О каноническом регулярном представлении s-тонких языков

П. С. Дергач

В работе изучаются спектры класса тонких языков. Дается классификация языков по их спектральным свойствам. Для спектра тонкого языка определяется два понятия размерности. Оказывается, что для произвольного натурального числа n можно привести пример спектра, у которого и первая, и вторая размерности равны n. Рассматривается вопрос о том, какие значения пар этих размерностей в совокупности бывают у спектров тонких языков. Для каждой пары m, n натуральных чисел таких, что $m = n = 1$ или $2 \leqslant m \leqslant n$ удается построить тонкий язык со спектром, у которого первая размерность равна m, а вторая равна n. Используемая при этом техника нетривиальна и позволяет поставить и решить сразу несколько теоретико-числовых задач, которые сами по себе представляют определенный интерес.

Ключевые слова: тонкие языки, регулярные выражения, алфавитное кодирование.

Введение

Работа является продолжением исследований автора в области теории кодирования регулярных текстов. Основной задачей этой теории является проблема декодирования при алфавитном кодировании (сокращенно ДПАК). Ее смысл в том, чтобы по произвольной паре, состоящей из регулярного языка $P \subseteq A^*$ и функции алфавитного кодирования $f : A \to B^*$, определить, является ли f однозначно декодируемой на P. Здесь A и B — произвольные фиксированные конечные алфавиты. О понятии регулярных языков и алфавитного кодирования можно прочитать в [1] и [2] соответственно. Данная проблема была успешна разрешена Ал. А. Марковым в [3]. В свою очередь, в [4]

автором было получено альтернативное решение проблемы ДПАК, сводящее ее к конечному перебору слов из P длины не более $L(R, f)$, где $L(R, f)$ — функция, явным образом зависящая от количества состояний задающего регулярный язык автомата и от длины схемы кодирования. Положительное решение проблемы ДПАК позволило поставить вопрос об эффективности функций алфавитного кодирования и ввести на них отношение строгого частичного порядка. А именно, говорим, что функция f_1 богаче, чем функция f_2, если класс регулярных языков, доставляющих положительное решение проблемы ДПАК для функции f_1, шире соответствующего класса для функции f_2. При помощи этого отношения на множестве функций алфавитного кодирования можно построить решетку. В данной работе изучается один из минимальных классов этой решетки. В дальнейшем он будет называться классом 1-тонких языков и обозначаться через T_1. Соответствующая T_1 функция кодирования имеет вид $\tilde{f}(a_i) = b$, где $i = 1, \ldots, |A|$ и $b \in B$. Выбор этого класса обусловлен, во-первых, тем, что $\tilde{f}$, в отличие от функций других минимальных классов, будет инвариантна относительно перестановки букв алфавита A. Это позволяет дать классу T_1 альтернативное определение, не использующее таких понятий, как кодирование и решетка. А именно, класс T_1 можно получить, если наложить на регулярные языки ограничение, запрещающее им содержать различные слова одинаковой длины. Во-вторых, алгоритм решения проблемы ДПАК, изложенный в [4], работает для T_1 за линейное от перебираемой длины время. Эти соображения позволяют сделать предположение о принципиальной роли класса 1-тонких языков в исследуемой модели. В работе описывается структура элементов класса T_1, а также приводится их каноническое универсальное представление в терминах регулярных выражений. О понятии регулярных выражений можно прочитать в [1]. Следует понимать, что T_1 является минимальным классом и использовать его для описания свойств регулярных языков общего вида затруднительно. Поэтому в работе делается обобщение класса T_1, для которого, с одной стороны, все еще можно быстро решать проблему ДПАК, и, с другой стороны, с помощью которого аппроксимируется уже произвольный регулярный язык. Это делается следующим образом. Для каждого натурального значения s рассматривается класс регулярных языков, в которых максимальное количество несовпадающих слов одинаковой длины меньше бесконечности и равно s. Такие языки на-

зываются s-тонкими а класс всех s-тонких языков обозначается через T_s. В качестве обобщения класса T_1 берется класс $\bigcup\limits_{i=1}^{\infty} T_s$. Он называется классом тонких языков и обозначается через T. Для класса T также удается описать его регулярную структуру и построить универсальное представление его элементов.

1. Основные понятия и результаты

Пусть A — непустое конечное множество. В дальнейшем будем называть его *алфавитом*, а его элементы — *буквами*. *Словами* называем произвольные конечные последовательности букв алфавита A. Для удобства рассматриваем при этом также *пустое* слово, не имеющее ни одной буквы и обозначаемое λ. Все слова кроме пустого называем *непустыми*. Множество слов в алфавите A обозначаем через A^*.

Пусть $\alpha = a(1)\dots a(k)$. Говорим, что k — *длина* слова α и обозначаем ее через $l(\alpha)$. Длина пустого слова равна 0. Пусть α, β, γ — слова такие, что $\alpha = \beta\gamma$. В этом случае говорим, что β — *начало* слова α, γ — *конец* слова α. Начало слова α, имеющее длину l, обозначаем через $S_l(\alpha)$; окончание слова α, имеющее длину n, обозначаем через $F_n(\alpha)$. Вводим обозначение $\alpha(n, m)$ для слова $F_{m-n}(S_m(\alpha))$, где $l(\alpha) \geqslant m > n \geqslant 0$.

Пусть β — непустое слово в алфавите A. Если существует слово α в алфавите A такое, что $\beta = \alpha^k$ для некоторого $k > 1$, то говорим, что β *измельчимо*. Иначе говорим, что β *неизмельчимо*. Здесь через α^k обозначается конкатенация k слов α.

Введем понятие спектра множеств. Пусть $P \subseteq A^*$ — произвольное множество слов. *Спектром* этого множества называем множество $\{l(\alpha) | \alpha \in P\}$ и обозначаем его через $Sp(P)$. Пусть $P_1, P_2 \subseteq A^*$. Будем говорить, что эти множества *спектрально независимы*, если их спектры не пересекаются. Пусть $P_1, \dots, P_r \subseteq A^*$, $r \geqslant 1$. Будем говорить, что эти множества *спектрально независимы в совокупности*, если любые два из них спектрально независимы. В противном случае говорим, что эти множества *спектрально зависимы в совокупности*.

Для произвольного множества $P \subseteq A^*$ через $|P|$ обозначаем его мощность. Если множество P конечно, то пишем $|P| < \infty$.

Введем понятие 1-тонкого множества в алфавите A. Регулярное множество $P \subseteq A^*$ называем *1-тонким в алфавите* A, если для про-

извольной пары слов $\alpha, \beta \in P$ из условия $|\alpha| = |\beta|$ следует условие $\alpha = \beta$. Другими словами, в P не должно быть двух несовпадающих слов одинаковой длины.

Пусть $(\alpha, \beta, \gamma, k, m) \in (A^*)^3 \times \mathbb{N} \times (\mathbb{N} \cup \{0\})$. Говорим, что $(\alpha, \beta, \gamma, m, n)$ — *порождающий след*, если выполнено одно из двух условий:

1. $\beta = \gamma = \lambda$, $k = 1, m = 0$;

2. $\beta \neq \lambda$, у α и β нет одинаковых непустых окончаний, β неизмельчимо и не является началом γ.

Для произвольного регулярного выражения $\mathfrak{P}$ в алфавите A через $|\mathfrak{P}|$ обозначаем соответствующее ему регулярное множество в алфавите A.

Говорим, что множество $P \subseteq A^*$ является *прогрессивным*, если оно представимо с помощью регулярного выражения $\alpha \cdot (\beta^k)^* \cdot \beta^m \cdot \gamma$ для некоторого порождающего следа $(\alpha, \beta, \gamma, k, m)$. В этом случае говорим также, что множество P *имеет порождающий след* $(\alpha, \beta, \gamma, k, m)$. Упорядоченную тройку (α, β, γ) называем *основанием* множества P. Позже будет доказано, что у любого прогрессивного множества может быть только один порождающий след, а значит, и только одно основание. Называем множество $P \subseteq A^*$ *общепрогрессивным*, если оно является конечным объединением прогрессивных множеств с одинаковым основанием.

Пусть A — непустой конечный алфавит и $s \in \mathbb{N}$. Введем понятие s-тонкого множества в алфавите A. Регулярное множество P, $P \subseteq A^*$, называем *s-тонким в алфавите A*, если

1) существуют s слов $\beta_1, \beta_2, \ldots, \beta_s \in P$ таких, что $l(\beta_1) = l(\beta_2) = \ldots = l(\beta_s)$ и для них не существуют $i, j \in \mathbb{N}$ такие, что $1 \leqslant i < j \leqslant s$ и $\beta_i = \beta_j$;

2) для любых $s+1$ слов $\alpha_1, \alpha_2, \ldots, \alpha_{s+1} \in P$ таких, что $l(\alpha_1) = l(\alpha_2) = \ldots = l(\alpha_{s+1})$ существуют $i, j \in \mathbb{N}$ такие, что $1 \leqslant i < j \leqslant s+1$ и $\alpha_i = \alpha_j$.

Другими словами, в P должно быть s несовпадающих слов одинаковой длины, но не должно быть $s+1$ несовпадающих слов одинаковой длины. Заметим, что при $s = 1$ мы получаем определение класса 1-тонких множеств, приведенное выше.

Для всех $s \in \mathbb{N}$ обозначаем через T_s множество всех s-множеств в алфавите A. Через T обозначаем множество $\bigcup\limits_{i=1}^{\infty} T_s$. Называем это

множество *классом тонких множеств*, а его элементы — *тонкими множествами*.

Теорема 1. *Любое конечное объединение спектрально независимых в совокупности общепрогрессивных множеств является 1-тонким множеством.*

Теорема 2. *Любое 1-тонкое множество представимо в виде конечного объединения спектрально независимых в совокупности общепрогрессивных множеств.*

Теорема 3. *Любое конечное объединение непересекающихся прогрессивных множеств является тонким множеством.*

Теорема 4. *Любое тонкое множество представимо в виде конечного непересекающегося объединения прогрессивных множеств.*

2. Доказательство вспомогательных утверждений

Лемма 1. *Пусть P — регулярное множество в алфавите A. Тогда оно представимо регулярным выражением вида*

$$\bigvee_{i=1}^{k} \alpha_{i,1} \cdot (\mathfrak{P}_{i,1})^* \cdot \alpha_{i,2} \cdot \ldots \cdot \alpha_{i,s(i)-1} \cdot (\mathfrak{P}_{i,s(i)-1})^* \cdot \alpha_{i,s(i)},$$

где $k, s(1), \ldots, s(k)$ — произвольные натуральные числа, $\alpha_{1,1}, \ldots, \alpha_{k,s(k)}$ — произвольные слова в алфавите A, $\mathfrak{P}_{1,1}, \ldots, \mathfrak{P}_{k,s(k)-1}$ — произвольные регулярные выражения в алфавите A.

Доказательство. Будем доказывать утверждение индукцией по минимальной длине l вывода P из λ и букв алфавита A с помощью операций $\vee, \cdot, *$.

База индукции ($l = 0$).

1. Если $R = \{\lambda\}$, то R представимо регулярным выражением λ. Здесь $k = 1, s(1) = 1, \alpha_{1,1} = \lambda$.

2. Если $R = \{a\}$, где $a \in A$, то R представимо регулярным выражением a. Здесь $k = 1, s(1) = 1, \alpha_{1,1} = a$.

Переход индукции $(1, \ldots, l \Rightarrow l+1)$.

1. Пусть минимальная длина вывода P равна l и $P = P_1^*$. Тогда минимальная длина вывода P_1 меньше l и по предположению индукции P_1 представимо регулярным выражением

$$\mathfrak{P}_1 = \bigvee_{i=1}^{k} \alpha_{i,1} \cdot (\mathfrak{P}_{i,1})^* \cdot \alpha_{i,2} \cdot \ldots \cdot \alpha_{i,s(i)-1} \cdot (\mathfrak{P}_{i,s(i)-1})^* \cdot \alpha_{i,s(i)}.$$

Значит, P представимо регулярным выражением $(\mathfrak{P}_1)^*$. Здесь $k = 1$, $s(1) = 2, \alpha_{1,1} = \alpha_{1,2} = \lambda, \mathfrak{P}_{1,1} = \mathfrak{P}_1$.

2. Пусть минимальная длина вывода P равна l и $P = P_1 \vee P_2$. Тогда минимальная длина вывода P_1 и P_2 меньше l и по предположению индукции P_1 и P_2 представимы регулярными выражениями

$$\mathfrak{P}_1 = \bigvee_{i=1}^{k} \alpha_{i,1} \cdot (\mathfrak{P}_{i,1})^* \cdot \alpha_{i,2} \cdot \ldots \cdot \alpha_{i,s(i)-1} \cdot (\mathfrak{P}_{i,s(i)-1})^* \cdot \alpha_{i,s(i)},$$

$$\mathfrak{P}_2 = \bigvee_{i=1}^{m} \beta_{i,1} \cdot (\mathfrak{C}_{i,1})^* \cdot \beta_{i,2} \cdot \ldots \cdot \beta_{i,t(i)-1} \cdot (\mathfrak{C}_{i,t(i)-1})^* \cdot \beta_{i,t(i)}$$

соответственно. Значит, P представимо регулярным выражением

$$(\mathfrak{P}_1 \vee \mathfrak{P}_2) = \bigvee_{i=1}^{k+m} \gamma_{i,1} \cdot (\mathfrak{I}_{i,1})^* \cdot \gamma_{i,2} \cdot \ldots \cdot \gamma_{i,u(i)-1} \cdot (\mathfrak{I}_{i,u(i)-1})^* \cdot \gamma_{i,u(i)},$$

где

$$u(i) = \begin{cases} s(i), & \text{если } k \geqslant i \geqslant 1; \\ t(i-k), & \text{если } k+m \geqslant i \geqslant k+1, \end{cases}$$

$$\gamma_{i,j} = \begin{cases} \alpha_{i,j}, & \text{если } k \geqslant i \geqslant 1, u(i) \geqslant j \geqslant 1; \\ \beta_{i-k,j}, & \text{если } k+m \geqslant i \geqslant k+1, u(i) \geqslant j \geqslant 1, \end{cases}$$

$$\mathfrak{I}_{i,j} = \begin{cases} \mathfrak{P}_{i,j}, & \text{если } k \geqslant i \geqslant 1, u(i)-1 \geqslant j \geqslant 1; \\ \mathfrak{C}_{i-k,j}, & \text{если } k+m \geqslant i \geqslant k+1, u(i)-1 \geqslant j \geqslant 1. \end{cases}$$

3. Пусть минимальная длина вывода P равна l и $P = P_1 \cdot P_2$. Тогда минимальная длина вывода P_1 и P_2 меньше l и по предположению

индукции P_1 и P_2 представимы регулярными выражениями

$$\mathfrak{P}_1 = \bigvee_{i=1}^{k} \alpha_{i,1} \cdot (\mathfrak{P}_{i,1})^* \cdot \alpha_{i,2} \cdot \ldots \cdot \alpha_{i,s(i)-1} \cdot (\mathfrak{P}_{i,s(i)-1})^* \cdot \alpha_{i,s(i)},$$

$$\mathfrak{P}_2 = \bigvee_{i=1}^{m} \beta_{i,1} \cdot (\mathfrak{C}_{i,1})^* \cdot \beta_{i,2} \cdot \ldots \cdot \beta_{i,t(i)-1} \cdot (\mathfrak{C}_{i,t(i)-1})^* \cdot \beta_{i,t(i)}$$

соответственно. Значит, P представимо регулярным выражением

$$(\mathfrak{P}_1 \cdot \mathfrak{P}_2) = \bigvee_{i=1}^{km} \gamma_{i,1} \cdot (\mathfrak{I}_{i,1})^* \cdot \gamma_{i,2} \cdot \ldots \cdot \gamma_{i,u(i)-1} \cdot (\mathfrak{I}_{i,u(i)-1})^* \cdot \gamma_{i,u(i)},$$

где

$$u(i) = s\left(\left[\frac{i-1}{m}\right]+1\right) + t\left(i - \left[\frac{i-1}{m}\right] m\right) - 1, \quad \text{при } km \geqslant i \geqslant 1,$$

$$\gamma_{i,j} = \begin{cases} \alpha_{\left[\frac{i-1}{m}\right]+1,j}, & \text{если } km \geqslant i \geqslant 1, s\left(\left[\frac{i-1}{m}\right]+1\right) - 1 \geqslant j \geqslant 1; \\ \alpha_{\left[\frac{i-1}{m}\right]+1,j} \cdot \beta_{i-\left[\frac{i-1}{m}\right]m,1}, & \text{если } km \geqslant i \geqslant 1, j = s\left(\left[\frac{i-1}{m}\right]+1\right); \\ \beta_{i-\left[\frac{i-1}{m}\right]m,j-s\left(\left[\frac{i-1}{m}\right]+1\right)+1}, & \text{если } km \geqslant i \geqslant 1, \\ & u(i) \geqslant j \geqslant s\left(\left[\frac{i-1}{m}\right]+1\right)+1, \end{cases}$$

$$\mathfrak{I}_{i,j} = \begin{cases} \mathfrak{P}_{\left[\frac{i-1}{m}\right]+1,j}, & \text{если } km \geqslant i \geqslant 1, s\left(\left[\frac{i-1}{m}\right]+1\right) - 1 \geqslant j \geqslant 1; \\ \mathfrak{C}_{i-\left[\frac{i-1}{m}\right]m,j-s\left(\left[\frac{i-1}{m}\right]+1\right)+1}, & \text{если } km \geqslant i \geqslant 1, \\ & u(i) - 1 \geqslant j \geqslant s\left(\left[\frac{i-1}{m}\right]+1\right). \end{cases}$$

Таким образом, утверждение индукции, а с ним и лемма 1 доказана.

Лемма 2. *Пусть $\alpha, \beta \in A^* \setminus \{\lambda\}$ и $\alpha^k = \beta^m$ для некоторых $k, m \in \mathbb{N}$. Тогда существует $\nu \in A^* \setminus \{\lambda\}$ такое, что $l(\nu) = (l(\alpha), l(\beta))$, $\alpha = \nu^{\frac{l(\alpha)}{l(\nu)}}$, $\beta = \nu^{\frac{l(\beta)}{l(\nu)}}$.*

Доказательство. Пусть $l(\alpha) = a$, $l(\beta) = b$, $(l(\alpha), l(\beta)) = c$. Через S будем обозначать множество $\{a \cdot i | \frac{b}{c} \geqslant i \geqslant 1\}$. Так как для любого $x \in S$ имеем $x \equiv 0 \pmod{c}$ и $b \equiv 0 \pmod{c}$, то числа из S могут принимать по модулю b только остатки, кратные c. Всего таких остатков $\frac{b}{c}$, как и чисел в S. Заметим, что никакие два разных числа $a \cdot i_1, a \cdot i_2$ из S не могут давать одинаковые остатки по модулю b, так как иначе

их разность $a \cdot (i_1 - i_2)$ делилась бы на b, что, очевидно, невозможно в силу $\frac{b}{c} > i_1 - i_2$. Значит, существует такое $x \in S$, что $x \equiv c \pmod b$. Другими словами, существуют такие $\frac{b}{c} \geqslant s \geqslant 1$, $\frac{a}{c} > t \geqslant 0$, что $as - bt = c$.

Обозначим через δ слово $\alpha^{2ks} = \beta^{2ms}$. Для любого $l(\delta) - a \geqslant i \geqslant 1$ очевидно, что $\delta(i-1, i) = \delta(i+a-1, i+a)$. Аналогично, для любого $l(\delta) \geqslant i \geqslant b+1$ верно, что $\delta(i-1, i) = \delta(i-b-1, i-b)$. Поэтому для любого $c \geqslant i \geqslant 1$ верно, что $\delta(i-1, i) = \delta(i+as-1, i+as) = \delta(i+as-bt-1, i+as-bt) = \delta(i+c-1, i+c)$. Здесь мы неявно воспользовались тем, что $i+as \leqslant c+as \leqslant a+as \leqslant 2aks = l(\delta)$. Значит, для любого $l(\delta)-2c \geqslant i \geqslant 0$ верно, что $\delta(i, i+c) = \delta(i+c, i+2c)$. Так как a делится на c, то $\alpha = \delta(0, a) = \delta(0, c)^{\frac{a}{c}}$. Аналогично получаем, что $\beta = \delta(0, b) = \delta(0, c)^{\frac{b}{c}}$. Осталось положить $\nu = \delta(0, c)$. Утверждение леммы 2 доказано.

Лемма 3. *Пусть*

$$x_1, \ldots, x_k \in \mathbb{N}, \quad r = НОД(x_1, \ldots, x_k),$$
$$H = \{a_1 \cdot x_1 + \ldots + a_k \cdot x_k \mid a_1, \ldots, a_k \in \mathbb{N} \cup \{0\}\}.$$

Тогда существует $n_0 \in \mathbb{N}$ такое, что для любых $n \geqslant n_0$ из $r|n$ следует $n \in H$.

Доказательство. Будем доказывать утверждение индукцией по k.

База индукции ($k = 1, 2$).

При $k = 1$ имеем $L = \{a \cdot x_1 \mid a \in \mathbb{N} \cup \{0\}\}$, $r = x_1$. Поэтому для любого $n \in \mathbb{N}$ если $r|n$, то $n \in H$. Значит, за n_0 можно взять, например, 1.

При $k = 2$ рассмотрим множество $S = \{a \cdot x_1 \mid \frac{x_2}{r} \geqslant a \geqslant 1\}$. Так как для любого $x \in S$ имеем $x \equiv 0 \pmod r$ и $x_2 \equiv 0 \pmod r$, то числа из S могут принимать по модулю x_2 только остатки, кратные r. Всего таких остатков $\frac{x_2}{r}$, как и чисел в S. Заметим, что никакие два разных числа $i_1 \cdot x_1, i_2 \cdot x_1$ из S не могут давать одинаковые остатки по модулю x_2, так как иначе их разность $(i_1 - i_2) \cdot x_1$ делилась бы на x_2, что, очевидно, невозможно в силу $\frac{x_2}{r} > i_1 - i_2$. Значит, в S встречаются все остатки по модулю x_2, кратные r. Возьмем $n_0 = \frac{x_1 \cdot x_2}{r}$. Тогда для всех $n \geqslant n_0$ из $r|n$ будет следовать, что существует $x \in S$, для которого верно $x \equiv n \pmod{x_2}$. Отсюда тривиально следует, что для всех $n \geqslant n_0$ из $r|n$ следует $n \in H$.

Переход индукции $(1, \ldots, k \Rightarrow k+1)$.
Пусть $x_1, \ldots, x_{k+1} \in \mathbb{N}$ и

$$H = \{a_1 \cdot x_1 + \ldots + a_{k+1} \cdot x_{k+1} \mid a_1, \ldots, a_{k+1} \in \mathbb{N} \cup \{0\}\},$$
$$H_1 = \{a_1 \cdot x_1 + \ldots + a_k \cdot x_k \mid a_1, \ldots, a_k \in \mathbb{N} \cup \{0\}\},$$
$$r = \text{НОД}(x_1, \ldots, x_{k+1}), r_1 = \text{НОД}(x_1, \ldots, x_k).$$

По предположению индукции, существует $n_0 \in \mathbb{N}$ такое, что для всех $n \geqslant n_0$ из $r_1 | n$ следует $n \in H_1$. Обозначим через p произвольное натуральное число такое, что $p \geqslant n_0$ и (p, x_{k+1})=1. Например, в качестве такого p можно взять произвольное простое число, большее n_0 и x_{k+1}. Обозначим через H_2 множество

$$\{a_1 \cdot r_1 \cdot p + a_2 \cdot x_{k+1} \mid a_1, a_2 \in \mathbb{N} \cup \{0\}\}.$$

Заметим, что $(r_1 \cdot p, x_{k+1}) = (r_1, x_{k+1}) = (\text{НОД}(x_1, \ldots, x_k), x_{k+1}) = r$. Из этого и предположения индукции, примененного к числам $r_1 \cdot p, x_{k+1}$ следует, что существует $n_1 \in \mathbb{N}$ такое, что для всех $n \geqslant n_1$ из $r | n$ получаем $n \in H_2$. Так как $r_1 | r_1 \cdot p$ и $r_1 \cdot p \geqslant n_0$, то $r_1 \cdot p \in H_1$. Значит, $r_1 \cdot p, x_{k+1} \in H$. Поэтому $H_2 \subseteq H$ и для всех $n \geqslant n_1$ из $r | n$ следует $n \in H$.

Таким образом, утверждение индукции, а с ним и лемма 3 доказана.

Лемма 4. *Пусть $R \subseteq A^*$ и R^* — 1-тонкое множество. Тогда существуют $\delta, \gamma \in A^*$ и $R_1 \subseteq A^*$ такие, что $|R_1| < \infty$, $R^* = R_1 \cup (\{\delta\} \cdot \{\gamma\}^*)$.*

Доказательство. Пусть $|R| = 1$, то есть $R = \{\mu\}$ для некоторого $\mu \in A^*$. Тогда полагаем $R_1 = \varnothing$, $\delta = \lambda$, $\gamma = \mu$.

Пусть $|R| > 1$. Так как $\{R \setminus \{\lambda\}\}^* = R^*$, то можно считать, что $\lambda \notin R$. Будем строить последовательность множеств $M_1 \subset M_2 \subset \ldots \subset M_l \subset R$ такую, что:
1. $M_i = \{\alpha_1, \alpha_2, \ldots, \alpha_{i+1}\}$ для всех $l \geqslant i \geqslant 1$;
2. $\alpha_i = \nu_j^{a_i(j)}$ для всех $l \geqslant j \geqslant 1$ и $j+1 \geqslant i \geqslant 1$, где $\nu_j \in A^* \setminus \{\lambda\}$ и $a_i(j) \in \mathbb{N}$;
3. $\text{НОД}(a_1(j), \ldots, a_{j+1}(j)) = 1$ для всех $l \geqslant j \geqslant 1$.

Построим M_1. Так как $|R| > 1$, то существуют $\alpha, \beta \in R$, $\alpha \neq \beta$. Обозначим $l(\alpha) = a$, $l(\beta) = b$. Тогда

$$\alpha^b, \beta^a \in R^*, \quad l(\alpha^b) = l(\beta^a) = ab.$$

Значит, $\alpha^b = \beta^a$. По лемме 2 существует $\nu \in A^* \setminus \{\lambda\}$ такое, что

$$l(\nu) = (a, b), \quad \alpha = \nu^{\frac{a}{(a,b)}}, \quad \beta = \nu^{\frac{b}{(a,b)}}.$$

Обозначим $\frac{a}{(a,b)} = m, \frac{b}{(a,b)} = n$. Заметим, что $(m, n) = 1$. Осталось положить $\alpha_1 = \alpha$, $\alpha_2 = \beta$, $\nu_1 = \nu$, $a_1(1) = m$, $a_2(1) = n$.

Пусть мы уже построили $M_1, \ldots, M_k$. Если для любого $\xi \in R$ $l(\xi)$ делится на $l(\nu_k)$, то $l = k$ и построение закончено. Пусть существует $\xi \in R$ такое, что $(l(\xi), l(\nu_k)) < l(\nu_k)$. Обозначим $l(\xi) = c$. Тогда

$$\alpha^c, \quad \xi^a \in R^* \setminus \{\lambda\}, \quad l(\alpha^c) = l(\xi^a) = ac.$$

Значит, $\alpha^c = \xi^a$. Но $\alpha = \alpha_1 = \nu_k^{a_1(k)}$, поэтому $\xi^a = \nu_k^{c \cdot a_1(k)}$. По лемме 2 существует $\sigma \in A^* \setminus \{\lambda\}$ такое, что

$$l(\sigma) = (c, l(\nu_k)), \quad \xi = \sigma^{\frac{c}{(c,l(\nu_k))}}, \quad \nu_k = \sigma^{\frac{l(\nu_k)}{(c,l(\nu_k))}}.$$

Обозначим $\frac{c}{(c,l(\nu_k))} = u$, $\frac{l(\nu_k)}{(c,l(\nu_k))} = v$. Получаем, что

$$\alpha_1 = \nu_k^{a_1(k)} = \sigma^{v \cdot a_1(k)}, \quad \alpha_2 = \nu_k^{a_2(k)} = \sigma^{v \cdot a_2(k)}, \quad \ldots,$$
$$\alpha_{k+1} = \nu_k^{a_{k+1}(k)} = \sigma^{v \cdot a_{k+1}(k)}, \quad \xi = \sigma^u.$$

Кроме того,

$$\text{НОД}(v \cdot a_1(k), v \cdot a_2(k), \ldots, v \cdot a_{k+1}(k), u) =$$
$$= (\text{НОД}(v \cdot a_1(k), \quad v \cdot a_2(k), \quad \ldots, \quad , v \cdot a_{k+1}(k)), u) = (v, u) = 1.$$

Осталось положить $\alpha_{k+2} = \xi$, $\nu_{k+1} = \sigma$, $a_1(k+1) = v \cdot a_1(k)$, ,…, $a_{k+1}(k+1) = v \cdot a_{k+1}(k)$, $a_{k+2}(k+1) = u$.

Заметим, что $l(\nu_{k+1}) = l(\sigma) = (c, l(\nu_k)) = (l(\xi), l(\nu_k)) < l(\nu_k)$. Значит, для некоторого l все длины слов из R будут делиться на $l(\nu_l)$ и процесс построения множеств M_i закончится. Возьмем произвольное $\rho \in R$. Обозначим $l(\rho) = d$. Так как

$$\rho^a, \alpha^d \in R^* \setminus \{\lambda\}, \quad l(\rho^a) = l(\alpha^d) = ad,$$

то $\rho^a = \alpha^d = \nu_l^{d\cdot a_1(l)}$. Из

$$a = l(\alpha) = l(\nu_l^{a_1(l)}) = l(\nu_l)\cdot a_1(l)$$

следует, что $\rho^{l(\nu_l)\cdot a_1(l)} = \nu_l^{d\cdot a_1(l)}$, то есть $\rho^{l(\nu_l)} = \nu_l^d$. Осталось заметить, что d делится на $l(\nu_l)$. Отсюда окончательно получаем $\rho = \nu_l^{\frac{d}{l(\nu_l)}}$. Значит, $R \subseteq \{\nu_l\}^*$. Отсюда следует, что и $R^* \subseteq \{\nu_l\}^*$. Но

$$\nu_l^{a_1(l)}, \ldots, \nu_l^{a_{l+1}(l)} \in R \ \text{ и } \ \text{НОД}(a_1(l), \ldots, a_{l+1}(l)) = 1.$$

По лемме 3, примененной для чисел $a_1(l), \ldots, a_{l+1}(l)$ получаем, что существует $n_0 \in \mathbb{N}$ такое, что для всех $t \geqslant n_0$ верно $t \in \{a_1\cdot a_1(l) + \ldots + a_k \cdot a_k(l) \mid a_1, \ldots, a_k \in \mathbb{N}\cup\{0\}\}$. Поэтому

$$\{\nu_l^{n_0}\}\cdot\{\nu_l\}^* \subseteq \{\nu_l^{a_1(l)}, \ldots, \nu_l^{a_{l+1}(l)}\}^*.$$

С другой стороны,

$$\{\nu_l^{a_1(l)}, \ldots, \nu_l^{a_{l+1}(l)}\}^* \subseteq R^*.$$

Получаем, что

$$\{\nu_l^{n_0}\}\cdot\{\nu_l\}^* \subseteq R^* \subseteq \{\nu_l\}^*.$$

Значит,

$$R^* = (R^* \setminus (\{\nu_l^{n_0}\}\cdot\{\nu_l\}^*)) \vee (\{\nu_l^{n_0}\}\cdot\{\nu_l\}^*).$$

При этом,

$$|R^* \setminus (\{\nu_l^{n_0}\}\cdot\{\nu_l\}^*)| \leqslant |\{\nu_l\}^* \setminus (\{\nu_l^{n_0}\}\cdot\{\nu_l\}^*)| < \infty.$$

Осталось положить $R^* \setminus (\{\nu_l^{n_0}\}\cdot\{\nu_l\}^*) = R_1$, $\nu_l^{n_0} = \delta$, $\nu_l = \gamma$.

Утверждение леммы 4 доказано.

Лемма 5. *Пусть $\alpha, \beta, \gamma \in A^* \setminus \{\lambda\}$ и $R = \{\alpha\}^*\cdot\{\beta\}\cdot\{\gamma\}^*$ — 1-тонкое множество. Тогда существуют $\delta, \nu, \mu \in A^* \setminus \{\lambda\}$ такие, что $R = \{\delta\}^*\cdot\{\nu\}^*\cdot\{\mu\}$.*

Доказательство. Будем доказывать утверждение индукцией по $k = l(\beta)$.

База индукции ($k = 1$).

Пусть $l(\beta) = 1$, то есть $\beta = a$ для некоторого $a \in A$. Тогда

$$R = \{\alpha\}^* \cdot \{a\} \cdot \{\gamma\}^*,$$
$$\alpha^{l(\gamma)} \cdot a, a \cdot \gamma^{l(\alpha)} \in S.$$

Но $l(\alpha^{l(\gamma)} \cdot a) = l(a \cdot \gamma^{l(\alpha)}) = l(\alpha) \cdot l(\gamma) + 1$. Значит, $\alpha^{l(\gamma)} \cdot a = a \cdot \gamma^{l(\alpha)}$. Поэтому в алфавите A существует слово γ_1 такое, что $\gamma = \gamma_1 \cdot a$. Отсюда

$$R = \{\alpha\}^* \cdot \{a\} \cdot \{\gamma_1 \cdot a\}^* = \{\alpha\}^* \cdot \{a \cdot \gamma_1\}^* \cdot a.$$

Осталось положить $\alpha = \delta$, $a \cdot \gamma_1 = \nu$, $a = \mu$. Очевидно, что $\beta, \nu, \mu \in A^* \setminus \{\lambda\}$.

Переход индукции ($k \Rightarrow k + 1$).

Пусть $|\beta| = k + 1$. Тогда для некоторых $a \in A$ и $\beta_1 \in A^* \setminus \{\lambda\}$, $l(\beta_1) = k$ верно, что

$$\beta = \beta_1 \cdot a,$$
$$R = \{\alpha\}^* \cdot \{\beta_1 \cdot a\} \cdot \{\gamma\}^*,$$
$$\alpha^{l(\gamma)} \cdot \beta_1 \cdot a, \beta_1 \cdot a \cdot \gamma^{l(\alpha)} \in R.$$

Так как

$$l(\alpha^{l(\gamma)} \cdot \beta_1 \cdot a) = l(\beta_1 \cdot a \cdot \gamma^{l(\alpha)}) = l(\alpha) \cdot l(\gamma) + k + 1,$$

то

$$\alpha^{l(\gamma)} \cdot \beta_1 \cdot a = \beta_1 \cdot a \cdot \gamma^{l(\alpha)}.$$

Поэтому существует $\gamma_1 \in A^*$ такое, что $\gamma = \gamma_1 \cdot a$. Отсюда

$$R = \{\alpha\}^* \cdot \{\beta_1\} \cdot \{a\} \cdot \{\gamma_1 \cdot a\}^* = \{\alpha\}^* \cdot \{\beta_1\} \cdot \{a \cdot \gamma_1\}^* \cdot \{a\}.$$

Пусть

$$S = \{\alpha\}^* \cdot \{\beta_1\} \cdot \{a \cdot \gamma_1\}^*,$$
$$\alpha_1, \alpha_2 \in S,$$
$$l(\alpha_1) = l(\alpha_2).$$

Тогда $\alpha_1 \cdot a, \alpha_2 \cdot a \in R$, $l(\alpha_1 \cdot a) = l(\alpha_2 \cdot a)$. Значит, $\alpha_1 \cdot a = \alpha_2 \cdot a$, то есть $\alpha_1 = \alpha_2$. Поэтому S будет 1-тонким множеством. Значит, по предположению индукции существуют $\delta_1, \nu_1, \mu_1 \in A^* \setminus \{\lambda\}$ такие, что

$$S = \{\delta_1\}^* \cdot \{\nu_1\}^* \cdot \{\mu_1\}.$$

Поэтому

$$R = \{\delta_1\}^* \cdot \{\nu_1\}^* \cdot \{\mu_1\} \cdot \{a\}.$$

Осталось положить $\delta_1 = \delta$, $\nu_1 = \nu$, $\mu_1 \cdot a = \mu$. Очевидно, что $\beta, \nu, \mu \in A^* \setminus \{\lambda\}$.

Утверждение индукции, а с ним и лемма 5 доказана.

Лемма 6. *Пусть $\alpha, \beta \in A^* \setminus \{\lambda\}$ и $R = \{\alpha\}^* \cdot \{\beta\}^*$ — 1-тонкое множество. Тогда существуют $\delta, \gamma \in A^* \setminus \{\lambda\}$ и $R_1 \subseteq A^*$ такие, что $|R_1| < \infty$, $R = R_1 \cup (\{\delta\} \cdot \{\gamma\}^*)$.*

Доказательство. Пусть $l(\alpha) = a$, $l(\beta) = b$. Заметим, что $\alpha^b, \beta^a \in R$, $l(\alpha^b) = l(\beta^a) = ab$. Поэтому $\alpha^b = \beta^a$ и по лемме 2 существует $\nu \in A^* \setminus \{\lambda\}$ такое, что

$$l(\nu) = (a, b),\ \alpha = \nu^{\frac{a}{(a,b)}},\ \beta = \nu^{\frac{b}{(a,b)}}.$$

Обозначим $\frac{a}{(a,b)} = m$, $\frac{b}{(a,b)} = n$. Тогда $(m, n) = 1$ и $R = \{\nu^m\}^* \cdot \{\nu^n\}^*$. Пусть

$$L = \{a_1 \cdot m + a_2 \cdot n \mid a_1, a_2 \in \mathbb{N} \cup \{0\}\}.$$

По лемме 3 существует $n_0 \in \mathbb{N}$ такое, что для всех $n \geqslant n_0$ верно $n \in L$. Значит,

$$\{\nu^{n_0}\} \cdot \{\nu\}^* \subseteq \{\nu^m\}^* \cdot \{\nu^n\}^*.$$

С другой стороны,

$$\{\nu^m\}^* \cdot \{\nu^n\}^* \subseteq \{\nu\}^*.$$

Получаем, что

$$\{\nu^{n_0}\} \cdot \{\nu\}^* \subseteq R \subseteq \{\nu\}^*.$$

Поэтому

$$R = (R \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*)) \vee (\{\nu^{n_0}\} \cdot \{\nu\}^*).$$

При этом,

$$|R \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*)| \leqslant |\{\nu\}^* \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*)| < \infty.$$

Осталось положить

$$R \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*) = R_1,\ \nu^{n_0} = \delta,\ \nu = \gamma.$$

Очевидно, что $\delta, \gamma \in A^* \setminus \{\lambda\}$.

Утверждение леммы 6 доказано.

Лемма 7. *У любого прогрессивного множества есть только один порождающий след.*

Доказательство. Пусть у некоторого прогрессивного множества P, $P \subseteq A^*$ есть два порождающих следа $(\alpha_1, \beta_1, \gamma_1, k_1, m_1)$, $(\alpha_2, \beta_2, \gamma_2, k_2, m_2)$, то есть P представимо с помощью регулярных выражений $\alpha_1 \cdot (\beta_1^{k_1})^* \cdot \beta_1^{m_1} \cdot \gamma_1$, $\alpha_2 \cdot (\beta_2^{k_2})^* \cdot \beta_2^{m_2} \cdot \gamma_2$. Будем по множеству P восстанавливать значения элементов порождающих следов и покажем, что эти следы совпадают.

Если P конечно, то $\beta_1 = \beta_2 = \lambda$, поэтому $\gamma_1 = \gamma_2 = \lambda$, $k_1 = k_2 = 1$, $m_1 = m_2 = 0$. Значит,

$$P = |\alpha_1 \cdot (\beta_1^{k_1})^* \cdot \beta_1^{m_1} \cdot \gamma_1| = |\alpha_1| = \{\alpha_1\}.$$

Аналогично,

$$P = |\alpha_2 \cdot (\beta_2^{k_2})^* \cdot \beta_2^{m_2} \cdot \gamma_2| = |\alpha_2| = \{\alpha_2\}.$$

Значит, $\alpha_1 = \alpha_2$, то есть в этом случае порождающие следы совпадают.

Пусть теперь P бесконечно. Тогда $\beta_1 \neq \lambda$ и $\beta_2 \neq \lambda$. Пусть α — самое маленькое по длине слово из P и β — самое маленькое по длине слово из $P \setminus \{\alpha\}$. Тогда

$$\alpha = \alpha_1 \cdot \beta_1^{m_1} \cdot \gamma_1, \ \beta = \alpha_1 \cdot \beta_1^{k_1} \cdot \beta_1^{m_1} \cdot \gamma_1.$$

Заметим, что оба слова имеют окончания $\beta_1^{m_1} \cdot \gamma_1$. Если $\alpha_1 = \lambda$, то слово α является окончанием слова β. Пусть $\alpha_1 \neq \lambda$. Тогда перед этим окончанием в α расположена последняя буква слова α_1, а в слове β — последняя буква слова β_1. По определению порождающего следа эти буквы различны. Таким образом, если слово α является окончанием слова β, то $\alpha_1 = \lambda$. Если же это не так, то в слове α можно отбросить максимальное по длине общее окончание слов α, β и получить слово α_1. В любом случае, по словам α и β однозначно восстанавливается слово α_1. Значит, $\alpha_1 = \alpha_2$. Но если мы знаем слово α_1, то, откинув его от начала слова α мы получаем слово $\beta_1^{m_1} \cdot \gamma_1$. Если его откинуть теперь от окончания слова β, то мы получим слово $\alpha_1 \cdot \beta_1^{k_1}$. Осталось убрать из его начала слово α_1 и мы получаем слово $\beta_1^{k_1}$. Таким образом, $\beta_1^{k_1} = \beta_2^{k_2}$. По лемме 2 существует $\nu \in A^* \setminus \{\lambda\}$ такое, что

$$l(\nu) = (l(\beta_1), l(\beta_2)), \ \beta_1 = \nu^{\frac{l(\beta_1)}{l(\nu)}}, \ \beta_2 = \nu^{\frac{l(\beta_2)}{l(\nu)}}.$$

Но слова β_1, β_2 неизмельчимы. Значит, $\beta_1 = \nu = \beta_2$. Но $\beta_1^{k_1} = \beta_2^{k_2}$, поэтому $k_1 = k_2$. Далее, нам известно слово $\beta_1^{m_1} \cdot \gamma_1$. По определению порождающего следа слово β_1 не является началом слова γ_1. Поэтому слово γ_1 получается из известного нам слова $\beta_1^{m_1} \cdot \gamma_1$ отбрасыванием из его начала слова β_1 до тех пор, пока это возможно. Количество таких отбрасываний однозначно задает нам значение m_1. Значит, $\gamma_1 = \gamma_2$, $m_1 = m_2$. Таким образом, мы показали, что и в этом случае порождающие следы совпадают.

Утверждение леммы 7 доказано.

Лемма 8. *Пусть $P = \{\alpha_1\} \cdot \{\alpha_2\}^* \cdot \{\alpha_3\}$, где $\alpha_1, \alpha_2, \alpha_3 \in A^*$. Тогда P — прогрессивное множество.*

Доказательство. Будем искать представление P в виде

$$\{\alpha\} \cdot \{\beta^k\}^* \cdot \{\beta^m\} \cdot \{\gamma\},$$

где $(\alpha, \beta, \gamma, k, m)$ — порождающий след.

Если P конечно, то $\alpha_2 = \lambda$. Значит, $P = \{\alpha_1\gamma_1\}$. Осталось положить

$$\alpha = \alpha_1\gamma_1,\ \beta = \gamma = \lambda,\ k = 1,\ m = 0.$$

Пусть P бесконечно. Тогда $\alpha_2 \neq \lambda$. Если $\alpha_1 = \lambda$, то $P = \{\alpha_2\}^* \cdot \{\alpha_3\}$. Рассмотрим множество

$$L = \{\delta \in A^* \setminus \{\lambda\} \mid \delta^k = \alpha_2,\ k \geqslant 2\}.$$

Пусть ν — минимальный по длине элемент множества L. Очевидно, что такое слово неизмельчимо. Заметим далее, что $\nu \neq \lambda$ и существует $r \geqslant 2$ такое, что $\alpha_2 = \nu^r$. Поэтому

$$P = \{\nu^r\}^* \cdot \{\alpha_3\}.$$

Выделяя из начала слова α_3 слова ν до тех пор, пока это возможно, получаем, что $\alpha_3 = \nu^s\alpha_4$. Здесь $s \in \mathbb{N} \vee \{0\}$ и слово ν не является началом слова α_4. Итак,

$$P = \{\nu^r\}^* \cdot \{\nu^s\} \cdot \{\alpha_4\}.$$

Осталось положить $\alpha = \lambda$, $\beta = \nu$, $\gamma = \alpha_4$, $k = r$, $m = s$. Все свойства из определения порождающего следа выполнены.

Пусть теперь $\alpha_1 \neq \lambda$. Обозначим через δ наибольшее общее окончание слов α_1 и α_2. Пусть $\alpha_1 = \mu\delta$, $\alpha_2 = \rho\delta$ и у слов μ, ρ нет одинаковых непустых окончаний. Тогда

$$P = \{\mu\delta\} \cdot \{\rho\delta\}^* \cdot \{\alpha_3\} = \{\mu\} \cdot \{\delta\rho\}^* \cdot \{\delta\alpha_3\}.$$

С оставшейся частью $\{\delta\rho\}^* \cdot \{\delta\alpha_3\}$ поступаем так же, как и в предыдущем случае с $\{\alpha_2\}^* \cdot \{\alpha_3\}$. А именно, находим неизмельчимое слово $\nu \neq \lambda$ такое, что $\delta\rho = \nu^r$, $r \geqslant 2$. Так как слова ν — окончание слова $\delta\rho$, то и у пары μ, ν нет одинаковых окончаний. Теперь выделим из начала слова $\delta\alpha_3$ все слова ν. Имеем

$$\delta\alpha_3 = \nu^s\alpha_4,$$

где $s \in \mathbb{N} \vee \{0\}$ и слово ν не является началом слова α_4. Итак,

$$P = \{\mu\} \cdot \{\nu^r\}^* \cdot \{\nu^s\} \cdot \{\alpha_4\}.$$

Осталось положить $\alpha = \mu$, $\beta = \nu$, $\gamma = \alpha_4$, $k = r$, $m = s$. Все свойства из определения порождающего следа выполнены.

Утверждение леммы 8 доказано.

Лемма 9. *Пусть P_1, P_2 — бесконечные спектрально зависимые прогрессивные множества и $P = P_1 \cup P_2$ — 1-тонкое множество. Тогда основания множеств P_1 и P_2 совпадают.*

Доказательство. Для начала заметим, что по лемме 7 у любого прогрессивного множества есть только один порождающий след. Поэтому и основания таких множеств определены однозначно.

Пусть P_1 имеет порождающий след $(\alpha_1, \beta_1, \gamma_1, k_1, m_1)$, P_2 имеет порождающий след $(\alpha_2, \beta_2, \gamma_2, k_2, m_2)$. Это значит, что P_1 и P_2 представимы с помощью регулярных выражений

$$\alpha_1 \cdot (\beta_1^{k_1})^* \cdot \beta_1^{m_1} \cdot \gamma_1,\ \alpha_2 \cdot (\beta_2^{k_2})^* \cdot \beta_2^{m_2} \cdot \gamma_2$$

соответственно. Из бесконечности P_1 и P_2 следует, что $\beta_1 \neq \lambda$, $\beta_2 \neq \lambda$. Так как P_1 и P_2 спектрально зависимы, то

$$Sp(P_1) \cap Sp(P_2) \neq \varnothing.$$

Поэтому существуют $\nu_1 \in P_1, \nu_2 \in P_2$ такие, что $|\nu_1| = |\nu_2|$. Поскольку $\nu_1, \nu_2 \in P$, то $\nu_1 = \nu_2$. Из

$$P_1 = |\alpha_1 \cdot (\beta_1^{k_1})^* \cdot \beta_1^{m_1} \cdot \gamma_1|, \quad P_2 = |\alpha_2 \cdot (\beta_2^{k_2})^* \cdot \beta_2^{m_2} \cdot \gamma_2|$$

получаем, что существуют $a_1, a_2 \in \mathbb{N} \vee \{0\}$, для которых

$$\alpha_1 \cdot \beta_1^{a_1 \cdot k_1 + m_1} \cdot \gamma_1 = \nu_1 = \nu 2 = \alpha_2 \cdot \beta_2^{a_2 \cdot k_2 + m_2} \cdot \gamma_2.$$

Докажем, что $\alpha_1 = \alpha_2$. Пусть это не так и $l_1 = l(\alpha_1)$, $l_2 = l(\alpha_2)$. Тогда $l_1 \neq l_2$. Без ограничения общности можно считать, что $l_1 < l_2$. Обозначим через μ_1, μ_2 слова

$$\alpha_1 \cdot \beta_1^{a_1 k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1} \cdot \gamma_1, \ \alpha_2 \cdot \beta_2^{a_2 k_2 + (l_2 + l(\beta_2))l(\beta_1) + m_2} \cdot \gamma_2$$

соответственно. Заметим, что $\mu_1, \mu_2 \in P$ и

$$l(\mu_1) = l(\nu_1) + (l_2 + l(\beta_2))l(\beta_1)l(\beta_2) = l(\nu_2) + (l_2 + l(\beta_2))l(\beta_1)l(\beta_2) = l(\mu_2).$$

Так как P — тонкое множество, то $\mu_1 = \mu_2$. Пусть $a = \mu_2(l_2 - 1, l_2)$. Это последняя буква слова α_2. С другой стороны, $a = \mu_1(l_2 - 1, l_2)$. Так как

$$l_1 < l_2 \leqslant (l_2 + l(\beta_2))l(\beta_2) \leqslant l_1 + (a_1 k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1)l(\beta_1),$$

то a является частью слова $\beta_1^{a_1 k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1}$. Далее,

$$l_1 < l_2 + l(\beta_1)l(\beta_2) \leqslant (l_2 + l(\beta_2))l(\beta_1)l(\beta_2) \leqslant \\ \leqslant l_1 + (a_1 k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1)l(\beta_1).$$

Отсюда следует, что $\mu_1(l_2 - 1, l_2 + l(\beta_1)l(\beta_2))$ также является частью слова $\beta_1^{a_1 k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1}$. Значит,

$$\mu_1(l_2 + l(\beta_1)l(\beta_2) - 1, l_2 + l(\beta_1)l(\beta_2)) = \\ = \mu_1(l_2 + l(\beta_1)(l(\beta_2) - 1) - 1, l_2 + l(\beta_1)(l(\beta_2) - 1)) = \dots \\ \dots = \mu_1(l_2 - 1, l_2) = a.$$

С другой стороны,

$$l_2 < l_2 + l(\beta_1)l(\beta_2) \leqslant l_2 + (a_2 k_2 + (l_2 + l(\beta_2))l(\beta_1) + m_2)l(\beta_2).$$

Значит,

$$\mu_2(l_2 + l(\beta_1)l(\beta_2) - 1, l_2 + l(\beta_1)l(\beta_2)) = \\ = \mu_2(l_2 + (l(\beta_1) - 1)l(\beta_2) - 1, l_2 + (l(\beta_1) - 1)l(\beta_2)) = \dots \\ \dots = \mu_2(l_2 + l(\beta_2) - 1, l_2 + l(\beta_2)).$$

Но это последняя буква слова $\alpha_2\beta_2$. Замечаем, что

$$a = \mu_1(l_2 + l(\beta_1)l(\beta_2) - 1, l_2 + l(\beta_1)l(\beta_2)) = \\ = \mu_2(l_2 + l(\beta_1)l(\beta_2) - 1, l_2 + l(\beta_1)l(\beta_2)).$$

Поэтому последние буквы слов α_2 и β_2 совпадают. Это противоречит определению порождающего следа для $(\alpha_2, \beta_2, \gamma_2, k_2, m_2)$. Отсюда заключаем, что $\alpha_1 = \alpha_2$.

Так как $\alpha_1 = \alpha_2$ и $\mu_1 = \mu_2$, то

$$\beta_1^{a_1k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1} \cdot \gamma_1 = \beta_2^{a_2k_2 + (l_2 + l(\beta_2))l(\beta_1) + m_2} \cdot \gamma_2.$$

Обозначим это слово через ρ. Так как

$$l(\rho) = (a_1k_1 + (l_2 + l(\beta_2))l(\beta_2) + m_1)l(\beta_1) + l(\gamma_1) \geqslant l(\beta_1)l(\beta_2),$$

то у ρ есть начало длины $l(\beta_1)l(\beta_2)$. Оно равно одновременно $\beta_1^{l(\beta_2)}$ и $\beta_2^{l(\beta_1)}$. По лемме 2 существует $\delta \in A^* \setminus \{\lambda\}$ такое, что

$$l(\delta) = (l(\beta_1), l(\beta_2)),\ \beta_1 = \delta^{\frac{l(\beta_1)}{l(\delta)}},\ \beta_2 = \delta^{\frac{l(\beta_2)}{l(\delta)}}.$$

Но, по определению порождающего следа, слова β_1, β_2 неизмельчимы. Значит,

$$\beta_1 = \nu = \beta_2.$$

Теперь пользуемся определением порождающего следа и замечаем, что слово γ_1 получается из слова ρ отбрасыванием из его начала слова β_1 до тех пор, пока это возможно. Аналогично, слово γ_2 получается из слова ρ отбрасыванием из его начала слова β_2 до тех пор, пока это возможно. Так как $\beta_1 = \beta_2$, то $\gamma_1 = \gamma_2$. Мы доказали, что

$$\alpha_1 = \alpha_2,\quad \beta_1 = \beta_2,\quad \gamma_1 = \gamma_2.$$

Значит, основания множеств P_1 и P_2 совпадают.

Утверждение леммы 9 доказано.

Лемма 10. *Пусть P_1, P_2 — спектрально независимые 1-тонкие множества. Тогда $P = P_1 \cup P_2$ тоже будет 1-тонким множеством.*

Доказательство. Так как P_1 и P_2 спектрально независимы, то

$$Sp(P_1 \cap P_2) = Sp(P_1) \cap Sp(P_2) = \varnothing.$$

Значит, $P_1 \cap P_2 = \varnothing$. Пусть α, β — произвольные слова из P и $l(\alpha) = l(\beta)$. Из спектральной независимости P_1 и P_2 следует, что они оба или из P_1, или из P_2. Тогда, по определению тонких множеств, $\alpha = \beta$. Поэтому P — тонкое множество.

Утверждение леммы 10 доказано.

Лемма 11. *Пусть $R, R_1, R_2 \subseteq A^*$, $R = R_1 \cup R_2$ и $R \in T$. Тогда $R_1, R_2 \in T$.*

Доказательство. Так как $R \in T$, то для некоторого $s \in \mathbb{N}$ верно, что $R \in T_s$. Пусть

$$\alpha_1, \alpha_2, \ldots, \alpha_{s+1} \in R_1, \quad l(\alpha_1) = l(\alpha_2) = \ldots = l(\alpha_{s+1}).$$

Тогда

$$\alpha_1, \alpha_2, \ldots, \alpha_{s+1} \in R.$$

Так как R — s-тонкое множество, то существуют $i, j \in \mathbb{N}$ такие, что

$$\alpha_i = \alpha_j, \quad 1 \leqslant i < j \leqslant s+1.$$

Таким образом, в R_1 нет s различных слов одинаковой длины. Обозначим через s_1 максимальное количество несовпадающих слов одинаковой длины из множества R_1. Мы показали, что такое s_1 существует и не превосходит s. Осталось заметить, что, в силу максимальности выбора числа s_1, в R_1 нет $s_1 + 1$ несовпадающих слов одинаковой длины. Поэтому R_1 — s_1-тонкое множество.

Аналогично показывается, что существует $s_2 \leqslant s$ такое, что R_2 — s_2-тонкое множество.

Утверждение леммы 11 доказано.

Лемма 12. *Пусть $R, R_1, R_2 \subseteq A^*$, $R = R_1 \cdot R_2$ и R — s-тонкое множество. Тогда для некоторых $s_1, s_2 \leqslant s$ верно, что R_1 — s_1-тонкое множество и R_2 — s_2-тонкое множество.*

Доказательство. Пусть

$$\alpha_1, \alpha_2, \ldots, \alpha_{s+1} \in R_1, \quad \beta \in R_2, \quad l(\alpha_1) = l(\alpha_2) = \ldots = l(\alpha_{s+1}).$$

Тогда

$$\alpha_1 \cdot \beta, \alpha_2 \cdot \beta, \quad \ldots \quad , \alpha_{s+1} \cdot \beta \in R.$$

Так как R — s-тонкое множество, то существуют $i, j \in \mathbb{N}$ такие, что

$$\alpha_i \cdot \beta = \alpha_j \cdot \beta, \quad 1 \leqslant i < j \leqslant s+1.$$

То есть, $\alpha_i = \alpha_j$. Таким образом, в R_1 нет s различных слов одинаковой длины. Обозначим через s_1 максимальное количество несовпадающих слов из множества R_1, имеющих одинаковую длину. Мы показали, что такое s_1 существует и не превосходит s. Осталось заметить, что, в силу максимальности выбора числа s_1, в R_1 нет $s_1 + 1$ различных слов одинаковой длины. Поэтому R_1 — s_1-тонкое множество.

Аналогично доказывается, что существует $s_2 \leqslant s$ такое, что R_2 — s_2-тонкое множество.

Утверждение леммы 12 доказано.

Лемма 13. *Пусть $R \subseteq A^*$ и R^* — s-тонкое множество. Тогда существуют $\delta, \gamma \in A^*$ и $R_1 \subseteq A^*$ такие, что $|R_1| < \infty$, $R^* = R_1 \cup (\{\delta\} \cdot \{\gamma\}^*)$.*

Доказательство. Пусть $|R| = 1$, то есть $R = \{\mu\}$ для некоторого $\mu \in A^*$. Тогда полагаем $R_1 = \varnothing$, $\delta = \lambda$, $\gamma = \mu$.

Пусть $|R| > 1$. Так как $\{R \setminus \{\lambda\}\}^* = R^*$, то можно считать, что $\lambda \notin R$. Будем строить последовательность множеств $M_1 \subset M_2 \subset \ldots \subset M_l \subset R$ такую, что:
(1) $M_i = \{\alpha_1, \alpha_2, \ldots, \alpha_{i+1}\}$ для всех $l \geqslant i \geqslant 1$;
(2) $\alpha_i = \nu_j^{a_i(j)}$ для всех $l \geqslant j \geqslant 1$ и $j + 1 \geqslant i \geqslant 1$, где $\nu_j \in A^* \setminus \{\lambda\}$ и $a_i(j) \in \mathbb{N}$;
(3) НОД$(a_1(j), \ldots, a_{j+1}(j)) = 1$ для всех $l \geqslant j \geqslant 1$.

Построим M_1. Так как $|R| > 1$, то существуют $\alpha, \beta \in R$, $\alpha \neq \beta$. Обозначим $l(\alpha) = a$, $l(\beta) = b$. Тогда для всех $0 \leqslant i \leqslant s$ верно

$$\alpha^{bi} \cdot \beta^{a(s-i)} \in R^*,$$
$$l(\alpha^{bi} \cdot \beta^{a(s-i)}) = abs.$$

Значит существуют t, u, $0 \leqslant t < u \leqslant s$ такие, что

$$\alpha^{bt} \cdot \beta^{a(s-t)} = \alpha^{bu} \cdot \beta^{a(s-u)}.$$

Отсюда получаем, что

$$\beta^{a(u-t)} = \alpha^{b(u-t)}, \quad \alpha^b = \beta^a.$$

По лемме 2 существует $\nu \in A^* \setminus \{\lambda\}$ такое, что

$$l(\nu) = (a, b), \quad \alpha = \nu^{\frac{a}{(a,b)}}, \quad \beta = \nu^{\frac{b}{(a,b)}}.$$

Обозначим $\frac{a}{(a,b)} = m, \frac{b}{(a,b)} = n$. Заметим, что $(m, n) = 1$. Осталось положить $\alpha_1 = \alpha$, $\alpha_2 = \beta$, $\nu_1 = \nu$, $a_1(1) = m$, $a_2(1) = n$.

Пусть мы уже построили $M_1, \ldots, M_k$. Если для любого $\xi \in R$ $l(\xi)$ делится на $l(\nu_k)$, то $l = k$ и построение закончено. Пусть существует $\xi \in R$ такое, что

$$(l(\xi), l(\nu_k)) < l(\nu_k).$$

Обозначим $l(\xi) = c$. Тогда для всех $0 \leqslant i \leqslant s$ верно

$$\alpha^{ci} \cdot \xi^{a(s-i)} \in R^* \setminus \{\lambda\},$$
$$l(\alpha^{ci} \cdot \xi^{a(s-i)}) = acs.$$

Значит, существуют t, u, $0 \leqslant t < u \leqslant s$ такие, что

$$\alpha^{ct} \cdot \xi^{a(s-t)} = \alpha^{cu} \cdot \xi^{a(s-u)}.$$

Отсюда получаем

$$\xi^{a(u-t)} = \alpha^{c(u-t)}, \quad \alpha^c = \xi^a.$$

Но $\alpha = \alpha_1 = \nu_k^{a_1(k)}$, поэтому

$$\xi^a = \nu_k^{c \cdot a_1(k)}.$$

По лемме 2 существует $\sigma \in A^* \setminus \{\lambda\}$ такое, что

$$l(\sigma) = (c, l(\nu_k)), \quad \xi = \sigma^{\frac{c}{(c, l(\nu_k))}}, \quad \nu_k = \sigma^{\frac{l(\nu_k)}{(c, l(\nu_k))}}.$$

Обозначим $\frac{c}{(c,l(\nu_k))} = u$, $\frac{l(\nu_k)}{(c,l(\nu_k))} = v$. Получаем

$$\alpha_1 = \nu_k^{a_1(k)} = \sigma^{v\cdot a_1(k)},$$
$$\alpha_2 = \nu_k^{a_2(k)} = \sigma^{v\cdot a_2(k)}, \ldots, \alpha_{k+1} = \nu_k^{a_{k+1}(k)} = \sigma^{v\cdot a_{k+1}(k)},$$
$$\xi = \sigma^u.$$

Кроме того,

$$\text{НОД}(v\cdot a_1(k), v\cdot a_2(k), \ldots, v\cdot a_{k+1}(k), u) =$$
$$= (\text{НОД}(v\cdot a_1(k), v\cdot a_2(k), \ldots, v\cdot a_{k+1}(k)), u) = (v,u) = 1.$$

Осталось положить

$$a_1(k+1) = v\cdot a_1(k), \ldots, a_{k+1}(k+1) = v\cdot a_{k+1}(k),$$
$$a_{k+2}(k+1) = u, \alpha_{k+2} = \xi,\ \nu_{k+1} = \sigma.$$

Заметим, что

$$l(\nu_{k+1}) = l(\sigma) = (c, l(\nu_k)) = (l(\xi), l(\nu_k)) < l(\nu_k).$$

Значит, для некоторого l все длины слов из R будут делиться на $l(\nu_l)$ и процесс построения множеств M_i закончится. Возьмем произвольное $\rho \in R$. Обозначим $l(\rho) = d$. Тогда для всех $0 \leqslant i \leqslant s$ верно

$$\rho^{ai}\cdot\alpha^{d(s-i)} \in R^* \setminus \{\lambda\},$$
$$l(\rho^{ai}\cdot\alpha^{d(s-i)}) = ads.$$

Значит, существуют t, u, $0 \leqslant t < u \leqslant s$ такие, что

$$\rho^{at}\cdot\alpha^{d(s-t)} = \rho^{au}\cdot\alpha^{d(s-u)}.$$

Отсюда получаем, что

$$\alpha^{d(u-t)} = \rho^{a(u-t)},\quad \rho^a = \alpha^d = \nu_l^{d\cdot a_1(l)}.$$

Так как

$$a = l(\alpha) = l(\nu_l^{a_1(l)}) = l(\nu_l)\cdot a_1(l),$$

то

$$\rho^{l(\nu_l)\cdot a_1(l)} = \nu_l^{d\cdot a_1(l)},\quad \rho^{l(\nu_l)} = \nu_l^d.$$

Осталось заметить, что d делится на $l(\nu_l)$. Отсюда окончательно получаем

$$\rho = \nu_l^{\frac{d}{l(\nu_l)}}.$$

Значит, $R \subseteq \{\nu_l\}^*$. Отсюда следует, что и $R^* \subseteq \{\nu_l\}^*$. Но $\nu_l^{a_1(l)}, \ldots, \nu_l^{a_{l+1}(l)} \in R$ и

$$\text{НОД}(a_1(l), \ldots, a_{l+1}(l)) = 1.$$

По лемме 3, примененной для чисел $a_1(l), \ldots, a_{l+1}(l)$, верно, что существует $n_0 \in \mathbb{N}$ такое, что для всех $t \geqslant n_0$ верно $t \in \{a_1 \cdot a_1(l) + \ldots + a_k \cdot a_k(l) | a_1, \ldots, a_k \in \mathbb{N} \cup \{0\}\}$. Значит,

$$\{\nu_l^{n_0}\} \cdot \{\nu_l\}^* \subseteq \{\nu_l^{a_1(l)}, \ldots, \nu_l^{a_{l+1}(l)}\}^*.$$

С другой стороны,

$$\{\nu_l^{a_1(l)}, \ldots, \nu_l^{a_{l+1}(l)}\}^* \subseteq R^*.$$

Получаем, что

$$\{\nu_l^{n_0}\} \cdot \{\nu_l\}^* \subseteq R^* \subseteq \{\nu_l\}^*.$$

То есть

$$R^* = (R^* \setminus (\{\nu_l^{n_0}\} \cdot \{\nu_l\}^*)) \vee (\{\nu_l^{n_0}\} \cdot \{\nu_l\}^*).$$

При этом,

$$|R^* \setminus (\{\nu_l^{n_0}\} \cdot \{\nu_l\}^*)| \leqslant |\{\nu_l\}^* \setminus (\{\nu_l^{n_0}\} \cdot \{\nu_l\}^*)| < \infty.$$

Осталось положить

$$R^* \setminus (\{\nu_l^{n_0}\} \cdot \{\nu_l\}^*) = R_1,\ \nu_l^{n_0} = \delta,\ \nu_l = \gamma.$$

Утверждение леммы 13 доказано.

Лемма 14. *Пусть $R = \{\alpha\}^* \cdot \{\beta\} \cdot \{\gamma\}^*$ для некоторых непустых слов $\alpha, \beta, \gamma \in A^*$ и R — s-тонкое множество. Тогда существуют $\delta, \nu, \mu \in A^* \setminus \{\lambda\}$ такие, что $R = \{\delta\}^* \cdot \{\nu\}^* \cdot \{\mu\}$.*

Доказательство. Будем доказывать утверждение индукцией по $k = l(\beta)$.

База индукции ($k = 1$).

Пусть $l(\beta) = 1$, то есть $\beta = a$ для некоторого $a \in A$. Тогда

$$R = \{\alpha\}^* \cdot \{a\} \cdot \{\gamma\}^*.$$

Для всех $0 \leqslant i \leqslant s$ верно

$$\alpha^{l(\gamma)i} \cdot a \cdot \gamma^{l(\alpha)(s-i)} \in R,$$
$$l(\alpha^{l(\gamma)i} \cdot a \cdot \gamma^{l(\alpha)(s-i)}) = l(\alpha) \cdot l(\gamma) \cdot s + 1.$$

Значит, существуют t, u такие, что

$$0 \leqslant t < u \leqslant s,$$
$$\alpha^{l(\gamma)t} \cdot a \cdot \gamma^{l(\alpha)(s-t)} = \alpha^{l(\gamma)u} \cdot a \cdot \gamma^{l(\alpha)(s-u)}.$$

Отсюда заключаем, что

$$a \cdot \gamma^{l(\alpha)(u-t)} = \alpha^{l(\gamma)(u-t)} \cdot a.$$

Поэтому существует $\gamma_1 \in A^*$ такое, что $\gamma = \gamma_1 \cdot a$. Отсюда

$$R = \{\alpha\}^* \cdot \{a\} \cdot \{\gamma_1 \cdot a\}^* = \{\alpha\}^* \cdot \{a \cdot \gamma_1\}^* \cdot \{a\}.$$

Осталось положить $\alpha = \delta$, $a \cdot \gamma_1 = \nu$, $a = \mu$. Очевидно, что $\beta, \nu, \mu \in A^* \setminus \{\lambda\}$.

Переход индукции ($k \Rightarrow k + 1$).

Пусть $l(\beta) = k + 1$, то есть $\beta = \beta_1 \cdot a$ для некоторых $a \in A$ и $\beta_1 \in A^* \setminus \{\lambda\}$, $l(\beta_1) = k$. Тогда

$$R = \{\alpha\}^* \cdot \{\beta_1 \cdot a\} \cdot \{\gamma\}^*,$$
$$\alpha^{l(\gamma)i} \cdot \beta_1 \cdot a \cdot \gamma^{l(\alpha)(s-i)} \in R,\ 0 \leqslant i \leqslant s,$$
$$l(\alpha^{l(\gamma)i} \cdot \beta_1 \cdot a \cdot \gamma^{l(\alpha)(s-i)}) = l(\alpha) \cdot l(\gamma) \cdot s + k + 1.$$

Значит, существуют t, u такие, что

$$\alpha^{l(\gamma)t} \cdot \beta_1 \cdot a \cdot \gamma^{l(\alpha)(s-t)} = \alpha^{l(\gamma)u} \cdot \beta_1 \cdot a \cdot \gamma^{l(\alpha)(s-u)},\quad 0 \leqslant t < u \leqslant s.$$

Отсюда заключаем, что

$$\beta_1 \cdot a \cdot \gamma^{l(\alpha)(u-t)} = \alpha^{l(\gamma)(u-t)} \cdot \beta_1 \cdot a.$$

Поэтому существует $\gamma_1 \in A^*$ такое, что $\gamma = \gamma_1 \cdot a$. Отсюда

$$R = \{\alpha\}^* \cdot \{\beta_1\} \cdot \{a\} \cdot \{\gamma_1 \cdot a\}^* = \{\alpha\}^* \cdot \{\beta_1\} \cdot \{a \cdot \gamma_1\}^* \cdot \{a\}.$$

Обозначим через S множество $\{\alpha\}^* \cdot \{\beta_1\} \cdot \{a \cdot \gamma_1\}^*$. Тогда $R = S \cdot \{a\}$. Пусть

$$\alpha_1, \alpha_2, \ldots, \alpha_{s+1} \in, \quad l(\alpha_1) = l(\alpha_2) = \ldots l(\alpha_{s+1}).$$

Тогда

$$\alpha_1 \cdot a, \alpha_2 \cdot a, \ldots, \alpha_{s+1} \cdot a \in R, \quad l(\alpha_1 \cdot a) = l(\alpha_2 \cdot a) = \ldots l(\alpha_{s+1} \cdot a).$$

Так как R — s-тонкое множество, то существуют t, u такие, что

$$1 \leqslant t < u \leqslant s+1, \quad \alpha_t \cdot a = \alpha_u \cdot a.$$

Поэтому $\alpha_t = \alpha_u$.

С другой стороны, так как R — s-тонкое множество, то существует s различных слов $\lambda_1, \lambda_2, \ldots, \lambda_s \in R$ таких, что

$$l(\lambda_1) = l(\lambda_2) = \ldots = l(\lambda_s) = p$$

для некоторого $p \in \mathbb{N}$. Вспомним, что $R = S \cdot \{a\}$. Значит,

$$F_1(\lambda_1) = F_1(\lambda_2) = \ldots = F_1(\lambda_s) = a \quad \text{и} \quad p > 1.$$

Обозначим $S_{p-1}(\lambda_1) = \delta_1$, $S_{p-1}(\lambda_2) = \delta_2$, $\ldots S_{p-1}(\lambda_s) = \delta_s$. Тогда

$$\lambda_1 = \delta_1 \cdot a, \ \lambda_2 = \delta_2 \cdot a, \ \ldots, \lambda_s = \delta_s \cdot a, \delta_1, \delta_2, \ldots, \delta_s \in S.$$

Так как слова $\lambda_1, \lambda_2, \ldots, \lambda_s$ попарно различны, то и слова $\delta_1, \delta_2, \ldots, \delta_s$ попарно различны. Но $l(\delta_1) = l(\delta_2) = \ldots = l(\delta_s) = p - 1$. Получили s различных слов одинаковой длины из множества S.

Объединяя полученные результаты, заключаем, что S — s-тонкое множество. По предположению индукции, существуют $\delta_1, \nu_1, \mu_1 \in A^* \setminus \{\lambda\}$ такие, что

$$S = \{\delta_1\}^* \cdot \{\nu_1\}^* \cdot \{\mu_1\}.$$

Поэтому

$$R = \{\delta_1\}^* \cdot \{\nu_1\}^* \cdot \{\mu_1\} \cdot \{a\}.$$

Осталось положить $\delta_1 = \delta$, $\nu_1 = \nu$, $\mu_1 \cdot a = \mu$. Очевидно, что $\beta, \nu, \mu \in A^* \setminus \{\lambda\}$.

Утверждение индукции, а с ним и лемма 14 доказана.

Лемма 15. *Пусть $\alpha, \beta \in A^* \setminus \{\lambda\}$ и $R = \{\alpha\}^* \cdot \{\beta\}^*$ — s-тонкое множество. Тогда существуют $\delta, \gamma \in A^* \setminus \{\lambda\}$ и $R_1 \subseteq A^*$ такие, что $|R_1| < \infty$, $R = R_1 \vee (\{\delta\} \cdot \{\gamma\}^*)$.*

Доказательство. Пусть $|\alpha| = a$, $|\beta| = b$. Тогда для всех $0 \leqslant i \leqslant s$ верно

$$\alpha^{bi} \cdot \beta^{a(s-i)} \in R^*, \quad l(\alpha^{bi} \cdot \beta^{a(s-i)}) = abs.$$

Значит существуют t, u, $0 \leqslant t < u \leqslant s$ такие, что

$$\alpha^{bt} \cdot \beta^{a(s-t)} = \alpha^{bu} \cdot \beta^{a(s-u)}.$$

Отсюда получаем, что

$$\beta^{a(u-t)} = \alpha^{b(u-t)}, \ \alpha^b = \beta^a.$$

По лемме 2 существует $\nu \in A^* \setminus \{\lambda\}$ такое, что

$$|\nu| = (a, b), \quad \alpha = \nu^{\frac{a}{(a,b)}}, \quad \beta = \nu^{\frac{b}{(a,b)}}.$$

Обозначим $\frac{a}{(a,b)} = m$, $\frac{b}{(a,b)} = n$. Тогда $(m, n) = 1$ и

$$R = \{\nu^m\}^* \cdot \{\nu^n\}^*.$$

Пусть

$$H = \{a_1 \cdot m + a_2 \cdot n \mid a_1, a_2 \in \mathbb{N} \cup \{0\}\}.$$

По лемме 3 существует $n_0 \in \mathbb{N}$ такое, что для всех $n \geqslant n_0$ верно $n \in H$. Значит,

$$\{\nu^{n_0}\} \cdot \{\nu\}^* \subseteq \{\nu^m\}^* \cdot \{\nu^n\}^*.$$

С другой стороны,

$$\{\nu^m\}^* \cdot \{\nu^n\}^* \subseteq \{\nu\}^*.$$

Получаем

$$\{\nu^{n_0}\} \cdot \{\nu\}^* \subseteq R \subseteq \{\nu\}^*.$$

То есть

$$R = (R \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*)) \cup (\{\nu^{n_0}\} \cdot \{\nu\}^*).$$

При этом,

$$|R \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*)| \leqslant |\{\nu\}^* \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*)| < \infty.$$

Осталось положить $R \setminus (\{\nu^{n_0}\} \cdot \{\nu\}^*) = R_1$, $\nu^{n_0} = \delta$, $\nu = \gamma$. Очевидно, что $\delta, \gamma \in A^* \setminus \{\lambda\}$.

Утверждение леммы 15 доказано.

Лемма 16. *Пусть A — конечный алфавит. Любое конечное объединение прогрессивных множеств в алфавите A представимо в виде конечного объединения непересекающихся прогрессивных множеств в алфавите A.*

Доказательство. Пусть для некоторого $k \in \mathbb{N}$ верно, что $R = \bigcup\limits_{i=1}^{k} R_i$ и все R_i — прогрессивные множества. Тогда

$$R = R_1 \bigsqcup (R_2 \setminus R_1) \bigsqcup ((R_3 \setminus R_2) \setminus R_1) \bigsqcup \ldots \bigsqcup ((\ldots (R_k \setminus R_{k-1}) \setminus \ldots \setminus R_1).$$

Осталось показать, что разность двух прогрессивных множеств тоже будет прогрессивным множеством. Пусть

$$R_1 = \{\alpha_1\} \cdot \{\beta_1\}^* \cdot \{\gamma_1\}, \quad R_2 = \{\alpha_2\} \cdot \{\beta_2\}^* \cdot \{\gamma_2\}$$

для некоторых

$$\alpha_1, \alpha_2, \beta_1, \beta_2, \gamma_1, \gamma_2 \in A^*.$$

Либо $R_1 \bigcap R_2 = \varnothing$, либо для некоторого $s \in \mathbb{N}$ верно

$$R_1 \bigcap R_2 = \{\alpha_1\} \cdot \{\beta_1^s\}^* \cdot \{\gamma_1\}.$$

В этом случае получаем

$$R_1 \setminus R_2 = \bigsqcup_{i=1}^{s-1} \{\alpha_1 \cdot \beta_1^i\} \cdot \{\beta_1^s\}^* \cdot \{\gamma_1\}.$$

Осталось заметить, что при всех $1 \leqslant i \leqslant s-1$ верно, что множество

$$\{\alpha_1 \cdot \beta_1^i\} \cdot \{\beta_1^s\}^* \cdot \{\gamma_1\}$$

является прогрессивным. Утверждение леммы 16 доказано.

3. Доказательство основных утверждений

Доказательство теоремы 1.

Из леммы 10 следует, что любое конечное объединение спектрально независимых в совокупности 1-тонких множеств является 1-тонким множеством. Осталось доказать, что любое общепрогрессивное множество является 1-тонким. Пусть P — общепрогрессивное множество. Оно является конечным объединением прогрессивных множеств с одинаковым основанием, то есть

$$P = \bigcup_{i=1}^{k} |\alpha \cdot (\beta^{k_i})^* \cdot \beta^{m_i} \cdot \gamma|$$

для некоторых

$$\alpha, \beta, \gamma \in A^*,\ k \in \mathbb{N},\ m \in \mathbb{N} \cup \{0\}.$$

Поэтому

$$P \subseteq |\alpha \cdot (\beta)^* \cdot \gamma|.$$

Значит, P — 1-тонкое множество.

Утверждение теоремы 1 доказано.

Доказательство теоремы 2.

Пусть P — 1-тонкое множество. Оно регулярно. Применив лемму 1, получаем, что для некоторых чисел $k, s(1), \ldots, s(k) \in \mathbb{N}$, слов $\alpha_{1,1}, \ldots, \alpha_{k,s(k)} \in A^*$ и регулярных выражений $\mathfrak{P}_{1,1}, \ldots, \mathfrak{P}_{k,s(k)-1}$ верно

$$P = \bigcup_{i=1}^{k} |\alpha_{i,1} \cdot (\mathfrak{P}_{i,1})^* \cdot \alpha_{i,2} \cdot \ldots \cdot \alpha_{i,s(i)-1} \cdot (\mathfrak{P}_{i,s(i)-1})^* \cdot \alpha_{i,s(i)}|.$$

Поэтому

$$P = \bigcup_{i=1}^{k} \{\alpha_{i,1}\} \cdot |\mathfrak{P}_{i,1}|^* \cdot \{\alpha_{i,2}\} \cdot \ldots \cdot \{\alpha_{i,s(i)-1}\} \cdot |\mathfrak{P}_{i,s(i)-1}|^* \cdot \{\alpha_{i,s(i)}\}. \quad (1)$$

Так как P — 1-тонкое множество, то все $|\mathfrak{P}_{i,j}|^*$ — 1-тонкие множества. Из леммы 4 следует, что

$$|\mathfrak{P}_{i,j}|^* = R_{i,j} \cup (\{\delta_{i,j}\} \cdot \{\gamma_{i,j}\}^*) \quad (2)$$

для некоторых конечных $R_{i,j} \subseteq A^*$ и $\delta_{i,j}, \gamma_{i,j} \in A^*$. Для любых $\alpha, \beta, \gamma \in A^*$ верно, что

$$(\{\alpha\} \cup \{\beta\}) \cdot \{\gamma\} = \{\alpha\gamma\} \cup \{\beta\gamma\}, \quad \{\gamma\} \cdot (\{\alpha\} \cup \{\beta\}) = \{\gamma\alpha\} \cup \{\gamma\beta\}. \quad (3)$$

Подставляя (2) в (1) и применяя (3), получаем:

$$P = \bigcup_{i=1}^{l} \{\beta_{i,1}\} \cdot \{\beta_{i,2}\}^* \cdot \{\beta_{i,3}\} \cdot \ldots \cdot \{\beta_{i,t(i)-1}\} \cdot \{\beta_{i,t(i)}\}^* \cdot \{\beta_{i,t(i)+1}\} \quad (4)$$

для некоторых $l, t(1), \ldots, t(l) \in \mathbb{N}$, $\beta_{1,1}, \ldots, \beta_{l,t(l)+1} \in A^*$. Применяя к (4) лемму 5 получаем, что:

$$P = \bigcup_{i=1}^{m} \{\gamma_{i,1}\} \cdot \{\gamma_{i,2}\}^* \cdot \{\gamma_{i,3}\}^* \cdot \ldots \cdot \{\gamma_{i,u(i)}\}^* \cdot \{\gamma_{i,u(i)+1}\} \quad (5)$$

для некоторых $m, u(1), \ldots, u(m) \in \mathbb{N}$, $\gamma_{1,1}, \ldots, \gamma_{m,u(m)+1} \in A^*$. Применяя к (5) лемму 6 и используя при этом (3), получаем:

$$P = \bigcup_{i=1}^{n} \{\delta_{i,1}\} \cdot \{\delta_{i,2}\}^* \cdot \{\delta_{i,3}\} \quad (6)$$

для некоторых $n \in \mathbb{N}$, $\delta_{1,1}, \ldots, \delta_{n,3} \in A^*$. Из леммы 8 получаем, что все множества $\{\delta_{i,1}\} \cdot \{\delta_{i,2}\}^* \cdot \{\delta_{i,3}\}$ являются прогрессивными. Поэтому P представимо в виде конечного объединения прогрессивных множеств. Обозначим это объединение через Σ. Если в Σ есть конечные множества, то они одноэлементны. Пусть P_1 — одно из таких множеств. Если $Sp(P_1)$ пересекается с $Sp(P_2)$ для какого-то другого $P_2 \in \Sigma$, то $P_1 \subset P_2$. Поэтому P_1 можно выкинуть из Σ. Значит, можно считать, что все конечные множества из Σ спектрально независимы с остальными множествами. При этом каждое из них, является общепрогрессивным. Таким образом, теперь нужно разложить объединение бесконечных множеств из Σ в объединение спектрально независимых общепрогрессивных множеств. Если P_1 и P_2 имеют непустое пересечение, то будем писать $P_1 \leftrightarrow P_2$. Если существуют $Q_1, \ldots, Q_h \in \Sigma$, $h \geqslant 1$ такие, что

$$Q_1 \leftrightarrow Q_2, \; Q_2 \leftrightarrow Q_3, \quad \ldots \quad Q_{h-1} \leftrightarrow Q_h,$$

то пишем $Q_1 \longleftrightarrow Q_2$. Очевидно, что это отношение эквивалентности. Оно разбивает элементы множества Σ на непересекающиеся классы эквивалентности $Z_1, \ldots, Z_n$. Из леммы 9 следует, что любые два элемента из общего класса имеют одинаковые основания. А любая пара элементов из разных классов спектрально независима, так как в противном случае эти элементы имели бы непустое пересечение и попали бы в один класс эквивалентности. Возьмем произвольный такой класс

$$Z_i = (P_1(i), \ldots, P_{w(i)}(i)).$$

Здесь $P_1, \ldots, P_w$ — элементы из этого класса. Обозначим

$$W_i = \bigcup_{j=1}^{w(i)} P_j,$$

где $n \geqslant i \geqslant 1$. Каждое W_i является конечным объединением прогрессивных множеств с одинаковым основанием, то есть оно общепрогрессивно. При этом, любые два разных множества W_i и W_j спектрально независимы, так как иначе нашлась бы пара $P_r(i) \in W_i$, $P_s(j) \in W_j$, которая была бы спектрально зависимой. Но, как уже отмечалось выше, это невозможно.

Таким образом, мы разбили 1-тонкое множество P в конечное объединение спектрально независимых в совокупности общепрогрессивных множеств.

Утверждение теоремы 2 доказано.

Доказательство теоремы 3.

Пусть $R = \bigsqcup_{i=1}^{k} R_i$, где R_i — прогрессивные множества. Тогда при всех $1 \leqslant i \leqslant k$ имеем:

$$R_i = \{\alpha_i\} \cdot \{\beta_i\}^* \cdot \{\gamma_i\}$$

для некоторых $\alpha_i, \beta_i, \gamma_i \in A^*$. Очевидно, что все R_i будут 1-тонкими множествами. Осталось заметить, что конечное объединение тонких множеств само является тонким множеством. Докажем это.

Достаточно доказать это для объединения двух множеств. Пусть

$$L = L_1 \cup L_2,\ L_1 \in T_s,\ L_2 \in T_r$$

для некоторых $s, r \in \mathbb{N}, s \geqslant r$. Тогда среди любых $2s+2$ слов одинаковой длины из L по принципу Дирихле будет хотя бы $s+1$ слов из L_1 или из L_2 и, по определению s-тонких и l-тонких множеств, хотя бы два из этих слов будут совпадать. Это и означает, что $L \in T$.

Утверждение теоремы 3 доказано.

Доказательство теоремы 4.

Пусть P — s-тонкое множество. Оно регулярно. Применив лемму 1, получаем, что для некоторых чисел $k, s(1), \ldots, s(k) \in \mathbb{N}$, слов $\alpha_{1,1}, \ldots, \alpha_{k,s(k)} \in A^*$ и регулярных выражений $\mathfrak{P}_{1,1}, \ldots, \mathfrak{P}_{k,s(k)-1}$ верно

$$P = \bigcup_{i=1}^{k} |\alpha_{i,1} \cdot (\mathfrak{P}_{i,1})^* \cdot \alpha_{i,2} \cdot \ldots \cdot \alpha_{i,s(i)-1} \cdot (\mathfrak{P}_{i,s(i)-1})^* \cdot \alpha_{i,s(i)}|.$$

Поэтому

$$P = \bigcup_{i=1}^{k} \{\alpha_{i,1}\} \cdot |\mathfrak{P}_{i,1}|^* \cdot \{\alpha_{i,2}\} \cdot \ldots \cdot \{\alpha_{i,s(i)-1}\} \cdot |\mathfrak{P}_{i,s(i)-1}|^* \cdot \{\alpha_{i,s(i)}\}. \quad (7)$$

Пользуясь леммами 11, 12 получаем, что существуют числа $l_{i,j} \leqslant s$ такие, что $|\mathfrak{P}_{i,j}|^*$ — $l_{i,j}$-тонкие множества. Здесь и далее i и j пробегают все значения, удовлетворяющие неравенствам

$$1 \leqslant i \leqslant k, \quad 1 \leqslant j \leqslant s(i) - 1.$$

Из леммы 13 следует, что

$$|\mathfrak{P}_{i,j}|^* = R_{i,j} \cup (\{\delta_{i,j}\} \cdot \{\gamma_{i,j}\}^*) \quad (8)$$

для некоторых конечных $R_{i,j} \subseteq A^*$ и $\delta_{i,j}, \gamma_{i,j} \in A^*$. Для любых $\alpha, \beta, \gamma \in A^*$ верно, что

$$(\{\alpha\} \cup \{\beta\}) \cdot \{\gamma\} = \{\alpha\gamma\} \cup \{\beta\gamma\}, \quad \{\gamma\} \cdot (\{\alpha\} \cup \{\beta\}) = \{\gamma\alpha\} \cup \{\gamma\beta\}. \quad (9)$$

Подставляя (8) в (7) и применяя (9), получаем:

$$P = \bigcup_{i=1}^{l} \{\beta_{i,1}\} \cdot \{\beta_{i,2}\}^* \cdot \{\beta_{i,3}\} \cdot \ldots \cdot \{\beta_{i,t(i)-1}\} \cdot \{\beta_{i,t(i)}\}^* \cdot \{\beta_{i,t(i)+1}\} \quad (10)$$

для некоторых $l, t(1), \ldots, t(l) \in \mathbb{N}$, $\beta_{1,1}, \ldots, \beta_{l,t(l)+1} \in A^*$. Применяя к (10) леммы 11, 12, 14 получаем, что:

$$P = \bigcup_{i=1}^{m} \{\gamma_{i,1}\} \cdot \{\gamma_{i,2}\}^* \cdot \{\gamma_{i,3}\}^* \cdot \ldots \cdot \{\gamma_{i,u(i)}\}^* \cdot \{\gamma_{i,u(i)+1}\} \qquad (11)$$

для некоторых $m, u(1), \ldots, u(m) \in \mathbb{N}$,$\gamma_{1,1}, \ldots, \gamma_{m,u(m)+1} \in A^*$. Применяя к (11) леммы 11, 12, 15 и используя при этом (9), получаем:

$$P = \bigcup_{i=1}^{n} \{\delta_{i,1}\} \cdot \{\delta_{i,2}\}^* \cdot \{\delta_{i,3}\} \qquad (12)$$

для некоторых $n \in \mathbb{N}$, $\delta_{1,1}, \ldots, \delta_{n,3} \in A^*$.

Для всех $1 \leqslant i \leqslant n$ обозначим через P_i множество $\{\delta_{i,1}\} \cdot \{\delta_{i,2}\}^* \cdot \{\delta_{i,3}\}$. В силу определения все P_i — прогрессивные множества. При этом, $P = \bigcup\limits_{i=1}^{n} P_i$. Доказательство теоремы завершает применение леммы 16.

Список литературы

[1] Кудрявцев В. Б., Алешин С. В., Подколзин А. С. Введение в теорию автоматов. — М.: Наука, 1985.

[2] Яблонский С. В. Введение в дискретную математику. — М.: Наука, 1986.

[3] Марков Ал. А. Введение в теорию кодирования. — М.: Наука, 1982.

[4] Дергач П. С. Об однозначности алфавитного декодирования // Интеллектуальные системы — 2011. Т. 15, вып. 1–4. — С. 349–361.