ДИСКРЕТНЫЙ АНАЛИЗ И ИССЛЕДОВАНИЕ ОПЕРАЦИЙ
Июль—сентябрь 2005. Серия 1. Том 12, № 3, 3–31

УДК 519.713

О СВОЙСТВАХ СТОХАСТИЧЕСКОГО КС-ЯЗЫКА, ПОРОЖДЁННОГО ГРАММАТИКОЙ С ДВУМЯ КЛАССАМИ НЕТЕРМИНАЛЬНЫХ СИМВОЛОВ

А. Е. Борисов

Рассматривается стохастический контекстно-свободный язык, порождённый грамматикой с двумя классами нетерминальных символов. Исследуется случай, когда перронов корень матрицы первых моментов меньше 1. Найдены асимптотические формулы для математического ожидания числа применений правила грамматики в дереве вывода фиксированной высоты. Получена асимптотика для энтропии множества слов, имеющих деревья вывода заданной высоты. Найдена нижняя оценка стоимости двоичного кодирования рассматриваемого языка.

Введение

В статье рассматривается язык, порождённый стохастической контекстно-свободной (КС-) грамматикой с двумя классами нетерминальных символов. Исследуется случай, когда матрица первых моментов грамматики разложима и её перронов корень меньше 1. Для такого языка установлены закономерности в деревьях вывода высоты t при $t \to \infty$. Особое внимание уделяется случаю кратного перронова корня, поскольку именно в этом случае появляются существенные отличия от неразложимого случая. Полученные результаты использованы для получения нижней оценки стоимости двоичного кодирования слов рассматриваемого языка.

В [4] рассматривалась КС-грамматика, матрица первых моментов которой удовлетворяет следующим ограничениям:

(1) неразложима и непериодична,

(2) максимальный по модулю неотрицательный собственный корень (перронов корень) строго меньше единицы.

Для слов большой длины стохастического КС-языка, порождаемого такой грамматикой, установлены свойства, аналогичные свойствам слов, генерируемых эргодическим конечным источником [9]. Кроме того, для множества деревьев вывода высоты t при $t \to \infty$ найдены математические ожидания числа применений произвольного правила грамматики на фиксированном ярусе дерева вывода и во всем дереве вывода.

В настоящей статье аналогичные вопросы исследуются для грамматики, содержащей два класса нетерминальных символов, с разложимой и непериодичной матрицей первых моментов, перронов корень которой меньше единицы. Показано, что для перронова корня кратности два математическое ожидание числа применений любого правила в дереве вывода высоты t на ярусе τ "почти линейно" зависит от τ, а для простого перронова корня оно стремится к константе. Также доказано, что среднее число правил, приходящихся на один ярус дерева вывода, стремится к константе при $t \to \infty$.

С помощью этих результатов получено, что энтропия $H(t)$ множества слов, имеющих деревья вывода высоты t, асимптотически линейно зависит от t. Это позволило найти нижнюю оценку стоимости кодирования слов рассматриваемого языка.

1. Определения

Будем использовать определения КС-грамматики и стохастического КС-языка, приведённые в [1, 8].

Языком над алфавитом $B = \{b_1, \ldots, b_n\}$ называется подмножество $L \subseteq B^*$. Стохастический язык $\mathcal{L}$ определяется как пара (L, P), где L — язык, а P — распределение вероятностей на множестве его слов, причём вероятность каждого слова языка больше нуля.

Стохастическая порождающая контекстно-свободная грамматика — это система $G = \langle V_T, V_N, R, s\rangle$, где V_T и V_N — множества терминальных и нетерминальных символов соответственно (в дальнейшем называемых просто терминалами и нетерминалами), $|V_N| = k$, $s \in V_N$ — аксиома, R — конечное множество правил, $R = \bigcup\limits_{i=1}^{k} R_i$, где $R_i = \{r_{i1}, \ldots, r_{in_i}\}$, и каждое правило в R_i имеет вид

$$r_{ij} : A_i \xrightarrow{p_{ij}} \beta_{ij},\ j = 1, \ldots, n_i, \text{ где } A_i \in V_N, \beta_{ij} \in (V_T \cup V_N)^*,$$

а p_{ij} — вероятность применения правила r_{ij}, удовлетворяющая условиям: $0 < p_{ij} \leqslant 1,\ \sum\limits_{j=1}^{n_i} p_{ij} = 1$.

Применение правила к слову состоит в замене в этом слове нетерминала, стоящего в левой части правила, на правую часть применяемого правила.

Будем говорить, что слово α *непосредственно выводимо* из слова β, если $\alpha = \alpha_1 A_i \alpha_2,\ \beta = \alpha_1 \beta_{ij} \alpha_2$ для некоторых слов $\alpha_1, \alpha_2 \in (V_T \cup V_N)^*$ и грамматика содержит правило $r_{ij} : A_i \overset{p_{ij}}{\to} \beta_{ij}$. Рефлексивное транзитивное замыкание отношения непосредственной выводимости назовём *отношением выводимости.* Через L_G обозначим множество слов в терминальном алфавите, выводимых из аксиомы s. Под выводом слова будем понимать последовательность правил грамматики, с помощью которой данное слово выводится из аксиомы. Левый вывод — это вывод, в котором каждое правило применяется к самому левому по порядку нетерминалу в слове.

Пусть $\alpha \in L_G$. Каждому левому выводу слова α соответствует дерево вывода, корень которого помечен аксиомой, а вершины — терминальными и нетерминальными символами [1]. При применении правила $A_i \overset{p_{ij}}{\to} h_1 h_2 \ldots h_n$ в вершине a, помеченной нетерминалом A_i, добавляются n дуг от a к вершинам следующего яруса, которые помечаются слева направо символами $h_1, h_2, \ldots, h_n$ ($h_i \in V_T \cup V_N$). Все листья дерева помечены терминальными символами; при этом само слово получается обходом листьев дерева слева направо. *Высотой* дерева вывода называется наибольшая длина пути от корня до листа.

Если α — некоторое слово, которое имеет левый вывод $\omega = r_{i_1 j_1} \ldots r_{i_k j_k}$, а d — соответствующее ему дерево вывода, то положим $p(d) = p_{i_1 j_1} \ldots p_{i_k j_k}$. Пусть $D(L_G)$ — множество всех деревьев вывода для слов из L_G.

Стохастическая КС-грамматика называется *согласованной,* если

$$\lim_{N\to\infty} \sum_{d \in D(L_G), |d| < N} p(d) = 1,$$

где через $|d|$ обозначается высота дерева d. В дальнейшем будем рассматривать только согласованные грамматики. Величину $p(d)$ будем называть *вероятностью* дерева вывода d.

Стохастическим КС-языком, соответствующим стохастической КС-грамматике G, называется множество слов L_G с индуцированным на нём распределением вероятностей (вероятность слова равна суммарной вероятности всех различных деревьев вывода для этого слова).

Введём некоторые обозначения и соглашения. Рассматривая вектор, будем считать, что мы имеем дело с вектором-столбцом, если противное не оговорено специально. В дальнейшем для вектора или матрицы X будем писать $X = c$ ($X \leqslant c,\ X \geqslant c,\ X < c,\ X > c$), где c — скаляр, если все компоненты вектора или матрицы X равны (соответственно меньше или равны, больше или равны, меньше или больше) c. Запись $|X| < c$ означает, что все компоненты вектора или матрицы X по модулю меньше скаляра c. Кроме того, для двух векторов или матриц X_1, X_2 одинаковой размерности будем писать, что $X_1 = X_2$ ($X_1 \leqslant X_2,\ X_1 \geqslant X_2,\ X_1 < X_2,\ X_1 > X_2$), если все компоненты вектора или матрицы X_1 равны (соответственно меньше или равны, больше или равны, меньше или больше) компонентам X_2. Через $\bar{c}$ обозначим вектор, все компоненты которого равны константе c. Записи $X/c,\ cX,\ X+c$, где X — вектор или матрица, обозначают соответственно деление, умножение всех элементов X на скаляр c или прибавление этого скаляра ко всем элементам X. Через X^T обозначим транспонирование вектора или матрицы X. Через $v(m:n)$, где $v = (v_1, \ldots, v_m, \ldots, v_n, \ldots, v_k)$ — вектор, будем обозначать вектор $(v_m, \ldots, v_n)$. Через (v_1, v_2), где v_1, v_2 — вектора-столбцы, будем обозначать объединённый вектор-столбец $(v_1^T, v_2^T)^T$.

Пусть k — число нетерминалов в рассматриваемой грамматике и $s = (s_1, \ldots, s_k)$. Введём многомерные производящие функции [6]:

$$F_i(s) = F_i(s_1, \ldots, s_k) = \sum_{j=1}^{n_i} p_{ij} \cdot s_1^{l_{ij}^1} \cdot \ldots \cdot s_k^{l_{ij}^k},\ i = 1, \ldots, k,$$

где l_{ij}^m — число вхождений нетерминального символа A_m в правую часть правила r_{ij} ($m = 1, \ldots, k$). Тогда первые моменты грамматики определяются как $a_j^i = \frac{\partial F_i(s)}{\partial s_j}|_{s=1}$, и матрица $A = (a_j^i)$ называется матрицей *первых моментов.*

Будем говорить, что нетерминал A_j следует за нетерминалом A_i, или A_i предшествует A_j, если из A_i выводимо хотя бы одно слово, содержащее нетерминал A_j (и обозначать $A_i \to A_j$). Грамматика называется *неразложимой,* если для любых двух различных нетерминалов A_i и A_j верно $A_i \to A_j$. В противном случае грамматика называется *разложимой.* Классом нетерминалов назовём максимальное по включению подмножество $K \in V_N$ такое, что $A_i \to A_j$ для любых $A_i, A_j \in K$. Будем говорить, что класс K_1 предшествует классу K_2 (и писать $K_1 \prec K_2$), если $A_1 \to A_2$ для любых $A_1 \in K_1,\ A_2 \in K_2$. Очевидно, что множество нетерминалов V_N является объединением конечного числа непересекающихся классов.

Любой стохастической КС-грамматике можно очевидным образом поставить в соответствие дискретный ветвящийся процесс, причём нетерминалу грамматики будет соответствовать тип частицы в ветвящемся процессе. Тогда классу нетерминалов грамматики будет соответствовать класс типов частиц в ветвящемся процессе.

Пусть G — стохастическая КС-грамматика. Через L_i обозначим язык, порождённый грамматикой, которая получена заменой аксиомы исходной грамматики на нетерминал A_i. Через D_i обозначим множество деревьев вывода слов из L_i, через D_i^t — множество деревьев вывода высоты t для слов из L_i. Если $d \in D_i^t$, то через $p_t(d)$ обозначим условную вероятность дерева d. Она равна $p(d)/P(D_i^t)$, где $P(D_i^t)$ — общая вероятность деревьев вывода из D_i^t. Обозначим через $Q_i(t)$ вероятность множества деревьев из D_i высоты более t. Эту величину назовём вероятностью продолжения по аналогии с теорией ветвящихся процессов. Очевидно, что $P(D_i^t) = Q_i(t-1) - Q_i(t)$. Для исходной грамматики G будем полагать, что аксиомой является нетерминал A_1. В дальнейшем в обозначениях $L_i,\ D_i^t,\ Q_i(t),\ P(D_i^t)$ индекс i будем опускать при $i = 1$.

Пусть $\Xi = (\xi_1, \ldots, \xi_k)$ — случайный вектор, $\alpha^* = (\alpha_1, \ldots \alpha_k)$ — фиксированный вектор с целочисленными неотрицательными компонентами и $\alpha = \alpha_1 + \ldots + \alpha_k$. Рассмотрим величину $\Xi^{\alpha^*} = \xi_1^{[\alpha_1]} \cdot \ldots \cdot \xi_k^{[\alpha_k]}$, где $\xi_i^{[\alpha_i]} = \xi_i(\xi_i - 1)\ldots(\xi_i - \alpha_i + 1)$. Математическое ожидание $M\Xi^{\alpha^*}$ называется *факториальным моментом* вектора Ξ порядка α.

Обозначим через $x_j^i(t)$ число нетерминальных символов A_j в дереве вывода из D_i на ярусе t. Факториальный момент вектора $X^i(t) = (x_1^i(t), \ldots, x_k^i(t))$ порядка α будем называть *факториальным моментом* грамматики порядка α. По аналогии с [6] введём обозначения $a_j^i(t)$ для факториальных моментов первого порядка, причём $a_j^i(1)$ совпадают с введёнными ранее величинами a_j^i. Для факториальных моментов второго порядка введём обозначения $b_{jj}^i(t) = Mx_j^i(t)(x_j^i(t) - 1)$ и $b_{jk}^i(t) = Mx_j^i(t)x_k^i(t)$ $(j \neq k)$. Факториальные моменты 3-го порядка будем обозначать через $c_{jkl}^i(t)$. Как и для первых моментов, положим $b_{jk}^i = b_{jk}^i(1)$, $c_{jkl}^i = c_{jkl}^i(1)$. Известно [6], что $b_{jk}^i = \left.\dfrac{\partial^2 F_i(s)}{\partial s_j \partial s_k}\right|_{s=1}$, $c_{jkl}^i = \left.\dfrac{\partial^3 F_i(s)}{\partial s_j \partial s_k \partial s_l}\right|_{s=1}$.

Производящая функция и вероятности продолжения ветвящегося процесса [6], соответствующего грамматике, совпадают с соответствующими характеристиками грамматики. Это позволяет использовать результаты теории ветвящихся процессов для стохастических КС-грамматик.

2. Вероятности продолжения

Рассмотрим разложимую грамматику с двумя классами нетерминалов K_1, K_2, причём $K_1 \prec K_2$, аксиома s принадлежит K_1. Введём обозначения: $k_1 = |K_1|,\ k_2 = |K_2|,\ k_1 + k_2 = k$. Будем считать, что нетерминалы из K_1 имеют номера $1, \ldots, k_1$, из K_2 — номера $k_1 + 1, \ldots, k$. Для такой грамматики матрица первых моментов A имеет блочный вид:

$$A = \begin{pmatrix} A^{(1)} & A^{(2)} \\ 0 & A^{(3)} \end{pmatrix}, \tag{1}$$

где $A^{(1)}$ — квадратная матрица порядка k_1, $A^{(3)}$ — квадратная матрица порядка k_2, а $A^{(2)}$ — матрица размера $k_1 \times k_2$. Дополнительно предположим, что матрица A непериодична [3]. Тогда можно считать, что $A^{(1)} > 0$, $A^{(2)} > 0$, $A^{(3)}>0$, если использовать метод укрупнения правил грамматики, описанный в [4]. Поскольку матрицы $A^{(1)}$, $A^{(3)}$ положительны, по теореме Перрона [3] для каждой из них существует простой максимальный по модулю действительный положительный собственный корень (перронов корень).

Будем рассматривать случай, когда соответствующий грамматике ветвящийся процесс является докритическим, т. е. когда неразложимый ветвящийся процесс, порождённый каждым из классов K_1, K_2, докритический. Условие докритичности эквивалентно тому, что перроновы корни r', r'' матриц $A^{(1)}$, $A^{(3)}$ меньше единицы. Таким образом, в дальнейшем будем полагать, что $0 < r' < 1$ и $0 < r'' < 1$.

В этом разделе будут получены асимптотические формулы для вероятностей продолжения $Q_i(t), i = 1, \ldots, k$. Отдельно будем рассматривать три случая : $r' > r''$, $r' < r''$, $r' = r''$. Обозначим через v', v'' левые, через u', u'' правые собственные векторы матриц $A^{(1)}$ и $A^{(3)}$, соответствующие их перроновым корням r' и r'' при нормировке $v'u' = v''u'' = 1$ (левый собственный вектор матрицы считаем вектором-строкой). Левый и правый собственные векторы для перронова корня r матрицы A при условии нормировки $vu = 1$ будем обозначать через v, u. Нам понадобится следующая

Лемма 1. *Пусть матрица A имеет вид* (1). *Тогда при $r' \neq r''$ максимальный собственный корень r матрицы A является простым и равен* $\max\{r', r''\}$. *Соответствующие ему левый и правый собственные векторы могут быть выбраны неотрицательными. Кроме того, справедливы утверждения*:

1) *при $r' > r''$ правый собственный вектор матрицы A может быть выбран в виде $u = (u', 0)$, а левый — в виде $v = (v', v^2)$, где $v^2 > 0$ — некоторый положительный вектор*:

2) *при $r' < r''$ правый собственный вектор матрицы A может быть выбран в виде $u = (u^1, u'')$, а левый — в виде $v = (0, v'')$, где $u^1 > 0$ — некоторый положительный вектор.*

Доказательство. Из равенства $|A - \lambda E| = |A^{(1)} - \lambda E| \cdot |A^{(3)} - \lambda E|$ следует, что любой собственный корень матрицы A является собственным корнем для $A^{(1)}$ или $A^{(3)}$ и наоборот. Отсюда следуют простота r при $r' \neq r''$ и равенство $r = \max\{r', r''\}$.

Сначала рассмотрим случай $r' > r''$. Очевидно, вектор вида $u = (u', 0)$ является правым собственным вектором для r', вектор $v = (v^1, v^2)$ является левым собственным вектором для перронова корня $r = r'$ матрицы A, если

$$v^1 A^{(2)} + v^2 A^{(3)} = r' v^2,\ v^1 A^{(1)} = r' v^1, \tag{2}$$

где $v^1 = (v^1_1, \ldots, v^1_{k_1})$ и $v^2 = (v^2_1, \ldots, v^2_{k_2})$. Таким образом, можно считать, что $v^1 = v' > 0$. Легко показать, что

$$v^2 = v^1 A^{(2)} \cdot (r'E - A^{(3)})^{-1} = (r')^{-1} v^1 A^{(2)} \cdot \sum_{n=0}^{\infty} ((r')^{-1} A^{(3)})^n. \tag{3}$$

Действительно, последний ряд сходится, поскольку известно [3], что элементы матрицы $(A^{(3)})^n$ имеют вид $O((r'')^n)$. Поэтому элементы матрицы $((r')^{-1}A^{(3)})^n$ имеют вид $O((r''/r')^n)$. Прямой проверкой устанавливается, что матрица $E + \sum\limits_{n=1}^{\infty} ((r')^{-1}A^{(3)})^n$ является обратной к $E - (r')^{-1}A^{(3)}$. Следовательно, вектор $v = (v', v^2)$ удовлетворяет равенствам (2). Из положительности матриц $A^{(2)}, A^{(3)}$ и равенства (3) следует, что $v^2 > 0$. Доказательство утверждения леммы для случая $r' < r''$ проводится аналогично. Лемма 1 доказана.

В дальнейшем в качестве собственных векторов для A будем рассматривать векторы со свойствами, о которых говорится в лемме 1. Асимптотика вероятностей продолжения описывается следующей теоремой.

Теорема 1. *Для разложимой грамматики с двумя классами нетерминалов, матрица первых моментов которой имеет вид* (1), *справедливы утверждения*:

1) *если* $r' \neq r''$ *и* $t \to \infty$, *то*

$$Q_i(t) = c_0 u_i r^t + o(r^t);$$

2) *если* $r = r' = r''$ *и* $t \to \infty$, *то*

$$Q_i(t) = bc_0 u'_i \cdot t r^t (1 + o(1)), i = 1, \ldots, k_1,$$
$$Q_i(t) = c_0 u''_{i-k_1} r^t (1 + o(1)), i = k_1 + 1, \ldots, k.$$

Здесь r', r'' *— перроновы корни матриц* $A^{(1)}$ *и* $A^{(3)}$ *соответственно,* $r = \max(r', r'')$ *— перронов корень матрицы первых моментов,* u, v *—*

соответствующие ему правый и левый собственные векторы, $c_0 > 0$ — некоторая константа и b определено формулой

$$b = v' A^{(2)} u''/r. \tag{4}$$

Для доказательства этой теоремы нам понадобится несколько лемм, аналогичных леммам, доказанным в [6].

Лемма 2. *Пусть матрица A имеет вид (1), r', r'' — перроновы корни матриц $A^{(1)}$ и $A^{(3)}$ соответственно и $r = \max(r', r'') \leqslant 1$. Тогда справедливы следующие асимптотические равенства*:

1) *если $r' \neq r''$, то*

$$A^n = uvr^n + O(r_1^n),$$

где $0 < r_1 < r$, $u \geqslant 0$ и $v \geqslant 0$;

2) *если $r' = r'' = r$, то*

$$A^n = \begin{pmatrix} u'v'r^n + O(r_1^n) & bu'v'' \cdot nr^n + o(nr^n) \\ 0 & u''v''r^n + O(r_1^n) \end{pmatrix},$$

где $r_1 < r$ и $b = v' A^{(2)} u''/r$.

Первая асимптотика доказывается так же, как и для неразложимой матрицы (см. [3] или [6]), существование и неотрицательность собственных векторов для перронова корня при $r' \neq r''$ доказана в лемме 1. Вторая асимптотика следует из равенства

$$A^n = \begin{pmatrix} (A^{(1)})^n & \sum\limits_{i=0}^{n-1} (A^{(1)})^i A^{(2)} (A^{(3)})^{n-1-i} \\ 0 & (A^{(3)})^n \end{pmatrix}$$

и соотношений $(A^{(1)})^t = u'v'r^t + O(r_1^t)$, $(A^{(3)})^t = u''v''r^t + O(r_1^t)$, $0 < r_1 < r$.

Из леммы 2 и равенства $A^n = (a_i^j)^n = (a_i^j(n))$ легко получить вид первых моментов грамматики.

Следствие 1. *Пусть матрица A имеет вид (1) и её перронов корень $r \leqslant 1$. Тогда (в обозначениях предыдущей леммы) при $r' \neq r''$*

$$a_j^i(t) = u_i v_j r^t + O(r_1^t),$$

а при $r = r' = r''$

$$\begin{aligned} a_j^i(t) &= u_i' v_j' r^t + O(r_1^t), && i, j \leqslant k_1; \\ a_j^i(t) &= b u_i' v_{j-k_1}'' t r^t + o(t r^t), && i \leqslant k_1,\ j > k_1; \\ a_j^i(t) &= u_{i-k_1}'' v_{j-k_1}'' r^t + O(r_1^t), && i, j > k_1, \end{aligned}$$

где $0 \leqslant r_1 < r$.

Лемма 3. *Пусть матрица A имеет вид* (1), r', r'' *— перроновы корни матриц $A^{(1)}$ и $A^{(3)}$ соответственно,* $r = \max(r', r'') \leqslant 1$ *и*

$$A_t = \begin{pmatrix} A_t^{(1)} & A_t^{(2)} \\ 0 & A_t^{(3)} \end{pmatrix}$$

— последовательность таких матриц, что $A_t^{(1)}$, $A_t^{(2)}$, $A_t^{(3)}$ *— матрицы размеров* $k_1 \times k_1$, $k_1 \times k_2$, $k_2 \times k_2$ *соответственно,* $0 \leqslant A_t \leqslant A$ *и* $A_t \to 0$ *при* $t \to \infty$. *Пусть* $A_t^* = (A - A_1)(A - A_2)\ldots(A - A_t)$. *Тогда для любого вектора* $x > 0$ *выполняются равенства*:

1) *если* $r' \neq r''$, *то* $\lim\limits_{t\to\infty} \dfrac{A_t^* x}{v A_t^* x} = u$;

2) *если* $r' = r'' = r$, *то*

$$\lim_{t\to\infty} \frac{(A_t^* x)(1 : k_1)}{v' \cdot (A_t^* x)(1 : k_1)} = u', \quad \lim_{t\to\infty} \frac{(A_t^* x)(k_1 + 1 : k)}{v'' \cdot (A_t^* x)(k_1 + 1 : k)} = u''.$$

Схема доказательства этой леммы аналогична схеме доказательства соответствующего утверждения для неразложимой матрицы A [6]. Очевидно, можно считать, что $r = \max(r', r'') = 1$ (домножив матрицы A, A_i на константу $1/r$).

Рассмотрим случай $r' \neq r''$. Очевидно, что

$$A_t^* x = (A - A_t)\ldots(A - A_{n+1}) \cdot A_n^* x \text{ при } t > n.$$

Положим

$$S_{ij} = \begin{cases} 0 \text{ при } k_1 < i \leqslant k, \ 0 < j \leqslant k_1, \\ 1 \text{ в остальных случаях.} \end{cases}$$

Зададим произвольно малые величины $\varepsilon_1, \varepsilon_2 > 0$, причём ε_2 возьмём настолько малым, чтобы выполнялось неравенство $A > \varepsilon_2 S$ (это возможно, поскольку матрица A имеет вид (1)), и выберем такое достаточно большое l, чтобы выполнялось неравенство $|A^l - uv| < \varepsilon_1$. Это возможно по первому утверждению леммы 2. Выберем $n = n(\varepsilon_2)$ таким, что $A_s < \varepsilon_2, \ i, j < k$ для всех $s > n$. Очевидно, что при $t = l + n$ верно неравенство

$$(A - \varepsilon_2 S)^{t-n} x \leqslant (A - A_t)\ldots(A - A_{n+1})x \leqslant A^{t-n} x, \tag{5}$$

так как $A - \varepsilon_2 S > 0$ и $A - A_k \geqslant 0$.

Пусть $m_0 = \max\limits_{i,j,d}(A^d)_{ij}$, причём $m_0 < \infty$ по лемме 2. Так как $SA \leqslant km_0 \cdot S$, $AS \leqslant km_0 \cdot S$, то $S^{d_1}A^{d_2} \leqslant k^{d_1}m_0 \cdot S$ и $A^{d_2}S^{d_1} \leqslant k^{d_1}m_0 \cdot S$. Поэтому

$$|A^l - (A - \varepsilon_2 S)^l| \leqslant \sum_{i=1}^{l} C_l^i (\varepsilon_2 k)^i \cdot m_0.$$

Преобразуя это неравенство, получаем

$$|A^l - (A - \varepsilon_2 S)^l| \leqslant ((1 + \varepsilon_2 k)^l - 1) \cdot m_0. \tag{6}$$

Очевидно, для заданных l и $\varepsilon_3 > 0$ можно выбрать такое $\varepsilon_2 > 0$ (и соответствующее n), что $((1+\varepsilon_2 k)^l - 1) \cdot m_0 < \varepsilon_3$. В силу неравенств (5), (6) и неравенства $|A^l - uv| < \varepsilon_1$ верно неравенство

$$\begin{aligned} |A_t^* x - uv \cdot A_n^* x| &\leqslant |(A^l - (A - \varepsilon_2 S)^l) \cdot A_n^* x| + |(uv - A^l) \cdot A_n^* x| \\ &< (\varepsilon_3 + \varepsilon_1) \cdot \sum_i (A_n^* x)_i \leqslant (\varepsilon_3 + \varepsilon_1) m_0 \cdot \sum_i x_i. \end{aligned}$$

Следовательно, можно найти такие достаточно большие n, t, что $|A_t^* x - uv \cdot A_n^* x| < \varepsilon$ для любого наперед заданного $\varepsilon > 0$. Обозначив скалярную величину $vA_n^* x$ через c_0 и учитывая условие нормировки $vu = 1$, можно записать, что

$$\frac{A_t^* x}{vA_t^* x} = \frac{c_0 u + O(\varepsilon)}{c_0 + O(\varepsilon)} = u + O(\varepsilon).$$

Тем самым первое утверждение леммы доказано. Доказательство леммы в случае $r' = r''$ проводится аналогично с использованием второго утверждения леммы 2. Лемма 3 доказана.

Пусть $F(t, s) = (F_1(t, s), \ldots, F_k(t, s))$ — (векторная) производящая функция, которая определяется как t-я итерация производящей функции $F(s)$ соотношениями:

$$\begin{aligned} &F(0, s) = s, \quad F(1, s) = F(s), \\ &F(t + 1, s) = F(F(t, s)). \end{aligned} \tag{7}$$

Очевидно, что $F(t, \bar{1}) = \bar{1}$ для всех t. Известно [6], что $Q(t) = \bar{1} - F(t, \bar{0})$, где $Q(t) = (Q_1(t), \ldots, Q_k(t))$. Пусть $R(t, s) = \bar{1} - F(t, s)$, в частности, $R_i(t, \bar{0}) = Q_i(t)$.

Лемма 4. *Для стохастической КС-грамматики с матрицей первых моментов A вида* (1) *верно равенство*

$$\bar{1} - F(s) = [A - E(s)](\bar{1} - s), \tag{8}$$

где $0 \leqslant E(s) \leqslant A$, *причём элементы матрицы* $E(s)$ *при* $0 \leqslant s \leqslant 1$ *удовлетворяют условиям*

$$E_{ij}(s) = \sum_l \delta^i_{jl}(s)(1 - s_l), \text{ где } 0 \leqslant \delta^i_{jl}(s) \leqslant b^i_{jl} \tag{9}$$

и $E(s)$ *имеет блочный вид* (1) *при любом* s, $0 \leqslant s \leqslant 1$.

Доказательство. Используя разложение производящей функции $F_i(s)$ в ряд Тейлора в окрестности 1, можно записать:

$$1 - F_i(s) = \sum_j \frac{\partial F_i(s)}{\partial s_j}|_{s=\theta^i}(1 - s_j) = \sum_j E_{ij}(s)(1 - s_j),$$

где $\theta^i = (\theta^i_1, \dots \theta^i_k)$, причём $\theta^i = \alpha^i \cdot \bar{1} + (1 - \alpha^i) \cdot s$ $(0 \leqslant \alpha^i \leqslant 1)$. Поскольку производящие функции $F_i(s)$ — многочлены с положительными коэффициентами, все их производные монотонно возрастают по всем s_j, и следовательно, $0 \leqslant \dfrac{\partial F_i(s)}{\partial s_j} \leqslant \dfrac{\partial F_i(s)}{\partial s_j}|_{s=\theta^i} \leqslant a^i_j$. Раскладывая $\dfrac{\partial F_i(s)}{\partial s_j}$ аналогичным образом в ряд Тейлора, получаем

$$\frac{\partial F_i(s)}{\partial s_j} = a^i_j - \sum_l \delta^i_{jl}(s)(1 - s_l),$$

где $0 \leqslant \delta^i_{jl}(s) \leqslant b^i_{jl} = \left.\dfrac{\partial^2 F_i(s)}{\partial s_j \partial s_l}\right|_{s=1}$. Отсюда следуют равенства (8) и (9). Утверждение о том, что матрица $E(s)$ имеет блочный вид (1), следует из того, что $a^i_j = b^i_{jl} = 0$ при $i > k_1$ и $j \leqslant k_1$. Лемма 4 доказана.

Подставляя в соотношение (8) в качестве s вектор $F(t, s)$ и используя равенство (7), получаем

$$\bar{1} - F(t + 1, s) = (A - E(F(t, s)))\,(\bar{1} - F(t, s)). \tag{10}$$

Обозначим $E(F(t, s))$ через $E_t(s)$, а $E_t(0)$ через E_t и применим формулу (10) рекурсивно. Тогда

$$R(t, s) = \bar{1} - F(t, s) = \prod_{i=1}^{t-1}(A - E_i(s))(\bar{1} - s). \tag{11}$$

Отсюда при $s = \bar{0}$ следует, что

$$Q(t) = \prod_{i=1}^{t-1}(A - E_i) \cdot \bar{1}. \tag{12}$$

Кроме того, из формулы (11) следует, что при любом s, $0 \leqslant s \leqslant 1$,

$$R(t,s) = \bar{1} - F(t,s) \leqslant A^t \cdot (\bar{1} - s). \tag{13}$$

Так как $A^t = O(tr^t)$, то $F(t,s) \to 1$ при $t \to \infty$. Следовательно, $\lim\limits_{t\to\infty} E_t(s)=0$ по лемме 4 (поэлементно).

Введём обозначение: $A - E_i = A_i$.

Лемма 5. *Пусть матрица первых моментов грамматики имеет вид* (1). *Тогда при любом* s, $0 \leqslant s \leqslant 1$ *справедливы следующие утверждения*:

1) *если* $r' \neq r''$, *то* $E_t(s) = O(r^t)$;

2) *если* $r' = r''$, *то* $E_t(s) = O(tr^t)$.

Кроме того, $0 \leqslant E_t(s) \leqslant A$ *при любом* $t > 0$.

Доказательство. Сначала докажем первое утверждение. Из неравенства (13) и леммы 2 следует, что если $r' \neq r''$, то $R(t,s) = O(r^t)$. По лемме 4 имеем

$$E_t(s)_{ij} = E(F(t,s))_{ij} \leqslant \sum_l \delta^i_{jl}(F(t,s))(1 - F_l(t,s)) = O(r^t).$$

Отсюда вытекает первое утверждение леммы. Второе утверждение доказывается аналогично с использованием соотношения $R(t,s) = O(tr^t)$ при $r' = r''$, которое также следует из формулы (13). Лемма 5 доказана.

Доказательство теоремы 1. Сначала рассмотрим случай $r' \neq r''$. Ввиду равенства (12) и леммы 3 верно соотношение $\lim\limits_{t\to\infty} Q(t)/vQ(t) = u$, или $Q_i(t) = u_i vQ(t)(1 + o(1))$. Следовательно, достаточно показать, что существует конечный предел $\lim\limits_{t\to\infty} vQ(t)/r^t > 0$. Так как $Q(t+1) = (A - E_t)Q(t)$, то $vQ(t+1) = r \cdot vQ(t) - vE_tQ(t)$. Поэтому $vQ(t+1) = \left(r - \dfrac{vE_tQ(t)}{vQ(t)}\right) \cdot vQ(t)$. В силу неравенства $0 \leqslant E_t(s) \leqslant A$ (см. лемму 5) доказываемое утверждение следует из сходимости бесконечного произведения $\prod\limits_{t=0}^{\infty}\left(1 - \dfrac{vE_tQ(t)}{rvQ(t)}\right)$, что по известному критерию [7] сходимости бесконечного произведения следует из сходимости ряда $\sum\limits_{t=0}^{\infty} \dfrac{vE_tQ(t)}{rvQ(t)}$, которая вытекает из соотношения $E_t(s)=O(r^t)$ леммы 5.

Теперь рассмотрим случай $r = r' = r''$. Введём обозначения:

$$Q(t)(1:k_1) = x^1_t, \quad Q(t)(k_1+1:k) = x^2_t.$$

Используя формулу из [6] для вероятностей продолжения неразложимого ветвящегося процесса, соответствующего классу K_2, получаем, что при $i > k_1$ и $t \to \infty$ выполняется асимптотическое равенство

$$Q_i(t) = c_0 u''_{i-k_1}(r'')^t(1 + o(1)), \tag{14}$$

где $c_0 > 0$. Поэтому $x_t^2 = c_0 u'' r^t (1 + o(1))$.

Пусть

$$E_t = \begin{pmatrix} E_t^{(1)} & E_t^{(2)} \\ 0 & E_t^{(3)} \end{pmatrix}.$$

Домножая равенство $x_{t+1}^1 = (A - E_t^{(1)})x_t^1 + (A^{(2)} - E_t^{(2)})x_t^2$ слева на v' и подставляя выражение для x_t^2, получаем, что $v'x_{t+1}^1 = rv'x_t^1 - v'E_t^{(1)}x_t^1 + v'(A^{(2)} - E_t^{(2)})u'' \cdot c_0 r^t(1 + o(1))$. Введём обозначение: $x_t = v'x_t^1/r^t$. Тогда предыдущее равенство переписывается таким образом:

$$x_{t+1} = x_t + v'(A^{(2)} - E_t^{(2)})u'' \cdot r^{-1}c_0(1 + o(1)) - v'E_t^{(1)}x_t^1/r^{t+1}. \tag{15}$$

Так как из равенства (15) вытекает, что

$$x_{t+1} \leqslant x_t + r^{-1}v'A^{(2)}u'' \cdot c_0(1 + o(1)),$$

то существует константа $m > 0$ такая, что $x_t < mt$. Используя полученную в лемме 5 оценку $E_t = O(tr^t)$, отсюда получаем, что $v'E_t^{(1)}x_t^1/r^{t+1} = O(t^2r^t) \to 0$ при $t \to \infty$. Следовательно, $x_t = c_0 v'A^{(2)}u''t/r + o(t)$. Используя вторую часть утверждения леммы 2, получаем утверждение теоремы. Теорема 1 доказана.

Следствие 2. *Вероятность деревьев вывода высоты t при $t \to \infty$ имеет следующий вид:*

1) *если $r' \neq r''$, то*

$$P(D_i^t) = c_0 u_i \cdot (1 - r)r^{t-1} + o(r^t);$$

2) *если $r' = r''$, то*

$$P(D_i^t) = bc_0 u'_i \cdot (1 - r)tr^{t-1} + o(tr^t), \qquad i = 1, \dots, k_1,$$
$$P(D_i^t) = c_0 u''_i \cdot (1 - r)r^{t-1} + o(r^t), \qquad i = k_1 + 1, \dots, k,$$

где b взято из формулы (4).

3. Закономерности в деревьях вывода слов

Введём следующие обозначения. Пусть $X = (x_1, \dots x_k)$ — вектор с неотрицательными целыми компонентами. Через $D_X(\tau)$ обозначим множество деревьев вывода, имеющих x_i нетерминалов A_i на ярусе τ, где $i = 1 \dots, k$, а через $D_X^t(\tau)$ — множество всех деревьев вывода из $D_X(\tau)$, имеющих высоту t. Через $\mathcal{D}_X^{\leqslant n}$ обозначим множество наборов деревьев вывода

$$d_X^{\leqslant n} = (d_{11}, \dots d_{1x_1}, \dots, d_{k1}, \dots, d_{kx_k}),$$

где корни деревьев $d_{m1}, \dots, d_{mx_m}$ помечены нетерминалом A_m, $m = 1, \dots, k$, причём каждое дерево в наборе имеет высоту не больше n. Через $p(d_X^{\leqslant n})$ обозначим произведение вероятностей деревьев набора d. Через $\mathcal{D}_X^n$ обозначим множество наборов из $\mathcal{D}_X^{\leqslant n}$, которые содержат хотя бы одно дерево высоты n. Наборы из $\mathcal{D}_X^n$ будем обозначать через d_X^n, а произведение вероятностей деревьев, образующих d_X^n, через $p(d_X^n)$. Пусть $Q_X(n) = \sum p(d_X^{\leqslant n})$, где суммирование производится по всем наборам из $\mathcal{D}_X^{\leqslant n}$, и $R_X(n) = \sum p(d_X^n)$, где суммирование производится по всем наборам из $\mathcal{D}_X^n$. Очевидно, справедливо равенство $P(D_X^t(\tau)) = P(D_X(\tau)) \cdot R_X(t-\tau)$. Далее величину $P(D_X^t(\tau))$ будем обозначать через $P_X(\tau)$. Произведем оценку величин $R_X(n)$ и $Q_X(n)$ для рассматриваемой разложимой грамматики.

Рассмотрим случайную величину $q_{ij}(t,\tau)$ — число применений правила r_{ij} на ярусе τ в деревьях вывода высоты t. Её математическое ожидание обозначим через $M(q_{ij}(t,\tau))$. Поскольку вероятности продолжения $Q_i(t)$ в случае $r' \neq r''$ для рассматриваемой разложимой грамматики имеют тот же вид, что и для неразложимой, все доказанные в [4] результаты относительно математического ожидания числа применений правила грамматики на фиксированном ярусе в деревьях вывода заданной высоты и математического ожидания числа применений правила грамматики на одном ярусе дерева вывода остаются в силе. Поэтому справедлива следующая

Теорема 2. *Пусть матрица первых моментов КС-грамматики имеет вид* (1) *и* $r' \neq r''$. *Тогда при* $\tau \to \infty$ *и* $t-\tau \to \infty$ *выполняется соотношение*

$$M(q_{ij}(t,\tau)) \to p_{ij} \cdot \left(\frac{v_i}{r} \sum_{l=1}^{k} u_l s_{ij}^l + \sum_{l=1}^{k} u_l g_{il} \right),$$

где u, v *— соответственно правый и левый собственные векторы матрицы первых моментов, соответствующие перронову корню,* s_{ij}^l *— число нетерминалов* A_l *в правой части правила* r_{ij}, *а* g_{ij} *— константы, определяемые* *равенствами*

$$b^i_{jn}(t) = u_i g_{jn} r^t + O(r_1^t),\ \ r_1 < r,$$

в которых $b^i_{jn}(t)$ *— вторые моменты грамматики.*

Далее, пусть $S_{ij}(t) = q_{ij}(t,1)+q_{ij}(t,2)+\ldots+q_{ij}(t,t)$, т. е. $S_{ij}(t)$ — число применений правила r_{ij} во всём дереве вывода. Тогда для рассматриваемой разложимой грамматики верна следующая теорема, аналогичная теореме из [4].

Теорема 3. *Пусть матрица первых моментов КС-грамматики имеет вид* (1) *и* $r' \neq r''$. *Тогда* $M(S_{ij}(t)/t) \to p_{ij} \cdot \left(\dfrac{v_i}{r}\sum\limits_{l=1}^{k} u_l s^l_{ij} + \sum\limits_{l=1}^{k} u_l g_{il}\right)$ *при* $t \to \infty$.

Рассмотрим случай кратного перронова корня $r = r' = r''$. Для оценки $Q_X(n)$ нам понадобится следующая

Лемма 6 [6]. *Пусть* s, d *— натуральные,* $n_1, \ldots, n_s$ *— целые неотрицательные числа,* $y = (y_1, \ldots y_s)$. *Тогда*

$$(1-y_1)^{n_1}\ldots(1-y_s)^{n_s} = \sum_{k_1+\ldots+k_s<d, k_i\geqslant 0} C^{k_1}_{n_1}\ldots C^{k_s}_{n_s}(-1)^{k_1+\ldots+k_s} y_1^{k_1}\ldots y_s^{k_s} + R_d(n_1,\ldots,n_s;y),$$

где остаточный член представим в виде

$$R_d(n_1,\ldots,n_s,y) = \sum_{k_1+\ldots+k_s=d, k_i\geqslant 0} (-1)^d \varepsilon_{k_1\ldots k_s}(n_1,\ldots,n_s;y)\cdot y_1^{k_1}\ldots y_s^{k_s},$$

причём

$$0 \leqslant \varepsilon_{k_1\ldots k_s}(n_1,\ldots,n_s;y') \leqslant \varepsilon_{k_1\ldots k_s}(n_1,\ldots,n_s;y) \leqslant C^{k_1}_{n_1}\ldots C^{k_s}_{n_s}$$

при $0 \leqslant y_i \leqslant y'_i \leqslant 1$ $(i = 1,\ldots s)$.

Используя эту лемму при $d = 2$ и выражения для $Q_i(t), i = 1,\ldots,k$ из теоремы 1, имеем

$$Q_X(n) = \prod_{i=1}^{k}(1-Q_i(n))^{x_i} = 1 - \sum_{i=1}^{k} x_i Q_i(n) + O\left(\sum_{i,j} x_i x_j Q_i(n) Q_j(n)\right)$$
$$= 1 - c_0\cdot\left(\sum_{i=1}^{k_1} x_i u'_i nb + \sum_{i=k_1+1}^{k} x_i u''_i\right) r^n(1+o(1)) + O(r^{2n})Y_X(n), \quad (16)$$

где $Y_X(n)=\sum_{i,j=1}^{k} x_i x_j Q_i(n)Q_j(n)/r^{2n} \leqslant O\biggl(\sum_{i,j} x_i x_j n^2\biggr)$.
Иногда нам будет достаточна более грубая оценка, которая является очевидным следствием равенства (16):

$$Q_X(n)=1-O(r^n)Z_X(n), \tag{17}$$

где $Z_X(n)=\sum_{i=1}^{k} x_i Q_i(n)/r^n \leqslant O\left(\sum_i x_i n\right)$. Чтобы вывести формулу для $R_X(n)$, используем доказанную в [4] оценку:

$$R_X(n)=\sum_{i=1}^{k} x_i\cdot(Q_i(n-1)-Q_i(n))+O(r^{2n}Y_X(n)).$$

Пользуясь этим соотношением, мы можем записать, что

$$\begin{aligned}R_X(n)=c_0\cdot(1-r)r^{n-1}\biggl(\sum_{i=1}^{k_1} x_i u_i' bn+\sum_{i=k_1+1}^{k} x_i u''_{i-k_1}\biggr)&\\ \times(1+o(1))+O(r^{2n})Y_X(n).&\end{aligned} \tag{18}$$

Для нахождения математического ожидания числа применений правила на ярусе дерева вывода нам понадобится вид первых и вторых моментов грамматики. Вид первых моментов $a_j^i(t)$ описывается в следствии 1. Для моментов второго порядка справедлива

Лемма 7. *Для КС-грамматики с матрицей первых моментов, имеющей вид* (1) *при* $r=r'=r''$ *и* $t\to\infty$, *вторые моменты* $b_{js}^i(t)$ *имеют следующий вид:*

$b_{js}^i(t)=u_i' g_{js} r^t(1+o(1))$ *при* $i\leqslant k_1$, *если* $j\leqslant k_1$ *либо* $s\leqslant k_1$;
$b_{js}^i(t)=bu_i' g_{js}\cdot tr^t(1+o(1))$ *при* $i\leqslant k_1$ *и* $j,s>k_1$;
$b_{js}^i(t)=u''_{i-k_1} g_{js} r^t(1+o(1))$ *при* $i,j,s>k_1$;
$b_{js}^i(t)=0$ *при* $i>k_1$, *если* $j\leqslant k_1$ *или* $s\leqslant k_1$,
где g_{js} — *некоторые константы*, $g_{js}=g_{sj}$ $(0<j\leqslant k,\ 0<s\leqslant k)$.

Для моментов третьего порядка справедлива

Лемма 8. *Для КС-грамматики с матрицей первых моментов, имеющей вид* (1) *при* $r=r'=r''$ *и* $t\to\infty$, *имеет место оценка* $c_{jkl}^i(t)=O(r^t)$, *кроме случая, когда* $i\leqslant k_1$, *а* $j,k,l>k_1$, *для которого верна оценка* $c_{jkl}^i(t)=O(tr^t)$.

Доказательства этих двух лемм чисто технические и поэтому приведены в приложении.

Исследуем асимптотическое поведение $M(q_{ij}(t,\tau))$ при $\tau \to \infty$ и $t-\tau \to \infty$. Зафиксируем вектор $X = (x_1, \ldots, x_k)$ и найдём вклад деревьев вывода из $D_X^t(\tau)$ в $M(q_{ij}(t,\tau))$. Он равен

$$\Delta_X = \frac{1}{P(D^t)} x_i \cdot (P^I + P^{II}),$$

где $P^I = P_X(\tau) p_{ij} R_S(t-\tau-1) Q_{X'}(t-\tau)$ — суммарная вероятность деревьев из $D_X^t(\tau)$, для которых последний ярус достигается через применение правила r_{ij} к первому по порядку нетерминалу A_i на ярусе τ, и $P^{II} = P_X(\tau) p_{ij} R_{X'}(t-\tau) Q_S(t-\tau-2)$ — суммарная вероятность деревьев из $D_X^t(\tau)$, для которых последний ярус достигается благодаря применению других правил (т. е. поддерево с корнем, помеченным самым левым нетерминалом A_i на ярусе τ, имеет высоту меньше $t-\tau$). Здесь $X' = (x_1, \ldots, x_{i-1}, x_i - 1, x_{i+1}, \ldots, x_k)$ и $S = (s_{ij}^1, \ldots, s_{ij}^k)$, где s_{ij}^l — число нетерминалов A_l в правой части правила r_{ij}.

Сначала рассмотрим случай $i \leqslant k_1$. Найдём $S_1 = R_S(t-\tau-1) Q_{X'}(t-\tau)$. Используя полученные выше оценки (17), (18) для $R_X(n)$ и $Q_X(n)$, получаем

$$S_1 = c_0 \cdot (1-r) r^{t-\tau-2} \times \left(b \cdot (t-\tau-1) \sum_{l=1}^{k_1} u'_l s_{ij}^l + \sum_{l=k_1+1}^{k} u''_{l-k_1} s_{ij}^l \right) \\ \times (1 + O(r^{t-\tau} \cdot Z_X(t-\tau-1)))(1+o(1)),$$

поскольку погрешность для $R_S(t-\tau-1)$ не зависит от X и равна $o(1)$. Если положить $\sum\limits_{l=1}^{k_1} u'_l s_{ij}^l = S'_{ij}$ и $\sum\limits_{l=k_1+1}^{k} u''_{l-k_1} s_{ij}^l = S''_{ij}$, то выражение для S_1 можно записать в виде

$$S_1 = c_0 \cdot (1-r) r^{t-\tau-2} \cdot \left(b \cdot (t-\tau-1) S'_{ij} + S''_{ij} \right) (1+o(1)) + O(r^{t-\tau} \cdot Z_X(t-\tau-1)).$$

Аналогично оценим величину $S_2 = R_{X'}(t-\tau) Q_S(t-\tau-2)$. Очевидно, что

$$S_2 = c_0 \cdot (1-r) r^{t-\tau-1} \times \Bigg((b \cdot (t-\tau) \sum_{l=1}^{k_1} x'_l u'_l + \sum_{l=k_1+1}^{k} x'_l u''_{l-k_1}) \\ + O(r^{t-\tau} \cdot \ Y_X(t-\tau-2)) \Bigg) (1+o(1)),$$

где $x_l' = x_l$ при $l \neq i$ и $x_i' = x_i - 1$. Здесь $Q_S(t-\tau-2)$ оценено как $1+o(1)$ с использованием (17). Подставляя формулы для S_1, S_2 в выражения для P^I, P^{II} и суммируя по X, а также учитывая вытекающие из следствия 1 и лемм 7, 8 оценки

$$\sum_X P_X(\tau) x_i Z_X(t-\tau-1) = O((t-\tau)\cdot r^\tau),$$
$$\sum_X P_X(\tau) x_i Y_X(t-\tau-2) = O((t-\tau)^2\cdot r^\tau)$$

для $i \leqslant k_1$, при $\tau \to \infty$ и $t-\tau \to \infty$ получаем

$$\begin{aligned} M(q_{ij}(t,\tau)) &= \frac{1}{P(D^t)} \sum_X x_i P_X(\tau) p_{ij} \cdot (S_1 + S_2) \\ &= \frac{p_{ij} c_0 \cdot (1-r) r^{t-\tau-2}(1+o(1))}{c_0 b t u_1' (1-r) r^{t-1}} \times \Bigg[\sum_X P_X(\tau) x_i S_{ij}' b \cdot (t-\tau-1) \\ &\quad + \sum_X P_X(\tau) x_i S_{ij}'' + \sum_X P_X(\tau) \cdot \sum_{l=1}^{k_1} u_i' x_i x_l' \cdot br \cdot (t-\tau) \\ &\quad + r \cdot \sum_X P_X(\tau) \cdot \sum_{l=k_1+1}^{k} u_{l-k_1}'' x_i x_l \Bigg]. \end{aligned}$$

Так как

$$\sum_X P_X(\tau) x_i = a_i^1(\tau) = O(r^\tau), \ \sum_X P_X(\tau) x_i x_l = b_{il}^1(\tau) = O(r^\tau),$$

то членами $S_{ij}'' \cdot \sum\limits_X P_X(\tau) x_i$ и $r \cdot \sum\limits_{l=k_1+1}^{k} u_{l-k_1}'' b_{il}^1(\tau)$ можно пренебречь при $t - \tau \to \infty$. Следовательно,

$$M(q_{ij}(t,\tau)) = \frac{p_{ij}}{u_1' t r^{\tau+1}} \left(a_i^1(\tau) \cdot (t-\tau-1) S_{ij}' + br(t-\tau) \cdot \sum_{l=1}^{k_1} u_l' b_{il}^1(\tau) \right) \times (1+o(1)).$$

Преобразуя выражение для $M(q_{ij}(t,\tau))$ и подставляя выражения для первых и вторых моментов из следствия 1 и леммы 7, получаем

$$M(q_{ij}(t,\tau)) = \frac{p_{ij}}{t} \cdot \left(\frac{(t-\tau-1) v_i' S_{ij}'}{r} + \sum_{l=1}^{k_1} u_l' g_{il} \cdot (t-\tau) \right) (1+o(1))$$

$$= \frac{p_{ij} \cdot (t-\tau)}{t} (G'_i + v'_i S'_{ij}/r)(1+o(1)),$$

где $G'_i = \sum\limits_{l=1}^{k_1} u'_l g_{il}$.

Теперь рассмотрим случай $i > k_1$. Поскольку $s^l_{ij} = 0$ при $i > k_1$ и $l \leqslant k_1$, то $S'_{ij} = 0$. Используя оценки

$$\sum_X P_X(\tau) x_i Z_X(t-\tau-1) = O(\tau(t-\tau) \cdot r^\tau),$$

$$\sum_X P_X(\tau) x_i Y_X(t-\tau-2) = O(\tau(t-\tau)^2 \cdot r^\tau),$$

получаем

$$M(q_{ij}(t,\tau)) = \frac{p_{ij} c_0 \cdot (1-r) r^{t-\tau-2}(1+o(1))}{c_0 b t u'_1 (1-r) r^{t-1}} \times \Big[\sum_X P_X(\tau) x_i S''_{ij}$$

$$+ \sum_X P_X(\tau) \cdot \sum_{l=1}^{k_1} u'_i x_i x_l \cdot br \cdot (t-\tau) + r \cdot \sum_X P_X(\tau) \cdot \sum_{l=k_1+1}^{k} u''_{l-k_1} x_i x'_l \Big]$$

$$= \frac{p_{ij}(1+o(1))}{u'_1 t r^{\tau+1}} \times \Big(a^1_i(\tau) \cdot S''_{ij} + br(t-\tau) \cdot \sum_{l=1}^{k_1} u'_l b^1_{il}(\tau)$$

$$+ r \cdot \sum_{l=k_1+1}^{k} u''_{l-k_1} b^1_{il}(\tau) \Big).$$

Используя следствие 1 и лемму 7 для первых и вторых моментов, имеем

$$M(q_{ij}(t,\tau)) = \frac{p_{ij}(1+o(1))}{t} \cdot \Big(\tau \cdot (v''_{i-k_1} S''_{ij}/r + G''_i) + (t-\tau) \cdot G'_i \Big),$$

где $G''_i = \sum\limits_{l=k_1+1}^{k} u''_{l-k_1} g_{il}$. Полученные результаты сформулируем в виде следующего утверждения.

Теорема 4. *Для КС-грамматики с матрицей первых моментов, имеющей вид (1), при $r = r' = r''$, $\tau \to \infty$ и $t - \tau \to \infty$ справедливы равенства: если $i \leqslant k_1$, то*

$$M(q_{ij}(t,\tau)) = \frac{p_{ij} \cdot (t-\tau)}{t} \cdot \left(G'_i + \frac{v'_i S'_{ij}}{r} \right) (1+o(1));$$

если $k_1 < i \leqslant k$, *то*

$$M(q_{ij}(t,\tau)) = \frac{p_{ij}}{t} \cdot \Big(\tau \cdot (v''_{i-k_1} S''_{ij}/r + G''_i) + (t-\tau) \cdot G'_i\Big)\,(1 + o(1)).$$

Здесь $j = 1, \ldots, n_i$, *а* n_i *— число правил грамматики с нетерминалом* A_i *в левой части.*

Таким образом, в случае кратного перронова корня для разложимой грамматики с двумя классами нетерминалов, так же как и в неразложимом случае, $M(q_{ij}(t,\tau))$ состоит из двух слагаемых. Первое слагаемое равно $p_{ij} \cdot (t-\tau) v'_i S'_{ij}/rt$ при $i \leqslant k_1$ или $p_{ij}\tau \cdot v''_{i-k_1} S''_{ij}/rt$ при $i > k_1$ и учитывает соответственно число нетерминалов из класса K_1 при $i \leqslant k_1$ или число нетерминалов из класса K_2 при $i > k_1$ в правой части правила r_{ij}. Второе слагаемое равно $p_{ij} \cdot (t-\tau) G'_i/t$ при $i \leqslant k_1$ или $p_{ij} \cdot (\tau G''_i + (t-\tau) \cdot G'_i)\,/t$ при $i > k_1$ и определяется вторыми моментами грамматики. При этом, как и в неразложимом случае, при $t \to \infty$ правила, порождающие больше нетерминальных символов, используются чаще в выводе слов, что, по-видимому, нужно для достижения большей высоты дерева вывода. Кроме того, величины $M(q_{ij}(t,\tau))$ ограничены сверху константами, поскольку отношения τ/t и $(t-\tau)/t$ не превосходят 1.

Возникают следующие отличия по сравнению с неразложимым случаем:

1) величины $M(q_{ij}(t,\tau))$ зависят линейно от соотношения τ/t,

2) $M(q_{ij}(t,\tau))$ убывают при $i \leqslant k_1$ и увеличении τ (см. пример раздела 4).

Теперь рассмотрим величину $S_{ij}(t)/t$ — среднее число правил r_{ij} на один ярус дерева вывода высоты t. Введём константы $\omega^{(m)}_{ij}$, $m = 1, 2$, $i, j = 1, \ldots, k$ равенствами

$$\begin{aligned} \omega^{(1)}_{ij} &= p_{ij} \cdot \Big(G'_i + v'_i S'_{ij}/r\Big) && \text{при } i \leqslant k_1, \\ \omega^{(1)}_{ij} &= p_{ij} G'_i && \text{при } i > k_1, \\ \omega^{(2)}_{ij} &= p_{ij} \cdot \Big(G''_i + v''_{i-k_1} S''_{ij}/r\Big) && \text{при } i > k_1, \end{aligned} \tag{19}$$

$$\omega_{ij} = \begin{cases} \omega^{(1)}_{ij} & \text{при } i \leqslant k_1, \\ \omega^{(1)}_{ij} + \omega^{(2)}_{ij} & \text{при } i > k_1. \end{cases} \tag{20}$$

Тогда

$$M(q_{ij}(t,\tau)) = \begin{cases} \omega^{(1)}_{ij} \cdot (t-\tau)/t & \text{при } i \leqslant k_1, \\ \omega^{(1)}_{ij} \cdot (t-\tau)/t + \omega^{(2)}_{ij} \cdot \tau/t & \text{при } i > k_1. \end{cases}$$

Теорема 5. *Для КС-грамматики с матрицей первых моментов, имеющей вид* (1), *при* $r = r' = r''$ *и* $t \to \infty$ *справедливо соотношение*

$$M(S_{ij}(t)/t) \to \omega_{ij}/2, \; i = 1, \ldots, k,$$

где величины ω_{ij} *взяты из* (20).

Доказательство. Теорема доказывается аналогично соответствующему утверждению для неразложимой грамматики в докритическом случае [4]. Рассмотрим случай $i \leqslant k_1$. Положим $\tau_0 = \lfloor \log \log t \rfloor$ (здесь и везде далее логарифм рассматривается по основанию 2). Представим $S_{ij}(t)$ в виде $S^1_{ij}(t) + S^2_{ij}(t) + S^3_{ij}(t)$, где

$$\begin{aligned}
S^1_{ij}(t) &= q_{ij}(t,1) + \ldots + q_{ij}(t,\tau_0),\\
S^2_{ij}(t) &= q_{ij}(t,\tau_0+1) + \ldots + q_{ij}(t,t-\tau_0),\\
S^3_{ij}(t) &= q_{ij}(t,t-\tau_0+1) + \ldots + q_{ij}(t,t).
\end{aligned}$$

В [4] (без предположений о неразложимости грамматики) были получены оценка $M(S^1_{ij}(t)) \leqslant \log^c t$, где $c = \log k_{max} + 1$ и k_{max} — максимальное число нетерминалов в правой части правил грамматики, и оценка для $S^3_{ij}(t)$:

$$S^3_{ij}(t) \leqslant \frac{a^1_i(t-\tau_0)}{P(D^t)} = O\left(\frac{1}{t \cdot r^{t-\tau_0}}\right) = \log^{c_1} t$$

при некоторой константе c_1.

Для $S^2_{ij}(t)$ имеем

$$M(S^2_{ij}(t)) = \omega^{(1)}_{ij} \cdot \sum_{\tau=\tau_0}^{t-\tau_0} \left(\frac{t-\tau}{t}\right)(1+o(1)) = \omega^{(1)}_{ij} \cdot \sum_{\tau=1}^{t} \left(\frac{t-\tau}{t}\right)(1+o(1))$$
$$= \omega^{(1)}_{ij} \cdot t/2 + o(t).$$

Пользуясь неравенством

$$M(S^2_{ij}(t)) \leqslant M(S_{ij}(t)) \leqslant M(S^1_{ij}(t)) + M(S^2_{ij}(t)) + M(S^3_{ij}(t)),$$

убеждаемся в том, что при $\tau_0 \to \infty$

$$\frac{1}{t}\left(\frac{t\omega^{(1)}_{ij}}{2}\right)(1+o(1)) \leqslant M(S_{ij}(t)/t) \leqslant \frac{1}{t}\left(\frac{t\omega^{(1)}_{ij}}{2} + O(\log^{c_2} t)\right)(1+o(1)),$$

где $c_2 = \max(c, c_1)$. Таким образом, устремляя τ_0 (а значит, и t) к бесконечности, получаем утверждение теоремы.

В случае $i > k_1$ доказательство проводится аналогично. Единственное отличие заключается в оценке $S_{ij}^2(t)$:

$$M(S_{ij}^2(t)) = \omega_{ij}^{(1)} \cdot \sum_{\tau=\tau_0}^{t-\tau_0} \frac{t-\tau}{t}(1+o(1)) + \omega_{ij}^{(2)} \cdot \sum_{\tau=\tau_0}^{t-\tau_0} \frac{\tau}{t}(1+o(1))$$

$$= \omega_{ij}^{(1)} \cdot \sum_{\tau=1}^{t} \frac{t-\tau}{t}(1+o(1)) + \omega_{ij}^{(2)} \cdot \sum_{\tau=1}^{t} \frac{\tau}{t}(1+o(1)) = (\omega_{ij}^{(1)} + \omega_{ij}^{(2)}) \cdot t/2 + o(t).$$

Теорема 5 доказана.

Эта теорема показывает, что величина $M(S_{ij}(t)/t)$ стремится к константе при $t \to \infty$, как и в неразложимом случае.

4. Пример грамматики с двумя классами нетерминалов

Приведем пример применения формул для математических ожиданий. Рассмотрим грамматику с двумя нетерминалами A_1, A_2, тремя терминалами $y, \bar{x}, x$ и следующими правилами:

$$r_{11} : A_1 \overset{p}{\to} A_1 y A_2, \qquad r_{12} : A_1 \overset{1-p}{\to} \lambda,$$
$$r_{21} : A_2 \overset{p/2}{\to} A_2 \bar{x} A_2 x, \qquad r_{22} : A_2 \overset{1-p/2}{\to} \lambda,$$

где λ — пустое слово. Такая грамматика порождает слова вида $y^n V_1 \dots V_n$, где V_i — правильные скобочные выражения от $\bar{x}, x$, если отождествить $\bar{x}$ с открывающей, а x с закрывающей скобкой, при этом $K_1 = \{A_1\}$, $K_2 = \{A_2\}$. Матрица первых моментов такой грамматики имеет вид

$$A = \begin{pmatrix} p & p \\ 0 & p \end{pmatrix},$$

вторые моменты b_{22}^2 и b_{12}^1 равны p, а все остальные моменты второго и более высоких порядков равны нулю. Легко вычислить вторые моменты грамматики: $b_{12}^1(t) = b_{22}^2(t) = \frac{1}{1-p}(1+o(1))$, $b_{11}^1(t) = b_{22}^1(t) = O(p^t)$. Используя теорему 4, при $\tau \to \infty$ и $t - \tau \to \infty$ получаем

$$M(q_{11}(t,\tau)) \sim \frac{t-\tau}{t}, \qquad M(q_{12}(t,\tau)) \to 0,$$
$$M(q_{21}(t,\tau)) \sim \frac{\tau}{t} + \frac{p}{2-2p}, \qquad M(q_{22}(t,\tau)) \to \frac{p}{2-2p}.$$

Математическое ожидание числа применений правила r_{11} убывает с ростом τ. Математическое ожидание числа применений правила r_{12} на один ярус дерева вывода равно нулю, поскольку это правило применяется только один раз в выводе любого слова. Математическое ожидание числа применений правила r_{21} линейно возрастает, а r_{22} асимптотически постоянно.

5. Оценка стоимости оптимального кодирования

В этом разделе будем рассматривать языки, порождённые КС-грамматиками, порождёнными грамматиками с однозначным выводом. КС-грамматика называется КС-грамматикой с *однозначным выводом,* если каждому слову в терминальном алфавите, выводимому из аксиомы, соответствует единственное дерево вывода.

Рассмотрим стохастический КС-язык $\mathcal{L} = (L, P)$, порождённый грамматикой с однозначным выводом. Под кодированием языка $\mathcal{L}$ будем понимать инъективное отображение $f : L \to \{0,1\}^+$. Образ слова при отображении f назовём его *кодом.* Множество всех кодирований языка $\mathcal{L}$ обозначим через $F(\mathcal{L})$.

Под энтропией языка $\mathcal{L}$ будем понимать величину

$$H(\mathcal{L}) = -\lim_{N\to\infty} \sum_{\alpha\in L, |\alpha|\leqslant N} p(\alpha)\cdot \log p(\alpha),$$

где через $p(\alpha)$ обозначается вероятность слова α. Если энтропия конечна, то будем писать $H(\mathcal{L}) = -\sum\limits_{\alpha\in L} p(\alpha)\cdot\log p(\alpha)$.

Через L^t обозначим множество слов языка L, деревья вывода которых имеют высоту t. Через $p_t(\alpha)$ обозначим условную вероятность слова α в множестве слов L^t. Очевидно, она равна $p(\alpha)/P(L^t)$. Через $\mathcal{L}^t = (L^t, p_t)$ обозначим множество слов L^t с определённым на нём выше распределением вероятностей. Стоимостью кодирования $f \in F(\mathcal{L})$ назовём величину

$$C(\mathcal{L}, f) = \lim_{t\to\infty} \frac{\sum_{\alpha\in L^t} p_t(\alpha)\cdot|f(\alpha)|}{\sum\limits_{\alpha\in L^t} p_t(\alpha)\cdot|\alpha|},$$

если предел существует. Через $F_*(\mathcal{L})$ обозначим множество всех кодирований, для которых величина $C(\mathcal{L}, f)$ определена. *Стоимостью* оптимального кодирования назовём величину $C^*(\mathcal{L}) = \inf\limits_{f\in F_*(\mathcal{L})} C(\mathcal{L}, f)$.

Сначала получим асимптотическую формулу для энтропии $H(\mathcal{L}^t)$ множества слов $\mathcal{L}^t$. В дальнейшем мы рассматриваем только случай кратного перронова корня, поскольку иначе оценки для $H(\mathcal{L}^t)$ и $C^*(\mathcal{L})$, как легко заметить, имеют такой же вид, как и в неразложимом случае. Поэтому в дальнейшем предполагаем, что матрица первых моментов грамматики имеет вид (1) и $r = r' = r''$.

По определению имеем $H(\mathcal{L}^t) = -\sum_{\alpha \in L^t} p_t(\alpha) \cdot \log p_t(\alpha)$. Следовательно,

$$H(\mathcal{L}^t) = -\sum_{\alpha \in L^t} p_t(\alpha)(\log p(\alpha) - \log P(L^t))$$
$$= \frac{1}{P(L^t)} \cdot \left(-\sum_{\alpha \in L^t} p(\alpha) \log p(\alpha) \right) + \log P(L^t).$$

Для слова α обозначим через $q_{ij}(\alpha)$ число применений правила r_{ij} при его выводе. Вероятность слова α равна $p(\alpha) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} p_{ij}^{q_{ij}(\alpha)}$. Следовательно, $\log p(\alpha) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} q_{ij}(\alpha) \cdot \log p_{ij}$. Поэтому

$$H(\mathcal{L}^t) = \frac{1}{P(L^t)} \cdot \left(-\sum_{\alpha \in L^t} p(\alpha) \cdot \sum_{i=1}^{k} \sum_{j=1}^{n_i} q_{ij}(\alpha) \cdot \log p_{ij} \right) + \log P(L^t)$$
$$= \frac{1}{P(L^t)} \cdot \left(-\sum_{i=1}^{k} \sum_{j=1}^{n_i} \log p_{ij} \cdot \sum_{\alpha \in L^t} p(\alpha) \cdot q_{ij}(\alpha) \right) + \log P(L^t).$$

Очевидно, что $\sum_{\alpha \in L^t} p(\alpha) \cdot q_{ij}(\alpha) = P(L^t) \cdot M(S_{ij}(t))$. По теореме 4 имеем $M(S_{ij}(t)) = t \cdot \omega_{ij}/2 + o(t)$, где ω_{ij} определены в (20). Поэтому выражение для энтропии можно переписать в виде

$$H(\mathcal{L}^t) = -t \cdot (1 + o(1)) \cdot \sum_{i=1}^{k} \sum_{j=1}^{n_i} \omega_{ij} \log p_{ij}/2 + \log P(L^t).$$

Ввиду однозначности вывода имеем $\log P(L^t) = \log P(D^t) = t \cdot \log r + O(\log t)$. Поэтому

$$H(\mathcal{L}^t) = t \cdot \left(\log r - \sum_{i=1}^{k} \sum_{j=1}^{n_i} \omega_{ij} \log p_{ij}/2 \right) + o(t). \tag{21}$$

Полученный результат сформулируем в виде теоремы.

Теорема 6. *Энтропия множества слов с деревьями вывода высоты*

t из стохастического КС-языка, порождённого разложимой КС-грамматикой с матрицей первых моментов вида (1) при $r = r' = r''$, имеет вид

$$H(\mathcal{L}^t) = t \cdot \left(\log r - \sum_{i=1}^{k} \sum_{j=1}^{n_i} \omega_{ij} \log p_{ij}/2 \right) + o(t).$$

Здесь величины ω_{ij} определены формулами (20), а p_{ij} — вероятность применения правила r_{ij}.

Таким образом, энтропия $H(\mathcal{L}^t)$ асимптотически линейно зависит от высоты дерева t как и в неразложимом случае.

Теперь оценим стоимость оптимального кодирования для слов КС-языка, порождённого рассматриваемой грамматикой. Обозначим через f^* кодирование множества слов $\mathcal{L}^t$, минимизирующее величину

$$M_t(f) = \sum_{\alpha \in L^t} p_t(\alpha) \cdot |f(\alpha)|.$$

Тогда, очевидно, для любого кодирования множества слов $L^t \subset L$ верно неравенство $M_t(f) \geqslant M_t(f^*)$. Величину $M^*(\mathcal{L}^t) = M_t(f^*)$ оценим, используя доказанную в [2] теорему.

Теорема 7. *Если последовательность стохастических языков $\mathcal{L}_k$ такова, что $H(\mathcal{L}_k) \to \infty$ при $k \to \infty$, то $\lim\limits_{k\to\infty} M^*(\mathcal{L}_k)/H(\mathcal{L}_k) = 1$.*

Поскольку $H(\mathcal{L}^t) \to \infty$ при $t \to \infty$, то из этой теоремы следует, что $M_t(f^*)/H(\mathcal{L}^t) \to 1$ при $t \to \infty$.

Теперь вычислим величину $\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha|$. Пусть правило r_{ij} в правой части содержит l_{ij} терминальных символов. Очевидно, что $|\alpha| = \sum\limits_{i=1}^{k} \sum\limits_{j=1}^{n_i} q_{ij}(\alpha) l_{ij}$. Поэтому

$$\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha| = \sum_{i=1}^{k} \sum_{j=1}^{n_i} l_{ij} M(S_{ij}(t)) = t \cdot \sum_{i,j} l_{ij} \omega_{ij}/2 + o(t).$$

Следовательно, справедлива

Теорема 8. *Пусть КС-язык $\mathcal{L}$ порождён разложимой грамматикой с матрицей первых моментов вида (1) с $r = r' = r''$. Тогда стоимость любого кодирования f этого языка удовлетворяет неравенству*

$$C(\mathcal{L}, f) \geqslant C^*(\mathcal{L}) = \frac{2 \log r - \sum\limits_{i,j} \omega_{ij} \log p_{ij}}{\sum\limits_{i,j} l_{ij} \omega_{ij}},$$

где ω_{ij} *определены в* (20), p_{ij} *— вероятность применения правила* r_{ij}, *а* l_{ij} *— число терминальных символов в правой части правила* r_{ij}.

Можно показать, что эта оценка является точной, т. е. существует кодирование слов из $\mathcal{L}^t$, стоимость которого сколь угодно близка к $C^*(\mathcal{L})$. Для этого будем использовать схему кодирования, предложенную в [5]. Как доказано выше, частота правила r_{ij} в деревьях вывода высоты $t \to \infty$ стремится к $p'_{ij} = \omega_{ij}/\omega_i$, где $\omega_i = \sum\limits_j \omega_{ij}$. Для каждого множества правил R_i с одинаковой левой частью A_i применяется схема двоичного префиксного кодирования Шеннона. При этом правилу r_{ij} будет соответствовать элементарный код длины $\lfloor -\log p'_{ij} \rfloor$.

Код слова $\alpha \in L^t$, имеющего левый вывод $\omega(\alpha) = r_{i_1 j_1} r_{i_2 j_2} \dots r_{i_m j_m}$, получается конкатенацией кодов правил $r_{i_s j_s}$. Такое кодирование обозначим через f_{sh}. По аналогии с [5] доказывается, что стоимость кодирования f_{sh} стремится к $C^*(\mathcal{L})$ при описанном в этой статье укрупнении правил грамматики.

6. Приложение. Доказательства лемм 7 и 8

Доказательство леммы 7. Для вторых моментов в [6] имеется формула

$$b^i_{jn}(t) = \sum_{\tau=1}^{t} \sum_{l,m,s} a^i_l(t-\tau) \cdot b^l_{ms} a^m_j(\tau-1) a^s_n(\tau-1). \qquad (22)$$

Поскольку $A(t) = a^i_j(t) = A^t$, а в [3] доказано, что в случае, когда перронов корень r матрицы A имеет кратность два, верно равенство

$$A^t = (A_0 t + A_1) r^t + \sum_{q=1}^{h} \phi^i_{jq}(t) r^t_q,$$

где r_q, $1 \leqslant q \leqslant h$, — собственный корень, отличный от r, $|r_q| < r$ и $\varphi^i_{jq}(t)$ — полином, степень которого меньше кратности корня r_q. В частности, для собственного корня r этот полином выписан здесь в виде $A_0 t + A_1$, где A_0, A_1 — квадратные матрицы порядка k. При этом из леммы 2 следует, что

$$A^i_{0j} = \begin{cases} b u'_i v''_{j-k_1}/r & \text{при } i \leqslant k_1 \text{ и } j > k_1, \\ 0 & \text{в остальных случаях,} \end{cases}$$

$$A^i_{1j} = \begin{cases} u'_i v'_j & \text{при } i, j \leqslant k_1, \\ u''_{i-k_1} v''_{j-k_1} & \text{при } i, j > k_1, \\ 0 & \text{при } i > k_1 \text{ и } j \leqslant k_1. \end{cases}$$

Мы ограничимся рассмотрением случая (3): $j > k_1$, $n > k_1$ и $i \leqslant k_1$; остальные случаи рассматриваются аналогично. Преобразуя правую часть равенства (22), получаем

$$b^i_{jn}(t) = \sum_{\tau=1}^{t} \sum_{l\leqslant k_1, 0<m,s\leqslant k} a^i_l(t-\tau)\cdot b^l_{ms} a^m_j(\tau-1) a^s_n(\tau-1)$$
$$+ \sum_{\tau=1}^{t} \sum_{l,m,s>k_1} a^i_l(t-\tau)\cdot b^l_{ms} a^m_j(\tau-1) a^s_n(\tau-1).$$

Остальные слагаемые равны нулю, поскольку $a^i_j(t) = 0$ при $i > k_1, j \leqslant k_1$, и $b^i_{js}(t) = 0$ при $i > k_1$ и $j \leqslant k_1$ или $s \leqslant k_1$. Первое слагаемое оценивается как

$$\sum_{\tau=1}^{t} \sum_{l\leqslant k_1, m,s} a^i_l(t-\tau)\cdot b^l_{ms} a^m_j(\tau-1) a^s_n(\tau-1) = O\left(\sum_{\tau=1}^{t} A^i_{1j}\tau^2\cdot r^{t-\tau} r^{2\tau}\right) = O(r^t).$$

Здесь использовался тот факт, что если $0 < r < 1, n \geqslant 0$ и $t \to \infty$, то

$$\sum_{k=n}^{t} C^n_k r^{k-n} = \frac{1+O(t^n r^t)}{(1-r)^n}. \tag{23}$$

Второе слагаемое можно переписать в виде

$$\sum_{\tau=1}^{t} \sum_{l,m,s>k_1} a^i_l(t-\tau)\cdot b^l_{ms} a^m_j(\tau-1) a^s_n(\tau-1)$$
$$= \sum_{\tau=1}^{t} \sum_{l,m,s>k_1} [bu'_i v''_{j-k_1}\cdot tr^{t-\tau-1} + O(r^t)] b^l_{ms}$$
$$\times \left[A^m_{1j} r^{\tau-1} + \sum_{q=1}^{h} \phi^m_{jq}(\tau-1) r_q^{\tau-1}\right] \times \left[A^s_{1n} r^{\tau-1} + \sum_{q=1}^{h} \phi^s_{nq}(\tau-1) r_q^{\tau-1}\right]$$
$$= tr^{t-1} bu'_i \cdot \sum_{\tau=1}^{t} \sum_{l,m,s>k_1} v''_{j-k_1} b^l_{ms} r^{-1} \times \left[A^m_{j1} + \sum_{q=1}^{h} \phi^m_{jq}(\tau-1) r_q^{\tau-1}\right]$$
$$\times \left[A^s_{1n} r^{\tau-1} + \sum_{q=1}^{h} \phi^s_{nq}(\tau-1) r_q^{\tau-1}\right] + O(r^t).$$

Обозначив через g_{jn} сумму сходящегося ряда

$$\sum_{\tau=1}^{t}\sum_{l,m,s>k_1} v''_{j-k_1} b^l_{ms} r^{-1} \times \left[A^m_{1j} + \sum_{q=1}^{h} \phi^m_{jq}(\tau-1) r_q^{\tau-1}\right] \times \left[A^s_{1n} r^{\tau-1} + \sum_{q=1}^{h} \phi^s_{nq}(\tau-1) r_q^{\tau-1}\right],$$

получаем утверждение леммы 7.

Доказательство леммы 8. Рассмотрим третьи моменты $c^i_{jkl}(t)$. Очевидно, что $c^i_{jkl}(t) = 0$ при $i > k_1$, если по крайней мере один из индексов j, k, l не превосходит k_1. В последнем случае $c^i_{jkl}(t)$ имеет вид $O(r^t)$ (для доказательства достаточно рассмотреть неразложимую грамматику, порождённую классом K_2).

Рассмотрим случай $i \leqslant k_1$, так как при $i > k_1$ соответствующие моменты являются моментами для неразложимой грамматики, порождённой нетерминалами из K_2. В [6] доказано соотношение

$$c^i_{jkl}(t) = \sum_{\tau=1}^{t}\sum_{s} a^i_s(t-\tau) z^s_{jkl}(\tau-1),$$

где $z^s_{jkl}(t) = c^s_{jkl}(t+1) - \sum_n a^s_n c^n_{jkl}(t)$. Преобразуя, получаем

$$c^i_{jkl}(t) = \sum_{\tau=1}^{t}\sum_{0<s\leqslant k_1} a^i_s(t-\tau) z^s_{jkl}(\tau-1) + \sum_{\tau=1}^{t}\sum_{k_1<s<k} a^i_s(t-\tau) z^s_{jkl}(\tau-1).$$

Поэтому если хотя бы один индекс j, k, l не превосходит k_1, то при $s > k_1$

$$z^s_{jkl}(t) = c^s_{jkl}(t+1) - \sum_n a^s_n c^n_{jkl}(t) = 0.$$

При этом $z^s_{jkl}(\tau-1)$ является суммой произведений моментов более низких порядков : $C \cdot a^m_j(\tau-1) a^n_k(\tau-1) a^q_l(\tau-1)$ и $C \cdot a^m_j(\tau-1) b^n_{kl}(\tau-1)$, где C — некоторая константа. Подставляя в сумму уже полученные выражения для первых и вторых моментов и применяя формулы (23), убеждаемся, что первая сумма равна $O(r^t)$. Оценим вторую сумму. Пусть $i \leqslant k_1$ и $s > k_1$. Тогда если по крайней мере один из индексов j, k, l не превосходит k_1, то величина $a^i_s z^s_{jkl}(t)$ равна 0. Если каждый индекс j, k, l больше k_1, то член $a^i_s z^s_{jkl}(t)$ равен $O(t \cdot r^t)$.

ЛИТЕРАТУРА

1. **Ахо А., Ульман Дж.** Теория синтаксического анализа, перевода и компиляции. Том 1. М.: Мир, 1978.
2. **Борисов А. Е.** Кодирование слов стохастического КС-языка, порождённого разложимой грамматикой с двумя нетерминалами // Вестник Нижегородского университета им. Н. И. Лобачевского. Сер. Математика. 2004. Вып. 1(2). С. 18–28.
3. **Гантмахер Ф. Р.** Теория матриц. М.: Наука, 1967.
4. **Жильцова Л. П.** Закономерности применения правил грамматики в выводах слов стохастического контекстно-свободного языка // Математические вопросы кибернетики. Вып.9. М.: Наука, 2000. С.101–126.
5. **Жильцова Л. П.** О нижней оценке стоимости кодирования и асимптотически оптимальном кодировании стохастического контекстно-свободного языка // Дискрет. анализ и исслед. операций. Сер. 1. 2001. Т. 8, № 3. С. 26–45.
6. **Севастьянов В. А.** Ветвящиеся процессы. М.: Наука, 1971.
7. **Фихтенгольц Г. М.** Основы математического анализа. Том 2. М.: Наука, 1968.
8. **Фу К.** Структурные методы в распознавании образов. М.: Мир, 1977.
9. **Шеннон К.** Математическая теория связи. М.: ИЛ, 1963.

Адрес автора:

ул. Лопатина, д. 3, кв. 199,
603163 Нижний Новгород,
Россия.
E-mail: abor1@rambler.ru,
alexander.borisov@intel.com

Статья поступила
8 декабря 2003 г.
Окончательный вариант —
28 июня 2005 г.