмически найти число* $d(I)$ *такое, что из условий* $k_1 > l_1 + d(I)$, $k_2 > l_2 + d(I)$ *и* $S(U^{k_1}, V^{l_1}) = S(U^{k_2}, V^{l_2})$ *следует* $S(U^{k_1+1}, V^{l_1+1}) = S(U^{k_2+1}, V^{l_2+1})$. *Равенство схем вхождения подслов следует понимать, как совпадение таблиц.*

Доказательство следует из двух лемм, доказанных выше:

Лемма 6.2.2. $S(U^k, V^{l+1})$ *можно однозначно определить по* $S(U^k, V^l)$.

Лемма 6.2.3. *Пусть известна схема* $S(U^k, V^l)$ *и кроме этого известно, что* $\min_i(|U^k|) > 2\max_j(|V_j|)$. *Тогда можно алгоритмически найти* $S(U^{k+1}, V^l)$.

Алгоритм для задачи 6.2.1

.

Размером схемы вхождения подслов будем называть наибольшее количество элементов в клетках соответствующей таблицы. Так, размер схемы из примера 6.1.1 равен 2.

Выберем $D(I)$ такое, что $D(I) > d(I)$ и $D(I) > \log_{\Theta_1}(\frac{K(I)K_2}{K_1})$.

Рассмотрим последовательность схем $S(U^{k+D(I)}, V^k)$ при $k = 1, 2, 3, \ldots$

Положим $N(I) = \lceil \frac{2K_2}{K_1} K(I) \Theta_2^{D(I)} \rceil$.

Из непериодичности $\psi(\varphi^{\infty(a_1)})$ следует

Лемма 6.2.4. *При любом k размер схемы $S(U^{k+D(I)}, V^k)$ не превосходит $N(I)$.*

Теперь можно привести псевдокод алгоритма, решающего задачу 6.2.1:

1. Вычислить $D(I)$ и $N(I)$;
2. Последовательно для всех натуральных k строить $S(U^{k+D(I)}, V^k)$, пока две из построенных схем не совпадут.
3. Если в двух совпавших схемах есть пустые клетки, то ответ «нет», иначе ответ «да».

Правильность работы алгоритма следует из лемм 6.2.1 и 6.2.1.

Глава 7

Заключение.

Хотелось бы упомянуть о возможных обобщениях и путей для дальнейшего исследования.

В главе 3 описывается, как устроена эволюция схем рози для равномерно рекуррентных морфических последовательностей. Интересно, можно ли как-нибудь описать поведение схем Рози для произвольных морфических последовательностей. Это позволило бы приблизиться к важной задаче алгоритмической классификации факторных языков подстановочных систем.

Изучение схем Рози кажется перспективным для исследования таких вопросов, как избегаемость паттернов. Кажется перспективным доказывать, что избегание любого паттерна реализуется некоторой морфической последовательностью.

Нашу теорему можно рассматривать как своего рода обобщение теоремы Вершика-Лившица, что позволяет решить ряд алгоритмичеких проблем. Другим направлением является попытка обобщить теорему Вершика-Лившица на многомерный случай. Некоторые частичные результаты в этом направлении докладывались автором на конференции в Люмини. Удалось обобщить результат Ю. Л. Притыкина с одной стороны и получена версия теоремы Вершика-Лившица с другой стороны для многомерного случая. Была установлена алгоритмическая разрешимость проблемы определения почти периодичности для многомерных подстановок в примитивном случае.

Этот путь представляется перспективным для решения ряда других задач: с одной стороны — для решения алгоритмических задач, с другой стороны — для исследования свойств объектов, управляемых высшими алгебраическими иррациональностями.

Рассматривая эволюцию схем как случайный процесс, можно построить вероятностную меру на факторных языках (и, следовательно, на субшифтах). Интересно изучить свойства типичного в смысле такой меры факторного языка.

Во введении были упомянуты некоторые открытые пока алгоритмические проблемы, связанные с морфическими последовательностями. Автор верит, что все они алгоритмически разрешимы.

Задачи, решаемые в разделах 4 и 5, можно пытаться обобщать на *многомерные самоподобные слова*. Метод, введенный в разделе 6, видится обобщаемым для изучения некоторого класса самоподобных мозаик на плоскости, что может быть интересно в свете работы [94].

Список литературы

1. С. И. Адян. *Проблема Бернсайда и тождества в группах.* Наука, М., 1975, 335 С.

2. С. И. Адян. *Проблема Бернсайда и связанные с ней вопросы.* УМН, **65**:5(395), 2010, С. 5–60.

3. С. В. Алешин, *О свободной группе конечного автомата.*//Вестник Моск. Унив. Сер 1. Мат. и Мех.1983, є4, 12–14.

4. С. В. Алешин, *О суперпозициях автоматных отображений.* // Кибернетика, Киев, 1975, є1, 29–34.

5. С. В. Алешин, *Конечные автоматы и проблема Бернсайда для периодических групп.* // Мат. Заметки 11 (1972), 319–328.

6. Адян С. И., Разборов А. А. *Периодические группы и алгебры Ли.* Успехи мат. наук, 1987, т. 42, no 2, стр. 2–68.

7. В. И. Арнольдь, *Малые знаменатели и проблемы устойчивости движения в классической и небесной механике*// Успехи Мат. Наук, 1963, 18:6(114), стр. 191–192

8. *Бабенко И. К. Проблемы роста и рациональности в алгебре и топологии.* Успехи мат. наук, 1986, vol 41, no 2, pages 96–142.

9. Ю.А. Бахтурин, *Тождества в алгебрах Ли* // Москва, Наука, 1985, 448 стр.

10. А. Я. Белов. *Проблемы бернсайдовского типа, теоремы о высоте и о независимости.* Фундамент. и прикл. матем., 13:5 (2007), С. 19–79; A. Ya. Belov. *Burnside-type problems, theorems on height, and independence.* J. Math. Sci., 156:2 (2009), С. 219–260.

11. А. Я. Белов. *Размерность Гельфанда–Кириллова относительно свободных ассоциативных алгебр.* Матем. сб., 195:12 (2004), С. 3–26.

12. А. Я. Белов, *Классификация слабо нетеровых мономиальных алгебр,* // Фундамент. и прикл. матем., 1:4 (1995), 1085–1089

13. А. Я. Белов, В. В. Борисенко, В. Н. Латышев, *Мономиальные алгебры* // Итоги науки и техники. Совр. Мат. Прил. Тем. Обзоры т. 26 (алг. 4), М. 2002. 35-214.

14. А. Я. Белов, М. И. Харитонов. *Субэкспоненциальные оценки в теореме Ширшова о высоте.* Мат. сб., No 4, 2012, 81–102.

15. А. Я. Белов, И. В. Митрофанов. *Фракталы Рози, цикл лекций.* Видеозаписи летней школа кСовременная математикањ, 2014

16. Голод Е. С., Шафаревич И. Р. *О башне полей классов.* Изв. АН СССР. Сер. мат., 964, т. 28, no 2, стр. 261–272.

17. Голод Е. С. *О нильалгебрах и финитно-аппроксимируемых p-группах.* Изв. АН СССР. Сер. мат., 1964, т. 28, no 2, стр. 273–276.

18. Григорчук Р. И. *К проблеме Милнора о групповом росте.* Докл. АН СССР, 1983, т. 271, є 1, стр. 53–54.

19. Григорчук Р. И. *Степени роста конечно порожденных групп и теория инвариантных средних.* Изв. АН СССР. Сер. мат., 1984, т. 48, no 5, стр. 939–985.

20. Григорчук Р. И. *О рядах Гильберта-Пуанкаре градуированных алгебр, ассоциированных с группами.* Мат. сб., 1989, т. 180, no 2, стр. 207–225.

21. Григорчук Р. И. *К проблеме Милнора о групповом росте.* Функц. анал. и его прил., 1980, 14, 1, стр. 53–54.

22. Зельманов Е. И., Кострикин А. И. *Теорема о сендвичевых алгебрах.* // Тр. Мат. ин-та АН СССР, 1990, 183, стр. 106–111. (РЖМат, 1991, 1А247)

23. Желваков К. А., Слинько А. М., Шестаков И. П., Ширшов А. И. *Кольца, близкие к ассоциативным.* // М.: Наука, 1978, 431 pp.

24. И.А.Иванов-Погодаев, *Машина Минского, свойства нильпотентности и размерность Гельфанда-Кириллова в конечно-определенных полугруппах,* Диссертация на соискание степени к.ф.-м.н. Москва, 2010.

25. Каргаполов М. И. Мерзляков Ю. И. *Основы теории групп,* (3е изд., Наука, 1982) 288стр.

26. А. Т. Колотов, *Апериодические последовательности и функции роста в алгебрах* // Алгебра и логика, 20 (1981), 138-154, 250.

27. А. Т. Колотов, *Алгебры и группы с периодической функцией роста* // Алгебра и логика, 19 (1980), 659-668, 745.

28. Кострикин А. И. *Вокруг Бернсайда.* М.: Наука, 1986, 232 стр.

29. А. Я. Белов, Г. В. Кондаков, *Обратные задачи символической динамики* // Фундаментальная и прикладная математика, T1, |1, 1995

30. Е. И. Зельманов. *Решение ослабленной проблемы Бернсайда для групп нечетного показателя.* Изв. АН СССР. Сер. матем., 54:1 (1990), С. 42–59.

31. В. Б. Кудрявцев, С. В. Алешин, А. С. Подколзин, *Введение в теорию автоматов* // Москва, "Наука", 1985. 320 стр.

32. В. Н. Латышев. *К теореме Регева о тождествах тензорного произведения PI-алгебр.* УМН, Т. 27, $No4$(166), 1972, С. 213–214.

33. В. Н. Латышев. *Нематричные многообразия ассоциативных алгебр.* Диссертация на соискание степени д. ф.-м. н. М., Изд-во Моск. ун-та, 1977.

34. Р. Линдон ,П. Шупп, *Комбинаторная теория групп* // М., Мир, 1980.

35. И. Г. Лысенок. *Бесконечность бернсайдовых групп периода 2^k при* $k > 13$. УМН, 47:2 (1992), 201–202.

36. И. Г. Лысенок. *Бесконечные бернсайдовы группы четного периода.* Изв. РАН. Сер. матем., 60:3 (1996), 3–224.

37. Медведев Ю. А. *Ниль–радикалы конечно порожденных йордановых* PI-*алгебр.* Препр. / Новосибирск: ИМ СО АН СССР, 1985, 24, 30 стр. (РЖ-Мат, 1986, 6А398)

38. Михалчв А. А. *Подалгебры свободных цветных супералгебр Ли.* // Мат. заметки.Ч 1985.ЧТ. 37, є 5.ЧС. 653Ч661.

39. Михалев А. А. *Техника композиции А. И. Ширшова в супералгебрах Ли (некоммутативные базисы Гребнера).* Тр. семинара им. И. Г. Петровского, 1994, т. 18,

40. Мищенко С. П., *Вариант теоремы о высоте для алгебр Ли.* Мат. заметки, 1990, ;7, no 4, стр. 83–89.

41. Ан. А. Мучник, Ю. Л. Притыкин, А. Л. Семенов, *Последовательности, близкие к периодическим*// УМН, 64:5(389) (2009), 21–96

42. П. С. Новиков, С. И. Адян. *О бесконечных периодических группах. I.* Изв. АН СССР. Сер. матем., 32:1 (1968), 212–244.

43. П. С. Новиков, С. И. Адян. *О бесконечных периодических группах. II.* Изв. АН СССР. Сер. матем., 32:2 (1968), 251–524.

44. П. С. Новиков, С. И. Адян. *О бесконечных периодических группах. III.* Изв. АН СССР. Сер. матем., 32:3 (1968), 709–731.

45. П. С. Новиков, С. И. Адян. *Определяющие соотношения и проблема тождества для свободных периодических групп нечетного порядка.* Изв. АН СССР. Сер. матем., 32:4 (1968), 971–979.

46. Ольшанский А. Ю. *Геометрия определяющих соотношений в группах.* сер. Соврем.алгебра. М.: Наука, 1989, 447 стр.

47. Ю. Л. Притыкин, *Алгоритмические свойства последовательностей, близких к периодическим*, Диссертация на соискание степени к.ф.-м.н. Москва, 2009.

48. Попов А. П. *О шпехтовости некоторых многообразий ассоциативных алгебр.* Плиска, 1981, 2, pp. 41–53.

49. М. А. Раскин, *Сверхслова, меры на них и их полупрямые произведения*, Диссертация на соискание степени к.ф.-м.н. Москва, 2014.

50. Я. Г. Синай,*Введение в эргодическую теорию* // М.: ФАЗИС, 1996. 144 с.

51. Скосырский В. Г. *Йордановы алгебры с условием минимальности для двусторонних идеалов.* Препр. / Новосибирск: ИМ СО АН СССР, 1985, 21,

52. В. А. Уфнаровский. *Комбинаторные и асимптотические методы в алгебре.* Итоги науки и техн., Соврем. пробл. мат. Фундам. направления, 1990, *No*57, С. 5–177.

53. В. А. Уфнаровский. *Теорема о независимости и ее следствия.* Матем. сб., 1985, 128(170):1(9), С. 124–132.

54. М.И.Харитонов, *Оценки, связанные с теоремой Ширшова о высоте* Диссертация на соискание степени к.ф.-м.н. Москва, 2015.

55. *Г. Р. Челноков О числе запретов, задающих периодическую последовательность* Моделирование и анализ информационных систем, Т.13, N3 (2007) 66–70

56. А. Л. Чернятьев, *Нормальные базисы и символическая динамика*, Диссертация на соискание степени к.ф.-м.н. Москва, 2008.

57. А. Э. Фрид, *О графах подслов DOL-последовательностей*,// Дискретн. анализ и исслед. опер., сер. 1, 6:4 (1999), 92–103

58. А. Э. Фрид. *Введение в комбинаторику слов.* Лекции, 2011.

59. A.Aberkane, *Words whose complexity satisfies* $\lim p(n)/n = 1$ // Theor. Comp. Sci., 307, (2003), 31-46.

60. J.-P. Allouche and J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations.* Cambridge University Press, 2003.

61. P.Ambro z, L. Hákov ; E. Pelantová, *Properties of 3iet preserving morphisms and their matrices* // Proceedings WORDS 2007, Eds. P. Arnoux, N. Bedaride, J. Cassaigne, (2007), 18–24.

62. P. Arnoux and G. Rauzy [1991], *Representation geometrique des suites the complexite* $2n + 1$ // Bull. Soc. Math. France 119, 199-215.

63. A. Belov. *Some estimations for nilpotence of nil-algebras over a field of an arbitrary characteristic and height theorem* // Communications in algebra, 20 (10):2919-2922, 1992.

64. Belov A., Gateva T., *Radicals of monomial algebras.*, // First International TainanЏMoscow Algebra Workshop. (Taiwan, Republic of China, July 23Џ-August 22, 1994), W de Greyer, Berlin, Proc. of the First Int. Conference held at National Cheng Kung University, Tainan., 1994, 159Џ-169.

65. *Alexei Kanel-Belov, Yakov Karasik, Louis Halle Rowen*, *Computational Aspects of Polynomial Identities: Volume l, Kemer's Theorems*, 2nd Edition, Monographs and Research Notes in Mathematics., Boca Raton, FL: CRC Press, 2015, ISBN: 978-1-4987-2008-3 , 418 pp.,

66. J. Berstel, P. Séébold, *Sturmian words*, in: M. Lothaire (Ed.) //*Algebraic Combinatorics on Words*, Encyclopedia of Mathematics and Its Applications, Vol. 90, Cambridge University Press, Cambridge, 2002 (Chap. 2).

67. J.Berstel, *Resent results on Sturmian words* // Developments in language theory II, 13–24, World Scientific, 1996.

68. V.Berthe, A.Siegel, J. Thuswaldner, *Substitutions, Rauzy fractals and tilings*, //Combinatorics, Automata and Number Theory: Cambridge University Press, pp. 248Џ323.

69. I.I.Bogdanov, Gr.R.Chelnokov. *The maximal length of the period of a periodic word defined by restrictions* arXiv:1305.0460

70. J.Cassaigne. *Special factors with linear subword complexity.* Developments in language theory, II (Magdeburg, 1995), 25-34, World Sci. Publl., River Edge, NJ, 1996.

71. *Cassaigne, J. (F-CNRS-IML); Hubert, P. [Hubert, Pascal] (F-CNRS-IML); Troubetzkoy, S. (F-CNRS-IML) Complexity and growth for polygonal billiards.* Ann. Inst. Fourier (Grenoble) 52 (2002), no. 3, 835–847.

72. *A.L. Chernyat'ev*, *Balanced Words and Dynamical Systems* // Fundamental and Applied Mathematics, 2007, vol. 13, No 5, pp. 213–224

73. *A.L. Chernyat'ev*, *Words with Minimal Growth Function* // Vestnik Mosk. Gos. Univ., 2008.

74. E. Coven, M. Keane, and M. Le Masurier. *A characterization of the morse minimal set up to topological conjugacy.* Ergodic Theory and Dynamical Systems 28 (2008), 1443–1451.

75. Drensky, V., *Free Algebras and PI-algebras: Graduate Course in Algebra*, Springer-Verlag, Singapore (2000).

76. X.Droubay, J.Justin, G.Pirillo,*Episturmian words and some construction of de Luca and Rauzy* // Theoret. Comp. Sci.,(2001) 539–553.

77. F.Durand, $\widehat{\text{HD}}$0L ω-equivalence and periodicity problems in the primitive case (to the memory of G. Rauzy), accepted in J. of Uniform Distribution Theory.

78. F.Durand. *A characterization of substitutive sequences using return words*, Discrete Mathematics 179 (1998), 89–101

79. F.Durand, *Cobham's theorem for substitutions*, J. Eur. Math. Soc. 13 (2011), 1797Ц1812.

80. F. Durand. *Decidability of the HD0L ultimate periodicity problem.* RAIRO – Theoretical Informatics and Applications 47 (2013), 201–214

81. F. Durand. *Decidability of uniform recurrence of morphic sequences*, International Journal of Foundations of Computer Science 24 (2013), 123–146.

82. A. Ehrenfeucht and G. Rozenberg. *Repetition of subwords in DOL languages.* Information and Control, 59(1–3): 13–35, 1983.

83. I. Fagnot. *On the subword equivalence problem for morphic words.* Discrete Appl. Math. 75 (1997), 231–253.

84. I. Fagnot. *Sur les facteurs des mots automatiques.* Theoret. Comput. Sci. 172 (1997), 67–89.

85. S. Ferenczi, L. Zamboni, *Languages of k-interval exchange transformations*, http://iml.univ-mrs.fr/ ferenczi/fz3.pdf

86. *H. Furstenberg. Poincaré reccurence and number theory.* Bull. Amer. Math. Soc., 5:211-234, 1981.

87. R. L. Graham, *Covering the Positive Integers by disjoints sets of the form* $\{[n\alpha+\beta] : n = 1, 2, \ldots\}$ // J. Combin. Theory Ser A15 (1973) 354-358.

88. T. Harju, M. Linna, *On the periodicity of morphisms on free monoids.*// RAIRO Inform. Theor. Appl. 20(1) (1986), pp. 47Ц54.

89. M.Hollander, B.Solomyak. *Two-symbol Pisot substitutions have pure discrete spectrum*, // Ergodic Theory and Dynamical Systems, Volume 23, Issue 02, April 2003, pp 533–540

90. J. Honkala, *A decision method for the recognizability of sets defined by number systems*, RAIRO Inform. Theor. Appl. 20 (1986) 395 40

91. J.Honkala, M.Rigo. *Decidability questions related to abstract numeration systems*, // Discrete Mathematics, Volume 285, Issues 1–3, 6 August 2004, Pages 329–333

92. P.Hubert, *Well balanced sequences* // Theoret. Comput. Sci. 242 (2000) 91Ц108.

93. S. V. Ivanov *The free Burnside groups of sufficiently large exponents.* // Internat. J. Algebra Comput. 4 (1994), no. 1–2

94. I.A.Ivanov-Pogodaev, A.Ya.Kanel-Belov. *Construction of infinite finitely presented nill-semigroup.* arXiv e-print (arXiv:1412.5221)

95. *A.Ya. Kanel-Belov, A.L. Chernyat'ev. Describing the set of words generated by interval exchange transformation.* Comm. in Algebra, Vol. 38, No 7, July 2010, pages 2588–2605.

96. Krause, G.; Lenagan, T.H.: Growth of algebra and Gelfand-Kirillov dimension // Research Notes in Math., Pitman, London, 1985.

97. A.I.Shirshov, *Selected papers of A.I.Shirshov*, eds. Zelmanov, Latyshev, Bokut, Shestakov, Birkhuser Verlag AG, 2009, 3–20, ISBN: 978-3-7643-8857-7/hbk; ISBN 978-3-7643-8858-4/ebook 242 pages

98. P. A. Lavrov. *Number of restrictions required for periodic word in the finite alphabet.* arXiv:1209.0220

99. Livshits, A. N. *Application of Adic representations in the investigations of metric, spectral and topological properties of dynamical systems.* Sanct-Petersburg, 1995, 176 pages.

100. M.Lothaire, *Combinatorics on Words.* // Encyclopedia of Mathematics and its Applications, Addison-Wesley, Reading, MA, 1983, Vol. 17.

101. M.Lothaire, *Algebraic combinatorics on words.* A collective work by Jean Berstel, Dominique Perrin, Patrice Seebold, Julien Cassaigne, Aldo De Luca, Steffano Varricchio, Alain Lascoux, Bernard Leclerc, Jean-Yves Thibon, Veronique Bruyere, Christiane Frougny, Filippo Mignosi, Antonio Restivo, Christophe Reutenauer, Dominique Foata, Guo-Niu Han, Jacques Desarmenien, Volker Diekert, Tero Harju, Juhani Karhumaki and Wojciech Plandowski. With a preface by Berstel and Perrin. Encyclopedia of Mathematics and its Applications, 90. Cambridge University Press, Cambridge, 2002, 504 pp.

102. A. de Luca, *Sturmian words: structure, combinatorics and their arithmetics* // Theoret. Comp. Sci., 183, (1997), 45-82.

103. A. Maes and M. Rigo. *More on generalized automatic sequences.*// Journal of Automata, Languages and Combinatorics, Vol. 7 Iss. 3, Jan 2002, pp. 351 – 376

104. M. Morse, G. A. Hedlund, *Symbolic dynamics II. Sturmian trajectories*, // Amer. J. Math. 62, 1–42. (1940)

105. A.Muchnik. *The definable criterion for definability in Presburger arithmetic and its applications.* Theoret. Comput. Sci. 290(3) (2003), 1433Ц-1444.

106. A.Muchnik, A.Semenov, M.Ushakov, *Almost periodic sequences*, Theoret. Comput. Sci., 304:1-3 (2003), 1–33

107. Francois Nicolas, Yuri Pritykin. *On uniformly recurrent morphic sequences*// International Journal of Foundations of Computer Science, Vol. 20, No. 5 (2009) 919–940

108. J.-J. Pansiot. *Complexit e des facteurs des mots in nis engendr es par morphismes it er es.* In Proceduings of ICALP'84, volume 172 of Lecture Notes in Computer Science, pages 380-389. Springer-Verlag, 1984.

109. J.-J. Pansiot, *Decidability of periodicity for infinite words.* RAIRO Inform. Theor. Appl. 20(1) (1986), pp. 43–46.

110. M. Queffelec. *Substitution dynamical systemsЧspectral analysis*, Vol. 1294 of Lecture Notes in Mathematics. Springer-Verlag, 1987.

111. G.Rauzy, *Nombres algebriques et substitutions*, Bull. Soc. Math. France, 110 (1982), p. 147–178.

112. G. Rauzy, *Mots infnis en arithmetique*, in: *Automata on Infnite Words* // Ecole de Printemps d'Informatique Thfeorique, Le Mont Dore, May 1984, ed. M. Nivat and D. Perrin, Lecture Notes in Computer Science, vol. 192, Springer-Verlag, Berlin etc., pp. 165-171, 1985

113. G. Rauzy, *Suites á termes dans un alphabet fini* // In *Sémin. Théorie des Nombres,* p. 25-01-25-16, 1982-1983, Exposé No 25.

114. G. Rauzy, *Échanges d'intervalles et transformations induites*, (in French), Acta Arith. 34 (1979), p. 315-328.

115. G.Rote, *Sequences with subword complexity* $2n$ // J. Number Theory 46 (1994) 196–213.

116. G. Rozenberg and A. Salomaa // *The Mathematical Theory of L-Systems*, Academic Press, New York etc., 1980

117. Arto Salomaa.*Jewels of Formal Language Theory* // Computer Science Press, 1981.

118. A.Salomaa and M.Soittola. *Automata-theoretic aspects of formal power series.* Springer-Verlag, New York, 1978. Texts and Monographs in Computer Science.

119. M. V. Sapir. *Combinatorial algebra: syntax and semantics.* Springer, 2014.

120. R.Tijdeman, *Fraenkel's conjecture for six sequences.* // Discrete Mathematics, Volume 222, Issue 1-3, 223–234, 2000

121. Vershik, A. M. *The adic realizations of the ergodic actions with the homeomorphisms of the Markov compact and the ordered Bratteli diagrams.* Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 223 (1995), Teor. Predstav. Din. Sistemy, Kombin. i Algoritm. Metody. I, 120–126, 338; translation in J. Math. Sci. (New York) 87 (1997), no. 6, 4054–4058.

122. Vershik, A. M.; Livshits, A. N. *Adic models of ergodic transformations, spectral theory, substitutions, and related topics. Representation theory and dynamical systems*, 185–204, Adv. Soviet Math., 9, Amer. Math. Soc., Providence, RI, 1992.

123. L. Vuillon, *Balanced words* // Bull. Belg. Math. Soc. Simon Stevin 10 (2003), no. 5, 787-805

124. L.Vuillon, *A characterization of Sturmian word by return words* // European Journal of Combinatorics (2001) 22, 263-275.

Работы автора по теме диссертации

125. A.Ya. Kanel-Belov, I.Mitrofanov. *Periodicity of Rauzy scheme and substitutional systems.* eprint arXiv:1107.0185

126. И. В. Митрофанов, *Периодичность морфических слов*, Фундамент. и прикл. матем., 18:4 (2013), 107–119

127. I.Mitrofanov. *Periodicity of Morphic Words.* May 2015, Journal of Mathematical Sciences, Volume 206, Issue 6, pp 679-687

128. A.Ya.Belov, G.V.Kondakov, I.Mitrofanov. *Inverse problems of symbolic dynamics.* Banach Center Publ. 94 (2011), 43 – 60.

129. И. В. Митрофанов, *Почти периодичность морфических слов*, Доклады Академии Наук, 2016, том 467, No 5, с. 519–522