1 Основные определения

Стохастической КС-грамматикой [5] называется система $G =< V_T, V_N, R, s >$, где V_T и V_N — соответственно алфавиты терминальных и нетерминальных символов (*терминалов* и *нетерминалов*), s — *аксиома* грамматики, R — множество *правил вывода*, представимое в виде $R = \cup_{i=1}^{k} R_i$, где $k = |V_N|$, и R_i — множество правил вида

$$r_{ij} : A_i \xrightarrow{p_{ij}} \beta_{ij} \quad (A_i \in V_N, \beta_{ij} \in (V_N \cup V_T)^*)\,, \tag{1}$$

где p_{ij} — вероятность применения правила r_{ij}, причём при фиксированном i вероятности r_{ij} задают вероятностное распределение на множестве R_i:

$$0 < p_{ij} \leqslant 1 \quad \text{и} \quad \sum_{j=1}^{n_i} p_{ij} = 1, \qquad i = 1, 2, \ldots, k, \tag{2}$$

где $n_i = |R_i|$.

Слово β называется *непосредственно выводимым* из α (обозначается $\alpha \Rightarrow \beta$), если существуют $\alpha_1, \alpha_2 \in (V_N \cup V_N)^*$, для которых $\alpha = \alpha_1 A_i \alpha_2$, $\beta = \alpha_1 \beta_{ij} \alpha_2$ и в R имеется правило $A_i \xrightarrow{p_{ij}} \beta_{ij}$.

Через $\Rightarrow_*$ обозначим рефлексивное транзитивное замыкание $\Rightarrow$. Если $\alpha \Rightarrow_* \beta$, говорят, что β *выводимо* из α. Язык L_G, *порождаемый* грамматикой G определяется как множество слов $\alpha \in V_T^*$, выводимых из аксиомы s грамматики G.

Последовательность правил грамматики $(r_1, r_2, \ldots, r_k)$ называется *выводом* слова α (обозначается $\omega(\alpha)$), если существует последовательность слов $(\alpha_0, \alpha_1, \alpha_2, \ldots, \alpha_k)$ таких, что для любго l слово α_l получается из α_{l-1} при замене одного из нетерминалов α_{l-1} согласно правилу r_l, и при этом $\alpha_0 = s$ и $\alpha_k = \alpha$. В случае, если каждое α_l получено из α_{l-1} заменой самого левого нетерминала в α_{l-1} согласно правилу r_l, такой вывод $\omega(\alpha)$ называют *левым*.

Вероятность вывода $\omega(\alpha)$ определяется как $p(\omega(\alpha)) = p(r_1) \cdot p(r_2) \cdot \ldots \cdot p(r_k)$, где $p(r_i)$ — вероятность применения правила r_i.

Каждому выводу слова α соответствует корневое дерево, называемое *деревом вывода*. Узлам дерева соответствуют терминалы и нетерминалы. Дерево вывода $\omega(\alpha) = (r_1, \ldots, r_k)$ может быть построено следующим образом. В корень дерева помещается аксиома s грамматики. Далее для правил $r_1, r_2, \ldots, r_k$ последовательно к листу дерева, соответствующему левой части правила r_i приписываются буквы правой части r_i, образующие листья на следующем ярусе дерева. После выполнения этого действия для всех правил полученное дерево будет являться деревом вывода слова α в грамматике G, соответствующее выводу α.

Одному дереву вывода, построенному для слову α могут соответствовать различные вывода этого слова. Известно, однако, что существует взаимно-однозначное соответствие между деревьями вывода слов и левыми выводами. Вероятность слова α в грамматике определяется как сумма вероятностей всех его левых выводов.

Грамматика G называется *согласованной*, если

$$\sum_{\alpha \in L_G} p(\alpha) = 1. \tag{3}$$

Согласованная грамматика G задаёт распределение вероятностей P на L_G, и определяет *стохастический КС-язык* $\mathfrak{L} = (L, P)$. В дальнейшем всюду предполагается, что грамматика согласованна.

Обозначим D_l^t — множество деревьев вывода высоты t, порождаемых грамматикой G при замене её аксиомы на A_l. Аналогично, $D_l^{\leqslant t}$ — множество деревьев вывода, высота которых не превосходит $t-1$.

Для исследования вероятностных характеристик стохастической КС-грамматики применяются производящие функции

$$F_i(s_1, s_2, \ldots, s_k) = \sum_{\substack{j=1 \\ r_{ij} \in R}}^{n_i} p_{ij} s_1^{l_1} s_2^{l_2} \ldots s_k^{l_k}, \tag{4}$$

где $l_m = l_m(i,j)$ — число вхождений нетерминала A_m в β_{ij}.

Величины

$$a_j^i = \left.\frac{\partial F_i(s_1, s_2, \ldots, s_k)}{\partial s_j}\right|_{s_1 = s_2 = \ldots = s_k = 1} \tag{5}$$

называются *первыми моментами* грамматики G. Матрица $A = (a_j^i)$ называется *матрицей первых моментов* грамматики G.

Матрица A, по построению, неотрицательна. По теореме Фробениуса, доказанной в [1], существует максимальный по модулю вещественный неотрицательный собственный корень r. Известно, что критерием согласованности стохастической КС-грамматики при отсутствии бесполезных нетерминалов является условие $r \leqslant 1$.

Говорят, что нетерминал A_j *непосредственно следует* за нетерминалом A_i (обозначается $A_i \to A_j$), если в R имеется правило $A_i \overset{(}{\to} p_{ij})\alpha_1 A_j \alpha_2$, где $\alpha_1, \alpha_2 \in (V_N \cup V_T)^*$. Транзитивное замыкание отношения $\to$ обозначается $\to_*$. Если $A_i \to_* A_j$, говорят, что A_j *выводится* из A_i.

Введём также отношение $\leftrightarrow_*$. Будем считать, что $A_i \leftrightarrow_* A_j$, если одновременно $A_i \to_* A_j$ и $A_j \to_* A_i$. Очевидно, отношение $\leftrightarrow_*$ есть отношение эквивалентности, и потому разбивает множество нетерминалов на классы $V_N = K_1 \cup K_2 \cup \ldots \cup K_m$: $K_i \cap K_j = \varnothing (i \neq j)$. Класс, содержащий ровно один нетерминал, будем называть *особым*. Множество классов $\{K_1, K_2, \ldots, K_m\}$ обозначим $\mathcal{K}$.

Если все нетерминалы грамматики образуют один класс, она называется *неразложимой*. В противном случае она называется *разложимой*. Очевидно, разложимой грамматике соответствует разложимая [1] матрица первых моментов.

Говорят, что класс K_j *непосредственно следует* за классом K_i (обозначается $K_i \prec K_j$), если существуют $A_1 \in K_i$ и $A_2 \in K_j$ такие, что $A_1 \to A_2$. Рефлексивное транзитивное замыкание $\prec$ обозначим $\prec_*$, и назовём отношением *следования*.

Будем говорить, что грамматика имеет вид *«цепочки»*, если она разложима, и граф, построенный на множестве $\mathfrak{K}$ по отношению $\prec$, имеет вид P_m. Пронумеруем классы грамматики таким образом, что $K_i \prec K_{i+1}, i = 1, 2, \ldots, m-1$. Пронумеруем нетерминалы так, что для любых $A_i \in K_p$ и $A_j \in K_q$ условия $i < j$ и $p < q$ равносильны. После этого матрица первых моментов грамматики приобретает вид:

$$A = \begin{pmatrix} A_{11} & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & A_{22} & A_{23} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{m-1,m-1} & A_{m-1,m} \\ 0 & 0 & 0 & \cdots & 0 & A_{m,m} \end{pmatrix} \qquad (6)$$

Блоки $A_{i,i}(i = 1, 2, \ldots, m)$ являются неразложимыми неотрицательными матрицами. Не уменьшая общности, будем считать их положительными и непериодичными [1]. Этого можно добиться с помощью метода укрупнения правил грамматики. Пусть r_i — перронов корень матрицы $A_{i,i}$. По построению матрицы A, $r = \max_i\{r_i\}$ и $r > 0$.

2 Свойства матрицы первых моментов

Обозначим $J = \{i : r_i = r\}$ — множество индексов i, таких что перронов корень матрицы A_i равен r. Обозначим $J = \{i_1, i_2, \ldots, i_q\}$ причём $i_1 < i_2 < \ldots < i_q$ Разобьём множество классов $\mathcal{K}$ на группы классов $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_w$. При этом $\mathcal{M}_1 = \{K_1, K_2, \ldots, K_{i_1}\}$, и $\mathcal{M}_l = \{K_{i_{l-1}+1}, \ldots, K_{i_l}\}$, где $l > 1$. Нетрудно видеть, что в каждой группе $\mathcal{M}_j$ содержится ровно один класс с номером из J.

Тогда матрицу первых моментов можно представить в виде

$$A = \begin{pmatrix} B_{11} & B_{12} & \cdots & 0 & 0 \\ 0 & B_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{w-1,w-1} & B_{w-1,w} \\ 0 & 0 & \cdots & 0 & B_{w,w} \end{pmatrix}, \qquad (7)$$

где B_{ij} — блок, находящийся на пересечении строк, соответствующих нетерминалам классов группы $\mathcal{M}_i$, и столбцов, соответствующим нетерминалам классов группы $\mathcal{M}_j$. Очевидно, каждой из матриц $B_{i,i}$ соответствует перронов корень равный r.

Рассмотрим матрицу

$$A^t = \begin{pmatrix} B_{11}^t & B_{12}^{(t)} & \cdots & B_{1,w-1}^{(t)} & B_{1,w}^{(t)} \\ 0 & B_{22}^t & \cdots & B_{2,w-1}^{(t)} & B_{2,w}^{(t)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{w-1,w-1}^t & B_{w-1,w}^{(t)} \\ 0 & 0 & \cdots & 0 & B_{w,w}^t \end{pmatrix}. \tag{8}$$

Её вид был установлен в [3], где приведено доказательство для случая $r < 1$, однако при $r = 1$ доказательство остаётся справедливым.

Сформулируем этот результат в виде теоремы.

Теорема 1 *В принятых обозначениях при* $t \to \infty$

$$B_{lh}^{(t)} = H_{lh} \cdot t^{s_{lh}-1} r^t (1 + o(1)) \quad \textit{при} \quad l \neq h, \tag{9}$$

где H_{lh} *не зависит от* t, *и* s_{lh} *— число классов с номерами из* J *среди* $K_l, K_{l+1}, \ldots, K_h$.

3 Вероятности продолжения

Вероятностью продолжения $Q_i(t)$ будем называть вероятность того, что дерево вывода с корнем A_i будет иметь высоту не менее t. Для исследования этих вероятностей, по аналогии с теорией ветвящихся процессов [2] определим *производящие функции*:

$$\begin{aligned} F(\mathbf{s}) = & F(0, \mathbf{s}) = \sum p_{ij} s_1^{l_i^1} s_2^{l_i^2} \ldots s_{n_i}^{l_i^{n_i}} \\ & F(t, \mathbf{s}) = F(F(t-1, \mathbf{s})) \end{aligned} \tag{10}$$

Раскладывая $F_i(\mathbf{s})$ в ряд Тейлора в окрестности точки $\mathbf{1} = (1, 1, \ldots, 1)$, получаем:

$$\begin{aligned} F_i(\mathbf{s}) = F_i(\mathbf{1}) + (\nabla F_i(\mathbf{1}), \mathbf{s} - \mathbf{1}) + \frac{1}{2}(\mathbf{s} - \mathbf{1})^{\mathrm{T}} \nabla^2 F_i(\mathbf{1})(\mathbf{s} - \mathbf{1}) = \\ = 1 + \sum_j a_j^i (s_j - 1) + \frac{1}{2} \sum_{j,l} b_{jl}^i (s_j - 1)(s_l - 1) + O(|s_{k^*(\mathbf{s})} - 1|^3), \end{aligned} \tag{11}$$

где $k^*(\mathbf{s}) = \arg\max_k |s_k - 1|$

Подставляя в это выражение вместо $\mathbf{s}$ вектор $F(t, \mathbf{s})$, получаем:

$$\begin{aligned} F_i(F(t, \mathbf{s})) = F_i(t + 1, \mathbf{s}) = 1 + \sum_i a_j^i (F_j(t, \mathbf{s}) - 1) + \\ + \frac{1}{2} \sum_{j,l} b_{j,l}^i (F_j(t, \mathbf{s})) - 1)(F_l(t, \mathbf{s}) - 1) + O(|F_{k^*(F(t,\mathbf{s}))}(t, \mathbf{s}) - 1|^3) \end{aligned} \tag{12}$$

Нетрудно показать [2], что $Q_i(t) = 1 - F_i(t, \mathbf{0})$. Отсюда получаем рекуррентное соотношение для вероятностей продолжения:

$$Q_i(t+1) = \sum_j a^i_j Q_j(t) - \frac{1}{2}\sum_{j,l} b^i_{j,l} Q_j(t) Q_l(t) + O(|Q^3_{k^*}(t)|), \tag{13}$$

где $Q_i(t)$ — вероятность продолжения для деревьев высоты t, и $k^* = k^*(Q(t)) = \arg\min_k |Q_k(t)|$.

Рассматривая лишь нетерминалы класса с номером m, получим:

$$\begin{aligned}\underset{i\in I_m}{Q_i}(t+1) = \sum_{j\in I_m} a^i_j Q_j(t) - \frac{1}{2}\sum_{j,l\in I_m} b^i_{jl} Q_j(t) Q_l(t) + \\ + \sum_{j\in I_{m+1}} a^i_j Q_j(t) - \frac{1}{2}\sum_{j,l\in I_{m+1}} b^i_{jl} Q_j(t) Q_l(t) - \\ - \frac{1}{2} b^i_{jl} Q_j(t) Q_l(t) + \underset{(k\in I_m\cup I_{m+1})}{O} (Q^3_k(t))\end{aligned} \tag{14}$$

Перемножая левый собственный вектор матрицы $A_{m,m}$ ($v^{(m)}$) и вектор $Q^{(m)}(t)$, составленный из вероятностей продолжения $Q_i(t) : i \in I_m$, и полагая $Q^{(m)}_*(t) = (v^{(m)}, Q^{(m)}(t))$, получаем, после преобразований:

$$Q^{(m)}_*(t+1) = r_m Q^{(m)}_*(t) + b_m Q^{(m+1)}_*(t)(1+o(1)) - \frac{1}{2} B_m (Q^{(m)}_*(t))^2 (1+o(1)) \tag{15}$$

Для цепочки, состоящей полностью из классов с перроновым корнем 1, результат известен. Рассмотрим теперь подцепочку классов $K_i, K_{i+1}, \ldots, K_M$, где K_M — последний из классов грамматики. Будем предполагать, что среди $A_i, \ldots, A_M$ есть матрица с перроновым корнем равным 1. Пусть $A_j (j \in i, i+1, \ldots, M)$ — первая из таких матриц, то есть $r_i, r_{i+1}, \ldots, r_{j-1}$ меньше 1.

Пусть асимптотика вероятностей продолжения для нетерминалов класса K_j известна: $Q^{(j)} = U_j V_j \cdot t^s(1 + o(1))$.

Известно, что асимптотика вероятностей продолжения определяется видом степени матрицы первых моментов, то есть, $Q(t) = A^{t-k} Q(k)(1 + o(1))$ для некоторого фиксированного k. Тогда для всех $l \in \{i, i+1, \ldots, j-1\}$

$$Q^{(l)}(t+k) = A^t_{l,l} Q^{(l)}(k) + A^{(t)}_{l,l+1} Q^{(l+1)}(k) + \ldots + A^{(t)}_{l,M} Q^{(M)}(k) \tag{16}$$

Поскольку в подцепочке $K_i, \ldots, K_{j-1}$ все перроновы корни меньше 1, выполняется неравенство

$$A^t_{l,l} Q^{(l)}(k) + A^{(t)}_{l,l+1} Q^{(l+1)}(k) + \ldots + A^{(t)}_{l,j-1} Q^{(j-1)}(k) \le O(r^t) \quad \text{при} \quad r < 1 \tag{17}$$

Это следует сразу из вида матрицы A^t. Кроме того $A^{(t)}_{l,h} = H_{l,h}t^{s_{lh}-1}(1+o(1))$ и $A^{(t)}_{j,h} = H_{j,h}t^{s_{jh}-1}(1+o(1))$, где $j \le h \le M$ и s_{ab} — число классов с максимальным перроновым корнем в подцепочке $K_a, \ldots, K_b$. Поскольку классы с максимальным перроновым корнем содержатся только в подцепочке $K_j, \ldots, K_M$, матрицы $A^{(t)}_{l,h}$ и $A^{(t)}_{j,h}$ имеют одинаковую асимптотику. Отсюда, $Q^{(l)} = O(t^s)$.

В ходе доказательства в [3] показано, что $\forall h \in \{j, j+1, \ldots, M\} \quad A^{(t)}_{l,h} = UV_{lh} \cdot t^{s_{lh}-1}(1+o(1))$, где $U = (r_l E - A_{l,l})^{-1} A_{l,l+1} \ldots A_{j-2,j-1}(r_{j-1}E - A_{j-1,j-1})^{-1} A_{j-1,j} u^{(j)}$, и соответственно, компонетны вектора $Q^{(l)}$ пропорциональны компонетнам U.

Объединяя результаты докритического случая, случая цепочки, полностью состоящей из критических классов, и полученные здесь, можно записать общий вид вероятностей продолжения для грамматики в виде цепочки.

Теорема 2 *Пусть грамматика имеет вид цепочки. Пусть также $Q_i(t)$ — вероятности продолжения для деревьев высоты не менее t, и $P_i(t) = Q_i(t) - Q_i(t-1)$. Тогда*

$$\begin{aligned} Q_n(t) &= c_\mu U^{(\mu)}_{n-k_{\mu-1}} t^{-\left(\frac{1}{2}\right)^{m-\mu}} \\ P_n(t) &= d_\mu U^{(\mu)}_{n-k_{\mu-1}} t^{-1-\left(\frac{1}{2}\right)^{m-\mu}}, \end{aligned} \tag{18}$$

где $n \in I_\mu$, c_μ и d_μ заданы.

4 Математические ожидания числа применений правила в деревьях вывода

Обозначим через $q^l_{ij}(t,\tau)$ и $\bar{q}^l_{ij}(t,\tau)$ случайные величины, равные числу применений правила r_{ij} в дереве вывода, соответственно, из D^t_l и $D^{\leqslant t}_l$. Пусть также

$$\begin{aligned} S^l_{ij}(t) &= \sum_{\tau=1}^{t-1} q^l_{ij}(t,\tau) \\ \bar{S}^l_{ij}(t) &= \sum_{\tau=1}^{t-1} \bar{q}^l_{ij}(t,\tau) \end{aligned}, \tag{19}$$

и $S^l_{ij}(t)$, $\bar{S}^l_{ij}$ — соответственно число применений правила r_{ij} в дереве из D^t_l, $D^{\leqslant t}_l$. Для удобства записи положим

$$\begin{aligned} S_{ij}(t) = S^1_{ij}(t), \quad \bar{S}_{ij}(t) = \bar{S}^1_{ij}(t) \\ q_{ij}(t,\tau) = q^1_{ij}(t,\tau), \quad \bar{q}_{ij}(t,\tau) = \bar{q}^1_{ij}(t,\tau) \end{aligned}. \tag{20}$$

Рассмотрим математические ожидания некоторых из введённых величин. Обозначим

$$M_{ij}^l(t) = M[S_{ij}^l(t)], \quad \bar{M}_{ij}^l(t) = M[\bar{S}_{ij}^l(t)]. \tag{21}$$

Задача данного раздела заключается в вычислении $\bar{M}_{ij}^l(t)$, $M_{ij}^l(t)$ для грамматик в виде «цепочки». Для их нахождения будет удобно использовать три леммы, доказанные в [4].

Лемма 1 *Пусть s, d — натуральные, $m = (m_1, \ldots, m_s)$ — вектор целых неотрицательных чисел, $y = (y_1, \ldots, y_s)$ — вектор, и $\bar{m} = \sum_{j=1}^{s} m_j$. Тогда*

$$(1-y_1)^{n_1} \ldots (1-y_s)^{n_s} = \sum_{\substack{\bar{m}<d \\ m \geqslant 0}} \binom{n_1}{m_1}\binom{n_2}{m_2} \ldots \binom{n_s}{m_s} (-1)^{\bar{m}} y^m + R_d(n_1, \ldots, n_s, y), \tag{22}$$

где $y^m = y_1^{m_1} \ldots y_s^{m_s}$, и остаточный член представим в виде

$$R_d(n_1, \ldots, n_s, y) = \sum_{\substack{\bar{m}=d \\ m \geqslant 0}} (-1)^d \varepsilon_m(n_1, \ldots, n_s, y) y^m, \tag{23}$$

причём

$$0 \leqslant \varepsilon_m(n_1, \ldots, n_s, y') \leqslant \varepsilon_m(n_1, \ldots, n_s, y) \leqslant \binom{n_1}{m_1} \ldots \binom{n_s}{m_s} \tag{24}$$

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

Лемма 2 *Пусть $A(t)$ — последовательность матриц размером $k \times k$, и $A(t) \to A$ при $t \to \infty$, причём $A > 0$, и её перронов корень $r = 1$. Пусть $b(t) = bt^\alpha(1 + o(1))$ — последовательность векторов длины k, где $b \geqslant 0$, $b \neq 0$, и α — действительное число. Тогда для последовательности векторов $x(t)$ при $t = 1, 2, \ldots$, определяемой рекуррентным соотношением $x(t) = b(t) + A(t)x(t-1)$ при $t \to \infty$ справедливо соотношение*

$$\frac{x_i(t)}{vx(t)} \to u_i, \tag{25}$$

при условии что $x(t_0) > 0$ для некоторого номера t_0, где $u, v > 0$ — соответственно правый и левый собственные векторы матрицы A при нормировке $vu = 1$.

Лемма 3 *Пусть последовательность x_t, $x_t > 0$ при любом $t \geqslant 0$, удовлетворяет рекуррентному соотношению*

$$x_{t+1} = at^\alpha(1 + \varepsilon_1(t)) + (1 - bt^\beta(1 + \varepsilon_2(t)))x_t, \tag{26}$$

где $\beta < 0$, $b > 0$, и $\varepsilon_1(t), \varepsilon_2(t) = o(1)$ при $t \to \infty$. Тогда верны следующие асимптотические равенства:

$$\begin{aligned} &(1)\ x_t = \frac{at^{\alpha+1}}{\alpha+1}(1+o(1)) \quad \text{при} \quad \beta < -1,\ \alpha \geqslant 0 \\ &(2)\ x_t = \frac{at^{\alpha+1}}{\alpha+b+1}(1+o(1)) \quad \text{при} \quad \beta = -1,\ \alpha > -1 \\ &(3)\ x_t = \frac{at^{\alpha-\beta}}{b}(1+o(1)) \quad \text{при} \quad -1 < \beta < 0 \end{aligned} \tag{27}$$

Вначале рассмотрим $\bar{M}^q_{ij}(t)$. Пусть $p(\cdot)$ — условная вероятность дерева d в грамматике G, при условии что $d \in D^{\leqslant t}_q$. Рассмотрим множество $D^{\leqslant t}_{ql}$ деревьев из $D^{\leqslant t}_q$, первый ярус которых получен применением правила r_{ql} к корню дерева. Пусть

$$\bar{P}^{ij}_{ql}(t) = \sum_{d \in D^{\leqslant t}_{ql}} p(d) q_{ij}(d), \tag{28}$$

где $q_{ij}(d)$ — число применений правила r_{ij} в дереве d, и $\bar{P}^{ij}_{ql}(t)$ — вклад деревьев из $D^{\leqslant t}_{ql}$ в матожидание $\bar{M}^q_{ij}(t)$. Для краткости, обозначим $\bar{P}_{ql} = \bar{P}^{ij}_{ql}$. Тогда, очевидно,

$$\bar{M}^q_{ij}(t) = \sum_{l=1}^{n_q} \bar{P}_{ql}(t). \tag{29}$$

Рассмотрим величину $\bar{P}_{ql}(t)$. Пусть

$$q_{ij}(d) = q^{(1)}_{ij}(d) + q^{(2)}_{ij}(d), \tag{30}$$

где $q^{(1)}_{ij}(d)$ — число применений правила r_{ql} в дереве d на первом его ярусе, а $q^{(2)}_{ij}(d)$ — на остальных ярусах. Тогда

$$\bar{P}_{ql}(t) = \sum_{d \in D^{\leqslant t}_{ql}} p(d) q_{ij}(d) = \sum_{d \in D^{\leqslant t}_{ql}} p(d) q^{(1)}_{ij}(d) + \sum_{d \in D^{\leqslant t}_{ql}} p(d) q^{(2)}_{ij}(d) = \bar{P}^{(1)}_{ql}(t) + \bar{P}^{(2)}_{ql}(t) \tag{31}$$

Очевидно, $q^{(1)}_{ij}(d) = \delta^q_i \delta^l_j$ (где $delta$ — символ Кронекера), и следовательно, учитывая что $p(\cdot)$ — условные вероятности, получаем

$$\bar{P}^{(1)}_{ql}(t) = \delta^q_i \delta^l_j \frac{p_{ij} Q_{s_{ij}}(t-1)}{1 - Q_q(t)}, \tag{32}$$

где $Q_X(t)$ — вероятность наборов деревьев вывода высоты не превосходящей $t-1$, набор корней которых задан характеристическим вектором $X \in \mathbb{N}^k$.

Обозначим также $\delta^i(n) = (\delta^i_k)|_{i=\overline{1,n}} \in \{0,1\}^n$.

Условную вероятность дерева $p(d)$ при $d \in D^{\leqslant t}_{ql}$ можно выразить как

$$p(d) = \frac{p_{ql}}{1 - Q_q(t)} p_1(d) p_2(d) \dots p_{\bar{s}_{ql}}(d), \tag{33}$$

где $p_j(d)$ — вероятность поддерева d с корнем в j-м узле первого яруса. Тогда

$$\bar{P}^{(2)}_{ql}(t) = \frac{p_{ql}}{1 - Q_q(t)} \sum_{d \in D^{\leqslant t}_{ql}} \prod_{n=1}^{\bar{s}_{ql}} p_n(d) \sum_{m=1}^{\bar{s}_{ql}} q'^m_{ij}(d), \tag{34}$$

где $q'^m_{ij}(d)$ — число применений правила r_{ij} в поддереве дерева d с корнем в m-том нетерминале первого яруса.

Выделим в d поддеревья $d_1, d_2, \dots, d_{\bar{s}_{ql}}$, где d_j — поддерево с корнем в j-м узле первого яруса d. Преобразуя (34), получаем

$$\bar{P}^{(2)}_{ql}(t) = \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\bar{s}_{ql}} \sum_{d \in D^{\leqslant t}_{ql}} \left(\prod_{n=1}^{\bar{s}_{ql}} p_n(d) \right) q'^m_{ij}(d) =$$

$$= \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\bar{s}_{ql}} \sum_{d_1, \dots, d_{m-1}, d_{m+1}, \dots, d_{\bar{s}_{ql}}} p_1(d_1) \dots p_{m-1}(d_{m-1}) p_{m+1}(d_{m+1}) \dots p_{\bar{s}_{ql}}(d_{\bar{s}_{ql}}) q'^m_{ij}(d) =$$

$$\frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\bar{s}_{ql}} Q_{s_{ql} - \delta^m} q_{ij}(d_m) = \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{k} s^m_{ql} \bar{M}^m_{ij}(t-1) Q_{s_{ql} - \delta^m}(t-1) \tag{35}$$

Зная $\bar{P}_{ql}(t) = \bar{P}^{(1)}_{ql}(t) + \bar{P}^{(2)}_{ql}(t)$, получаем

$$\bar{M}^q_{ij}(t) = \frac{1}{1 - Q_q(t)} \left[\delta^q_i p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} s^m_{ql} \bar{M}^m_{ij}(t-1) Q_{s_{ql} - \delta^m}(t-1) \right] \tag{36}$$

Обозначая

$$\bar{M}'^q_{ij}(t) = \bar{M}^q_{ij}(t)(1 - Q_q(t)), \tag{37}$$

имеем

$$\bar{M}'^q_{ij}(t) = \delta^q_i p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} s^m_{ql} \bar{M}'^m_{ij}(t-1) Q_{s_{ql} - \delta^m}(t-1) \tag{38}$$

Рекуррентное соотношение (38) является опорной точкой для вычисления $\bar{M}^q_{ij}(t)$. Получим аналогичное уравнение для $M^q_{ij}(t)$.

$M_{ij}^q(t) = \sum_{l=1}^{n_q} P_{ql}(t)$, где $P_{ql}(t)$ — вклад деревьев из D_{ql}^t в $M_{ij}^q(t)$. Аналогично тому, как это сделано в (31), полагаем $P_{ql}(t) = P_{ql}^{(1)}(t) + P_{ql}^{(2)}(t)$. При этом

$$P_{ql}^{(1)}(t) = \delta_i^q \delta_j^l \frac{p_{ij} R_{s_{ij}}(t-1)}{P_q(t)}, \tag{39}$$

где $R_X(t)$ — вероятность наборов деревьев из $D^{\leqslant t}$, набор корней которых задан характеристическим вектором X, и высота хотя бы одного из которых достигает $t-1$. $P_{ql}^{(2)}(t)$ можно представить в виде

$$P_{ql}^{(2)}(t) = \sum_{m=1}^{\bar{s}_{ql}} P_{ql}^{(2)m}(t), \tag{40}$$

где $P_{ql}^{(2)m}(t)$ — вклад деревьев с m-м корнем на первом ярусе в $M_{ij}^q(t)$.

Обозначим через S_1 вклад в $P_{ql}^{(2)m}(t)$ наборов деревьев, в которых ярус t достигается деревом с корнем в m-м нетерминале первого яруса. Очевидно,

$$S_1 = \frac{(1 - Q_{z_m}(t-1)) Q_{s_{ql}-\delta^{z_m}}(t-1) M_{ij}^{z_m}(t-1)}{P_q(t)}, \tag{41}$$

где z_m — m-й нетерминал первого яруса.

Пусть S_2 — вклад наборов, где ярус t достигается через другие деревья. Тогда

$$S_2 = \frac{(1 - Q_{z_m}(t-1)) R_{s_{ql}-\delta^m}(t-1) \bar{M}_{ij}^{z_m}(t-1)}{P_q(t)}. \tag{42}$$

В результате, для M_{ij}^q получаем

$$\begin{aligned} M_{ij}^q &= \sum_{l=1}^{n_q} \left(P_{ql}^{(1)}(t) + \sum_{m=1}^{\bar{s}_{ql}} P_{ql}^{(2)m}(t) \right) = \\ &= \frac{1}{P_q(t)} [\delta_i^q p_{ij} R_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} (P_m(t-1) Q_{s_{ql}-\delta^m}(t-1) M_{ij}^m(t-1) + \\ &\qquad + (1 - Q_m(t-1)) R_{s_{ql}-\delta^m}(t-1) \bar{M}_{ij}^m(t-1))] \end{aligned} \tag{43}$$

Таким образом, получено рекуррентное соотношение для $M_{ij}^q(t)$, аналогичное (38).

Из леммы 1 следуют равенства для $Q_X(t$ и $R_X(t)$:

$$\begin{aligned} Q_X(t) &= \prod_{i=1}^{k} (1 - Q_i(t))^{x_i} = 1 - \sum_{i=1}^{k} x_i Q_i(t) + \Theta \left(\sum_{i,j=1}^{k} x_i x_j Q_i(t) Q_j(t) \right) \\ R_X(t) &= Q_X(t) - Q_X(t-1) = \sum_{i=1}^{k} x_i P_i(t) + \Theta \left(\sum_{i,j=1}^{k} x_i x_j Q_i(t) Q_j(t) \right) \end{aligned} \tag{44}$$

Теперь можно приступить к вычислению $\bar{M}'^{q}_{ij}(t)$ и $M'^{q}_{ij}(t)$.

4.1 Случай критического класса

Рассмотрим вначале случай, когда $I(q) \in J$.

Пусть $q, i \in I_\mu$. Тогда при $m \in I_\nu : \nu > \mu$ $\bar{M}'^{m}_{ij}(t) = 0$, и для $\bar{M}'^{q}_{in}(t$ получаем:

$$\bar{M}'^{q}_{ij}(t) = \delta^i_q p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m \in I_\mu} s^m_{ql} \bar{M}'^{m}_{ij}(t-1) Q_{s_{ql}-\delta^m}(t-1) \tag{45}$$

Подставляя выражения (44) где это необходимо, и учитывая, что $I(q) \in J$, имеем

$$\bar{M}'^{q}_{ij}(t) = \delta^i_q + \sum_{m \in I_\mu} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^{m}_{ij}(t-1) - \sum_{m \in I_\mu} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^{m}_{ij}(t-1) \cdot \sum_{n \in I_\mu} (s^n_{ql} - \delta^m_n) Q_n(t-1)(1+o(1)). \tag{46}$$

Непосредственным взятием производных от производящих функций проверяются выражения для первых и вторых моментов:

$$\begin{aligned} a^q_m &= \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \\ b^q_{mn} &= \sum_{l=1}^{n_q} p_{ql} s^m_{ql} (s^n_{ql} - \delta^m_n) \end{aligned} \tag{47}$$

Подставляя их в (46), получаем

$$\begin{aligned} \bar{M}'^{q}_{ij}(t) = \delta^i_q p_{ij} + \sum_{m \in I_\mu} a^q_m \bar{M}'^{m}_{ij}(t-1) - \\ - c_\mu t^{\xi(\mu)-1} \sum_{\substack{m \in I_\mu \\ n \in I}} b^q_{mn} u^{I(n)}_{n-k_{I(n)-1}} \bar{M}'^{m}_{ij}(t-1)(1+o(1)) \end{aligned} \tag{48}$$

где $\xi(\mu)$ — число классов с перроновым корнем, равным 1 в цепочке $K_\mu, K_{\mu+1}, \ldots, K_m$, и $K_{I(n)} \ni A_n$.

Применяя лемму 2, получаем:

$$\bar{M}'^{q}_{ij}(t) = u^{(\mu)}_{q-k_{\mu-1}} \cdot \sum_{l \in I_\mu} v^{(\mu)}_{n-k_{\mu-1}} \bar{M}'^{l}_{ij}(t) = u^{(\mu)}_{q-k_{\mu-1}} M^{(\mu)}_*(t), \tag{49}$$

где $M^{(\mu)}_*(t) = \sum_{l \in I_\mu} v^{(\mu)}_{n-k_{\mu-1}} \bar{M}'^{l}_{ij}(t)$, и $v^{(\mu)}$ — левый собственный вектор матрицы $A_{\mu,\mu}$.

Домножая (48) на $v^{(\mu)}_{q-k_{\mu-1}}$ и суммируя по q, получаем:

$$\delta \bar{M}^{(\mu)}_*(t) = v^{(\mu)}_{i-k_{\mu-1}} p_{ij} - c_\mu t^\alpha \sum_{q,m,n \in I_\mu} v^{(\mu)}_{q-k_{\mu-1}} b^q_{mn} u^{(\mu)}_{m-k_{\mu-1}} u^{(\mu)}_{n-k_{\mu-1}} = v^{(\mu)}_{i-k_{\mu-1}} p_{ij} - c_\mu t^\alpha B_\mu, \quad (50)$$

где $\alpha = -\left(\frac{1}{2}\right)^{\xi(\mu)-1}$.

Нетрудно видеть, что величина $\bar{M}^{(\mu)}_*(t)$ удовлетворяет условиям леммы 3. Применяя её, получаем:

$$\begin{matrix} \bar{M}^{(\mu)}_*(t) = \dfrac{v^{(\mu)}_{i-k_{\mu-1}} p_{ij}}{c_\mu B_\mu + 1} t(1+o(1)), \quad \text{если } \alpha = -1 \\ \bar{M}^{(\mu)}_*(t) = \dfrac{v^{(\mu)}_{i-k_{\mu-1}} p_{ij}}{c_\mu B_\mu} t^{-\alpha}(1+o(1)), \quad \text{если } \alpha > -1 \end{matrix}, \quad (51)$$

где $\alpha = -\left(\frac{1}{2}\right)^{\xi(\mu)-1}$.

Пусть теперь $I(q) < I(i)$ (q и i в различных классах). Тогда $\bar{M}'^q_{ij}(t)$ выражается следующим образом:

$$\bar{M}'^q_{ij}(t) = \delta^i_q p_{ij} Q_{s_{ij}}(t-1) + \sum_{m \in I_\mu \cup I_{\mu+1}} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^m_{ij}(t-1) Q_{s_{ql}-\delta^m}(t-1) \quad (52)$$

Учитывая малость $Q_n(t)$ при $I(n) > I(q)$, получаем

$$\bar{M}'^q_{ij}(t) = O(p_{ij}) + \sum_{m \in I_\mu} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^m_{ij}(t-1) \cdot \left(1 - \sum_{n \in I_\mu} (s^n_{ql} - \delta^m_n) Q_n(t-1)\right)(1+o(1)) + \\ + \sum_{m \in I_{\mu+1}} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^m_{ij}(t-1)(1+o(1)) \quad (53)$$

Положим, $\bar{M}'^m_{ij}(t-1) = \overline{\mathcal{M}}'_{\mu+1} t^{\gamma(\mu+1)}(1+o(1))$. Это выполняется для $\mu+1 = m$, что видно из полученных соотношений (51). Исходя из этого, получим выражение для $\bar{M}'^q_{ij}(t)$. Подставляя дополнительно выражения для первых и вторых моментов, а также для вероятностей продолжения $Q_n(t-1)$, получаем:

$$\bar{M}'^q_{ij}(t) = \sum_{m \in I_\mu} a^q_m \bar{M}'^m_{ij}(t-1) - c_\mu t^{\alpha(\mu)} \sum_{m,n \in I_\mu} b^q_{mn} u^{(\mu)}_{n-k_{\mu-1}} \bar{M}'^m_{ij}(t-1)(1+o(1)) + \\ + \overline{\mathcal{M}}'_{\mu+1} t^{\gamma(\mu+1)} \sum_{m \in I_{\mu+1}} a^q_m u^{(\mu+1)}_{m-k_\mu}(1+o(1)) \quad (54)$$

Домножая на $v^{(\mu)}_{q-k_{\mu-1}}$ и суммируя по q, получаем

$$\delta\bar{M}^{(\mu)}_*(t) = \overline{\mathcal{M}}'_{\mu+1}\left(\sum_{\substack{q\in I_\mu\\ m\in I_{\mu+1}}} v^{(\mu)}_{q-k_{\mu-1}}a^q_m u^{(\mu+1)}_{m-k_\mu}\right)\cdot t^{\gamma(\mu+1)}-$$
$$-c_\mu t^{\alpha(\mu)}\cdot\left(\sum_{q,m,n\in I_\mu} v^{(\mu)}_{q-k_{\mu-1}}b^q_{mn}u^{(\mu)}_{m-k_{\mu-1}}u^{(\mu)}_{n-k_{\mu-1}}\right)\cdot\bar{M}^{(\mu)}_*(t-1)(1+o(1)) =$$
$$= \overline{\mathcal{M}}'_{\mu+1}b_{\mu+1}t^{\gamma(\mu+1)}(1+o(1)) - c_\mu B_\mu t^{\alpha(\mu)}\bar{M}^{(\mu)}_*(t-1)(1+o(1)) \quad (55)$$

Случай $\alpha(\mu) = -1$ рассматривать не имеет смысла, так как это означает, что не существует критических классов с номерами, превышающими μ. Полагая $\alpha(\mu) > -1$ и применяя лемму 2, получаем

$$\bar{M}^{(\mu)}_*(t) = \frac{\overline{\mathcal{M}}'_{\mu+1}b_{\mu+1}t^{\gamma(\mu+1)-\alpha(\mu)}}{c_\mu B_\mu} \quad (56)$$

Из полученных формул (51) и (56) нетрудно получить общее выражение для величины $\bar{M}^{(\mu)}_*(t)$, при условии что грамматика имеет вид «цепочки» и состоит только из критических классов ($J = \{1, 2, \dots, m\}$).

$$\bar{M}^{(\mu)}_*(t) = \prod_{j=\mu}^{\nu-1}\left(\frac{b_{j+1}}{c_j B_j}\right)\cdot\left(\frac{v^{(\nu)}_{i-k_{\nu-1}}p_{ij}}{c_\nu B_\nu + \delta^m_\nu}\right)\cdot t^{\frac{1}{2}^{m-\nu}\left(2-\frac{1}{2}^{\nu-\mu}\right)}, \quad (57)$$

где $\mu = I(q)$, $\nu = I(i)$. Подставляя (49), и затем (37), непосредственно получаем

$$\bar{M}^q_{ij}(t) = \frac{u^{(\mu)}_{q-k_{\mu-1}}}{1-Q_q(t)}\prod_{j=\mu}^{\nu-1}\left(\frac{b_{j+1}}{c_j B_j}\right)\cdot\left(\frac{v^{(\nu)}_{i-k_{\nu-1}}p_{ij}}{c_\nu B_\nu + \delta^m_\nu}\right)\cdot t^{\frac{1}{2}^{m-\nu}\left(2-\frac{1}{2}^{\nu-\mu}\right)} \quad (58)$$

Перейдём к вычислению $M^q_{ij}(t)$. Пусть $I(q) = I(i)$. Полагая $M'^q_{ij}(t) = M^q_{ij}(t)P_q(t)$, из (43) получаем

$$M'^q_{ij}(t) = O(p_{ij})t^{\beta(\mu)} + \sum_{m\in I_\mu}\sum_{l=1}^{n_q} p_{ql}s^m_{ql}M'^m_{ij}(t-1)-$$
$$-\sum_{m,n\in I_\mu}\sum_{l=1}^{n_q} p_{ql}s^m_{ql}(s^n_{ql}-\delta^m_n)Q_n(t-1)M'^m_{ij}(t-1)+$$
$$+\sum_{m,n\in I_\mu}\sum_{l=1}^{n_q} p_{ql}s^m_{ql}(s^n_{ql}-\delta^m_n)P_n(t-1)\bar{M}'^m_{ij}(t-1) \quad (59)$$

Обозначим

$$Q^{(\mu)}(t) = c_\mu u^{(\mu)} t^{\alpha(\mu)}$$
$$P^{(\mu)}(t) = d_\mu u^{(\mu)} t^{\beta(\mu)} \tag{60}$$

Подставляя выражения (47) для первых и вторых моментов в (59), имеем

$$M'^q_{ij}(t) = \sum_{m\in I_\mu} a^q_m M'^m_{ij}(t-1) - c_\mu t^{\alpha(\mu)} \sum_{m,n\in I_\mu} b^q_{mn} u^{(\mu)}_{n-k_{\mu-1}} M'^m_{ij}(t-1)(1+o(1)) + \\ + d_\mu t^{\beta(\mu)} \sum_{m,n\in I_\mu} b^q_{mn} u^{(\mu)}_{n-k_{\mu-1}} \bar{M}^m_{ij}(t-1)(1+o(1)) \tag{61}$$

Обозначая дополнительно $M^q_{ij}(t) = \mathcal{M}_q t^{\left(\frac{1}{2}\right)^{m-\mu}}$, а также учитывая $\beta(\mu) = -1-\left(\frac{1}{2}\right)^{m-\mu}$ и выражения для первых моментов, получаем

$$M'^q_{ij}(t) = \sum_{m\in I_\mu} a^q_m M'^m_{ij}(t-1) - c_\mu t^{\alpha(\mu)} \sum_{m,n\in I_\mu} b^q_{mn} u^{(\mu)}_{n-k_{\mu-1}} M'^m_{ij}(t-1)(1+o(1)) + \\ + d_\mu \mathcal{M}_\mu t^{-1} \sum_{m,n\in I_\mu} b^q_{mn} u^{(\mu)}_{m-k_{\mu-1}} u^{(\mu)}_{n-k_{\mu-1}} (1+o(1)) \tag{62}$$

Применяя лемму 2, получаем

$$M'^q_{ij}(t) = u^{(\mu)}_{q-k_{\mu-1}} M^{(\mu)}_*(t)(1+o(1))$$
$$M^{(\mu)}_*(t) = \sum_{m\in I_\mu} v^{(\mu)}_{n-k_{\mu-1}} M'^n_{ij}(t) \tag{63}$$

Домножая (62) на $v^{(\mu)}_{q-k_{\mu-1}}$ и суммируя по q, имеем

$$\delta M^{(\mu)}_*(t) = d_\mu \mathcal{M}_\mu B_\mu t^{-1} - c_\mu t^{\alpha(\mu)} B_\mu M^{(\mu)}_*(t-1)(1+o(1)) \tag{64}$$

Применяя лемму 3, получаем в результате

$$M^{(\mu)}_*(t) = \begin{cases} d_\mu \overline{\mathcal{M}}'_\mu B_\mu (1+o(1)), & \text{при } \mu = m, \\ \frac{d_\mu \overline{\mathcal{M}}'_\mu}{c_\mu} t^{-1-\alpha(\mu)} (1+o(1)), & \text{при } \mu < m \end{cases} \tag{65}$$

Пусть теперь $I(q) = \mu$, $I(i) = \nu$, $\mu < \nu$, тогда из (43) получаем

$$M'^q_{ij}(t) = \delta^q_i p_{ij} R_{s_{ij}}(t-1) + \sum_{m\in I_\mu\cup I_{\mu+1}} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \cdot [Q_{s_{ql}-\delta^m}(t-1) M'^m_{ij}(t-1) + \\ + (1-Q_m(t-1)) R_{s_{ql}-\delta^m}(t-1) \bar{M}^m_{ij}(t-1)], \tag{66}$$

откуда

$$M'^q_{ij}(t) = O\left(t^{\beta(\mu)}\right) + \sum_m a^q_m M'^m_{ij}(t-1) - \sum_{\substack{m\in I_\mu\cup I_{\mu+1}\\ n\in I_\mu}} b^q_{mn} Q_n(t-1) M'^m_{ij}(t-1)(1+o(1)) + \\ + \left(\sum_{\substack{m\in I_\mu\cup I_{\mu+1}\\ n\in I_\mu}} b^q_{mn} P_n(t-1)\bar{M}^m_{ij}(t-1)\right)(1+o(1)) \quad (67)$$

Можем записать

$$M'^q_{ij}(t) = \sum_{m\in I_\mu} a^q_m M'^m_{ij}(t-1) - \sum_{m,n\in I_\mu} b^q_{mn} Q_n(t-1) M'^m_{ij}(t-1)(1+o(1)) + \\ + \sum_{m,n\in I_\mu} b^q_{mn} P_n(t-1)\bar{M}^m_{ij}(t-1)(1+o(1)) \quad (68)$$

Домножая на $v^{(\mu)}_{q-k_{\mu-1}}$ и суммируя по q, имеем

$$\delta M^{(\mu)}_*(t) = d_\mu \overline{\mathcal{M}}_\mu B_\mu t^{\beta(\mu)+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)}(1+o(1)) - c_\mu B_\mu t^{\alpha(\mu)} \cdot M^{(\mu)}_*(t-1)(1+o(1)) \quad (69)$$

Так как $\mu < m$, $\alpha(\mu) > -1$, поэтому по лемме 3 получаем

$$M^{(\mu)}_*(t) = \frac{d_\mu \overline{\mathcal{M}}_\mu}{c_\mu} t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)} \quad (70)$$

Объединяя результаты (65) и (70), получаем

$$M^{(\mu)}_*(t) = \frac{d_\mu \overline{\mathcal{M}}_\mu B_\mu}{\delta^m_\mu (c_\mu B_\mu - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)}(1+o(1)), \quad (71)$$

после чего из (62)

$$M^q_{ij}(t) = \frac{u^{(\mu)}_{q-k_{\mu-1}}}{P_q(t)} \frac{d_\mu \overline{\mathcal{M}}_\mu B_\mu}{\delta^m_\mu (c_\mu B_\mu - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)}(1+o(1)) \quad (72)$$

Сформулируем полученный результат в виде теоремы.

Теорема 3 *Пусть стохастическая КС-грамматика имеет вид цепочки, и для неё известны вторые моменты b^q_{mn}. Пусть также для перронов корень подматрицы каждого класса равен* 1.

$$M^q_{ij}(t) = \frac{u^{(\mu)}_{q-k_{\mu-1}}}{P_q(t)} \frac{d_\mu \overline{\mathcal{M}}_\mu B_\mu}{\delta^m_\mu (c_\mu B_\mu - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)}(1+o(1)), \quad (73)$$

где
$$B_\mu = \sum_{m,n \in I_\mu} b^q_{mn} u^{(\mu)}_{m-k_{\mu-1}} u^{(\mu)}_{n-k_{\mu-1}}, \tag{74}$$
$P_q(t)$ — вероятности деревьев высоты t в данной грамматике, и c_μ, d_μ — некоторые константы.

5 Энтропия

Пусть L^t — множество слов языка L_G, которым соответствуют деревья вывода из D^t. Будем рассматривать грамматики с однозначным выводом, то есть, положим что каждому слову из L^t соответствует единственное дерево вывода из D^t.

По определению, энтропия языка L^t есть
$$H(L^t) = -\sum_{\alpha \in L^t} p_t(\alpha) \log p_t(\alpha), \tag{75}$$
где $p_t(\alpha) = p(\alpha | \alpha \in L^t) = \frac{p(\alpha)}{P(L^t)}$. Используя это выражение для $p_t(\alpha)$, получаем:
$$\begin{aligned} H(L^t) = -\frac{1}{P(L^t)} \sum_{\alpha \in L^t} p(\alpha) \left(\log p(\alpha) - \log P(L^t)\right) = \\ = \frac{\log P(L^t)}{P(L^t)} \sum_{\alpha \in L^t} p(\alpha) - \frac{1}{P(L^t)} \sum_{\alpha \in L^t} p(\alpha) \log p(\alpha) = \\ = \log P(L^t) - \frac{1}{P(L^t)} \sum_{\alpha \in L^t} p(\alpha) \log p(\alpha). \end{aligned} \tag{76}$$

Выразим вероятность слова α через вероятности правил вывода r_{ij}. Будем рассматривать грамматику с однозначным выводом и считать, что каждому слову α из L^t соответствует единственное дерево $d(\alpha)$ из D^t, и, следовательно, единственный левый вывод $\omega(\alpha) = r_1 \cdot r_2 \cdot \ldots \cdot r_\mu$. Получаем:
$$p(\alpha) = p(r_1) \cdot p(r_2) \cdot \ldots \cdot p(r_\mu) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} p_{ij}^{q_{ij}(\alpha)}, \tag{77}$$
где $q_{ij}(\alpha)$ — число применений правила r_{ij} при выводе слова α (в грамматике с однозначным выводом это число определяется единственным образом). Тогда:
$$\begin{aligned} \sum_{\alpha \in L^t} p(\alpha) \log p(\alpha) = \sum_{\alpha \in L^t} p(\alpha) \sum_{i=1}^{k} \sum_{j=1}^{n_i} q_{ij}(\alpha) \log p_{ij} = \\ = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \log p_{ij} \sum_{\alpha \in L^t} q_{ij}(\alpha) p(\alpha) \end{aligned} \tag{78}$$

Пользуясь определением $M(S_{ij}(t))$, получаем:

$$\sum_{\alpha\in L^t} p(\alpha)\log p(\alpha) = \sum_{i=1}^{k}\sum_{j=1}^{n_i} \log p_{ij} M(S_{ij}(t))P(L^t) \tag{79}$$

Подставляя это выражение в (76), получаем:

$$H(L^t) = \log P(L^t) - \sum_{i=1}^{k}\sum_{j=1}^{n_i} M(S_{ij}(t))\log p_{ij} \tag{80}$$

По определению, $P(L^t) = P_1(t) = O(t^{-1-\frac{1}{2}^{q_j-1}})$, и $\log P(L^t) = O(\log t)$. Подставляя выражение для $M(S_{ij}(t)) = M_{ij}(t)$ в (80), получаем:

$$\begin{aligned} H(L^t) = O(\log t) - \sum_{i=1}^{k}\sum_{j=1}^{n_i} p_{ij} log p_{ij} d_i t^{\frac{1}{2}^{q_l^*-q}} = \\ = \sum_{i=1}^{k}\sum_{j=1}^{n_i} H(R_i) d_i t^{\frac{1}{2}^{q_l^*-q}}(1+o(1)), \end{aligned} \tag{81}$$

где $H(R_i) = -\sum_{j=1}^{n_i} p_{ij}\log p_{ij}$ — энтропия множества R_i правил вывода.

Обозначим $l' = maxl : l \in J$ — номер последнего критического класса. Элементы суммы при $i \in I_{l'}$ имеют вид $O(t^2)$, остальные имеют вид $o(t^2)$. Поэтому:

$$H(L^t) = \sum_{i\in I_{l'}}\sum_{j=1}^{n_i} d_i H(R_i) t^2 (1+o(1)) \tag{82}$$

Сформулируем теорему:

Теорема 4 *Энтропия языка L^t, состоящего из слов длины t, порождаемых разложимой стохастической контекстно-свободной грамматикой, имеющей вид "цепочки выражается формулой*

$$H(L^t) \sim \sum_{i\in I_{l'}}\sum_{j=1}^{n_i} d_i H(R_i)\cdot t^2, \tag{83}$$

где $d_i > 0$, $H(R_i) = \sum_{j=1}^{n_i} p_{ij}\log p_{ij}$ — энтропия множества R_i правил вывода с нетерминалом A_i в левой части, и l' — номер критического класса, наиболее удалённого от начала цепочки.

Список литературы

[1] Гантмахер Ф.Р. **Теория матриц.** — 5-е изд., — М.: ФИЗМАТЛИТ, 2010 — 560 с.

[2] Севастьянов Б.А. **Ветвящиеся процессы** — М.: Наука, 1971 — 436 с.

[3] Жильцова Л.П. О матрице первых моментов разложимой стохастической КС-грамматики // Учёные записки Казанского государственного университета, 2009

[4] Борисов А.Е. Закономерности в словах стохастических контекстно-свободных языков, порождённых грамматиками с двумя классами нетерминальных символов. Вопросы экономного кодирования.

[5] Фу К. Структурные методы в распознавании образов.

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