Министерство образования и науки Российской Федерации
Государственное образовательное учреждение высшего профессионального образования «Нижегородский государственный университет им. Н.И. Лобачевского»

Факультет вычислительной математики и кибернетики

Кафедра математической логики и высшей алгебры

Направление: Прикладная математика и информатика

ВЫПУСКНАЯ КВАЛИФИКАЦИОННАЯ РАБОТА БАКАЛАВРА

Тема:
«Исследование одного класса стохастических контекстно-свободных грамматик»

Заведующий кафедрой:

Проф., д. ф.-м. н. ____________________
Шевченко Валерий Николаевич

Выполнил:

студент группы 84-03 ________________
Мартынов Игорь Михайлович

Научный руководитель:

Проф., д. ф.-м. н. ____________________
Жильцова Лариса Павловна

Нижний Новгород
2011

Аннотация

В работе рассматривается один класс стохастических контекстно-свободных грамматик и исследуются свойства порождаемых ими языков. Исследование проводится с использованием аппарата дискретного анализа и теории ветвящихся процессов. Полученные результаты могут быть применены построения асимптотически оптимального алгоритма кодирования слов рассматриваемых языков, а также для исследования более общих классов грамматик.

Содержание

1 Введение

При передаче и хранении информации часто возникает необходимость кодирования данных таким образом, чтобы обеспечить наибольшую степень сжатия. Сжатие данных может быть достигнуто использованием статистических данных, таких как частоты появления букв в сообщениях. Если, кроме этого, учитывать структурные свойства языка сообщений, можно дополнительно увеличить эффективность сжатия.

К. Шеннон в статье "Математическая теория связи" [1] рассматривал задачу экономного кодирования, моделируя источник сообщений автоматом с конечным числом состояний.

А.А. Марков поставил задачу экономного кодирования на множестве слов, порождаемых конечным автоматом и доказал [2], что учитывая таким образом структуру источника сообщений, можно увеличить эффективность сжатия и уменьшить вычислительную сложность алгоритма кодирования.

Ближайшим обобщением регулярных языков (языков, порождаемых конечными автоматами) являются контекстно-свободные языки. При рассмотрении таких языков удобно моделировать источник сообщений с помощью стохастической контекстно-свободной грамматики, и большую роль приобретает исследование вероятностных свойств таких грамматик.

Л.П. Жильцова изучила задачу экономного кодирования на множестве слов контекстно-свободного языка, и построила алгоритм асимптотически оптимального кодирования с полиномиальной временной сложностью для некоторых классов грамматик [6][7]. Кроме того, Жильцова показала, что перронов корень [3] матрицы первых моментов [4] грамматики существенно влияет на её вероятностные свойства и эффективность кодирования.

Изучение стохастических контекстно-свободных грамматик было продолжено А.Е. Борисовым. Он изучил грамматику с разложимой матрицей первых моментов (разложимую грамматику), с двумя классами нетерминалов [8]. В частности, Борисов рассмотрел случай, когда перронов корень матрицы первых моментов грамматики равен единице. По аналогии с теорией ветвящихся процессов такой случай называется критическим.

В данной работе рассматриваются критический случай для разложимых грамматик, классы нетерминалов в которых расположены в виде «цепочки». В результате проведённого исследования изучены вероятностные свойства матрицы первых моментов таких грамматик, получена асимптотика вероятностей продолжения и вероятностей деревьев вывода фиксированной высоты, а также асимптотика математических ожиданий числа применений некоторого правила в дереве вывода фиксированной высоты. Полученные результаты будут использованы для построения асимптотически оптимального алгоритма кодирования для рассматриваемых классов грамматик.

2 Основные определения

Стохастической КС-грамматикой называется система $G =< V_T, V_N, R, s >$, где V_T и V_N — алфавиты терминальных и нетерминальных символов (*терминалов* и *нетерминалов*) соответственно, 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

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

и 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 $\omega(\alpha) = (r_1, r_2, \ldots, r_k)$, последовательное применение которых к s даёт слово α, называется *выводом* этого слова. Если на каждом шаге правило применяется к самому левому нетерминалу в слове, вывод назвается *левым.*

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

Каждому выводу слова соответствует *дерево вывода*, получаемое следующим образом. В корень дерева помещается аксиома s грамматики. Далее на каждом ярусе дерева ко всем нетерминалам этого яруса последовательно, слева направо, применяются правила левого вывода соответствующего слова. Так происходит до тех пор, пока на очередном ярусе не останутся только терминальные символы. При этом будут применены последовательно все правила левого вывода соответствующего слова.

По выводу слова может быть построено *дерево вывода*. В корень дерева помещается аксиома s, далее на каждом ярусе дерева ко всем нетерминалам этого яруса применяется правило, соответствующее выводу. Символы этого слова записываются слева направо в дереве, присоединяясь к исходному нетерминалу как к родителю.

Пример дерева вывода

Рассмотрим грамматику $G = \langle \{E, S, M\}, \{a, b, (,), +, -, *, /\}, E, R \rangle$, со следующим набором правил R:

$$\begin{aligned}
&r_{11} : E \to E + S \\
&r_{12} : E \to E - S \\
&r_{13} : E \to S \\
&r_{21} : S \to S * M \\
&r_{22} : S \to S/M \\
&r_{23} : S \to M \\
&r_{31} : M \to (E) \\
&r_{32} : M \to a \\
&r_{33} : M \to b
\end{aligned} \tag{4}$$

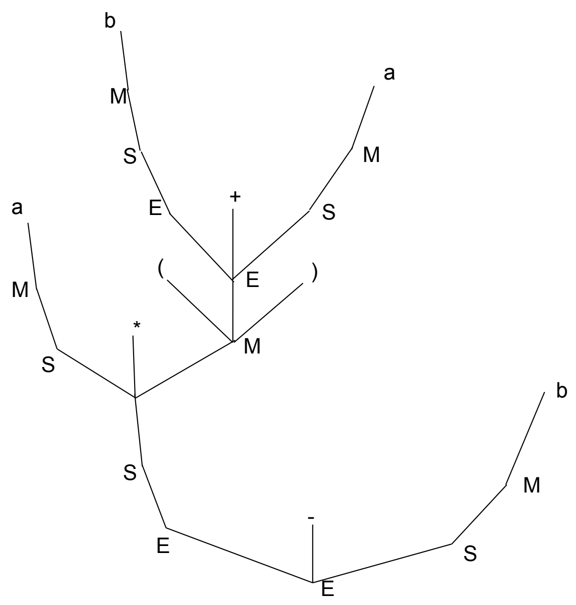

Рис. 1: Пример дерева вывода

Нетрудно видеть, что такая КС-грамматика задаёт язык арифметических выражений, содержащих арифметические операции, круглые скобки и аргументы a и b.

Рассмотрим слово $a * (b + a) - b$, порождаемое грамматикой. Ему соответствует дерево вывода на рис. 1.

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 $t - 1$.

Одними из наиболее важных, с точки зрения кодирования, величин, являются *вероятности продолжения.* В соответствии с введённым в [4] определением вероятность продолжения $Q_i(t)$ есть вероятность того, что в заданной грамматике, при выборе в качестве аксиомы нетерминала A_i полученное дерево вывода будет иметь высоту более t. Используется также величина $P_i(t) = Q_i(t-1) - Q_i(t)$, равная вероятности деревьев вывода высоты t.

Для исследования вероятностных характеристик стохастической КС-грамматики применяются производящие функции

$$F_i(s_1, s_2, \ldots, s_k) = \sum_{\substack{j=1 \\ r_{ij} \in R}}^{n_i} p_{ij} s_1^{l_1} s_2^{l_2} \ldots s_k^{l_k}, \tag{5}$$

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

Величины

$$a_j^i = \left. \frac{\partial F_i(s_1, s_2, \ldots, s_k)}{\partial s_j} \right|_{s_1 = s_2 = \ldots = s_k = 1} \tag{6}$$

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 [3], 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 $A_i \xrightarrow{(} p_{ij})\alpha_1 A_j \alpha_2$, 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 [3] 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 классы её нетерминалов связаны соотношением

$$K_1 \prec K_2 \prec \ldots \prec K_m, \tag{7}$$

где m — общее число классов нетерминалов.

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:

$$A = \begin{pmatrix} A_{11} & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & A_{22} & A_{23} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{m-1,m-1} & A_{m-1,m} \\ 0 & 0 & 0 & \cdots & 0 & A_{m,m} \end{pmatrix} \tag{8}$$

Блоки $A_{i,i} (i = 1, 2, \ldots, m)$ являются неразложимыми неотрицательными матрицами. Не уменьшая общности, будем считать их положительными и непериодичными [3]. 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 $r > 0$.

3 Свойства матрицы первых моментов

Обозначим $J = \{i : r_i = r\}$ — множество индексов i, таких что перронов корень матрицы A_i равен r. В случае, если $r = 1$, классы с номерами из множества J будем называть *критическими*. 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

$$A = \begin{pmatrix} B_{11} & B_{12} & \cdots & 0 & 0 \\ 0 & B_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{w-1,w-1} & B_{w-1,w} \\ 0 & 0 & \cdots & 0 & B_{w,w} \end{pmatrix}, \tag{9}$$

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

$$A^t = \begin{pmatrix} B_{11}^t & B_{12}^{(t)} & \cdots & B_{1,w-1}^{(t)} & B_{1,w}^{(t)} \\ 0 & B_{22}^t & \cdots & B_{2,w-1}^{(t)} & B_{2,w}^{(t)} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & B_{w-1,w-1}^t & B_{w-1,w}^{(t)} \\ 0 & 0 & \cdots & 0 & B_{w,w}^t \end{pmatrix}. \tag{10}$$

Её вид был установлен в [5], 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 $t \to \infty$*

$$B_{lh}^{(t)} = H_{lh} \cdot t^{s_{lh}-1} r^t (1 + o(1)) \quad \textit{при} \quad l \neq h, \tag{11}$$

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

4 Вероятности продолжения

Вероятность продолжения $Q_i(t)$ определяется как вероятность того, что дерево вывода, имеющее корень A_i и построенное по правилам рассматриваемой грамматики, будет иметь высоту более t.

Для получения вероятностей продолжения [4] рассматриваемой стохастической КС-грамматики, определим *производящие функции* следующим образом:

$$\begin{aligned} F(\mathbf{s}) &= \sum p_{ij} s_1^{l_i^1} s_2^{l_i^2} \ldots s_{n_i}^{l_i^{n_i}} \\ F(t, \mathbf{s}) &= F(F(t-1, \mathbf{s})) \end{aligned} \tag{12}$$

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

$$\begin{aligned} F_i(\mathbf{s}) = F_i(\mathbf{1}) + (\nabla F_i(\mathbf{1}), \mathbf{s} - \mathbf{1}) + \frac{1}{2}(\mathbf{s} - \mathbf{1})^{\mathrm{T}} \nabla^2 F_i(\mathbf{1})(\mathbf{s} - \mathbf{1}) = \\ = 1 + \sum_j a_j^i (s_j - 1) + \frac{1}{2} \sum_{j,l} b_{jl}^i (s_j - 1)(s_l - 1) + \underset{(k \in 1,2,\ldots,n)}{O} (|s_k - 1|^3) \end{aligned} \tag{13}$$

Учитывая

$$F(t+1, \mathbf{s}) = F(F(t, \mathbf{s})), \tag{14}$$

можно записать рекуррентное соотношение для вероятностей продолжения:

$$Q_i(t+1) = \sum_j a_j^i Q_j(t) - \frac{1}{2} \sum_{j,l} b_{j,l}^i Q_j(t) Q_l(t) + O(|Q_k^3(t)|), \tag{15}$$

где $Q_i(t)$ — вероятность продолжения для деревьев высоты t.

Для каждого из классов разложимости K_m будем рассматривать соответствующий его нетерминалам вектор вероятностей продолжения $Q^{(m)}(t)$. Тогда

$$Q(t) = \begin{pmatrix} Q^{(1)}(t) \\ Q^{(2)}(t) \\ \dots \\ Q^{(M)}(t) \end{pmatrix} : \quad Q^{(j)}(t) \in \mathbb{R}^{k_j}$$

В обозначениях $Q^{(m)}(t)$ уравнение (15) перепишется в виде:

$$Q_i^{(m)}(t+1) = \sum_{j \in I_m} a_j^i Q_{j-\sigma_{m-1}}^{(m)}(t) + \sum_{j \in I_{m+1}} a_j^i Q_{j-\sigma_m}^{(m+1)}(t) - \\ - \frac{1}{2} \sum_{j,l \in I_m} b_{jl}^i Q_j^{(m)}(t) Q_l^{(m)}(t) + O\left(Q_{\max}^{(m)} Q_{\max}^{(m+1)} + \left(Q_{\max}^{(m)} \right)^3 \right), \quad (16)$$

где $Q_{\max}^{(j)}(t) = \max\limits_{i=\overline{1,k_j}} \left\{ Q_i^{(j)}(t) \right\}$.

Уравнение (15) может быть переписано в матричной форме:

$$Q(t+1) = (A - A(t))Q(t) = \prod_{k=n}^{t} (A - A(k))Q(n), \quad (17)$$

где $A(t) = \left\{ \frac{1}{2} \sum\limits_{l=1}^{N} b_{jl}^i Q_l(t) \right\}$, $(i, j = \overline{1, N})$. Поскольку рассматривается согласованная грамматика, $Q(t) \to \mathbf{0}$, и следовательно, $A(t) \to 0$ покомпонентно.

Покажем пропорциональность элементов вектора $Q^{(m)}(t)$ элементам соответствующего собственного вектора $u^{(m)}$. Для этого проведём рассуждения аналогичные тем, что были использованы в [8] для случая грамматики с двумя классами разложимости.

Введём обозначения:

$$\begin{aligned} &A^*(t) = (A - A(t)) \cdot \ldots \cdot (A - A(1)) \\ &B_{ij}^*(t) = \frac{A_{ij}^*(t)}{t^{j-i}} \\ &B_{ij}(t) = \frac{A_{ij}^{(t)}}{t^{j-i}} \end{aligned} \quad (18)$$

Из исследования матрицы первых моментов известно, что $B_{ij}(t) \to b_{ij} u^{(i)} v^{(j)}$. Выберем произвольные $0 < \varepsilon_1, \varepsilon_2 < 1$. Зададим $l(\varepsilon_1)$ и $n(\varepsilon_2)$, для которых выполняются условия:

$$\begin{aligned} &\left| B_{ij}(l) - b_{ij} u^{(i)} v^{(j)} \right| < \varepsilon_1 E \\ &\forall\, t \ge n \quad A(t) < \varepsilon_2 A \end{aligned} \quad (19)$$

Это всегда можно сделать, так как $B_{ij}(t) \to b_{ij} u^{(i)} v^{(j)}$ и $A(t) \to 0$.

Рассмотрим произвольный вектор $x > \mathbf{0}$. Тогда $A^*(t)x$, при условии $t > n$, можно оценить так:

$$(1 - \varepsilon_2)^l A^l x^{(n)} \le (A - A(t)) \cdot \ldots \cdot (A - A(n+1)) A^*(n) x \le A^l x^{(n)}, \quad (20)$$

где $x^{(n)} = A^*(t)x$. Используя эту оценку, можем записать:

$$\left|B_{ij}^*(t)x - b_{ij}u^{(i)}v^{(j)}x^{(n)}\right| \le \left|B_{ij}^*(t)x - B_{ij}(l)x^{(n)}\right| + \\ + \left|B_{ij}(l)x^{(n)} - b_{ij}u^{(i)}v^{(j)}\right| \le \left(1-(1-\varepsilon_2)^l\right) B_{ij}(l)x^{(n)} + \varepsilon_1 u^{(i)}v^{(j)}x^{(n)}$$

В силу $B_{ij}(t) \to b_{ij}u^{(i)}v^{(j)}$, существует $m_0 = \max\limits_{l,i,j} B_{ij}(l)$. Используя это, получаем:

$$\left|B_{ij}^*(t)x - b_{ij}u^{(i)}v^{(j)}x^{(n)}\right| \le k_j\left(\left(1-(1-\varepsilon_2)^l\right)m_0 + \varepsilon_1\right)\max x^{(n)}$$

В этом выражении устремляем ε_1 к нулю, затем ε_2 к нулю таким образом, что $\varepsilon_2 = o\left(\frac{1}{l(\varepsilon_1)}\right)$. В результате, можем записать:

$$\left|B_{ij}^*(t)x - b_{ij}u^{(i)}v^{(j)}x^{(n)}\right| \le \varepsilon \max x^{(n)} \tag{21}$$

для произвольного $\varepsilon > 0$. Домножая это неравенство слева на $v^{(i)}$, получаем:

$$\left|v^{(i)}B_{ij}^*(t)x - b_{ij}v^{(j)}x^{(n)}\right| \le k\varepsilon \max v^{(i)} \max x^{(n)} \le \varepsilon^* v^{(j)}x^{(n)}, \tag{22}$$

где $\varepsilon^* = \frac{k\varepsilon \max v^{(i)}}{\min v^{(i)}}$. Принимая во внимание (21) и (22), можем записать:

$$\left|\frac{B_{ij}^*(t)x}{v^{(i)}B_{ij}^*(t)x} - \frac{b_{ij}u^{(i)}v^{(j)}x^{(n)}}{b_{ij}v^{(j)}x^{(n)}}\right| = \left|\frac{B_{ij}^*(t)x}{v^{(i)}B_{ij}^*(t)x} - u^{(i)}\right| \to 0$$

Последнее равнозначно

$$\left|\frac{A_{ij}^*(t)x}{v^{(i)}A_{ij}^*(t)x} - u^{(i)}\right| \to 0,$$

или же

$$A_{ij}^*(t)x = u^{(i)}v^{(i)}A_{ij}^*(t)x \cdot (1+o(1)) \tag{23}$$

Применим полученное соотношение для вектора $Q(n)$. Так как

$$A^*(t)Q(n) = \begin{pmatrix} A_{11}^*(t)Q^{(1)}(n) + \ldots + A_{M,M}^*(t)Q^{(M)}(n) \\ A_{22}^*(t)Q^{(2)}(n) + \ldots + A_{M,M}^*(t)Q^{(M)}(n) \\ \vdots \\ A_{M,M}Q^{(M)}(n) \end{pmatrix}, \tag{24}$$

рассмотрим отдельно некоторую строчку вектора $A^*(t)Q(t)$, принимая во внимание (23). Получим:

$$Q^{(j)}(t) = A^*(t)Q(n)|_{I_j} = A_{jj}^*(t)Q^{(1)}(n) + \ldots + A_{M,M}^*(t)Q^{(M)}(n) = \\ = u^{(j)}v^{(j)} \cdot \left(A_{jj}^*(t)Q^{(1)}(n) + \ldots + A_{M,M}^*(t)Q^{(M)}(n)\right) = \\ = u^{(j)} \cdot \left(v^{(j)}, Q^{(j)}(t)\right) \tag{25}$$

Таким образом, компоненты каждого из векторов $Q^{(j)}(t)$ пропорциональны компонентам соответствующего собственного вектора $u^{(j)}$.

Далее оценим асимптотику элементов вектора $Q^{(m)}(t)$ при $t \to \infty$.

Положим теперь $v^{(m)}Q^{(m)}(t) = Q_*^{(m)}(t) \in R^1$, и домножим уравнение скалярно на $v^{(m)}$. Тогда

$$Q_*^{(m)}(t+1) = Q_*^{(m)}(t) + v^{(m)}A_{m,m+1}u^{(m+1)}Q_*^{(m+1)}(t)(1+o(1)) - \\ - \frac{1}{2}\sum_{i,j,l=1}^{k_m} v_i^{(m)} b_{jl}^i u_j^{(m)} u_l^{(m)} (Q_*^{(m)}(t))^2(1+o(1)) \quad (26)$$

Обозначим $\delta Q_*^{(m)}(t) = Q_*^{(m)}(t+1) - Q_*^{(m)}(t)$, а также

$$b_m = v^{(m)}A_{m,m+1}u^{(m+1)}$$

$$B_m = \sum_{i,j,l=1}^{k_m} v_i^{(m)} b_{jl}^i u_j^{(m)} u_l^{(m)}$$

Тогда уравнение (26) перепишется как

$$\delta Q_*^{(m)}(t) = b_m Q_*^{(m+1)}(t)(1+o(1)) - \frac{1}{2}B_m (Q_*^{(m)}(t))^2(1+o(1)) \quad (27)$$

Выражение для $\delta Q_*^{(m)}(t)$ также можно получить из (16), вычитая это уравнение из него же с заменой $t \to t+1$:

$$\delta Q_i^{(m)}(t+1) = \sum_{j\in I_m} a_j^i \delta Q_{j-\sigma_{m-1}}^{(m)}(t) + \sum_{j\in I_{m+1}} a_j^i \delta Q_{j-\sigma_m}^{(m+1)}(t)(1+o(1)) - \\ - \frac{1}{2}\sum_{j,l\in I_m} b_{jl}^i \left(Q_{j-\sigma_{m-1}}^{(m)}(t+1) Q_{l-\sigma_{m-1}}^{(m)}(t+1) - \\ - Q_{j-\sigma_{m-1}}^{(m)}(t) Q_{l-\sigma_{m-1}}^{(m)}(t) \right)(1+o(1)) \quad (28)$$

Домножая скалярно на $v^{(m)}$, получаем

$$\delta Q_*^{(m)}(t+1) = \delta Q_*^{(m)}(t) + b_m \delta Q_*^{(m+1)}(t)(1+o(1)) - \\ - \frac{1}{2}B_m \delta Q_*^{(m)}(t) \left(Q_*^{(m)}(t+1) + Q_*^{(m)}(t) \right)(1+o(1)) \quad (29)$$

Из исследования неразложимого случая в [4] известна асимптотика вероятностей продолжения для последнего класса:

$$Q_*^{(M)}(t) = k_M t^{-1}(1+o(1)) \quad (30)$$

Здесь, для удобства, все константы перед $\frac{1}{t}$ обозначены через k_M.

Далее проведём рассуждение по индукции. Пусть для некоторого класса с номером $m+1$

$$Q_*^{(m+1)}(t) = k_{m+1}t^{-\alpha}(1+o(1))\colon 0 < \alpha \le 1 \quad (31)$$

Положим

$$z(t) = t^\alpha \delta Q_*^{(m)}(t) \quad (32)$$

Производя замену в уравнении (29), и имея в виду $Q_*(t+1) = O\left(Q_*(t)\right)$, получаем:

$$\frac{z(t+1)}{(t+1)^\alpha} - \frac{z(t)}{t^\alpha} = b_m \delta Q_*^{(m+1)}(t)(1+o(1)) - \frac{1}{2}B_m \frac{z(t)}{t^\alpha}\cdot 2Q_*^{(m)}(t)(1+o(1))$$

Преобразуем выражение в левой части уравнения.

$$\frac{z(t+1)}{(t+1)^\alpha} - \frac{z(t)}{t^\alpha} = \frac{t^\alpha z(t+1) - (t+1)^\alpha z(t)}{t^\alpha (t+1)^\alpha} =$$
$$= \frac{t^\alpha z(t+1) - t^\alpha \left(1 + \frac{\alpha}{t} + o\left(\frac{1}{t}\right)\right) z(t)}{t^\alpha (t+1)^\alpha} = \frac{\delta z(t)}{(t+1)^\alpha} - \frac{\alpha z(t)(1+o(1))}{t(t+1)^\alpha}$$

Тогда

$$\frac{\delta z(t)}{(t+1)^\alpha} - \frac{\alpha z(t)(1+o(1))}{t(t+1)^\alpha} = b_m \delta Q_*^{(m+1)}(t)(1+o(1)) - \frac{B_m}{t^\alpha} Q_*^{(m)}(t) z(t)(1+o(1))$$

По предположению индукции (31), $\delta Q_*^{(m+1)}(t) = -\frac{k_{m+1}\alpha}{t(t+1)^\alpha}(1+o(1))$, и тогда

$$\frac{\delta z(t)}{(t+1)^\alpha} - \frac{\alpha z(t)(1+o(1))}{t(t+1)^\alpha} = -\frac{b_m \alpha k_{m+1}}{t(t+1)^\alpha}(1+o(1)) - \frac{B_m}{t^\alpha} Q_*^{(m)}(t) z(t)(1+o(1))$$

Домножая на $(t+1)^\alpha$, получаем

$$\delta z(t) - \frac{\alpha z(t)}{t} = \frac{-b_m \alpha k_{m+1}}{t}(1+o(1)) - B_m Q_*^{(m)}(t) z(t)(1+o(1))$$

Заметим, что, в силу предположения индукции, $\frac{1}{t} \le Q_*^{(m+1)}(t) = o(Q_*^{(m)}(t))$, поэтому уравнение можно переписать в виде:

$$\delta z(t) = -\frac{b_m \alpha k_{m+1}}{t}(1+o(1)) - B_m Q_*^{(m)}(t)(1+o(1)) \tag{33}$$

В [8] была доказана следующая лемма:

Лемма 1 *Пусть последовательность $z(t)(t = 1, 2, \ldots)$ удовлетворяет рекуррентному соотношению $\delta z(t) = f(t) - g(t)z(t)$, где при $t \to \infty$ выполняются условия*

$$g(t) \to 0, \frac{f(t)}{g(t)} \to 0, \sum_{i=1}^{t} g(i) \to \infty. \tag{34}$$

Пусть $g(t) > 0$ при любом $t > t_0$ для некоторого t_0. Тогда $z(t) \to 0$ при $t \to \infty$.

Полагая в уравнении (33) $f(t) = \frac{-b_m \alpha k_{m+1}}{t}(1+o(1))$, $g(t) = B_m Q_*^{(m)}(t)(1+o(1))$, замечаем, что для $\delta z(t)$ выполняются все условия леммы (1), и соответственно, $z(t) \to 0$. Из определения $z(t)$ получаем:

$$\delta Q_*^{(m)}(t) = o\left(\frac{1}{t^\alpha}\right) \tag{35}$$

Подставляя эту оценку в (27), получаем:

$$o\left(\frac{1}{t^\alpha}\right) = \frac{b_m k_{m+1}}{t^\alpha}(1+o(1)) - \frac{B_m}{2}\left(Q_*^{(m)}(t)\right)^2 (1+o(1)) \tag{36}$$

Отсюда:

$$\frac{b_m k_{m+1}}{t^\alpha}(1+o(1)) = \frac{B_m}{2}\left(Q_*^{(m)}(t)\right)^2 (1+o(1)) \tag{37}$$

Тогда для $Q_*^{(m)}(t)$ получаем оценку:

$$Q_*^{(m)}(t) = \sqrt{\frac{2b_m}{B_m} k_{m+1} \cdot \frac{1}{t^\alpha}}(1 + o(1)) = \sqrt{\frac{2b_m}{B_m} k_{m+1}} \cdot t^{-\frac{\alpha}{2}}(1 + o(1))$$

При этом, полагая $k_m = \sqrt{\frac{2b_m}{B_m} k_{m+1}}$, мы остаёмся в рамках предположения индукции.

Применяя это соотношение при $m = M-1, M-2, \ldots$, получаем оценку для произвольного класса цепочки. Если $Q_*^{(M)}(t)$ представлена в виде

$$Q_*^{(M)}(t) = \frac{2}{B_M t}(1 + o(1))$$

то $Q_*^{(m)}(t)$ для произвольного класса выражается следующим образом:

$$Q_*^{(m)}(t) = \sqrt{\frac{2b_m}{B_m}\sqrt{\frac{2b_{m+1}}{B_{m+1}}\sqrt{\frac{2b_{m+2}}{B_{m+2}} \cdots \sqrt{\frac{2b_{M-1}}{B_{M-1}B_M}}}}} \cdot t^{-(\frac{1}{2})^{q-1}} =$$
$$= \prod_{k=m}^{M-1} \left(\frac{2b_k}{B_k}\right)^{(\frac{1}{2})^{k-m+1}} \cdot t^{-(\frac{1}{2})^{q-1}} \qquad (38)$$

где $q = M - m + 1$ - длина цепочки, соединяющей классы K_m и K_M.

Полученные результаты можно сформулировать в виде следующей теоремы:

Теорема 2 *Пусть разложимая грамматика G имеет вид «цепочки», и перронов корень, соответствующий каждому из классов, равен единице. Тогда вероятность продолжения i-го нетерминала m-го класса имеет вид:*

$$Q_i^{(m)}(t) = \frac{k_m u_i^{(m)}(1 + o(1))}{t^{(\frac{1}{2})^{q-1}}}, \qquad (39)$$

где

$$k_m = \prod_{k=m}^{M-1} \left(\frac{2b_k}{B_k}\right)^{(\frac{1}{2})^{k-m+1}},$$
$$q = M - m + 1$$

Полученный результат нетрудно обобщить на случай произвольной «цепочки», содержащей как критические, так и докритические классы.

Рассматривая нетерминалы класса с номером m, получим:

$$\underset{i \in I_m}{Q_i}(t+1) = \sum_{j \in I_m} a_j^i Q_j(t) - \frac{1}{2} \sum_{j,l \in I_m} b_{jl}^i Q_j(t) Q_l(t) +$$
$$+ \sum_{j \in I_{m+1}} a_j^i Q_j(t) - \frac{1}{2} \sum_{j,l \in I_{m+1}} b_{jl}^i Q_j(t) Q_l(t) -$$
$$- \frac{1}{2} b_{jl}^i Q_j(t) Q_l(t) + \underset{(k \in I_m \cup I_{m+1})}{O} (Q_k^3(t)) \qquad (40)$$

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:

$$Q_*^{(m)}(t+1) = r_m Q_*^{(m)}(t) + b_m Q_*^{(m+1)}(t)(1 + o(1)) - \frac{1}{2} B_m (Q_*^{(m)}(t))^2 (1 + o(1)) \qquad (41)$$

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

$$Q^{(l)}(t+k) = A^t_{l,l}Q^{(l)}(k) + A^{(t)}_{l,l+1}Q^{(l+1)}(k) + \ldots + A^{(t)}_{l,M}Q^{(M)}(k) \tag{42}$$

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

$$A^t_{l,l}Q^{(l)}(k) + A^{(t)}_{l,l+1}Q^{(l+1)}(k) + \ldots + A^{(t)}_{l,j-1}Q^{(j-1)}(k) \leq O(r^t) \quad \text{при} \quad r < 1 \tag{43}$$

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 $j \leq h \leq M$ и s_{ab} — число классов с максимальным перроновым корнем в подцепочке $K_a, \dots, K_b$. Поскольку классы с максимальным перроновым корнем содержатся только в подцепочке $K_j, \dots, K_M$, матрицы $A^{(t)}_{l,h}$ и $A^{(t)}_{j,h}$ имеют одинаковую асимптотику. Отсюда, $Q^{(l)} = O(t^s)$.

В ходе доказательства в [5] показано, что $\forall h \in \{j, j+1, \dots, M\} \quad A^{(t)}_{l,h} = UV_{lh} \cdot t^{s_{lh}-1}(1+o(1))$, где $U = (r_l E - A_{l,l})^{-1}A_{l,l+1}\dots A_{j-2,j-1}(r_{j-1}E - A_{j-1,j-1})^{-1}A_{j-1,j}u^{(j)}$, 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 величин $Q_n(t)$ и $P_n(t)$. Сформулируем этот общий результат в виде теоремы.

Теорема 3 *Пусть грамматика G имеет вид «цепочки» (7). Тогда*

$$\begin{aligned} Q_n(t) &= c_\mu U^{(\mu)}_{n-k_{\mu-1}} t^{-\left(\frac{1}{2}\right)^{q_\mu - 1}} \\ P_n(t) &= d_\mu U^{(\mu)}_{n-k_{\mu-1}} t^{-1-\left(\frac{1}{2}\right)^{q_\mu - 1}}, \end{aligned} \tag{44}$$

где $A_n \in K_\mu$, c_μ и d_μ — некоторые константы, q_μ — количество классов с перроновым корнем равным 1 в подцепочке $K_\mu, K_{\mu+1}, \dots, K_m$, и $U^{(\mu)}$ — вектор.
Замечание: *При $\mu \in J$ вектор $U^{(\mu)}$ является собственным для матрицы $A_{\mu,\mu}$.*

5 Математические ожидания числа применений правил 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

$$\begin{aligned} S^l_{ij}(t) &= \sum_{\tau=1}^{t-1} q^l_{ij}(t,\tau) \\ \bar{S}^l_{ij}(t) &= \sum_{\tau=1}^{t-1} \bar{q}^l_{ij}(t,\tau) \end{aligned} \tag{45}$$

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

$$\begin{aligned} S_{ij}(t) = S^1_{ij}(t), \quad \bar{S}_{ij}(t) = \bar{S}^1_{ij}(t) \\ q_{ij}(t,\tau) = q^1_{ij}(t,\tau), \quad \bar{q}_{ij}(t,\tau) = \bar{q}^1_{ij}(t,\tau) \end{aligned}. \tag{46}$$

Рассмотрим математические ожидания некоторых из введённых величин. Обозначим

$$M^l_{ij}(t) = M[S^l_{ij}(t)], \quad \bar{M}^l_{ij}(t) = M[\bar{S}^l_{ij}(t)]. \tag{47}$$

Задача данного раздела заключается в вычислении $\bar{M}^l_{ij}(t)$, $M^l_{ij}(t)$ для грамматик в виде «цепочки». Для их нахождения будет удобно использовать три леммы, доказанные в [8].

Лемма 2 [4] *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*

$$(1-y_1)^{n_1} \ldots (1-y_s)^{n_s} = \sum_{\substack{\bar{m}<d \\ m\geqslant 0}} \binom{n_1}{m_1}\binom{n_2}{m_2} \ldots \binom{n_s}{m_s} (-1)^{\bar{m}} y^m + R_d(n_1, \ldots, n_s, y), \tag{48}$$

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

$$R_d(n_1, \ldots, n_s, y) = \sum_{\substack{\bar{m}=d \\ m\geqslant 0}} (-1)^d \varepsilon_m(n_1, \ldots, n_s, y) y^m, \tag{49}$$

причём

$$0 \leqslant \varepsilon_m(n_1, \ldots, n_s, y') \leqslant \varepsilon_m(n_1, \ldots, n_s, y) \leqslant \binom{n_1}{m_1} \ldots \binom{n_s}{m_s} \tag{50}$$

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

Лемма 3 [8] *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*

$$\frac{x_i(t)}{vx(t)} \to u_i, \tag{51}$$

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 4 [8] *Пусть последовательность x_t, $x_t > 0$ при любом $t \geqslant 0$, удовлетворяет рекуррентному соотношению*

$$x_{t+1} = at^\alpha(1+\varepsilon_1(t)) + (1 - bt^\beta(1+\varepsilon_2(t)))x_t, \tag{52}$$

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

$$\begin{aligned} &(1)\ x_t = \frac{at^{\alpha+1}}{\alpha+1}(1+o(1)) \quad \text{при} \quad \beta < -1,\ \alpha \geqslant 0 \\ &(2)\ x_t = \frac{at^{\alpha+1}}{\alpha+b+1}(1+o(1)) \quad \text{при} \quad \beta = -1,\ \alpha > -1 \\ &(3)\ x_t = \frac{at^{\alpha-\beta}}{b}(1+o(1)) \quad \text{при} \quad -1 < \beta < 0 \end{aligned} \tag{53}$$

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

$$\bar{P}^{ij}_{ql}(t) = \sum_{d \in D^{\leqslant t}_{ql}} p(d) q_{ij}(d), \tag{54}$$

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,

$$\bar{M}^q_{ij}(t) = \sum_{l=1}^{n_q} \bar{P}_{ql}(t). \tag{55}$$

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

$$q_{ij}(d) = q^{(1)}_{ij}(d) + q^{(2)}_{ij}(d), \tag{56}$$

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

$$\bar{P}_{ql}(t) = \sum_{d \in D^{\leqslant t}_{ql}} p(d) q_{ij}(d) = \sum_{d \in D^{\leqslant t}_{ql}} p(d) q^{(1)}_{ij}(d) + \sum_{d \in D^{\leqslant t}_{ql}} p(d) q^{(2)}_{ij}(d) = \bar{P}^{(1)}_{ql}(t) + \bar{P}^{(2)}_{ql}(t) \tag{57}$$

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

$$\bar{P}^{(1)}_{ql}(t) = \delta^q_i \delta^l_j \frac{p_{ij} Q_{s_{ij}}(t-1)}{1 - Q_q(t)}, \tag{58}$$

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

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

$$p(d) = \frac{p_{ql}}{1 - Q_q(t)} p_1(d) p_2(d) \ldots p_{\bar{s}_{ql}}(d), \tag{59}$$

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

$$\bar{P}^{(2)}_{ql}(t) = \frac{p_{ql}}{1 - Q_q(t)} \sum_{d \in D^{\leqslant t}_{ql}} \prod_{n=1}^{\bar{s}_{ql}} p_n(d) \sum_{m=1}^{\bar{s}_{ql}} q'^m_{ij}(d), \tag{60}$$

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 (60), получаем

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

$$\cdot \sum_{m=1}^{\bar{s}_{ql}} \left(\sum_{d_1, \dots, d_{m-1}, d_{m+1}, \dots, d_{\bar{s}_{ql}}} p_1(d_1) \dots p_{m-1}(d_{m-1}) p_{m+1}(d_{m+1}) \dots p_{\bar{s}_{ql}}(d_{\bar{s}_{ql}}) q_{ij}'^{m}(d) \right) =$$

$$\frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\bar{s}_{ql}} Q_{s_{ql} - \delta^m} q_{ij}(d_m) = \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{k} s_{ql}^m \bar{M}_{ij}^m(t-1) Q_{s_{ql} - \delta^m}(t-1) \quad (61)$$

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

$$\bar{M}_{ij}^q(t) = \frac{1}{1 - Q_q(t)} \left[\delta_i^q p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} s_{ql}^m \bar{M}_{ij}^m(t-1) Q_{s_{ql} - \delta^m}(t-1) \right] \quad (62)$$

Обозначая

$$\bar{M}_{ij}'^q(t) = \bar{M}_{ij}^q(t)(1 - Q_q(t)), \quad (63)$$

имеем

$$\bar{M}_{ij}'^q(t) = \delta_i^q p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} s_{ql}^m \bar{M}_{ij}'^m(t-1) Q_{s_{ql} - \delta^m}(t-1) \quad (64)$$

Рекуррентное соотношение (64) является опорной точкой для вычисления $\bar{M}_{ij}^q(t)$. Получим аналогичное уравнение для $M_{ij}^q(t)$.

$M_{ij}^q(t) = \sum_{l=1}^{n_q} P_{ql}(t)$, где $P_{ql}(t)$ — вклад деревьев из D_{ql}^t в $M_{ij}^q(t)$. Аналогично тому, как это сделано в (57), полагаем $P_{ql}(t) = P_{ql}^{(1)}(t) + P_{ql}^{(2)}(t)$. При этом

$$P_{ql}^{(1)}(t) = \delta_i^q \delta_j^l \frac{p_{ij} R_{s_{ij}}(t-1)}{P_q(t)}, \quad (65)$$

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

$$P_{ql}^{(2)}(t) = \sum_{m=1}^{\bar{s}_{ql}} P_{ql}^{(2)m}(t), \quad (66)$$

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,

$$S_1 = \frac{(1 - Q_{z_m}(t-1)) Q_{s_{ql} - \delta^{z_m}}(t-1) M_{ij}^{z_m}(t-1)}{P_q(t)}, \quad (67)$$

где z_m — m-й нетерминал первого яруса.

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

$$S_2 = \frac{(1 - Q_{z_m}(t-1)) R_{s_{ql} - \delta^m}(t-1) \bar{M}_{ij}^{z_m}(t-1)}{P_q(t)}. \quad (68)$$

В результате, для M_{ij}^q получаем

$$M_{ij}^q = \sum_{l=1}^{n_q}\left(P_{ql}^{(1)}(t) + \sum_{m=1}^{\bar{s}_{ql}} P_{ql}^{(2)m}(t)\right) =$$
$$= \frac{1}{P_q(t)}[\delta_i^q p_{ij} R_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} (P_m(t-1) Q_{s_{ql}-\delta^m}(t-1) M_{ij}^m(t-1) +$$
$$+ (1 - Q_m(t-1)) R_{s_{ql}-\delta^m}(t-1) \bar{M}_{ij}^m(t-1))] \quad (69)$$

Таким образом, получено рекуррентное соотношение для $M_{ij}^q(t)$, аналогичное (64). Из леммы 2 следуют равенства для $Q_X(t$ и $R_X(t)$:

$$\begin{aligned} Q_X(t) &= \prod_{i=1}^{k}(1-Q_i(t))^{x_i} = 1 - \sum_{i=1}^{k} x_i Q_i(t) + \Theta\left(\sum_{i,j=1}^{k} x_i x_j Q_i(t) Q_j(t)\right) \\ R_X(t) &= Q_X(t) - Q_X(t-1) = \sum_{i=1}^{k} x_i P_i(t) + \Theta\left(\sum_{i,j=1}^{k} x_i x_j Q_i(t) Q_j(t)\right) \end{aligned} \quad (70)$$

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

5.1 Случай критического класса

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

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

$$\bar{M}_{ij}'^q(t) = \delta_q^i p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m \in I_\mu} s_{ql}^m \bar{M}_{ij}'^m(t-1) Q_{s_{ql}-\delta^m}(t-1) \quad (71)$$

Подставляя выражения (70) где это необходимо, и учитывая, что $I(q) \in J$, имеем

$$\bar{M}_{ij}'^q(t) = \delta_q^i + \sum_{m \in I_\mu} \sum_{l=1}^{n_q} p_{ql} s_{ql}^m \bar{M}_{ij}'^m(t-1) -$$
$$- \sum_{m \in I_\mu} \sum_{l=1}^{n_q} p_{ql} s_{ql}^m \bar{M}_{ij}'^m(t-1) \cdot \sum_{n \in I_\mu} (s_{ql}^n - \delta_n^m) Q_n(t-1)(1+o(1)). \quad (72)$$

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

$$\begin{aligned} a_m^q &= \sum_{l=1}^{n_q} p_{ql} s_{ql}^m \\ b_{mn}^q &= \sum_{l=1}^{n_q} p_{ql} s_{ql}^m (s_{ql}^n - \delta_n^m) \end{aligned} \quad (73)$$

Подставляя их в (72), получаем

$$\bar{M}_{ij}'^q(t) = \delta_q^i p_{ij} + \sum_{m \in I_\mu} a_m^q \bar{M}_{ij}'^m(t-1) -$$
$$- c_\mu t^{\xi(\mu)-1} \sum_{\substack{m \in I_\mu \\ n \in I}} b_{mn}^q u_{n-k_{I(n)-1}}^{I(n)} \bar{M}_{ij}'^m(t-1)(1+o(1)) \quad (74)$$

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 3, получаем:

$$\bar{M}'^q_{ij}(t) = u^{(\mu)}_{q-k_{\mu-1}} \cdot \sum_{l\in I_\mu} v^{(\mu)}_{n-k_{\mu-1}} \bar{M}'^l_{ij}(t) = u^{(\mu)}_{q-k_{\mu-1}} M^{(\mu)}_*(t), \tag{75}$$

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

$$\delta\bar{M}^{(\mu)}_*(t) = v^{(\mu)}_{i-k_{\mu-1}} p_{ij} - c_\mu t^\alpha \sum_{q,m,n\in I_\mu} v^{(\mu)}_{q-k_{\mu-1}} b^q_{mn} u^{(\mu)}_{m-k_{\mu-1}} u^{(\mu)}_{n-k_{\mu-1}} = v^{(\mu)}_{i-k_{\mu-1}} p_{ij} - c_\mu t^\alpha B_\mu, \tag{76}$$

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

Нетрудно видеть, что величина $\bar{M}^{(\mu)}_*(t)$ удовлетворяет условиям леммы 4. Применяя её, получаем:

$$\begin{array}{ll} \bar{M}^{(\mu)}_*(t) = \dfrac{v^{(\mu)}_{i-k_{\mu-1}} p_{ij}}{c_\mu B_\mu + 1} t(1+o(1)), & \text{если } \alpha = -1 \\ \bar{M}^{(\mu)}_*(t) = \dfrac{v^{(\mu)}_{i-k_{\mu-1}} p_{ij}}{c_\mu B_\mu} t^{-\alpha}(1+o(1)), & \text{если } \alpha > -1 \end{array}, \tag{77}$$

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:

$$\bar{M}'^q_{ij}(t) = \delta^i_q p_{ij} Q_{s_{ij}}(t-1) + \sum_{m\in I_\mu \cup I_{\mu+1}} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^m_{ij}(t-1) Q_{s_{ql}-\delta^m}(t-1) \tag{78}$$

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

$$\bar{M}'^q_{ij}(t) = O(p_{ij}) + \sum_{m\in I_\mu} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^m_{ij}(t-1) \cdot \left(1 - \sum_{n\in I_\mu} (s^n_{ql} - \delta^m_n) Q_n(t-1)\right)(1+o(1)) + \\ + \sum_{m\in I_{\mu+1}} \sum_{l=1}^{n_q} p_{ql} s^m_{ql} \bar{M}'^m_{ij}(t-1)(1+o(1)) \tag{79}$$

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 (77). 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:

$$\bar{M}'^q_{ij}(t) = \sum_{m\in I_\mu} a^q_m \bar{M}'^m_{ij}(t-1) - c_\mu t^{\alpha(\mu)} \sum_{m,n\in I_\mu} b^q_{mn} u^{(\mu)}_{n-k_{\mu-1}} \bar{M}'^m_{ij}(t-1)(1+o(1)) + \\ + \overline{\mathcal{M}}'_{\mu+1} t^{\gamma(\mu+1)} \sum_{m\in I_{\mu+1}} a^q_m u^{(\mu+1)}_{m-k_\mu}(1+o(1)) \tag{80}$$

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

$$\delta\bar{M}^{(\mu)}_*(t) = \overline{\mathcal{M}}'_{\mu+1}\left(\sum_{\substack{q\in I_\mu\\ m\in I_{\mu+1}}} v^{(\mu)}_{q-k_{\mu-1}} a^q_m u^{(\mu+1)}_{m-k_\mu}\right)\cdot t^{\gamma(\mu+1)} -$$

$$- c_\mu t^{\alpha(\mu)}\cdot\left(\sum_{q,m,n\in I_\mu} v^{(\mu)}_{q-k_{\mu-1}} b^q_{mn} u^{(\mu)}_{m-k_{\mu-1}} u^{(\mu)}_{n-k_{\mu-1}}\right)\cdot \bar{M}^{(\mu)}_*(t-1)(1+o(1)) =$$

$$= \overline{\mathcal{M}}'_{\mu+1} b_{\mu+1} t^{\gamma(\mu+1)}(1+o(1)) - c_\mu B_\mu t^{\alpha(\mu)}\bar{M}^{(\mu)}_*(t-1)(1+o(1)) \quad (81)$$

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 3, получаем

$$\bar{M}^{(\mu)}_*(t) = \frac{\overline{\mathcal{M}}'_{\mu+1} b_{\mu+1} t^{\gamma(\mu+1)-\alpha(\mu)}}{c_\mu B_\mu} \quad (82)$$

Из полученных формул (77) и (82) нетрудно получить общее выражение для величины $\bar{M}^{(\mu)}_*(t)$, при условии что грамматика имеет вид «цепочки» и состоит только из критических классов ($J=\{1,2,\ldots,m\}$).

$$\bar{M}^{(\mu)}_*(t) = \prod_{j=\mu}^{\nu-1}\left(\frac{b_{j+1}}{c_j B_j}\right)\cdot\left(\frac{v^{(\nu)}_{i-k_{\nu-1}} p_{ij}}{c_\nu B_\nu + \delta^m_\nu}\right)\cdot t^{\frac{1}{2}^{m-\nu}\left(2-\frac{1}{2}^{\nu-\mu}\right)}, \quad (83)$$

где $\mu=I(q)$, $\nu=I(i)$. Подставляя (75), и затем (63), непосредственно получаем

$$\bar{M}^q_{ij}(t) = \frac{u^{(\mu)}_{q-k_{\mu-1}}}{1-Q_q(t)}\prod_{j=\mu}^{\nu-1}\left(\frac{b_{j+1}}{c_j B_j}\right)\cdot\left(\frac{v^{(\nu)}_{i-k_{\nu-1}} p_{ij}}{c_\nu B_\nu + \delta^m_\nu}\right)\cdot t^{\frac{1}{2}^{m-\nu}\left(2-\frac{1}{2}^{\nu-\mu}\right)} \quad (84)$$

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 (69) получаем

$$M'^q_{ij}(t) = O(p_{ij})t^{\beta(\mu)} + \sum_{m\in I_\mu}\sum_{l=1}^{n_q} p_{ql}s^m_{ql}M'^m_{ij}(t-1) -$$

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

$$+ \sum_{m,n\in I_\mu}\sum_{l=1}^{n_q} p_{ql}s^m_{ql}(s^n_{ql}-\delta^m_n)P_n(t-1)\bar{M}'^m_{ij}(t-1) \quad (85)$$

Обозначим

$$\begin{aligned} Q^{(\mu)}(t) &= c_\mu u^{(\mu)} t^{\alpha(\mu)} \\ P^{(\mu)}(t) &= d_\mu u^{(\mu)} t^{\beta(\mu)} \end{aligned} \quad (86)$$

Подставляя выражения (73) для первых и вторых моментов в (85), имеем

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

$$+ d_\mu t^{\beta(\mu)}\sum_{m,n\in I_\mu} b^q_{mn} u^{(\mu)}_{n-k_{\mu-1}} \bar{M}^m_{ij}(t-1)(1+o(1)) \quad (87)$$

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

$$M_{ij}'^q(t)=\sum_{m\in I_\mu} a_m^q M_{ij}'^m(t-1)-c_\mu t^{\alpha(\mu)}\sum_{m,n\in I_\mu} b_{mn}^q u_{n-k_{\mu-1}}^{(\mu)} M_{ij}'^m(t-1)(1+o(1))+ \\ +d_\mu\mathcal{M}_\mu t^{-1}\sum_{m,n\in I_\mu} b_{mn}^q u_{m-k_{\mu-1}}^{(\mu)} u_{n-k_{\mu-1}}^{(\mu)}(1+o(1)) \quad (88)$$

Применяя лемму 3, получаем

$$\begin{aligned} M_{ij}'^q(t)&=u_{q-k_{\mu-1}}^{(\mu)} M_*^{(\mu)}(t)(1+o(1)) \\ M_*^{(\mu)}(t)&=\sum_{m\in I_\mu} v_{n-k_{\mu-1}}^{(\mu)} M_{ij}'^n(t) \end{aligned} \quad (89)$$

Домножая (88) на $v_{q-k_{\mu-1}}^{(\mu)}$ и суммируя по q, имеем

$$\delta M_*^{(\mu)}(t)=d_\mu\mathcal{M}_\mu B_\mu t^{-1}-c_\mu t^{\alpha(\mu)} B_\mu M_*^{(\mu)}(t-1)(1+o(1)) \quad (90)$$

Применяя лемму 4, получаем в результате

$$M_*^{(\mu)}(t)=\begin{cases} d_\mu\overline{\mathcal{M}}_\mu' B_\mu(1+o(1)), & \text{при } \mu=m, \\ \frac{d_\mu\overline{\mathcal{M}}_\mu'}{c_\mu} t^{-1-\alpha(\mu)}(1+o(1)), & \text{при } \mu<m \end{cases} \quad (91)$$

Пусть теперь $I(q)=\mu$, $I(i)=\nu$, $\mu<\nu$, тогда из (69) получаем

$$M_{ij}'^q(t)=\delta_i^q p_{ij} R_{s_{ij}}(t-1)+\sum_{m\in I_\mu\cup I_{\mu+1}}\sum_{l=1}^{n_q} p_{ql} s_{ql}^m\cdot[Q_{s_{ql}-\delta^m}(t-1)M_{ij}'^m(t-1)+ \\ +(1-Q_m(t-1))R_{s_{ql}-\delta^m}(t-1)\bar{M}_{ij}^m(t-1)], \quad (92)$$

откуда

$$M_{ij}'^q(t)=O\left(t^{\beta(\mu)}\right)+\sum_m a_m^q M_{ij}'^m(t-1)-\sum_{\substack{m\in I_\mu\cup I_{\mu+1}\\ n\in I_\mu}} b_{mn}^q Q_n(t-1)M_{ij}'^m(t-1)(1+o(1))+ \\ +\left(\sum_{\substack{m\in I_\mu\cup I_{\mu+1}\\ n\in I_\mu}} b_{mn}^q P_n(t-1)\bar{M}_{ij}^m(t-1)\right)(1+o(1)) \quad (93)$$

Можем записать

$$M_{ij}'^q(t)=\sum_{m\in I_\mu} a_m^q M_{ij}'^m(t-1)-\sum_{m,n\in I_\mu} b_{mn}^q Q_n(t-1)M_{ij}'^m(t-1)(1+o(1))+ \\ +\sum_{m,n\in I_\mu} b_{mn}^q P_n(t-1)\bar{M}_{ij}^m(t-1)(1+o(1)) \quad (94)$$

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

$$\delta M_*^{(\mu)}(t)=d_\mu\overline{\mathcal{M}}_\mu B_\mu t^{\beta(\mu)+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)}(1+o(1))-c_\mu B_\mu t^{\alpha(\mu)}\cdot M_*^{(\mu)}(t-1)(1+o(1)) \quad (95)$$

Так как $\mu < m$, $\alpha(\mu) > -1$, поэтому по лемме 4 получаем

$$M_*^{(\mu)}(t) = \frac{d_\mu \overline{\mathcal{M}}_\mu}{c_\mu} t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)} \tag{96}$$

Объединяя результаты (91) и (96), получаем

$$M_*^{(\mu)}(t) = \frac{d_\mu \overline{\mathcal{M}}_\mu B_\mu}{\delta_\mu^m (c_\mu B_\mu - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)} (1 + o(1)), \tag{97}$$

после чего из (88)

$$M_{ij}^q(t) = \frac{u_{q-k_{\mu-1}}^{(\mu)}}{P_q(t)} \frac{d_\mu \overline{\mathcal{M}}_\mu B_\mu}{\delta_\mu^m (c_\mu B_\mu - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{m-\nu}\left(2-\left(\frac{1}{2}\right)^{\nu-\mu}\right)} (1 + o(1)) \tag{98}$$

Подставляя из (84) выражение для $\bar{M}_{ij}^q(t)$, а также полагая $\mu = 1$, для $M_{ij}(t) = M_{ij}^1(t)$ получаем:

$$M_{ij}(t) = d_i p_{ij} t^{\left(\frac{1}{2}\right)^{m-\nu-1}} \cdot (1 + o(1)), \tag{99}$$

где ν — номер класса, содержащего нетерминал A_i, а m — номер последнего класса в цепочке.

5.2 Случай докритического класса

Аналогично рассмотрим случай, когда $I(q) \notin J$. Рассмотрение этого случая полностью аналогично рассмотрению предыдущего случая, за исключением выражений для $Q_i(t)$ и $P_i(t)$. Проводя те же рассуждения для этих выражений, получаем:

$$M_{ij}(t) = d_i p_{ij} t^{\left(\frac{1}{2}\right)^{q_\mu-1-\xi}} \cdot (1 + o(1)), \tag{100}$$

где $A_i \in K_\mu$, q_μ — число классов с номерами из J в цепочке $K_\mu, K_{\mu+1}, \ldots, K_m$ и

$$\xi = \begin{cases} 1, & \text{если } \mu \in J \\ 0, & \text{если } \mu \notin J \end{cases} \tag{101}$$

Объединяя полученные результаты, можем сформулировать теорему:

Теорема 4 *Пусть грамматика G имеет вид «цепочки» (7). Тогда математическое ожидание $M_{ij}(t)$ числа применений правила r_{ij} в деревьях вывода высоты t имеет вид:*

$$M_{ij}(t) = d_i p_{ij} t^{\left(\frac{1}{2}\right)^{q_\mu-1-\xi}} \cdot (1 + o(1)), \tag{102}$$

где μ — номер класса, содержащего нетерминал A_i, p_{ij} — вероятность правила r_{ij} в грамматике, d_i — некоторая константа и

$$\xi = \begin{cases} 1, & \textit{если } \mu \in J \\ 0, & \textit{если } \mu \notin J \end{cases}, \tag{103}$$

где J — множество номеров классов с перроновым корнем, равным 1.

Как видно из формулы, величина $M_{ij}(t)$ определяется удалённостью класса K_μ, содержащего нетерминал A_i из левой части правила r_{ij}, от конца цепочки. Чем больше число критических классов, следующих за классом K_μ, тем больше применений правила r_{ij} будет в дереве вывода. В случае рассмотрения только двух классов данный результат согласуется с полученным в [8].

6 Заключение

В результате проведённого исследования были изучены основные вероятностные характеристики грамматик заданного класса. Полученные асимптотические оценки позволяют непосредственно перейти к построению алгоритма асимптотически оптимального кодирования для рассматриваемого класса языков сообщений, а также существенно упрощают исследование этой задачи для КС-языков в общем случае.

Список литературы

[1] Шеннон К. Математическая теория связи. М.: ИЛ, 1963

[2] Марков А.А. Введение в теорию кодирования. М.: Наука, 1982

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

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

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

[6] Жильцова Л.П. Закономерности применения правил грамматики в выводах слов стохастического контекстно-свободного языка // Математические вопросы кибернетики. Вып. 9. М.: Наука, 2000. С. 100–126.

[7] Жильцова Л.П. О нижней оценке стоимости кодирования и асимптотически оптимальном кодировании стохастического контекстно-свободного языка // Дискретный анализ и исследование операций. Серия 1, т. 8, №3. Новосибирск: Издательство Института математики СО РАН, 2001. С. 26–45.

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

[9] Борисов А.Е. О свойствах стохастического КС-языка, порождённого грамматикой с двумя классами нетерминальных символов // Дискретный анализ и исследование операций. Серия 1, т. 12, №3. Новосибирск: Издательство Института математики СО РАН, 2005. С. 3–31.