Содержание

1 Введение

При передаче и хранении информации часто возникает необходимость её экономного кодирования. Сжатие информации может быть достигнуто путём использования статистических данных об источнике сообщений, таких как частоты появления определённых символов в сообщении. Если, кроме этого, учитывать структурные особенности языка сообщений, можно дополнительно увеличить эффективность сжатия.

К. Шеннон [1] рассмотрел 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] [?]. 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 [9]. В частности, Борисов рассмотрел случай, когда перронов корень матрицы первых моментов грамматики равен единице. По аналогии с теорией ветвящихся процессов такой случай называется критическим.

При построении алгоритма кодирования сообщений контекстно-свободного языка важную роль играют *вероятности продолжения* для деревьев вывода слов в грамматике. В работе рассматривается стохастическая КС-грамматика, задающая распределение вероятностей на множестве слов порождаемого ею языка. Получены оценки для вероятностей продолжения и вероятностей деревьев вывода фиксированной высоты.

2 Основные определения

Стохастической контекстно-свободной грамматикой [3] называется система

$$G = < V_T, V_N, R, s >,$$

где $V_N = \{A_1, A_2, \ldots, A_n\}$ — конечный алфавит нетерминальных символов (нетерминалов), V_T — конечный алфавит терминальных символов, $s \in V_N$

— аксиома, и $R = \cup_{i=1}^n R_i$ — множество правил вывода (продукций), где $R_i = \{r_{i1}, \ldots, r_{in_i}\}$. Каждое правило r_{ij} из R_i имеет вид

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

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

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

Для $\alpha_1, \alpha_2 \in (V_N \cup V_T)^*$ будем говорить, что α_2 непосредственно выводится из α_1, если существуют такие $\gamma_1, \gamma_2 \in (V_N \cup V_T)^*$, что $\alpha_1 = \gamma_1 A_i \gamma_2$, $\alpha_2 = \gamma_1 \beta_{ij} \gamma_2$, и грамматика содержит правило вывода $A_i \xrightarrow{p_{ij}} \beta_{ij}$. Будем говорить, что α_2 выводится из α_1, если существуют такие $\alpha'_1, \alpha'_2, \ldots, \alpha'_k \in (V_N \cup V_T)^*$, что $\alpha_1 = \alpha'_1$, $\alpha_2 = \alpha'_k$, и α'_{i+1} непосредственно выводится из α'_i для $1 \leqslant i \leqslant k-1$. Языком, порождаемым грамматикой, называется множество слов, выводимых из её аксиомы.

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

$$\lim_{n \to \infty} \sum_{\substack{\alpha \in L_G \\ |\alpha| \leqslant n}} p(\alpha) = 1. \tag{2}$$

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

$$K_1, K_2, \ldots, K_m. \tag{3}$$

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,

$$(F_i(\mathbf{s}))^k = \left(\sum_j p_{ij} s_1^{l_1^{ij}} \dots s_n^{l_n^{ij}}\right)^k = \sum p_{ij_1} p_{ij_2} \dots p_{ij_k} s_1^{l_1^{ij_1}+\ldots+l_1^{ij_k}} \dots s_n^{l_n^{ij_1}+\ldots l_n^{ij_k}} \tag{4}$$

Каждое слагаемое с коэффициентом $p_{ij_1} \dots p_{ij_k}$ 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.

$$\begin{aligned} a_j^i &= \left.\frac{\partial F_i(s_1, s_2, \ldots, s_k)}{\partial s_j}\right|_{s_1=\ldots=s_k=1} \\ b_{jl}^i &= \left.\frac{\partial^2 F_i(s_1, s_2, \ldots, s_k)}{\partial s_l \partial s_j}\right|_{s_1=\ldots=s_k=1} \end{aligned} \tag{5}$$

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 $F_i(t-1, \mathbf{s}) = \sum_k p_k s_1^{l_1} s_2^{l_2} \dots s_n^{l_n}$, 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 (4), получаем сумму, каждый компонент которой соответствует возможному характеристическому вектору.

Рассмотрим пару классов $K_i \prec_* K_j$ и все соединяющие эту пару цепочки классов $K_{i_1} \prec K_{i_2} \prec \ldots \prec K_{i_k}$, где $i_1 = i, i_k = j$. Для каждой такой цепочки найдём максимальный перронов корень среди $r_{i_1}, r_{i_2}, \ldots, r_{i_k}$, и будем обозначать $\tilde{r}_{ij}$. Кроме того, найдём такую цепочку, в которой число классов, соответствующих перронову корню $\tilde{r}_{ij}$ максимально. Этот максимум будем обозначать $\tilde{s}_{ij}$.

$$\tilde{r}_{ij} = \max_{\substack{k, i_1, i_2, \ldots, i_k \\ i_1 = i, i_k = j \\ K_{i_1} \prec K_{i_2} \prec \ldots \prec K_{i_k}}} \{r_{i_1}, r_{i_2}, \ldots, r_{i_k}\} \tag{6}$$

$$\tilde{s}_{ij} = \max_{\substack{i_1, i_2, \ldots, i_k \\ i_1 = i, i_2 = j \\ K_{i_1} \prec K_{i_2} \prec \ldots \prec K_{i_k}}} \{\{i_1, i_2, \ldots, i_k\} \cap \{i : r_i = \tilde{r}_{ij}\}\} \tag{7}$$

Кроме этого, будем обозначать

$$\tilde{r}_i = \max_j \{\tilde{r}_{ij} : K_i \prec_* K_j\} \tag{8}$$

$$\tilde{s}_i = \max_j \{\tilde{s}_{ij} : K_i \prec_* K_j\} \tag{9}$$

4 Матрица первых моментов

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

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1,m-1} & A_{1,m} \\ 0 & A_{22} & \cdots & A_{2,m-1} & A_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{m-1,m-1} & A_{m-1,m} \\ 0 & 0 & \cdots & 0 & A_{m,m} \end{pmatrix}. \tag{10}$$

Блок A_{ii} соответствует классу K_i и является неразложимой неотрицательной матрицей. По определению (5), 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 $r \leqslant 1$.

Теорема 1 *Пусть согласованная стохастическая КС-грамматика G содержит m классов $K_1, K_2, \ldots, K_m$ нетерминалов, причём $K_j \not\prec_* K_i$ для любых $i < j$.*

$$A^t = \begin{pmatrix} A_{11}^t & A_{12}^{(t)} & \cdots & A_{1,m}^{(t)} \\ 0 & A_{22}^t & \cdots & A_{2,m}^{(t)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{m-1,m}^{(t)} \\ 0 & 0 & \cdots & A_{m,m}^{(t)} \end{pmatrix}$$

Тогда для любых $1 \leqslant i, j \leqslant m$ при $t \to \infty$ верно, что:

$$\begin{aligned} A_{ij}^{(t)} &\sim U^{(i)} V^{(j)} \cdot t^{\tilde{s}_{ij}-1} \cdot \tilde{r}_{ij}^t, && \text{при } K_i \prec_* K_j \\ A_{ij}^{(t)} &= 0, && \text{при } K_i \not\prec_* K_j \end{aligned}$$

где $U^{(i)}$ — вектор-столбец длины n_i, $V^{(j)}$ — вектор-строка длины j, $\tilde{r}_{ij}$ и $\tilde{s}_{ij}$ определены формулами (6) и (7) соответственно.

Теорема следует из результата Л. П. Жильцовой [7]. Для её доказательства будем использовать ту же технику.

Проведём индукцию по числу m классов грамматики. При $m = 1$ матрица A неразложима, и $A = (A_{11})$ в терминах теоремы 1. Асимптотика A^t для этого случая известна из [5]:

$$A^t \sim U \cdot V \cdot r^t,$$

где r — перронов корень матрицы A, U и V — правый и левый собственные векторы матрицы A, соответствующие корню r, и $V \cdot U = 1$.

Пусть теорема 1 верна для грамматики с $m - 1$ классами. Докажем её для m классов. Матрицу A можно представить в виде

$$A = \begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1,m-1} & A_{1,m} \\ 0 & A_{22} & \cdots & A_{2,m-1} & A_{2,m} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & A_{m-1,m-1} & A_{m-1,m} \\ 0 & 0 & \cdots & 0 & A_{m,m} \end{pmatrix} = \begin{pmatrix} A_{11} & E_1 \\ 0 & D_1 \end{pmatrix} = \begin{pmatrix} D_2 & E_2 \\ 0 & A_{m+1,m+1} \end{pmatrix}.$$

Для блоков D_1 и D_2 выполняются условия теоремы 1, поэтому её утверждение остаётся верным для блоков A_{ij} входящих в D_1 или D_2. Остаётся доказать его для $A_{1,m+1}$.

$$A_{1,m}^{(t)} = \sum_{l=1}^{m-1} \sum_{j=0}^{t-1} A_{1,l}^{(j)} A_{l,m} A_{m,m}^{t-j-1}$$

Очевидно, $K_1 \prec_* K_j$ тогда и только тогда, когда $K_1 \prec_* K_l \prec K_m$ для некоторого $1 \leqslant l \leqslant m - 1$. Если $K_1 \not\prec_* K_j$, такого l не найдётся, и все слагаемые в сумме равны нулю.

5 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 $\mathbf{Q}(t) = (Q_1(t), Q_2(t), \ldots Q_n(t))^T$.

Теорема 2

$$\begin{array}{ll} \begin{array}{l} Q_i(t) \sim c_n \cdot u_i \cdot t^{\tilde{s}_i - 1} \cdot \tilde{r}_i^t \\ P_i(t) \sim \tilde{c}_n \cdot u_i \cdot t^{\tilde{s}_i - 1} \cdot \tilde{r}_i^t \end{array} , & \textit{при } \tilde{r}_i < 1 \\ \begin{array}{l} Q_i(t) \sim \frac{c_n \cdot u_i}{t^{\left(\frac{1}{2}\right)^{\tilde{s}_i - 1}}} \\ P_i(t) \sim \frac{\tilde{c}_n \cdot u_i}{t^{1 + \left(\frac{1}{2}\right)^{\tilde{s}_i - 1}}} \end{array} , & \textit{при } \tilde{r}_i = 1 \end{array}$$

где $\tilde{r}_i$ и $\tilde{s}_i$ определены формулами (8) *и* (9) *соответственно.*

Лемма 1 *Компоненты вектора $Q(t) = (Q_1(t), Q_2(t), \ldots, Q_n(t))$ пропорциональны некоторому вектору u при $t \to \infty$, с точностью до главного члена:*

$$Q_i(t) \sim C \cdot u_i \cdot Q_*(t),$$

где C — некоторая константа, u_i определяется только номером i нетерминала, и $Q_*(t)$ — скалярная функция от t.

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

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

$$Q^{(n)}(t) = U^{(n)} Q_*^{(n)}(t)(1 + o(1)). \tag{11}$$

$$Q_*^{(n)}(t+1) = Q(n)_*(t) + V^{(n)} B_{n,n+1} U^{(n+1)} Q_*^{(n+1)}(t) - \\ - \frac{1}{2} \sum_{1 \leqslant i,j,l \leqslant k_n} V_i^{(n)} b_{jl}^i(n) U_j^{(n)} U_l^{(n)} \left(Q_*^{(n)}(t) \right)^2 (1 + o(1)). \tag{12}$$

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

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

Тогда уравнение (12) перепишется как

$$\delta Q_*^{(n)}(t) = b_n Q_*^{(n+1)}(t) - \frac{1}{2} B_n (Q_*^{(n)}(t))^2 (1 + o(1)) \tag{13}$$

Выражение для $\delta Q_*^{(n)}(t)$ также можно получить из (**??**), 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

$$\delta Q_*^{(n)}(t+1) = \delta Q_*^{(n)}(t) + b_n \delta Q_*^{(n+1)}(t) -$$
$$- \frac{1}{2} B_n \delta Q_*^{(n)}(t) \left(Q_*^{(n)}(t+1) + Q_*^{(n)}(t) \right) (1 + o(1)) \tag{14}$$

Для последнего класса

$$Q_*^{(w)}(t) = c_w t^{-1} (1 + o(1)), \tag{15}$$

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 (14), 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

$$\delta z(t) = -\frac{b_n \alpha c_{n+1}}{t} - B_n Q_*^{(n)}(t)(1+o(1)) \tag{16}$$

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

Лемма 2 *Пусть последовательность $z(t)$ ($t = 1,2,\dots$) 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 (16) $f(t) = -\frac{b_n \alpha c_{n+1}}{t}(1+o(1))$, $g(t) = B_n Q_*^{(n)}(t)(1+o(1))$, замечаем, что для $z(t)$ выполняются все условия леммы (2), 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 (13), 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 (15), 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 (11), 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

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

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] **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.