Московский государственный университет имени М. В. Ломоносова
Механико-математический факультет

На правах рукописи
УДК 510.5+519.1

Притыкин Юрий Львович

Алгоритмические свойства последовательностей, близких к периодическим

специальность 01.01.06 —
математическая логика, алгебра и теория чисел

Автореферат
диссертации на соискание ученой степени
кандидата физико-математических наук

Москва — 2009

Работа выполнена на кафедре математической логики и теории алгоритмов Механико-математического факультета Московского государственного университета имени М. В. Ломоносова.

Научный руководитель:	член-корреспондент РАН, профессор Семёнов Алексей Львович
Официальные оппоненты:	доктор физико-математических наук, профессор Волков Михаил Владимирович
	кандидат физико-математических наук Вялый Михаил Николаевич
Ведущая организация:	Институт математики имени Л. С. Соболева СО РАН

Защита диссертации состоится 2 октября 2009 г. в 16 ч. 45 мин. на заседании диссертационного совета Д.501.001.84 при МГУ имени М. В. Ломоносова по адресу: РФ, 119991, Москва, ГСП-1, Ленинские горы, д. 1, МГУ, Механико-математический факультет, ауд. 14-08.

С диссертацией можно ознакомиться в библиотеке Механико-математического факультета МГУ (Главное здание, 14 этаж).

Автореферат разослан 2 сентября 2009 г.

Ученый секретарь
диссертационного совета
Д.501.001.84 при МГУ
доктор физико-математических наук,
профессор А. О. Иванов

Общая характеристика работы

Актуальность темы

Впервые почти периодические последовательности были рассмотрены в связи с символической динамикой. Продолжая и ссылаясь на работы Адамара и Пуанкаре, Морс в 1921 году опубликовал работу[1], в которой определил почти периодические последовательности (он называл их рекуррентными). В 1938 — 1940 годах вышли работы Морса и Хедлунда "Символическая динамика"[2,3], в которых многие свойства почти периодических последовательностей были подробно исследованы.

Последовательность символов конечного алфавита *почти периодическая*, если для каждого конечного подслова u последовательности найдётся такое натуральное число n, что в каждом подслове длины n последовательности найдётся вхождение слова u.

Определение почти периодической последовательности в достаточной степени является комбинаторным. Комбинаторные результаты появляются уже в работах Морса и Хедлунда. Кроме того, конечные и бесконечные слова и до этого изучались с комбинаторной точки зрения. Зарождением соответствующей области — комбинаторики слов — принято считать работы Туэ 1906 и 1912 годов[4,5], в которых он, в частности, изучает свойства последовательности, теперь называемой последовательность Туэ — Морса, и доказывает её бескубность. Отметим, что Туэ не имел в виду каких-то конкретных применений своих результатов и считал рассматриваемые им вопросы представляющими самостоятельный интерес[6]. Сейчас комбинаторика слов — активная область, и в настоящей работе к ней можно отнести многие результаты.

---

[1] M. Morse. Recurrent geodesics on a surface of negative curvature. *Transactions of the American Mathematical Society*, 22:84–100, 1921.

[2] M. Morse, G. A. Hedlund. Symbolic dynamics. *American Journal of Mathematics*, 60(4):815–866, 1938.

[3] M. Morse, G. A. Hedlund. Symbolic dynamics ii: Sturmian trajectories. *American Journal of Mathematics*, 62(1):1–42, 1940.

[4] A. Thue. Über die gegenseitige Lage gleicher Teile gewisser Zeichenreihen. In *Selected mathematical papers of Axel Thue*, pp. 413–478. Oslo, Norway: Universitetsforlaget, 1977.

[5] A. Thue. Über unendliche Zeichenreihen. In *Selected mathematical papers of Axel Thue*, pp. 139–158. Oslo, Norway: Universitetsforlaget, 1977.

[6] J.-P. Allouche, J. Shallit. The ubiquitous Prouhet–Thue–Morse sequence. In *Sequences and their applications, Proceedings of SETA'98*, pp. 1–16, Springer-Verlag, 1999.

Впервые рассмотренные в символической динамике, почти периодические последовательности нашли применение и в математической логике, а именно при рассмотрении вопросов разрешимости теории первого порядка и монадической теории натуральных чисел с отношением порядка, расширенной некоторой конечно-значной функцией, то есть последовательностью букв конечного алфавита. Оказывается, что благодаря применению теории автоматов для достаточно широких классов последовательностей можно получить критерий разрешимости их монадических теорий.

Интересно, что свойства типа почти периодичности полезны и, более того, естественно возникают при решении такого типа задач. Следующее понятие было введено Семёновым[7]. Назовём последовательность *обобщённо почти периодической*, если каждое конечное слово либо входит в неё бесконечное количество раз с ограниченными интервалами между соседними вхождениями, либо входит лишь конечное количество раз. Будем называть последовательности, некоторый суффикс которых почти периодичен, *заключительно почти периодическими*. Заметим, что естественное предположение о том, что обобщённо почти периодические последовательности исчерпываются заключительно почти периодическими, оказывается неверным[8].

В работе Бюхи 1962 года[9] была доказана разрешимость теории $\mathrm{MT}\langle\mathbb{N},<\rangle$ — монадической теории натуральных чисел с отношением порядка, при помощи сопоставления формул теории конечным автоматам на бесконечных последовательностях.

После этого возникает естественный вопрос о разрешимости теорий вида $\mathrm{MT}\langle\mathbb{N},<,x\rangle$, то есть $\mathrm{MT}\langle\mathbb{N},<\rangle$, расширенной некоторой последовательностью x. Конечно, теория $\mathrm{MT}\langle\mathbb{N},<,x\rangle$ разрешима, если последовательность x выразима в исходной теории, но запас таких последовательностей невелик — это периодические последовательности, возможно, с предпериодом. Оказывается, что можно получить критерий разрешимости теории $\mathrm{MT}\langle\mathbb{N},<,x\rangle$, если ограничиться рассмотрением обобщённо

[7] А. Л. Семёнов. О некоторых расширениях арифметики сложения натуральных чисел. *Известия АН СССР, серия математическая*, 43(5):1175–1195, 1979.

[8] Ю. Л. Притыкин. Конечно-автоматные преобразования строго почти периодических последовательностей. *Математические заметки*, 80(5):751–756, 2006.

[9] J. R. Büchi. On a decision method in restricted second-order arithmetic. In *Proceedings of International Congress for Logic, Methodology and Philosophy of Science*, pp. 1–11. Stanford University Press, 1962.

почти периодических последовательностей.

Пусть последовательность x обобщённо почти периодическая. Аналогом периода и предпериода для таких последовательностей является регулятор почти периодичности. *Регулятором почти периодичности* последовательности x называется такая функция $\mathrm{R}_x\colon \mathbb{N} \to \mathbb{N}$, которая на числе n равна минимальному такому l, что всякое слово длины n, входящее в x конечное количество раз, может входить только в начальный отрезок x длины l и не может входить далее, а также каждое слово длины n, входящее в x бесконечно много раз, входит в каждый отрезок длины l. Назовём последовательность *эффективно обобщённо почти периодической*, если она является вычислимой обобщённо почти периодической последовательностью, и некоторая оценка сверху на регулятор почти периодичности этой последовательности вычислима. Как доказал Семёнов[10], теория $\mathrm{MT}\langle\mathbb{N},<,x\rangle$ обобщённо почти периодической последовательности x разрешима тогда и только тогда, когда x эффективно обобщённо почти периодична.

Для доказательства этого результата необходимо разобраться в том, как ведёт себя конечный автомат, на вход которому подаётся обобщённо почти периодическая последовательность. Этот вопрос связан со следующей серией вопросов — понять, как меняются или насколько сохраняются различные свойства последовательностей при применении к ним различного вида преобразований. Естественно рассматривать простейшие алгоритмические преобразования — *конечно-автоматные*, то есть осуществляемые машиной с конечной памятью.

Основным средством в получении вышеприведённого критерия разрешимости монадических теорий обобщённо почти периодических последовательностей стал результат о сохраняемости множества обобщённо почти периодических последовательностей при конечно-автоматных преобразованиях, а также эффективная его версия, дающая оценку на регулятор образа при имеющемся регуляторе исходной последовательности[10,11].

Однако этот результат представляет и самостоятельный интерес. Известны и другие результаты о сохранении классов последовательностей под действием конечно-автоматных преобразований. В частности,

[10] А. Л. Семёнов. Логические теории одноместных функций на натуральном ряде. *Известия АН СССР, серия математическая*, 47(3):623–658, 1983.

[11] An. Muchnik, A. Semenov, M. Ushakov. Almost periodic sequences. *Theoretical Computer Science*, 304(1–3):1–33, 2003.

известно[12], что класс автоматных последовательностей (см. ниже) сохраняется под действием равномерных конечно-автоматных преобразований. Под равномерными мы понимаем такие конечно-автоматные преобразования, которые, получив на вход букву, всегда выдают на выход ровно одну букву. Морфические последовательности (см. ниже) также сохраняются под действием конечно-автоматных преобразований[13].

Таким образом, представляется естественным более детально изучить свойства замкнутости классов последовательностей с различными свойствами типа почти периодичности относительно конечно-автоматных преобразований.

Другие важные для нас классы последовательностей, обладающих свойствами, близкими к свойствам периодических последовательностей, — это классы автоматных и морфических последовательностей.

Любая периодическая последовательность может быть порождена машиной с конечной памятью — достаточно иметь информацию о предпериоде и периоде последовательности. И наоборот, любая машина с конечной памятью, печатающая символы конечного алфавита, порождает заключительно периодическую последовательность. Действительно, во время её работы в какой-то момент конфигурация машины полностью совпадёт с какой-то из уже встречавшихся, и так начнётся период в выходной последовательности. Если разрешить машине обращаться к символам, написанным ранее, класс порождаемых последовательностей существенно возрастёт. При некоторых ограничениях на порядок, в котором ранее написанные символы читаются снова, мы получим класс автоматных последовательностей, введённый Кобхэмом[12].

Формально определим автоматные последовательности двумя эквивалентными способами.

Рассмотрим конечный автомат, действующий на словах в алфавите $\Sigma_k = \{0, 1, \ldots, k-1\}$. Каждому состоянию автомата соответствует буква в некотором другом алфавите A (разным состояниям могут соответствовать одинаковые буквы). Автомат действует так: получает на вход слово в алфавите Σ_k, производит вычисления и выдаёт ту букву алфавита A, которая соответствует последнему состоянию в вычислении. Последовательность x букв алфавита A называется *k-автоматной*, если

[12] A. Cobham. Uniform tag sequences. *Mathematical Systems Theory*, 6:164–192, 1972.

[13] F. M. Dekking. Iteration of maps by an automaton. *Discrete Mathematics*, 126(1–3):81–86, 1994.

существует конечный автомат вышеуказанного вида, который, будучи запущенным на числе n, записанном в k-ичной системе счисления, выдаёт букву $x(n)$. Когда ясно или неважно, о каком k идёт речь, приставку "k-" мы будем опускать.

Теперь опишем другой подход. Пусть A, B — конечные алфавиты. Рассмотрим операцию подстановки, которая каждую букву алфавита A отображает в слово алфавита B. Результатом применения подстановки к конечному слову или бесконечной последовательности в алфавите A будет конкатенация образов букв, составляющих это слово или последовательность. Такие отображения подстановки называются *морфизмами*. Морфизм называется *k-равномерным*, если длины образов всех букв совпадают и равны k. Морфизм *нестирающий*, если образы всех букв непусты.

Ограничимся теперь рассмотрением ситуации, когда алфавиты A и B совпадают. Нас интересуют бесконечные последовательности, являющиеся неподвижными точками морфизмов, причём неподвижными точками достаточно общего вида. Пусть ϕ — морфизм в алфавите A, и s — буква алфавита A, такая что слово $\phi(s)$ начинается с s. Будем итерировать ϕ на s, получим слова s, $\phi(s)$, $\phi^2(s)$, $\phi^3(s)$, $\phi^4(s)$, $\ldots$, каждое из которых начинается с предыдущего. Если длина слова $\phi^n(s)$ стремится к бесконечности с ростом n, можно корректно определить бесконечную последовательность $\phi^\infty(s)$, началами которой являются слова $\phi^n(s)$ для любого n. Последовательности $x = \phi^\infty(s)$, которые можно получить таким способом, называются *чисто морфическими*. Несложно видеть, что чисто морфическая последовательность, порождённая морфизмом ϕ, является неподвижной точкой под действием этого морфизма. Последовательности, получающиеся из чисто морфических отождествлением некоторых символов, называются *морфическими*. Как доказал Кобхэм[12], k-автоматные последовательности — это в точности морфические последовательности, полученные из порождённых k-равномерными морфизмами чисто морфических последовательностей.

Морфические последовательности, точнее, тесно связанные с ними $D0L$-последовательности и, более общо, L-системы (системы Линденмайера), впервые появились в работах математика и биолога Линденмайера для описания процессов развития живых организмов[14]. Однако

[14]G. Rozenberg, A. Salomaa, editors. *Lindenmayer Systems: Impacts on Theoretical Computer Science, Computer Graphics, and Developmental Biology*, Springer-Verlag,

впоследствии морфические последовательности стали рассматриваться и в теории динамических систем, комбинаторике слов, теоретической информатике[15,16,17]. Автоматные последовательности неявно были впервые рассмотрены в работе Бюхи о связях монадических теорий и конечных автоматов[18], первой работой, посвящённой их систематическому изучению, была работа Кобхэма[12].

Множества морфических и почти периодических последовательностей находятся в общем положении. Заметим также, что морфические последовательности можно конечно описать — для этого нужно описать алфавит, морфизм, букву, на которой итерируется морфизм, и последующий способ отождествления букв в чисто морфической последовательности. Поэтому можно ставить алгоритмические вопросы о морфических последовательностях. Довольно естественным является вопрос о существовании алгоритма для следующей проблемы.

Проблема распознавания почти периодичности морфических последовательностей (ППМ):

Дано: конечное описание морфической последовательности.

Определить: является ли эта последовательность почти периодической.

Многие алгоритмические задачи, связанные с морфизмами, являются довольно сложными и интересными[16]. Один из самых известных примеров такой задачи — проблема соответствий Поста (Post correspondence problem), которая неразрешима[19,16]. Про проблему ППМ до сих пор неизвестно, является ли она разрешимой, хотя гипотеза о разрешимости является довольно правдоподобной. В связи с этим представляется интересным исследовать некоторые частные случаи проблемы ППМ, возникающие при ограничении множества входов на морфические последовательности специального вида.

1992.

[15]J.-P. Allouche, J. Shallit. *Automatic Sequences: Theory, Applications, Generalizations.* Cambridge University Press, 2003.

[16]T. Harju, J. Karhumäki. Morphisms. In G. Rozenberg, A. Salomaa, editors, *Handbook of formal languages*, vol. 1, pp. 439–510, Springer-Verlag, 1997.

[17]M. Queffélec. *Substitution Dynamical Systems—Spectral Analysis*, *Lecture Notes in Mathematics*, vol. 1284, Springer-Verlag, 1987.

[18]J. R. Büchi. Weak second-order arithmetic and finite automata. *Z. Math. Logik Grundlagen Math.*, 6:66–92, 1960.

[19]M. Sipser. *Introduction to the Theory of Computation.* PWS Publishing Company: Boston, 2nd edition, 2005.

Тесто связанный с проблемой ППМ вопрос или даже вариант формулировки — получение по возможности как можно более компактного и эффективно проверяемого критерия почти периодичности для морфических последовательностей. Вопрос о получении такого критерия для чисто морфических последовательностей был поставлен как открытый в монографии Аллуша и Шаллита[15], раздел 10.12, проблема 5.

Одной из простейших и наиболее естественных характеристик бесконечных последовательностей является их подсловная сложность. *Подсловной сложностью* последовательности x называется такая функция $\mathrm{p}\colon \mathbb{N} \to \mathbb{N}$, что $\mathrm{p}(n)$ равно количеству слов длины n, входящих в последовательность x. Для последовательностей в алфавите из m символов подсловная сложность может варьироваться от 1 до m^n. Как доказано в работе Морса и Хедлунда[2], подсловная сложность последовательности ограничена тогда и только тогда, когда последовательность является заключительно периодической.

Асимптотическое поведение подсловной сложности бесконечных последовательностей широко изучалось, в том числе и в контексте теории динамических систем[20,21]. В серии работ, завершающейся работой Пансио 1984 года[22], доказано, что подсловная сложность чисто морфической последовательности может удовлетворять одной из пяти следующих асимптотик: $O(1)$, $\Theta(n)$, $\Theta(n \log\log n)$, $\Theta(n \log n)$, $\Theta(n^2)$. В работе Пансио 1985 года[23] показано, что существуют примеры морфических последовательностей с подсловной сложностью вида $\Theta(n^{1+\frac{1}{k}})$ для каждого $k \in \mathbb{N}$. После этого про подсловную сложность морфических последовательностей произвольного вида долгое время не было известно ничего нового. Наконец, Девятовым было доказано[24], что подсловная сложность морфи-

[20] J.-P. Allouche. Sur la complexité des suites infinies. *Bulletin of the Belgian Mathematical Society—Simon Stevin*, 1(2):133–143, 1994.

[21] S. Ferenczi. Complexity of sequences and dynamical systems. *Discrete Mathematics*, 206(1):145–154, 1999.

[22] J.-J. Pansiot. Complexité des facteurs des mots infinis engendrés par morphismes itérés. In *Proceedings of the 11th International Colloquium on Automata, Languages, and Programming (ICALP'84), Antwerp, Belgium, Lecture Notes in Computer Science*, vol. 172, pp. 380–389, Springer-Verlag, 1984.

[23] J.-J. Pansiot. Subword complexities and iteration. *Bulletin of the European Association for Theoretical Computer Science*, 26:55–62, 1985.

[24] R. Devyatov. On subword complexity of morphic sequences. In *Proceedings of the 3rd International Computer Science Symposium in Russia (CSR 2008), Moscow, Russia, Lecture Notes in Computer Science*, vol. 5010, pp. 146–157, Springer-Verlag, 2008.

ческой последовательности имеет один из следующих видов: $O(n \log n)$ или $\Theta(n^{1+\frac{1}{k}})$ для некоторого $k \in \mathbb{N}$. Таким образом, для полного описания возможных асимптотик подсловной сложности морфической последовательности осталось разобраться подробнее со случаем $O(n \log n)$.

В случае автоматных последовательностей ситуация существенно проще: подсловная сложность автоматной последовательности не более чем линейна[12].

Несложно показать, что подсловная сложность почти периодической последовательности в алфавите из m символов не может быть больше $m^{\alpha n}$ для некоторого фиксированного $\alpha < 1$ [21]. При этом можно показать, что для любого $\beta < 1$ существует почти периодическая последовательность в алфавите из m символов с подсловной сложностью не меньше $m^{\beta n}$ [25].

Оказывается, подсловная сложность последовательности, являющейся одновременно морфической и почти периодической, устроена гораздо проще. Как мы доказываем в настоящей работе, подсловная сложность таких последовательностей не более чем линейна. Отметим, что линейность подсловной сложности почти периодических чисто морфических последовательностей была известна ранее[17].

Наконец, заключительной темой настоящей работы является изучение свойств понятия меры апериодичности бесконечной последовательности.

Периодические последовательности имеют самую простую структуру среди всех последовательностей над конечным алфавитом, поэтому естественно было бы научиться измерять, насколько данная последовательность далека от любой периодической. Мы вводим[26] понятие меры апериодичности бесконечной последовательности. Неформально, мера апериодичности последовательности — это максимальное такое число α, что последовательность с любым своим нетривиальным сдвигом имеет хотя бы долю α различий. Насколько известно автору, систематического исследования этого понятия и его свойств ранее не проводилось, извест-

[25] Yu. Pritykin. Information in infinite words. In *Proceedings of the 6th International Conference on Words (Words 2007), Marseille, France*, pp. 254–261. Institute de Mathématiques de Luminy, Marseille, France, 2007.

[26] Yu. Pritykin, J. Ulyashkina. Aperiodicity measure for infinite sequences. In *Proceedings of the 4th International Computer Science Symposium in Russia (CSR 2009), Novosibirsk, Russia*, Lecture Notes in Computer Science, vol. 5675, pp. 274–285. Springer-Verlag, 2009.

ны только некоторые отдельные результаты. Мы получаем некоторые простейшие свойства и вычисляем меру апериодичности некоторых конкретных последовательностей.

Цель работы

1) Изучение свойств замкнутости различных классов последовательностей со свойствами типа почти периодичности относительно действия конечно-автоматных преобразований;
2) исследование возможностей эффективного распознавания и вычисления свойств и характеристик последовательностей со свойствами типа почти периодичности;
3) получение эффективных алгоритмов разрешения свойства почти периодичности для некоторых подклассов класса морфических последовательностей;
4) изучение поведения подсловной сложности последовательностей, являющихся одновременно почти периодическими и морфическими;
5) изучение простейших свойств понятия меры апериодичности бесконечной последовательности.

Основные методы исследования

В диссертации используются методы теории алгоритмов, теории автоматов и комбинаторики на словах.

Научная новизна

Результаты работы являются новыми и состоят в следующем.

1) Доказано, что классы заключительно почти периодических и заключительно рекуррентных последовательностей сохраняются под действием конечно-автоматных преобразований.
2) Получена оценка сверху на длину конечного слова, которое достаточно отрезать от образа почти периодической последовательности под действием конечно-автоматного преобразования, чтобы получить почти периодическую последовательность.
3) Доказано, что по обобщённо почти периодической последовательности и её регулятору почти периодичности невозможно распознать,

является ли последовательность заключительно почти периодической, и получены другие результаты о невычислимости некоторых свойств обобщённо почти периодических последовательностей.

4) Получены полиномиальные по времени работы алгоритмы, определяющие по автоматной или чисто морфической последовательности, полученной из нестирающего морфизма, является ли эта последовательность почти периодической.
5) Доказано, что подсловная сложность морфической почти периодической последовательности не более чем линейна.
6) Доказаны простейшие свойства меры апериодичности бесконечных последовательностей и вычислены значения меры апериодичности для некоторых хорошо известных последовательностей.

Теоретическая и практическая ценность

Работа носит теоретический характер. Результаты диссертации могут быть использованы при дальнейшем изучении свойств почти периодических, морфических и автоматных последовательностей.

Апробация результатов

Результаты работы докладывались на следующих научно-исследовательских семинарах и научных конференциях.

- Колмогоровский семинар кафедры математической логики и теории алгоритмов под руководством профессора Н. К. Верещагина, к. ф.-м. н. А. Е. Ромащенко, члена-корреспондента РАН, профессора А. Л. Семёнова, к. ф.-м. н. А. Шеня, 2006 — 2009 гг.
- Семинар “Алгоритмические вопросы алгебры и логики” под руководством академика РАН С. И. Адяна, 2006 — 2007 гг.
- Научно-исследовательский семинар кафедры математической логики и теории алгоритмов под руководством академика РАН С. И. Адяна, члена-корреспондента РАН Л. Д. Беклемишева и профессора В. А. Успенского, 2009 г.
- XXVIII Конференция молодых учёных, механико-математический факультет МГУ им. М. В. Ломоносова, 2006 г.
- Семинар по словам и автоматам в рамках 1-го международного симпозиума по компьютерным наукам в России, Санкт-Петербург, 2006 г.

- Workshop on Algorithms on Words, Турку, Финляндия, 2007 г.
- International Conference “Automata: from Mathematics to Applications” (AutoMathA 2007), Палермо, Италия, 2007 г.
- 11-th International Conference “Developments in Language Theory” (DLT 2007), Турку, Финляндия, 2007 г.
- Семинар по словам, автоматам и динамике в рамках 2-го международного симпозиума по компьютерным наукам в России, Екатеринбург, 2007 г.
- 6-th International Conference on Combinatorics on Words (Words 2007), Марсель, Франция, 2007 г.
- Конференция “Фундаментальная математика в работах молодых учёных”, посвящённая десятилетию всероссийского конкурса Мёбиуса математических работ, Москва, 2007 г.

Публикации

Основные результаты диссертации опубликованы в 6 работах автора, список которых приведён в конце автореферата [1–6].

Структура и объем диссертации

Диссертация состоит из 6 глав и списка литературы. Полный объём диссертации — 96 страниц, список цитируемой литературы содержит 48 наименований.

Краткое содержание работы

Глава 1 является введением диссертации. Во введении описываются цели и задачи работы, обосновывается её актуальность и научная новизна.

В **главе 2** вводятся необходимые понятия и обсуждаются их основные свойства.

Для конечного алфавита A через A^* обозначим множество всех слов, а через $A^{\mathbb{N}}$ множество всех последовательностей в этом алфавите. Подслова последовательности x вида $x[0,i]$ называются *префиксами* x, последовательности вида $x(i)x(i+1)x(i+2)\dots$ — *суффиксами* x и обозначаются $x[i,\infty)$. Через $|u|$ будем обозначать длину слова u.

Последовательность x называется *почти периодической*, если для каждого её подслова u найдётся такое натуральное l, что на каждом отрезке длины l последовательности x найдётся вхождение слова u. Тем самым, любое слово, входящее в почти периодическую последовательность, входит в неё бесконечное количество раз. Будем называть последовательность x *заключительно почти периодической*, если некоторый её суффикс почти периодичен. Последовательность x называется *обобщённо почти периодической*, если для каждого её подслова u, входящего в неё бесконечное число раз, найдётся такое натуральное l, что на каждом отрезке длины l последовательности x найдётся вхождение слова u. Отметим, что множество заключительно почти периодических последовательностей является собственным подмножеством множества обобщённо почти периодических последовательностей.

Регулятором почти периодичности обобщённо почти периодической последовательности x назовём функцию $\mathrm{R}_x\colon \mathbb{N} \to \mathbb{N}$, которая на числе n равна минимальному такому l, что каждое слово длины n, которое входит в x бесконечное количество раз, встретится на любом отрезке длины l последовательности x, а также любое слово длины n, которое входит в x конечное количество раз, не входит в $x[l, \infty)$. Часто вместо регулятора достаточно рассматривать только какую-то верхнюю оценку на него, то есть функцию $f\colon \mathbb{N} \to \mathbb{N}$, такую что $f(n) \geqslant \mathrm{R}_x(n)$ для всех n — в этом случае пишут $f \geqslant \mathrm{R}_x$.

Пусть A, B — конечные алфавиты. Отображение $\phi\colon A^* \to B^*$ называется *морфизмом*, если для любых $u, v \in A^*$ выполнено $\phi(uv) = \phi(u)\phi(v)$. Ясно, что морфизм полностью определяется своими значениями на однобуквенных словах. Морфизм *нестирающий*, если $|\phi(a)| \geqslant 1$ для всех $a \in A$. Морфизм называется *k-равномерным*, если $|\phi(a)| = k$ для всех $a \in A$. 1-равномерный морфизм называется *кодированием*. Естественным образом морфизмы продолжаются на бесконечные последовательности: если x — последовательность букв алфавита A, по определению положим $\phi(x) = \phi(x(0))\phi(x(1))\phi(x(2))\ldots$

Пусть $\phi\colon A^* \to A^*$ — некоторый морфизм, и пусть $\phi(s) = su$ для некоторых $s \in A$, $u \in A^*$. Тогда для всех натуральных $m < n$ слово $\phi^n(s)$ начинается со слова $\phi^m(s)$, так что можно корректно определить $\phi^\infty(s) = \lim_{n\to\infty} \phi^n(s) = su\phi(u)\phi^2(u)\phi^3(u)\ldots$. Если $\phi^n(u) \neq \Lambda$ для всех n, то $\phi^\infty(s)$ бесконечна. В этом случае говорят, что морфизм ϕ *продолжаем на s*. Последовательности вида $h(\phi^\infty(s))$, где $h\colon A \to B$ — кодирование, называются *морфическими*, вида $\phi^\infty(s)$ — *чисто морфическими*.

Последовательности вида $h(\phi^\infty(s))$, где $h\colon A \to B$ — кодирование, а ϕ — k-равномерный морфизм, называются *k-автоматными*. Когда ясно или неважно, о каком k идёт речь, приставку "k-" опускают.

Конечно-автоматным преобразователем назовём совокупность $M = \langle A, B, Q, \tilde{q}, \lambda, \mu \rangle$, где A и B — конечные множества, называемые соответственно входной и выходной алфавит, Q — конечное множество состояний, $\tilde{q} \in Q$ — выделенное состояние, называемое начальным, и

$$\lambda\colon Q \times A \to B^*, \qquad \mu\colon Q \times A \to Q$$

— функции переходов. Функции переходов можно естественным образом продолжить на $Q \times A^*$ по индукции: $\mu(q, ua) = \mu(\mu(q, u), a)$ и $\lambda(q, ua) = \lambda(q, u)\lambda(\mu(q, u), a)$ для $q \in Q$, $u \in A^*$, такого что $u \neq \Lambda$, и $a \in A$.

Пусть $x \in A^{\mathbb{N}}$. Последовательность $(p_n)_{n=0}^\infty$ элементов множества Q назовём *ходом преобразователя M на x*, если $p_0 = \tilde{q}$ и для каждого n выполняется $p_{n+1} = \mu(p_n, x(n))$. Ясно, что ход существует и единствен. Последовательность $M(x)$, определяемую как $M(x)(n) = \lambda(p_n, x(n))$, где $(p_n)_{n=0}^\infty$ — ход преобразователя M на x, назовём *образом последовательности x под действием M*.

Если для каждых $a \in A$, $q \in Q$ выполнено $|\lambda(q, a)| = 1$, то преобразователь M называется *равномерным*. Несложно видеть, что применение произвольного конечно-автоматного преобразователя к последовательности можно представить как последовательное применение равномерного конечно-автоматного преобразователя и некоторого морфизма. Поэтому часто для упрощения ситуации мы ограничиваемся рассмотрением равномерных конечно-автоматных преобразователей. Если $[i, j]$ — вхождение слова u в последовательность x, причём $p_i = q$, где $(p_n)_{n=0}^\infty$ — ход преобразователя M на x, то будем говорить, что преобразователь M *подходит* к этому вхождению слова u *в состоянии* q.

Назовём конечно-автоматный преобразователь $M = \langle A, B, Q, \tilde{q}, \lambda, \mu \rangle$ *обратимым*, если для каждых $q \in Q$ и $a \in A$ существует ровно одно состояние $q' \in Q$, такое что $\mu(q', a) = q$. Другими словами, в таком преобразователе каждая буква входного алфавита осуществляет взаимно однозначное отображение множества состояний в себя. Находясь в некотором состоянии и зная последовательность предыдущих входных символов, можно восстановить и последовательность пройденных состояний (в этом и заключается свойство обратимости).

Подсловной сложностью последовательности x называется такая функция $\mathrm{p}_x\colon \mathbb{N} \to \mathbb{N}$, что $\mathrm{p}_x(n)$ равно количеству слов длины n, входящих в последовательность x.

В последующих главах излагаются результаты диссертации.

Глава 3 посвящена конечно-автоматным преобразованиям. Из теоремы Семёнова следует, что образ почти периодической последовательности под действием конечно-автоматного преобразования является обобщённо почти периодической последовательностью. Оказывается, верно более сильное утверждение.

Теорема 1. *Почти периодические последовательности под действием конечно-автоматных преобразований переходят в заключительно почти периодические.*

Рассмотрим такой вопрос об эффективизации теоремы 1. Допустим, нам известны конечно-автоматный преобразователь, почти периодическая последовательность и оценка на её регулятор почти периодичности. В соответствии с теоремой 1 от образа этой последовательности под действием этого преобразователя можно отрезать некоторый начальный отрезок, так что получится почти периодическая последовательность. Можно ли оценить сверху длину этого отрезка? Ан. А. Мучник в 2005 г. высказал гипотезу о том, что это невозможно сделать эффективно, однако результат следующей теоремы 2 показывает, что такая оценка существует. Для произвольной функции $g\colon \mathbb{N} \to \mathbb{N}$ обозначим $\underbrace{g \circ g \circ \dots \circ g}_{m}$ через g^m.

Теорема 2. *Пусть M — равномерный конечно-автоматный преобразователь с m состояниями, и x — почти периодическая последовательность. Тогда образ x под действием M становится почти периодической последовательностью при удалении префикса длиной не более $\mathrm{R}_x^m(1) + \mathrm{R}_x^{m-1}(1) + \dots + \mathrm{R}_x(1)$.*

Одним из основных шагов в доказательстве теоремы 2, представляющим и самостоятельный интерес, является следующий результат.

Предложение 3. *Почти периодические последовательности под действием обратимых конечно-автоматных преобразований переходят в почти периодические.*

Далее мы доказываем замкнутость относительно конечно-автоматных преобразований для рекуррентных последовательностей, а также относительно почти обратимых конечно-автоматных преобразований для обобщённо почти периодических последовательностей. Мы доказываем, однако, что множество обобщённо почти периодических последовательностей не замкнуто относительно более широкого класса стековых конечно-автоматных преобразований.

Глава 4 посвящена вопросам вычислимости на бесконечных последовательностях. В связи с теоремой 2 возникает следующий более общий вопрос: пусть нам дана заключительно почти периодическая последовательность и оценка сверху на её регулятор почти периодичности. Можно ли только на основании этих данных получить оценку на тот префикс, который достаточно отрезать от этой последовательности, чтобы получить почти периодическую последовательность? Мы доказываем, что такой оценки не существует. Другие теоремы главы 4 являются дальнейшими примерами результатов, в которых доказывается невозможность по обобщённо почти периодической, заключительно почти периодической или почти периодической последовательности и её регулятору почти периодичности найти некоторую характеристику или проверить некоторое свойство. Наиболее общий результат утверждает, что для любого счётного множества M последовательностей, содержащего все периодические последовательности, по обобщённо почти периодической последовательности и её регулятору невозможно определить, принадлежит ли эта последовательность множеству M. В качестве следствий из этих результатов доказывается, что некоторые множества последовательностей не являются регулярными.

В **главе 5** исследуются свойства морфических последовательностей. Основной обсуждаемый вопрос — разрешимость проблемы распознавания почти периодичности для морфических последовательностей. Вопрос по-прежнему остаётся открытым, но мы получаем следующие два более слабых результата об эффективной разрешимости почти периодичности для морфических последовательностей специального вида.

Теорема 4. *Существует полиномиальный по времени алгоритм, определяющий по чисто морфической последовательности, является ли она почти периодической.*

Теорема 5. *Существует полиномиальный по времени алгоритм, определяющий по автоматной последовательности, является ли она по-*

чти периодической.

Перед доказательством этих результатов мы получаем критерии почти периодичности для чисто морфических и автоматных последовательностей.

В заключительном разделе 5.4 главы 5 мы получаем следующий результат.

Теорема 6. *Подсловная сложность почти периодических морфических последовательностей не более чем линейна.*

Глава 6 посвящена мере апериодичности бесконечных последовательностей.

Расстояние Безиковича определяется как $d_B(x, y) = \liminf_{n\to\infty} \frac{1}{n}\#\{i : 0 \leqslant i \leqslant n-1, x(i) \neq y(i)\}$. Определим *меру апериодичности* $\mathrm{AM}(x) = \inf\{d_B(x, L^n x) : n \geqslant 1\}$, где L обозначает операцию левого сдвига. Другими словами, $\mathrm{AM}(x)$ — это максимальное такое число от 0 до 1, что можно утверждать, что у последовательности x с любым её собственным сдвигом хотя бы доля $\mathrm{AM}(x)$ символов различны.

Мы приводим результаты работы [6], которые получены автором диссертации. А именно, мы получаем общие верхнюю и нижнюю оценки для меры апериодичности последовательностей для алфавита фиксированного размера. Мы доказываем, что мера апериодичности последовательностей Штурма (ещё одно простейшее обобщение периодических последовательностей) равна 0. После этого мы доказываем, что мера апериодичности последовательности Туэ — Морса равна 1/3, а также получаем верхнюю оценку на меру апериодичности одного естественного обобщения последовательности Туэ — Морса — последовательностей Пруэ.

Благодарности

Автор выражает глубокую благодарность своему научному руководителю чл.-корр. РАН, проф. Алексею Львовичу Семёнову за постановку задач, поддержку и постоянное внимание к работе, а также к. ф.-м. н. Андрею Альбертовичу Мучнику (1958 — 2007) за постановку задач, многочисленные обсуждения результатов настоящей работы.

Автор благодарен всем сотрудникам кафедры математической логики и теории алгоритмов за прекрасную доброжелательную атмосферу.

Публикации автора по теме диссертации

[1] Притыкин Ю. Л. Конечно-автоматные преобразования строго почти периодических последовательностей // *Математические заметки*, 80(5):751–756, 2006.

[2] Притыкин Ю. Л. Конечно-автоматные преобразования почти периодических последовательностей и алгоритмическая неразрешимость // *Труды XXVIII Конференции молодых учёных*, сс. 177–181, мех.-мат. ф-т МГУ им. М. В. Ломоносова, 2006.

[3] Притыкин Ю. Л. О нерегулярности некоторых множеств бесконечных слов // *Труды 30-й конференции молодых учёных Информационные Технологии и Системы (ИТиС 2007)*, сс. 134–137, Институт проблем передачи информации РАН, Москва, 2007.

[4] Pritykin Yu. On almost periodicity criteria for morphic sequences in some particular cases // *Proceedings of the 11th International Conference on Developments in Language Theory (DLT 2007), Turku, Finland, Lecture Notes in Computer Science*, vol. 4588, pp. 361–370, Springer-Verlag, 2007.

[5] Pritykin Yu. Information in infinite words // *Proceedings of the 6th International Conference on Words (Words 2007), Marseille, France*, pp. 254–261. Institute de Mathématiques de Luminy, Marseille, France, 2007.

[6] Pritykin Yu., Ulyashkina J. Aperiodicity measure for infinite sequences // *Proceedings of the 4th International Computer Science Symposium in Russia (CSR 2009), Novosibirsk, Russia*, Lecture Notes in Computer Science, vol. 5675, pp. 274–285. Springer-Verlag, 2009.

В этой работе Притыкину Ю. Л. принадлежит общая идея систематического изучения свойств понятия меры апериодичности, а также результаты теорем 1 — 5. Уляшкиной Ю. С. принадлежит результат теоремы 6.