УДК 519.713

Часть I
Закономерности в деревьях вывода слов и оптимальное кодирование. Критический случай

В данной главе мы рассмотрим разложимую грамматику с матрицей первых моментов вида (**??**), перронов корень r которой равен 1. Этот случай называется критическим по аналогии с теорией ветвящихся процессов. Будем пользоваться теми же обозначениями для матрицы первых моментов, их подматриц и собственных векторов, что и в третьей главе для докритического случая. Будем считать, что собственные вектора матрицы A удовлетворяют условиям и нормировке, указанным в лемме **??**.

Как и в докритическом случае, рассмотрим отдельно случай кратного ($r' = r''$) и некратного ($r' \neq r''$) перронова корня.

I.1 Случай кратного перронова корня

В этом разделе мы рассмотрим случай $r' = r'' = 1$. Сначала выведем асимптотические формулы для вероятностей продолжения $Q_i(t)$, $i = 1, \ldots, k$, и вероятностей $P_i(t)$ деревьев вывода высоты t. Затем с помощью этих результатов найдем асимптотику для математических ожиданий количества применений правил r_{ij} в деревьях вывода. С помощью этих результатов будет найдена асимптотика энтропии множества слов языка с деревьями вывода фиксированной высоты и получена нижняя оценка для стоимости оптимального кодирования языка, порожденного грамматикой. Затем построим схему кодирования, стоимость которого равна полученной нижней оценке (такое кодирование, однако, не является эффективным). В конце раздела покажем, что эффективный алгоритм кодирования, предложенный для докритического случая, и здесь является асимптотически оптимальным.

I.1.1 Вероятности продолжения

Выведем асимптотические формулы для вероятностей продолжения $Q_i(t)$, $i = 1, \ldots, k$. Приведем одно необходимое утверждение из теории ветвящихся процессов [16].

Теорема I.1.1 *Пусть* $F(s) = \sum\limits_{\alpha} P\{\xi = \alpha\} s^{\alpha}$- *вероятностная производящая функция случайного вектора* $\xi = \{\xi_1, \ldots, \xi_n\}$ *со значениями из* N^n. *Пусть* $d > 0$

- целое число. Предположим, что все β-моменты d-го порядка $m_\beta = M\xi^{[\beta]}$ конечны. Тогда имеет место разложение

$$F(s) = \sum_{\bar{\beta}<d} \frac{(-1)^{\bar{\beta}} m_\beta}{\beta!} (1-s_1)^{\beta_1} \cdot \ldots \cdot (1-s_n)^{\beta_n} + R(d,s),$$

где $\bar{\beta} = \beta_1 + \ldots + \beta_n$, а остаточный член имеет вид

$$R(d,s) = (-1)^d \sum_{\bar{\beta}=d} \varepsilon_\beta(s)(1-s_1)^{\beta_1} \cdot \ldots \cdot (1-s_n)^{\beta_n},$$

причем при $0 \le s \le s' \le 1$ верны оценки

$$0 \le \varepsilon_\beta(s) \le \varepsilon_\beta(s') \le \frac{m_\beta}{\beta!}, \ \lim_{s\to 1-0} \varepsilon_\beta(s) = \frac{m_\beta}{\beta!}.$$

Все моменты ветвящегося процесса, соответствующего КС-грамматике, конечны, так как для КС-грамматики производящая функция $F(s_1,\ldots,s_k)$ является полиномом от $s_1,\ldots,s_k$. Применяя теорему I.1.1 к производящей функции $F(s)$, получаем следующее разложение для $1-F_i(s),\ i=1,\ldots,k$, в окрестности 1:

$$1-F_i(s_1,\ldots,s_k) = \sum_{j=1}^{k} a^i_j \cdot (1-s_j) - \frac{1}{2} \cdot \sum_{j,l\le k_1} b^i_{jl} \cdot (1-s_j)(1-s_l) + R^1_i(d,s),\ i=1,\ldots,k_1, \tag{I.1.1}$$

$$1-F_i(s_1,\ldots,s_k) = \sum_{j>k_1} a^i_j \cdot (1-s_j) - \frac{1}{2} \cdot \sum_{j,l>k_1} b^i_{jl} \cdot (1-s_j)(1-s_l) + R^2_i(d,s),\ i=k_1+1,\ldots,k, \tag{I.1.2}$$

где $d=3,4,\ldots$, и $R^1_i(d,s), R^2_i(d,s)$ имеют вид

$$R^1_i(d,s) = -\frac{1}{2} \cdot \sum_{l>k_1 \text{или} j>k_1} b^i_{jl} \cdot (1-s_j)(1-s_l)$$

$$+\sum_{\bar{\beta}=3}^{d} \frac{(-1)^{\bar{\beta}} m^i_\beta}{\beta!}(1-s)^\beta + O\left(\sum_{\bar{\beta}=d+1} m^i_\beta \cdot (1-s)^\beta\right),\ i=1,\ldots,k_1$$

$$R^2_i(d,s) = \sum_{\bar{\beta}=3}^{d} \frac{(-1)^{\bar{\beta}} m^{i+k_1}_\beta}{\beta!}(1-s)^\beta + O\left(\sum_{\bar{\beta}=d+1} m^{i+k_1}_\beta \cdot (1-s)^\beta\right),\ i=1,\ldots k_2.$$

Очевидно, $m^{i+k_1}_\beta = 0$, если хотя бы одно $\beta_j > 0$ при $j \le k_1$. Поэтому выражение для $R^2_i(d,s)$ можно переписать как

$$R^2_i(d,s) = \sum_{2<\bar{\beta}\le d,\ \beta(1:k_1)=0} \frac{(-1)^{\bar{\beta}} m^{i+k_1}_\beta}{\beta!} \cdot (1-s_{k_1+1})^{\beta_{k_1+1}} \ldots (1-s_k)^{\beta_k}$$

$$+O\left(\sum_{\bar{\beta}=d+1,\ \beta(1:k_1)=0} m^{i+k_1}_\beta \cdot (1-s_{k_1+1})^{\beta_{k_1+1}} \ldots (1-s_k)^{\beta_k}\right),\ i=1,\ldots,k_2.$$

Обозначим

$$x_i(t) = 1 - F_i(t,0) = Q_i(t),\ i = 1, \dots k_1,$$
$$y_i(t) = 1 - F_{i+k_1}(t,0) = Q_{i+k_1}(t),\ i = 1, \dots k_2,$$
$$x(t) = (x_1(t), \dots, x_{k_1}(t)),\ y(t) = (y_1(t), \dots, y_{k_2}(t)).$$

Подставляя $s_i = x_i(t)$ при $i = 1, \dots, k_1$, и $s_{i+k_1} = y_i(t)$ при $i = 1, \dots, k_2$ в равенства (I.1.1), (I.1.2), и пользуясь равенством $F(t+1, s) = F(F(t, s))$, получаем соотношения

$$x_i(t+1) = \sum_{j \le k_1} a^i_j x_j(t) + \sum_{j > k_1} a^i_j y_{j-k_1}(t)$$

$$-\frac{1}{2} \cdot \sum_{j,l \le k_1} b^i_{jl} x_j(t) x_l(t) + R^1_i(d, x(t), y(t)),\ i = 1, \dots, k_1, \qquad \text{(I.1.3)}$$

$$y_{i-k_1}(t+1) = \sum_{j > k_1} a^i_j y_{j-k_1}(t)$$

$$-\frac{1}{2} \cdot \sum_{j,l > k_1} b^i_{jl} y_{j-k_1}(t) y_{l-k_1}(t) + R^2_i(d, y(t)),\ i = k_1 + 1, \dots, k, \qquad \text{(I.1.4)}$$

где $d = 3, 4, \dots$. Здесь записано $R^2_i(d, x(t), y(t)) = R^2_i(d, y(t))$, поскольку этот член на самом деле не зависит от $x(t)$.

Из ранее выписанной оценки для $R^1_i(d, s)$ следует, что в выражении для $R^1_i(d, x(t), y(t))$ все члены не превосходят по модулю $O\Big(x_i(t)y_j(t) + y_i(t)y_j(t) + x_i(t)x_j(t)x_l(t)\Big)$.

Поскольку производящие функции $F_i(s)$, $i = k_1 + 1, \dots, k$, являются производящими функциями для грамматики, порожденной нетерминалами класса K_2, то при $i > k_1$ формулы для вероятностей продолжения и вероятности деревьев вывода высоты t имеют вид [10]:

$$y_i(t) = Q_i(t) = \frac{2u''_{i-k_1} \cdot (1 + o(1))}{B_2 t},$$

$$y_i(t) - y_i(t+1) = P_i(t) = \frac{2u''_{i-k_1} \cdot (1 + o(1))}{B_2 t^2}, \qquad \text{(I.1.5)}$$

где $i = k_1 + 1, \dots, k$, и

$$B_2 = \sum_{i,l,m > k_1} v''_{i-k_1} b^i_{lm} u''_{l-k_1} u''_{m-k_1}. \qquad \text{(I.1.6)}$$

Пусть $z(t), t = 1, 2, \dots$ - некоторая последовательность векторов или скаляров. Обозначим через $\delta z(t)$ последовательность, члены которой имеют вид $\delta z(t) = z(t + 1) - z(t)$.

Лемма I.1.1 *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*

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

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

Возьмем некоторое число $C > 2$. Очевидно, можно считать без ограничения общности, что $0 < g(t) < 1/2$ для любого t (взяв элемент последовательности $z(t)$ с достаточно большим номером в качестве первого). Рассмотрим два случая. В первом случае для некоторого t_1 при всех $t_2 > t_1$ выполняется неравенство $|z(t_2)| \geq C \cdot |f(t_2)/g(t_2)|$. Во втором случае предполагаем обратное, т.е. для любого t_1 существует номер $t_2 > t_1$, такой что $|z(t_2)| < C \cdot |f(t_2)/g(t_2)|$. Докажем, что в обоих случаях утверждение леммы выполнено. Сначала рассмотрим первый случай. Заметим, что при $|z(t)| \geq C \cdot |f(t)/g(t)|$ из соотношения $\delta z(t) = f(t) - g(t)z(t)$ вытекают неравенства

$$|z(t+1) - z(t)| = |f(t) - g(t)z(t)| < |f(t)| + |g(t)z(t)| < g(t)|z(t)|(1 + 1/C),$$

$$\begin{aligned} &|z(t+1) - z(t)| = |f(t) - g(t)z(t)| \\ \geq &\,||g(t)z(t)| - |f(t)|| = |g(t)z(t)| - |f(t)| > g(t)|z(t)|(1 - 1/C). \end{aligned}$$

При получении второй оценки было использовано неравенство $|z(t)g(t)| \geq C|f(t)| > |f(t)|$. Из этого же неравенства следует, что $z(t)(z(t+1) - z(t)) < 0$. Таким образом, $z(t+1)$ получается прибавлением к $z(t)$ противоположной по знаку величины $\delta z(t)$. Следовательно, справедливы неравенства

$$|z(t)| \cdot (1 - (C+1)g(t)/C) \leq |z(t+1)| \leq |z(t)| \cdot (1 - (C-1)g(t)/C)\,. \qquad \text{(I.1.7)}$$

Разделив эти неравенства на $|z(t)|$ и перемножив получившиеся оценки по $t = t_1, \ldots, n-1$, получим неравенство

$$|z(t_1)| \cdot \prod_{t=t_1}^{n-1} \bigl(1 - (C+1)g(t)/C\bigr) \leq |z(n)| \leq |z(t_1)| \cdot \prod_{t=t_1}^{n-1} \bigl(1 - (C-1)g(t)/C\bigr),$$

т.е $|z(n)| \to 0$ при $n \to \infty$, поскольку по известному критерию [18] сходимости бесконечного произведения оба произведения сходятся к нулю при $n \to \infty$ ввиду того, что ряд $\sum\limits_{t=1}^{n} g(t)$ расходится и $g(t) > 0$.

Теперь рассмотрим второй случай. Зафиксируем произвольно малое $\varepsilon > 0$. Возьмем номер t_1 такой, что $g(t) < \varepsilon,\ |f(t)/g(t)| < \varepsilon$ при $t > t_1$, и соответствующий ему номер $t_2 > t_1$, такой что $|z(t_2)| < C \cdot |f(t_2)/g(t_2)|$. Тогда для любого $t > t_2$ справедливо одно из двух утверждений:
1) $|z(t)| < C \cdot |f(t)/g(t)| < C\varepsilon$;
2) Существует $t_3,\ t_2 \leq t_3 < t$, такое что $|z(n)| > C \cdot |f(n)/g(n)|$ при $n = t, t-1, \ldots, t_3+1$, и $|z(t_3)| < C \cdot |f(t_3)/g(t_3)| < C\varepsilon$.
Поскольку при $|z(t)| > C \cdot |f(t)/g(t)|$ из неравенства (I.1.7) следует, что $|z(t+1)| < |z(t)|$, то во втором случае $|z(t)| < |z(t_3+1)| < |z(t_3)| + |f(t_3)| < (C+1)\varepsilon$. Устремляя $\varepsilon \to 0$, получаем, что $|z(t)| \to 0$ при $t \to \infty$. Лемма доказана.

Для решения рекуррентных соотношений (I.1.3) нам понадобится также следующая лемма.

Лемма I.1.2 *Пусть последовательность $\{x_t\}$, $x_t > 0$ при любом $t \geq 0$, удовлетворяет рекуррентному соотношению*

$$x_{t+1} = at^{\alpha}(1 + \varepsilon_1(t)) + (1 - bt^{\beta} \cdot (1 + \varepsilon_2(t)))x_t, \qquad \text{(I.1.8)}$$

где $\beta < 0$, $b > 0$, и $\varepsilon_1(t), \varepsilon_2(t) = o(1)$ при $t \to \infty$. Тогда верны следующие асимптотические равенства:
(1) $x_t = \frac{at^{\alpha+1}}{\alpha+1}(1+o(1))$ при $\beta < -1,\ \alpha \geq 0$,
(2) $x_t = \frac{at^{\alpha+1}}{\alpha+b+1}(1+o(1))$ при $\beta = -1,\ \alpha > -1$,
(3) $x_t = \frac{at^{\alpha-\beta}}{b}(1+o(1))$ при $-1 < \beta < 0$.

Доказательство. Сначала докажем (1). Обозначим $x_*(t) = \frac{at^{\alpha+1}}{\alpha+1}$. Просуммировав по $i = 0, \ldots, t$ равенство (I.1.8), получаем

$$x_{t+1} = x_0 + a \cdot \sum_{i=0}^{t} i^\alpha \cdot (1+o(1)) - b \cdot \sum_{i=0}^{t} i^\beta (1+o(1)) x_i \leq$$

$$\leq a \cdot \int_0^t i^\alpha di \cdot (1+o(1)) = \frac{at^{\alpha+1}}{\alpha+1} \cdot (1+o(1)) = x_*(t) \cdot (1+o(1)).$$

Пользуясь этой оценкой, получаем, что

$$\sum_{i=0}^{t} i^\beta x_i = O(t^{\alpha+\beta+2}).$$

Подставляя эту оценку в предыдущее равенство, получаем, что $x_{t+1} = x_*(t) \cdot (1+o(1))$.

Равенство (2) доказывается применением леммы I.1.1 к величине $z_t = (x_t t^{-\alpha-1} - a/(\alpha+b+1))$. Действительно, подставив $x_t = (z_t + a/(\alpha+b+1)) \cdot t^{\alpha+1}$ в соотношение (I.1.8) и разделив на $(t+1)^{\alpha+1}$, получим соотношение

$$\delta z_t = c_0 \varepsilon_3(t)/t - z_t \cdot (\alpha + b + 1 + \varepsilon_4(t))/t$$

для некоторой константы c_0, где $\varepsilon_3(t), \varepsilon_4(t) = o(1)$ при $t \to \infty$, из которого по лемме I.1.1 следует, что $z(t) \to 0$.

Равенство (3) доказывается применением леммы I.1.1 к величине $z_t = (x_t t^{\beta-\alpha} - a/b)$. Аналогичным образом подставив $x_t = (z_t + a/b)t^{\alpha-\beta}$ в (I.1.8) и разделив на $t^{\alpha-\beta}$, получим соотношение

$$\delta z_t = c_0 \varepsilon_3(t) t^{\beta+1} - b z_t t^{\beta+1}(1 + \varepsilon_4(t))$$

для некоторой константы c_0 и $\varepsilon_3(t), \varepsilon_4(t) = o(1)$.

Введем следующие обозначения:

$$\begin{aligned} x_*(t) &= \sum_{i=1}^{k_1} v_i' x_i(t), \\ y_*(t) &= \sum_{i=1}^{k_2} v_i'' y_i(t). \end{aligned} \tag{I.1.9}$$

В силу соотношений (I.1.3), (I.1.4) справедливо равенство

$$Q(t+1) = (x(t), y(t)) = (A - E_t) \cdot Q(t),$$

где $E_t = O(Q(t)) = O\Big(\max(x(t), y(t))\Big)$, причем $0 \leq E_t \leq A$ при достаточно большом t.

Пусть n_0 таково, что $E_t \geq 0$ при $t > n_0$. Тогда можно записать:

$$Q(n) = \prod_{t=n_0}^{n-1} (A - E_t) \cdot Q(n_0).$$

Поскольку вероятности продолжения $Q(t)$ для рассматриваемой грамматики стремятся к нулю при $t \to 0$ в силу согласованности грамматики, то $x(t), y(t) \to 0$ при $t \to \infty$. Следовательно, $E_t \to 0$ при $t \to \infty$ ввиду оценки $E_t = O(\max(x(t), y(t))$. Поскольку $Q(t) \geq 0$, то, применяя утверждение (2) леммы **??**, получаем, что при $t \to \infty$

$$x_i(t) = u'_i x_*(t) \cdot (1 + o(1)), \tag{I.1.10}$$

и $y_i(t)/x_j(t) = o(1)$.

Вычитая из соотношения (I.1.3) такое же соотношение для $t+1$, и пользуясь равенством

$$\delta\big(z_1(t) \cdot \ldots \cdot z_d(t)\big) = \sum_{i=1}^{d} z_1(t+1) \cdot \ldots \cdot z_{i-1}(t+1) \cdot \delta z_i(t) \cdot z_{i+1}(t) \cdot \ldots \cdot z_d(t), \tag{I.1.11}$$

получаем соотношение

$$P(t+1) = -(\delta x(t), \delta y(t)) = (A - E'_t) \cdot P(t),$$

где $E'_t = O(x(t))$, и $E'_t \geq 0$ при достаточно большом t. Отсюда, пользуюсь тем, что $P(t) > 0$, получаем при $t \to \infty$ соотношение

$$\delta x(t) = u'_i \delta x_*(t) \cdot (1 + o(1)). \tag{I.1.12}$$

Из соотношений (I.1.5) для $y_i(t), \delta y_i(t)$ следуют соотношения

$$\begin{aligned} y_i(t) &= u''_{i-k_1} y_*(t) \cdot (1 + o(1)), \\ \delta y_i(t) &= u''_{i-k_1} \delta y_*(t) \cdot (1 + o(1)). \end{aligned} \tag{I.1.13}$$

Лемма I.1.3 *Для величины $x(t)$, заданной рекуррентными соотношениями (I.1.3), (I.1.4), где $x(t) \to 0, y(t)/x(t) \to 0$ при $t \to \infty$, выполнены асимптотическое равенства*

$$x_i(t) \sim u'_i k_0 t^{-1/2},\ x_i(t) - x_i(t+1) \sim u'_i k_0 t^{-3/2}/2,$$

где $k_0 = \sqrt{\frac{4b}{B_1 B_2}}$, и величины B_1, B_2 заданы равенствами

$$\begin{aligned} B_1 &= \textstyle\sum_{i,j,l \leq k_1} v'_i b^i_{jl} u'_j u'_l, \\ B_2 &= \textstyle\sum_{i,j,l > k_1} v''_{i-k_1} b^i_{jl} u''_{j-k_1} u''_{l-k_1}, \end{aligned} \tag{I.1.14}$$

причем $B_1 B_2 > 0$, и $b = v' A^{(2)} u''$ введено ранее для докритического случая.

Сначала получим оценку $x_*(t) \sim k_0 t^{-1/2}$. Для этого покажем, что $t \delta x_*(t) \to 0$ при $t \to \infty$.

Домножим (скалярно) равенство (I.1.3) слева на v', и произведем подстановки величин $x(t), y(t)$ из соотношений (I.1.10), (I.1.13). Пользуясь тем, что $v'A^{(1)} = v'$, получаем соотношение

$$\delta x_*(t) = v'A^{(2)}u'' \cdot y_*(t) \cdot (1 + o(1)) - \sum_{i,j,l \le k_1} v'_i b^i_{jl} u'_j u'_l \cdot x^2_*(t)(1 + o(1))/2. \tag{I.1.15}$$

Подставим в равенство (I.1.3) $t = t + 1$ и вычтем из него аналогичное равенство для t. Домножая слева на v', и учитывая, что $v'A^{(1)} = v'$, получаем соотношение

$$\delta x_*(t+1) = \delta x_*(t) + v'A^{(2)}\delta y(t) \cdot (1 + O(x_*(t)))$$
$$- \sum_{i,j,l \le k_1} v'_i b^i_{jl} \cdot \Big(x_j(t+1)\delta x_l(t) + x_l(t)\delta x_j(t) \Big) \cdot \Big(1 + O\big(x_*(t) + y_*(t)/x_*(t)\big) \Big)/2.$$

Подставляя выражения (I.1.10),(I.1.12) для $x_i(t), \delta x_i(t)$ и (I.1.13) для $y_i(t), \delta y_i(t)$, и используя оценку $\delta x_*(t) = O(x^2_*(t) + y_*(t))$, вытекающую из равенства (I.1.15), получаем

$$\delta x_*(t+1) = \delta x_*(t) + v'A^{(2)}u''\delta y_*(t) \cdot \Big(1 + o(1) + O(x_*(t)) \Big)$$
$$-x_*(t)\delta x_*(t) \cdot \Big(1 + o(1) + O\big(x_*(t) + y_*(t)/x_*(t)\big) \Big) ot \sum_{i,j,l \le k_1} v'_i b^i_{jl} u'_j u'_l. \tag{I.1.16}$$

Рассмотрим величину $z(t) = t\delta x_*(t)$. Подставляя $\delta x_*(t) = z(t)/t$, выражения (I.1.10),(I.1.12),(I.1.13) для $x_i(t), \delta x_i(t), y_i(t), \delta y_i(t)$ в уравнение (I.1.16), и учитывая, что $y_*(t)/x_*(t) \to 0$, $\delta x_*(t) = o(x_*(t))$, имеем

$$\frac{\delta z(t)}{t+1} - \frac{z(t)}{t(t+1)} = v'A^{(2)}u''\delta y_*(t) \cdot (1 + o(1)) - B_1 x_*(t) z(t) \cdot (1 + o(1))/t.$$

Домножая на $t + 1$, учитывая оценку $\frac{1}{tx(t)} = o(1)$, вытекающую из соотношения $y_*(t)/x_*(t) \to 0$ и оценки (I.1.5) для $y_*(t)$, и подставляя $\delta y_*(t) = (2 + o(1))/(B_2 t^2)$, получаем соотношение

$$\delta z(t) = -2b \cdot (1 + o(1))/(B_2 t) - B_1 x_*(t) z(t) \cdot (1 + o(1)). \tag{I.1.17}$$

Опять пользуясь оценкой $\frac{1}{tx(t)} = o(1)$, получаем, что к величине $z(t)$ можно применить лемму I.1.1. Поэтому $z(t) = o(1)$ при $t \to \infty$. Подставляя $z(t) = t\delta x_*(t)$ в уравнение (I.1.15), имеем:

$$by_*(t) - B_1 x^2_*(t)/2 = o(1/t),$$

откуда следует, что $x^2_*(t) = 4b \cdot (1 + o(1))/(B_1 B_2 t)$. Используя соотношения (I.1.10), получаем требуемую асимптотику для $x(t)$. Подставляя найденную оценку для $x_*(t)$ в равенство (I.1.17) и пользуясь леммой I.1.2, получаем, что $z(t) = 2b \cdot (1 + o(1))/(B_1 B_2 x_*(t) t)$, откуда вытекает оценка

$$\delta x_*(t) = -k_0 \cdot (1 + o(1))/(2t^{-3/2}).$$

Из соотношения (I.1.12) находим:

$$x_i(t) - x_{i+1}(t) = -\delta x_i(t) = u'_i k_0 t^{-3/2}(1 + o(1)).$$

Лемма доказана.

как Из этой леммы сразу же следует

Теорема I.1.2 *Для КС-грамматики с двумя классами нетерминалов и матрицей первых моментов, имеющей вид (**??**) при $r' = r'' = 1,\ B_1B_2 > 0$, где B_1, B_2 введены формулами (I.1.14), при $t \to \infty$ верны следующие асимптотические равенства:*

$$Q_i(t) = u_i' k_0 t^{-1/2} \cdot (1 + o(1)),$$

$$P_i(t) = P(D_i^t) = \frac{u_i' k_0 \cdot (1 + o(1))}{2t^{3/2}},\ i = 1, \ldots, k_1,$$

*где $k_0 = \sqrt{\frac{4b}{B_1B_2}}$, и b введено формулой (**??**).*

I.1.2 Математические ожидания числа применений правил в деревьях вывода

Рассмотрим случайные величины $q_{ij}^l(t,\tau), \bar{q}_{ij}^l(t,\tau)$ - соответственно число применений правила r_{ij} в деревьях из D_l^t и $D_l^{\leq t}$ на ярусе τ. Далее, пусть

$$S_{ij}^l(t) = q_{ij}^l(t,1) + q_{ij}^l(t,2) + \ldots + q_{ij}^l(t,t-1),$$

$$\bar{S}_{ij}^l(t) = \bar{q}_{ij}^l(t,1) + \bar{q}_{ij}^l(t,2) + \ldots + \bar{q}_{ij}^l(t,t-1)$$

- число применений правила r_{ij} в дереве вывода из D_l^t и $D_l^{\leq t}$ соответственно. Будем обозначать

$$S_{ij}(t) = S_{ij}^1(t),\ \bar{S}_{ij}(t) = \bar{S}_{ij}^1(t),$$

$$q_{ij}(t,\tau) = q_{ij}^1(t,\tau),\ \bar{q}_{ij}(t,\tau) = \bar{q}_{ij}^1(t,\tau).$$

Для краткости обозначим $M_{ij}^l(t) = M(S_{ij}^l(t)),\ \bar{M}_{ij}^l(t) = M(\bar{S}_{ij}^l(t))$.

Используя установленные в теореме I.1.2 свойства деревьев вывода, можно построить рекуррентные соотношения для величин $M_{ij}^l(t),\ \bar{M}_{ij}^l(t)$ при $l, i = 1, \ldots, k,\ j = 1, \ldots, n_i$.

Для решения этих соотношений нам понадобится следующая лемма, в некотором смысле усиленный вариант леммы **??**:

Лемма I.1.4 *Пусть $A(t) \to A$ при $t \to \infty$ - последовательность матриц размером $k \times k$, причем матрица $A > 0$ и ее перронов корень $r = 1$. Пусть $b(t) = bt^\alpha(1 + o(1))$ при $t \to \infty$ - последовательность векторов длины k, где $b \geq 0, b \neq 0$ - вектор длины k, а α - произвольное действительное число. Тогда для последовательности векторов $x(t), t = 0, 1, \ldots$, 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 $v, u > 0$ - соответственно левый и правый собственные вектора матрицы A при нормировке $vu = 1$.

Для доказательства утверждения леммы достаточно показать, что для любого $\varepsilon > 0$ существует номер $t_1 > t_0$, такой, что $|x(t) - uvx(t)| < \varepsilon uvx(t)$ при $t > t_1$. Зафиксируем произвольно малые величины $0 < \varepsilon_1, \varepsilon_2 < 1$. Возьмем такое $t_2 = t_2(\varepsilon_1)$, что $|A(t) - A| < \varepsilon_1 A$, $|b(t) - bt^\alpha| < bt^\alpha \varepsilon_1$ при $t \geq t_2$. Поскольку $A^t \to uv$ при $t \to \infty$ для $A > 0$ [16], то существует $n_1 = n_1(\varepsilon_2)$, такое что $|A^t - uv| < uv\varepsilon_2$ при $t \geq n_1$. Положим $n_2 = n_2(\varepsilon_2) = n_1/\varepsilon_2$.

Введем обозначение $A^*(m,n) = \prod\limits_{i=0}^{m-n-1} A(m-i)$ (при $m > n$), причем считаем, что $A^*(n,n) = E$. Очевидно,

$$x(t+t_2) = A^*(t+t_2, t_2)x(t_2) + \sum_{m=1}^{t} A^*(t+t_2, m+t_2)b(m+t_2). \tag{I.1.18}$$

В силу неравенства $(1-\varepsilon_1)A < A(t) < (1+\varepsilon_1)A$, выполняющегося при $t \geq t_2$, получаем, что при $n \geq t_2$ справедливо неравенство

$$(1-\varepsilon_1)^m \cdot A^m < A^*(m+n, n) < (1+\varepsilon_1)^m \cdot A^m. \tag{I.1.19}$$

Положим в равенстве (I.1.18) $t = n_2$. Обозначим

$$\begin{aligned}
S_0 &= A^*(n_2+t_2, t_2)x(t_2),\\
S_1 &= \sum_{m=1}^{n_2-n_1} A^*(n_2+t_2, m+t_2)b(m+t_2),\\
S_2 &= \sum_{m=n_2-n_1+1}^{n_2} A^*(n_2+t_2, m+t_2)b(m+t_2).
\end{aligned}$$

Вклад первого слагаемого S_0 в $x(n_2+t_2) - uvx(n_2+t_2)$ можно оценить как

$$|A^*(n_2+t_2, t_2)x(t_2) - uvA^*(n_2+t_2, t_2)x(t_2)| \leq |A^*(n_2+t_2, t_2)x(t_2) - A^{n_2}x(t_2)|$$

$$+|A^{n_2}x(t_2) - uvA^{n_2}x(t_2)| + |uvA^{n_2}x(t_2) - uvA^*(n_2+t_2, t_2)x(t_2)|.$$

Ввиду неравенства $|A^{n_2} - uv| < uv\varepsilon_2$ и неравенств (I.1.19) можно записать

$$|A^{n_2}x(t_2) - uvA^{n_2}x(t_2)| < |(1+\varepsilon_2)uvx(t_2) - (1-\varepsilon_2)uvx(t_2)| < 2\varepsilon_2(1-\varepsilon_2)^{-1}A^{n_2}x(t_2)$$

$$\leq 2\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2}A^*(n_2+t_2, t_2)x(t_2).$$

Используя неравенство $uvx > \min(uv)x$, и обозначая $M_{uv} = (\min(uv))^{-1}$, имеем неравенство

$$|A^{n_2}x(t_2) - uvA^{n_2}x(t_2)| < 2M_{uv}\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2}uvA^*(n_2+t_2, t_2)x(t_2).$$

Опять применяя неравенства (I.1.19), получаем оценку

$$|A^*(n_2+t_2, t_2)x(t_2) - A^{n_2}x(t_2)| + |uvA^{n_2}x(t_2) - uvA^*(n_2+t_2, t_2)x(t_2)|$$

$$\begin{aligned}
&< (1+M_{uv}) \cdot |uvA^*(n_2+t_2, t_2)x(t_2) - uvA^{n_2}x(t_2)|\\
&< (1+M_{uv})\left((1-\varepsilon_1)^{-n_2} - 1\right) uvA^*(n_2+t_2, t_2)x(t_2).
\end{aligned}$$

Поэтому

$$|S_0 - uvS_0| < \Big(2M_{uv}\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2} + (1+M_{uv}) \cdot \big((1-\varepsilon_1)^{-n_2} - 1\big)\Big) \cdot uvS_0. \tag{I.1.20}$$

Вклад в $x(n_2+t_2)-uvx(n_2+t_2)$ слагаемых из S_1 оценивается с помощью неравенств (I.1.19) таким же образом как

$$|A^*(n_2+t_2,m+t_2)b(m+t_2)-uvA^*(n_2+t_2,m+t_2)b(m+t_2)|\leq$$

$$\Big(2M_{uv}\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2}+(1+M_{uv})\cdot\big((1-\varepsilon_1)^{-n_2}-1\big)\Big)\cdot uvA^*(n_2+t_2,m+t_2)b(m+t_2),$$

откуда получаем оценку

$$|S_1-uvS_1|<\Big(2M_{uv}\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2}+(1+M_{uv})\cdot\big((1-\varepsilon_1)^{-n_2}-1\big)\Big)\cdot uvS_1. \quad \text{(I.1.21)}$$

Поскольку матрицы A^m, $m\geq 0$, положительны и ограничены, можно считать, что $M_1<A^m<M_2$ для некоторых констант $M_1,M_2>0$ при любом $m\geq 0$. Оценим вклад в $x(n_2+t_2)-uvx(n_2+t_2)$ слагаемых из S_2. Докажем, что он мал по сравнению со всей суммой $S_1+S_2=\sum\limits_{m=1}^{n_2}A^*(n_2+t_2,m+t_2)b(m+t_2)$. Опять применяя оценку (I.1.19), и пользуясь оценкой $b(m+t_2)<(1+\varepsilon_1)b\cdot(m+t_2)^{\alpha}$, получим:

$$|S_2-uvS_2|\leq(1+k\max(uv))\cdot\max(S_2)$$

$$\leq(1+k\max(uv))\cdot(1+\varepsilon_1)^{n_1}\max\left(\sum_{m=0}^{n_1-1}A^mb\right)\cdot\max_{m=n_2-n_1+1,\dots,n_2}(m+t_2)^{\alpha}$$

$$\leq\big(1+k\max(uv)\big)kM_2n_1\cdot(1+\varepsilon_1)^{n_1}\max(b)\cdot x_{m=n_2-n_1+1,\dots,n_2}(m+t_2)^{\alpha}. \quad \text{(I.1.22)}$$

Оценим сумму S_1+S_2 снизу:

$$S_1+S_2=\sum_{m=1}^{n_2}A^*(n_2+t_2,m+t_2)b(m+t_2)>(1-\varepsilon_1)^{n_2}\cdot\min_{m=1,\dots,n_2}(m+t_2)^{\alpha}\cdot\sum_{m=0}^{n_2-1}A^mb$$

$$>(1-\varepsilon_1)^{n_2}n_2M_1b\cdot\min_{m=1,\dots,n_2}(m+t_2)^{\alpha}.$$

Поэтому

$$uv(S_1+S_2)>(1-\varepsilon_1)^{n_2}M_{uv}n_2M_1\max(b)\cdot\min_{m=1,\dots,n_2}(m+t_2)^{\alpha}.$$

в силу неравенства $uvx>\min(uv)\max(x)$. Умножая и деля правую часть оценки (I.1.22) на правую часть полученного неравенства, получаем оценку

$$|S_2-uvS_2|\leq\frac{n_1kM_2(1+\varepsilon_1)^{n_1}\cdot(1+k\max(uv))\cdot\max\limits_{m=n_2-n_1+1,\dots,n_2}(m+t_2)^{\alpha}}{M_{uv}n_2M_1(1-\varepsilon_1)^{n_2}\cdot\min\limits_{m=1,\dots,n_2}(m+t_2)^{\alpha}}otuv(S_1+S_2)$$

Учитывая, что $n_1/n_2=\varepsilon_2$, и применив очевидное неравенство

$$\frac{\max\limits_{m=n_2-n_1+1,\dots,n_2}(m+t_2)^{\alpha}}{\min\limits_{m=1,\dots,n_2}(m+t_2)^{\alpha}}\leq\Big(\frac{t_2+n_2}{t_2}\Big)^{|\alpha|},$$

имеем оценку

$$|S_2-uvS_2|\leq\frac{\varepsilon_2kM_2\cdot(1+k\max(uv))}{M_{uv}M_1}\cdot\Big((1+\varepsilon_1)^{n_1}/(1-\varepsilon_1)^{n_2}\Big)$$

$$\cdot\Big(\frac{t_2+n_2}{t_2}\Big)^{|\alpha|} uv(S_1+S_2). \qquad \text{(I.1.23)}$$

Суммируя неравенства (I.1.20), (I.1.21), (I.1.23), и пользуясь неравенством $|A+B| < |A|+|B|$, получим оценку

$$|x(n_2+t_2) - uvx(n_2+t_2)| \le |S_0 - uvS_0| + |S_1 - uvS_1| + |S_2 - uvS_2| <$$

$$\Big(2M_{uv}\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2} + (1+M_{uv})\cdot\big((1-\varepsilon_1)^{-n_2}-1\big)\Big)\cdot uvS_0 +$$

$$\Big[\frac{\varepsilon_2 k M_2\cdot(1+k\max(uv))}{M_{uv}M_1}\cdot\Big((1+\varepsilon_1)^{n_1}/(1-\varepsilon_1)^{n_2}\Big)\cdot\Big(\frac{t_2+n_2}{t_2}\Big)^{|\alpha|} +$$

$$\Big(2M_{uv}\varepsilon_2(1-\varepsilon_2)^{-1}(1-\varepsilon_1)^{-n_2} + (1+M_{uv})\cdot\big((1-\varepsilon_1)^{-n_2}-1\big)\Big)\Big]\cdot uv(S_1+S_2).$$

Учитывая, что $x(n_2+t_2) = S_0+S_1+S_2$, получим для некоторой константы $C>0$, не зависящей от t_2, n_2:

$$|x(n_2+t_2) - uvx(n_2+t_2)| < C\cdot\Big(\big((1-\varepsilon_1)^{-n_2}-1\big) + 2\varepsilon_2(1-\varepsilon_2)^{-1}\cdot\Big(\frac{1+\varepsilon_1}{1-\varepsilon_1}\Big)^{n_2}$$

$$\times\Big(\frac{t_2+n_2}{t_2}\Big)^{|\alpha|}\Big) uvx(n_2+t_2).$$

Устремляя сначала $\varepsilon_2 \to 0$, а затем $\varepsilon_1 \to 0$ и $t_2 = t_2(\varepsilon_1) \to \infty$, получим соответствующие номера t_2, n_2, такие что $|x(n_2+t) - uvx(n_2+t)| < \varepsilon uvx(n_2+t)$ для сколь угодно малого ε при $t > t_2$. Лемма доказана.

Рассмотрим величину $\bar{M}^q_{ij}(t)$. Пусть $D^{\le t}_{ql}$ - множество деревьев из $D^{\le t}_q$, корень которых помечен нетерминалом A_q, и в нем применено правило r_{ql}. Обозначим через $\bar{P}^{ij}_{ql}(t)$ величину $\sum\limits_{d\in D^{\le t}_{ql}} p(d)q_{ij}(d)$, где $p(d)$ - условная вероятность дерева d в $D^{\le t}_q$, а $q_{ij}(d)$ - общее число применений правила r_{ij} в дереве d. Очевидно, $\bar{M}^q_{ij}(t) = \sum_l \bar{P}^{ij}_{ql}(t)$. В дальнейшем в обозначении $\bar{P}^{ij}_{ql}(t)$ будем для краткости опускать индексы i,j.

Напомним, что вероятность наборов деревьев вывода $D^{\le t}_X$ (где $X=(x_1,\dots,x_k)$), в которых ровно x_i деревьев имеют корень, помеченный нетерминалом A_i для $i=1,\dots,k$, и высоты которых не превосходят t, обозначается $Q_X(t)$, а вероятность наборов деревьев D^t_X из $D^{\le t}_X$, в которых хотя бы одно достигает яруса t, обозначается $R_X(t)$. Обозначим за $s_{ij} = (s^1_{ij},\dots,s^k_{ij})$ вектор нетерминалов в правой части правила r_{ij}.

Рассмотрим величину $\bar{P}_{ql}(t)$. Можно записать $q_{ij}(d) = q^{(1)}_{ij}(d) + q^{(2)}_{ij}(d)$, где $q^{(1)}_{ij}(d)$ равно единице, если $r_{ij} = r_{ql}$, и нулю в противном случае, а $q^{(2)}_{ij}(d)$ обозначает количество применений правила r_{ij} в поддеревьях с корнем на первом ярусе. Очевидно, что

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

Таким образом, $\bar{P}^1_{ql}(t)$ учитывает только применение правила r_{ij} в корне дерева, а $\bar{P}^2_{ql}(t)$ учитывает количество применений правила r_{ij} к остальным вершинам дерева.

Очевидно, что $\bar{P}^1_{ql}(t) = p_{ij}Q_{s_{ij}}(t-1)/(1-Q_q(t))$ при $i = q, j = l$, и $\bar{P}^1_{ql}(t) = 0$ в противном случае.

Обозначим через $\delta^i(n) = (\delta^i_1, \ldots, \delta^i_n)$ вектор длины n, в котором на позиции i стоит единица, а все остальные компоненты равны нулю. Параметр n будет опускаться, если из контекста длина вектора известна.

Вероятность дерева $d \in D^{\leq t}_{ql}$ можно представить в виде $p(d) = p_{ql}p_1(d)p_2(d) \cdot \ldots \cdot p_{\bar{s}_{ql}}(d)$, где $p_i(d)$ - вероятность поддерева d с i-м по порядку корнем на втором ярусе. Количество применений правила r_{ij} в дереве d на ярусах с номерами, большими единицы $q'_{ij}(d)$ можно представить в виде суммы $q'_{ij}(d) = \sum\limits_{m=1}^{\bar{s}_{ql}} q'^m_{ij}(d)$, где $q'^m_{ij}(d)$ - количество применений правила r_{ij} в поддереве с m-м по порядку корнем на втором ярусе.

Очевидно,

$$\bar{P}^2_{ql}(t) = \frac{1}{1-Q_q(t)} \cdot \sum_d p(d)q'_{ij}(d) = \frac{p_{ql}}{1-Q_q(t)} \cdot \sum_d \prod_{n=1}^{\bar{s}_{ql}} p_n(d) \sum_{m=1}^{\bar{s}_{ql}} q'^m_{ij}(d)$$

$$= \frac{p_{ql}}{1-Q_q(t)} \cdot \sum_{m=1}^{\bar{s}_{ql}} \sum_d p_1(d) \ldots p_{m-1}(d)p_{m+1}(d) \ldots p_{\bar{s}_{ql}}(d) \cdot p_m(d)\bar{M}^m_{ij}(t-1)$$

$$= \frac{p_{ql}}{1-Q_q(t)} \cdot \sum_m s^m_{ql}\bar{M}^m_{ij}(t-1) \cdot Q_{s_{ql}}(t-1).$$

Суммируя по l, получаем следующее рекуррентное соотношение для $\bar{M}^q_{ij}(t)$:

$$\bar{M}^q_{ij}(t) = \frac{1}{1-Q_q(t)} \cdot \left(\delta^q_i p_{ij}Q_{s_{ij}}(t-1) + \sum_l p_{ql} \sum_m s^m_{ql}\bar{M}^m_{ij}(t-1)Q_{s_{ql}}(t-1) \right). \quad \text{(I.1.24)}$$

Обозначим $\bar{M}'^q_{ij}(t) = \bar{M}^q_{ij}(t) \cdot (1-Q_q(t))$. Тогда соотношение (I.1.24) можно записать в виде

$$\bar{M}'^q_{ij}(t) = \delta^q_i p_{ij}Q_{s_{ij}}(t-1) + \sum_l p_{ql} \sum_m s^m_{ql}\bar{M}'^m_{ij}(t-1)Q_{s_{ql}-\delta^{z_m}}(t-1). \quad \text{(I.1.25)}$$

Отметим, что $\bar{M}'^q_{ij}(t) = \bar{M}^q_{ij}(t) \cdot (1 + o(1))$.

Аналогично соотношению для $\bar{M}^q_{ij}(t)$ получаем рекуррентное соотношение для $M^q_{ij}(t)$. Вклад $P^1_{ql}(t)$ правила, примененного к корню дерева, равен $\delta^q_i p_{ij}R_{s_{ij}}(t-1)/P_q(t)$. Вклад $P^2_{ql}(t)$ деревьев из $D^t_{s_{ql}}$ можно представить в виде суммы $\sum_m P^{2m}_{ql}(t)$, где $P^{2m}_{ql}(t)$ учитывает вклад дерева, корень которого помечен m-м по порядку нетерминалом $z_m, m = 1, \ldots, \bar{s}_{ql}$, стоящим в правой части правила r_{ql}. Вклад в $P^{2m}_{ql}(t)$ наборов деревьев из $D^{t-1}_{s_{ql}}$, в которых ярус t достигается через дерево с корнем на первом ярусе, помеченным нетерминалом z_m, равен

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

а вклад наборов, в которых ярус t достигается через другие деревья, равен

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

Подставляя $P_{ql}^{2m}(t) = S_1 + S_2$, получаем равенство

$$M_{ij}^q(t) = \sum_{l=1}^{n_q} \left(P_{ql}^1(t) + \sum_{m=1}^{\bar{s}_{ql}} P_{ql}^{2m}(t) \right)$$

$$= \frac{1}{P_q(t)} \cdot \Big(\delta_i^q p_{ij} R_{s_{ij}}(t-1) + \sum_l p_{ql} \sum_m s_{ql}^m \cdot \big(P_m(t-1) Q_{s_{ql}-\delta^m}(t-1) M_{ij}^m(t-1)$$

$$+ (1 - Q_m(t-1)) R_{s_{ql}-\delta^m}(t-1) \bar{M}_{ij}^m(t-1) \big) \Big). \tag{I.1.26}$$

Из леммы **??** вытекают равенства

$$\begin{aligned} Q_X(t) &= \prod_{i=1}^k (1 - Q_i(t))^{x_i} = 1 - \sum_{i=1}^k x_i Q_i(t) + O\left(\sum_{i,j} x_i x_j Q_i(t) Q_j(t) \right), \\ R_X(t) &= Q_X(t) - Q_X(t-1) = \sum_{i=1}^k x_i P_i(t) + O\left(\sum_{i,j} x_i x_j Q_i(t) P_j(t) \right), \end{aligned} \tag{I.1.27}$$

применяя которые к выражениям для $Q_i(t), P_i(t)$, полученным в теореме I.1.2, получаем оценки

$$Q_{s_{ij}-\delta^m}(t) = \begin{cases} 1 - \sum_{l \le k_1} (s_{ij}^l - \delta_l^m) Q_l(t) + o(t^{-1/2}), \ i \le k_1, \\ 1 - \sum_{l > k_1} (s_{ij}^l - \delta_l^m) Q_l(t) + o(1/t), \ i > k_1, \end{cases} \tag{I.1.28}$$

$$R_{s_{ij}-\delta^m}(t) = \begin{cases} \sum_{l \le k_1} (s_{ij}^l - \delta_l^m) P_l(t) + o(t^{-3/2}), \ i \le k_1, \\ \sum_{l > k_1} (s_{ij}^l - \delta_l^m) P_l(t) + o(1/t^2), \ i > k_1. \end{cases} \tag{I.1.29}$$

Найдем сначала $\bar{M}_{ij}^q(t), \bar{M}_{ij}^{\prime q}(t)$ при $i, q > k_1$. Подставляя оценки (I.1.28), (I.1.29) и полученные ранее выражения для $Q_i(t)$ в соотношение (I.1.25), и учитывая, что $(1 - Q_i(t))^{-1} = 1 + Q_i(t) + O(Q_i^2(t))$, получаем следующее соотношение для величин $\bar{M}_{ij}^{\prime q}(t)$:

$$\bar{M}_{ij}^{\prime q}(t) = \delta_i^q p_{ij} + \sum_{m > k_1} \sum_l p_{ql} s_{ql}^m M_{ij}^{\prime m}(t-1)$$

$$- \sum_{m,n > k_1} \sum_l p_{ql} s_{ql}^m (s_{ql}^n - \delta_n^m) \cdot Q_n(t-1) \bar{M}_{ij}^{\prime m}(t-1) \cdot (1 + o(1))$$

$$= \delta_i^q p_{ij} + \sum_{m > k_1} a_m^q \bar{M}_{ij}^{\prime m}(t-1) - \frac{2 + o(1)}{B_2 t} \cdot \sum_{m,n > k_1} b_{mn}^q u_{n-k_1}'' \bar{M}_{ij}^{\prime m}(t-1). \tag{I.1.30}$$

В последнем преобразовании были использованы равенства

$$a_m^q = \sum_l p_{ql} s_{ql}^m, \ b_{mn}^q = \sum_l p_{ql} s_{ql}^m (s_{ql}^n - \delta_n^m).$$

Применяя лемму I.1.4, получаем, что

$$\bar{M}_{ij}^{\prime q}(t) = u_{q-k_1}'' \cdot (1 + o(1)) \cdot \sum_{n > k_1} v_{n-k_1}'' \bar{M}_{ij}^n(t). \tag{I.1.31}$$

Если обозначить $\bar{M}^{*}_{ij}(t) = \sum_{n>k_1} v''_{n-k_1}\bar{M}'^{n}_{ij}(t)$, то, домножив соотношение (I.1.30) слева на v''_{q-k_1}, суммируя по $q > k_1$ и используя равенства (I.1.31), получим:

$$\delta\bar{M}^{*}_{ij}(t) = v''_{i-k_1}p_{ij} - \frac{(2+o(1))\bar{M}^{*}_{ij}(t)}{B_2 t}\cdot\sum_{m,n,q>k_1} v''_{q-k_1}b^{q}_{mn}u''_{m-k_1}u''_{n-k_1}$$

$$= v''_{i-k_1}p_{ij} - \bar{M}^{*}_{ij}(t)\cdot(2+o(1))/t.$$

Применяя лемму I.1.2, получаем $\bar{M}^{*}_{ij}(t) = v''_{i-k_1}p_{ij}t\cdot(1+o(1))/3$, то есть

$$\begin{aligned} \bar{M}'^{q}_{ij}(t) &= u''_{q-k_1}v''_{i-k_1}p_{ij}t\cdot(1+o(1))/3 \\ \bar{M}^{q}_{ij}(t) &= u''_{q-k_1}v''_{i-k_1}p_{ij}t\cdot(1+o(1))/3. \end{aligned} \tag{I.1.32}$$

Теперь найдем $M^{q}_{ij}(t)$ при $i, q > k_1$. Обозначим далее $M'^{q}_{ij}(t) = M^{q}_{ij}(t)P_q(t)$. Из равенств (I.1.26) с использованием оценок (I.1.28), (I.1.29), выражений для $Q_i(t), P_i(t)$, и доказанного равенства (I.1.32) для $\bar{M}^{m}_{ij}(t)$, получаем соотношение

$$M'^{q}_{ij}(t) = O(p_{ij}t^{-3/2}) + \sum_{m>k_1}\sum_{l} p_{ql}s^{m}_{ql}M'^{m}_{ij}(t-1)$$

$$-\sum_{m,n>k_1}\sum_{l} p_{ql}s^{m}_{ql}(s^{n}_{ql}-\delta^{m}_{n})\cdot Q_n(t-1)M'^{m}_{ij}(t-1)$$

$$+\sum_{m,n>k_1}\sum_{l} p_{ql}s^{m}_{ql}(s^{n}_{ql}-\delta^{m}_{n})\cdot P_n(t-1)\bar{M}^{m}_{ij}(t-1)$$

$$= \sum_{m>k_1} a^{q}_{m}M'^{m}_{ij}(t-1) - \frac{2+o(1)}{B_2 t}\cdot\sum_{m,n>k_1} b^{q}_{mn}u''_{n-k_1}M'^{m}_{ij}(t-1)$$

$$+\frac{p_{ij}v''_{i-k_1}\cdot(2+o(1))}{3B_2 t}\cdot\sum_{m,n>k_1} b^{q}_{mn}u''_{n-k_1}u''_{m-k_1}.$$

Применяя лемму I.1.4, имеем:

$$M'^{q}_{ij}(t) = u''_{q-k_1}\cdot\sum_{n>k_1} v''_{n-k_1}M'^{n}_{ij}(t)\cdot(1+o(1)).$$

Домножая полученное рекуррентное соотношение для $M'^{q}_{ij}(t)$ на v''_{q-k_1}, суммируя по $q > k_1$ и обозначая $M^{*}_{ij}(t) = \sum_{n>k_1} v''_{n-k_1}M'^{n}_{ij}(t)$, аналогично предыдущему случаю получаем рекуррентное соотношение

$$\delta M^{*}_{ij}(t) = (2+o(1))p_{ij}v''_{i-k_1}/(3t) - (2+o(1))M^{*}_{ij}(t)/t,$$

откуда получаем $M^{*}_{ij}(t) = p_{ij}v''_{i-k_1}\cdot(1+o(1))/3$, и

$$M^{q}_{ij}(t) = \frac{p_{ij}}{3P_q(t)}u''_{q-k_1}v''_{i-k_1}\cdot(1+o(1)) = p_{ij}v''_{i-k_1}B_2t^2\cdot(1+o(1))/6.$$

Эта формула совпадает с соответствующей формулой, выведенной в [10] для неразложимой грамматики.

Аналогично для $\bar{M}_{ij}^{'q}(t)$ при $i, q \le k_1$ получаем соотношение

$$\bar{M}_{ij}^{'q}(t) = \delta_i^q p_{ij} + \sum_{m\le k_1} a_m^q \bar{M}_{ij}^{'m}(t-1)$$

$$-k_0(t-1)^{-1/2}(1+o(1)) \cdot \sum_{m,n\le k_1} b_{mn}^q u_n' \bar{M}_{ij}^{'m}(t-1).$$

Применяя лемму I.1.4, имеем:

$$\bar{M}_{ij}^{'q}(t) = u_q' \cdot \sum_{n\le k_1} v_n' \bar{M}_{ij}^{'n}(t) \cdot (1+o(1)).$$

Домножая рекуррентное соотношение для $\bar{M}_{ij}^{'q}(t)$ на v' и обозначая $\bar{M}_{ij}^*(t) = \sum_{n\le k_1} v_n' \bar{M}_{ij}^{'n}(t)$, получаем для $\bar{M}_{ij}^*(t)$ рекуррентное соотношение

$$\delta \bar{M}_{ij}^*(t) = v_i' p_{ij} - B_1 k_0 \bar{M}_{ij}^*(t) \cdot (1+o(1)) t^{-1/2},$$

откуда находим $\bar{M}_{ij}^*(t) = v_i' p_{ij} t^{1/2} \cdot (1+o(1))/(B_1 k_0)$, и

$$M_{ij}^{'q}(t) = u_q' v_i' p_{ij} \cdot t^{1/2}(1+o(1))/(B_1 k_0),$$
$$M_{ij}^{q}(t) = u_q' v_i' p_{ij} \cdot t^{1/2}(1+o(1))/(B_1 k_0).$$

Рассмотрим величины $M_{ij}^q(t),\ i, q \le k_1$. Рекуррентное соотношение для $M_{ij}^{\prime q}(t)$ имеет вид:

$$M_{ij}^{'q}(t) = O(p_{ij} t^{-3/2}) + \sum_{m\le k_1} a_m^q M_{ij}^{'m}(t-1) - k_0 t^{-1/2}(1+o(1)) \cdot \sum_{m,n\le k_1} b_{mn}^q u_n' M_{ij}^{'m}(t-1)$$

$$+ \sum_{m,n\le k_1} b_{mn}^q P_n(t-1) \bar{M}_{ij}^m(t-1)(1+o(1)) =$$

$$= \sum_{m\le k_1} a_m^q M_{ij}^{\prime m}(t-1) - k_0 t^{-1/2}(1+o(1)) \cdot \sum_{m,n\le k_1} b_{mn}^q u_n' M_{ij}^{\prime m}(t-1)$$

$$+ \frac{v_i' p_{ij} \cdot (1+o(1))}{2tB_1} \cdot \sum_{m,n>k_1} b_{mn}^q u_n' u_m'.$$

По лемме I.1.4,

$$M_{ij}^{\prime q}(t) = u_q' \cdot \sum_{n\le k_1} v_n' M_{ij}^{\prime n}(t) \cdot (1+o(1)).$$

Обозначая $M_{ij}^*(t) = \sum_{n\le k_1} v_n' M_{ij}^{\prime n}(t)$, получаем соотношение

$$\delta M_{ij}^*(t) = p_{ij} v_i' \cdot (1+o(1))/(2t) - B_1 k_0 t^{-1/2}(1+o(1)) \cdot M_{ij}^*(t),$$

откуда следует асимптотическое равенство

$$M_{ij}^q(t) = p_{ij} u_q' v_i' \cdot (1+o(1))/(2B_1 k_0 t^{1/2} \cdot P_q(t)) = v_i' p_{ij} t \cdot (1+o(1))/(B_1 k_0^2).$$

Для $\bar{M}'^{q}_{ij}(t)$ при $q \le k_1,\ i > k_1$ имеем:

$$\bar{M}'^{q}_{ij}(t) = O(p_{ij}) + \sum_{m \le k_1} \sum_{l} p_{ql} s^{m}_{ql} \bar{M}'^{m}_{ij}(t-1) \left(1 - \sum_{n \le k_1} (s^{n}_{ql} - \delta^{m}_{n}) Q_n(t-1) \right) \cdot (1 + o(1))$$

$$+ \sum_{m > k_1} \sum_{l} p_{ql} s^{m}_{ql} \bar{M}'^{m}_{ij}(t-1) \cdot (1 + o(1)).$$

Подставляя полученные выше выражения для $\bar{M}'^{m}_{ij}(t),\ \ m, i > k_1$, и производя аналогичные предыдущим случаям преобразования, получаем равенство:

$$\bar{M}'^{q}_{ij}(t) = \sum_{m \le k_1} a^{q}_{m} \bar{M}'^{m}_{ij}(t-1) - k_0 t^{-1/2} (1 + o(1)) \cdot \sum_{m,n \le k_1} b^{q}_{mn} u'_n \bar{M}'^{m}_{ij}(t-1)$$

$$+ \frac{p_{ij} v''_{i-k_1} t \cdot (1 + o(1))}{3} \cdot \sum_{m > k_1} a^{q}_{m} u''_{m-k_1}.$$

Отсюда, домножая на v'_q и суммируя, для $\bar{M}^{*}_{ij}(t) = \sum_{n \le k_1} v'_n \bar{M}'^{n}_{ij}(t)$ получаем соотношение

$$\delta \bar{M}^{*}_{ij}(t) = p_{ij} v''_{i-k_1} b t \cdot (1 + o(1))/3 - B_1 k_0 t^{-1/2} \bar{M}^{*}_{ij}(t) \cdot (1 + o(1)),$$

(напомним, что $b = \sum_{q \le k_1, m > k_1} a^{q}_{m} v'_q u''_{m-k_1}$), откуда следует,что

$$\begin{aligned} \bar{M}'^{q}_{ij}(t) &= p_{ij} b u'_q v''_{i-k_1} (1 + o(1)) t^{3/2} / (3 B_1 k_0), \\ \bar{M}^{q}_{ij}(t) &= p_{ij} b u'_q v''_{i-k_1} (1 + o(1)) t^{3/2} / (3 B_1 k_0). \end{aligned}$$

И, наконец, найдем величины $M^{q}_{ij}(t),\ \ q \le k_1, i > k_1$. Рекуррентное соотношение для $M'^{q}_{ij}(t)$ имеет вид:

$$M'^{q}_{ij}(t) = O(t^{-3/2}) + \sum_{m} a^{q}_{m} M'^{m}_{ij}(t-1) - \sum_{m < k, n \le k_1} b^{q}_{mn} Q_n(t) M'^{m}_{ij}(t-1) \cdot (1 + o(1))$$

$$+ \left(\sum_{m < k, n \le k_1} b^{q}_{mn} P_n(t-1) \bar{M}^{m}_{ij}(t-1) \right) \cdot (1 + o(1)).$$

Заметим, что

$$\sum_{m > k_1, n \le k_1} b^{q}_{mn} Q_n(t) M'^{m}_{ij}(t-1) + \sum_{m > k_1, n \le k_1} b^{q}_{mn} P_n(t-1) \bar{M}^{m}_{ij}(t-1) = O(t^{-1/2}$$

в силу полученных ранее оценок для $\bar{M}^{m}_{ij}(t-1), M^{m}_{ij}(t-1)$ при $m, i > k_1$. Поэтому, подставляя $M'^{m}_{ij}(t-1) = P_m(t-1) M^{m}_{ij}(t-1)$ для $m > k_1$, получаем соотношение

$$M'^{q}_{ij}(t) = O(t^{-1/2}) + \sum_{m \le k_1} a^{q}_{m} M'^{m}_{ij}(t-1) - \sum_{m,n \le k_1} b^{q}_{mn} Q_n(t) M'^{m}_{ij}(t-1) \cdot (1 + o(1))$$

$$+ \left(\sum_{m,n \le k_1} b^{q}_{mn} P_n(t-1) \bar{M}^{m}_{ij}(t-1) + \sum_{m > k_1} a^{q}_{m} P_m(t-1) M^{m}_{ij}(t-1) \right) \cdot (1 + o(1)).$$

Используя найденные ранее выражения для $\bar{M}_{ij}^m(t),\ m,i \le k_1$ и $M_{ij}^m(t),\ m > k_1, i \le k_1$, и применяя лемму I.1.4, получаем для $M_{ij}^*(t) = \sum_{n\le k_1} v'_n M'^n_{ij}(t)$ следующее рекуррентное соотношение:

$$\delta M_{ij}^*(t) = O(t^{-1/2}) + \frac{bp_{ij}v''_{i-k_1}t^{3/2}}{3B_1k_0} \cdot \frac{k_0}{2t^{3/2}} \cdot \sum_{q,m\le k_1}\sum_{n\le k_1} v'_q b^q_{mn} u'_n u'_m + v''_{i-k_1} \cdot \sum_{m>k_1, q\le k_1} v'_q a^q_m u''_{m-k_1}$$

$$\cdot \frac{2}{B_2t^2} \cdot p_{ij}B_2t^2 \cdot (1+o(1))/6 - M_{ij}^*(t)B_1k_0t^{-1/2} \cdot (1+o(1))$$

$$= bp_{ij}v''_{i-k_1} \cdot (1+o(1))/2 - M_{ij}^*(t)B_1k_0t^{-1/2} \cdot (1+o(1)),$$

откуда получаем

$$M_{ij}^*(t) = v''_{i-k_1} bp_{ij} t^{1/2} \cdot (1+o(1))/(2B_1k_0),$$

$$M_{ij}^q(t) = v''_{i-k_1} bp_{ij} t^2 \cdot (1+o(1))/(B_1k_0^2).$$

Полученные результаты сформулируем в виде следующей теоремы.

Теорема I.1.3 *Для разложимой грамматики с двумя классами нетерминалов, матрица первых моментов которой имеет вид (**??**) и $r' = r'' = 1,\ B_1B_2 > 0$, где B_1, B_2 введены формулами (I.1.14), величины $M(S_{ij}(t))$ имеют следующий вид при $t \to \infty$:*

$$M(S_{ij}(t)) = \frac{p_{ij}v'_iB_2t \cdot (1+o(1))}{4b},\ i \le k_1$$

$$M(S_{ij}(t)) = p_{ij}v''_{i-k_1}B_2t^2 \cdot (1+o(1))/4,\ i > k_1,$$

*где b дано формулой (**??**), а p_{ij} - вероятность применения правила r_{ij} в грамматике.*

Кратко суммируем полученные для случае кратного перронова корня $r = 1$ свойства деревьев вывода t при $t \to \infty$.

1) Количество правил, примененных к нетерминалам второго класса, имеет такую же ($O(t^2)$) асимптотику, как и в неразложимом случае, а правил, примененных к первому нетерминалу, значительно меньше ($O(t)$).

2) Вероятность продолжения $Q_i(t)$ для $i \le k_1$ больше, чем $Q_i(t)$ для $i > k_1$.

3) Как и в неразложимом критическом случае, математическое ожидание количества применений правила r_{ij} в деревьях вывода пропорционально его вероятности p_{ij}, то есть, если за $M_j^i(t) = \sum_l M_{jl}^i(t)$ обозначить число нетерминалов A_j в дереве вывода из D_i^t, то $M_j^i(t)/M_{jl}^i(t) \to p_{jl}$ при $t \to \infty$.

4) Величины $M(S_{ij}(t))$ не меняются при замене аксиомы грамматики на любой нетерминал из класса K_1 .

I.1.3 Энтропия и стоимость оптимального кодирования

В этом параграфе мы будем рассматривать только КС-грамматики с однозначным выводом. Сначала выведем асимптотическую формулу для энтропии множества слов $\mathcal{L}^t$ грамматики, имеющих деревья вывода высоты t. По определению, $H(\mathcal{L}^t) = -\sum\limits_{\alpha\in L^t} p_t(\alpha) \cdot \log p_t(\alpha)$. Как и в докритическом случае, мы можем записать, что

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

Во всех случаях, в силу доказанных выше оценок для $P(L^t) = P_1(t)$, имеем $\log P(L^t) = O(\log t)$. Пользуясь асимптотиками для $M(S_{ij}(t))$, полученными в теореме I.1.3, получаем:

$$H(\mathcal{L}^t) = -\sum_{i\le k_1}\sum_{j=1}^{n_j}\log p_{ij}\cdot M(S_{ij}(t)) - \sum_{i>k_1}\sum_{j=1}^{n_j}\log p_{ij}\cdot M(S_{ij}(t)) + O(\log t)$$

$$= -\sum_{i>k_1} v''_{i-k_1}\cdot\sum_{j=1}^{n_j} p_{ij}\log p_{ij}\cdot B_2t^2\cdot(1+o(1))/4 + O(t)$$

$$= \sum_{i>k_1} v''_{i-k_1} H(R_i)\cdot B_2t^2\cdot(1+o(1))/4,$$

где

$$H(R_i) = -\sum_{j=1}^{n_j} p_{ij}\log p_{ij} \qquad \text{(I.1.33)}$$

- энтропия множества правил R_i.

Таким образом, справедлива следующая теорема:

Теорема I.1.4 *Для КС-грамматики с матрицей первых моментов вида (**??**) с* $r = r' = r'' = 1$, $B_1B_2 > 0$, *где* B_1, B_2 *введены формулами (I.1.14), энтропия множества слов* $\mathcal{L}^t$ *имеет вид*

$$H(\mathcal{L}^t) = \sum_{i>k_1} v''_{i-k_1} H(R_i)\cdot B_2t^2\cdot(1+o(1))/4,$$

где $H(R_i)$ *определено формулой (I.1.33).*

Теперь найдем стоимость оптимального кодирования для случая кратного перронова корня $r = 1$. Пусть f^* - кодирование множества слов $\mathcal{L}^t$, минимизирующее величину $M_t(f)$, т.е. $M_t(f) \ge M_t(f^*)$ для любого кодирования f, определенного и однозначного на множестве слов $\mathcal{L}^t$. Как и в докритическом случае, поскольку $H(\mathcal{L}^t) \to \infty$ при $t \to \infty$, то из теоремы **??** следует, что $M_t(f^*)/H(\mathcal{L}^t) \to 1$ при $t \to \infty$.

Рассмотрим величину $\sum_{\alpha\in L^t} p_t(\alpha)\cdot|\alpha|$. Пусть l_{ij} - количество терминальных символов в правой части правила r_{ij}. Ранее было показано, что

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

Подставляя полученные в теореме I.1.3 асимптотики для $M(S_{ij}(t))$, находим, что

$$\sum_{\alpha\in L^t} p_t(\alpha)\cdot|\alpha| = \sum_{i>k_1} v''_{i-k_1} L(R_i)\cdot B_2t^2(1+o(1))/4, \qquad \text{(I.1.34)}$$

где

$$L(R_i) = \sum_{j=1}^{n_j} l_{ij}p_{ij} \qquad \text{(I.1.35)}$$

- среднее число терминалов в правой части правил из R_i.

Как следствие теоремы I.1.4, замечания о том, что $M_t(f^*)/H(\mathcal{L}^t) \to 1$ при $t \to \infty$ и асимптотического равенства (I.1.34) верна следующая теорема:

Теорема I.1.5 *Пусть $\mathcal{L}$ - КС-язык, порожденный разложимой грамматикой с матрицей первых моментов вида (**??**) при $r = r' = r'' = 1$, $B_1B_2 > 0$. Тогда стоимость любого кодирования f этого языка удовлетворяет неравенству*

$$C(\mathcal{L}, f) \geq C^*(\mathcal{L}),$$

где

$$C^*(\mathcal{L}) = \frac{\sum_{i>k_1} v''_{i-k_1} H(R_i)}{\sum_{i>k_1} v''_{i-k_1} L(R_i)},$$

величины $H(R_i), L(R_i)$ введены формулами (I.1.33),(I.1.35), а B_1, B_2 - формулами (I.1.14).

Заметим, что в случае кратного перронова корня нижняя оценка стоимости оптимального кодирования определяется свойствами правил, применяемых к нетерминалам из класса K_2, поскольку в этом случае математическое ожидание количества нетерминалов из K_2 в дереве вывода имеет больший порядок, чем математическое ожидание количества нетерминалов из K_1.

I.1.4 Алгоритм оптимального кодирования

Можно показать, что полученная в предыдущем параграфе оценка для стоимости кодирования является точной, т. е. существует кодирование множества слов языка, стоимость которого равна $C^*(\mathcal{L})$. Такое кодирование строится следующим образом. Упорядочим слова языка в порядке невозрастания вероятностей, так что $p_1 \geq p_2 \geq \ldots$, где p_i-вероятность i-го слова. Поскольку

$$\sum_i 2^{-\lceil -\log p_i \rceil} \leq \sum_i 2^{\log p_i} = 1,$$

то из выполнения неравенства Мак-Миллана [22] следует, что существует взаимно-однозначное кодирование f_s множества слов языка, при котором длина кода i-го слова равна $\lceil -\log p_i \rceil$. Величина $M_t(f_s)$ для такого кодирования оценивается как

$$M_t(f_s) = \sum_{\alpha \in L^t} p_t(\alpha) \lceil -\log p(\alpha) \rceil$$

$$= -\sum_{\alpha \in L^t} p_t(\alpha) \log p_t(\alpha) - \Big(\log P(L^t) + O(1)\Big) \cdot \sum_{\alpha \in L^t} p_t(\alpha).$$

Поскольку

$$\log P(L^t) = O(\log t), \ \sum_{\alpha \in L^t} p_t(\alpha) = 1,$$

$$-\sum_{\alpha \in L^t} p_t(\alpha) \log p_t(\alpha) = H(\mathcal{L}^t) = O(t^2),$$

то справедливо равенство

$$C(f_s) = \lim_{t \to \infty} \frac{H(\mathcal{L}^t) + O(\log(t))}{\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha|} = C^*(\mathcal{L}).$$

Указанный алгоритм кодирования не является эффективным ввиду его большой вычислительной сложностью, которая обусловлена необходимостью упорядочивания слов языка по возрастанию вероятностей. Однако описанный в главе 3 эффективный алгоритм блочного кодирования f_{sh}, использованный в докритическом случае, будет асимптотически оптимальным и в данном случае при укрупнении грамматики $G(n)$ с $n \to \infty$, что доказывается аналогично неразложимому случае. Напомним вкратце идею доказательства. Избыточность кодирования f_{sh} может быть оценена как

$$\Delta(f_{sh}) = C(\mathcal{L}, f_{sh}) - C^*(\mathcal{L}) < \frac{\sum_{i,j} M(S_{ij}(t))\big(\lceil -logp_i j\rceil + \log p_i j\big)}{\sum_{\alpha \in L^t} p_t(\alpha) \cdot |\alpha|}$$

$$< \frac{\sum_{i \geq k_1} v''_{i-k_1} \sum_j p_{ij}(1 + o(1))}{\sum_{i \geq k_1} v''_{i-k_1} L(R_i)}.$$

Учитывая нормировку $\sum_{i \geq k_1} v''_{i-k_1} = 1$ и условие $\sum_j p_{ij} = 1$, получаем, что

$$\Delta(f_{sh}) < \frac{1}{\sum_{i \geq k_1} v''_{i-k_1} L(R_i)}.$$

Избыточность кодирования f_{sh} при использовании укрупненной грамматики $G(n)$ можно оценить как

$$\Delta(f_{sh}, n) < \frac{1}{\sum_{i \geq k_1} v''_{i-k_1} L(R'_i)},$$

где

$$L(R'_i) = \sum_{j,q=1}^{k} M(S^i_{jq}(n)) l_{jq}.$$

Из этого неравенства и ранее доказанных оценок $M(S^i_{jq}(n)) = O(n^2)$ следует, что $\Delta(f_{sh}, n) \to 0$ при $t \to \infty$.

случае.

I.2 Случай простого перронова корня

В этом параграфе будет исследована грамматика с двумя классами нетерминалов в случае, когда матрица первых моментов имеет вид (**??**), причем $r' \neq r''$, и $r = \max(r', r'') = 1$. Оказывается, что в этом случае, как и при $r < 1$, асимптотика для $Q_i(t), P_i(t)$, а также математических ожиданий числа применений правил $M(q_{ij}(t, \tau))$, $M(S_{ij}(t))$ на фиксированном ярусе τ дерева вывода из D^t_i и во всем дереве имеют такой же вид, как и в неразложимом случае. Доказательства проводятся аналогично неразложимому случаю [10], [16] с некоторыми небольшими отличиями, показанными далее. Случаи $r' < r''$ и $r' > r''$ рассматриваются отдельно, поскольку содержат некоторые отличия в доказательствах.

I.2.1 Вероятности продолжения и вероятности деревьев вывода фиксированной высоты

Заметим, что в случае $r' < r''$ верны соотношения $v = (0, v'')$, $u > 0, v'' > 0$, а при $r' > r''$ - соотношения $u = (u', 0)$, $v > 0, u' > 0$. Для величины $R(t, s) = (R_1(t, s), \ldots, R_k(t, s)) = 1 - F(t, s)$ справедливы следующие результаты.

Лемма I.2.1 *Для грамматики с двумя классами нетерминалов, имеющей матрицу первых моментов вида (**??**) при $r' < r'' = 1$ верна оценка*

$$R_n(t,s) = \frac{u_n \sum\limits_{i=k_1+1}^{k} v_i \cdot (1-s_i)}{1 + Bt \cdot \sum\limits_{i=k_1+1}^{k} v_i \cdot (1-s_i)/2} \cdot (1+o(1))$$

при $n = 1, \dots, k$, где B задано формулой

$$B = \sum_{i,j,l=1}^{k} v_i b^i_{jl} u_j u_l \tag{I.2.1}$$

и $B > 0$. Оценка является равномерной по $0 \le s \le 1, s(k_1+1:k) < 1$ в области $\max(1-s)/\min_{j>k_1}(1-s_j) < C$ для заданной константы $C > 0$

Этот результат доказывается точно так же, как и соответствующее утверждение в неразложимом случае [16]. Действительно, в равенстве

$$\bar{1} - F(t,s) = [A - E(F(t-1,s))][A - E(F(t-2,s))] \dots [A - E(1,s)](\bar{1} - s)$$

величины $A_t(s) = E_{jn}(F(t,s)) \le \sum_l b^l_{jn} \cdot (1 - F(l,0)) = A_t$ ограничены равномерно по s, причем $F_l(t,0) \to 1$ при $t \to \infty$. Поэтому равномерная сходимость

$$\frac{R_i(t,s)}{\sum\limits_{i=k_1+1}^{k} v_i R_i(t,s)} \to u_i$$

вытекает из лемм **??** и **??** и замечания в конце доказательства леммы **??** о равномерности оценок в области $\max(1-s)/\min_{j>k_1}(1-s_j) < C$ и $A_t(s) < A_t$, где $A_t \to 0$ при $t \to \infty$.

Лемма I.2.2 *Для грамматики с двумя классами нетерминалов, имеющей матрицу первых моментов вида (**??**), при $1 = r' > r''$ верны оценки*

$$R_n(t,s) = \frac{u_n \sum\limits_{i=1}^{k} v_i \cdot (1-s_i)}{1 + Bt \cdot \sum\limits_{i=1}^{k} v_i \cdot (1-s_i)/2} \cdot (1+o(1))$$

при $n = 1, \dots, k_1$, где $B > 0$ введено формулой (I.2.1). Для $n = k_1+1, \dots, k$ верны оценки $R_n(t,s) = O((r'')^t) \cdot \sum\limits_{i=k_1+1}^{k} v_i \cdot (1-s_i)$. Обе оценки выполнены равномерно по $0 \le s \le 1,\ s \ne 1$.

Оценки для $R_n(t,s),\ n = 1, \dots, k_1$ доказываются таким же образом с помощью лемм **??** и **??**, для $k_1 < n \le k$ следуют из оценки $R(t,s)(k_1+1:k) \le (A^{(3)})^t(\bar{1}-s) = O((r'')^t(\bar{1}-s))$ для неразложимой грамматики, порожденной нетерминалами из K_2.

Из этих двух лемм сразу же получаются оценки для величин $Q_i(t)$ и $P_i(t) = Q_i(t-1) - Q_i(t)$. Соотношение $\frac{P_i(t)}{\sum_i v_i P_i(t)} \to u_i$ получается из леммы **??** таким же образом, как и в неразложимом случае.

Сформулируем эти результаты в виде следующей теоремы.

Теорема I.2.1 *Для грамматики с двумя классами нетерминалов, имеющей матрицу первых моментов вида (**??**) при $r' \neq r''$, $r = 1$, для $i = 1, \ldots, k$ верны асимптотические равенства*

$$Q_i(t) = \frac{2u_i + o(1)}{Bt},$$

$$P_i(t) = \frac{2u_i + o(1)}{Bt^2},$$

где B задано формулой (I.2.1) и $B > 0$.

Заметим, что $u_i = 0$ при $r' > r''$ для $i > k_1$, поэтому указанные асимптотики в этом случае имеют вид $Q_i(t) = o(1/t)$, $P_i(t) = o(1/t^2)$. Поскольку при $i > k_1$ величины $Q_i(t), P_i(t)$ совпадают с соответствующими величинами для неразложимой грамматики, порожденной нетерминалами из K_2, то при $r' > r''$ более точные асимптотики для $Q_i(t), P_i(t)$ имеют вид $O((r'')^t)$.

I.2.2 Распределение нетерминалов на фиксированном ярусе дерева вывода

Для нахождения асимптотики величин $M(q_{ij}(t,\tau))$ нам необходимо исследовать распределение случайного вектора $x(t) = (x_1(t), \ldots, x_k(t))$ при условии $x(t) > 0$, где $x_i(t), i = 1, \ldots, k$ - число нетерминалов A_i на последнем ярусе в дереве вывода высоты t. Рассмотрим случайные величины $\xi_j^i(t) = \frac{2x_j^i(t)}{v_j Bt}$, где $j = 1, \ldots, k$ при $r' > r''$ и $j = k_1 + 1, \ldots, k$ при $r' < r''$, а $x_j^i(t)$ - количество нетерминалов A_i на ярусе t в деревьях вывода слов из L с корнем, помеченным нетерминалом A_j. Сформулируем теоремы, аналогичные теореме 6.3.3 из [16], для случаев $r' < r''$ и $r'' < r'$.

Теорема I.2.2 *Для грамматики с двумя классами нетерминалов, имеющей матрицу первых моментов вида (**??**) при $r' < r'' = 1$ условное распределение векторов $\xi''^i(t) = (\xi_{k_1+1}^i(t), \ldots, \xi_k^i(t))$, $i = 1, \ldots, k_1$ при условии $\xi^i(t) = (\xi_1^i(t), \ldots, \xi_k^i(t)) \neq 0$ сходится при $t \to \infty$ по распределению к случайному вектору $\xi'' = (\xi''_{k_1+1}, \ldots, \xi''_k)$, не зависящему от i, при этом*

$$P(\xi''_1 < x) = 1 - e^{-x}, \ x \geq 0,$$

и с вероятностью 1 справедливо равенство $\xi''_{k_1+1} = \xi''_{k_1+2} = \ldots = \xi''_k$.

Доказательство этого утверждения аналогично доказательству соответствующего утверждения для неразложимого случая. Рассмотрим преобразование Лапласа

$$\phi_t^i(\tau) = M\Big(\exp(-\sum_{j=k_1+1}^{k} \tau_j \xi_j''^i(t) | \xi^i(t) > 0)\Big) = 1 - \frac{R_i(t, s^\tau(t))}{Q_i(t)},$$

где $\tau = (\tau_{k_1+1}, \ldots, \tau_k)$, $s_j^\tau(t) = 0$ при $j \leq k_1$, и $s_j^\tau(t) = \exp\big(-2\tau_j/(v_j Bt)\big)$ при $j > k_1$. Заметим, что $1 - s_j^\tau(t) = (2 + o(1))\tau_j/(v_j Bt)$ при $j > k_1$ ввиду оценки $e^{-x} = 1 - x + O(x^2)$, $x > 0$. Следовательно, существует константа C, такая что $(1 - s_i^\tau(t))/(1 -$

$s_j^\tau(t)) < C,\ i,j = k_1+1,\dots,k$ для любого t. Поэтому, применяя оценки для $R_i(t,s)$, полученные в лемме I.2.1, получаем

$$R_n(t, s^\tau(t)) = \frac{2u_n\bar{\tau}/(Bt)}{1+2Bt\bar{\tau}/(2Bt)}\cdot(1+\varepsilon(t)) = \frac{2u_n\bar{\tau}}{Bt(1+\bar{\tau})}\cdot(1+\varepsilon(t)),$$

где $\bar{\tau} = \tau_{k_1+1}+\ldots+\tau_k$, а $\varepsilon(t)\to 0$ при $t\to\infty$. Поэтому, используя оценку $Q_n(t) = (2+o(1))u_n/Bt$, полученную в лемме I.2.1, имеем

$$\lim_{t\to\infty}\phi_t^i(\tau) = 1/(1+\bar{\tau}).$$

Поскольку полученная функция является преобразованием Лапласа для экспоненциального распределения, то, применяя результаты об однозначности соответствия между распределением и его преобразованием Лапласа (см. например [17]), получаем требуемое утверждение.

Для оценки количества нетерминалов из класса K_1 на ярусе дерева вывода нам потребуется следующая оценка для факториальных моментов $m_\alpha(t) = (m_\alpha^1(t),\dots,m_\alpha^k(t))$ порядка α.

Лемма I.2.3 *Для разложимой КС-грамматики с матрицей первых моментов, имеющей вид (**??**), при $r' < r'' = 1$ верна оценка $m_\alpha^i(t) = O({}_1(\bar{\alpha})(r')^t)$, если $i\le k_1$ и $\alpha_j\le k_1$ хотя бы для одного индекса j, и $m_\alpha^i(t) = O({}_2(\bar{\alpha})t^{\bar{\alpha}-1})$, если $\alpha > k_1$. Здесь через ${}_{1,2}(\bar{\alpha})$ обозначена константа, не зависящая от t, но, возможно, разная для разных значений $\bar{\alpha}$*

Лемма доказывается индукцией по $n = \bar{\alpha}$. При $n = 1$ верны оценки $a_j^i(t) = O((r')^t)$ при $i,j\le k_1$ и $a_j^i(t) = O(1)$ при $j > k_1$, т.к $(A^{(1)})^t = O((r')^t)$, где $A^{(1)}$ - верхний левый блок матрицы первых моментов A, и $A^t = O((r'')^t) = O(1)$. Пусть утверждение доказано для некоторого n. Докажем его для $\bar{\alpha} = n+1$.

Используем соотношение, доказанное в [16]:

$$m_\alpha^i(t) = \sum_{\tau=1}^{t}\sum_{s=1}^{k} a_s^i(t-\tau)z_\alpha^s(\tau-1), \tag{I.2.2}$$

где $z_\alpha^s(\tau-1)$ является суммой конечного числа произведений моментов более низких порядков, причем

$$z_\alpha^s(t) = m_\alpha^s(t+1) - \sum_n a_n^s m_\alpha^n(t). \tag{I.2.3}$$

Используя оценки $a_j^i(t) = O(1)$ при $i\le k_1,\ j > k_1$, из этой формулы и оценок $m_{\alpha_1}^i(t) = O(t^{\bar{\alpha}_1})$ для $\alpha_1 > k_1,\ \bar{\alpha}_1 < n+1$ сразу следует $m_\alpha^i(t) = O(t^n)$ для $\alpha > k_1,\ \bar{\alpha} = n+1$.

Из равенства (I.2.3) вытекает, что $z_\alpha^s(t) = 0$, если $n > k_1$ и если $\alpha_j\le k_1$ для некоторого индекса j. Поэтому можно считать, что в этом случае $n\le k_1$. Используя оценки для $m_\alpha^i(t)$ при $\bar{\alpha}\le n$, получаем

$$m_\alpha^i(t) = O((r')^t)\cdot\sum_{\tau=1}^{t}\tau^{n-1}(r')^\tau = O((r')^t).$$

Пусть $x(t) = (x_1(t),\dots,x_k(t))$ - случайная величина, где $x_i(t)$ - количество нетерминалов A_i на ярусе t в деревьях вывода слов языка. Из предыдущей леммы вытекает

Следствие I.2.1 *При $i \le k_1$ справедливы оценки*

$$M(x_i(t)) = O(t(r')^t),\ P(x_i(t) > 0) = O(t \cdot (r')^t).$$

Первое утверждение сразу следует из предыдущей леммы, поскольку

$$M(x_i(t)) = a_i^1(t)/Q(t) = O(t(r')^t),\ M(x_i(t)^2) = \left(b_{ii}^1(t) + a_i^1(t)\right)/Q(t) = O(t \cdot (r')^t).$$

Второе утверждение следует из неравенства Чебышева $P(|x - M(x)| > \varepsilon) < D(x)/\varepsilon^2$, примененного к случайной величине $x_i(t)$ с $\varepsilon = 1 - M(x_i(t))$, поскольку для такого ε неравенство $|x_i(t) - M(x_i(t))| < \varepsilon$ эквивалентно условию $x_i(t) = 0$.

Это утверждение показывает, что в случае $r' < r''$ в векторе $x(t)$ пропорция $x_i(t)/x_j(t),\ i \le k_1, j > k_1$ стремится к нулю при $t \to \infty$.

Теорема I.2.3 *Для разложимой грамматики, имеющей матрицу первых моментов вида (**??**) при $1 = r' > r''$, условное распределение векторов $\xi^i(t) = (\xi_1^i, \ldots, \xi_k^i)$, $i = 1, \ldots, k_1$, при условии $\xi^i(t) \ne 0$ сходится при $t \to \infty$ по распределению к случайному вектору $\xi = (\xi_1, \ldots, \xi_k)$, не зависящему от i. При этом*

$$P(\xi_1 < x) = 1 - e^{-x},\ x \ge 0,$$

и с вероятностью 1 верно равенство $\xi_1 = \xi_2 = \ldots = \xi_k$.

Доказательство этого утверждения полностью повторяет доказательство аналогичного утверждения для неразложимого случая [16], поскольку в нем не используется существенным образом неразложимость, а только асимптотики для $R(t, s)$ и $Q(t)$, полученные в леммах I.2.2 и I.2.1 которые имеют такой же вид, как и в неразложимом случае.

I.2.3 Математические ожидания числа применения правил в деревьях вывода

В этом параграфе мы получим оценки для $M(q_{ij}(t, \tau))$ и $M(S_{ij}(t))$ для случая $r' \ne r''$. Заметим, что асимптотики для $Q_i(t), P_i(t)$ имеют тот же вид, что и в неразложимом случае, и распределение нетерминалов на фиксированном ярусе дерева вывода имеет такой же вид. Поэтому теоремы 1,2,3 из [11] для неразложимого критического случая справедливы и в нашем случае, так как в них не используются предположения о неразложимости грамматики и строгой положительности собственных векторов, а нужны только асимптотики $Q_i(t), P_i(t)$ и утверждение о пропорциональном составе нетерминалов на ярусах дерева вывода (теоремы I.2.2 и I.2.3). Для случая $r' < r''$ влиянием первого нетерминала можно пренебречь ввиду следствия I.2.1.

Поэтому при $r'' \ne r'$ асимптотики величин $M(q_{ij}(t, \tau))$ и $M(S_{ij}(t))$ такие же, как и в неразложимом случае.

Теорема I.2.4 *Для разложимой стохастической КС-грамматики с матрицей первых моментов, имеющей вид (**??**) с перроновым корнем $r = 1$ и $r' \ne r''$, при $B > 0,\ t \to \infty$ выполняются асимптотические равенства:*

$$M(q_{ij}(t, \tau)) = (v_i + O(\varepsilon^2))p_{ij} \cdot B\tau(t - \tau)/t$$

при $\varepsilon t \le \tau \le (1-\varepsilon)t$ *и*

$$M(S_{ij}(t)) = p_{ij} v_i B t^2/6 + o(t^2),$$

где B *введено формулой (I.2.1).*

В случае $r' < r''$ влиянием нетерминалов из K_1 можно пренебречь ввиду оценок моментов, полученных в лемме I.2.3.

В случае $r' < r''$, когда $v(1:k_1) = 0$, могут быть получены более точные оценки для величин $M(q_{ij}(t,\tau))$ и $M(S_{ij}(t))$ при $i \le k_1$.

Пользуясь леммой I.2.3 и следствием I.2.1, оценим величины $M(x_i(t,\tau)), M(q_{ij}(t,\tau))$ при $i \le k_1$. Очевидно, число применений правила r_{ij} на ярусе τ меньше, чем $x_i(t)$. Поэтому

$$M(q_{ij}(t,\tau)) < M(x_i(t,\tau)) < \frac{p_{ij}}{P_1(t)} \cdot \sum_X P_X(\tau) x_i(\tau) R_X(t-\tau),$$

где $X = (x_1, \ldots, x_k)$. Доказанная для неразложимого критического случая в [11] оценка $R_X(t-\tau) = O(1/(t-\tau))$ верна и в нашем случае, поскольку при ее получении использовались только асимптотики для $Q_i(t), P_i(t)$. Поэтому справедливы оценки

$$M(x_i(t,\tau)) = O\left(t^2 r'^{\tau}/(t-\tau)\right),\ M(q_{ij}(t,\tau)) = O\left(t^2 r'^{\tau}/(t-\tau)\right).$$

Величина $M(S_{ij}(t))$ при $i \le k_1,\ r' < r''$ может быть оценена по аналогии с доказательством теоремы I.1.3 для случая кратного перронова корня.

Подставляя в соотношения (I.1.24),(I.1.26) оценки $Q_i(t) = o(1),\ Q_X(t) = 1 + o(1),\ P_i(t)/P_j(t) = u_i/u_j + o(1)$ и пользуясь формулой (I.1.27) для вычисления $R_X(t)$, получаем, что величины $\bar{M}_{ij}^q$ удовлетворяют следующим рекуррентным соотношениям при $i, j \le k_1$:

$$\bar{M}_{ij}^q(t) = \delta_q^i p_{ij} \cdot (1 + o(1)) + \sum_l \sum_{m=1}^{k_1} p_{ql} s_{ql}^m \bar{M}_{ij}^m(t-1)(1+o(1)) =$$

$$\delta_q^i p_{ij}(1+o(1)) + \sum_{m=1}^{k_1} a_m^q \bar{M}_{ij}^m(t-1)(1+o(1)), \tag{I.2.4}$$

аналогично для величин M_{ij}^q получаем:

$$M_{ij}^q(t) = \frac{1+o(1)}{u_q} \cdot \left(\delta_q^i p_{ij} \cdot \sum_{m=1}^{k} s_{ij}^m u_m + \sum_l \sum_{m \le k_1} p_{ql} s_{ql}^m \sum_{n=1}^{k} (s_{ql}^n - \delta_n^m) u_m \bar{M'}_{ij}^m(t-1) \right)$$

$$+ \sum_l \sum_{m=1}^{k_1} p_{ql} s_{ql}^m u_m M_{ij}^m(t-1)$$

$$= \frac{1+o(1)}{u_q} \cdot \left(\delta_q^i p_{ij} S_{ij} + \sum_{m=1}^{k_1} G_m^q \bar{M}_{ij}^m(t-1) + \sum_{m \le k_1} a_m^q u_m M_{ij}^m(t-1) \right), \tag{I.2.5}$$

где обозначено

$$S_{ij} = \sum_{m=1}^{k} s_{ij}^m u_m,\ G_m^q = \sum_{n=1}^{k} b_{mn}^q u_n.$$

Для решения этих рекуррентных соотношений нам потребуется следующая лемма :

Лемма I.2.4 *Пусть $A_i \to A$ при $i \to \infty$ - последовательность матриц размером $k \times k$, причем матрица $A > 0$ и ее перронов корень $0 < r < 1$. Пусть $b_i \to b$ при $i \to \infty$ - последовательность векторов размером k, где $b \geq 0$, $b \neq 0$. Тогда для любого вектора x_0 последовательность $x_i, i = 0, 1, \ldots$, определяемая рекуррентным соотношением $x_t = b_t + A_t x_{t-1}$, сходится к положительному вектору $x_* = (E - A)^{-1} b$.*

Идея доказательства такая же как в лемме I.1.4. Докажем, что для любого $\varepsilon > 0$ существует номер n_0, такой, что $|x_* - x_n| < \varepsilon$ при $n > n_0$. Сначала докажем, что x_n ограничено некоторой константой C. Возьмем такой номер n_1, что $A_n < A/r_1$, $b_n < b/r_1$ для некоторого $r_1 > r, r_1 < 1$ при $n \geq n_1$. Используя обозначение $A^*(t,n) = A_t A_{t-1} \ldots A_{n+1}$ при $t > n$, и $A^*(t,n) = E$ при $t = n$, можем записать, что

$$x_{t+n_1} = A^*(t+n_1, n_1) x_{n_1} + \sum_{m=1}^{t} A^*(t+n_1, m+n_1) b_{m+n_1}. \qquad \text{(I.2.6)}$$

Из этого соотношения сразу следует, что величины x_t ограничены, поскольку $A^*(t + n_1, m + n_1) = O((r/r_1)^{t-m})$ при $m < t$, то есть можно считать, что $x(t) < C$ для некоторой константы $C > 0$.

Матричный ряд $\sum\limits_{m=0}^{\infty} A^m$ сходится (поэлементно), поскольку его члены положительны и имеют порядок $O(r^m)$. Прямой проверкой показывается, что его сумма дает обратную матрицу к $E - A$ (которая в силу этого рассуждения положительна). Зададим произвольно малые величины $0 < \varepsilon_1, \varepsilon_2 < 1$. Возьмем такое m_0, что сумма остатка этого ряда $\sum\limits_{m=m_0}^{\infty} A^m < \varepsilon_1$, и $A^m/(r_1)^m < \varepsilon_1$ при $m \geq m_0$. Возьмем такое $n_2 > n_1$, что $|b_n - b| \leq \varepsilon_2 b$, и $|A_n - A| < \varepsilon_2 A$ при $n \geq n_2$. Тогда, пользуясь равенством (I.2.6), получаем, что при $n \geq n_2$ справедливо неравенство

$$|x_{m_0+n} - x_*| = \Big| A^*(m_0+n, n) x_n + \sum_{m=0}^{m_0} \big(A^*(m_0+n, m+n) b_{m+n} - A^{m_0-m} b \big)$$

$$+ \left(\sum_{i=0}^{m_0} A^i b - (E-A)^{-1} b \right) \Big|.$$

Из неравенства $|A - A_n| < \varepsilon_2 A$ и положительности A_n при $n > n_2$ следуют оценки

$$|A^*(m_0+n, n) x_n| < A^{m_0} k \cdot (1+\varepsilon_2)^{m_0} \max x_n < kC \cdot (1+\varepsilon_2)^{m_0} \varepsilon_1,$$

$$\Big| \sum_{i=0}^{m_0} A^i b - (E-A)^{-1} b \Big| = \sum_{i > m_0} A^m b < k \max(b) \varepsilon_1,$$

$$|A^*(m_0+n, m+n) b_{m+n} - A^{m_0-m} b| \leq \big((1-\varepsilon_2)^{-m_0} - 1 \big) \cdot A^{m_0-m} b,$$

Поэтому

$$|x_{m_0+n} - x_*| < kC \cdot (1+\varepsilon_2)^{m_0} \varepsilon_1 + \big((1-\varepsilon_2)^{-m_0} - 1 \big) \cdot \max \Big(\sum_{i=0}^{m_0} A^i b \Big) + k \max(b) \varepsilon_1.$$

Поскольку элементы матриц $A^i = O(r^i)$ ограничены сверху некоторой константой C_1, то

$$|x_{m_0+n} - x_*| < \Big(C\varepsilon_1(1+\varepsilon_2)^{m_0(\varepsilon_1)} + \max(b)\varepsilon_1 + \big((1-\varepsilon_2)^{-m_0(\varepsilon_1)} - 1\big)m_0(\varepsilon_1)C_1\max(b)\Big)k.$$

Устремляя сначала ε_1, затем ε_2 к нулю, получаем соответствующие номера m_0, n_2, такие что $|x_* - x_{m_0+n}| < \varepsilon$ для любого сколь угодно малого $\varepsilon > 0$ при $n > n_2$. Лемма доказана.

Решая рекуррентные соотношения (I.2.4) для $\bar{M}^q_{ij}(t)$ с использованием леммы I.2.4, получаем, что при $i \le k_1,\ t \to \infty$

$$\bar{M}_{ij}(t) \to (\bar{M}^1_{ij}(t), \ldots, \bar{M}^{k_1}_{ij}(t)) = p_{ij} \cdot (1 + o(1)) \cdot (E - A^{(1)})^{-1}\delta^i = \bar{M}^*_{ij},$$

и соответственно для величин $M_{ij}(t) = \big(u_1 M^1_{ij}(t), \ldots, u_{k_1} M^{k_1}_{ij}(t)t\big)$ из рекуррентных соотношений (I.2.5) с помощью I.2.4 находим асимптотику

$$M_{ij}(t) = p_{ij} \cdot (1 + o(1)) \cdot \left(S_{ij} + \sum_m G^q_m \bar{M}^{*m}_{ij}\right) \cdot (E - A^{(1)})^{-1}\delta^i.$$

Таким образом, верна следующая:

Теорема I.2.5 *Для КС-грамматики с матрицей первых моментов вида (??), при $r' < r'' = 1,\ B > 0$, где B дано формулой (I.2.1), при $i \le k_1$ и $t, \tau \to \infty$ имеют место оценки:*

$$M(x_i(t,\tau)) < C_1 \cdot \Big(t^2 (r')^\tau / (t - \tau)\Big),$$

$$M(q_{ij}(t,\tau)) < C_2 \cdot \Big(t^2 (r')^\tau / (t - \tau)\Big).$$

для некоторых $C_1, C_2 > 0$. Кроме того, для $i \le k_1$ при $t \to \infty$

$$M(\bar{S}^q_{ij}(t)) \to p_{ij} \cdot (1 + o(1)) z^i_q,$$

$$M(S^q_{ij}(t)) \to p_{ij} \cdot \left(S_{ij} + \sum_m G^q_m p_{ij} z^i_m\right) z^i_q (1 + o(1))/u_q,$$

где обозначено

$$S_{ij} = \sum_{m=1}^{k} s^m_{ij} u_m,\ G^q_m = \sum_{n=1}^{k} b^q_{mn} u_n,\ z^i = (E - A^{(1)})^{-1}\delta^i.$$

Из этой теоремы видно, что в случае $0 < r' < r'' = 1$ в дереве вывода на каждом ярусе нетерминалов из класса K_2 значительно больше (O(t)), чем из K_1. Кроме того, общее число применений правил $r_{ij},\ i \le k_1$ в деревьях вывода высоты t стремится к константе.

I.2.4 Алгоритм оптимального кодирования

Поскольку асимптотика для величин $M(S_{ij}(t))$ и $P(D^t)$ для случая $r' \neq r''$ такая же, как и в неразложимом случае, то выражения для энтропии $H(\mathcal{L}^t)$ множества слов $\mathcal{L}^t$ и нижней оценки стоимости оптимального кодирования $C^*(\mathcal{L})$ имеют такой же вид, как и в неразложимом критическом случае, причем схема кодирования, предложенная для случая кратного перронова корня в предыдущем разделе, является оптимальной и здесь, так как $H(\mathcal{L}^t) = O(t^2), P(t) = O(t^2)$. Эффективный алгоритм асимптотически оптимального при укрупнении правил грамматики кодирования, описанный в главе 3, будет асимптотически оптимальным и в этом случае.

Можно заметить, что в случае $r' < r''$ оценки для $H(\mathcal{L}^t)$ и $C^*(\mathcal{L})$ имеют тот же вид, что и для грамматики, порожденной только нетерминалами из K_2 (это следует из того, что в этом случае $v(1 : k_1) = 0$) - как и для случая кратного перронова корня.

Часть II

Заключение

А это заключение.

Список литературы

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

[2] Борисов А.Е. Кодирование слов стохастического КС- языка, порожденного разложимой грамматикой с двумя нетерминалами // Вестник Нижегородского университета им. Н.И. Лобачевского. Серия Математика вып. 1(2), 2004. С. 18-28.

[3] Борисов А.Е. Закономерности в деревьях вывода для стохастической разложимой КС - грамматики.// Труды V международной научной конференции "Дискретные модели в теории управляющих систем". М.: Изд. отдел ВМиК МГУ, 2003. С. 15-17

[4] Борисов А.Е. О свойствах стохастического КС-языка, порожденного разложимой грамматикой.// Материалы научной конференции "Синтез и сложность управляющих систем". Н. Новгород, 2003. С. 15-18

[5] Борисов А.Е. О свойствах слов языка, порожденного разложимой стохастической КС-грамматикой с двумя нетерминалами. Критический случай.// Материалы VIII Международного семинара "Дискретная математика и ее приложения". М.: Изд. мех-мат. ф-та МГУ, 2004. С. 408-410.

[6] Борисов А.Е. О свойствах стохастического КС-языка, порожденного грамматикой с двумя классами нетерминальных символов.// Дискретный

анализ и исследование операций. Серия 1, т.12, N3. Новосибирск: Издательство Института математики СО РАН, 2005. С.3-31.

[7] Борисов А.Е. О числе применений правил стохастической КС-грамматики.// Тезисы докладов XIV Международной конференции "Проблемы теоретической кибернетики". Изд. мех-мат. ф-та МГУ, 2005. С. 22.

[8] Гантмахер Ф.Р. Теория матриц. М.: Наука, 1967.

[9] Жильцова Л.П. Закономерности применения правил грамматики в выводах слов стохастического контекстно-свободного языка // Математические вопросы кибернетики. Вып.9. М.: Наука, 2000. С.101- 126.

[10] Жильцова Л.П. Закономерности в деревьях вывода слов стохастического контекстно-свободного языка и нижняя оценка стоимости кодирования. Критический случай.// Дискретный анализ и исследование операций. Серия 1, т.10, N3. Новосибирск: Издательство Института математики СО РАН, 2003. С.23-53.

[11] 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, N3. Новосибирск: Издательство Института математики СО РАН, 2001. С.26- 45.

[12] Кричевский Р.Е. Сжатие и поиск информации. М.:Радио и связь, 1989.

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

[14] Марков А.А., Смирнова Т.Г. Алгоритмические основания обобщенно-префиксного кодирования.// Доклады АН СССР,т. 274 N4, С.790-793, 1984.

[15] Марков А.А. О неукоторых мерах сложности м эффективности в алфавитном кодировании. Матем. вопросы кибернетики вып. 6, с.348-352, М.: Наука, Физматгиз, 1996.

[16] Севастьянов В.А. Ветвящиеся процессы. М.: Наука, 1971.

[17] Феллер В. Введение в теорию вероятностей и ее приложения, том 1. - М.:Мир, 1984.

[18] Фихтенгольц Г.М. Основы математического анализа. Rnl 2. М.: Наука, 1968.

[19] Фу К. Структурные методы в распознавании образов. М.: Мир, 1977.

[20] Харрис Т. Теория ветвящихся случайных процессов. М.: Мир, 1966.

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

[22] Яблонский С.В. Введение в дискретную математику. М.:Наука, 1986.

[23] Сагитов С., Ватутин В.А. Разложимый критический ветвящийся процесс с двумя типами частиц.// Вероятностные проблемы дискретной математики. Труды мат.инст. Стеклов. 177 (1986), стр. 3-20.

Докл.

[24] Сагитов С., Ватутин В.А. Разложимый критический ветвящийся процесс Беллмана-Харриса с двумя типами частиц. I.//Теор. Вероятности. 33(1988), N3, стр. 460-472.

[25] Сагитов С., Ватутин В.А. Разложимый критический ветвящийся процесс Беллмана-Харриса с двумя типами частиц.II.//Теор. Вероятности. 34(1989), N2, стр. 216-227.

[26] Borisov A.E. Optimal coding cost for stochastic CF- language induced by decomposable grammar.// VI International Conference on Mathematical Modeling/Book of abstracts, N.Novgorod, 2004. pp. 72.

[27] Ziv J.,Lempel A. Compression of individual sequences via variable-rate coding. // IEEE Trans.Inf.Theory IT-24,5 (Sept. 1978), p.530-536.

[28] Ziv J.,Lempel A. A universal algorithm for sequential data compression. IEEE Trans. Inf. Theory IT 23,3 (1977), p.337-343.

[29] Jorma Rissanen, Glen G.Langdon. Universal modeling and coding. // IEEE Transactions on Information Theory, V.21, N 1,pp 12-23,1981.

[30] Jorma Rissanen. Ggeneralized Kraft inequality and arithmetic coding. // IBM Journal Res.Develop, 1976. V.20, N3, p.198-203.

[31] Haffman D.A. A method for construction of minimum-redundancy codes// Proc. IRE 1952, V.40, N.10, p1098-1101.

[32] Zhiltsova L. On Entropy and Optimal Coding Cost for Stochastic Language.// Fundamenta Informaticae,V.36,pp.285-305,1998.