Почти периодичность и конечные автоматы

Ю. Притыкин*

October 12, 2006

Аннотация

В работе рассмотрены различные понятия, обобщающие понятие периодической последовательности. Мы вводим класс существенно строго почти периодических последовательностей, таких что некоторый их суффикс строго почти периодичен (=равномерно рекуррентен). Класс почти периодических последовательностей включает в себя класс существенно строго почти периодических последовательностей, и мы доказываем, что это включение строгое. Мы доказываем, что класс существенно строго почти периодических последовательностей замкнут относительно преобразований, осуществляемых конечными автоматами и конечными преобразователями. Более того, мы также доказываем эффективный вариант этого утверждения. В заключение мы рассматриваем некоторые алгоритмические вопросы, связанные с почти периодичностью.

1 Введение

Строго почти периодические (=равномерно рекуррентные) последовательности изучались в работах Морса и Хедлунда [7, 8] и многих других авторов (например, см. [3, 9]). Последовательность называется строго почти периодической, если каждый её фактор встречается в ней бесконечное количество раз с ограниченными интервалами. Это понятие возникло в символической динамике, но впоследствии оказалось продуктивным в математической логике, теории алгоритмов, комбинаторике слов. Почти периодические последовательности были введены Семёновым в [18] при изучении логических теорий одноместных функций на натуральном ряде. Последовательность почти периодична, если каждый её фактор либо встречается бесконечное количество раз с ограниченными интервалами, либо встречается лишь конечное количество раз. Мы вводим новый класс последовательностей, называемых существенно строго почти периодические, у которых некоторый суффикс строго почти периодичен, и изучаем свойства этого класса.

Данная работа организована следующим образом.

В разделе 2 мы приводим различные варианты определений, обобщающих понятие периодичности. Класс почти периодических последовательностей включает в себя класс существенно строго почти периодических последовательностей. Мы доказываем, что это включение строгое (теорема 1).

Раздел 3 посвящен автоматным преобразованиям. Почти периодические последовательности подробно изучались в [19, 9]. В частности, авторы доказывают, что класс почти периодических последовательностей замкнут относительно конечно-автоматных преобразований.

*Московский Государственный Университет, e-mail: yura@mccme.ru, yuri.pritykin@gmail.com. Работа частично поддержана грантами РФФИ 06-01-00122, 05-01-02803, грантом Научных Школ 358.2003.1 и Колмогоровским грантом Института Новых Технологий.

Класс конечно-автоматных преобразований строго почти периодических последовательностей содержит класс существенно строго почти периодических последовательностей. Один из основных результатов статьи (теорема 3) заключается в том, что указанные классы совпадают. Другими словами, теорема 3 утверждает, что свойство существенной строгой почти периодичности сохраняется при конечно-автоматных преобразованиях. Более того, мы доказываем эффективный вариант этого утверждения (теорема 4) при помощи любопытной конструкции из теории конечных автоматов и комбинаторики слов. После этого мы рассматриваем обобщение понятия конечного автомата — конечный преобразователь, и для него доказываем аналогичное утверждение.

В разделе 4 мы рассматриваем различные алгоритмические вопросы, связанные с почти периодичностью. А именно, мы доказываем, что некоторые свойства, связанные с почти периодичностью, не имеют эффективных аналогов (в противоположность теореме 4 — эффективному аналогу теоремы 3). Например, мы показываем, что по существенно строго почти периодической последовательности и её регулятору невозможно найти префикс, который достаточно отрезать от последовательности, чтобы получить строго почти периодическую последовательность. В действительности, доказательства соответствующих фактов — это просто конкретные конструкции в комбинаторике слов.

Введём некоторые основные понятия и обозначения.

Стандартный алфавит из двух символов $\{0,1\}$ обозначим $\mathbb{B}$, множество натуральных чисел $\{0,1,2,\dots\}$ через $\mathbb{N}$. Пусть Σ — конечный алфавит, содержащий по крайней мере два символа. Будем рассматривать последовательности над этим алфавитом — отображения $\omega\colon \mathbb{N}\to\Sigma$. Множество всех таких последовательностей образует стандартное канторовское компактное метрическое пространство, которое мы обозначаем $\Sigma^{\mathbb{N}}$. Таким образом, $\lim_{n\to\infty} x_n = \omega$, если $\forall i\,\exists n\,\forall m>n\ x_m(i)=\omega(i)$ (это определение годится и для конечных слов x_n).

Через Σ^* обозначим множество всех слов над алфавитом Σ, включая пустое слово Λ. Если $i\leqslant j$ натуральные, через $[i,j]$ обозначим отрезок натурального ряда с концами в i и j, то есть множество $\{i,i+1,i+2,\dots,j\}$. Через $\omega[i,j]$ обозначим отрезок последовательности ω — слово $\omega(i)\omega(i+1)\dots\omega(j)$. Говорят, что $[i,j]$ — вхождение в последовательность ω слова $x\in\Sigma^*$, если $\omega[i,j]=x$. Слово $x\neq\Lambda$ называется фактором последовательности ω, если x входит в ω. Факторы вида $\omega[0,i]$ называются префиксами ω, последовательности вида $\omega(i)\omega(i+1)\omega(i+2)\dots$ — суффиксами ω и обозначаются $\omega[i,\infty)$. Через $|x|$ будем обозначать длину слова x. Будем говорить, что вхождение $x=\omega[i,j]$ в последовательность k-выровнено, если $k|i$. Мы представляем себе последовательность расположенной горизонтально и идущей слева направо до бесконечности, поэтому будем использовать выражения "левее" и "правее" для меньших и бо́льших индексов соответственно.

2 Почти периодичность

Последовательность ω периодическая, если для некоторого $T\in\mathbb{N}$ (называемого периодом) имеем $\omega(i)=\omega(i+T)$ для любого $i\in\mathbb{N}$. Множество всех периодических последовательностей обозначим за $\mathcal{P}$. Рассмотрим некоторые расширения этого класса.

Последовательность ω называется *почти периодической*, если для каждого её фактора x, входящего в неё бесконечное число раз, найдётся такое натуральное l, что на каждом отрезке длины l последовательности ω найдётся вхождение слова x. Класс всех таких последовательностей обозначим через $\mathcal{AP}$ (almost periodic).

Последовательность ω называется *строго почти периодической*, если для каждого её фактора x найдётся такое натуральное l, что на каждом отрезке длины l последовательности ω

найдётся вхождение слова x. Тем самым, любое слово, входящее в строго почти периодическую последовательность, входит в неё бесконечное количество раз. Через $\mathcal{SAP}$ (strongly almost periodic) будем обозначать класс всех таких последовательностей. Ясно, что для проверки строгой почти периодичности достаточно только убедиться в повторяемости с ограниченными интервалами всех префиксов, а не всех факторов.

Введём также ещё одно естественное определение. Будем называть последовательность ω *существенно строго почти периодической*, если некоторый её суффикс строго почти периодичен. Класс всех таких последовательностей обозначим $\mathcal{EAP}$ (eventually strongly almost periodic).

Если $\omega \in \mathcal{EAP}$, то минимальное такое n, что $\omega[n,\infty) \in \mathcal{SAP}$, обозначим $\mathrm{pr}(\omega)$. Заметим, что для любого $m \geqslant \mathrm{pr}(\omega)$ имеем $\omega[m,\infty) \in \mathcal{SAP}$.

Регулятором почти периодичности последовательности $\omega \in \mathcal{AP}$ назовём функцию $\mathrm{R}_\omega\colon \mathbb{N} \to \mathbb{N}$, которая на числе n равна такому l, что каждое слово длины n, которое входит в ω бесконечное количество раз, встретится на любом отрезке длины l последовательности ω, а также любое слово длины n, которое входит в ω конечное количество раз, не входит правее l (второе важно только для почти периодических последовательностей, не являющихся строго почти периодическими). Заметим, что регулятор определён неоднозначно: если какая-то функция больше регулятора, то она тоже является регулятором. Мы будем также использовать буквы $f, g, \ldots$ для обозначения регулятора, когда ясно, о какой последовательности идёт речь.

Несложно видеть, что $\mathcal{P} \subset \mathcal{SAP} \subset \mathcal{EAP} \subset \mathcal{AP}$. Оказывается, все эти включения строгие. Например, известная последовательность Туэ–Морса $\omega_T = 0110100110010110\ldots$ (см. [13, 1] или раздел 4) — пример последовательности из $\mathcal{SAP}$, но не из $\mathcal{P}$ (более того, класс $\mathcal{SAP}$ континуален, тогда как $\mathcal{P}$ счётный, доказательства см. в [5] или [9]). Неравенство $\mathcal{SAP} \subsetneq \mathcal{EAP}$ очевидно. Докажем $\mathcal{EAP} \subsetneq \mathcal{AP}$.

Теорема 1. *Существует двоичная последовательность $\omega \in \mathcal{AP} \setminus \mathcal{EAP}$.*

Доказательство. Построим цепочку двоичных слов: $a_0 = 1$, $a_1 = 10011$,
$a_2 = 1001101100011001001110011$, и так далее. Слово a_{n+1} получается из a_n по такому правилу:

$$a_{n+1} = a_n \overline{a}_n \overline{a}_n a_n a_n,$$

где $\overline{x}$ обозначает слово, полученное из слова x заменой всех символов 1 на 0 и 0 на 1. Введём обозначение

$$c_n = a_n a_n a_n a_n$$

и рассмотрим последовательность

$$\omega = c_0 c_1 c_2 c_3 \ldots$$

Докажем, что она и является искомой.

Длина a_n равна 5^n, поэтому длина начального отрезка $c_0 c_1 \ldots c_{n-1}$ последовательности ω равна $4(1 + 5 + \cdots + 5^{n-1}) = 5^n - 1$. Для удобства введём обозначение

$$l_n = 5^n - 1 = |c_0 c_1 \ldots c_{n-1}|.$$

Покажем, что ω почти периодична. Пусть $x \neq \Lambda$ встречается в ω бесконечное количество раз. Возьмём такое n, что $|x| < 5^n$. Пусть $[i,j]$ — вхождение слова x в ω, такое что $i \geqslant l_n$. Из построения последовательности следует, что для каждого k суффикс $\omega[l_k,\infty)$ можно рассматривать не только как конкатенацию символов 0 и 1, но и как конкатенацию слов a_k и $\overline{a}_k$. Поэтому в силу выбора i слово x является подсловом одного из четырёх слов $a_n a_n$, $a_n \overline{a}_n$, $\overline{a}_n a_n$, $\overline{a}_n \overline{a}_n$. Заметим, что слово 10011 содержит всевозможные слова длины два 00, 01, 10, 11, поэтому слово a_{n+1} содержит каждое из слов $a_n a_n$, $a_n \overline{a}_n$, $\overline{a}_n a_n$, $\overline{a}_n \overline{a}_n$. Отсюда получаем, что слово x является подсловом слова a_{n+1}. Аналогично, x является подсловом слова $\overline{a}_{n+1}$. На каждом отрезке длины $2|a_{n+1}|$, расположенном правее позиции l_{n+1}, найдётся вхождение слова a_{n+1} или $\overline{a}_{n+1}$, поэтому для $l = (5^{n+1} - 1) + 2 \cdot 5^{n+1}$ выполнено, что на каждом отрезке длины l найдётся вхождение слова x.

Докажем теперь для каждого натурального $n > 0$, что слово c_n не встречается в последовательности ω правее позиции l_{n+1}. Тем самым мы покажем, что для суффикса $\omega[l_n, \infty)$ найдётся слово, а именно, c_n, которое входит в неё ненулевое конечное количество раз, то есть такая последовательность не является строго почти периодической. Таким образом, будет доказано, что ω не является существенно строго почти периодической.

Пусть $\nu = \omega[l_{n+1}, \infty)$. Как уже было замечено, для каждого k, $1 \leqslant k \leqslant n+1$, её можно представить как конкатенацию слов a_k и $\overline{a}_k$. Предположим, c_n входит в ν, и пусть $[i, j]$ — одно из таких вхождений. При $n > 0$ слово c_n начинается с a_1, поэтому $[i, i+4]$ является вхождением слова a_1 в ν. Заметим, что если $a_1 = 10011$ входит в $a_1 a_1 = 1001110011$, $a_1 \overline{a}_1 = 1001101100$, $\overline{a}_1 a_1 = 0110010011$ или $\overline{a}_1 \overline{a}_1 = 0110001100$, то только с нулевой или с пятой позиции. Поэтому $[i, j]$ 5-выровнено, то есть можно считать c_n и ν составленными из слов a_1 и $\overline{a}_1$, и c_n входит в ν. Индукцией по m несложно доказать, что при $1 \leqslant m \leqslant n$ вхождение $[i, j]$ 5^m-выровнено, то есть если представить ν и c_n в виде конкатенации слов a_m и $\overline{a}_m$, то c_n входит в ν. База для $m = 1$ уже доказана. Зная это утверждение для $m = k$, мы можем представить ν и c_n составленными из "букв" a_k и $\overline{a}_k$, причём c_n входит в ν. Тогда, чтобы доказать утверждение для $m = k+1$, можно применить точно такое же рассуждение, как и в случае $m = 1$, только заменить 1 и 0 на a_k и $\overline{a}_k$, и a_1 и $\overline{a}_1$ на a_{k+1} и $\overline{a}_{k+1}$, и при этом использовать, что c_n начинается с a_{k+1}.

Таким образом, мы показали, что $[i, j]$ 5^n-выровнено, то есть если представлять себе ν и c_n составленными из "букв" a_n и $\overline{a}_n$, то мы доказали, что $c_n = a_n a_n a_n a_n$ входит в ν. Но заметим, что в любой последовательности, составленной из $a_1 = 10011$ и $\overline{a}_1 = 01100$, четырёх подряд идущих символов 0 или 1. Аналогично c_n не может входить в ν. Противоречие. $\square$

Более того, из доказательства видно, как построить континуальное множество почти периодических последовательностей, не являющихся существенно строго почти периодическими. Можно действовать, например, так: каждой последовательности $\tau \colon \mathbb{N} \to \{4, 5\}$ сопоставить последовательность, построенную так же, как и ω из доказательства теоремы 1, только в качестве c_n брать слово

$$c_n^{(\tau)} = \underbrace{a_n a_n \ldots a_n}_{\tau(n)}.$$

Ясно, что построенные таким образом ω_τ будут различны для различных τ. Осталось заметить, что всевозможных τ континуум.

3 Конечно-автоматные преобразования

Представляется интересным рассматривать преобразования последовательностей и пытаться понять, сохраняют ли эти преобразования свойства почти периодичности. Простейшим алгоритмическим преобразованием можно считать конечно-автоматное. Другая мотивация, несколько менее абстрактная, заключается в том, что конечно-автоматные преобразования были одним из основных средств в [19] при изучении почти периодичности и получении критерия разрешимости теорий первого порядка и монадических теорий одноместных функций на натуральном ряде. В [19, 9] было показано, что класс почти периодических последовательностей замкнут относительно конечно-автоматных преобразований. Ещё из подобных результатов можно привести, например, результат о том, что класс k-автоматных последовательностей замкнут относительно конечно-автоматных преобразований (см. [4], а также [2]). Мы доказываем замкнутость относительно этого вида преобразований класса существенно строго почти периодических последовательностей.

Конечным автоматом назовём совокупность $F = \langle \Sigma, \Delta, Q, \tilde{q}, f \rangle$, где Σ и Δ — конечные множества, называемые соответственно входной и выходной алфавит, Q — конечное множество состояний, $\tilde{q} \in Q$ — выделенное состояние, называемое начальным, и

$$f\colon Q \times \Sigma \to Q \times \Delta$$

— функция переходов. Для $\alpha \in \Sigma^{\mathbb{N}}$ рассмотрим последовательность $\langle p_n, \beta(n) \rangle_{n=0}^{\infty}$, где $p_n \in Q$, $\beta(n) \in \Delta$, и пусть $p_0 = \tilde{q}$ и для каждого n выполняется $\langle p_{n+1}, \beta(n) \rangle = f(p_n, \alpha(n))$. Последовательность $\beta = F(\alpha)$ назовём тогда результатом преобразования α автоматом F. Если $[i, j]$ — вхождение слова x в последовательность α, причём $p_i = q$, то будем говорить, что автомат F подходит к этому вхождению слова x в состоянии q.

В [19, 9] была доказана следующая теорема.

Теорема 2. *Пусть F — конечный автомат. Тогда если $\omega \in \mathcal{AP}$, то $F(\omega) \in \mathcal{AP}$.*

Мы можем доказать аналог этой теоремы для существенно строго почти периодических последовательностей.

Теорема 3. *Пусть F — конечный автомат. Тогда если $\omega \in \mathcal{EAP}$, то $F(\omega) \in \mathcal{EAP}$.*

Доказательство. Очевидно, что достаточно доказать утверждение для $\omega \in \mathcal{SAP}$, так как существенно строго почти периодическая последовательность (в нашем случае $F(\omega)$) остаётся существенно строго почти периодической при приписывании к ней произвольного префикса.

Итак, пусть $\omega \in \mathcal{SAP}$. По теореме 2 последовательность $F(\omega)$ почти периодическая. Предположим, что она не является существенно строго почти периодической. Это означает, что для любого натурального N можно найти такое слово, которое входит в $F(\omega)$ правее позиции N, и после этого уже не входит. Действительно, отрезав от $F(\omega)$ начальный отрезок $[0, N]$, мы не получим строго почти периодическую последовательность, значит, найдётся слово, входящее в неё ненулевое конечное количество раз. Тогда возьмём самое правое вхождение этого слова.

Пусть $[i_0, j_0]$ — самое правое вхождение некоторого слова y_0 в последовательность $F(\omega)$. Для некоторого натурального l_0 слово $x_0 = \omega[i_0, j_0]$ входит в любой отрезок длины l_0 последовательности ω (в силу строгой почти периодичности). При этом если q_0 — состояние, в котором автомат F подошёл к позиции i_0, то ко всем дальнейшим вхождениям слова x_0 в ω автомат не может подойти в состоянии q_0, иначе бы он выдал полностью слово y_0.

Пусть теперь $[r, s]$ — самое правое вхождение некоторого слова a в последовательность $F(\omega)$, причём $r > i_0 + l_0$. На отрезке $\omega[r - l_0, r]$ найдётся вхождение $[r', s']$ слова x_0, причём в силу выбора r будет выполнено $r' > i_0$. Положим тогда

$$i_1 = r',\ j_1 = s,\ x_1 = \omega[i_1, j_1],\ y_1 = F(\omega)[i_1, j_1].$$

Поскольку слово a не входит в $F(\omega)$ правее позиции r, то и слово y_1, которое содержит a в качестве подслова, не входит в $F(\omega)$ правее позиции i_1. Это значит, что если q_1 — состояние, в котором автомат подошёл к позиции i_1, то больше никогда в состоянии q_1 к вхождению слова x_1 в последовательность ω автомат не подойдёт. Поскольку x_1 начинается со слова $\omega[r', s'] = x_0$, и $r' > i_0$, то $q_1 \neq q_0$. Таким образом, мы нашли слово x_1, ко всем вхождениям которого, расположенным правее i_1, автомат не может подойти, находясь в состояниях q_0 или q_1.

Пусть $m = |Q|$ — количество возможных состояний автомата. Продолжая так дальше рассуждать по индукции, мы построим цепочку слов $x_k = \omega[i_k, j_k]$ и соответствующих различных состояний q_k, где $k < m$, так что ко всем вхождениям x_k в ω правее i_k автомат не может подойти в состояниях $q_0, q_1, \ldots, q_k$. При $k = m$ получаем противоречие. □

Отметим, что приведённое доказательство не эффективно в следующем смысле. Допустим, мы знаем $\omega \in \mathcal{SAP}$ и её регулятор R_ω. Тогда по теореме 3 существует оценка сверху на $\mathrm{pr}(F(\omega))$, так как $F(\omega) \in \mathcal{EAP}$, но имеющееся доказательство не позволяет по данным объектам найти никакую такую оценку эффективно (и даже до некоторого момента не позволяло надеяться, что такой эффективный способ вообще существует).

Теорема 3 была доказана в [15]. Следующий эффективный вариант этой теоремы был объявлен в [16] и доказан в [11].

Для произвольной функции g обозначим $\underbrace{g \circ g \circ \cdots \circ g}_{n}$ через g^n.

Теорема 4. *Пусть F — конечный автомат с n состояниями, $\omega \in \mathcal{SAP}$. Тогда*

$$\mathrm{pr}(F(\omega)) \leqslant \mathrm{R}_\omega^n(1) + \mathrm{R}_\omega^{n-1}(1) + \cdots + \mathrm{R}_\omega(1).$$

Чтобы доказать эту теорему, мы сначала рассмотрим специальный класс автоматов, называемых обратимыми, для которых утверждение теоремы несложно. Потом при помощи некоторой конструкции из комбинаторики слов мы сведём общий случай к случаю обратимого автомата.

Назовём конечный автомат $F = \langle \Sigma, \Delta, Q, \tilde{q}, f \rangle$ *обратимым*, если для каждого его состояния q и каждого символа $a \in \Sigma$ существует ровно одно состояние q' и ровно один символ $b \in \Delta$, для которых $f(q', a) = \langle q, b \rangle$. Другими словами, в таком автомате каждая буква входного алфавита осуществляет взаимно однозначное отображение множества состояний в себя (выходные символы не имеют значения). Находясь в некотором состоянии и зная последовательность предыдущих входных символов, можно восстановить и последовательность пройденных состояний (в этом и заключается свойство обратимости).

Лемма 5. *Пусть F — обратимый конечный автомат. Тогда если $\omega \in \mathcal{SAP}$, то $F(\omega) \in \mathcal{SAP}$.*

Доказательство. Предположим, x входит в ω, и F подходит к этому вхождению в состоянии q. Мы хотим доказать, что следующий раз, когда F подойдёт к x в ω, находясь в состоянии q, произойдёт на некотором ограниченном расстоянии от текущей такой ситуации, причём мы можем оценить это расстояние в терминах $|x|$ и R_ω. Это означает, что такая оценка годится для любой ситуации в ω, когда F подходит к x в состоянии q. Таким образом, этого достаточно для доказательства теоремы.

Пусть $x_0 = \omega[r, s]$, и автомат подходит к этому вхождению в состоянии q. Пусть $[i_0, j_0]$ — следующее вхождение x_0 в ω, так что $j_0 \leqslant r + \mathrm{R}_\omega(|x|) + 1$. Если F подходит к этому вхождению в состоянии q, мы нашли, что искали. В противном случае F подходит к i_0 в состоянии $q_0 \neq q$. Пусть $x_1 = \omega[r, j_0]$, и пусть $[i_1, j_1]$ — следующее вхождение x_1 в ω, так что $j_1 \leqslant r + \mathrm{R}_\omega(\mathrm{R}_\omega(|x|) + 1) + 1$. Пусть F подходит к $i_1 + i_0$ в состоянии q_1. Если $q_1 = q$, то $\omega[i_1 + i_0, j_1] = x_0$. Если

$q_1 = q_0$, то F подходит к i_1 в состоянии q в силу обратимости автомата, и мы опять нашли, что искали. Иначе $q_1 \neq q$ и $q_1 \neq q_0$.

Аналогично, для $x_2 = \omega[r, j_1]$ и его вхождения $[i_2, j_2]$ в ω, такого что $i_2 > r$ и $j_2 \leqslant r + \mathrm{R}_\omega(\mathrm{R}_\omega(\mathrm{R}_\omega(|x|) + 1) + 1) + 1$, в худшем случае мы получаем, что F подходит к $i_2 + i_1 + i_0$ в состоянии q_2, где $q_2 \neq q$, $q_2 \neq q_0$ и $q_2 \neq q_1$. Продолжая так дальше рассуждать, для $k < m = |Q|$ мы строим слова $x_0, x_1, \ldots, x_k$, их вхождения $[i_0, j_0], [i_1, j_1], \ldots, [i_k, j_k]$ и различные состояния $q_0, q_1, \ldots, q_{k-1}$, такие что в худшем случае F не может подойти к $i_k + i_{k-1} + \cdots + i_0$ в состояниях $q, q_0, \ldots, q_{k-1}$. Таким образом, для $k = m$ мы заведомо найдём искомую вхождение x_0 в состоянии q, и оценка на расстояние равна $f(f(\ldots(|x|)\ldots))$, где $f = \mathrm{R}_\omega + 1$, а количество итераций в формуле равно m. □

Для $\omega \in \Sigma^{\mathbb{N}}$, $\nu \in \Delta^{\mathbb{N}}$ определим $\omega \times \nu \in (\Sigma \times \Delta)^{\mathbb{N}}$, так что $(\omega \times \nu)(i) = \langle \omega(i), \nu(i) \rangle$.

Следствие 6. *Если $\omega \in \mathcal{SAP}$ и $\nu \in \mathcal{P}$, то $\omega \times \nu \in \mathcal{SAP}$.*

Доказательство. Операция “$\times$” с периодической последовательностью может быть осуществлена циклическим автоматом, который, очевидно, обратим. □

Замечание 7. Сформулируем некоторые открытые вопросы, связанные со следствием 6. Пусть $\omega, \nu \in \mathcal{SAP}$. Интересно выяснить тогда, что можно сказать про $\omega \times \nu$. Несложно построить пример, в котором $\omega \times \nu \notin \mathcal{AP}$ или $\omega \times \nu \in \mathcal{EAP} \setminus \mathcal{SAP}$. Можно ли построить пример, в котором $\omega \times \nu \in \mathcal{AP} \setminus \mathcal{EAP}$? Существует ли какой-то критерий для определения, верно ли, что $\omega \times \nu \in \mathcal{AP}$? Пример $\omega, \nu \in \mathcal{AP}$, для которого $\omega \times \nu \notin \mathcal{AP}$, можно найти в [9].

Теперь рассмотрим следующую конструкцию. Пусть $\omega \in \Sigma^{\mathbb{N}}$, и символ $a \in \Sigma$ входит в ω бесконечное количество раз. Разрежем ω на блоки вида xa, где $x \in (\Sigma \backslash \{a\})^*$, то есть на слова, содержащие ровно один символ a на последнем месте. Чтобы получить такое разбиение, надо провести разрез после каждого символа a. Если символ a встречается в ω с ограниченными интервалами, то количество всевозможных таких блоков конечно (например, если $\omega \in \mathcal{AP}$, то их длины не больше $\mathrm{R}_\omega(1)$). Закодируем однозначно эти блоки буквами некоторого конечного алфавита, обозначим этот алфавит $\mathrm{b}_{a,\omega}(\Sigma)$. Таким образом, мы из ω получили новую последовательность в этом алфавите. Последовательность, полученную из неё отрезанием первого символа, назовём *a-разбиением* последовательности ω и обозначим $\mathrm{s}_a(\omega)$. Например, 0-разбиение последовательности 3200122403100110... — это (0)(12240)(310)(0)(110)...

Лемма 8. *Пусть $\omega \in \mathcal{SAP}$, и символ a входит в ω. Тогда $\mathrm{s}_a(\omega) \in \mathcal{SAP}$.*

Доказательство. Пусть k — максимальная длина блоков разбиения. Рассмотрим слово x — произвольный префикс последовательности $\mathrm{s}_a(\omega)$. Ему соответствует слово y длины не более $k|x|$ в ω, из которого x получается операцией разбиения. Пусть $z = ay$, $|z| \leqslant k|x| + 1$. Слово z входит в ω (задевая первым символом последний символ первого блока разбиения, отбрасываемого при переходе к $\mathrm{s}_a(\omega)$). Тогда на любом отрезке длины $l = \mathrm{R}_\omega(k|x| + 1)$ последовательности ω найдётся вхождение слова z. Поскольку z начинается и заканчивается символом a, любое такое вхождение “правильно” выровнено, то есть границы разбиения этого вхождения на блоки совпадут с границами разбиения последовательности ω. Поэтому каждому вхождению z в ω соответствует вхождение x в $\mathrm{s}_a(\omega)$. Таким образом, на каждом отрезке длины $\mathrm{R}_\omega(k|x| + 1)$ в $\mathrm{s}_a(\omega)$ найдётся вхождение x. □

Замечание 9. В связи с леммой 8 возникает интересный вопрос, частный случай сформулированного в замечании 7. Вместо разбиения последовательности $\omega \in \mathcal{SAP}$ на блоки, оканчивающиеся фиксированным символом, можно рассматривать произвольные разбиения на блоки различной длины. Когда такое разбиение будет строго почти периодическим или хотя бы почти периодическим?

Теперь мы можем доказать обещанную теорему 4.

Доказательство теоремы 4. Пусть $F = \langle \Sigma, \Delta, Q, \tilde{q}, f\rangle$ и $|Q| = n$. Мы построим алгоритм вычисления некоторого

$$l \geqslant \mathrm{pr}(F(\omega)),$$

и одновременно с этим докажем оценку

$$l \leqslant \mathrm{R}^n_\omega(1) + \mathrm{R}^{n-1}_\omega(1) + \cdots + \mathrm{R}_\omega(1).$$

Будем считать, что у всех автоматов, встречающихся в доказательстве, выходной алфавит является максимально возможным — "входной алфавит"$\times$"множество состояний" (произвольный случай легко получается из такого проекцией). Например, для F это $\Sigma \times Q$. Соответствующим образом изменяется и функция переходов f — в выходную последовательность она пишет максимальную информацию — пару из входного символа и текущего состояния. Далее мы всюду будем опускать вторую компоненту значения функции переходов (то есть, например, вместо $f(p, a) = \langle q, b\rangle$ мы будем писать теперь просто $f(p, a) = q$, имея в виду $f(p, a) = \langle q, \langle p, a\rangle\rangle$).

Положим $\omega_0 = \omega$. Можно считать, что каждый символ из Σ входит в ω_0, иначе ограничим F только на символы, встречающиеся в ω_0; какие это будут символы, можно выяснить эффективно, просмотрев начальный отрезок ω_0 длины $\mathrm{R}_{\omega_0}(1)$.

Если F — обратимый, то по лемме 5 получаем, что $\mathrm{pr}(F(\omega_0)) = 0$. В противном случае найдётся такой символ $a_1 \in \Sigma$, который на множестве состояний действует не взаимно однозначно, то есть множество

$$Q_1 = \{q \ : \ \exists q' \ \ f(q', a_0) = q\}$$

строго вложено в Q. Рассмотрим разбиение

$$\omega_1 = \mathrm{s}_{a_0}(\omega_0),$$

которое по лемме 8 строго почти периодично. Заметим, что в каком бы состоянии автомат не запускался на ω_0, к каждому блоку разбиения $\mathrm{s}_{a_0}(\omega_0)$ он подходит в состоянии из множества Q_1, поскольку каждый такой блок заканчивается символом a_0.

Построим новый автомат F_1 (построение эффективно по F). Положим входной алфавит автомата F_1 равным $\mathrm{b}_{a_0,\omega_0}(\Sigma)$, множество состояний равным Q_1, и значение функции переходов на паре $x \in \mathrm{b}_{a_0,\omega_0}(\Sigma)$, $q \in Q_1$ положим равным выходу автомата F, запущенного на слове x, записанном в алфавите Σ, и состоянии q. Начальным состоянием автомата F_1 объявим то состояние, в котором F оказывается после работы на начале ω до первого вхождения a_0 (на том куске, который мы отбрасываем при переходе к $\mathrm{s}_{a_0}(\omega_0)$). Теперь действие автомата F на ω_0 моделируется действием автомата F_1 на ω_1. При этом ω_1 получается из ω отрезанием не более чем $\mathrm{R}_{\omega_0}(1)$ символов, если считать их в алфавите Σ.

Мы имеем последовательность ω_1 (в большем, чем исходный, алфавите), на которой действует автомат F_1 с меньшим, чем в исходном автомате F, количеством состояний. Если F_1 не обратимый, то повторив процедуру, описанную в предыдущем абзаце, ещё один раз, мы получим последовательность ω_2 в каком-то алфавите $\mathrm{b}_{a_1,\omega_1}(\mathrm{b}_{a_0,\omega_0}(\Sigma))$, на которой действует автомат с ещё меньшим количеством состояний. Она получена из ω_1 отрезанием не более чем $\mathrm{R}_{\omega_1}(1)$ символов, считая в символах алфавита $\flat_{a_0,\omega_0}(\Sigma)$. Значит, ω_2, записанная в исходном алфавите Σ, получается из ω отрезанием начального отрезка длины не более $\mathrm{R}_{\omega_0}(\mathrm{R}_{\omega_0}(1)) + \mathrm{R}_{\omega_0}(1)$, считая в символах алфавита Σ.

Автомат с одним состоянием и с произвольным количеством входных символов всегда обратимый (и даже тождественный). Поэтому повторив описанную процедуру некоторое количество раз k ($k \leqslant n$, так как при каждом запуске процедуры количество состояний автомата уменьшается), мы получим в конце концов ситуацию, в которой возвратный автомат F_k действует на строго почти периодической последовательности ω_k в некотором алфавите. Символы в этом алфавите получены из блоков исходной последовательности. Тогда $F_k(\omega_k) \in \mathcal{SAP}$.

Записав ω_k в исходном алфавите Σ, получим некоторый суффикс ω' исходной последовательности ω, полученный из ω отрезанием начала длины не более

$$\mathrm{R}_\omega^k(1) + \mathrm{R}_\omega^{k-1}(1) + \cdots + \mathrm{R}_\omega(1) \leqslant \mathrm{R}_\omega^n(1) + \mathrm{R}_\omega^{n-1}(1) + \cdots + \mathrm{R}_\omega(1).$$

Осталось понять, почему из $F_k(\omega_k) \in \mathcal{SAP}$ следует, что $F(\omega') \in \mathcal{SAP}$. Объясним это в простом частном случае, когда полученный уже на первом шаге автомат F_1 является обратимым (общая ситуация сводится к этому случаю по индукции). Тогда ω' получается из ω отбрасыванием всех символов до первого вхождения символа a_0 включительно. Пусть начальное состояние F_1 (то, в котором F подходит к ω') равно q. Чтобы доказать $F(\omega') \in \mathcal{SAP}$, необходимо и достаточно убедится в том, что для каждого префикса ω' вхождения его копий в ω', к которым F подходит в состоянии q, достаточно регулярны — встречаются на каждом отрезке длины l для некоторого l (достаточность очевидна, а необходимость следует из нашего условия о том, что автомат всегда выдаёт пару $\langle$входной символ, текущее состояние$\rangle$).

Пусть x — произвольный префикс ω', который заканчивается символом a_0, и тем самым его можно правильно разбить на блоки (произвольный префикс обязательно содержится в каком-то префиксе такого вида). Обозначим это разбиение через y. Поскольку F_1 возвратный, выполнено $F_1(\omega_1) \in \mathcal{SAP}$. Поэтому в силу вышеуказанного необходимого и достаточного условия автомат F_1 регулярно — на каждом отрезке длины t для некоторого t — подходит к y в состоянии q. Каждая такая ситуация соответствует в ω' подходу F к некоторому вхождению x в состоянии q, и это происходит на каждом отрезке длины tk, где $k \leqslant \mathrm{R}_\omega(1)$ — максимальная длина блока. $\square$

Доказательство теоремы 2, предложенное в [9], даёт оценку сверху примерно R_ω^{2n} для регулятора выходной последовательности $F(\omega)$ при действии автомата F с n состояниями на последовательности ω. Однако если обобщить рассуждения при доказательстве теоремы 4 на случай почти периодических последовательностей (рассматривая в них слова, входящие бесконечное количество раз), то можно улучшить приведённую оценку примерно до R_ω^n. Как показано в [17], эту оценку нельзя улучшить существенно: можно построить примеры, в которых регулятор выходной последовательности есть как минимум αn раз проитерированный регулятор исходной последовательности (для некоторого $\alpha < 1$, правда, предложенное значение α довольно далеко от 1). После незначительных модификаций конструкции в [17] можно показать, что верхняя оценка на $\mathrm{pr}(F(\omega))$ в теореме 4 также существенно неулучшаема (в вышеприведённом смысле).

Замечательно, что, таким образом, мы имеем два различных доказательства теоремы 3, связь между которыми неясна. Поскольку приведённые рассуждения можно приспособить и для доказательства теоремы 2, то с ней ситуация аналогична. Рассуждение в доказательстве теоремы 4 использует любопытную конструкцию в теории конечных автоматов. Возможно, она поможет в решении каких-либо других вопросов, связанных с конечными автоматами.

Интересно понять, что произойдёт, если мы несколько расширим класс рассматриваемых преобразований.

По-видимому, простейшим обобщением можно считать стековый автомат. Его неформально можно описать как конечный автомат, к которому добавлен стек. На каждом шаге следующее состояние автомата и выходной символ определяются текущим состоянием, символом входной последовательности и символом в голове стека. Кроме того, на каждом шаге автомат должен решить, что делать со стеком: добавить символ, удалить символ или оставить стек без изменений. Параллель здесь такая: рассматриваемым нами конечным автоматам соответствуют обычные принимающие автоматы без выхода, задающие регулярные языки, а рассматриваемым нами стековым автоматам (которые мы так называем для простоты, а на самом деле это преобразователи) соответствуют принимающие стековые автоматы, распознающие контекстно-свободные языки. Подробнее см., например, в [12].

Оказывается, стековые автоматы не сохраняют свойство почти периодичности. Действительно, рассмотрим строго почти периодическую последовательность из 0 и 1, обладающую следующим свойством: для любого $n \in \mathbb{N}$ найдётся такой её начальный отрезок, в котором нулей на n больше, чем единиц, а также найдётся такой начальный отрезок, в котором единиц на n больше, чем нулей (такую последовательность легко построить с помощью блочного произведения, см. [5, 6]). Применим к этой последовательности следующий стековый автомат. У него два режима, 0 (начальный) и 1. В режиме 0 он действует так: увидев в последовательности 0, он кладёт его в стек, а увидев 1, убирает 0 из стека. Когда стек опустошается (что соответствует начальному отрезку последовательности с равным количеством 0 и 1), автомат переключается в режим 1. В режиме 1 автомат действует противоположным образом: при входе 1 кладёт 1 в стек, при входе 0 убирает 1 из стека, при опустошении стека переключается в режим 0. В выходную последовательность автомат подаёт всегда номер своего режима. Видно тогда, что в выходной последовательности найдутся сколь угодно длинные отрезки подряд идущих 0, и значит, она не является почти периодической.

Результаты о конечно-автоматных преобразованиях могут быть продолжены на более широкий класс конечных преобразователей (finite transducers).

Пусть Σ, Δ — конечные алфавиты. Отображение $h\colon \Sigma^* \to \Delta^*$ называется *гомоморфизмом*, если для любых $u, v \in \Sigma^*$ выполнено $h(uv) = h(u)h(v)$. Ясно, что гомоморфизм полностью определяется своими значениями на однобуквенных словах. Если ω — последовательность букв алфавита Σ, по определению положим

$$h(\omega) = h(\omega(0))h(\omega(1))h(\omega(2))\ldots$$

Пусть $h\colon \Sigma^* \to \Delta^*$ — гомоморфизм, ω почти периодическая последовательность над алфавитом Σ. В [9] было показано, что если последовательность $h(\omega)$ бесконечна, то она почти периодична. Ясно тогда, что для ω — строго почти периодической $h(\omega)$ будет строго почти периодической (если бесконечна). Действительно, достаточно показать, что для любого слова $u = \omega[i,j]$ слово $h(u) = h(\omega(i))\ldots h(\omega(j))$ встречается в $h(\omega)$ бесконечно много раз. Но это следует из определения $h(\omega)$ и из того, что ω строго почти периодична, и значит, слово u встречается в ней бесконечно много раз. Очевидно, для ω — существенно строго почти периодической $h(\omega)$ будет снова существенно строго почти периодической (если она бесконечна).

Естественным обобщением понятия конечного автомата является понятие *конечного преобразователя* (подробнее см. [2, 9, 12, 14]). Отличие состоит в том, что конечный преобразователь может выдавать слово произвольной длины, прочитав один символ входа. Формально, меняется только определение функции переходов: теперь она имеет вид

$$f\colon Q \times \Sigma \to Q \times \Delta^*.$$

Если последовательность $\langle p_n, v_n \rangle_{n=0}^{\infty}$, где $p_n \in Q$, $v_n \in \Delta^*$, является преобразованием последовательности α, то результатом преобразования назовём последовательность $v_0 v_1 v_2 \ldots$

Действие конечного преобразователя можно представить в виде композиции автоматного преобразования и гомоморфизма. Каждый из этих типов преобразований, как мы знаем, сохраняет свойство почти периодичности, поэтому получаем следствие: почти периодические последовательности под действием конечных преобразователей переходят в почти периодические. Аналогично, из теоремы 3 и сохранения существенной строгой почти периодичности при гомоморфных отображениях получаем следствие:

Следствие 10. *Пусть F — конечный преобразователь, $\omega \in \mathcal{EAP}$, и, допустим, $F(\omega)$ бесконечна. Тогда $F(\omega) \in \mathcal{EAP}$.*

4 Эффективность

Во многих вопросах, связанных с почти периодическими последовательностями, естественным образом возникает алгоритмическая составляющая: можно ли ту или иную характеристику или то или иное свойство проверять алгоритмически, получая на вход последовательность. Иногда эти вопросы являются эффективными аналогами уже известных результатов, например, теорема 4 — эффективный аналог теоремы 3. В основном нас будет интересовать ситуация, когда на эти вопросы ответ отрицательный. Мы доказываем, что некоторые свойства почти периодичности не имеют эффективных аналогов.

Формально, мы рассматриваем алгоритм A, которому на вход подаётся оракул для последовательности ω. Алгоритм должен завершать работу на любом оракуле и выдавать ответ в виде конечного бинарного слова или любого другого конструктивного объекта. Основное свойство такого алгоритма — непрерывность: перед тем, как выдать ответ, он может прочитать лишь конечное количество элементов последовательности. Таким образом, для доказательства алгоритмической неразрешимости некоторого свойства таким алгоритмом достаточно показать, что это свойство не является непрерывным.

Почти никакие осмысленные свойства нельзя алгоритмически распознать, имея на входе только саму последовательность. Например, про $\omega \in \mathbb{B}^{\mathbb{N}}$ нельзя даже сказать за конечное время, входит ли в неё символ 1: если алгоритм проверил некоторое конечное количество символов, и все они оказались 0, он не может гарантировать, что далее не встретится 1. Вопрос об эффективности станет более интересным, если мы на вход будем подавать какую-то дополнительную информацию. В случае с почти периодическими последовательностями в качестве этой информации естественно взять регулятор почти периодичности.

Ясно, что функции $\mathbb{N} \to \mathbb{N}$, а также пары $\langle$последовательность, функция$\rangle$ можно однозначным образом кодировать отдельными бинарными последовательностями. Поэтому мы можем корректно рассматривать алгоритмы, получающие на вход почти периодическую последовательность ω и её регулятор f.

Теперь мы видим, что при таком подходе сформулированная выше задача может быть решена эффективно: прочитав первые $f(1)$ символов последовательности, мы можем понять, входит ли в неё 1, а прочитав следующие $f(1)$ символов, сможем сказать даже, входит ли в неё 1 бесконечно много раз.

Следующие несколько теорем — примеры результатов о неэффективности. Особенно интересно, что результаты теоремы 11 прямо противоположны результатам теоремы 4. Теорема 14 также связана с теоремой 4. Все дальнейшие теоремы были впервые объявлены в [16].

Под сходимостью $f_n \to f$, где $f_n, f\colon \mathbb{N} \to \mathbb{N}$, будем понимать условие $\forall i\ \exists n\ \forall m > n\ f_m(i) = f(i)$.

Теорема 11. *По $\omega \in \mathcal{EAP}$ и её регулятору f невозможно алгоритмически определять какое-либо $l \geqslant \mathrm{pr}(\omega)$.*

Напомним, что ω_T — известная последовательность Туэ–Морса — может быть получена следующим образом: положим $a_0 = 0$, $a_{n+1} = a_n\overline{a}_n$, и далее $\omega_T = \lim_{n\to\infty} a_n$. Заметим, что $|a_n| = 2^n$. Последовательность Туэ–Морса обладает огромным количеством интересных свойств (см. [1]), но сейчас нам особенно интересно следующее: ω_T бескубная, то есть для любого $a \in \mathbb{B}^*$, $a \neq \Lambda$ слово aaa не входит в ω_T (см. [1, 13]).

Доказательство теоремы 11. Для доказательства достаточно построить $\omega_n \in \mathcal{EAP}$, $\omega \in \mathcal{SAP}$ с регуляторами f_n и f, такие что $\omega_n \to \omega$, $f_n \to f$, но $\mathrm{pr}(\omega_n) \to \infty$. Действительно, предположим, указанный в условии теоремы алгоритм существует. Пусть, получив на вход $\langle \omega, f\rangle$, он выдаёт число $l \geqslant 0$ (оно может быть произвольным, так как $\omega \in \mathcal{SAP}$). Во время вычисления l алгоритм прочитал лишь конечное количество символов ω и значений f, поэтому существует такое $N > l$, что он не знает $\omega(k)$ и $f(k)$ для $k > N$ (формально это означает, что на всех входах, различие с которыми только в позициях $k > N$, алгоритм работает так же, как и на $\langle \omega, f\rangle$). Поскольку $\mathrm{pr}(\omega_n) \to \infty$, найдётся n, для которого $\mathrm{pr}(\omega_n) > N$. Ясно тогда, что на входе $\langle \omega_n, f_n\rangle$ алгоритм будет работать так же, как и на $\langle \omega, f\rangle$, и значит, выдаст l, но $\mathrm{pr}(\omega_n) > N > l$.

Положим $\omega = \omega_T$, $\omega_n = a_na_na_n\omega$. Докажем, что $\mathrm{pr}(\omega_n) \geqslant 2^n$. Действительно, если $\mathrm{pr}(\omega_n) < 2^n$, то $a_na_n\omega = a_na_na_n\bar{a}_n\bar{a}_na_n\ldots \in \mathcal{SAP}$, и значит, $a_na_na_n$ входит в ω_T — противоречие с бескубностью ω_T.

Осталось показать, что можно выбрать регуляторы f_n, f, так что $f_n \to f$. Для этого достаточно найти общий регулятор g для всех ω_n (потом мы можем увеличить g и добиться того, чтобы g была регулятором для всех ω_n и для ω). Зафиксируем некоторый регулятор R_ω и положим $g = 4\cdot \mathrm{R}_\omega$. Пусть слово v, $|v| = k$ входит в $\omega_n = a_na_na_n\omega$ бесконечно много раз. Возьмём отрезок $[i,j]$ последовательности ω длины $4\cdot \mathrm{R}_\omega(k)$ и покажем, что v в него входит. Если $j \geqslant 3\cdot 2^n + \mathrm{R}_\omega(k)$, то v встречается на отрезке $\omega[3\cdot 2^n, 3\cdot 2^n + \mathrm{R}_\omega(k)]$ (по определению регулятора). Иначе $j < 3\cdot 2^n + \mathrm{R}_\omega(k)$, откуда $i \leqslant 3\cdot 2^n - 3\,\mathrm{R}_\omega(k)$. Но $i \geqslant 0$, значит, в этом случае $\mathrm{R}_\omega(k) \leqslant 2^n = |a_n|$. Значит, $\omega_n[i, i+\mathrm{R}_\omega(k)]$ целиком содержится в a_na_n. Но a_na_n входит в ω, и значит, $\omega_n[i, i+\mathrm{R}_\omega(k)]$ тоже, поэтому v в него входит.

Однако g — ещё не искомая. Необходимо проследить за словами, которые встречаются в ω_n конечное количество раз. Но ясно, что если какое-то v входит в ω_n конечное количество раз, то $|v| = k > 2^n$ (иначе v входит в блок из двух последовательных слов a_n или $\bar{a}_n$, а значит, и в ω), то есть такое может быть лишь для конечного множества различных n. Поэтому, рассмотрев все ситуации, когда слова фиксированной длины k входят в какие-то ω_n конечное количество раз, нам, возможно, придётся увеличить значение $g(k)$, но лишь конечное количество раз. Значит, искомая оценка регуляторов найдётся. □

Мы уже видели, что $\mathcal{EAP} \subsetneq \mathcal{AP}$ (теорема 1). Используя ту же конструкцию, покажем, что эти классы невозможно разделять эффективно.

Теорема 12. *По $\omega \in \mathcal{AP}$ и её регулятору $f \geqslant \mathrm{R}_\omega$ невозможно алгоритмически определять, верно ли, что $\omega \in \mathcal{EAP}$.*

В [9] был предложен следующий универсальный метод построения строго почти периодических последовательностей. Он основан на алгебре блоков, введённой в [6] и впоследствии рассматривавшейся также в [5]. Этот универсальный метод является частным случаем предложенной в [19] общей конструкции разложения произвольной последовательности в блочное произведение в некотором обобщённом смысле.

Последовательность $\langle A_n, l_n \rangle$, где $A_n \subset \Sigma^*$ для некоторого конечного алфавита Σ, $l_n \in \mathbb{N}$, называется *строгой Σ-схемой*, если выполнены следующие условия:

(1) все слова из A_n имеют длину l_n;

(2) любое слово $u \in A_{n+1}$ представимо в виде $u = v_1 v_2 \dots v_k$, где $v_i \in A_n$, и для любого $w \in A_n$ существует i, такое что $v_i = w$.

Мы говорим, что $\alpha \in \Sigma^{\mathbb{N}}$ порождается строгой Σ-схемой $\langle A_n, l_n \rangle$, если для любых i и n имеем

$$\alpha[il_n, (i+1)l_n - 1] \in A_n.$$

Как несложно видеть из соображений компактности, каждая строгая схема порождает какую-то последовательность. В [9] авторы доказывают, что любая последовательность, порождённая строгой схемой, является строго почти периодической. Более того, каждая строго почти периодическая последовательность порождается некоторой схемой.

Доказательство теоремы 12. Для доказательства достаточно построить такие $\omega_n \in \mathcal{EAP}$, $\omega \in \mathcal{AP} \setminus \mathcal{EAP}$ и общий регулятор f для них, что $\omega_n \to \omega$.

Положим $a_0 = 1$, $a_1 = 10011$, и далее по правилу: $a_{n+1} = a_n \bar{a}_n \bar{a}_n a_n a_n$. Обозначим слово $a_n a_n a_n a_n$ через c_n. Введём для удобства обозначение $l_n = 5^n - 1 = |c_0 c_1 \dots c_{n-1}|$. Рассмотрим $\omega = c_0 c_1 c_2 c_3 \dots$ и $\nu = \lim_{n\to\infty} a_n$. Как было доказано в [15], $\omega \in \mathcal{AP} \setminus \mathcal{EAP}$. Пусть $\omega_n = c_0 c_1 \dots c_n \nu$. Поскольку $\nu \in \mathcal{SAP}$ (так как порождено строгой схемой $\langle \{a_n, \bar{a}_n\}, 5^n \rangle$), то $\omega_n \in \mathcal{EAP}$. Ясно, что $\omega_n \to \omega$, и осталось выбрать общий для всех ω_n регулятор f. Мы получим конечное количество требований вида $f(k) \geqslant \alpha$, после чего в качестве $f(k)$ можем взять максимум по всем таким α.

Пусть $v = \omega_n[i, j]$, $|v| = k$ входит в $\omega_n = c_0 c_1 \dots c_n \nu$ бесконечно много раз. Тогда v входит в ν, а значит, и в некоторое a_m. Поэтому v входит в ω бесконечно много раз, и достаточно взять $f(k) \geqslant \mathrm{R}_\omega(k) + \mathrm{R}_\nu(k)$.

Пусть $v = \omega_n[i, j]$, $|v| = k$ входит в ω_n конечное количество раз. Ясно тогда, что $i < l_n$. Если $j > l_n$, то тогда $k > 5^n$, так как иначе v содержится в некотором a_m, а значит, входит в ν бесконечно много раз. Но неравенство $k > 5^n$ выполнимо только для конечного количества значений n, что даёт лишь конечное количество условий на $f(k)$. Пусть теперь $j \leqslant l_n$. Но тогда v содержится в $c_0 c_1 \dots c_n$, причём входит в ω конечное количество раз (иначе бы v содержалось в каком-то a_m), поэтому в этом случае подойдёт $f(k) \geqslant \mathrm{R}_\omega(k)$. $\square$

Следующая теорема показывает, что невозможно даже эффективно разделять классы $\mathcal{SAP}$ и $\mathcal{P}$.

Теорема 13. *По $\omega \in \mathcal{SAP}$ и её регулятору $f \geqslant \mathrm{R}_\omega$ невозможно алгоритмически определять, верно ли, что $\omega \in \mathcal{P}$.*

Доказательство. Для доказательства достаточно построить такие $\omega_n \in \mathcal{P}$, $\omega \in \mathcal{SAP} \setminus \mathcal{P}$ с общим для всех ω_n регулятором $f \geqslant \mathrm{R}_{\omega_n}$, что $\omega_n \to \omega$.

Любая строго почти периодическая последовательность может быть получена из строгой Σ-схемы $\langle A_n, l_n\rangle$. Усилим главное условие на A_n, а именно, будем рассматривать строгие схемы, для которых верно: для любого $n \in \mathbb{N}$ каждое $u \in A_{n+1}$ имеет вид $u = v_1 v_2 \dots v_k$, где $v_i \in A_n$, причём для каждых $w_1, w_2 \in A_n$ найдётся $i < k$, такое что $v_i v_{i+1} = w_1 w_2$. Заметим, что такие схемы существуют и могут порождать непериодические последовательности, например, $\langle\{a_n, \bar{a}_n\}, 2^n\rangle$, порождающая ω_T.

Пусть $\langle A_n, l_n\rangle$ — строгая схема, удовлетворяющая вышеуказанному усиленному условию, и пусть она порождает $\omega \notin \mathcal{P}$. Определим $p_n = \omega[0, l_n]$, таким образом, $p_n \in A_n$ и $\lim_{n\to\infty} p_n = \omega$. Положим $\omega_n = p_n p_n p_n \ldots \in \mathcal{P}$. Ясно, что $\omega_n \to \omega$, и осталось выбрать общий регулятор f для всех ω_n.

Пусть $v = \omega_n[i, j]$, $|v| = k$ (это сразу означает, что v входит в ω_n бесконечно много раз, так как $\omega_n \in \mathcal{P}$). Неравенство $k \geqslant |p_n| = l_n$ может выполняться лишь для конечного количества различных n, что даёт лишь конечное количество условий на $f(k)$. Теперь можно считать, что $k < l_n$. Возьмём такое t, что $l_{t-1} < k \leqslant l_t$ (важно, что t не зависит от n и однозначно определяется по k). Тогда $t < n$. Существует m, такое что $ml_t \leqslant i$ и $j \leqslant (m+2)l_t$, то есть v содержится в некотором ab, где $a, b \in A_t$, и значит, благодаря свойству схемы, v входит в любое $c \in A_{t+1}$. Но на каждом отрезке длины $2l_{t+1}$ последовательности ω_n найдётся вхождение какого-нибудь $c \in A_{t+1}$ (целиком входящего в какое-то p_n), а значит, и вхождение v. Таким образом, достаточно $f \geqslant 2l_{t+1}$. □

Из конструкции в доказательстве теоремы 13 следует существование бесконечного количества периодических последовательностей с общим регулятором (априори не очевидное, так как самая естественная оценка на регулятор — период последовательности — в этой ситуации стремится к бесконечности). Эта конструкция пригодится нам в следующей теореме об алгоритмической неразрешимости. Она говорит о том, что даже приписав к строго почти периодической последовательности один символ, мы, вообще говоря, уже не можем проверять, является ли она по-прежнему строго почти периодической.

Теорема 14. *По $\omega \in \mathcal{EAP}$, её регулятору $f \geqslant \mathrm{R}_\omega$ и некоторому $l \geqslant \mathrm{pr}(\omega)$ невозможно алгоритмически находить $\mathrm{pr}(\omega)$.*

Лемма 15. *Если $a\omega \in \mathcal{SAP}$, где $a \in \Sigma^*$, и ω периодическая с периодом l, то $a\omega$ тоже периодическая с периодом l.*

Доказательство. Достаточно доказать лемму для однобуквенного слова a. Пусть $\alpha = 012\ldots(l-1)012\ldots(l-1)012\ldots(l-1)\ldots$ — периодическая последовательность, составленная из букв алфавита $\Sigma_l = \{0, 1, 2, \ldots, l-1\}$. Тогда по следствию 6 получаем, что $a\omega \times \alpha \in \mathcal{SAP}$. В этой последовательности символ $\langle a, 0\rangle$ встречается бесконечно много раз, откуда ясно, что $a = \omega(l)$, что и требовалось. □

Доказательство теоремы 14. Для доказательства достаточно построить такие $\omega_n \in \mathcal{EAP}$, $\omega \in \mathcal{SAP}$ и общий регулятор f для всех ω_n, что $\omega_n \to \omega$ и $\mathrm{pr}(\omega_n) = 1$ (из $\omega \in \mathcal{SAP}$ следует $\mathrm{pr}(\omega) = 0$).

Заметим, что $1\omega_T \in \mathcal{SAP}$. Для этого достаточно проверить, что любой префикс этой последовательности входит в неё бесконечно много раз. Но действительно, для любого n слова $a_n a_n$ и $\bar{a}_n a_n$ входят в ω_T, а значит, и $1a_n$ тоже. Аналогично показывается, что $0\omega_T \in \mathcal{SAP}$.

Как видно из доказательства теоремы 13, можно выбрать такую последовательность $k_n \to \infty$, что все периодические последовательности вида

$\omega(0)\ldots\omega(k_n)\omega(0)\ldots\omega(k_n)\omega(0)\ldots$ имеют общий регулятор f. Выберем такую подпоследовательность m_n последовательности k_n, что все символы $\omega(m_n)$ одинаковы. Предположим, что они равны 0.

Положим $\omega_n = 1\omega(0)\ldots\omega(m_n)\omega(0)\ldots\omega(m_n)\omega(0)\ldots$, $\omega = 1\omega_T$. Ясно, что можно выбрать для них общую оценку сверху g на регулятор. Для этого достаточно, чтобы было выполнено $g(k) \geqslant f(k)+1$ (исходя из рассмотрения слов, входящих бесконечное количество раз) и $g(k) \geqslant k$ (исходя из рассмотрения слов, входящих лишь конечное количество раз, что возможно только для префиксов, входящих ровно один раз).

Если бы $\omega_n \in \mathcal{SAP}$, то по лемме 15 последовательность ω_n периодична с периодом m_n, но мы специально подобрали $\omega_n(0) = 1 \neq \omega_n(m_n) = 0$. Таким образом, $\mathrm{pr}(\omega_n) = 1$.

Случай, когда все символы $\omega(m_n)$ равны 1, разбирается аналогично (тогда ω_n начинается с 0). $\square$

5 Благодарности

Автор благодарен А. Л. Семёнову и Ан. А. Мучнику за помощь в работе и привлечение моего внимания к этой теме, а также М. Раскину, А. Румянцеву, А. Шеню и остальным участникам Колмогоровского семинара за интересные и полезные обсуждения.

Литература

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

[2] J.-P. Allouche, J. Shallit. *Automatic Sequences.* Cambridge University Press, 2003.

[3] J. Cassaigne. *Recurrence in infinite words.* Proceedings of the 18th Symposium on Theoretical Aspects of Computer Science (STACS 2001), Springer Verlag, pp. 1–11, 2001.

[4] A. Cobham. *Uniform taq sequences.* Math. Systems Theory, 6, pp. 164–192, 1972.

[5] K. Jacobs. *Maschinenerzeugte 0-1-Folgen.* Selecta Mathematica II. Springer Verlag: Berlin, Heidelberg, New York, 1970.

[6] M. Keane. *Generalized Morse sequences.* Z. Wahrscheinlichkeitstheorie verw. Geb., 10, pp. 335–353, 1968.

[7] M. Morse, G. A. Hedlund. *Symbolic dynamics.* American Journal of Mathematics, 60, pp. 815–866, 1938.

[8] M. Morse, G. A. Hedlund. *Symbolic dynamics II: Sturmian trajectories.* American Journal of Mathematics, 62, pp. 1–42, 1940.

[9] An. Muchnik, A. Semenov, M. Ushakov. *Almost periodic sequences.* Theoretical Computer Science, vol. 304, pp. 1–33, 2003.

[10] A. Rumyantsev, M. Ushakov. *Forbidden Substrings, Kolmogorov Complexity and Almost Periodic Sequences.* Lecture Notes in Computer Science, vol. 3884, STACS 2006, pp. 396–407.

[11] Yu. Pritykin. *Almost Periodicity, Finite Automata Mappings and Related Effectiveness Issues.* Proceedings of WoWA (student session), St. Petersburg, Russia, June 7th, 2006 (satellite to CSR'06), to appear. See also `http://arXiv.org/abs/cs.DM/0607009`.

[12] M. Sipser. *Introduction to the Theory of Computation.* PWS Publishing Company: Boston, 1997.

[13] A. Thue. *Über unendliche Zeichenreihen.* Norske vid. Selsk. Skr. Mat. Nat. Kl., 7, pp. 1–22, 1906. Reprinted in *Selected mathematical papers of Axel Thue*, Universitetsforlaget, Oslo, pp. 139–158, 1977.

[14] A. Weber. *On the valuedness of finite transducers.* Acta Informatica, 27, pp. 749–780, 1989.

[15] Ю. Л. Притыкин. *Конечно-автоматные преобразования строго почти периодических последовательностей.* Математические заметки, т. 80, N 5, с. 751–756, 2006. Английская версия на `http://arXiv.org/abs/cs.DM/0605026`.

[16] Ю. Л. Притыкин. *Конечно-автоматные преобразования почти периодических последовательностей и алгоритмическая неразрешимость.* Труды XXVIII Конференции молодых учёных, мех.-мат. ф-т МГУ им. Ломоносова, 2006, готовится к публикации.

[17] М. А. Раскин. *Об оценке регулятора автоматного образа почти периодической последовательности.* Труды XXVIII Конференции молодых учёных, мех.-мат. ф-т МГУ им. Ломоносова, 2006, готовится к публикации.

[18] А. Л. Семёнов. *О некоторых расширениях арифметики сложения натуральных чисел.* Известия АН СССР, серия математическая, т. 43, N 5, с. 1175–1195, 1979.

[19] А. Л. Семёнов. *Логические теории одноместных функций на натуральном ряде.* Известия АН СССР, серия математическая, т. 47, N 3, с. 623–658, 1983.