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

# КУРСОВАЯ РАБОТА

Тема:
**«Энтропия множества деревьев вывода в разложимой стохастической КС-грамматике, имеющей вид "цепочки"»**

**Заведующий кафедрой:**
Д. ф.-м. н. Шевченко Валерий Николаевич

**Выполнил:** студент группы 85М1
Мартынов Игорь Михайлович

**Научный руководитель:**
Д. ф.-м. н. Жильцова Лариса Павловна

Нижний Новгород
2012

# Содержание

**1 Введение** **2**

**2 Основные определения** **2**

**3 Вероятности продолжения** **5**

**4 Математические ожидания числа применений правила в деревьях вывода** **11**

**5 Энтропия** **18**

**6 Заключение** **20**

# 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 структурые 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 [8] [9]. Кроме того, она показала, что перронов корень [6] матрицы первых моментов [5] 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 [10]. 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», причём среди классов могут присутствовать как критические, так и докритические. 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. Кроме того, получена асимптотика энтропии множества деревьев фиксированной высоты, которая будет использована для построения асимптотически оптимального алгоритма кодирования.

# 2 Основные определения

*Стохастической КС-грамматикой* [3] называется система $G =< V_T, V_N, R, s >$, 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

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

где $A_i \in V_N$, $\beta_{ij} \in (V_N \cup V_T)^*$ и $p_{ij}$ — вероятность применения правила $r_{ij}$, причём

$$0 < p_{ij} \leqslant 1, \qquad \sum_{j=1}^{n_i} p_{ij} = 1. \tag{2}$$

Для $\alpha, \gamma \in (V_N \cup V_T)^*$ будем обозначать $\alpha \Rightarrow \gamma$, если существуют $\alpha_1, \alpha_2 \in (V_N \cup V_T)^*$, для которых $\alpha = \alpha_1 A_i \alpha_2$, $\gamma = \alpha_1 \beta_{ij} \alpha_2$ и в грамматике имеется правило $A_i \xrightarrow{p_{ij}} \beta_{ij}$. Через $\Rightarrow_*$ обозначим рефлексивное транзитивное замыкание отношения $\Rightarrow$. Грамматика $G$ задаёт контекстно-свободный язык $L_G = \{\alpha \in V_T^* : s \Rightarrow_* \alpha\}$.

Выводом слова $\alpha$ назовём последовательность правил $\omega(\alpha) = (r_{i_1 j_1}, r_{i_2 j_2}, \ldots, r_{i_q j_q})$, с помощью последовательного применения которых слово $\alpha$ выводится из аксиомы $s$. Если при этом каждое правило применяется к самому левому нетерминалу в слове, такой вывод называется левым. Для вывода $\omega(\alpha) = (r_{i_1 j_1}, \ldots, r_{i_q j_q})$ определим величину $p(\omega(\alpha)) = p_{i_1 j_1} \cdot \ldots \cdot p_{i_q j_q}$.

Важное значение имеет понятие *дерева вывода* [4]. Дерево вывода для слова $\alpha$ строится следующим образом. Корень дерева помечается аксиомой $s$. Далее последовательно рассматриваются правила левого вывода $\omega(\alpha) = r_{i_1 j_1}, r_{i_2 j_2}, \ldots, r_{i_q j_q}$. Пусть на очередном шаге рассматривается правило $A_i \xrightarrow{p_{ij}} b_{i1} b_{i2} \ldots b_{i,m}$, где $b_{i,l} \in (V_N \cup V_T)$ $(1 \leqslant l \leqslant m)$. Тогда из самой левой вершины-листа дерева, помеченной символом $A_i$, проводится $m$ дуг в вершины следующего яруса, которые помечаются слева направо символами $b_{i1}, \ldots, b_{i,m}$ соответственно. После построения дуг и вершин для всех правил в выводе листья дерева помечены терминальными символами (либо пустым словом $\lambda$, если применяется правило вида $A_i \xrightarrow{p_{ij}} \lambda$) и само слово получается при обходе листьев дерева слева направо. *Высотой* дерева вывода будем называть максимальную длину пути от корня к листу.

Обозначим $p(\alpha) = \sum \omega(\alpha)$, где сумма берётся по всем левым выводам слова $\alpha$. Грамматика $G$ называется *согласованной*, если

$$\lim_{n \to \infty} \sum_{\substack{\alpha \in L_G \\ |\alpha| \leqslant n}} p(\alpha) = 1. \tag{3}$$

Согласованная грамматика $G$ задаёт распределение вероятностей $P$ на множестве $L_G$ и определяет *стохастический КС-язык* $\mathcal{L} = (L, P)$. В этом случае величина $p(\alpha)$ — вероятность слова $\alpha$ в $L_G$. В дальнейшем будем рассматривать только согласованные грамматики.

Определим многомерные производящие функции [3]:

$$F_i(s_1, s_2, \ldots, s_k) = \sum_{j=1}^{n_i} p_{ij} s_1^{l_1} s_2^{l_2} \ldots s_k^{l_k} \quad (1 \leqslant i \leqslant k), \tag{4}$$

где $n_i$ — число правил вывода с нетерминалом $A_i$ в левой части, и $l_m = l_m(i, j)$ — число вхождений нетермина $A_m$ в правую часть правила $A_i \xrightarrow{p_{ij}} \beta_{ij}$.

Определим первые и вторые моменты грамматики ($a^i_j$ и $b^i_{jl}$ соответственно):

$$a^i_j = \left.\frac{\partial F_i(s_1, s_2, \ldots, s_k)}{\partial s_j}\right|_{s_1=\ldots=s_k=1} \qquad (1 \leqslant i, j, l \leqslant k) \tag{5}$$
$$b^i_{jl} = \left.\frac{\partial^2 F_i(s_1, s_2, \ldots, s_k)}{\partial s_l \partial s_j}\right|_{s_1=\ldots=s_k=1}$$

Для нетерминалов $A_i, A_j$ будем обозначать $A_i \to A_j$, если в грамматике имеется правило $A_i \xrightarrow{p_{ij}} \alpha_1 A_j \alpha_2$, где $\alpha_1, \alpha_2 \in (V_N \cup V_T)^*$. Рефлексивное транзитивное замыкание отношения $\to$ обозначим $\to_*$. Если одновременно $A_i \to_* A_j$ и $A_j \to_* A_i$, будем обозначать $A_i \leftrightarrow_* A_j$. Отношение $\leftrightarrow_*$ разбивает множество нетерминалов грамматики на классы $K_1, K_2, \ldots, K_m$. Множества номеров нетерминалов, входящих в класс $K_j$ обозначим через $I_j$. При $m \geqslant 2$ грамматика называется *разложимой*.

Разложимой грамматике соответствует разложимая матрица [6] первых моментов. Обозначим $K_i \prec K_j$ , если $i \neq j$ и существуют такие $A_1 \in K_i$ и $A_2 \in K_j$, что $A_1 \to A_2$. Будем говорить, что грамматика имеет вид «цепочки», если она разложима, и для множества классов выполняется соотношение $K_1 \prec K_2 \prec \ldots \prec K_m$. При этом граф, построенный на множестве классов по отношению $\prec$, имеет вид $P_m$:

$$K_{i_1} \longrightarrow K_{i_2} \longrightarrow \quad \ldots \quad \longrightarrow K_{i_m}$$

Назовём класс $K$ *особым*, если он содержит ровно один нетерминал $A_i$, и в грамматике отсутствует правило вида $A_i \xrightarrow{p_{ij}} \alpha_1 A_i \alpha_2$, где $\alpha_1, \alpha_2 \in (V_N \cup V_T)^*$.

В работе рассматриваются грамматики, имеющие вид «цепочки», без особых классов.

Матрица первых моментов такой грамматики имеет вид

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

Блок $A_{ii}$ соответствует классу $K_i$ и является неразложимой неотрицательной матрицей.

По определению (5), матрицы $A_{11}, A_{22}, \ldots, A_{m,m}$ неразложимы и неотрицательны. Согласно теореме Фробениуса [6], неразложимая неотрицательная матрица всегда имеет простое положительное собственное число (перронов корень), максимальное по модулю. Обозначим перроновы корни матриц $A_{11}, \ldots, A_{m,m}$ через $r_1, \ldots, r_m$ соответственно. Тогда $r = \max r_1, \ldots, r_m$ — перронов корень всей матрицы $A$. В данной работе рассматривается случай $r = 1$. По аналогии с теорией ветвящихся процессов [5], будем называть этот случай *критическим*. Классы, для которых перронов корень соответствующих подматриц равен 1, также будем называть критическими.

Обозначим $s_{lh}$ (при $l \leqslant h$) — число критических классов среди подцепочки $K_l, K_{l+1}, \ldots, K_h$. Разобьём последовательность классов на группы $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_w$, где $w = s_{1m}$. При этом к группе $\mathcal{M}_j$ отнесём докритические классы $K_l$, для которых $s_{1l} < l$ и $l$-й критический класс. Докритические классы $K_l$, для которых $s_{1l} = w$, также отнесём к классу $\mathcal{M}_w$. Таким образом, каждая группа содержит в себе ровно один критический класс.

Тогда матрицу $A$ можно представить в виде:

$$A = \begin{pmatrix} B_{11} & B_{12} & 0 & \cdots & 0 & 0 \\ 0 & B_{22} & B_{23} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & B_{w-1,w-1} & B_{w-1,w} \\ 0 & 0 & 0 & \cdots & 0 & B_{w,w} \end{pmatrix},$$

где матрица $B_{ij}$ находится на пересечении строк для классов группы $\mathcal{M}_i$ и столбцов для классов группы $\mathcal{M}_j$. Матрицы $B_{ij}$, в свою очередь, имеют вид

$$B_{ij} = \begin{pmatrix} C_{11} & C_{12} & 0 \\ 0 & C_{22} & C_{23} \\ 0 & 0 & C_{33} \end{pmatrix},$$

где $C_{22} = A_{lh}$ — блок, стоящий на пересечении строк $i$-го критического класса и столбцов $j$-го критического класса. Матрицы $C_{11}$ и $C_{33}$ содержат в себе, вообще говоря, несколько блоков, стоящих на пересечении строк и столбцов соответствующих докритических классов.

Для такого разбиения матрицы $A$ из [7] следует

**Теорема 1** *При* $t \to \infty$

$$B_{lh}^{(t)} = U^{(h)} V^{(l)} t^{s_{lh}-1} r^t (1 + o(1)),$$

*где* $U^{(h)} = \left(U_I^{(h)}, U_{II}^{(h)}, U_{III}^{(h)}\right)$ *и* $V^{(l)} = \left(V_I^{(l)}, V_{II}^{(l)}, V_{III}^{(l)}\right)$ *— соответственно правый и левый собственные векторы матрицы* $B_{lh}$.

*При этом* $U_{II}^{(h)}$ *— правый собственный вектор матрицы* $A_{ii}$*, соответствующей* $h$*-му критическому классу, а* $V_{II}^{(l)}$ *— левый собственный вектор матрицы* $A_{jj}$*, соответствующей* $l$*-му критическому классу.*

# 3 Вероятности продолжения

Введём обозначения

$$\begin{aligned} F_i(\mathbf{s}) &= F_i(0, \mathbf{s}) = \sum_{j=1}^{n_i} p_{ij} s_1^{l_1} s_2^{l_2} \ldots s_k^{l_k} \quad (1 \leqslant i \leqslant k), \\ F_i(t, \mathbf{s}) &= F_i(\mathbf{F}(t-1, \mathbf{s})) \quad (t > 0,\ 1 \leqslant i \leqslant k), \end{aligned} \tag{7}$$

где $\mathbf{s} = (s_1, \ldots, s_k)$, $0 \leqslant s_j \leqslant 1$ и $\mathbf{F}(t, \mathbf{s}) = (F_1(t, \mathbf{s}), \ldots, F_k(t, \mathbf{s}))$.

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

$$1 - F_i(\mathbf{s}) = \sum_{j=1}^{n_i} a_j^i (1 - s_j) - \frac{1}{2} \sum_{1 \leqslant j, l \leqslant n_i} b_{jl}^i (1 - s_j)(1 - s_l) + O\left(\sum_{j=1}^{k} |1 - s_j|^3\right), \tag{8}$$

Подставляя в качестве $\mathbf{s}$ вектор $\mathbf{F}(t,s) = (F_1(t,s), F_2(t,s), \ldots, F_k(t,s))$, получаем:

$$1 - F_i(t+1,s) = \sum_{i=1}^{k} a_j^i (1 - F_j(t,s)) - \frac{1}{2} \sum_{1 \leqslant j,l \leqslant k} b_{jl}^i (1 - F_j(t,s))(1 - F_l(t,s)) + \\ + O\left(\sum_{j=1}^{k} |1 - F_j(t,s)|^3\right) \quad (9)$$

Введём вектор *вероятностей продолжения* $Q(t) = (Q_1(t), Q_2(t), \ldots, Q_k(t))^T$, положив

$$Q_i(t) = 1 - F_i(t,\mathbf{s})\big|_{s_1 = s_2 = \ldots = s_k = 0} \quad (10)$$

Тогда уравнение (9) перепишется в виде

$$Q_i(t+1) = \sum_{i=1}^{k} a_j^i Q_i(t) - \frac{1}{2} \sum_{1 \leqslant j,l \leqslant k} b_{jl}^i (1 - Q_j(t))(1 - Q_l(t)) + O\left(\sum_{j=1}^{k} |Q_j(t,s)|^3\right) \quad (11)$$

Для каждого из классов $K_n$ будем рассматривать вектор $Q^{(n)}(t)$ — вектор-столбец, содержащий вероятности продолжения для нетерминалов $K_n$ в порядке их нумерации. Тогда

$$Q(t) = \begin{pmatrix} Q^{(1)}(t) \\ Q^{(2)}(t) \\ \vdots \\ Q^{(m)}(t) \end{pmatrix}, \quad Q^{(j)}(t) \in \mathbb{R}^{k_j}, \quad (12)$$

где $k_j = |K_j|$. Тогда уравнение (11) можно записать в виде

$$Q_i(t+1) = \sum_{j \in I_n} a_j^i Q_j(t) + \sum_{i \in I_{n+1}} a_j^i Q_j(t) \cdot (1 + o(1)) \qquad (i \in I_n, n < m) \quad (13)$$

$$Q_i(t+1) = \sum_{j \in I_m} a_j^i Q_j(t) \cdot (1 + o(1)) \qquad (i \in I_m) \quad (14)$$

или, в матричном виде,

$$Q^{(n)}(t+1) = A_{n,n} Q^{(n)}(t) + A_{n,n+1} Q^{(n+1)}(t)(1 + o(1)) \quad (15)$$

Для всего вектора $Q(t)$ верно равенство

$$Q(t+1) = (A - A(t))Q(t), \quad (16)$$

где $A(t)$ — матрица, составленная из элементов $a_{ij} = \frac{1}{2}\sum_{l=1}^{k} b_{jl}^i Q_l(t)$ $(1 \leqslant i,j \leqslant k)$. В силу согласованности грамматики $Q(t) \to 0$ и, следовательно, $A(t) \to 0$ при $t \to \infty$.

Докажем, что компоненты вектора $Q^{(n)}(t)$ пропорциональны некоторому вектору $U^{(n)}$. Доказательство аналогичного факта для случая двух классов принадлежит А. Борисову. Здесь мы проведём похожие рассуждения.

Зафиксируем некоторое $\tau \geqslant 0$. Тогда из (16) получаем

$$Q(t+1) = (A - A(t)) \cdot \ldots \cdot (A - A(\tau))Q(\tau) \quad (17)$$

Обозначим

$$
\begin{aligned}
&A^*(t) = (A - A(t)) \cdot (A - A(t-1)) \cdot \ldots \cdot (A - A(\tau+1)) \\
&\tilde{A}^*_{ij} = \frac{A^*_{ij}(t)}{t^{s_{ij}}} \\
&\tilde{A}_{ij} = \frac{A^{(t)}_{ij}}{t^{s_{ij}}},
\end{aligned}
\tag{18}
$$

где $A^{(t)}_{ij}$ — блоки, расположенные на месте блоков $A_{ij}$ в матрице $A^t$ и $s_{ij}$ — число критических классов в подцепочке $K_i, K_{i+1}, \ldots, K_j$.

Из исследования асимптотики матрицы $A^t$ известно [7], что $\tilde{A}_{ij}(t) \to \tilde{a}_{ij} U^{(i)} V^{(j)}$, где $\tilde{a}_{ij}$ — некоторые константы, $U^{(i)}$ — вектор-строка длины $k_i$, а $V^{(j)}$ — вектор-столбец длины $k_j$.

Выберем произвольные $\varepsilon_1, \varepsilon_2$, такие что $0 < \varepsilon_1, \varepsilon_2 < 1$. Тогда существуют функции $l(\varepsilon_1)$ и $n(\varepsilon_2)$, такие что

$$
\begin{array}{l}
\left|\tilde{A}_{ij}(l(\varepsilon_1)) - \tilde{a}_{ij} U^{(i)} V^{(j)}\right| < \varepsilon_1 E \\
\forall t \geqslant n(\varepsilon_2) \quad A(t) < \varepsilon_2 A
\end{array}.
\tag{19}
$$

Рассмотрим произвольный вектор-столбец $x > \mathbf{0}$ длины $k$. Тогда выполняется оценка

$$
(1-\varepsilon_2)^l A^l x^{(\tau)} \leqslant A^*(t) x^{(\tau)} \leqslant A^l x^{(\tau)},
\tag{20}
$$

где $x^{(\tau)} = (A - A(\tau))x$. Записывая это неравенство отдельно для блоков $A_{ij}$, получаем

$$
(1-\varepsilon_2)^l A^l_{ij} x^{(\tau)}_j \leqslant A^*_{ij}(l) x^{(\tau)}_j \leqslant A^{(l)}_{ij} x^{(\tau)}_j,
\tag{21}
$$

откуда

$$
(1-\varepsilon_2)^l \tilde{A}_{ij}(l) x^{(\tau)} \leqslant \tilde{A}^*_{ij}(l) x^{(\tau)}_j \leqslant \tilde{A}_{ij}(l) x^{(\tau)}
\tag{22}
$$

Вычитая из всех частей неравенства $\tilde{A}_{ij}(l) x^{(\tau)}_j$, получаем оценку

$$
\left|\left(\tilde{A}^*_{ij}(l) - \tilde{A}_{ij}(l)\right) x^{(\tau)}_j\right| \leqslant (1 - (1-\varepsilon_2)^l) \tilde{A}_{ij}(l) x^{(\tau)}
\tag{23}
$$

Используя эту оценку, можем записать

$$
\begin{aligned}
\left|\tilde{A}^*_{ij}(t) - \tilde{a}_{ij} U^{(i)} V^{(j)} x^{(\tau)}_j\right| &\leqslant \left|\left(\tilde{A}^*_{ij}(t) - \tilde{A}_{ij}(t)\right) x^{(\tau)}\right| + \\
&+ \left|\left(\tilde{A}_{ij}(l) - \tilde{a}_{ij} U^{(i)} V^{(j)}\right) x^{(\tau)}_j\right| \leqslant (1 - (1-\varepsilon_2)^l) \tilde{A}_{ij}(l) x^{(\tau)}_j + \varepsilon_1 x^{(\tau)}_j \leqslant \\
&\leqslant (1 - (1-\varepsilon_2)^l) h k_j x^{(\tau)}_j + \varepsilon_1 x^{(\tau)}_j \leqslant \left((1 - 1 - \varepsilon_2)^l) h k_j + \varepsilon_1\right) x^*_j(\tau),
\end{aligned}
\tag{24}
$$

где $h = \max_{i,j,l}\left\{\tilde{A}_{ij}(l)\right\}$ и $x^*_j(\tau) = \max_i (x^{(\tau)}_j)_i$.

Устремляем $\varepsilon_2$ к нулю, затем $\varepsilon_1$ к нулю таким образом, чтобы выполнялось условие

$$
l(\varepsilon_1) \log(1-\varepsilon_2) \to -\infty
\tag{25}
$$

Тогда

$$
\left|\tilde{A}^*_{ij}(t) - \tilde{a}_{ij} U^{(i)} V^{(j)} x^{(\tau)}_j\right| \leqslant \varepsilon x^*_j(\tau) \quad (\varepsilon \to 0).
\tag{26}
$$

Домножая слева на $V^{(i)}$, имеем

$$\left|V^{(i)}\tilde{A}^*_{ij}(t)x_j^{(\tau)} - \tilde{a}_{ij}V^{(j)}x_j^{(\tau)}\right| \leqslant \varepsilon k_i \max\left\{(V^{(i)}\right\} x^*_j(\tau) \leqslant \varepsilon^* V^{(j)}x_j^{(\tau)}. \tag{27}$$

Отсюда,

$$\left|\frac{\tilde{A}^*_{ij}(t)x_j^{(\tau)}}{V^{(i)}\tilde{A}^*_{ij}(t)x_j^{(\tau)}} - \frac{\tilde{a}_{ij}U^{(i)}V^{(i)}x_j^{(\tau)}}{\tilde{a}_{ij}V^{(j)}x_j^{(\tau)}}\right| = \left|\frac{\tilde{A}^*_{ij}(t)x_j^{(\tau)}}{V^{(i)}\tilde{A}^*_{ij}(t)x_j^{(\tau)}} - U^{(i)}\right| \to 0 \tag{28}$$

или же

$$\left|\frac{A^*_{ij}(t)x_j^{(\tau)}}{V^{(i)}A^*_{ij}(t)x_j^{(\tau)}} - U^{(i)}\right| \to 0, \tag{29}$$

откуда

$$(A - A(t)) \cdot \ldots \cdot (A - A(\tau)) \cdot x_j = U^{(i)}V^{(i)}(A - A(t)) \cdot \ldots \cdot (A - A(\tau)) \cdot x_j \cdot (1 + o(1)) \tag{30}$$

Ввиду полученного выражения и (17) компоненты каждого из векторов $Q^{(n)}(t)$ пропорциональны компонентам вектора $U^{(n)}$.

Оценим теперь асимптотику элементов вектора $Q^{(n)}(t)$ при $t \to \infty$.

Положим $V^{(n)}Q^{(n)}(t) = Q^{(n)}_*(t)$, и домножим уравнение (11) скалярно на $V^{(n)}$. Заметим, что

$$Q^{(n)}(t) = U^{(n)}Q^{(n)}_*(t)(1 + o(1)). \tag{31}$$

$$\begin{aligned} Q^{(n)}_*(t+1) = Q(n)_*(t) + V^{(n)}B_{n,n+1}U^{(n+1)}Q^{(n+1)}_*(t) - \\ - \frac{1}{2}\sum_{1\leqslant i,j,l\leqslant k_n} V_i^{(n)}b^i_{jl}(n)U_j^{(n)}U_l^{(n)}\left(Q^{(n)}_*(t)\right)^2(1 + o(1)). \end{aligned} \tag{32}$$

Обозначим $\delta Q^{(n)}_*(t) = Q^{(n)}_*(t+1) - Q^{(n)}_*(t)$, а также

$$\begin{aligned} b_n &= V^{(n)}B_{n,n+1}U^{(n+1)} \\ B_n &= \sum_{1\leqslant i,j,l\leqslant k_n} V_i^{(n)}b^i_{jl}(n)U_j^{(n)}U_l^{(n)} \end{aligned}$$

Тогда уравнение (32) перепишется как

$$\delta Q^{(n)}_*(t) = b_n Q^{(n+1)}_*(t) - \frac{1}{2}B_n(Q^{(n)}_*(t))^2(1 + o(1)) \tag{33}$$

Выражение для $\delta Q^{(n)}_*(t)$ также можно получить из (11), вычитая это уравнение из себя с заменой $t \to t+1$:

$$\begin{aligned} \delta Q^{(n)}_*(t+1) = \sum_{j=1}^{k_n} a^i_j(n)\delta Q^{(n)}_j(t) + \sum_{j=1}^{k_{n+1}} a^i_j(n)\delta Q^{(n+1)}_j(t) - \\ - \frac{1}{2}\sum_{1\leqslant j,l\leqslant k_n} b^i_{jl}(n)\left(Q^{(n)}_j(t+1)Q^{(n)}_l(t+1) - Q^{(n)}_j(t)Q^{(n)}_l(t)\right)(1 + o(1)) \end{aligned}$$

Скалярно домножая на $V^{(n)}$, получим

$$\delta Q_*^{(n)}(t+1) = \delta Q_*^{(n)}(t) + b_n \delta Q_*^{(n+1)}(t) - \\ - \frac{1}{2} B_n \delta Q_*^{(n)}(t) \left( Q_*^{(n)}(t+1) + Q_*^{(n)}(t) \right) (1+o(1)) \quad (34)$$

Для последнего класса

$$Q_*^{(w)}(t) = c_w t^{-1}(1+o(1)), \quad (35)$$

что следует из неразложимого случая. Проведём рассуждение по индукции. Пусть для группы с номером $n+1$ верно

$$Q_*^{(n+1)}(t) = c_{n+1} t^{-\alpha}(1+o(1)),$$

где $0 < \alpha \leqslant 1$. Положим

$$z(t) = t^{\alpha} \delta Q_*^{(n)}(t)$$

Произведя замену в уравнении (34), и имея в виду, что $Q_*^{(n)}(t+1) = O(Q_*^{(n)}(t))$, получаем

$$\frac{z(t+1)}{(t+1)^{\alpha}} - \frac{z(t)}{t^{\alpha}} = b_n \delta Q_*^{(n+1)}(t)(1+o(1)) - \frac{1}{2} B_n \frac{z(t)}{t^{\alpha}} \cdot 2 Q_*^{(n)}(t)(1+o(1))$$

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

$$\frac{\delta z(t)}{(t+1)^{\alpha}} - \frac{\alpha z(t)(1+o(1))}{t(t+1)^{\alpha}} = b_n \delta Q_*^{(n)}(t) - \frac{B_n}{t^{\alpha}} Q_*^{(n)}(t) z(t)(1+o(1))$$

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

$$\frac{\delta z(t)}{(t+1)^{\alpha}} - \frac{\alpha z(t)(1+o(1))}{t(t+1)^{\alpha}} = -\frac{b_n \alpha c_{n+1}}{t(t+1)^{\alpha}} - \frac{B_n}{t^{\alpha}} Q_*^{(n)}(t) z(t)(1+o(1))$$

Домножая на $(t+1)^{\alpha}$, получаем

$$\delta z(t) - \frac{\alpha z(t)}{t} = -\frac{b_n \alpha c_{n+1}}{t} - B_n Q_*^{(n)}(t) z(t)(1+o(1))$$

Заметим, что, в силу предположения индукции, $\frac{1}{t} \leqslant Q_*^{(n+1)}(t) = o(Q_*^{(n)}(t))$, поэтому можно записать

$$\delta z(t) = -\frac{b_n \alpha c_{n+1}}{t} - B_n Q_*^{(n)}(t)(1+o(1)) \quad (36)$$

Известна следующая лемма (доказательство леммы принадлежит А. Борисову).

**Лемма 1** *Пусть последовательность z(t) (t = 1,2,...) удовлетворяет рекуррентному соотношению*

$$\delta z(t) = f(t) - g(t)z(t),$$

*где при $t \to \infty$ выполняются условия*

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

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

Полагая в уравнении (36) $f(t) = -\frac{b_n \alpha c_{n+1}}{t}(1 + o(1))$, $g(t) = B_n Q_*^{(n)}(t)(1 + o(1))$, замечаем, что для $z(t)$ выполняются все условия леммы (1), и соответственно, $z(t) \to 0$ при $t \to \infty$. Из определения $z(t)$ получаем:

$$\delta Q_*^{(n)}(t) = o\left(\frac{1}{t^\alpha}\right).$$

Подставляя эту оценку в (33), получаем

$$o\left(\frac{1}{t^\alpha}\right) = \frac{b_n c_{n+1}}{t^\alpha}(1 + o(1)) - \frac{B_n}{2}\left(Q_*^{(n)}(t)\right)^2(1 + o(1))$$

Отсюда

$$\frac{b_n c_{n+1}}{t^\alpha}(1 + o(1)) = \frac{B_n}{2}\left(Q_*^{(n)}(t)\right)^2(1 + o(1))$$

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

$$Q_*^{(n)}(t) = \sqrt{\frac{2b_n}{B_n}c_{n+1}\frac{1}{t^\alpha}}(1 + o(1)) = \sqrt{\frac{2b_n}{B_n}k_{n+1}} \cdot t^{-\frac{\alpha}{2}}(1 + o(1))$$

При этом, полагая $c_n = \sqrt{\frac{2b_n}{B_n}c_{n+1}}$, мы остаёмся в рамках предположения индукции. Учитывая (35), можем записать асимптотику $Q_*^{(n)}(t)$ для произвольной группы $n$:

$$Q_*^{(n)}(t) = \sqrt{\frac{2b_n}{B_n}\sqrt{\frac{2b_{n+1}}{B_{n+1}}\cdots\sqrt{\frac{2b_{w-1}}{B_{w-1}B_w}\cdot t^{-\left(\frac{1}{2}\right)^{w-n}}}}} =$$

$$= \prod_{k=n}^{w-1}\left(\frac{2b_n}{B_n}\right)^{\left(\frac{1}{2}\right)^{w-n+1}} \cdot \left(\frac{1}{B_w}\right)^{\left(\frac{1}{2}\right)^{w-n}} \cdot t^{-\left(\frac{1}{2}\right)^{w-n}}$$

Учитывая (31), получаем

$$Q_i(t) = c_n U_j^{(n)} t^{-\left(\frac{1}{2}\right)^{w-n}} \cdot (1 + o(1))$$
$$P_i(t) = \tilde{c}_n U_j^{(n)} t^{-1-\left(\frac{1}{2}\right)^{w-n}} \cdot (1 + o(1))$$

где нетерминал $A_i$ находится в последнем критическом классе цепочки или в одном из предшествующих классов, $n$ — номер группы, в которую входит класс, содержащий $A_i$, $w$ — число групп, и

$$c_n = \prod_{k=n}^{w-1}\left(\frac{2b_n}{B_n}\right)^{\left(\frac{1}{2}\right)^{w-n+1}} \cdot \left(\frac{1}{B_w}\right)^{\left(\frac{1}{2}\right)^{w-n}}$$

# 4 Математические ожидания числа применений правила в деревьях вывода

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

$$S^l_{ij}(t) = \sum_{\tau=1}^{t-1} q^l_{ij}(t,\tau)$$
$$\overline{S}^l_{ij}(t) = \sum_{\tau=1}^{t-1} \overline{q}^l_{ij}(t,\tau)$$

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

$$S_{ij}(t) = S^l_{ij}(t), \quad \overline{S}_{ij}(t) = \overline{S}^l_{ij}(t),$$
$$q_{ij}(t,\tau) = q^l_{ij}(t,\tau), \quad \overline{q}_{ij}(t,\tau) = \overline{q}^l_{ij}(t,\tau)$$

Рассмотрим математические ожидания некоторых из введённых величин. Обозначим

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

Для нахождения величин $\overline{M}^l_{ij}(t)$ и $M^l_{ij}(t)$ будут использованы следующие три леммы.

**Лемма 2** *[5] Пусть $s, d$ — натуральные числа, $m = (m_1, \ldots, m_s)$ — вектор целых неотрицательных чисел, $y = (y_1, \ldots, y_s)$ — вектор, и $\overline{m} = \sum_{j=1}^{s} m_j$. Тогда*

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

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

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

*причём*

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

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

**Лемма 3** *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 $\alpha$ — 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*

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

*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* $vu = 1$.

Доказательство леммы принадлежит А. Борисову.

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

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

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

$$(1) \quad x_t = \frac{\alpha t^{\alpha+1}}{\alpha+1}(1+o(1)) \qquad \textit{при} \quad \beta < -1, \alpha \geqslant 0 \tag{37}$$

$$(2) \quad x_t = \frac{\alpha t^{\alpha+1}}{\alpha+b+1}(1+o(1)) \qquad \textit{при} \quad \beta = -1, \alpha > -1 \tag{38}$$

$$(3) \quad x_t = \frac{\alpha t^{\alpha-\beta}}{b}(1+o(1)) \qquad \textit{при} \quad -1 < \beta < 0 \tag{39}$$

Доказательство леммы принадлежит А. Борисову.

Вначале рассмотрим $\overline{M}_{ij}^{q}(t)$. Пусть $p(\cdot)$ — вероятность дерева $d$ в грамматике $G$. Рассмотрим множество $D_{ql}^{\leqslant t}$ деревьев из $D_q^{\leqslant t}$, первый ярус которых получен применением правила $r_{ql}$ к корню дерева. Пусть

$$\overline{P}_{ql}^{ij}(t) = \sum_{d \in D_{ql}^{\leqslant t}} p(d) q_{ij}(d),$$

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

$$\overline{M}_{ij}^{q}(t) = \sum_{l=1}^{n_q} \overline{P}_{ql}(t). \tag{40}$$

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

$$q_{ij}(d) = q_{ij}^{(1)}(d) + q_{ij}^{(2)}(d),$$

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

$$\overline{P}_{ql}(t) = \sum_{d \in D_{ql}^{\leqslant t}} p(d) q_{ij}(d) = \sum_{d \in D_{ql}^{\leqslant t}} p(d) q_{ij}^{(1)}(d) + \sum_{d \in D_{ql}^{\leqslant t}} p(d) q_{ij}^{(2)}(d) = \overline{P}_{ql}^{(1)} + \overline{P}_{ql}^{(2)}(t)$$

Очевидно, $q^{(1)}_{ij}(d) = \delta^q_i \delta^l_j$ (где $\delta$ — символ Кронекера). Тогда

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

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

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

$$p(d) = \frac{p_{ql}}{1 - Q_q(t)} p_1(d) p_2(d) \dots p_{\overline{s_{ql}}}(d),$$

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

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

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 $d_1, d_2, \dots, d_{\overline{s}_{ql}}$, где $d_j$ — поддерево с корнем в $j$-м узле первого яруса дерева $d$. Тогда

$$\begin{aligned}
\overline{P}^{(2)}_{ql}(t) &= \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\overline{s}_{ql}} \sum_{d \in D^{\leqslant t}_{ql}} \left( \prod_{n=1}^{\overline{s}_{ql}} p_n(d) \right) q'^m_{ij}(d) = \\
&= \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\overline{s}_{ql}} \sum_{d_j : j \neq m} p_1(d) \dots p_{m-1}(d_{m-1}) p_{m+1}(d_{m+1}) \dots p_{\overline{s}_{ql}}(d_{\overline{s}_{ql}}) q'^m_{ij}(d) = \\
&= \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\overline{s}_{ql}} Q_{s_{ql} - \delta^m} q_{ij}(d_m) = \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{k} s^m_{ql} \overline{M}^m_{ij}(t-1) Q_{s_{ql} - \delta^m}(t-1)
\end{aligned}$$

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

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

Обозначая

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

имеем

$$\overline{M}'^q_{ij}(t) = \delta^q_i p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} s^m_{ql} \overline{M}'^m_{ij}(t-1) Q_{s_{ql} - \delta^m}(t-1) \tag{41}$$

Рекуррентное соотношение (41) является опорной точкой для вычисления $\overline{M}^q_{ij}(t)$. Получим аналогичное уравнение для $M^q_{ij}(t)$.

Аналогично (40)

$$M_{ij}^q(t) = \sum_{l=1}^{n_q} P_{ql}(t),$$

где $P_{ql}(t)$ — вклад деревьев из $D_{ql}^t$ в матожидание $M_{ij}^q(t)$. Положим $P_{ql}(t) = P_{ql}^{(1)}(t) + P_{ql}^{(2)}(t)$, аналогично тому, как это сделано для $\overline{P}_{ql}(t)$. При этом

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

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

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

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,

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

где $z_m$ — $m$-й нетерминал первого яруса.

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

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

В результате, для $M_{ij}^q(t)$ получаем

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

Из леммы (2) следуют выражения для $Q_X(t)$ и $R_X(t)$

$$\begin{aligned} Q_X(t) &= \prod_{i=1}^{k} (1 - Q_i(t))^{x_i} = 1 - \sum_{i=1}^{k} x_i Q_i(t) + \Theta \left( \sum_{1 \leqslant i,j \leqslant k} x_i x_j Q_i(t) Q_j(t) \right) \\ R_X(t) &= Q_X(t) - Q_X(t-1) = \sum_{i=1}^{k} x_i P_i(t) + \Theta \left( \sum_{1 \leqslant i,j \leqslant k} x_i x_j Q_i(t) Q_j(t) \right) \end{aligned} \tag{42}$$

Теперь приступим к вычислению $\overline{M}'^{q}_{ij}(t)$ и $M'^{q}_{ij}(t)$.

Пусть вначале $A_q$ и $A_i$ принадлежат классам, находящимся в одной группе $\mathcal{M}_n$. Тогда $\overline{M}'^{(\alpha)}_{ij}(t) = 0$ для всех $A_\alpha \in K \in \mathcal{M}_m$, таких что $m > n$. Для $\overline{M}'^{(n)}_{ij}(t)$ получаем:

$$\overline{M}'^{q}_{ij}(t) = \delta^q_i p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} s^\alpha_{ql} \overline{M}'^{\alpha}_{ij}(t-1) Q_{s_{ql}-\delta^\alpha}(t-1)$$

Подставляя выражение (42) для $Q_X(t)$, получаем

$$\overline{M}'^{q}_{ij}(t) = \delta^q_i p_{ij}(1+o(1)) \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}'^{\alpha}_{ij}(t-1) -$$
$$- \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}'^{\alpha}_{ij}(t-1) \sum_{\beta: A_\beta \in K \in \mathcal{M}_n} (s^\beta_{ql} - \delta^\alpha_\beta) Q_\beta(t-1)(1+o(1)) \quad (43)$$

Непосредственной проверкой устанавливается, что

$$\begin{aligned} a^q_\alpha &= \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \\ b^q_{\alpha\beta} &= \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} (s^\beta_{ql} - \delta^\alpha_\beta) \end{aligned} \quad (44)$$

Заменяя соответствующие выражения в (43), а также подставляя асимптотику для $Q^{(n)}(t)$, получаем

$$\overline{M}'^{q}_{ij}(t) = \delta^q_i p_{ij}(1+o(1)) + \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} a^q_\alpha \overline{M}'^{\alpha}_{ij}(t-1) -$$
$$- c_n t^{-\left(\frac{1}{2}\right)^{w-n}} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\beta \overline{M}'^{\alpha}_{ij}(t-1)(1+o(1)) \quad (45)$$

Применяя лемму (3) для вектора $(\overline{M}'^{q_1}_{ij}(t), \overline{M}'^{q_2}_{ij}(t), \ldots, \overline{M}'^{q_\gamma}(t))$, где $A_{q_1}, A_{q_2}, \ldots, A_{q_\gamma}$ — нетерминалы классов группы $\mathcal{M}_n$, имеем

$$\overline{M}'^{q}_{ij}(t) = U_q \sum_{l: A_l \in K \in \mathcal{M}_n} V_l \overline{M}'^{l}_{ij}(t) = U_q M^{(n)}_*(t).$$

Домножая (45) на $V^{(n)}$ слева, получаем

$$\delta \overline{M}^{(n)}_*(t) = V_i p_{ij}(1+o(1)) - c_n t^{-\left(\frac{1}{2}\right)^{w-n}} \sum_{q,\alpha,\beta} V_q b^q_{\alpha\beta} U_\alpha U_\beta = V_i p_{ij} - c_n t^{-\left(\frac{1}{2}\right)^{w-n}} B_n$$

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

$$\overline{M}^{(n)}_*(t) = \left(\frac{V_i p_{ij}}{c_n B_n + 1}\right) \cdot t \cdot (1+o(1)), \qquad \text{если} \quad n = w$$
$$\overline{M}^{(n)}_*(t) = \left(\frac{V_i p_{ij}}{c_n B_n}\right) \cdot t^{-\left(\frac{1}{2}\right)^{w-n}} \cdot (1+o(1)), \qquad \text{если} \quad n < w$$

Пусть теперь классы, содержащие $A_q$ и $A_i$, находятся в различных группах. Тогда

$$\overline{M}'^q_{ij}(t) = \delta^q_i p_{ij} Q_{s_{ij}}(t-1) + \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n \cup \mathcal{M}_{n+1}} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}'^\alpha_{ij}(t-1) Q_{s_{ql}-\delta^\alpha}(t-1)$$

Учитывая $Q^{(n+1)}(t) = o(Q^{(n)}(t))$, получаем

$$\overline{M}'^q_{ij}(t) = O(p_{ij}) +$$
$$+ \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}' \alpha_{ij}(t-1) \left(1 - \sum_{\beta: A_\beta \in K \in \mathcal{M}_n} (s^\alpha_{ql} - \delta^\alpha_\beta) Q_\beta(t-1)\right)(1 + o(1)) +$$
$$+ \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_{n+1}} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}'^\alpha_{ij}(t-1)(1 + o(1))$$

Положим $\overline{M}'^\alpha_{ij}(t-1) = \overline{d}'_{n+1} t^{\gamma(n+1)}(1 + o(1))$, что выполняется для $n + 1 = w$. Подставляя это выражение, а также (44), получаем

$$\overline{M}'^q_{ij}(t) = \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} a^q_\alpha \overline{M}'^\alpha_{ij}(t-1) - c_n t^{-\left(\frac{1}{2}\right)^{w-n}} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\beta \overline{M}'^\alpha_{ij}(t-1)(1 + o(1)) +$$
$$+ \overline{d}'_{n+1} t^{\gamma(n+1)} \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_{n+1}} a^q_\alpha U_\alpha (1 + o(1))$$

Домножая на $V_q$, имеем

$$\delta \overline{M}^{(n)}_*(t) = \overline{d}'_{n+1} \left( \sum_{\substack{q: A_q \in K \in \mathcal{M}_n \\ \alpha: A_\alpha \in K \in \mathcal{M}_{n+1}}} V_q a^q_\alpha U_\alpha \right) \cdot t^{\gamma(n+1)} -$$
$$- c_n t^{-\left(\frac{1}{2}\right)^{w-n}} \cdot \left( \sum_{q,\alpha,\beta} V_q b^q_{\alpha\beta} U_\alpha U_\beta \right) \cdot \overline{M}^{(n)}_*(t-1)(1 + o(1)) =$$
$$= \overline{d}'_{n+1} b_{n+1} t^{\gamma(n+1)}(1 + o(1)) - c_n B_n t^{-\left(\frac{1}{2}\right)^{w-n}} \overline{M}^{(n)}_*(t-1)(1 + o(1))$$

Рассматривать случай $n = w$ не имеет смысла, поэтому в выражении $t^{-\left(\frac{1}{2}\right)^{w-n}}$ показатель всегда будет больше $-1$. Учитывая это, и применяя лемму (4), получаем

$$\overline{M}^{(n)}_*(t) = \frac{\overline{d}'_{n+1} b_{n+1} t^{\gamma(n+1) + \left(\frac{1}{2}\right)^{w-n}}}{c_n B_n}$$

Отсюда,

$$\overline{M}^{(n)}_*(t) = \prod_{j=n}^{h-1} \left( \frac{b_{j+1}}{c_j B_j} \right) \cdot \left( \frac{V_i p_{ij}}{c_h B_h + \delta^\alpha_h} \right) \cdot t^{\left(\left(\frac{1}{2}\right)^{\alpha-h-1} - \left(\frac{1}{2}\right)^{\alpha-n}\right)}$$

Подставляя (4) и $\overline{M}'^q_{ij}(t) = \overline{M}^q_{ij}(t)(1 - Q_q(t))$, получаем

$$\overline{M}^{(n)}_*(t) = \frac{U_q}{1 - Q_q(t)} \prod_{j=n}^{h-1} \left( \frac{b_{j+1}}{c_j B_j} \right) \cdot \left( \frac{V_i p_{ij}}{c_h B_h + \delta^\alpha_h} \right) \cdot t^{\left(\left(\frac{1}{2}\right)^{\alpha-h-1} - \left(\frac{1}{2}\right)^{\alpha-n}\right)}$$

Перейдём к вычислению $M^q_{ij}(t)$. Вначале пусть нетерминалы $A_q$ и $A_i$ принадлежат классам из одной группы $\mathcal{M}_n$. Полагая $M'q_{ij}(t) = M^q_{ij}(t)P_q(t)$, получаем

$$M'q_{ij}(t) = O(t^{-1-\left(\frac{1}{2}\right)}) + \sum_{\alpha: Q_\alpha \in K \in \mathcal{M}} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} M'^\alpha_{ij}(t-1) -$$
$$- \sum_{\alpha,\beta: A_\alpha, A_\beta \text{в группе} \mathcal{M}_n} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} (s^\beta_{ql} - \delta^\alpha_\beta) Q_n(t-1) M'^\alpha_{ij}(t-1) +$$
$$+ \sum_{\alpha,\beta} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} (s^\beta_{ql} - \delta^\alpha_\beta) P_n(t-1) \overline{M}'^\alpha_{ij}(t-1)$$

Подставляя выражение (44) для первых и вторых моментов, получаем

$$M'^q_{ij}(t) = \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} q^q_\alpha M'^\alpha_{ij}(t-1) - c_n t^{-\left(\frac{1}{2}\right)^{w-n}} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\beta M'^\alpha_{ij}(t-1)(1+o(1)) +$$
$$+ \tilde{c}_n t^{-1-\left(\frac{1}{2}\right)^{w-n}} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\beta \overline{M}^\alpha_{ij}(t-1)(1+o(1))$$

Подставляя выражение для $\overline{M}'^\alpha_{ij}(t-1)$, имеем

$$M'^q_{ij}(t) = \sum_\alpha a^q_\alpha M'^\alpha_{ij}(t-1) - c_n t^{-\left(\frac{1}{2}\right)} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\beta M'^\alpha_{ij}(t-1)(1+o(1)) +$$
$$+ \tilde{c}_n d_n t^{-1} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\alpha U_\beta (1+o(1)) \quad (46)$$

Применяя лемму (3), получаем

$$M'q_{ij}(t) = U_q M^{(n)}_*(t)(1+o(1))$$
$$M^{(n)}_*(t) = \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} V_\alpha M'^\alpha_{ij}(t)$$

Домножая (46) на $V^{(n)}$, получаем

$$\delta V^{(n)}_*(t) = \tilde{c}_n d_n B_n t^{-1} - c_n t^{-\left(\frac{1}{2}\right)^{w-n}} B_n V^{(n)}_*(t)(t-1)(1+o(1))$$

Применяя лемму (4), получаем в результате

$$M^{(n)}_*(t) = \begin{cases} \tilde{c}_n \overline{d}'_n B_n (1+o(1)), & \text{при} \quad \alpha = n \\ \dfrac{\tilde{c}_n \overline{d}'_n}{c_n} t^{-1-\left(\frac{1}{2}\right)^{w-n}} (1+o(1)), & \text{при} \quad \alpha > n \end{cases}$$

Пусть теперь $A_q$ и $A_i$ находятся в классах, принадлежащих разным группам $\mathcal{M}_n$ и $\mathcal{M}_h$ $(h > m)$. Тогда

$$M'^q_{ij}(t) = \delta^q_i p_{ij} R_{s_{ij}}(t-1) + \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n \cup \mathcal{M}_{n+1}} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} [Q_{s_{ql}}(t-1) M'^\alpha_{ij}(t-1) +$$
$$+ (1 - Q_\alpha(t-1)) R_{s_{ql} - \delta^\alpha}(t-1) \overline{M}^\alpha_{ij}(t-1)]$$

откуда

$$M'^q_{ij}(t) = \sum_{\alpha: A_\alpha \in K \in \mathcal{M}_n} a^q_\alpha M'^\alpha_{ij}(t-1) - \sum_{\alpha,\beta \text{ в группе } \mathcal{M}_n} b^q_{\alpha\beta} Q_\beta(t-1) M'^\alpha_{ij}(t-1)(1+o(1)) + \\ + \sum_{\alpha,\beta \text{ в группе } \mathcal{M}_n} b^q_{\alpha\beta} P_\beta(t-1) \overline{M}^\alpha_{ij}(t-1)(1+o(1))$$

Домножая на $V^{(n)}$, имеем

$$\delta M^{(n)}_*(t) = \tilde{c}_n \overline{d}_n B_n t^{-1-\left(\frac{1}{2}\right)^{w-n} + \left(\frac{1}{2}\right)^{w-h}\left(2-\left(\frac{1}{2}\right)^{h-n}\right)}(1+o(1)) - \\ - c_n B_n t^{-\left(\frac{1}{2}\right)^{w-n}} M^{(n)}_*(t-1)(1+o(1))$$

Поскольку $n < w$, показатель степени в выражении $t^{-\left(\frac{1}{2}\right)}$ всегда больше $-1$, и по лемме (4) получаем

$$M^{(n)}_*(t) = \frac{\tilde{c}_n \overline{d}_n}{c_n} t^{-1+\left(\frac{1}{2}\right)^{w-h}\left(2-\left(\frac{1}{2}\right)^{h-n}\right)}$$

Объединяя результаты для $n < w$ и $n = w$, получаем

$$M^{(n)}_*(t) = \frac{\tilde{c}_n \overline{d}_n B_n}{\delta^n_w (c_n B_n - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{w-h}\left(2-\left(\frac{1}{2}\right)^{h-n}\right)}(1+o(1))$$

Откуда

$$M^{(n)}_*(t) = \frac{U_q}{P_q(t)} \frac{\tilde{c}_n \overline{d}_n B_n}{\delta^n_w (c_n B_n - 1) + 1} \cdot t^{-1+\left(\frac{1}{2}\right)^{w-h}\left(2-\left(\frac{1}{2}\right)^{h-n}\right)}(1+o(1))$$

# 5 Энтропия

Пусть $L^t$ — множество слов языка $L_G$, порождаемых деревьями вывода из $D^t$. Будем рассматривать грамматики с однозначным выводом.

По определению, энтропия языка $L^t$ есть

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

где $p_t(\alpha) = p(\alpha : \alpha \in L^t) = p(\alpha)/p(L^t)$. Используя это выражение для $p_t(\alpha)$, получаем

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

Выразим вероятность слова $\alpha$ через вероятности правил вывода $r_{ij}$. Поскольку рассматривается грамматика с однозначным выводом, каждому слову $\alpha$ из $L^t$ соответствует единственное дерево $d(\alpha)$ из $D^t$ и единственный левый вывод $\omega_l(\alpha) = (r_{i_1,j_1}, r_{i_2,j_2}, \dots, r_{i_s,j_s})$. Получаем

$$p(\alpha) = p(r_{i_1,j_1}) \cdot \ldots \cdot p(r_{i_s,j_s}) = \prod_{i=1}^{k} \prod_{j=1}^{n_i} p_{ij}^{q_{ij}(\alpha)},$$

где $q_{ij}(\alpha)$ — число применений правила $r_{ij}$ при выводе слова $\alpha$ (учитывая единственность дерева вывода, это число определяется единственным образом). Тогда

$$\sum_{\alpha \in L^t} p(\alpha) \log p(\alpha) = \sum_{\alpha \in L^t} p(\alpha) \sum_{i=1}^{k} \sum_{j=1}^{n_i} q_{ij}(\alpha) \log p_{ij} = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \log p_{ij} \sum_{\alpha \in L^t} q_{ij}(\alpha) p(\alpha)$$

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

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

Отсюда

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

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

$$H(L^t) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} H(R_i) d_i t^2 (1 + o(1)),$$

где $H(R_i) = -\sum_{j=1}^{n_i} p_{ij} \log p_{ij}$ — энтропия множества $R_i$ правил вывода. Асимптотика $t^2$ задаётся величиной $M_{ij}^q(t)$ для последнего критического класса и классов, следующих за ними.

Сформулируем теорему:

**Теорема 2** *Энтропия языка $L^t$, состоящего из слов, порождаемых в разложимой стохастической КС-грамматике вида «цепочки» с однозначным выводом деревьями высоты $t$, выражается формулой*

$$H(L^t) \sum_{i \in I} \sum_{j=1}^{n_i} d_i H(R_i) \cdot t^2,$$

*где $d_i > 0$, $H(R_i) = \sum_{j=1}^{n_i} p_{ij} \log p_{ij}$ — энтропия множества $R_i$ правил вывода с нетерминалов $A_i$ в левой части, и $I$ — множество индексов нетерминалов, содержащихся в последнем критическом классе, а также классах, следующих за ним.*

# 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] **Фу К.** Структурные методы в распознавании образов. М.: Мир, 1977

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

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

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

[7] **Жильцова Л. П.** О матрице первых моментов разложимой стохастической КС-грамматики. УЧЁНЫЕ ЗАПИСКИ КАЗАНСКОГО ГОСУДАРСТВЕННОГО УНИВЕРСИТЕТА, Том 151, кн. 2, 2009

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

[9] **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.

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