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

Л. П. Жильцова

Автором в [4, 5] рассматривались вопросы, связанные с кодированием слов стохастического контекстно-свободного языка (стохастического КС-языка), при условии, что матрица первых моментов порождающей грамматики неразложима 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 [1, 8].

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}, удовлетворяющая 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. В работе рассматриваются бесконечные КС-языки.

Каждому слову α КС-языка соответствует последовательность правил грамматики (вывод), с помощью которой α выводится из аксиомы s. Для определенности в качестве вывода будем рассматривать левый вывод, когда очередное правило грамматики применяется к самому левому вхождению нетерминала в слово. Вероятность вывода определяется как произведение вероятностей правил, его образующих.

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 $A_i \xrightarrow{p_{ij}} b_{i_1}b_{i_2}\dots b_{i_m}$, где $b_{i_l} \in V_N \cup V_T$ $(l = 1,\dots,m)$. 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 $b_{i_1}, b_{i_2}, \dots, b_{i_m}$ 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 кроны 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

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

где переменная s_i соответствует нетерминальному символу A_i [7]. Функция $F_i(s_1, s_2, \dots, s_k)$ строится по множеству правил R_i с одинаковой левой частью A_i следующим образом.

Для каждого правила $A_i \xrightarrow{p_{ij}} \beta_{ij}$ 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 $(m = 1,\dots,k)$. Тогда

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

Пусть

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

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 [3] 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 $A_i \overset{p_{il}}{\to} \alpha_1 A_j \alpha_2$, 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 разложимая матрица [3] первых моментов A имеет следующий вид:

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

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 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 [4].

Обозначим через r_i перронов корень матрицы A_{ii}. Для неразложимой матрицы перронов корень является действительным и простым [3]. Очевидно, $r = \max_i\{r_i\}$, и $r > 0$, ввиду положительности матрицы A_{ii} для любого i.

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}. В случае $K_l \not\prec_* K_h$ положим $s_{lh} = 0$.

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

Разобьем полученную последовательность классов $K_1, K_2, \dots, K_m$ 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 $\max_{K_l \in \mathcal{M}_i, K_h \in \mathcal{M}_j}\{s_{lh}\}$.

Среди последовательностей вида

$$K_{i_1} \prec K_{i_2} \prec \ldots \prec K_{i_s}, \tag{2}$$

где $i_1 = l$ и i_s принимает всевозможные значения, выберем последовательности с наибольшим числом классов с номерами из J. Это число обозначим через q_l. Максимальным путем назовем последовательность вида (2) при $i_1 = 1$, содержащую q_1 классов с номерами из J. Множество всех классов с номерами из J, принадлежащих максимальным путям, обозначим J_{MAX}.

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

$$A = \begin{pmatrix} B_{11} & B_{12} & \ldots & B_{1w} \\ 0 & B_{22} & \ldots & B_{2w} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & B_{ww} \end{pmatrix}, \tag{3}$$

где B_{lh} – подматрица на пересечении строк для классов из группы $\mathcal{M}_l$ и столбцов для классов из $\mathcal{M}_h$. Заметим, что любая группа содержит хотя бы один класс с номером из определяющего множества J. Поэтому каждая матрица B_{ll} имеет перронов корень r. Запись $B_{lh}^{(t)}$ будем применять для обозначения соответствующей подматрицы матрицы A^t.

Теорема 1 [6]. *При* $t \to \infty$

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

где H_{lh} *– неотрицательная 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 $\mathcal{M}_{11}$, $\mathcal{M}_{12}$ и $\mathcal{M}_{13}$. К $\mathcal{M}_{11}$ отнесем классы нетерминалов с $s_{1l} = 0$, к $\mathcal{M}_{12}$ — классы с номерами из множества J, к $\mathcal{M}_{13}$ — 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

$$B_{ll} = \begin{pmatrix} C_{11} & C_{12} & C_{13} \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix}, \tag{4}$$

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 $C_{33}^t = o(r^t)$ по теореме 1.

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 [7], матрица C_{22}^t имеет вид

$$\begin{pmatrix} U_{j_1+1}V_{j_1+1} & 0 & \dots & 0 & 0 \\ 0 & U_{j_1+2}V_{j_1+2} & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & U_{j_1+j_2}V_{j_1+j_2} \end{pmatrix} \cdot r^t(1+o(1)), \tag{5}$$

где U_i –правый собственный положительный вектор (вектор-столбец) и V_i – левый собственный положительный вектор (вектор-строка) матрицы A_{ii}, соответствующие r, при нормировке $V_i \cdot U_i = 1$, $i = j_1+1, \dots, j_1+j_2$ (здесь $j_1 = |\mathcal{M}_{l1}|$).

Обозначим матрицу

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

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

$$U_i^{(2l)} = (0, \dots, 0, U_i, 0, \dots, 0)^T, \ \ V_i^{(2l)} = (0, \dots, 0, V_i, 0, \dots, 0), \tag{6}$$

и U_i и V_i расположены на местах, соответствующих классу K_i в $\mathcal{M}_{l2}$. Заметим, что $U_i^{(2l)}$ и $V_i^{(2l)}$ являются соответственно правым и левым собственными векторами матрицы C_{22} для корня r.

Кроме того, каждому классу $K_i \in \mathcal{M}_{l2}$ соответствуют правый и левый собственные векторы всей матрицы B_{ll} для перронова корня r.

Компоненты правого собственного вектора можно представить 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 [6] 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 [6] также показано, что $V_i^{(1l)} = 0$ и

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

Используя описанные векторы, уточним вид матрицы B_{ll} [6]:

$$B_{ll}^t \sim \begin{pmatrix} 0 & \sum_{i=j_1+1}^{j_2} U_i^{(1l)} V_i^{(2l)} & \sum_{i=j_1+1}^{j_2} U_i^{(1l)} V_i^{(3l)} \\ 0 & \sum_{i=j_1+1}^{j_2} U_i^{(2l)} V_i^{(2l)} & \sum_{i=j_1+1}^{j_2} U_i^{(2l)} V_i^{(2l)} \\ 0 & 0 & 0 \end{pmatrix} \cdot r^t. \tag{7}$$

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 $l < h$. Матрицу B_{lh} представим в виде

$$B_{lh} = \begin{pmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \\ D_{31} & D_{32} \end{pmatrix}, \tag{8}$$

где разбиение по строкам сделано в соответствии с подгруппами $\mathcal{M}_{l1}$, $\mathcal{M}_{l2}$ и $\mathcal{M}_{l3}$, а по столбцам – в соответствии с подгруппами $\mathcal{M}_{h2}$ и $\mathcal{M}_{h3}$. Так как $h > 1$, для группы нетерминалов $\mathcal{M}_h$ подгруппа $\mathcal{M}_{h1}$ пустая и не представлена в матрице B_{lh}. В [6] 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 $U_j'^{(sl)} = \frac{1}{s_{lh}^*-1} \sum_i \left(V_i^{(2l)} D_{21} + V_i^{(3l)} D_{31} \right) U_j^{(2h)} U_i^{(sl)}$, $s = 1, 2$.

2 Моменты

Пусть $\Xi = (\xi_1, \ldots, \xi_k)$– случайный вектор, $\alpha^* = (\alpha_1, \ldots, \alpha_k)$ — фиксированный вектор с целочисленными неотрицательными компонентами и $\alpha = \alpha_1 + \ldots + \alpha_k$. Обозначим

$$\Xi^{[\alpha^*]} = \xi_1^{[\alpha_1]} \ldots \xi_n^{[\alpha_k]},$$

где $x^{[a]} = x(x-1)\ldots(x-a+1)$. Математическое ожидание $M\Xi^{[\alpha^*]}$ будем называть α^*-*моментом* Ξ [Севаст].

Пусть A_i - некоторый нетерминальный символ грамматики G. Через L_i обозначим язык, порожденный грамматикой G_i, которая получается из G заменой аксиомы на A_i. Будем считать, что аксиомой исходной грамматики является нетерминал A_1 и $L = L_1$ для исходного языка L. Через D_i обозначим множество деревьев вывода для слов из L_i.

Пусть $x_j^i(t)$ — число нетерминалов A_j в дереве вывода из D_i на ярусе t. Через $M_{\alpha^*}^i(t)$ обозначим α^*−момент вектора $X^i(t) = (x_1^i(t), \ldots, x_k^i(t))$.

Примем специальные обозначения для моментов первых четырех порядков. Факториальные моменты первого порядка будем обозначать через $a_j^i(t)$. Для факториальных моментов второго порядка введем обозначения $b_{jn}^i(t)$. Таким образом, $b_{jj}^i(t) = Mx_j^i(t)(x_j^i(t)-1)$ и $b_{jn}^i(t) = Mx_j^i(t)x_n^i(t)$ при $j \neq n$. Для факториальных моментов третьего и четвертого порядков введем обозначения $c_{jnq}^i(t)$ и $f_{jnql}^i(t)$ соответственно.

Нетрудно заметить, что $a_j^i(1)$ – элементы матрицы первых моментов, для которых мы ввели ранее обозначения a_j^i.

Будем также применять далее обозначения b_{jn}^i для $b_{jn}^i(1)$.

Нас интересуют оценки для первых четырех моментов.

Свойства первых моментов исследованы в [6], так как $a_j^i(t)$ - элемент матрицы A^t.

Для вторых моментов известна следующая формула из [7]:

$$b_{jn}^i(t) = \sum_{\tau=1}^{t} \sum_{l,m,s} a_l^i(t-\tau) b_{ms}^l a_j^m(\tau-1) a_n^s(\tau-1). \tag{9}$$

Пусть a_l^i принадлежит подматрице $A_{h_i h_l}$, a_j^m — подматрице $A_{h_m h_j}$, и a_n^s — подматрице $A_{h_s h_n}$. Подставим в (9) представление для первых моментов:

$$b_{jn}^i(t) = \sum_{\tau=1}^{t} \sum_{l,m,s} c_{il} \cdot (t-\tau)^{\delta_1} \cdot \left(1 + O\left(\frac{1}{t-\tau}\right)\right) \cdot r^{t-\tau} \cdot b_{ms}^l \cdot c_{mj} \cdot (\tau-1)^{\delta_2} \cdot r^{\tau-1} \times$$

$$c_{sn} \cdot (\tau-1)^{\delta_3} \cdot r^{\tau-1} \cdot \left(1 + O\left(\frac{1}{\tau}\right)\right).$$

Здесь $\delta_1 = s_{h_i h_l} - 1$, $\delta_2 = s_{h_m h_j} - 1$, и $\delta_3 = s_{h_s h_n} - 1$, и c_{il}, c_{mj} и c_{sn} — коэффициенты в соответствующих элементах матрицы A^t.

Проведем несложные преобразования в полученном равенстве:

$$b_{jn}^i(t) = r^t \sum_{l} c_{il}\, t^{\delta_1} \sum_{\tau=1}^{t} \left(1 - \frac{\tau}{t}\right)^{\delta_1} \left(1 + O\left(\frac{1}{t-\tau}\right)\right) \sum_{m,s} b_{ms}^l c_{mj} c_{sn} \cdot \tau^{\delta_2+\delta_3} \times$$

$$\left(1 - \frac{1}{\tau}\right)^{\delta_2+\delta_3} r^{\tau-2} \cdot \left(1 + O\left(\frac{1}{\tau}\right)\right).$$

Ряд

$$\sum_{\tau=1}^{\infty} \left(1 - \frac{\tau}{t}\right)^{\delta_1} \left(1 + O\left(\frac{1}{t-\tau}\right)\right) \sum_{m,s} b_{ms}^l c_{mj} c_{sn} \cdot \tau^{\delta_2} \left(1 - \frac{1}{\tau}\right)^{\delta_2+\delta_3} r^{\tau-2} \cdot \left(1 + O\left(\frac{1}{\tau}\right)\right)$$

сходится. Обозначим его сумму через $g_{jn}^i(l)$. Отметим, что $g_{jn}^i(l) > 0$ в тех случаях, когда существуют m и s такие, что $b_{ms}^l > 0$,

$$K_{h_i} \prec_* K_{h_l} \prec_* K_{h_m} \prec_* K_{h_j} \ \text{ и } \ K_{h_i} \prec_* K_{h_l} \prec_* K_{h_s} \prec_* K_{h_n}.$$

Условие $b_{ms}^l > 0$ выполняется тогда и только тогда, когда в грамматике существует правило с нетерминалом A_l в левой части, содержащее в правой части оба нетерминала A_m и A_s.

При $t \to \infty$

$$b_{jn}^i(t) = \sum_{l} g_{jn}^i(l) \cdot t^{\delta_1} r^t (1 + o(1)),$$

где суммирование ведется по тем l, для которых $g_{jn}^i(l) > 0$.

Очевидно, определяющими в этой сумме являются слагаемые с теми значениями l, для которых δ_1 имеет наибольшее значение. Обозначим его через δ_{jn}^i. Поэтому формулу для $b_{jn}^i(t)$ можно записать в следующем виде:

$$b_{jn}^i(t) = g_{jn}^i t^{\delta_{jn}^i} r^t \cdot (1 + o(1)). \tag{10}$$

Здесь $g_{jn}^i = \sum_l g_{jn}^i(l)$, где суммирование ведется по значениям l, удовлетворяющим перечисленным выше условиям. Так как $l \le j$ и $l \le n$, то $\delta_{jn}^i \le \max\{s_{h_i h_j} - 1, s_{h_i h_n} - 1\}$. Поэтому $b_{jn}^i(t) \le O\left(\, a_j^i(t) + a_n^i(t)\, \right)$.

Используя результаты из [7], запишем формулу для третьего момента:

$$c_{jnq}^i(t) = \sum_{\tau=1}^{t} \sum_{l} a_l^i(t-\tau) \cdot z_{jnq}^l(\tau-1).$$

В этой формуле $z^l_{jnq}(\tau-1)$ состоит из конечного числа слагаемых двух типов. Слагаемые первого типа имеют вид: $Ca^s_q(\tau-1)\cdot a^m_n(\tau-1)\cdot a^l_j(\tau-1)$ для некоторых s, m, l, где C — некоторая константа, зависящая от слагаемого; слагаемые второго типа имеют вид: $Ca^l_j(\tau-1)\cdot b^m_{nq}(\tau-1)$ для некоторых l, m и константы C.

Поэтому вычисление $c^i_{jnq}(t)$ сводится к вычислению конечного числа сумм вида

$$S_1(t)=\sum_{\tau=1}^{t} a^i_l(t-\tau)\cdot a^l_j(\tau-1)\cdot a^s_q(\tau-1)\cdot a^m_n(\tau-1)$$

и вида

$$S_2(t)=\sum_{\tau=1}^{t} a^i_l(t-\tau)\cdot a^s_j(\tau-1)\cdot b^m_{nq}(\tau-1)$$

для некоторых значений l, m, s.

Оценим $S_1(t)$ и $S_2(t)$, используя оценки $a^i_j(t)=O(t^{s_{ij}-1}r^t)$, $b^i_{jl}(t)\le O\left(a^i_j(t)+a^i_l(t)\right)$. Применяя очевидное неравенство $s_{ij}\le w$, где w –число групп, получим, что $S_1(t)\le O(t^{w-1}r^t)$ и $S_2(t)\le O(t^{w-1}r^t)$. Поэтому

$$c^i_{jnq}(t)\le O\left(t^{w-1}r^t\right).$$

Аналогичный результат может быть получен и для четвертого момента. Для него известна формула [7]:

$$f^i_{j_1j_2j_3j_4}(t)=\sum_{\tau=1}^{t}\sum_{j=1}^{k} a^i_j(t-\tau)z^j_{j_1j_2j_3j_4}(\tau-1),$$

где $z^j_{j_1j_2j_3j_4}(\tau-1)$ состоит из конечного числа слагаемых следующих четырех видов:

$$S_1(\tau-1)=a^{i_1}_{j_1}(\tau-1)\cdot a^{i_2}_{j_2}(\tau-1)\cdot a^{i_3}_{j_3}(\tau-1)\cdot a^{i_4}_{j_4}(\tau-1);$$

$$S_2(\tau-1)=a^{i_1}_{j_1}(\tau-1)\cdot a^{i_2}_{j_2}(\tau-1)\cdot b^{i_3}_{j_3j_4}(\tau-1);$$

$$S_3(\tau-1)=b^{i_1}_{j_1j_2}(\tau-1)\cdot b^{i_2}_{j_3j_4}(\tau-1);$$

$$S_4(\tau-1)=a^{i_1}_{j_1}(\tau-1)\cdot c^{i_2}_{j_2j_3j_4}(\tau-1)$$

для некоторых значений i_1, i_2, i_3, i_4.

Используя оценки для первых трех моментов и проведя элементарные преобразования, получаем оценку по порядку для четвертого момента:

$$f^i_{j_1j_2j_3j_4}(t)\le O\left(t^{w-1}r^t\right).$$

3 Вероятности продолжения

Через $Q_l(t)$ обозначим вероятность множества деревьев вывода из D_l, высота которых больше t. 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 $j+1$ – номер первого по порядку нетерминала в $\mathcal{M}_i$.

Через $Q^{(i)}(t)$ обозначим вектор вероятностей продолжения $Q^{(i)}(t) = (Q_{j+1}(t), Q_{j+2}(t), \ldots, Q_{j+k_i}(t))^T$. Представим $Q^{(i)}(t)$ в виде

$$Q^{(i)}(t) = \left(Q_1^{(i)}(t), Q_2^{(i)}(t), Q_3^{(i)}(t)\right)^T, \tag{11}$$

где $Q_l^{(i)}(t)$ соответствует подгруппе $\mathcal{M}_{i,l}$ ($l = 1, 2, 3$). Напомним, что подгруппа $\mathcal{M}_{i,1}$ может быть пустой, в этом случае вектор $Q_1^{(i)}(t)^T$ имеет нулевую размерность.

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 приведем несколько лемм.

Лемма 1 [7]. *Пусть A – неотрицательная неразложимая матрица, r – ее перронов корень, $r \le 1$, и A_t – последовательность матриц, для которых $0 \le A_t \le A$, 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:

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

$$F(t + 1, s) = F(F(t, s)). \tag{12}$$

Здесь k – общее число нетерминалов в грамматике.

Очевидно, что $F(t, \bar{1}) = \bar{1}$ для всех t. Известно [7], 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 2. *Для стохастической КС-грамматики с матрицей первых моментов A вида* (1) *справедливо равенство*

$$\bar{1} - F(s) = (A - E(s))\left(\bar{1} - s\right), \tag{13}$$

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

$$E(s)_{ij} = \frac{1}{2} \sum_l \delta_{jl}^i(s)\left(1 - s_l\right), \ \textit{где}\ \ 0 \le \delta_{jl}^i(s) \le b_{jl}^i, \tag{14}$$

$E(s)$ имеет блочный вид (1) при любом s, $\bar 0 \le s \le \bar 1$ и b^i_{jl} – вторые моменты.

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 (13) и (14). 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 (3). Лемма доказана.

Подставляя в соотношение (14) в качестве s вектор $F(t,s)$ и используя равенство (13), получаем

$$\bar 1 - F(t+1, s) = (A - E(F(t,s)))\,(\bar 1 - F(t,s)). \tag{15}$$

Обозначим $E(F(t,s))$ через $E_t(s)$, а $E_t(0)$ через E_t, и применим формулу (15) рекурсивно. Тогда

$$R(t,s) = \bar 1 - F(t,s) = \prod_{l=n}^{t-1}(A - E_l(s))R(n,s) = \prod_{l=1}^{t-1}(A - E_l(s))(\bar 1 - s). \tag{16}$$

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 (15) и (16) при $s = \bar 0$ следует, что

$$Q(t) = (A - E_{t-1})\cdot Q(t-1) = \prod_{l=n}^{t-1}(A - E_l)\cdot Q(n) = \prod_{l=1}^{t-1}(A - E_l)\cdot \bar 1. \tag{17}$$

Кроме того, из (16) следует, что при любом s, $\bar 0 \le s \le \bar 1$,

$$R(t,s) = \bar 1 - F(t,s) \le A^{t-1}\cdot(\bar 1 - s). \tag{18}$$

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 3. *Для любого s, $0 \le s \le 1$, справедлива оценка $E_t(s)_{ij} = O(t^{s_{lh}-1}r^t)$ при $s_{lh} \ge 1$, где l и h – номера классов, которым принадлежат нетерминалы A_i и A_j соответственно.*

Доказательство. Из леммы 2 следует, что

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

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

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

Из (18) следует оценка

$$E_t(s)_{ij} \leq O\left(t^{s_{lh}-1} r^t\right).$$

Лемма доказана.

Доказательство теоремы 2.

Доказательство проведем методом математической индукции по числу групп w. Пусть $w = 1$. Применим представление (17) для $Q(t)$:

$$Q(t+1) = \prod_{i=n}^{t} (A - E_i) \cdot Q(n). \tag{19}$$

Положим $n = t - \lfloor \log_2 t \rfloor$. Выберем r' и t_1 такие, что $r < r' < 1$ и $E_t \leq (r')^{t_1} \cdot A$ при $t \geq t_1$. Тогда $\left(1 - (r')^{t_1}\right)^l \cdot A^l \leq \prod_{i=n}^{t} (A - E_i) \leq A^l$, где $l = t - n + 1$. Поэтому $A^l - \prod_{i=n}^{t} (A - E_i) \leq A^l \cdot \left(1 - \left(1 - (r')^n\right)^l\right)$.

Оценим разность:

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

Выберем $t_2 \geq t_1$ и r'' такие, что $r' < r'' < 1$ и $n \cdot (r')^n < (r'')^n$ при $t \geq t_2$. Тогда

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

и уравнение (19) можно переписать в следующем виде:

$$Q(t+1) = A^l \cdot \left(1 + O((r'')^n)\right) \cdot Q(n). \tag{20}$$

Для $w = 1$ из представления (3) следует $A = B_{11}$, поэтому

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

Из (18) следует, что $Q^{(1)}(n) \leq B_{11}^{n-1} \cdot \overline{1} = O(r^n)$. Поэтому, с учетом (11) и (4), справедлива оценка $Q_3^{(1)}(t) \leq C_{33}^{t-1} \cdot \overline{1}$.

Известно следующее представление для степени произвольной матрицы C [3]:

$$C^t = \sum_{l=1}^{s} \left(\lambda_l^t Z_{l1} + \left(\lambda_l^t\right)' \cdot Z_{l2} + \ldots + \left(\lambda_l^t\right)^{(m_l - 1)} Z_{lm_l}\right), \tag{21}$$

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

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

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

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 (22) 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. С учетом оценки (16) для $Q[K_i](t)$ и оценки из леммы 3 для E_t^i при $w = 1$ получим 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 так:

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

Просуммировав уравнение от 1 до t, найдем $x_t = x_1 + c + o(1)$, 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 $c_i^{(1)} = x_1 + c$.

Применяя лемму 1 к неразложимой матрице A_{ii} с учетом оценки для E_t^i, 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 $U''^{(1)} = \sum_i c_i^{(1)} U_i^{(21)}$ и $U_i^{(21)}$ определяется в соответствии с (6).

Используя (20), 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 C^l_{11}, $C^{(l)}_{12}$, $C^{(l)}_{13}$, следующих из теоремы 1, и оценок для $Q_1^{(1)}(n)$ и $Q_3^{(1)}(n)$ получим:

$$Q_1^{(1)}(t+1) = \left(C^{(l)}_{12} \cdot Q_2^{(1)}(n)\right) \cdot (1+o(1)) = \sum_i c_i^{(1)} U_i^{(11)} r^{t+1} \cdot (1+o(1)).$$

Здесь мы учли тот факт, что $V_i^{(21)} U''^{(1)} = V_i^{(21)} U_i^{(21)} = 1$.

Введем обозначение $U'^{(1)} = \sum_i c_i^{(1)} U_i^{(11)}$. Тогда

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

Таким образом,

$$Q^{(1)}(t) = \begin{pmatrix} Q_1^{(1)}(t) \\ Q_2^{(1)}(t) \\ Q_3^{(1)}(t) \end{pmatrix} = \begin{pmatrix} U'^{(1)} r^t \cdot (1+o(1)) \\ U''^{(1)} r^t \cdot (1+o(1)) \\ o(r^t) \end{pmatrix}, \tag{23}$$

и утверждение теоремы при $w = 1$ справедливо.

Пусть теперь $w = 2$.

Используя (20), запишем систему для $Q(t+1)$:

$$\begin{cases} Q^{(1)}(t+1) = \left(B^l_{11} Q^{(1)}(n) + B^{(l)}_{12} Q^{(2)}(n)\right)(1 + O((r'')^n)) \\ Q^{(2)}(t+1) = \left(B^l_{22} Q^{(2)}(n)\right)(1 + O((r'')^n)). \end{cases} \tag{24}$$

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 (17), 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 (8) для B_{11} и B_{12}, а также учитывая оценки для $Q^{(1)}(t)$ и $Q^{(2)}(t)$, следующие из (18), запишем уравнение для $Q_2^{(1)}(t)$:

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

$$C_{22} Q_2^{(1)}(t) + C_{23} U'^{(2)} r^t + D_{21} U''^{(2)} r^t + o(r^t) + O(t^{s^*_{12}-1} \cdot r^{2t}). \tag{25}$$

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

$$V_i^{(21)}Q_2^{(1)}(t+1) = c_i^{(1)}tr^{t+1}(1 + o(1)). \tag{26}$$

Используя первое уравнение системы (24), запишем уравнение для $Q_2^{(1)}(t)$:

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

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

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

Учитывая (26), а также равенство $t - n = \lfloor \log_2 t \rfloor$, получим, что

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

Отсюда следует, что для любого 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 (24):

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

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

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, найдем, что

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

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 $Q^{(1)}(t)$:

$$Q^{(1)}(t+1) = B_{11}Q^{(1)}(t) + B_{12}Q^{(2)}(t) + B_{13}Q^{(3)}(t) + \ldots + B_{1w}Q^{(w)}(t)(1 + O(t^{q_1^*-1}r^t)). \quad (27)$$

Для последовательности групп $\mathcal{M}_2, \mathcal{M}_3, \ldots, \mathcal{M}_w$ утверждение теоремы справедливо по предположению индукции. Подставим значения $Q^{(l)}(t)$ $(l = 2, \ldots, w)$ в (27). Так как q_l^* имеет наибольшее значение для $l = 2$, определяющим будет слагаемое $B_{12}Q^{(2)}(t)$. Поэтому уравнение (27) можно записать в следующем виде:

$$Q^{(1)}(t+1) = B_{11}Q^{(1)}(t) + B_{12}Q^{(2)}(t)(1 + O(t^{q_1^*-1}r^t)).$$

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.

Обозначим D_i^t множество всех деревьев вывода высоты t для слов из L_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

$$P(D_i^t) = d_i t^{q_l^*-1} r^{t-1} \cdot (1-r) \cdot (1 + o(1)), \quad (28)$$

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 $\mathcal{M}_{l+1,1}$.

4 Закономерности в деревьях вывода слов стохастического КС-языка

Для доказательства основного результата раздела предварительно докажем две леммы.

Через $R_X(n)$ обозначим выражение

$$\prod_{j=1}^{k} (1 - Q_j(n))^{x_j} - \prod_{j=1}^{k} (1 - Q_j(n-1))^{x_j},$$

где $X = (x_1, \ldots, x_k)$ — целочисленный неотрицательный вектор, $Q_j(n)$ — вероятности продолжения $(j = 1, \ldots, k)$, k — общее число нетерминалов в грамматике.

Лемма 4. *При* $n \to \infty$

$$\sum_{l=1}^{k} x_l P(D_l^n) \cdot (1+\varphi_l(n-1)) \cdot \prod_{j=1}^{k} (1-Q_j(n-1))^{x_j} \le R_X(n) \le$$

$$\sum_{l=1}^{k} x_l P(D_l^n) \cdot (1+\varphi_l(n)) \cdot \prod_{j=1}^{k} (1-Q_j(n))^{x_j}\,, \qquad (29)$$

где $\varphi_l(n) = \frac{Q_l(n)}{1-Q_l(n)}$.

Доказательство.

Проведем доказательство индукцией по k. При $k=1$ имеем:

$$R_X(n) = (1-Q_1(n))^{x_1} - (1-Q_1(n-1))^{x_1} =$$

$$((1-Q_1(n)) - (1-Q_1(n-1))) \cdot \sum_{l=0}^{x_1-1} (1-Q_1(n))^{x_1-1-l} \cdot (1-Q_1(n-1))^l\,.$$

Так как

$$(1-Q_1(n)) - (1-Q_1(n-1)) = Q_1(n-1) - Q_1(n) = P(D_1^n),$$

то

$$R_X(n) = P(D_1^n) \cdot \sum_{l=0}^{x_1-1} (1-Q_1(n))^{x_1-1-l} \cdot (1-Q_1(n-1))^l\,.$$

Положим $\varphi_j(n) = \frac{1}{1-Q_j(n)} - 1 = \frac{Q_j(n)}{1-Q_j(n)}$ $(j=1,\ldots,k)$. Поскольку $Q_1(n-1) \ge Q_1(n)$, можно записать систему неравенств

$$P(D_1^n) \cdot (1+\varphi_1(n-1)) \cdot x_1 \cdot (1-Q_1(n-1))^{x_1} \le R_X(n) \le$$

$$P(D_1^n) \cdot (1+\varphi_1(n)) \cdot x_1 \cdot (1-Q_1(n))^{x_1}.$$

Таким образом, доказана справедливость (29) для $k=1$.

Предположим, что соотношения (29) справедливы при $k-1$. Добавим к $R_X(n)$ и вычтем слагаемое $(1-Q_k(n))^{x_k} \prod_{j=1}^{k-1} (1-Q_j(n-1))^{x_j}$. Тогда

$$R_X(n) = \prod_{j=1}^{k} (1-Q_j(n))^{x_j} - (1-Q_k(n))^{x_k} \prod_{j=1}^{k-1} (1-Q_j(n-1))^{x_j} +$$

$$(1-Q_k(n))^{x_k} \prod_{j=1}^{k-1} (1-Q_j(n-1))^{x_j} - \prod_{j=1}^{k} (1-Q_j(n-1))^{x_j} =$$

$$(1-Q_k(n))^{x_k} \cdot \left[\prod_{j=1}^{k-1} (1-Q_j(n))^{x_j} - \prod_{j=1}^{k-1} (1-Q_j(n-1))^{x_j}\right] +$$

$$((1-Q_k(n))^{x_k} - (1-Q_k(n-1))^{x_k}) \cdot \prod_{j=1}^{k-1} (1-Q_j(n-1))^{x_j}\,.$$

Очевидно,

$$P(D_k^n)\cdot(1+\varphi_k(n-1))\cdot x_k\cdot(1-Q_k(n-1))^{x_k}\le(1-Q_k(n))^{x_k}-(1-Q_k(n-1))^{x_k}\le$$

$$P(D_k^n)\cdot(1+\varphi_k(n))\cdot x_k\cdot(1-Q_1(n))^{x_k},$$

так как доказательство этого факта полностью совпадает с доказательством (29) при $k=1$. Кроме того, выражение в квадратных скобках есть $R_X(n)$ при $k-1$. Поэтому

$$R_X(n)\le(1-Q_k(n))^{x_k}\cdot\sum_{l=1}^{k-1}x_lP(D_l^n)\cdot(1+\varphi_l(n))\cdot\prod_{j=1}^{k-1}(1-Q_j(n))^{x_j}+$$

$$P(D_k^n)\cdot(1+\varphi_k(n))\cdot x_k\cdot(1-Q_k(n))^{x_k}\cdot\prod_{j=1}^{k-1}(1-Q_j(n-1))^{x_j}\le$$

$$\prod_{j=1}^{k}(1-Q_j(n))^{x_j}\cdot\sum_{l=1}^{k}x_lP(D_l^n)\cdot(1+\varphi_l(n)).$$

Аналогично доказывается неравенство

$$R_X(n)\ge\prod_{j=1}^{k}(1-Q_j(n-1))^{x_j}\cdot\sum_{l=1}^{k}x_lP(D_l^n)\cdot(1+\varphi_l(n-1)).$$

Лемма доказана.

Заметим, что в доказательстве леммы не используется вид функции $Q_j(n)$, а используется лишь условие $Q_j(n-1)\ge Q_j(n)$.

Лемма 5. *Пусть $X=(x_1,\dots,x_k)$ – неотрицательный целочисленный вектор и n – натуральное число. Тогда при $n\to\infty$*

$$R_X(n)=(1+\psi_X(n))\sum_{l=1}^{k}x_lP(D_l^n),$$

где $-\tilde{c}_1n^{q_1^*-1}r^n\cdot\sum x_j\le\psi_X(n)\le\tilde{c}_2n^{q_1^*-1}r^n$ и $\tilde{c}_1$ и $\tilde{c}_2$ – некоторые положительные константы.

Доказательство.

Найдем верхнюю оценку для $R_X(n)$. Используя (29), можно записать, что

$$R_X(n)\le\prod_{j=1}^{k}(1-Q_j(n))^{x_j}\sum_{l=1}^{k}x_lP(D_l^n)\cdot(1+\varphi_l(n)).$$

Заметим, что $\prod_{j=1}^{k}(1-Q_j(n))^{x_j}\le1$ и $1+\varphi_l(n)=\frac{1}{1-Q_l(n)}$. Поэтому

$$R_X(n)\le\sum_{l=1}^{k}x_lP(D_l^n)\cdot\frac{1}{1-Q_l(n)}.$$

Применим теорему 2 для $Q_l(n)$, учитывая, что $q_1^* = \max_{\mathcal{M}_i}\{q_i^*\}$. Тогда при $n \to \infty$

$$R_X(n) \le \sum_{l=1}^{k} x_l P(D_l^n) \cdot (1 + \tilde{c}_2 n^{q_1^*-1} r^n), \text{ где } \tilde{c}_2 > 0.$$

Получим теперь нижнюю оценку для $R_X(n)$. Используя (29), можно записать, что

$$R_X(n) \ge \prod_{j=1}^{k} (1 - Q_j(n-1))^{x_j} \sum_{l=1}^{k} x_l P(D_l^n) \cdot (1 + \varphi_l(n-1)).$$

Для оценки выражения $\prod_{j=1}^{k} (1 - Q_j(n-1))^{x_j}$ используем следующее равенство, доказанное в [7]:

$$(1-y_1)^{n_1} \ldots (1-y_k)^{n_k} = 1 - \Delta_1, \text{ где } 0 \le \Delta_1 \le \sum_j n_j y_j. \tag{30}$$

Применяя (30), получаем, что

$$R_X(n) \ge \left(1 - \sum_{j=1}^{k} x_j Q_j(n-1)\right) \sum_{l=1}^{k} x_l P(D_l^n) \cdot (1 + \varphi_l(n-1)).$$

Так как $(1 + \varphi_l(n-1)) = \frac{1}{1 - Q_l(n-1)} \ge 1$, то

$$R_X(n) \ge \left(1 - \sum_{j=1}^{k} x_j Q_j(n-1)\right) \sum_{l=1}^{k} x_l P(D_l^n) \ge \left(1 - \tilde{c}_1 n^{q_1^*-1} r^n \sum_{j=1}^{k} x_j\right) \sum_{l=1}^{k} x_l P(D_l^n)$$

для некоторой положительной константы $\tilde{c}_1$.

Лемма доказана.

Будем полагать, как и ранее, что аксиомой исходной грамматики G является нетерминал A_1. Рассмотрим D_1^t — множество деревьев из D_1 высоты t. Для $d \in D_1^t$ через $p_t(d)$ будем обозначать условную вероятность дерева d, т.е. $p_t(d) = \frac{p(d)}{P(D_1^t)}$.

Через $M_i(t, \tau)$ обозначим условное математическое ожидание числа вершин на ярусе τ, помеченных нетерминалом A_i, в деревьях вывода высоты t.

Для нетерминала $A_l \in K_j$ положим $q_l' = q_j$ и $s_{1l}' = s_{1j}$.

Теорема 3. *Пусть G — стохастическая КС-грамматика с разложимой матрицей первых моментов, для которой перронов корень $r < 1$, и D_1^t – множество деревьев вывода высоты t.*

Тогда для любого $i \in \{1, \ldots, k\}$ при $\tau \to \infty$ и $t - \tau \to \infty$ выполняется асимптотическое равенство

$$M_i(t, \tau) \sim \sum_{l=1}^{k} \frac{f_{il} \cdot (t-\tau)^{q_l'-1} \cdot \tau^{\delta_{il}^1}}{t^{q_1-1}} + \frac{f_i \cdot (t-\tau)^{q_i'-1} \cdot \tau^{s_{1i}'-1}}{t^{q_1-1}},$$

в котором f_{il}, f_i - неотрицательные константы и δ_{il}^1 определено в (30).

Доказательство. Представим $M_i(t,\tau)$ в виде

$$M_i(t,\tau) = \sum_{d\in D_1^t} p_t(d) z_i(d,\tau) = \frac{1}{P(D_1^t)} \sum_{d\in D_1^t} p(d) z_i(d,\tau),$$

где $z_i(d,\tau)$ – число вершин на ярусе τ дерева d, помеченных A_i.

Рассмотрим неотрицательный целочисленный вектор $X = (x_1,\ldots,x_k)$, который будем называть далее вектором нетерминалов. Используя вектор X, мы можем записать, что

$$M_i(t,\tau) = \frac{1}{P(D_1^t)} \sum_{X\neq 0} \Delta_X,$$

где Δ_X – вклад в математическое ожидание тех деревьев вывода из D_1^t, которые на ярусе τ содержат x_j вершин, помеченных нетерминалом A_j $(j=1,\ldots,k)$. Множество таких деревьев обозначим через $D_X^t(\tau)$.

Пусть $d \in D_X^t(\tau)$. Выделим в d поддерево d_0 и последовательность поддеревьев $(d_1,\, d_2,\ldots d_n)$, где $n = \sum_{l=1}^k x_l$. Поддерево d_0 получено из d удалением всех вершин на ярусах $\tau+1,\, \tau+2,\ldots,t$ и инцидентных им дуг. Последовательность $(d_1,\, d_2,\ldots d_n)$ образуют все поддеревья, корни которых расположены на ярусе τ дерева d. При этом корни поддеревьев $d_1,\, d_2,\ldots d_m$ расположены в дереве d последовательно в порядке обхода вершин яруса τ слева направо, и каждое дерево d_l $(l=1,\ldots,n)$ содержит все дуги и вершины дерева d, лежащие на путях от корня d_l к листьям дерева d.

Выделим в $D_X^t(\tau)$ множество деревьев, имеющих в качестве поддерева d_0 одно и то же дерево. Обозначим это множество через D_0. Нетрудно понять, что

$$P\left(D_0\right) = p(d_0)\cdot\left(\prod_{l=1}^{k}(1-Q_l(t-\tau))^{x_l} - \prod_{l=1}^{k}(1-Q_l(t-\tau-1))^{x_l}\right), \qquad (31)$$

где $Q_l(n)$ — суммарная вероятность деревьев из множества D_l, высота которых больше n, и, следовательно, $(1-Q_l(n))$ – суммарная вероятность деревьев из D_l, высота которых не превосходит n.

Обозначим через $\delta_1(X)$ выражение $\prod_{l=1}^{k}(1-Q_l(t-\tau))^{x_l}$ и через $\delta_2(X)$ — выражение $\prod_{l=1}^{k}(1-Q_l(t-\tau-1))^{x_l}$.

В (31) величина $p(d_0)\cdot\delta_1(X)$ есть суммарная вероятность деревьев, определяемых поддеревом d_0, высота которых не превосходит t, так как каждое поддерево с корнем на ярусе τ имеет высоту, не превосходящую $(t-\tau)$.

Вторая величина $p(d_0)\cdot\delta_2(X)$ есть суммарная вероятность деревьев, определяемых поддеревом d_0, высота которых не превосходит $(t-\tau-1)$.

Разность этих величин равна, очевидно, суммарной вероятности деревьев высоты t, определяемых деревом d_0, и значение $\delta_1(X)-\delta_2(X)$ не зависит от порядка следования вершин на ярусе τ, помеченных нетерминалами. Поэтому

$$P(D_X^t(\tau)) = \sum_{d_0} p(d_0)\cdot(\delta_1(X)-\delta_2(X)) = (\delta_1(X)-\delta_2(X))\sum_{d_0} p(d_0),$$

где суммирование ведется по всем возможным поддеревьям d_0 деревьев из $D_X^t(\tau)$.

Для каждой вершины, помеченной некоторым нетерминалом A_l, суммарная вероятность возможных деревьев с корнем в этой вершине и листьями, помеченными только терминалами, равна $P(D_l)$. Ввиду согласованности исходной грамматики $P(D_l) = 1$ для любого l. Поэтому $\sum_{d_0} p(d_0)$ равна вероятности деревьев вывода из D_1, имеющих x_l вершин на ярусе τ, помеченных нетерминалом A_l $(l = 1, \ldots, k)$:

$$\sum_{d_0} p(d_0) = \sum_{d_0} p(d_0) \cdot P(D_1)^{x_1} \cdot P(D_2)^{x_2} \cdot \ldots \cdot P(D_k)^{x_k} = \sum_{d \in D_X(\tau)} p(d),$$

где $D_X(\tau)$ — множество деревьев из D_1, имеющих x_j вершин на ярусе τ, помеченных A_j $(j = 1, \ldots, k)$.

Будем обозначать $\sum_{d \in D_X(\tau)} p(d)$ через $P_X(\tau)$. Таким образом,

$$M_i(t,\tau) = \frac{1}{P(D_1^t)} \sum_{X \neq 0} P_X(\tau) \cdot (\delta_1(X) - \delta_2(X)) \cdot x_i.$$

Ранее величина $(\delta_1(X) - \delta_2(X))$ была обозначена через $R_X(t-\tau)$. Применим лемму 5 для представления $R_X(t-\tau)$. Получим, что

$$M_i(t,\tau) = \frac{1}{P(D_1^t)} \sum_{X \neq 0} P_X(\tau) \cdot x_i \cdot (1 + \psi_X(t-\tau)) \sum_{l=1}^{k} x_l P(D_l^{t-\tau}) =$$

$$\sum_{l=1}^{k} \frac{P(D_l^{t-\tau})}{P(D_1^t)} \sum_{X \neq 0} P_X(\tau) \cdot x_i \cdot x_l \cdot (1 + \psi_X(t-\tau)).$$

Отдельно вычислим $S_1 = \sum_{X \neq 0} P_X(\tau) \cdot x_i x_l$ и $S_2 = \sum_{X \neq 0} P_X(\tau) \cdot x_i \cdot x_l \cdot \psi_X(t-\tau)$.

Используя первые и вторые моменты, мы можем записать, что $S_1 = b_{il}^1(\tau)$ при $i \neq l$ и $S_1 = b_{ii}^1(\tau) + a_i^1(\tau)$ при $l = i$.

Учитывая оценку из леммы 5 для $\psi_X(n)$ и используя первые три момента, получим нижнюю и верхнюю оценки для S_2 :

$$S_2 = \sum_{X \neq 0} P_X(\tau) \cdot x_i \cdot x_l \cdot \psi_X(t-\tau) \geq$$

$$-c_2 \tau^{q_1^*-1} r^\tau \sum_{X \neq 0} P_X(\tau) \cdot x_i \cdot x_l \cdot \sum_j x_j = -c_2 \tau^{q_1^*-1} r^\tau \sum_j c_{ilj}^{1*}(\tau),$$

где

$$c_{ilj}^{1*}(\tau) = c_{ilj}^1(\tau) \ \text{ при } i \neq l, \ i \neq j \ \text{ и } j \neq l,$$
$$c_{iii}^{1*}(\tau) = c_{iii}^1(\tau) + 3b_{ii}^1(\tau) - a_i^1(\tau),$$

и

$$c_{ijj}^{1*}(\tau) = c_{ijj}^1(\tau) + b_{ij}^1(\tau), \ c_{iij}^{1*}(\tau) = c_{iij}^1(\tau) + b_{ij}^1(\tau).$$

Применяя оценки для первых трех моментов, получим, что $S_2 \geq -c \cdot \tau^{2q_1^*-2} r^{2\tau}$, где c — некоторая положительная константа.

С другой стороны,

$$S_2 \leq c_1 \tau^{q_1^*-1} r^\tau \sum_{X \neq 0} P_X(\tau) \cdot x_i \cdot x_l = c_1 \tau^{q_1^*-1} r^\tau \cdot S_1.$$

Так как $S_1 = b^1_{il}(\tau)$ при $i \neq l$ и $S_1 = b^1_{ii}(\tau) + a^1_i(\tau)$, то, с учетом оценок для моментов, получаем, что $S_2 = O(\tau^{2q_1^*-2} r^{2\tau})$.

Вернемся к вычислению $M_i(t, \tau)$:

$$M_i(t,\tau) = \sum_{l \neq i} \frac{P(D_l^{t-\tau})}{P(D_1^t)} b^1_{il}(\tau) + \frac{P(D_i^{t-\tau})}{P(D_1^t)} \left(b^1_{ii}(\tau) + a^1_i(\tau)\right) + \sum_{l=1}^{k} \frac{P(D_l^{t-\tau})}{P(D_1^t)} \cdot O\left(\tau^{2q_1^*-2} r^{2\tau}\right).$$

Раскрывая моменты и используя лемму 5, после несложных преобразований получим, что

$$M_i(t,\tau) =$$

$$\sum_{l=1}^{k} \frac{d_l \cdot (1-r) \cdot (t-\tau)^{q'_l-1} \cdot r^{t-\tau-1} \cdot (1+\phi_l(t-\tau))}{d_1 \cdot (1-r) \cdot t^{q_1-1} \cdot r^{t-1} \cdot (1+\phi_1(t))} \cdot \left(g^1_{il} \cdot r^\tau \cdot \tau^{\delta^1_{il}} \left(1 + \psi_{il}(\tau)\right)\right) +$$

$$\frac{d_i \cdot (1-r) \cdot (t-\tau)^{q'_i-1} r^{t-\tau-1} (1+\phi_i(t-\tau)) \cdot c_{1i} \cdot \tau^{s'_{1i}-1} r^\tau (1+\varphi_{1i}(n)(\tau))}{d_1 \cdot (1-r) \cdot t^{q_1-1} r^{t-1} \cdot (1+\phi_1(t))} + O\left(\tau^{2q_1^*-2} r^{2\tau}\right),$$

где $\phi_i(n) = o(1)$, $\psi_{il}(n) = o(1)$, $\varphi_{1i}(n) = o(1)$, $q'_l = q_j$ для $A_l \in K_j$, и $s'_{1i} = s_{1m}$ для $A_i \in K_m$.

Отсюда следует, что

$$M_i(t,\tau) = \frac{1}{d_1} \left(\sum_{l=1}^{k} \frac{d_l \cdot g^1_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta^1_{il}}}{t^{q_1-1}} + \frac{d_i \cdot c_{1i} \cdot (t-\tau)^{q'_i-1} \cdot \tau^{s'_{1i}-1}}{t^{q_1-1}} \right) \left(1 + \xi(\tau, t-\tau)\right), \qquad (32)$$

где $\xi(\tau, t-\tau) \to 0$ при $\tau, t-\tau \to \infty$. Теорема доказана.

Рассмотрим подробнее слагаемые в (32). Определяющими в сумме являются те значения l, для которых $g^1_{il} > 0$ и $q'_l + \delta^1_{il} = q_1$. Равенство справедливо при одновременном выполнении следующих условий:

1) нетерминал A_l принадлежит классу K_{j_1} с $j_1 \in J_{MAX}$,

2) $A_i \in K_{j_2}$, для которого $K_{j_1} \prec_* K_{j_2}$.

Обозначим N_i множество номеров l, для которых выполняются условия 1) и 2).

Отметим, что слагаемое $\frac{d_i \cdot c_{1i} \cdot (t-\tau)^{q'_i-1} \cdot \tau^{s'_{1i}-1}}{t^{q_1-1}}$ влияет на значение $M_i(t,\tau)$ при $s'_{1i} + q'_i - 1 = q_1$. Это равенство выполняется в случае, если $A_i \in K_{j_2}$, где $j_2 \in J_{MAX}$. Поэтому равенство (32) при $N_i \neq \emptyset$ можно записать в виде

$$M_i(t,\tau) = \left(\sum_{l \in N_i} \frac{f_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta^1_{il}}}{t^{q_1-1}} + \frac{f_i \cdot (t-\tau)^{q'_i-1} \cdot \tau^{s'_{1i}-1}}{t^{q_1-1}} \right) \left(1 + \xi(\tau, t-\tau)\right),$$

где $f_{il} = \frac{d_l \cdot g^1_{il}}{d_1}$, $f_i = \frac{d_i \cdot c_{1i}}{d_1}$ и $\xi(\tau, t-\tau) \to 0$ при $\tau, t-\tau \to \infty$. Очевидно, $M_i(t,\tau) \leq O(1/t)$ при $N_i = \emptyset$. Поэтому справедливо

Следствие.

$$1) M_i(t,\tau) = \left(\sum_{l \in N_i} \frac{f_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta^1_{il}}}{t^{q_1-1}} + \frac{f_i \cdot (t-\tau)^{q'_i-1} \cdot \tau^{s'_{1i}-1}}{t^{q_1-1}} \right) \left(1 + \xi(\tau, t-\tau)\right)$$

при $N_i \neq \emptyset$;

2) $M_i(t,\tau) \le O(1/t)$ *при* $N_i = \emptyset$.

Пусть r_{ij} — произвольное правило грамматики G. Через $s_l^{(ij)}$ обозначим число нетерминалов A_l в правой части правила r_{ij}. Условное математическое ожидание числа применений правила r_{ij} в деревьях вывода высоты t на ярусе τ будем обозначать через $M_{ij}(t,\tau)$.

Теорема 4. *Пусть G — стохастическая КС-грамматика с разложимой матрицей первых моментов, для которой перронов корень $r < 1$, и D_1^t – множество деревьев вывода высоты t.*

Тогда при $\tau \to \infty$ и $t-\tau \to \infty$ выполняется следующее асимптотическое равенство:

$$M_{ij}(t,\tau) \sim \frac{p_{ij}}{t^{q_1-1}} \left(\sum_{l=1}^{k} f_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta^1_{il}} + \frac{1}{r} \sum_{m=1}^{k} f_m \cdot s_m^{(ij)} \cdot (t-\tau)^{q'_m-1} \tau^{s'_{1i}-1} \right).$$

В формулировке теоремы p_{ij} — вероятность правила r_{ij}, задаваемая в исходной грамматике, $s_m^{(ij)}$ – число нетерминалов A_m в правой части правила r_{ij}, а величины q'_l, δ^1_{il}, f_{il} и f_m имеют тот же смысл, что и в теореме 3.

Доказательство. Обозначим $z_{ij}(d,\tau)$ число вершин на ярусе τ дерева d, помеченных нетерминалом A_i, к которым применено правило r_{ij}. Используя неотрицательный целочисленный вектор $X = (x_1, \ldots, x_k)$, можно записать, что

$$M_{ij}(t,\tau) = \sum_{X \neq 0} \sum_{d \in D_X^t(\tau)} p_t(d) z_{ij}(d,\tau),$$

где $D_X^t(\tau)$ введено в доказательстве теоремы 3.

Представим $z_{ij}(d,\tau)$ в виде суммы случайных величин $I_1 + I_2 + \ldots + I_{x_i}$, где $I_m = 1$, если к m-й по порядку вершине среди вершин, помеченных нетерминалом A_m на ярусе τ, применено правило r_{ij}, и $I_m = 0$ в противном случае $(m = 1, 2, \ldots, x_i)$. Тогда

$$M_{ij}(t,\tau) = \sum_{X \neq 0} \sum_{d \in D_X^t(\tau)} p_t(d) \cdot (I_1 + I_2 + \ldots + I_{x_i}).$$

Очевидно, что случайные величины I_m $(m = 1, 2, \ldots, x_i)$ – одинаково распределены на $D_X^t(\tau)$, поэтому

$$M_{ij}(t,\tau) = \frac{1}{P(D_1^t)} \sum_{X \neq 0} P\left(D_{X,1}^t(\tau)\right) \cdot x_i,$$

где $P(D_{X,1}^t(\tau))$ – суммарная вероятность тех деревьев из $D_X^t(\tau)$, в которых правило r_{ij} применено к первой по порядку вершине на ярусе τ, помеченной A_i.

Подсчитаем вероятность $P\left(D_{X,1}^t(\tau)\right)$:

$$P\left(D_{X,1}^t(\tau)\right) = p_{ij} \cdot P_X(\tau) \times \left[\prod_{m=1}^{k} \left(1 - Q_m(t-\tau)\right)^{x'_m} \cdot \prod_{m=1}^{k} \left(1 - Q_m(t-\tau-1)\right)^{s_m^{(ij)}} - \right.$$

$$\prod_{m=1}^{k}(1-Q_m(t-\tau-1))^{x'_m}\cdot\prod_{m=1}^{k}(1-Q_m(t-\tau-2))^{s_m^{(ij)}}\Bigg]. \qquad (33)$$

Здесь $X'=(x'_1,\dots,x'_k)=(x_1,\dots,x_{i-1},x_i-1,x_{i+1},\dots,x_k)$ и $S=(s_1^{(ij)},\dots,s_k^{(ij)})$, где $s_m^{(ij)}$ равно числу нетерминалов A_m в правой части правила r_{ij} ($m=1,\dots,k$). Величина $P_X(\tau)$ имеет тот же смысл, что и в доказательстве теоремы 3. Выражение в квадратных скобках в (33) аналогично выражению $R_X(t-\tau)$. При этом с помощью множителей $(1-Q_m(t-\tau-1))^{s_m^{(ij)}}$ и $(1-Q_m(t-\tau-2))^{s_m^{(ij)}}$ учитывается тот факт, что к первому нетерминалу A_i на ярусе τ применено правило r_{ij}, которому на ярусе $\tau+1$ соответствует $s_m^{(ij)}$ вершин, помеченных нетерминалом A_m ($m=1,\dots,k$).

Проведем несложные преобразования в (33):

$$P\left(D_{X,1}^{t}(\tau)\right)=p_{ij}\cdot P_X(\tau)\cdot\frac{1}{1-Q_i(t-\tau)}\cdot\prod_{m=1}^{k}(1-Q_m(t-\tau-1))^{s_m^{(ij)}}\times$$

$$\left[\prod_{m=1}^{k}(1-Q_m(t-\tau))^{x_m}-\frac{1-Q_i(t-\tau)}{1-Q_i(t-\tau-1)}\times\right.$$

$$\left.\prod_{m=1}^{k}\left((1-Q_m(t-\tau-1))^{x_m}\cdot\frac{(1-Q_m(t-\tau-2))^{s_m^{(ij)}}}{(1-Q_m(t-\tau-1))^{s_m^{(ij)}}}\right)\right].$$

Очевидно, что

$$\frac{1-Q_i(t-\tau)}{1-Q_i(t-\tau-1)}=1+\frac{Q_i(t-\tau-1)-Q_i(t-\tau)}{1-Q_i(t-\tau-1)}=$$

$$1+\frac{P(D_i^{t-\tau})}{1-Q_i(t-\tau-1)}=1+P(D_i^{t-\tau})+\frac{P(D_i^{t-\tau})\cdot Q_i(t-\tau-1)}{1-Q_i(t-\tau-1)}.$$

Применим теорему 2 и следствие из нее для оценки $Q_i(t-\tau-1)$ и $P(D_i^{t-\tau})$. Получим, что

$$\frac{1-Q_i(t-\tau)}{1-Q_i(t-\tau-1)}=1+P(D_i^{t-\tau})+O((t-\tau)^{2(q_1-1)}\cdot r^{2(t-\tau)}).$$

Проводя аналогичные преобразования и учитывая, что $s_m^{(ij)}$ – константа, определяемая правой частью правила r_{ij}, мы можем записать, что

$$\prod_{m=1}^{k}\frac{(1-Q_m(t-\tau-2))^{s_m^{(ij)}}}{(1-Q_m(t-\tau-1))^{s_m^{(ij)}}}=1-\sum_{m=1}^{k}s_m^{(ij)}\cdot P\left(D_m^{t-\tau-1}\right)+O\left((t-\tau)^{2(q_1-1)}\cdot r^{2(t-\tau)}\right).$$

Поэтому

$$P\left(D_{X,1}^{t}(\tau)\right)=p_{ij}\cdot P_X(\tau)\cdot\left(1+O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\right)\right)\left[\prod_{m=1}^{k}(1-Q_m(t-\tau))^{x_m}-\right.$$

$$\prod_{m=1}^{k}(1-Q_m(t-\tau-1))^{x_m}\cdot\left(1+P(D_i^{t-\tau})+O((t-\tau)^{2(q_1-1)}\cdot r^{2(t-\tau)})\right)\times$$

$$\left.\left(1-\sum_{m=1}^{k}s_m^{(ij)}\cdot P\left(D_m^{t-\tau-1}\right)+O\left((t-\tau)^{2(q_1-1)}\cdot r^{2(t-\tau)}\right)\right)\right]=$$

$$p_{ij}\cdot P_X(\tau)\left[R_X(t-\tau)+\prod_{m=1}^{k}(1-Q_m(t-\tau-1))^{x_m}\times\right.$$

$$\left.\left(\sum_{m=1}^{k}s_m^{(ij)}P(D_m^{t-\tau-1})-P(D_i^{t-\tau})\right)\right]\cdot\left(1+O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\right)\right).$$

(Здесь $R_X(t-\tau)$ – величина, рассмотренная в лемме 4.)

Вернемся к вычислению $M_{ij}(t,\tau)$, учитывая оценку

$$\prod_{m=1}^{k}(1-Q_m(t-\tau-1))^{x_m}=1-O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\sum_{m=1}^{k}x_m\right),$$

следующую из (30). Тогда

$$M_{ij}(t,\tau)=\frac{1}{P(D_1^t)}\left[\sum_{X\neq0}p_{ij}\cdot P_X(\tau)\cdot R_X(t-\tau)\cdot x_i+\left(\sum_{m=1}^{k}s_m^{(ij)}P\left(D_m^{t-\tau-1}\right)-P\left(D_i^{t-\tau}\right)\right)\times\right.$$

$$\left.\sum_{X\neq0}p_{ij}\cdot P_X(\tau)\cdot x_i\cdot\left(1-O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\sum_{m=1}^{k}x_m\right)\right)\right]\cdot\left(1+O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\right)\right).$$

Величина

$$\frac{1}{P(D_1^t)}\cdot\sum_{X\neq0}P_X(\tau)\cdot R_X(t-\tau)\cdot x_i$$

есть $M_i(t,\tau)$ из теоремы 3, и

$$\sum_{X\neq0}P_X(\tau)\cdot x_i=a_i^1(\tau)=c_{1i}\cdot\tau^{s'_{1i}-1}r^{\tau}(1+o(1)),$$

где $a_i^1(\tau)$ – элемент матрицы A^{τ}, A – матрица первых моментов, и s'_{1i} имеет тот же смысл, что и в теореме 3.

Кроме того,

$$\sum_{X\neq0}P_X(\tau)\cdot x_i\cdot O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\sum_{m=1}^{k}x_m\right)=$$

$$O\left((t-\tau)^{q_1-1}\cdot r^{t-\tau}\right)\sum_{m=1}^{k}b_{im}^1(\tau)=O\left(\tau^{q_1-1}\cdot(t-\tau)^{q_1-1}r^{t}\right),$$

где $b_{im}^1(\tau)$ – вторые моменты. Следовательно,

$$M_{ij}(t,\tau)=\left(M_i(t,\tau)\cdot p_{ij}+\frac{p_{ij}\cdot c_{1i}\cdot\tau^{s'_{1i}-1}r^{\tau}\cdot(1+o(1))}{P(D_1^t)}\times\right.$$

$$\left(\sum_{m=1}^{k} s_m^{(ij)} \cdot P(D_m^{t-\tau-1}) - P(D_i^{t-\tau})\right)\Bigg) \cdot \left(1 + O\left((t-\tau)^{q_1-1} r^{t-\tau}\right)\right).$$

Применяя теорему 3 к $M_i(t,\tau)$ и формулу (28) к $P(D_m^n)$, после проведения несложных преобразований $M_{ij}(t,\tau)$ можем представить в следующем виде:

$$M_{ij}(t,\tau) = \frac{p_{ij}}{t^{q_1-1}}\left(\sum_{l=1}^{k} f_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta_{il}^1} + \frac{1}{r}\sum_{m=1}^{k} f_m \cdot s_m^{(ij)} \cdot (t-\tau)^{q'_m-1} \tau^{s'_{1i}-1}\right) +$$

$$\xi_{ij}^1(t) + \xi_{ij}^2(\tau) + \xi_{ij}^3(t-\tau),$$

где $\xi_{ij}^1(t) = o(1)$, $\xi_{ij}^2(\tau) = o(1)$ и $\xi_{ij}^3(t-\tau) = o(1)$.

Обозначим сумму $\xi_{ij}^1(t) + \xi_{ij}^2(\tau) + \xi_{ij}^3(t-\tau)$ через $\xi_{ij}(t,\tau)$. Очевидно, $\xi_{ij}(t,\tau) \to 0$ при $\tau \to \infty$ и $t-\tau \to \infty$. Поэтому

$$M_{ij}(t,\tau) =$$

$$\frac{p_{ij}}{t^{q_1-1}}\left(\sum_{l=1}^{k} f_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta_{il}^1} + \frac{1}{r}\sum_{m=1}^{k} f_m \cdot s_m^{(ij)} \cdot (t-\tau)^{q'_m-1} \tau^{s'_{1i}-1}\right) + \xi_{ij}(t,\tau). \qquad (34)$$

Теорема доказана.

Сделаем несколько выводов из теоремы 4.

1. $M_{ij}(t,\tau)$ ограничено константой при $\tau \to \infty$, $t-\tau \to \infty$.

2. Величина $\sum_{m=1}^{k} f_m \cdot s_m^{(ij)} \cdot (t-\tau)^{q'_m-1} \tau^{s'_{1i}-1}$ имеет большее значение для тех правил, которые содержат в правой части большее количество нетерминальных символов.

3. Величина $\sum_{l \in N_i} f_{il} \cdot (t-\tau)^{q'_l-1} \cdot \tau^{\delta_{il}^1}$ имеет одно и то же значение для всех правил грамматики с одинаковой левой частью A_i.

Пусть $S_{ij}(t) = q_{ij}(t,0) + q_{ij}(t,1) + \ldots + q_{ij}(t,t-1)$, где $q_{ij}(t,\tau)$ — число правил r_{ij} на ярусе τ в дереве из D_1^t; $S_{ij}(t)$ — число правил r_{ij} в дереве вывода из D_1^t.

Рассмотрим случайную величину $\frac{S_{ij}(t)}{t}$ — среднее число правил r_{ij}, приходящееся на один ярус дерева вывода из D_1^t.

Теорема 5. *Пусть G — стохастическая КС-грамматика с разложимой матрицей первых моментов, для которой перронов корень $r < 1$, и D_1^t – множество деревьев вывода высоты t.*

Тогда при $t \to \infty$ выполняется следующее асимптотическое равенство:

$$M\left(\frac{S_{ij}(t)}{t}\right) \sim w_{ij},$$

где w_{ij} - константа, определяемая грамматикой G.

Доказательство.

Разобьем $S_{ij}(t)$ на три части:

$$S_{ij}(t) = S_{ij}^{(1)}(t) + S_{ij}^{(2)}(t) + S_{ij}^{(3)}(t),$$

где

$$S_{ij}^{(1)}(t) = q_{ij}(t,0) + \ldots + q_{ij}(t,\tau_0 - 1),$$

$$S_{ij}^{(2)}(t) = q_{ij}(t, \tau_0) + \ldots + q_{ij}(t, t - \tau_0 - 1),$$

$$S_{ij}^{(3)}(t) = q_{ij}(t, t - \tau_0) + \ldots + q_{ij}(t, t - 1),$$

и положим $\tau_0 = \lfloor \log\log t \rfloor$ (здесь и далее логарифм берется по основанию 2). Число слагаемых в $S_{ij}^{(1)}(t)$ и в $S_{ij}^{(3)}(t)$ равно $\lfloor \log\log t \rfloor$, а в $S_{ij}^{(2)}(t)$ равно $t - 2\lfloor \log\log t \rfloor$.

Найдем математические ожидания $M\left(S_{ij}^{(1)}(t)\right)$, $M\left(S_{ij}^{(2)}(t)\right)$ и $M\left(S_{ij}^{(3)}(t)\right)$.

Величину $M\left(S_{ij}^{(1)}(t)\right)$ можно представить в следующем виде:

$$M\left(S_{ij}^{(1)}(t)\right) = M_{ij}(t, 0) + M_{ij}(t, 1) + \ldots + M_{ij}(t, \tau_0 - 1).$$

Число правил r_{ij} на ярусе τ в дереве из D_1^t обозначим $q_{ij}(t, \tau)$. Оценим $q_{ij}(t, \tau)$ для $\tau < \tau_0$. Обозначим через k_{max} максимальное число нетерминалов в правой части правил грамматики G. Тогда $q_{ij}(t, \tau) \le k_{max}^{\tau} < k_{max}^{\tau_0}$. Поэтому $M_{ij}(t, \tau) < k_{max}^{\tau_0}$ и

$$M\left(S_{ij}^{(1)}(t)\right) \le k_{max}^{\tau_0}\tau_0 \le k_{max}^{\log\log t} \log\log t = \log^{c_1} t \, \log\log t \le \log^{c_2} t,$$

где $c_1 = \log k_{max},\ c_2 = c_1 + 1$.

Для $t - \tau_0 \le \tau < t$ имеем:

$$M_{ij}(t, \tau) \le M_i(t, \tau) = \frac{1}{P(D^t)} \sum_X P_X(\tau) R_X(t - \tau) x_i \le \frac{1}{P(D^t)} \sum_X P_X(\tau) x_i =$$

$$\frac{1}{P(D^t)} a_i^1(\tau) \le O\left(\frac{\tau^{q_1 - 1}}{t^{q_1 - 1} \cdot r^{t-\tau}}\right) \le O\left(\frac{1}{r^{t-\tau}}\right).$$

Поэтому

$$M\left(S_{ij}^{(3)}(t)\right) \le \sum_{t-\tau_0}^{t-1} O\left(\frac{1}{r^{t-\tau}}\right) = O\left(\frac{\tau_0}{r^{\tau_0}}\right) = O\left(\frac{\log\log t}{r^{\log\log t}}\right) = O\left(\log^{c_3} t\right)$$

для некоторой константы $c_3 > 0$.

Для τ, удовлетворяющего условию $\tau_0 \le \tau \le t - \tau_0 - 1$, применим теорему 4:

$$M\left(S_{ij}^{(2)}(t)\right) = \sum_{\tau = \lfloor \log\log t \rfloor}^{t - \lfloor \log\log t \rfloor - 1} \frac{p_{ij}}{t^{q_1 - 1}} \left(\sum_{l=1}^{k} f_{il} \cdot (t - \tau)^{q'_l - 1} \cdot \tau^{\delta_{il}^1} + \right.$$

$$\left. \frac{1}{r} \sum_{m=1}^{k} f_m \cdot s_m^{(ij)} \cdot (t - \tau)^{q'_m - 1} \tau^{s'_{1i} - 1}\right) + \sum_{\tau = \lfloor \log\log t \rfloor}^{t - \lfloor \log\log t \rfloor - 1} \xi(t, \tau).$$

Оценим величину $\delta = \frac{1}{t^{n_1 + n_2}} \cdot \sum_{\tau = \lfloor \log\log t \rfloor}^{t - \lfloor \log\log t \rfloor - 1} (t - \tau)^{n_1} \cdot \tau^{n_2}$:

$$\delta = \sum_{\tau = \lfloor \log\log t \rfloor}^{t - \lfloor \log\log t \rfloor - 1} \left(1 - \frac{\tau}{t}\right)^{n_1} \left(\frac{\tau}{t}\right)^{n_2} =$$

$$\sum_{\tau=\lfloor\log\log t\rfloor}^{t-\lfloor\log\log t\rfloor-1}\sum_{n=0}^{n_1}(-1)^n C_{n_1}^n\left(\frac{\tau}{t}\right)^{n+n_2}=\left(\sum_{n=0}^{n_1}(-1)^n C_{n_1-1}^n\cdot\frac{t}{n+n_2+1}\right)\cdot(1+o(1)).$$

Очевидно, величина $\sum_{n=0}^{n_1}(-1)^n C_{n_1-1}^n\cdot\frac{1}{n+n_2+1}$ является константой, зависящей от n_1 и n_2, обозначим ее $\alpha(n_1,n_2)$. Применяя обозначение $\alpha(n_1,n_2)$, мы можем записать:

$$\delta=\alpha(n_1,n_2)\cdot t\cdot(1+o(1)).$$

Применим полученную оценку к вычислению $M\left(S_{ij}^{(2)}(t)\right)$, учитывая равенства $q_l' + \delta_{il}^1 = q_1$ и $q_i'+s_{1i}'-1=q_1$:

$$M\left(S_{ij}^{(2)}(t)\right)=p_{ij}\cdot\left[\sum_{l=1}^{k}f_{il}\cdot\alpha(q_l'-1,\delta_{il}^1)+\frac{1}{r}\sum_{m=1}^{k}f_m\cdot s_m^{(ij)}\cdot\alpha(q_m'-1,s_{1i}'-1)\right]t\cdot(1+o(1)).$$

Константу в квадратных скобках обозначим w_{ij}.

Применяя полученные оценки для $M\left(S_{ij}^{(1)}(t)\right)$, $M\left(S_{ij}^{(2)}(t)\right)$ и $M\left(S_{ij}^{(3)}(t)\right)$, находим, что при $t\to\infty$

$$M\left(\frac{S_{ij}(t)}{t}\right)=w_{ij}+o(1)+O\left(\frac{\log^{c_2}t}{t}\right)+O\left(\frac{\log^{c_3}t}{t}\right)=w_{ij}+o(1).$$

Теорема доказана.

5 Энтропия и нижняя оценка стоимости кодирования

Пусть L - стохастический язык, т.е. язык, на множестве слов которого задано распределение вероятностей.

Под энтропией стохастического языка L будем понимать величину

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

Если энтропия конечна, будем применять запись $H(L)=-\sum_{\alpha\in L}p(\alpha)\log p(\alpha)$.

Кодированием языка L назовем инъективное отображение

$$f:L\to\{0,1\}^+.$$

В качестве L рассмотрим язык, порождаемый стохастической КС-грамматикой с однозначным выводом, т.е. грамматикой, в которой каждое слово из L имеет единственное дерево вывода. Через L^t обозначим множество всех слов из L, каждое из которых имеет дерево вывода высоты t. Для $\alpha\in L^t$ через $p_t(\alpha)$ обозначим условную вероятность появления слова α, т.е. $p_t(\alpha)=\frac{p(\alpha)}{P(L^t)}$. В силу однозначности вывода $P(L^t)=P(D_1^t)$.

Стоимостью кодирования f назовем величину

$$C(L,f) = \lim_{t\to\infty} \frac{\sum_{\alpha\in L^t} p_t(\alpha)\cdot|f(\alpha)|}{\sum_{\alpha\in L^t} p_t(\alpha)\cdot|\alpha|} \tag{35}$$

(здесь $|x|$ -длина последовательности x).

Величина $C(L,f)$ характеризует число двоичных разрядов, приходящихся на кодирование одного символа слова языка.

Через $F(L)$ обозначим класс всех инъективных отображений из L в $\{0,1\}^+$, для которых существует $C(L,f)$.

Стоимостью оптимального кодирования языка L назовем величину

$$C_0(L) = \inf_{f\in F(L)} C(L,f).$$

Предварительно получим асимптотическую формулу для энтропии множества слов L^t. По определению имеем

$$H(L^t) = -\sum_{\alpha\in L^t} p_t(\alpha)\log p_t(\alpha).$$

Следовательно,

$$H(L^t) = -\sum_{\alpha\in L^t} p_t(\alpha)\left(\log p(\alpha) - \log P(L^t)\right) =$$

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

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

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

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

Очевидно, что $\sum_{\alpha\in L^t} p(\alpha)q_{ij}(\alpha) = P(L^t)\cdot M(S_{ij}(t))$. Используя теорему 5, выражение для энтропии можно переписать в виде

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

Ввиду однозначности вывода, с использованием (28), имеем

$$\log P(L^t) = \log P(D_1^t) = t\log r + O(\log t).$$

Поэтому

$$H(L^t) = t \cdot \left(\log r - \sum_{j=1}^{n_i} w_{ij} \log p_{ij} \right) + o(t).$$

Полученный результат сформулируем в виде следующей теоремы.

Теорема 6. *Пусть G — однозначная стохастическая КС-грамматика с разложимой матрицей первых моментов, для которой перронов корень $r < 1$, и L^t - множество всех слов из L, порождаемого G, с деревьями вывода высоты t. Тогда*

$$H(L^t) = t \cdot \left(\log r - \sum_{j=1}^{n_i} w_{ij} \log p_{ij} \right) + o(t),$$

где w_{ij} определяются теоремой 5.

Таким образом, энтропия $H(L^t)$ линейно зависит от высоты t дерева вывода, как и в неразложимом случае [5].

Используя энтропию, оценим стоимость оптимального кодирования $C_0(L)$. Обозначим через f^* кодирование множества L^t, минимизирующее величину

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

Очевидно, для любого кодирования $f \in F(L)$ верно неравенство $M_t(f) \geq M_t(f^*)$. Оценим $M^*(L^t) = M_t(f^*)$, используя следующую теорему, доказанную в [2].

Теорема 7. *Пусть L_k – последовательность стохастических языков, для которой $H(L_k) \to \infty$ при $k \to \infty$. Тогда*

$$\lim_{k \to \infty} \frac{M^*(L_k)}{H(L_k)} = 1.$$

Поскольку $H(L)^t \to \infty$ при $t \to \infty$, из теоремы 7 следует, что $M_t(f^*)/H(L^t) \to 1$ при $t \to \infty$.

Найдем величину $\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha|$. Пусть правило r_{ij} содержит в правой части l_{ij} терминальных символов. Очевидно, $|\alpha| = \sum_{ij} q_{ij}(\alpha) \cdot l_{ij}$. Поэтому

$$\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha| = \sum_{ij} l_{ij} M(S_{ij}(t)) = t \cdot \sum_{ij} l_{ij} w_{ij} + o(t).$$

Следовательно, справедлива

Теорема 8. *Пусть L - стохастический КС-язык, порожденный разложимой стохастической КС - грамматикой с однозначным выводом, для которой перронов корень r матрицы первых моментов меньше* 1. *Тогда стоимость любого кодирования $f \in F(L)$ удовлетворяет неравенству*

$$C(L, f) \geq C_0(L) = \frac{\log r - \sum_{ij} w_{ij} \log p_{ij}}{\sum_{ij} l_{ij} w_{ij}}.$$

Список литературы

[1] Ахо А., Ульман Дж. Теория синтаксического анализа, перевода и компиляции. Том 1. — М.: Мир, 1978.

[2] Борисов А.Е. О свойствах стохастического КС-языка, порожденного грамматикой с двумя классами нетерминальных символов // Дискретный анализ и исследование операций. 2005. — Серия 1, том 12, N3. — Новосибирск: Издательство Института математики СО РАН. — С.3 – 31.

[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. — М.: Наука. — 2000. — Вып.9. — С. 101 – 126.

[5] Жильцова Л.П. О нижней оценке стоимости кодирования и асимптотически оптимальном кодировании стохастического контекстно-свободного языка // Дискретный анализ и исследование операций. — 2001.— Серия 1. — Том 8, N3. — Новосибирск: Издательство Института математики СО РАН. — С. 26 – 45.

[6] Л.П. Жильцова. О матрице первых моментов разложимой стохастической КС-грамматики // Ученые записки Казанского государственного университета. Физико-математические науки. — Том 151, — книга 2, — 2009. — С. 80 – 89.

[7] Севастьянов Б. А. Ветвящиеся процессы. — М.: Наука, 1971.

[8] Фу К. Структурные методы в распознавании образов. — М.: Мир, 1977.