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

Факультет вычислительной математики и кибернетики

Кафедра математической логики и высшей алгебры

МАГИСТЕРСКАЯ ДИССЕРТАЦИЯ

Тема:

«Исследование свойств стохастических КС-грамматик вида "цепочки"»

Допущена к защите:
Заведующий кафедрой МЛиВА
д.ф.-м.н., проф.
_____ Шевченко В.Н.
Подпись
«_____»_____2013г.

Выполнил: студент группы 86М1
_____ Мартынов И.М.
Подпись

Научный руководитель:
Д.ф.-м.н., проф.
_____ Жильцова Л.П.
Подпись

Нижний Новгород
2013

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 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 $R = \cup_{i=1}^{n} R_i$, где n — мощность алфавита V_N и $R_i = \{r_{i1}, \ldots, r_{in_i}\}$. 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 α 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 α 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 α строится следующим образом. Корень дерева помечается аксиомой s. Далее последовательно рассматриваются правила левого вывода слова α. Пусть на очередном шаге рассматривается правило $A_i \xrightarrow{p_{ij}} b_{i_1} b_{i_2} \ldots b_{i_m}$, где $b_{i_l} \in (V_N \cup V_T)$ $(l = 1, \ldots, m)$. Тогда из самой левой вершины-листа дерева, помеченной символом A_i, 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 λ, 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 α. Грамматика G называется *согласованной*, если

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

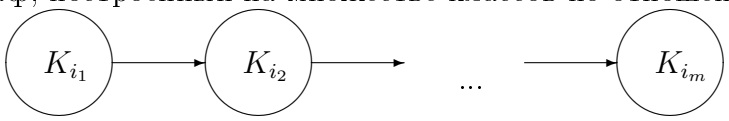
Согласованная грамматика G задаёт распределение вероятностей P на множестве L_G, при этом $p(\alpha)$ — вероятность слова α. Пара $\mathcal{L} = (L_G, P)$ называется *стохастическим КС-языком*. В дальнейшем будем всюду предполагать, что рассматривается согласованная грамматика.

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

$$K_1, K_2, \ldots, K_m. \tag{4}$$

Множества номеров нетерминалов, входящих в класс K_j обозначим через I_j. При $m \geqslant 2$ грамматика называется *разложимой*.

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:



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 $\alpha_1, \alpha_2 \in (V_N \cup V_T)^*$. Не уменьшая общности, будем считать, что грамматика не имеет особых классов.

3 Производящие функции. Моменты

Определим многомерные производящие функции [3]:

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

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

Для краткости будем обозначать

$$\begin{aligned} &\mathbf{s} = (s_1, s_2, \ldots s_n)^T \\ &F_i(\mathbf{s}) = F_i(s_1, s_2, \ldots, s_n) \\ &\mathbf{F}(\mathbf{s}) = (F_1(\mathbf{s}), F_2(\mathbf{s}), \ldots, F_n(\mathbf{s}))^T \end{aligned}$$

Производящую функцию $F_i(\mathbf{s})$ можно интерпретировать следующим образом. Выберем нетерминал A_i в качестве аксиомы грамматики. Затем применим к нему случайным образом какое-нибудь правило из множетсва R_i согласно распределению вероятностей на этом множестве. В полученной строке подсчитаем количество нетерминалов каждого вида и запишем в виде характеристического вектора $L = (l_1, l_2, \ldots, l_n)$, где l_j — количество нетерминалов A_j в полученной строке. Каждому характеристическому вектору, который мы можем таким образом получить, функция $F_i(\mathbf{s})$ ставит в соответствие его вероятность p_{ij}.

Степень производящей функции $(F_i(\mathbf{s}))^k$ соответствует ситуации, когда мы строим одновременно k деревьев вывода из нетерминала A_i, в каждом дереве применяя случайным образом одно из правил вывода, и затем подсчитываем количество нетерминалов разных типов в листьях всех деревьев. В самом деле,

$$(F_i(\mathbf{s}))^k = \left(\sum_j p_{ij} s_1^{l_1^{ij}} \ldots s_n^{l_n^{ij}} \right)^k = \sum p_{ij_1} p_{ij_2} \ldots p_{ij_k} s_1^{l_1^{ij_1} + \ldots + l_1^{ij_k}} \ldots s_n^{l_n^{ij_1} + \ldots l_n^{ij_k}} \qquad (5)$$

Каждое слагаемое с коэффициентом $p_{ij_1} \ldots p_{ij_k}$ соответствует случаю, когда к дереву вывода с индексом l было применено правило r_{ij_l} $(1 \leqslant l \leqslant k)$. При этом в каждой

компоненте характеристического вектора суммируется количество нетерминалов соответствующего типа в каждом из деревьев.

Аналогично, выражение $F_1^{k_1}(\mathbf{s}) \cdot \ldots \cdot F_n^{k_n}(\mathbf{s})$ соответствует случаю, когда одновременно строятся деревья вывода из нетерминалов разных типов, причём деревьев с корнем A_l имеется ровно k_l штук.

Величина

$$\left.\frac{\partial^n F_i(\mathbf{s})}{\partial s_{k_1} \partial s_{k_2} \cdots \partial s_{k_n}}\right|_{\mathbf{s}=\mathbf{1}}$$

где $\mathbf{1} = (1, 1, \ldots, 1)^T$, называется n-м моментом. Поскольку $F_i(\mathbf{s})$ является полиномом, порядок дифференцирования не имеет значения.

Первые и вторые моменты будем обозначать следующим образом.

$$\begin{aligned} a_j^i &= \left.\frac{\partial F_i(s_1, s_2, \ldots, s_k)}{\partial s_j}\right|_{s_1=\ldots=s_k=1} \\ b_{jl}^i &= \left.\frac{\partial^2 F_i(s_1, s_2, \ldots, s_k)}{\partial s_l \partial s_j}\right|_{s_1=\ldots=s_k=1} \end{aligned} \tag{6}$$

Определим многомерные производящие функции $F(t, \mathbf{s})$, где $t \geqslant 1$, следующим образом.

$$F_i(t, \mathbf{s}) = \begin{cases} F_i(\mathbf{s}), & t = 1 \\ F_i(t-1, \mathbf{F}(\mathbf{s})), & t > 1 \end{cases}$$

Функцию $F_i(t, \mathbf{s})$ можно интерпретировать следующим образом. Выберем в качестве аксиомы грамматики нетерминал A_i и будем строить дерево вывода. На каждом шаге в уже построенном дереве выберем какой-нибудь нетерминал A_k, находящийся на ярусе выше t, применим к нему какое-нибудь правило r_{kj} из R_k в соответствии с распределением вероятностей и добавим символы β_{kj} в качестве потомков A_k. Будем продолжать этот процесс до тех пор, пока в дереве вывода не останется нетерминалов на ярусах выше t. Количество нетерминалов различного типа в полученном слове вновь обозначим характеристическим вектором $L = (l_1, l_2, \ldots, l_n)$. Тогда функция $F(t, \mathbf{s})$ ставит в соответствие каждому из возможных векторов L его вероятность.

Это можно показать индукцией по t. При $t = 1$ это верно в силу определения $F_i(\mathbf{s})$. Пусть это верно для $F_i(t-1, \mathbf{s}) = \sum_k p_k s_1^{l_1} s_2^{l_2} \ldots s_n^{l_n}$, где сумма берётся по всем возможным характеристическим векторам $(l_1, \ldots l_n)$, и p_k — вероятность соответствующего вектора. При переходе от $F_i(t-1, \mathbf{s})$ к $F_i(t, \mathbf{s})$ каждое произведение вида $p_k s_1^{l_1} \ldots s_n^{l_n}$ приобретает вид $p_k \cdot F_1^{l_1}(\mathbf{s}) \ldots F_n^{l_n}(\mathbf{s})$. Принимая во внимание представление (5), получаем сумму, каждый компонент которой соответствует возможному характеристическому вектору.

Матрица A, составленная из первых моментов a_j^i, называется *матрицей первых моментов*. Для разложимой грамматики она имеет следующий блочно-ленточный вид.

$$A = \begin{pmatrix} A_{11} & A_{12} & 0 & \cdots & 0 & 0 \\ 0 & A_{22} & A_{23} & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & A_{m-1,m-1} & A_{m-1,m} \\ 0 & 0 & 0 & \cdots & 0 & A_{m,m} \end{pmatrix}. \tag{7}$$

Блок A_{ii} соответствует классу K_i и является неразложимой неотрицательной матрицей. По определению (6), матрицы $A_{11}, A_{22}, \ldots, A_{m,m}$ неотрицательны. Они также неразложимы, так как любой нетерминал может быть с ненулевой вероятностью выведен из любого нетерминала того же класса. Обозначим перронов корень [6] матрицы A_{ii} через r_i. 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*.

Обозначим через J множество индексов i, таких что классы K_i имеют перронов корень $r_i = 1$. Будем также обозначать через $\overline{J}$ дополнение к J.

Обозначим s_{lh} (при $l \leqslant h$) — число критических классов среди подцепочки $K_l, K_{l+1}, \ldots, K_h$. Разобьём последовательность классов $K_1, K_2, \ldots, K_m$ на группы $\mathcal{M}_1, \mathcal{M}_2, \ldots, \mathcal{M}_w$, где $w = s_{1m}$. Класс K_l отнесём к группе $\mathcal{M}_w$ при $s_{lw} <= 1$ и к группе M_{w-j+1} при $s_{lw} = j$ $(j = 2, \ldots, w)$.

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_{lh} находится на пересечении строк для классов из группы $\mathcal{M}_l$ и столбцов для классов из группы $\mathcal{M}_h$. Матрицы B_{lh}, в свою очередь, имеют вид

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

где C_{22} — блок, стоящий на пересечении строк для l-го критического класса и столбцов для h-го критического класса. При $l = h$ этот блок является неразложимой матрицей. Блоки C_{11} и C_{33} стоят на пересечении строк и столбцов, соответствующих докритическим классам. При $l, h < w$ блок B_{lh} имеет вид

$$B_{lh} = \begin{pmatrix} C_{11} & C_{12} \\ 0 & C_{22} \end{pmatrix}.$$

Блок, находящийся на позиции блока B_{lh} в матрице A^t, обозначим $B_{lh}^{(t)}$.

В [7] доказана следующая теорема.

Теорема 1 *При* $t \to \infty$

$$B_{lh}^{(t)} = \begin{pmatrix} 0 & b \cdot U_I^{(l)} V_{II}^{(h)} & b \cdot U_I^{(l)} V_{III}^{(h)} \\ 0 & b \cdot U_{II}^{(l)} V_{II}^{(h)} & b \cdot U_{II}^{(l)} V_{III}^{(h)} \\ 0 & 0 & 0 \end{pmatrix} = b \cdot U^{(l)} V^{(h)} t^{s_{lh}-1} r^t \cdot (1 + o(1)),$$

где $U^{(q)}$ *и* $V^{(q)}$ *— правый и левый собственные векторы матрицы* B_{qq}, *и* $b = V^{(l)} B_{lh} U^{(h)}$.

4 Вероятности продолжения

Вероятностью продолжения $Q_i(t)$ будем называть функцию

$$Q_i(t) = 1 - F_i(t, \mathbf{0})$$

По смыслу функции $F_i(t, \mathbf{s})$ вероятность продолжения $Q_i(t)$ есть вероятность того, что при построении дерева вывода из нетерминала A_i случайным образом это дерево будет иметь высоту более t. Будем обозначать $\mathbf{Q}(t) = (Q_1(t), Q_2(t), \dots Q_n(t))^T$.

В силу согласованности грамматики $Q_i(t) \to 0$ при $t \to \infty$. В самом деле, по смыслу $F_i(t, \mathbf{s})$

$$F_i(t, \mathbf{0}) = \sum_{d \in D^{\leqslant t}} p(d) \xrightarrow[t \to \infty]{} 1$$

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

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

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

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

Переходя от $1 - F_i(t, \mathbf{s})$ к $Q_i(t)$, получаем

$$Q_i(t+1) = \sum_{i=1}^{k} a_j^i Q_i(t) - \frac{1}{2} \sum_{1 \leqslant j,l \leqslant k} b_{jl}^i Q_j(t) Q_l(t) + O\left(\|\mathbf{Q}_j(t)\|^3\right) \tag{10}$$

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

$$Q(t) = \begin{pmatrix} Q^{(1)}(t) \\ Q^{(2)}(t) \\ \vdots \\ Q^{(m)}(t) \end{pmatrix}, \quad Q^{(j)}(t) \in \mathbb{R}^{k_j}, \tag{11}$$

где $k_j = |K_j|$. Обозначим через I_n иножество индексов нетерминалов, входящих в класс K_n. Используя это обозначение, уравнение (10) можно записать в виде

$$Q_i(t+1) = \sum_{j \in I_n} a_j^i Q_j(t) + \sum_{i \in I_{n+1}} a_j^i Q_j(t) \cdot (1 + o(1)) \qquad (i \in I_n, n < m) \tag{12}$$

$$Q_i(t+1) = \sum_{j \in I_m} a_j^i Q_j(t) \cdot (1 + o(1)) \qquad (i \in I_m) \tag{13}$$

или, используя вид (7) матрицы первых моментов,

$$Q^{(n)}(t+1) = A_{n,n}Q^{(n)}(t) + A_{n,n+1}Q^{(n+1)}(t)(1+o(1)) \tag{14}$$

Для всего вектора $Q(t)$ верно равенство

$$Q(t+1) = (A - A(t))Q(t), \tag{15}$$

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 (15) получаем

$$Q(t+1) = (A - A(t)) \cdot \ldots \cdot (A - A(\tau))Q(\tau) \tag{16}$$

Обозначим

$$\begin{aligned} A^*(t) &= (A - A(t)) \cdot (A - A(t-1)) \cdot \ldots \cdot (A - A(\tau+1)) \\ \tilde{A}^*_{ij} &= \frac{A^*_{ij}(t)}{t^{s_{ij}}} \\ \tilde{A}_{ij} &= \frac{A^{(t)}_{ij}}{t^{s_{ij}}}, \end{aligned} \tag{17}$$

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

$$\begin{aligned} &\left|\tilde{A}_{ij}(l(\varepsilon_1)) - \tilde{a}_{ij}U^{(i)}V^{(j)}\right| < \varepsilon_1 E \\ &\forall t \geqslant n(\varepsilon_2) \quad A(t) < \varepsilon_2 A \end{aligned}. \tag{18}$$

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

$$(1-\varepsilon_2)^l A^l x^{(\tau)} \leqslant A^*(t)x^{(\tau)} \leqslant A^l x^{(\tau)}, \tag{19}$$

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

$$(1-\varepsilon_2)^l A^l_{ij} x^{(\tau)}_j \leqslant A^*_{ij}(l)x^{(\tau)}_j \leqslant A^{(l)}_{ij} x^{(\tau)}_j, \tag{20}$$

откуда

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

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

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

Используя эту оценку, можем записать

$$\left|\tilde{A}^*_{ij}(t) - \tilde{a}_{ij}U^{(i)}V^{(j)}x_j^{(\tau)}\right| \leqslant \left|\left(\tilde{A}^*_{ij}(t) - \tilde{A}_{ij}(t)\right)x^{(\tau)}\right| + \\ + \left|\left(\tilde{A}_{ij}(l) - \tilde{a}_{ij}U^{(i)}V^{(j)}\right)x_j^{(\tau)}\right| \leqslant (1-(1-\varepsilon_2)^l)\tilde{A}_{ij}(l)x_j^{(\tau)} + \varepsilon_1 x_j^{(\tau)} \leqslant \\ \leqslant (1-(1-\varepsilon_2)^l)hk_jx_j^{(\tau)} + \varepsilon_1 x_j^{(\tau)} \leqslant \left((1-1-\varepsilon_2)^l\right)hk_j + \varepsilon_1\big)\,x_j^*(\tau), \quad (23)$$

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

Устремляем ε_2 к нулю, затем ε_1 к нулю таким образом, чтобы выполнялось условие

$$l(\varepsilon_1)\log(1-\varepsilon_2) \to -\infty \quad (24)$$

Тогда

$$\left|\tilde{A}^*_{ij}(t) - \tilde{a}_{ij}U^{(i)}V^{(j)}x_j^{(\tau)}\right| \leqslant \varepsilon x_j^*(\tau) \quad (\varepsilon \to 0). \quad (25)$$

Домножая слева на $V^{(i)}$, имеем

$$\left|V^{(i)}\tilde{A}^*_{ij}(t)x_j^{(\tau)} - \tilde{a}_{ij}V^{(j)}x_j^{(\tau)}\right| \leqslant \varepsilon k_i \max\left\{(V^{(i)}\right\}x_j^*(\tau) \leqslant \varepsilon^* V^{(j)}x_j^{(\tau)}. \quad (26)$$

Отсюда,

$$\left|\frac{\tilde{A}^*_{ij}(t)x_j^{(\tau)}}{V^{(i)}\tilde{A}^*_{ij}(t)x_j^{(\tau)}} - \frac{\tilde{a}_{ij}U^{(i)}V^{(i)}x_j^{(\tau)}}{\tilde{a}_{ij}V^{(j)}x_j^{(\tau)}}\right| = \left|\frac{\tilde{A}^*_{ij}(t)x_j^{(\tau)}}{V^{(i)}\tilde{A}^*_{ij}(t)x_j^{(\tau)}} - U^{(i)}\right| \to 0 \quad (27)$$

или же

$$\left|\frac{A^*_{ij}(t)x_j^{(\tau)}}{V^{(i)}A^*_{ij}(t)x_j^{(\tau)}} - U^{(i)}\right| \to 0, \quad (28)$$

откуда

$$(A-A(t))\cdot\ldots\cdot(A-A(\tau))\cdot x_j = U^{(i)}V^{(i)}(A-A(t))\cdot\ldots\cdot(A-A(\tau))\cdot x_j\cdot(1+o(1)) \quad (29)$$

Ввиду полученного выражения и (16) 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 (10) скалярно на $V^{(n)}$. Заметим, что

$$Q^{(n)}(t) = U^{(n)}Q^{(n)}_*(t)(1+o(1)). \quad (30)$$

$$Q^{(n)}_*(t+1) = Q(n)_*(t) + V^{(n)}B_{n,n+1}U^{(n+1)}Q^{(n+1)}_*(t) - \\ - \frac{1}{2}\sum_{1\leqslant i,j,l\leqslant k_n} V_i^{(n)}b^i_{jl}(n)U_j^{(n)}U_l^{(n)}\left(Q^{(n)}_*(t)\right)^2(1+o(1)). \quad (31)$$

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

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

Тогда уравнение (31) перепишется как

$$\delta Q_*^{(n)}(t) = b_n Q_*^{(n+1)}(t) - \frac{1}{2} B_n (Q_*^{(n)}(t))^2 (1 + o(1)) \tag{32}$$

Выражение для $\delta Q_*^{(n)}(t)$ также можно получить из (10), вычитая это уравнение из себя с заменой $t \to t+1$:

$$\delta Q_*^{(n)}(t+1) = \sum_{j=1}^{k_n} a_j^i(n) \delta Q_j^{(n)}(t) + \sum_{j=1}^{k_{n+1}} a_j^i(n) \delta Q_j^{(n+1)}(t) - \\ - \frac{1}{2} \sum_{1 \leqslant j,l \leqslant k_n} b_{jl}^i(n) \left(Q_j^{(n)}(t+1) Q_l^{(n)}(t+1) - Q_j^{(n)}(t) Q_l^{(n)}(t) \right) (1 + o(1))$$

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

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

Для последнего класса

$$Q_*^{(w)}(t) = c_w t^{-1} (1 + o(1)), \tag{34}$$

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 (33), 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

$$\delta z(t) = -\frac{b_n \alpha c_{n+1}}{t} - B_n Q_*^{(n)}(t)(1+o(1)) \tag{35}$$

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 (35) $f(t) = -\frac{b_n \alpha c_{n+1}}{t}(1+o(1))$, $g(t) = B_n Q_*^{(n)}(t)(1+o(1))$, 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 (32), 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 (34), 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 (30), 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

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

5 Математические ожидания числа применений правила в деревьях вывода

Обозначим через $q_{ij}^l(t,\tau)$ и $\overline{q}_{ij}^l(t,\tau)$ случайные величины, равные числу применений правила r_{ij} в дереве вывода, соответственно, из D_l^t и $D_l^{\leqslant t}$, на ярусе τ. Пусть также

$$S_{ij}^l(t) = \sum_{\tau=1}^{t-1} q_{ij}^l(t,\tau)$$
$$\overline{S}_{ij}^l(t) = \sum_{\tau=1}^{t-1} \overline{q}_{ij}^l(t,\tau)$$

и $S_{ij}^l(t)$, $\overline{S}_{ij}^l(t)$ — соответственно число применений правила r_{ij} в дереве из D_l^t, $D_l^{\leqslant t}$. Для удобства записи положим

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

Рассмотрим математические ожидания некоторых из введённых величин. Обозначим

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

Для нахождения величин $\overline{M}_{ij}^l(t)$ и $M_{ij}^l(t)$ 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*

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

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

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

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* 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*

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

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:*

$$(1) \quad x_t = \frac{\alpha t^{\alpha+1}}{\alpha + 1}(1 + o(1)) \qquad \text{при} \quad \beta < -1, \alpha \geqslant 0 \tag{36}$$

$$(2) \quad x_t = \frac{\alpha t^{\alpha+1}}{\alpha + b + 1}(1 + o(1)) \qquad \text{при} \quad \beta = -1, \alpha > -1 \tag{37}$$

$$(3) \quad x_t = \frac{\alpha t^{\alpha-\beta}}{b}(1 + o(1)) \qquad \text{при} \quad -1 < \beta < 0 \tag{38}$$

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

$$\overline{M}_{ij}^q(t) = \sum_{l=1}^{n_q} \overline{P}_{ql}(t). \tag{39}$$

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, $q_{ij}^{(1)}(d) = \delta_i^q \delta_j^l$ (где δ — символ Кронекера). Тогда

$$\overline{P}_{ql}^{(1)}(t) = \delta_i^q \delta_j^l \frac{p_{ij} Q_{s_{ij}}(t-1)}{1 - Q_q(t)},$$

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

Вероятность дерева $p(d)$ при $d \in D_{ql}^{\leqslant t}$ 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

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

где $q_{ij}'^m(d)$ — 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

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

$$= \frac{p_{ql}}{1 - Q_q(t)} \sum_{m=1}^{\overline{s}_{ql}} \sum_{d_j : j \neq m} p_1(d) \dots p_{m-1}(d_{m-1}) p_{m+1}(d_{m+1}) \dots p_{\overline{s}_{ql}}(d_{\overline{s}_{ql}}) q_{ij}'^m(d) =$$

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

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

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

Обозначая

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

имеем

$$\overline{M}'^q_{ij}(t) = \delta^q_i p_{ij} Q_{s_{ij}}(t-1) + \sum_{l=1}^{n_q} p_{ql} \sum_{m=1}^{k} s^m_{ql} \overline{M}'^m_{ij}(t-1) Q_{s_{ql}-\delta^m}(t-1) \tag{40}$$

Рекуррентное соотношение (40) является опорной точкой для вычисления $\overline{M}^q_{ij}(t)$. Получим аналогичное уравнение для $M^q_{ij}(t)$.

Аналогично (39)

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

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

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

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 $t-1$. $P^{(2)}_{ql}(t)$ можно представить в виде

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

где $P^{(2)m}_{ql}(t)$ — вклад деревьев с m-м корнем на первом ярусе в $M^q_{ij}(t)$.

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

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

где z_m — m-й нетерминал первого яруса.

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

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

В результате, для $M^q_{ij}(t)$ получаем

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

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

$$\begin{aligned} Q_X(t) &= \prod_{i=1}^{k}(1-Q_i(t))^{x_i} = 1 - \sum_{i=1}^{k} x_i Q_i(t) + \Theta\left(\sum_{1\leqslant i,j\leqslant k} x_i x_j Q_i(t) Q_j(t)\right) \\ R_X(t) &= Q_X(t) - Q_X(t-1) = \sum_{i=1}^{k} x_i P_i(t) + \Theta\left(\sum_{1\leqslant i,j\leqslant k} x_i x_j Q_i(t) Q_j(t)\right) \end{aligned} \tag{41}$$

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 (41) для $Q_X(t)$, получаем

$$\begin{aligned} \overline{M}'^{q}_{ij}(t) = \delta^q_i p_{ij}(1+o(1)) &\sum_{\alpha:A_\alpha\in K\in\mathcal{M}_n} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}'^{\alpha}_{ij}(t-1) - \\ - \sum_{\alpha:A_\alpha\in K\in\mathcal{M}_n} \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \overline{M}'^{\alpha}_{ij}(t-1) &\sum_{\beta:A_\beta\in K\in\mathcal{M}_n} (s^\beta_{ql} - \delta^\alpha_\beta) Q_\beta(t-1)(1+o(1)) \end{aligned} \tag{42}$$

Непосредственной проверкой устанавливается, что

$$\begin{aligned} a^q_\alpha &= \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} \\ b^q_{\alpha\beta} &= \sum_{l=1}^{n_q} p_{ql} s^\alpha_{ql} (s^\beta_{ql} - \delta^\alpha_\beta) \end{aligned} \tag{43}$$

Заменяя соответствующие выражения в (42), а также подставляя асимптотику для $Q^{(n)}(t)$, получаем

$$\begin{aligned} \overline{M}'^{q}_{ij}(t) = \delta^q_i p_{ij}(1+o(1)) + \sum_{\alpha:A_\alpha\in K\in\mathcal{M}_n} & a^q_\alpha \overline{M}'^{\alpha}_{ij}(t-1) - \\ & - c_n t^{-\left(\frac{1}{2}\right)^{w-n}} \sum_{\alpha,\beta} b^q_{\alpha\beta} U_\beta \overline{M}'^{\alpha}_{ij}(t-1)(1+o(1)) \end{aligned} \tag{44}$$

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 (44) 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

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

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

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

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 (43), получаем

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

Домножая на V_q, имеем

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

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,

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

Подставляя (5) и $\overline{M}_{ij}'^q(t) = \overline{M}_{ij}^q(t)(1-Q_q(t))$, получаем

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

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

$$\begin{aligned}M'q_{ij}(t) = O(t^{-1-\left(\frac{1}{2}\right)}) + \sum_{\alpha:Q_\alpha\in K\in\mathcal{M}}\sum_{l=1}^{n_q} p_{ql}s_{ql}^\alpha M_{ij}'^\alpha(t-1) - \\ - \sum_{\alpha,\beta:A_\alpha,A_\beta\text{ в группе}\mathcal{M}_n}\sum_{l=1}^{n_q} p_{ql}s_{ql}^\alpha(s_{ql}^\beta-\delta_\beta^\alpha)Q_n(t-1)M_{ij}'^\alpha(t-1) + \\ + \sum_{\alpha,\beta}\sum_{l=1}^{n_q} p_{ql}s_{ql}^\alpha(s_{ql}^\beta-\delta_\beta^\alpha)P_n(t-1)\overline{M}_{ij}'^\alpha(t-1)\end{aligned}$$

Подставляя выражение (43) для первых и вторых моментов, получаем

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

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

$$\begin{aligned}M_{ij}'^q(t) = \sum_\alpha a_\alpha^q M_{ij}'^\alpha(t-1) - c_n t^{-\left(\frac{1}{2}\right)}\sum_{\alpha,\beta} b_{\alpha\beta}^q U_\beta M_{ij}'^\alpha(t-1)(1+o(1)) + \\ + \tilde{c}_n d_n t^{-1}\sum_{\alpha,\beta} b_{\alpha\beta}^q U_\alpha U_\beta(1+o(1)) \quad (45)\end{aligned}$$

Применяя лемму (3), получаем

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

Домножая (45) 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

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

Теорема 2 *Пусть матрица мервых моментов грамматики G имеет вид* (7)*, и r_{ij} — правило вывода, для которого $A_i \in K_l$. Тогда при $t \to \infty$*

$$M_{ij}(t) \sim d_i \cdot p_{ij} \cdot t^{\left(\frac{1}{2}\right)^{q^*_l - 1}},$$

где $q^*_l = q_l - 1$ при $l \in J$, и $q^*_l = q_l$ при $l \notin J$, $d_i > 0$ – некоторая константа, и p_{ij} — вероятность правила r_{ij}.

6 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 α через вероятности правил вывода r_{ij}. Поскольку рассматривается грамматика с однозначным выводом, каждому слову α из L^t соответствует единственное дерево $d(\alpha)$ из D^t и единственный левый вывод $\omega_l(\alpha) = (r_{i_1,j_1}, r_{i_2,j_2}, \ldots, r_{i_s,j_s})$. 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 α (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 3 *Энтропия языка L^t, состоящего из слов, порождаемых в разложимой стохастической КС-грамматике вида «цепочки» с однозначным выводом деревьями высоты t, выражается формулой*

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

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.

7 Стоимость оптимального кодирования

Будем обозначать через L^t множество слов, порождённых деревьями высоты t. Через f^* обозначим кодирование языка L^t, минимизирующее величину

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

где величины $p_t(\alpha)$ задают распределение вероятностей на множестве слов L^t. По определению f^*, для любого кодирования f множества слов L^t справедливо $M_t(f) \geqslant M_t(f^*)$.

В [10] была доказана

Теорема 4 *Если имеется последовательность стохастических языков $\mathcal{L}_t$, такая, что $H(\mathcal{L}_t) \to \infty$ при $t \to \infty$, то $\lim\limits_{t \to \infty} M^*(\mathcal{L}_t)/H(\mathcal{L}_t) = 1$.*

Рассматривая в качестве $\mathcal{L}_t$ язык L^t, при $t \to \infty$ получим

$$M^*(L^t) = H(L^t) \cdot (1 + o(1)). \tag{46}$$

Оценим теперь величину $\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha|$. Обозначим через l_{ij} число терминальных символов в правой части правила r_{ij}. Каждое правило r_{ij}, применённое при выводе слова α, добавляет в это слово l_{ij} нетерминалов, поэтому

$$\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha| = \sum_{i=1}^{n} \sum_{j=1}^{n_i} l_{ij} M_{ij}(t).$$

Подставляя асимптотику (2) для $M_{ij}(t)$, получаем

$$\sum_{\alpha in L^t} p_t(\alpha) \cdot |\alpha| = \sum_{i \in I_l^+} d_i L(R_i) t^2 \cdot (1 + o(1)),$$

где $L(R_i) = \sum_{j=1}^{n_i} l_{ij} p_{ij}$, K_l — критический класс цепочки, наиболее удалённый от её начала, I_l^+ — множество индексов нетерминалов, находящихся в классе K_l и более удалённых от начала цепочки классах, и $d_i > 0$ — некоторая константа.

Учитывая соотношение (46), получаем следующий результат.

Теорема 5 *Пусть матрица первых моментов G имеет вид (7) и её перронов корень $r = 1$. Тогда стоимость любого кодирования f языка L, порождаемого этой грамматикой, удовлетворяет неравенству*

$$C(L, f) \geqslant C^*(L),$$

где

$$C^*(L) = \frac{\sum_{i \in I_l^+} d_i H(R_i)}{\sum_{i \in I_l^+} d_i L(R_i)},$$

где, в свою очередь, $H(R_i) = -\sum_{j=1}^{n_i} p_{ij} \log p_{ij}$, $L(R_i) = \sum_{j=1}^{n_i} l_{ij} p_{ij}$, l_{ij} — число терминальных символов в правой части правила r_{ij}, и $d_i > 0$ — некоторая константа.

8 Заключение

В результате проведённого исследования были изучены основные вероятностные характеристики стохастических КС-грамматик вида «цепочки». Получены асимптотические оценки вероятностей продолжения и вероятностей деревьев фиксированной высоты, оценка математического ожидания числа применений некоторого правила вывода в деревьях фиксированной высоты. Также получены оценки для энтропии множества слов языка, порождённых деревьями фиксированной высоты, и для стоимости оптимального кодирования.

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. Вып. 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. // Диссертация на соискание учёной степени кандидата физико-математических наук. Нижний Новгород, 2006.