О свойствах вероятностей деревьев вывода в разложимых стохастических КС-грамматиках. Докритический случай

Л. П. Жильцова

20 марта 2010 г.

Автором в [?, ?] рассматривались вопросы, связанные с кодированием сообщений, являющихся словами стохастического контекстно-свободного языка (стохастического КС-языка), при условии, что матрица первых моментов грамматики неразложима, непериодична, и ее максимальный по модулю собственный корень (перронов корень) строго меньше единицы (докритический случай). При неразложимой матрице первых моментов нетерминальные символы грамматики образуют один класс.

В настоящей работе рассматриваются стохастические КС-грамматики с произвольным числом классов нетерминальных символов без ограничений на порядок следования классов.

1 Предварительные сведения

Для изложения результатов о контекстно-свободных языках будем использовать определения КС-языка и стохастического КС-языка из [?, ?].

Стохастической КС-грамматикой называется система $G = \langle V_T, V_N, R, s\rangle$, где V_T и V_N - конечные множества терминальных и нетерминальных символов (терминалов и нетерминалов) соответственно; $s \in V_N$ - аксиома, R – множество правил. Множество R можно представить в виде $R = \cup_{i=1}^{k} R_i$, где k - мощность алфавита V_N и $R_i = \{r_{i1}, \ldots, r_{i,n_i}\}$. Каждое правило r_{ij} из R_i имеет вид

$$r_{ij} : A_i \xrightarrow{p_{ij}} \beta_{ij},\ j = 1, ..., n_i,$$

где $A_i \in V_N, \beta_{ij} \in (V_T \cup V_N)^*$ и p_{ij} - вероятность применения правила r_{ij} (вероятность правила r_{ij}), которая удовлетворяет следующим условиям:

$$0 < p_{ij} \leq 1 \text{ и } \sum_{j=1}^{n_i} p_{ij} = 1.$$

Применение правила грамматики к слову в алфавите $V_T \cup V_N$ состоит в замене вхождения нетерминала из левой части правила на слово, стоящее в его правой части. КС-язык определяется как множество всех слов в алфавите V_T, выводимых из аксиомы s с помощью конечного числа применений правил грамматики.

Каждому слову α КС-языка соответствует последовательность правил грамматики (вывод), с помощью которой α выводится из аксиомы s. Вероятность вывода определяется как произведение вероятностей правил, образующих вывод.

Дерево вывода строится по левому выводу слова следующим образом. Корень дерева помечается аксиомой s. Пусть при выводе слова α на очередном шаге в процессе левого вывода применяется правило $A_i \stackrel{p_{ij}}{\to} b_{i_1} b_{i_2} \ldots b_{i_m}$, где $b_{i_l} \in V_N \cup V_T$ $(l = 1, \ldots, m)$. Тогда из самой левой вершины-листа дерева, помеченной символом A_i (при обходе листьев дерева слева направо), проводится m дуг в вершины следующего яруса, которые помечаются слева направо символами $b_{i_1}, b_{i_2}, \ldots, b_{i_m}$ соответственно. После построения дуг и вершин для всех правил грамматики в выводе слова языка все листья дерева помечены терминальными символами и само слово получается при обходе листьев дерева слева направо. Высотой дерева называется максимальная длина пути от корня к листу.

Важной характеристикой стохастической КС-грамматики является матрица первых моментов, которая строится по грамматике.

Рассмотрим многомерные производящие функции

$$F_i(s_1, s_2, \ldots, s_k),\ i = 1, \ldots, k,$$

где переменная s_i соответствует нетерминальному символу A_i [?]. Функция $F_i(s_1, s_2, \ldots, s_k)$ строится по множеству правил R_i с одинаковой левой частью A_i следующим образом.

Для каждого правила $A_i \stackrel{p_{ij}}{\to} \beta_{ij}$ выписывается слагаемое

$$q_{ij=}p_{ij} \cdot s_1^{l_1} \cdot s_2^{l_2} \cdot \ldots \cdot s_k^{l_k},$$

где l_m - число вхождений нетерминального символа A_m в правую часть правила $(m = 1, \ldots, k)$. Тогда

$$F_i(s_1, s_2, \ldots, s_k) = \sum_{j=1}^{n_i} q_{ij}.$$

Пусть

$$a_j^i = \frac{\partial F_i(s_1, \ldots, s_k)}{\partial s_j} \mid_{s_1 = s_2 = \ldots = s_k = 1.}$$

Квадратная матрица A порядка k, образованная элементами a_j^i, называется *матрицей первых моментов* грамматики G.

Так как матрица A неотрицательна, существует максимальный по модулю действительный неотрицательный собственный корень (перронов корень) [3]. Обозначим этот корень через r.

В работе рассматривается докритический случай, т.е. случай, когда $r < 1$. Основные результаты относятся к стохастическим КС-грамматикам с разложимой матрицей [Гантмахер] первых моментов.

Введем некоторые обозначения. Будем говорить, что нетерминал A_j непосредственно следует за нетерминалом A_i, (и обозначать $A_i \to A_j$), если в грамматике существует правило вида $A_i \stackrel{p_{il}}{\to} \alpha_1 A_j \alpha_2$, где $\alpha_1, \alpha_2 \in (V_T \cup V_N)^*$. Рефлексивное транзитивное замыкание отношения "$\to$"обозначим $\to_*$.

Грамматика называется неразложимой, если для любых двух различных нетерминалов A_i и A_j верно $A_i \to_* A_j$. В противном случае она называется разложимой. Классом нетерминалов назовем максимальное по включению подмножество $K \in V_N$, такое, что $A_i \to_* A_j$ для любых $A_i, A_j \in K$.

Для различных классов K_1 и K_2 будем говорить, что класс K_2 непосредственно следует за классом K_1 (и обозначать $K_1 \prec K_2$), если существуют $A_1 \in K_1$ и $A_2 \in K_2$, такие, что $A_1 \to A_2$. Рефлексивное транзитивное замыкание отношения $\prec$ обозначим через $\prec_*$ и назовем отношением следования.

Очевидно, множество классов нетерминалов является разбиением множества V_N и отношение $\prec$ устанавливает на множестве классов нетерминалов частичный порядок.

Будем полагать, что классы нетерминалов перенумерованы таким образом, что $K_i \prec K_j$ тогда и только тогда, когда $i < j$.

Соответствующая разложимой грамматике матрица первых моментов A имеет следующий вид:

$$A = \begin{pmatrix} A_{11} & A_{12} & \dots & A_{1m-1} & A_{1m} \\ 0 & A_{22} & \dots & A_{2m-1} & A_{2m} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{m-1m-1} & A_{m-1m} \\ 0 & 0 & \dots & 0 & A_{mm} \end{pmatrix}. \tag{1}$$

Один класс нетерминалов в матрице первых моментов представлен множеством подряд идущих строк и соответствующим множеством столбцов с теми же номерами. Для класса K_i квадратная подматрица, образованная соответствующими строками и столбцами, обозначается через A_{ii}. Блоки, расположенные ниже главной диагонали, нулевые в силу упорядоченности классов нетерминалов. Подматрица A_{ij} является нулевой, если $K_i \not\prec K_j$.

Для каждого класса K_i матрица A_{ii} неразложима. Без ограничения общности будем считать, что она строго положительна и непериодична. Этого всегда можно добиться, применяя метод укрупнения правил грамматики, описанный в [мат.вопросы киб.]

Обозначим через r_i перронов корень матрицы A_{ii}. Для неразложимой матрицы перронов корень является действительным и простым [Гантмахер]. Очевидно, в силу структуры матрицы первых моментов, $r = \max_i\{r_i\}$, и $r > 0$, так как $r_i > 0$ для любого i ввиду положительности матрицы A_{ii}.

Пусть $J = \{i_1, i_2, \dots, i_l\}$ — множество всех номеров i_j классов, для которых $r_{i_j} = r$. Назовем J определяющим множеством.

Зафиксируем пару (l, h), $l, h \in \{1, 2, \dots, m\}$, и рассмотрим всевозможные последовательности классов $K_{i_1} \prec K_{i_2} \prec \dots \prec K_{i_s}$, где $i_1 = l, i_s = h$. Среди всех таких последовательностей выберем ту, которая содержит наибольшее число классов с номерами из J. Это число обозначим через s_{lh}.

Дополнительно переупорядочим классы по неубыванию величины s_{1l}, причем при одинаковых значениях s_{1l} сначала поставим классы с номерами из множества J.

Разобьем последовательность классов на группы классов $\mathcal{M}_1, \mathcal{M}_2, \dots, \mathcal{M}_w$, при этом класс K_l отнесем к группе $\mathcal{M}_1$ при $s_{1l} \le 1$, и к группе $\mathcal{M}_j$ при $s_{1l} = j$ $(j > 1)$.

Для групп $\mathcal{M}_i$ и $\mathcal{M}_j$ определим s^*_{ij} как $\max_{K_l \in \mathcal{M}_i, K_h \in \mathcal{M}_j}\{s_{lh}\}$. Положим $s^*_{11} = 1$.

Среди последовательностей $K_{i_1} \prec K_{i_2} \prec \dots \prec K_{i_s}$, где $i_1 = l$ и i_s принимает всевозможные значения, выберем ту, которая содержит наибольшее число классов с номерами из J. Это число обозначим через q_l.

Отметим, что для любого $K_i \in \mathcal{M}_j$ при $q_i > 0$ найдется такой класс $K_l \in \mathcal{M}_{j+1}$, что $K_i \prec_* K_l$. Для группы $\mathcal{M}_j$ через q_j^* обозначим $\max_{K_i \in \mathcal{M}_j}\{q_i\}$.

Матрицу первых моментов будем также представлять в виде

$$A = \begin{pmatrix} B_{11} & B_{12} & \dots & B_{1w} \\ 0 & B_{22} & \dots & B_{2w} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & B_{ww} \end{pmatrix}, \tag{2}$$

где B_{lh} – подматрица на пересечении строк для классов из группы $\mathcal{M}_l$ и столбцов для классов из $\mathcal{M}_h$. Очевидно, каждая матрица B_{ll} имеет перронов корень r. Запись $B_{lh}^{(t)}$ будем применять для обозначения соответствующей подматрицы матрицы A^t.

Теорема 1 [Казань]. *При* $t \to \infty$

$$B_{lh}^{(t)} = H_{lh} \cdot t^{s_{lh}^* - 1} r^t (1 + o(1)),$$

где H_{lh} *– матрица, не зависящая от* t.

Рассмотрим более подробно некоторые необходимые в дальнейшем свойства матриц B_{lh} и $B_{lh}^{(t)}$ при $t \to \infty$.

Группу $\mathcal{M}_1$ разобьем на три подгруппы. К первой подгруппе $\mathcal{M}_{11}$ отнесем классы нетерминалов с $q_l = 0$, ко второй подгруппе $\mathcal{M}_{12}$ — классы с номерами из множества J, к третьей подгруппе $\mathcal{M}_{13}$ — все остальные классы. Подгруппа $\mathcal{M}_{11}$ может быть пустой в том случае, если для класса K_1 матрица A_{11} имеет перронов корень r. Каждая следующая группа $\mathcal{M}_h$ в силу упорядоченности классов начинается с класса с номером из J. Поэтому $\mathcal{M}_h$ при $h > 1$ разобьем на две подгруппы $\mathcal{M}_{h2}$ и $\mathcal{M}_{h3}$, где $\mathcal{M}_{h2}$ содержит классы с номерами из J, и $\mathcal{M}_{h3}$ — все остальные классы. Для единообразия для $\mathcal{M}_h$ будем рассматривать пустую подгруппу $\mathcal{M}_{h1}$.

В соответствии с этим разбиением B_{ll} представим в виде

$$B_{ll} = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix}, \tag{3}$$

где C_{ij} — подматрица со строками для классов из $\mathcal{M}_{li}$ и столбцами для классов из $\mathcal{M}_{lj}$.

Так как каждому классу K_i из $\mathcal{M}_{l1}$ или из $\mathcal{M}_{l3}$ соответствует перронов корень $r_i < r$, для C_{11}^t и C_{33}^t справедливы оценки $C_{11}^t = o(r^t)$ и $C_{33}^t = o(r^t)$.

Пусть $\mathcal{M}_{l2}$ содержит j_2 классов. Любому классу K_i из $\mathcal{M}_{l2}$ соответствует неразложимая подматрица A_{ii} в представлении (1) и классы из $\mathcal{M}_{l2}$ попарно несравнимы. Поэтому в силу свойств неразложимых матриц [Севаст], матрица C_{22}^t имеет вид

$$\begin{pmatrix} U_{j_1+1}V_{j_1+1} & 0 & \dots & 0 & 0 \\ 0 & U_{j_1+2}V_{j_1+2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & U_{j_1+j_2}V_{j_1+j_2} \end{pmatrix} \cdot r^t (1 + o(1)), \tag{4}$$

где U_i и V_i – правый и левый собственные положительные векторы матрицы A_{ii}, соответствующие r, при нормировке $V_i \cdot U_i = 1$, $i = j_1+1, \ldots, j_1+j_2$.

Обозначим матрицу

$$\begin{pmatrix} U_{j_1+1}V_{j_1+1} & 0 & \ldots & 0 & 0 \\ 0 & U_{j_1+2}V_{j_1+2} & \ldots & 0 & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & 0 & U_{j_1+j_2}V_{j_1+j_2} \end{pmatrix}$$

через D. Очевидно, D можно представить в виде $\sum_{i=j_1+1}^{j_2} U_i^{(2l)} V_i^{(2l)}$, где

$$U_i^{(2l)} = (0, \ldots, 0, U_i, 0, \ldots, 0)^T, \;\; V_i^{(2l)} = (0, \ldots, 0, V_i, 0, \ldots, 0), \tag{5}$$

и U_i и V_i расположены на местах, соответствующих классу K_i. Заметим, что $U_i^{(2l)}$ и $V_i^{(2l)}$ являются соответственно правым и левым собственными корнями матрицы C_{22} для корня r.

Каждому классу K_i соответствуют правый и левый собственные вектора всей матрицы B_{ll}, соответствующие r.

Компоненты правого собственного вектора представим в виде $U_i^{(l)} = (U_i^{(1l)}, U_i^{(2l)}, U_i^{(3l)})^T$, где $U_i^{(jl)}$ соответствует $\mathcal{M}_{lj}$, $j = 1, 2, 3$. В [Казань] установлено, что $U_i^{(3l)} = 0$ и

$$U_i^{(1l)} = (rE - C_{11})^{-1} C_{12} U_i^{(2l)}.$$

Компоненты левого собственного вектора матрицы B_{ll} для перронова корня r представим в виде $V_i^{(l)} = (V_i^{(1l)}, V_i^{(2l)}, V_i^{(3l)})$. В [Казань] также показано, что $V_i^{(1l)} = 0$ и

$$V_i^{(3l)} = V_i^{(2l)} C_{23} (rE - C_{33})^{-1}.$$

Используя описанные вектора, уточним вид матрицы B_{ll} [Казань]:

$$B_{ll}^t \sim \begin{pmatrix} 0 & \sum_{i=j_1+1}^{j_2} U_i^{(1l)} V_i^{(2l)} & \sum_{i=j_1+1}^{j_2} U_i^{(1l)} V_i^{(3l)} \\ 0 & \sum_{i=j_1+1}^{j_2} U_i^{(2l)} V_i^{(2l)} & \sum_{i=j_1+1}^{j_2} U_i^{(2l)} V_i^{(2l)} \\ 0 & 0 & 0 \end{pmatrix} \cdot r^t. \tag{6}$$

Отметим, что строки матрицы B_{ll}, соответствующие классам $K_i \in \mathcal{M}_{l2}$, т.е. классам, для которых $i \in J$, пропорциональны компонентам правого собственного вектора $U_i^{(2l)}$, а столбцы, соответствующие классам $K_j \in \mathcal{M}_{l2}$, пропорциональны компонентам левого собственного вектора $V_j^{(2l)}$.

Рассмотрим случай $l \neq h$. Матрицу B_{lh} представим в виде

$$B_{lh} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \\ D_{31} & D_{32} \end{pmatrix}, \tag{7}$$

где разбиение по строкам сделано в соответствии с подгруппами группы $\mathcal{M}_l$, а по столбцам – в соответствии с подгруппами группы $\mathcal{M}_h$. В [Казань] доказано следующее асимптотическое равенство:

$$B_{lh}^{(t)} \sim \begin{pmatrix} \sum_j U_j'^{(1l)} V_j^{(2h)} & \sum_j U_j'^{(1l)} V_j^{(3h)} \\ \sum_j U_j'^{(2l)} V_j^{(2h)} & \sum_j U_j'^{(2l)} V_j^{(3h)} \\ 0 & 0 \end{pmatrix} \cdot t^{s_{lh}^*-1} r^t,$$

где $U_j^{\prime(sl)} = \frac{1}{s^*_{lh}-1} \sum_i \left(V_i^{(2l)} D_{21} + V_i^{(3l)} D_{31} \right) U_j^{(2h)} U_i^{(sl)}$, $s = 1, 2$.

2 Вероятности продолжения

Через $Q_l(t)$ обозначим вероятность множества деревьев вывода, высота которых больше t и корень помечен нетерминалом A_l. Эту вероятность назовем вероятностью продолжения по аналогии с теорией ветвящихся процессов. Пусть $(A_{j+1}, A_{j+2}, \ldots, A_{j+k_i})$ - последовательность нетерминалов, образующих группу $\mathcal{M}_i$, где k_i – число нетерминалов в $\mathcal{M}_i$ и j – номер первого по порядку нетерминала в $\mathcal{M}_i$.

Через $Q^{(i)}(t)$ обозначим вектор вероятностей продолжения $Q^{(i)}(t) = (Q_{j+1}(t), Q_{j+2}(t), \ldots, Q_{j+k_i}(t))$. При $i = 1$ представим $Q^{(i)}(t)$ в виде $Q^{(i)}(t) = \left(Q_1^{(i)}(t), Q_2^{(i)}(t), Q_3^{(i)}(t) \right)^T$, и при $i > 1$ в виде $Q^{(i)}(t) = \left(Q_2^{(i)}(t), Q_3^{(i)}(t) \right)^T$, где $Q_l^{(i)}(t)$ соответствует подгруппе $\mathcal{M}_{i,l}$ $(l = 1, 2, 3)$.

Теорема 2. *При* $t \to \infty$

$$\begin{pmatrix} Q_1^{(i)}(t) \\ Q_2^{(i)}(t) \\ Q_3^{(i)}(t) \end{pmatrix} = \begin{pmatrix} U'^{(i)} \cdot t^{q_i^*-1} \cdot r^t \cdot (1 + o(1)) \\ U''^{(i)} \cdot t^{q_i^*-1} \cdot r^t \cdot (1 + o(1)) \\ o\left(t^{q_i^*-1} \cdot r^t\right) \end{pmatrix}.$$

Для доказательства теоремы предварительно докажем несколько лемм.

Лемма 2.[Севаст] *Пусть A – неотрицательная неразложимая матрица, r – ее перронов корень, $r \leq 1$, и A_t – последовательность матриц, для которых $0 \leq A_t \leq A$, и $A_t \to 0$ при $t \to \infty$. Пусть $A_t^* = (A - A_t)(A - A_{t-1}), \ldots, (A - A_1)$. Тогда для любого вектора $x > 0$ выполняется равенство*

$$\lim_{t\to\infty} \frac{A_t^* x}{v A_t^* x} = u,$$

где u и v – соответственно правый и левый собственные положительные векторы, соответствующие r, при нормировке $vu = 1$.

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:

$$F(0, s) = s, \; F(1, s) = F(s),$$

$$F(t + 1, s) = F(F(t, s)). \tag{8}$$

Здесь k – общее число нетерминалов в грамматике.

Очевидно, что $F(t, \bar{1}) = \bar{1}$ для всех t. Известно [Севаст], что $Q(t) = \bar{1} - F(t, \bar{0})$, где $Q(t) = (Q_1(t), \ldots, Q_k(t))^T$. Пусть $R(t, s) = \bar{1} - F(t, s)$, в частности, $R_i(t, \bar{0}) = Q_i(t)$.
Лемма 3. *Для стохастической КС-грамматики с матрицей первых моментов A вида (2) справедливо равенство*

$$\bar{1} - F(s) = (A - E(s)) \left(\bar{1} - s\right), \tag{9}$$

где $0 \le E(s) \le A$, *причем элементы матрицы* $E(s)$ *при* $\bar{0} \le s \le \bar{1}$ *удовлетворяют условиям*

$$E_{ij}(s) = \frac{1}{2}\sum_{l} \delta^{i}_{jl}(s)\,(1-s_l)\,,\ \textit{где}\ \ 0 \le \delta^{i}_{jl}(s) \le b^{i}_{jl} \qquad (10)$$

и $E(s)$ *имеет блочный вид (2) при любом* $s,\ \bar{0} \le s \le \bar{1}$.

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

$$1 - F_i(s) = \sum_{j} \left.\frac{\partial F_i(s)}{\partial s_j}\right|_{s=\theta^i} (1-s_j),\ \text{где}\ \bar{0} \le \theta^i \le \bar{1}.$$

Поскольку производящие функции $F_i(s)$ – многочлены с положительными коэффициентами, все их производные являются многочленами с неотрицательными коэффициентами и, следовательно, $0 \le \left.\frac{\partial F_i(s)}{\partial s_j}\right|_{s=\theta^i} \le a^{i}_{j}$. Раскладывая $\frac{\partial F_i(s)}{\partial s_j}$ аналогичным образом, получаем

$$\frac{\partial F_i(s)}{\partial s_j} = a^{i}_{j} - \frac{1}{2}\sum_{l} \delta^{i}_{jl}(s)(1-s_l) = a^{i}_{j} - E_{ij}(s),$$

где $0 \le \delta^{i}_{jl}(s) \le \left.\frac{\partial^2 F_i(s)}{\partial s_j \partial s_l}\right|_{s=\bar{1}}$. Отсюда следуют равенства (12) и (13). Из неотрицательности и монотонности по s всех производных производящей функции $F_i(s)$ следует, что $0 \le E(s) \le A$ при любом $\bar{0} \le s \le \bar{1}$, и матрица $E(s)$ имеет блочный вид (2). Лемма доказана.

Подставляя в соотношение (13) в качестве s вектор $F(t,s)$ и используя равенство (9), получаем

$$\bar{1} - F(t+1,s) = (A - E(F(t,s)))\,(\bar{1} - F(t,s)). \qquad (11)$$

Обозначим $E(F(t,s))$ через $E_t(s)$, а $E_t(0)$ через E_t, и применим формулу (13) рекурсивно. Тогда

$$R(t,s) = \bar{1} - F(t,s) = \prod_{l=n}^{t-1}(A - E_l(s))R(n,s) = \prod_{l=1}^{t-1}(A - E_l(s))(\bar{1} - s). \qquad (12)$$

Здесь и далее будем применять запись $\prod_{l=m}^{n}(A - E_l(s))$ для выражения $(A - E_n(s))\,(A - E_{n-1}(s))\ldots(A - E_m(s))$.

Из (13) и (14) при $s = \bar{0}$ следует, что

$$Q(t) = (A - E_{t-1}) \cdot Q(t-1) = \prod_{l=n}^{t-1}(A - E_l) \cdot Q(n) = \prod_{l=1}^{t-1}(A - E_l) \cdot \bar{1}. \qquad (13)$$

Кроме того, из формулы (13) следует, что при любом $s,\ \bar{0} \le s \le \bar{1}$,

$$R(t,s) = \bar{1} - F(t,s) \le A^{t-1} \cdot (\bar{1} - s). \qquad (14)$$

Из теоремы 1 следует, что $A^t \to 0$ при $t \to \infty$, поэтому $F(t,s) \to \bar{1}$ и $\lim_{t\to\infty} E_t(s) = 0$ (поэлементно).

Лемма 4. *Для любого* s, $0 \le s \le 1$, *справедлива оценка* $E_t(s)_{ij} = O(t^{s_{1h}-1} r^t)$ *при* $s_{1h} \ge 1$, *где* h *– номер класса, которому принадлежит нетерминал* A_i.

Доказательство. Из доказательства леммы 3 следует, что

$$E_{ij}(s) \le \frac{1}{2} \sum_l b^i_{jl} \cdot (1 - s_l).$$

Подставляя в качестве s значение $F(t,s)$, получаем неравенство

$$E_t(s)_{ij} \le \frac{1}{2} \sum_l b^i_{jl} \left(1 - F_l(t,s)\right) = \frac{1}{2} \sum_l b^i_{jl} \cdot R_l(t,s).$$

Из (15) следует оценка

$$E_t(s)_{ij} \le O\left(t^{s_{1h}-1} r^t\right),$$

где h – номер класса, которому принадлежит нетерминал A_i.

Лемма доказана.

Доказательство теоремы 2.

Доказательство проведем методом математической индукции по числу групп w. Пусть $w = 1$. Применим представление (15) для $Q(t)$:

$$Q(t+1) = \prod_{i=n}^{t} (A - E_i) \cdot Q(n). \tag{15}$$

Выберем такие r' и n, что $r < r' < 1$ и $E_t \le (r')^n \cdot A$ при $t \ge n$. Тогда

$$\left(1 - (r')^n\right)^{t-n+1} \cdot A^{t-n+1} \le \prod_{i=n}^{t} (A - E_i) \le A^{t-n+1}.$$

Поэтому справедливо неравенство

$$A^l - \prod_{i=n}^{t} (A - E_i) \le A^l \cdot \left(1 - (1 - (r')^n)^l\right), \text{ где } l = t - n + 1.$$

Оценим разность $1 - (1 - (r')^n)^l$:

$$\left(1 - (1 - (r')^n)^l\right) = (r')^n \cdot \sum_{i=0}^{l} ((r')^n)^i < (l+1) \cdot (r')^n.$$

Следовательно, при $l \le n$ справедливо равенство

$$\prod_{i=n}^{t} (A - E_i) = A^l \cdot \left(1 + O((r'')^n)\right),$$

где r'' удовлетворяет условию $r' < r'' < 1$.

Уравнение (16) можно переписать в следующем виде:

$$Q(t+1) = A^l \cdot (1 + O((r'')^n)) \cdot Q(n), \tag{16}$$

и для $w = 1$

$$Q^{(1)}(t+1) = B_{11}^l \cdot (1 + O((r'')^n)) \cdot Q^{(1)}(n).$$

Представим $Q^{(1)}(t)$ в виде $Q^{(1)}(t) = \left(Q_1^{(1)}(t), Q_2^{(1)}(t), Q_3^{(1)}(t)\right)^T$, где $Q_l^{(1)}(t)$ соответствует подгруппе $\mathcal{M}_{1,l}$ $(l = 1, 2, 3)$.

Из (15) следует, что $Q^{(1)}(n) \leq B_{11}^{n-1} \cdot \overline{1} = O(r^n)$. Поэтому, с учетом представления (4) для B_{11}, справедлива оценка $Q_3^{(1)}(t) \leq C_{33}^{t-1} \cdot \overline{1}$.

Известно следующее представление для степени произвольной матрицы C [Гантмахер]:

$$C^t = \sum_{l=1}^{s} \left(\lambda_l^t Z_{l1} + \left(\lambda_l^t\right)' \cdot Z_{l2} + \ldots + \left(\lambda_l^t\right)^{(m_l-1)} Z_{lm_l}\right), \tag{17}$$

где λ_l — корни минимального многочлена $\psi(\lambda)$ матрицы C $(l = 1, \ldots, s)$, $s < k$, m_l — кратность корня λ_l для минимального многочлена, $(\lambda_l^t)^{(n)}$ — n-я производная по λ_l от λ_l^t, матрицы Z_{lj} вполне определяются заданием матрицы C и не зависят от t.

Применяя это представление для C_{33}^t, получим, что $Q_3^{(1)}(t) \leq O\left(t^m \cdot (r')^t\right)$, где r' – перронов корень для C_{33} и m – его кратность. Отметим, что $r' < r$.

Используя (15) и оценку для $Q_3^{(1)}(t)$, запишем уравнение для $Q_2^{(1)}(t)$:

$$Q_2^{(1)}(t+1) = (C_{22} - E'_t) \cdot Q_2^{(1)}(t) + O\left(t^m \cdot (r')^t\right). \tag{18}$$

Здесь E'_t – подматрица матрицы E_t, соответствующая C_{22}.

Через $Q[K_i](t)$ обозначим вектор вероятностей продолжения для класса нетерминалов K_i. Рассмотрим $K_i \in \mathcal{M}_{12}$. Для него из (19) следует уравнение

$$Q[K_i](t+1) = (A_{ii} - E_t^i) \cdot Q[K_i](t) + O\left(t^m \cdot (r')^t\right),$$

где E_t^i – подматрица матрицы E'_t, соответствующая подстрокам для класса K_i.

Умножим слева обе части уравнения на вектор V_i – левый собственный вектор матрицы A_{ii}, соответствующий перронову корню r. Тогда с учетом оценок при $w = 1$ для E_t^i и $Q[K_i](t)$ получим следующее уравнение:

$$V_i \cdot Q[K_i](t+1) = rV_i \cdot Q[K_i](t) + O\left(r^{2t}\right) + O\left(t^m \cdot (r')^t\right).$$

Введем обозначение $x_t = \frac{V_i Q[K_i](t)}{r^t}$. Тогда предыдущее уравнение перепишется таким образом:

$$x_{t+1} = x_t + O\left(r^t\right) + O\left(t^m \cdot \left(\frac{r'}{r}\right)^t\right).$$

Просуммировав уравнение от 1 до t, получим $x_t = x_1 + c + o(1)$, где константа c получается в результате суммирования сходящихся рядов $\sum_{j=1}^{t} O\left(r^j\right)$ и $\sum_{j=1}^{t} O\left(j^m \cdot \left(\frac{r'}{r}\right)^j\right)$ при $t \to \infty$. Отсюда $V_i \cdot Q[K_i](t) = c_i^{(1)} \cdot r^t \cdot (1 + o(1))$, где $c_i^{(1)} = x_1 + c$.

Поскольку матрица A_{ii} неразложима, то, учитывая оценку для E_t^i из леммы 3 и применяя лемму 2, найдем, что

$$Q[K_i](t) = c_i^{(1)} U_i r^t \cdot (1 + o(1)),$$

где U_i – правый собственный вектор матрицы A_{ii}, соответствующий r.

Таким образом, поскольку подгруппу $\mathcal{M}_{12}$ составляют несравнимые классы с перроновым корнем r, мы можем записать формулу для $Q_2^{(1)}(n)$ в следующем виде:

$$Q_2^{(1)}(t) = U''^{(1)} r^t \cdot (1 + o(1)),$$

где $U''^{(1)} = \sum_i c_i^{(1)} U_i^{(2)}$.

Используя (17), запишем уравнение для $Q_1^{(1)}(t)$:

$$Q_1^{(1)}(t+1) = \left(C_{11}^l \cdot Q_1^{(1)}(n) + C_{12}^{(l)} \cdot Q_2^{(1)}(n) + C_{13}^{(l)} \cdot Q_3^{(1)}(n) \right) \cdot (1 + O(1)).$$

После подстановки значений для $Q_2^{(1)}(n)$, C_{11}^l, $C_{12}^{(l)}$, $C_{13}^{(l)}$ и оценок для $Q_1^{(1)}(n)$ и $Q_3^{(1)}(n)$ получим:

$$Q_1^{(1)}(t) = \sum_i c_i^{(1)} U_i^{(1)} r^t \cdot (1 + o(1)).$$

Здесь мы учли тот факт, что $V_i^{(2)} U''^{(1)} = V_i^{(2)} U_i^{(2)} = 1$.

Введем обозначение $U'^{(1)} = \sum_i c_i^{(1)} U_i^{(1)}$. Тогда

$$Q_1^{(1)}(t) = U'^{(1)} r^t \cdot (1 + o(1)).$$

Таким образом,

$$\begin{pmatrix} Q_1^{(1)}(t) \\ Q_2^{(1)}(t) \\ Q_3^{(1)}(t) \end{pmatrix} = \begin{pmatrix} U'^{(1)} r^t \cdot (1 + o(1)) \\ U''^{(1)} r^t \cdot (1 + o(1)) \\ o(r^t) \end{pmatrix} \tag{19}$$

и утверждение теоремы при $w = 1$ справедливо.

Пусть теперь $w = 2$.

Используя (15), запишем систему для $Q(t+1)$:

$$\begin{cases} Q^{(1)}(t+1) = \left(B_{11}^{t-n} Q^{(1)}(n) + B_{12}^{(t-n)} Q^{(2)}(n) \right) (1 + O((r'')^n)) \\ Q^{(2)}(t+1) = \left(B_{22}^{t-n} Q^{(2)}(n) \right) (1 + O((r'')^n)). \end{cases} \tag{20}$$

Представим $Q^{(2)}(t)$ в виде $Q^{(2)}(t) = \left(Q_2^{(2)}(t), Q_3^{(2)}(t) \right)^T$, где $Q_l^{(2)}(t)$ соответствует подгруппе $\mathcal{M}_{2,l}$ $(l = 2, 3)$. Применяя доказательство теоремы для $w = 1$ ко второму уравнению, получим, что

$$Q^{(2)}(t) = \begin{pmatrix} Q_2^{(2)}(t) \\ Q_3^{(2)}(t) \end{pmatrix} = \begin{pmatrix} U''^{(2)} r^t \cdot (1 + o(1)) \\ o(r^t) \end{pmatrix},$$

где $U''^{(2)} = \sum_i c_i^{(2)} U_i^{(22)}$, и $c_i^{(2)}$ – константа, соответствующая классу нетерминалов $K_i \in \mathcal{M}_{22}$.

Перейдем к оценке компонент вектора $Q^{(1)}(t)$. Применяя равенство (15), запишем уравнение для $Q^{(1)}(t)$:

$$Q^{(1)}(t+1) = \left(B_{11}Q^{(1)}(t) + B_{12}Q^{(2)}(t)\right) \cdot (1 + O(t^{s_{12}-1} \cdot r^t)).$$

Заметим, что $Q_3^{(1)}(t)$ можно рассматривать и в качестве $Q_1^{(2)}(t)$, поэтому отсюда следует оценка $Q_3^{(1)}(t) = U'^{(21)} \cdot r^t \cdot (1 + o(1))$.

Применяя представления (4) и (9) для B_{11} и B_{12}, а также учитывая оценки для $Q^{(1)}(t)$ и $Q^{(2)}(t)$, следующие из (15), запишем уравнение для $Q_2^{(1)}(t)$:

$$Q_2^{(1)}(t+1) = C_{22}Q_2^{(1)}(t) + C_{23}Q_3^{(1)}(t) + D_{21}Q_2^{(2)}(t) + o(r^t) + O(t^{s_{12}-1} \cdot r^{2t}) =$$

$$C_{22}Q_2^{(1)}(t) + C_{23}U'^{(2)}r^t + D_{21}U''^{(2)}r^t + o(r^t) + O(t^{s_{12}-1} \cdot r^{2t}). \tag{21}$$

Далее в верхнем индексе собственных векторов будем добавлять номер группы. Умножим обе части уравнения на левый собственный вектор $V_i^{(21)}$ матрицы C_{22}, соответствующий r. Получим, что

$$V_i^{(21)}Q_2^{(1)}(t+1) = rV_i^{(21)}Q_2^{(1)}(t) + V_i^{(21)} \cdot (C_{23}U'^{(2)} + D_{21}U''^{(2)})r^t + o(r^t) + O(t^{s_{12}-1} \cdot r^{2t}).$$

Как и в случае $w = 1$, введем обозначение $x_t = \frac{V_i^{(21)}Q_2^{(1)}(t)}{r^t}$. Тогда предыдущее уравнение перепишется таким образом:

$$x_{t+1} = x_t + \frac{V_i^{(21)} \cdot (C_{23}U'^{(2)} + D_{21}U''^{(2)})}{r} \cdot (1 + o(1)) + O(t^{s_{12}-1} \cdot r^t).$$

Просуммировав обе части уравнения по t от 1 до t, получим:

$$x_{t+1} = x_1 + c_i^{(1)}t + O(1) = c_i^{(1)}t \cdot (1 + o(1)),$$

где $c_i^{(1)}t$ – сумма ряда $\sum_{l=1}^{t} \frac{V_i^{(21)} \cdot (C_{23}U'^{(2)} + D_{21}U''^{(2)})}{r}$. Поэтому

$$V_i^{(21)}Q_2^{(1)}(t) = c_i^{(1)}tr^t(1 + o(1)). \tag{22}$$

Используя второе уравнение системы (21), запишем уравнение для $Q_2^{(1)}(t)$:

$$Q_2^{(1)}(t+1) = \left(C_{22}^{t-n}Q_2^{(1)}(n) + C_{23}^{(t-n)}Q_3^{(1)}(n) + D_{21}^{(t-n)}Q_2^{(2)}(n) \cdot (1 + o(1))\right)(1 + O((r'')^n)).$$

Применим оценки для C_{22}^{t-n}, $C_{23}^{(t-n)}$ и $D_{21}^{(t-n)}$, следующие из теоремы 1, и оценки для $Q_3^{(1)}(n)$ и $Q_2^{(2)}(n)$:

$$Q_2^{(1)}(t+1) = \left(\sum_i U_i^{(21)}V_i^{(21)}r^{t-n}Q_2^{(1)}(n) + O\left((t-n) \cdot r^t\right)\right)(1 + O((r'')^n)).$$

Подставляя оценку (23) для $V_i^{(21)}Q_2^{(1)}(n)$, а также учитывая, что $t - n = \lfloor \log t \rfloor$, получим, что

$$Q_2^{(1)}(t+1) = \sum_i c_i^{(1)}U_i^{(21)}tr^t \cdot (1 + o(1)).$$

Отсюда следует, что для класса $K_i \in \mathcal{M}_{12}$ вектор $Q[K_i](t)$ пропорционален правому собственному вектору $U_i^{(21)}$ матрицы A_{ii}.

Наконец, запишем уравнение для $Q_1^{(1)}(t+1)$, используя систему (21):

$$Q_1^{(1)}(t+1) =$$

$$\left(C_{11}^{t-n}Q_1^{(1)}(n) + C_{12}^{t-n}Q_2^{(1)}(n) + C_{13}^{t-n}Q_3^{(1)}(n) + D_{11}^{t-n}Q_2^{(2)}(n) + D_{12}^{t-n}Q_3^{(2)}(n)\right)(1+O((r'')^n)).$$

Подставляя в уравнение полученные ранее оценки для $Q_2^{(1)}(n)$, $Q_3^{(1)}(n)$, $Q_2^{(2)}(n)$ и $Q_3^{(2)}(n)$, а также применяя теорему 1 для входящих в уравнение матриц и учитывая, что $t-n = \lfloor \log t \rfloor$, получим, что

$$Q_1^{(1)}(t) = \sum_i c_i^{(1)} U_i^{(11)} t r^t \cdot (1+o(1)) = U'^{(11)} t r^t \cdot (1+o(1)).$$

Таким образом, доказана справедливость теоремы для $w = 2$.

Предположим, что теорема 2 справедлива для $w-1$ групп. Докажем, что тогда она справедлива и для w групп. Запишем уравнение для $Q_1^{(1)}(t)$:

$$Q_1^{(1)}(t+1) = B_{11}Q_1^{(1)}(t) + B_{12}Q_1^{(2)}(t) + B_{13}Q_1^{(3)}(t) + \ldots + B_{1w}Q_1^{(w)}(t)(1+O(t^q r^t)). \quad (23)$$

Для последовательности групп $\mathcal{M}_2, \mathcal{M}_3, \ldots, \mathcal{M}_w$ утверждение теоремы справедливо по предположению индукции. Подставим значения $Q_1^{(l)}(t)$ $l = 2, \ldots, w$ в (25). Так как q_l^* имеет наибольшее значение для $l = 2$, определяющим будет слагаемое $B_{12}Q_1^{(2)}(t)$. Поэтому уравнение (25) можно записать в следующем виде:

$$Q_1^{(1)}(t+1) = B_{11}Q^{(1)}(t) + B_{12}Q^{(2)}(t)(1+O(t^q r^t)).$$

Это уравнение аналогично уравнению для $w = 2$. Повторяя рассуждения для $w = 2$, при этом учитывая, что $Q^{(2)}(t) = O(t^{q_2^*-1} r^t)$ и рассматривая в качестве x_t значение $\frac{V_i^{(21)} Q_2^{(1)}(t)}{t^{q_2^*-1} r^t}$, получим утверждение теоремы для w групп нетерминалов.

Теорема доказана.

Через $P(D_i^t)$ обозначим вероятность деревьев вывода высоты t, корень которых помечен нетерминалом A_i. Очевидно,

$$P(D_i^t) = Q_i(t-1) - Q_i(t).$$

Из теоремы 2 вытекает

Следствие. Пусть нетерминал $A_i \in \mathcal{M}_{l1}$, либо $A_i \in \mathcal{M}_{l2}$. Тогда

$$P(D_i^t) = d_i t^{q_l^*-1} r^{t-1} \cdot (1-r) \cdot (1+o(1)), \quad (24)$$

где d_i – компонента вектора U' либо вектора U'', соответствующая нетерминалу A_i.

Заметим, что в случае, когда $A_i \in \mathcal{M}_{l3}$, для нахождения $Q_i(t)$ и $P(D_i^t)$ следует присоединить $\mathcal{M}_{l3}$ к следующей группе в качестве $\mathcal{M}_{l+1,1}$.

Работа поддержана грантом РФФИ 07-01-00739-а.

ЛИТЕРАТУРА

1. 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.

2. Фу К. Структурные методы в распознавании образов. — М.: Мир, 1977.