Министерство высшего и среднего специального
образования Латвийской ССР
Латвийский ордена Трудового Красного Знамени
государственный университет имени Петра Стучки
Вычислительный центр

ТЕОРИЯ АЛГОРИТМОВ И ПРОГРАММ

Выпуск III

Республиканский межведомственный сборник
научных трудов

Латвийский государственный университет им.П.Стучки
Рига, 1977

Probabilistic program synthesis

K.M.Podnieks

The program computing φ is guessed from given values $\varphi(0), \ldots, \varphi(m)$. As $m \to \infty$, the hypothetic programs h_m may converge to the limit-program really computing φ. If $\tau(n,x)$ is a recursive function, we say: the program n computes the function $\tau_n = \lambda x\, \tau(n,x)$. For any τ a probabilistic strategy M_τ can be constructed such that for all n in the case of τ_n: 1) the M_τ-hypotheses converge to a τ-program of τ_n almost surely, 2) the number of alternations in the sequence is $\lesssim \ln n$ with probability $\to 1$ as $n \to \infty$. The estimate $\ln n$ is the best possible. Deterministic strategies do the same with the estimate $\sim n$.

ВЕРОЯТНОСТНЫЙ СИНТЕЗ ПРОГРАММ

К.М.Подниекс
ВЦ ЛГУ им. П.Стучки

По данным значениям $\varphi(0), \varphi(1), \dots, \varphi(m)$ требуется восстановить программу, вычисляющую функцию φ. Задача считается решенной, если последовательность программ-гипотез $\{h_m\}$ стабилизируется на верной гипотезе.

§ I. Аппарат

Следуя [I], методом программирования будем называть любую 2-местную ч.р. функцию $\tau(n,x)$. Программа n "вычисляет" функцию $\tau_n(x)$. В этой статье рассматриваются только о.р. методы программирования, т.е. по существу, нумерованные классы о.р. функций. Нумерованный класс $(\mathcal{U},\tau)$ состоит из эффективно перечислимого класса $\mathcal{U}$ о.р. функций вместе с некоторой вычислимой нумерацией τ класса $\mathcal{U}$. Понятие τ-программы и τ-номера данной функции тождественны.

Определение вероятностных стратегий (В-стратегий) см. в [2]. В случае синтеза программ (τ-номеров) случайная величина $M(\varphi,m)$ соответствующая согласно В-стратегии M начальному куску $\varphi(0), \dots, \varphi(m)$, именуется уже не прогнозом, а гипотезой. Гипотеза n считается верной (относительно функции φ и нумерации τ), если $\tau_n=\varphi$, т.е. если n является τ-программой φ. Гипотеза $M(\varphi,m)$, как правило, верна лишь с некоторой вероятностью.

Пусть В-стратегия M определена над вероятностным пространством (Ω, B, P). Осуществление любого элементарного события $w\in\Omega$ превращает гипотезу $M(\varphi,m)$ в некоторое натуральное число или в ∞. Каждой функции φ и элементарному событию $w\in\Omega$ соответствует поэтому вполне определенная последовательность гипотез В-стратегии M.

На пространстве (Ω, B, P) определены следующие вероятности:

(а) вероятность $P_m(M, \varphi)$ того, что на шаге m $(m \geqslant -1)$ стратегия M исправляет гипотезу, т.е.

$$P_m(M, \varphi) = P\left\{M(\varphi, m) \neq M(\varphi, m+1)\right\};$$

(б) вероятность $P\left\{M, \varphi, \leqslant \kappa\right\}$ того, что на функции φ стратегия M сделает не более κ исправлений гипотезы;

(в) вероятность $P\left\{M, \varphi, \mathcal{T}\right\}$ того, что последовательность гипотез В-стратегии M, выданная на функции φ, стабилизируется на $\mathcal{T}$ -программе φ.

Приведем одно достаточное условие для $P\left\{M, \varphi, \mathcal{T}\right\} = 1$. Будем говорить, что В-стратегия M $\mathcal{T}$-нормальна на функции φ, если для всех m, n таких, что

$$n = \infty \ \vee (\exists x \leqslant m)\ \mathcal{T}_n(x) \neq \varphi(x),$$

из $P\left\{M(\varphi, m) = n\right\} > 0$ следует

$$P\left\{(\forall_\kappa^\infty)\ M(\varphi, \kappa) \neq n \mid M(\varphi, m) = n\right\} = 1.$$

Грубо говоря, нормальная В-стратегия с вероятностью 1 "убирает" всякую гипотезу, неверность которой стала очевидной.

ПРЕДЛОЖЕНИЕ I. Если В-стратегия M $\mathcal{T}$-нормальна на функции φ, то $\sum\limits_m P_m(M, \varphi) < \infty$ влечет $P\left\{M, \varphi, \mathcal{T}\right\} = 1$.

ДОКАЗАТЕЛЬСТВО. По лемме Бореля-Кантелли [3], условие

$$\sum_m P_m(M, \varphi) < \infty$$

гарантирует, что с вероятностью 1 гипотезы M на φ стабилизируются, начиная с некоторого места. Из определения нормальности следует, что стабилизация на ∞ или на неверном $\mathcal{T}$-номере возможна лишь с вероятностью 0.

Предложение I доказано.

Определение рекурсивных, финитных на U, всюду финитных В-стратегий см. в [2]. Следуя Я.М.Барздиню, В-стратегию M будем называть $\mathcal{T}$ - регулярной, если для всех $\varphi \in U$ и всех m, n условие $P\left\{M(\varphi, m) = n\right\} > 0$ влечет

$$n \in N \wedge (\forall x \leqslant m)\ \mathcal{T}_n(x) = \varphi(x).$$

Таким образом, $\mathcal{T}$ -регулярная В-стратегия на функции $\varphi \in U$ всегда с вероятностью 1 выдает гипотезу, согласованную с имеющейся информацией о функции φ. Всякая $\mathcal{T}$-регулярная В-стратегия $\mathcal{T}$ -нормальна на любой функции $\varphi \in U$.

В § 3,4 нам потребуются два вспомогательных утверждения.

ПРЕДЛОЖЕНИЕ 2. Пусть L - не более чем конечное множество натуральных чисел и пусть $K(L)$ -множество всех кортежей $(i_1 \ldots i_s)$ таких, что

$$(i_1 < i_2 < \cdots < i_s) \wedge (\forall x \leqslant s)\ i_x \in L,$$

включая сюда и пустой кортеж Λ. Тогда для любой числовой последовательности $\{x_n\}$ имеет место равенство

$$\sum_{\tau \in K(L)} (x_{i_1} - 1) \ldots (x_{i_s} - 1) = \prod_{j \in L} x_j.$$

ПРЕДЛОЖЕНИЕ 3. Пусть $\{x_n(t)\}$ - последовательность независимых случайных величин, принимающих только значения 0,1 и зависящих от какого-либо параметра t. Обозначим $\sum\limits_n P\left\{x_n(t) = 1\right\} = a(t)$.

(I) Если в каком-либо процессе изменения t: $a(t) \leqslant b(t) < \infty$, где $b(t) \to \infty$, то

$$P\left\{\sum_n x_n \leqslant b(t) + \sqrt{b(t) \log b(t)}\right\} \to 1.$$

(2) Если $a(t) \geqslant b(t) < \infty$ и $b(t) \to \infty$, то

$$P\left\{\sum_n x_n \geqslant b(t) - \sqrt{b(t) \log b(t)}\right\} \to 1.$$

ДОКАЗАТЕЛЬСТВО легко вытекает из неравенства Чебышева.

§ 2. Результаты

Первым результатом о вероятностном синтезе τ -номеров является следующая теорема Я.М.Барздиня и Р.В.Фрейвалда [4].

Для любого нумерованного класса $(\mathcal{U}, \tau)$ можно построить рекурсивную В-стратегию M такую, что $P\{M, \varphi, \tau\} = 1$ для всех $\varphi \in \mathcal{U}$ и при этом, если $n \to \infty$, то

$$P\{M, \tau_n, < \log_2 n\} \to 1.$$

Аналогичная оценка для детерминированных стратегий имеет порядок n (см. [4,5]).

В настоящей статье показывается, что для В-стратегий оценку $\log_2 n$ можно улучшить до $\ln n$ (т.е. приблизительно до $0.69 \log_2 n$) и что эта новая оценка асимптотически точна. Таким образом, если для восстановления программ (по данным значениям функции) используются В-стратегии, то каждое исправление гипотезы восстанавливает в среднем 1.45 битов программы.

ТЕОРЕМА I. Для любого нумерованного класса $(\mathcal{U}, \tau)$ можно построить рекурсивную В-стратегию M такую, что $P\{M, \varphi, \tau\} = 1$ для всех $\varphi \in \mathcal{U}$ и при этом, если $n \to \infty$, то

$$P\{M, \tau_n, \leqslant \ln n + \sqrt{\log n}\, \log \log n\} \to 1.$$

Стратегию M можно сделать либо всюду финитной, либо τ -регулярной и финитной на $\mathcal{U}$.

ТЕОРЕМА 2. Можно построить нумерованный класс $(\mathcal{U}, \tau)$ такой, что для любой рекурсивной В-стратегии M, если $P\{M, \varphi, \tau\} = 1$ для всех $\varphi \in \mathcal{U}$, то существует возрастающая последовательность $\{n_k\}$ такая, что при $k \to \infty$:

$$P\{M, \tau_{n_k}, \leqslant \ln n_k - \sqrt{\log n_k}\, \log \log n_k\} \to 0.$$

Эти теоремы доказываются в § 3,4,5,6.

§ 3. Доказательство теоремы I

Верхняя оценка $\ln n$ получается с помощью В-стратегии Барздиня-Фрейвалда, для которой ими получена оценка $\log_2 n$ (см. [4]). Предлагаемый новый метод оценки возник при рассмотрении несколько более общей ситуации, в которой игнорируется вычислимость.

Пусть $(\mathcal{U}, \tau)$ – произвольный (счетный) нумерованный класс всюду определенных функций типа $N \to N$. Возьмем некоторое распределение вероятностей на N

$$\pi = \{\pi_n \mid n \in N\},$$

где $\pi_n > 0$ для всех n и $\sum \pi_n = 1$. Через $\text{БФ}_{\tau\pi}$ обозначим следующую В-стратегию.

Определение гипотезы $\text{БФ}_{\tau\pi}(\varphi, m)$ при условии $\text{БФ}_{\tau\pi}(\varphi, m-1) = n$. Если $n = \infty$, то и $\text{БФ}_{\tau\pi}(\varphi, m) = \infty$. Если $n \in N \wedge \tau_n(m) = \varphi(m)$ (т.е. гипотеза n еще не устарела), то $\text{БФ}_{\tau\pi}(\varphi, m) = n$ с вероятностью 1. Если же $\tau_n(m) \neq \varphi(m)$, берется сумма $\mathfrak{S} = \sum\{\pi_i \mid i \in \text{Э}\}$, где Э – множество всех подходящих (пока) гипотез, т.е.

$$\text{Э} = \{i \mid (\forall x \leqslant m)\, \tau_i(x) = \varphi(x)\}.$$

Если $\mathfrak{S} = 0$, то $\text{БФ}_{\tau\pi}(\varphi, m) = \infty$. Если же $\mathfrak{S} > 0$ то полагаем для всех $i \in \text{Э}: \text{БФ}_{\tau\pi}(\varphi, m) = i$ с вероятностью $\pi_i / \mathfrak{S}$.

Легко видеть, что в случае, когда $(\mathcal{U}, \tau)$ – нумерованный класс о.р. функций (с вычислимой нумерацией), а $\{\pi_n\}$ – рекурсивная последовательность конструктивных действительных чисел, то $\text{БФ}_{\tau\pi}$ – рекурсивная В-стратегия.

Грубо говоря, "тактика" стратегии $\text{БФ}_{\tau\pi}$ состоит в следующем. Она "считает", что ей подаются значения функции φ, которая "случайно" (в соответствии с распределением π) выбрана из нумерации τ. Гипотезу, которая согласуется с появляющимися новыми данными о функции φ, стратегия не старается исправлять. Но если эта гипотеза приходит

дит в противоречие с новыми данными, стратегия $Б\Phi_{\mathcal{T}\mathcal{R}}$ вычисляет вероятность $\mathcal{R}_i/\mathcal{G}$ того, что подается функция $\mathcal{T}_i$ из тех, которые неотличимы от φ на основе полученных до сих пор данных. Новая гипотеза стратегии $Б\Phi_{\mathcal{T}\mathcal{R}}$ получается имитацией вычисленного распределения $\{\mathcal{R}_i/\mathcal{G}\}$.

В-стратегия $Б\Phi_{\mathcal{T}\mathcal{R}}$, очевидно, $\mathcal{T}$-регулярна. Поэтому для доказательства того, что $P\{Б\Phi_{\mathcal{T}\mathcal{R}}, \varphi, \mathcal{T}\}=1$ для всех $\varphi\in U$, достаточно установить для всех $n\in N$ конечность суммы $\sum_m P_m(Б\Phi_{\mathcal{T}\mathcal{R}}, \mathcal{T}_n)$.

ЛЕММА I. $\sum_m P_m(Б\Phi_{\mathcal{T}\mathcal{R}}, \mathcal{T}_n) \leqslant \ln\frac{1}{\mathcal{R}_n}$.

Кроме того, нам нужна будет следующая

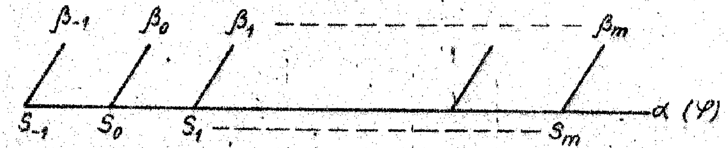
ЛЕММА 2. Пусть фиксирована функция $\varphi\in U$. Тогда независимы события

$$A_m = \{Б\Phi_{\mathcal{T}\mathcal{R}}(\varphi,m)\ne Б\Phi_{\mathcal{T}\mathcal{R}}(\varphi,m+1)\}.$$

Любопытно отметить, что события $\{A_m\}$ (т.е. "стратегия $Б\Phi_{\mathcal{T}\mathcal{R}}$ на шаге m исправляет гипотезу") не обнаруживают никаких содержательных признаков независимости. Формальному критерию независимости они, тем не менее, удовлетворяют, и этого достаточно для применения результатов теории вероятностей.

Сначала докажем лемму 2, после этого – лемму I.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ 2. Рассмотрим дерево функций, до некоторого места совпадающих с φ.



Каждой точке этого дерева можно приписать вероятность того, что произвольная функция ψ, выбранная из нумерации $\mathcal{T}$ в соответствии с распределением $\mathcal{R}$, пройдет через эту

точку. Пусть β_m – вероятность того, что ψ пройдет через точку S_m в сторону от функции φ. Через B_m обозначим сумму

$$B_m = \beta_m + \beta_{m+1} + \dots .$$

Тогда точке S_m приписана вероятность $\alpha + B_m$, где α вероятность, приписанная всей ветке φ.

Через β_{ij} $(i\geqslant -1,\ j\geqslant i)$ обозначим вероятность того, что в случае исправления гипотезы в точке S_i стратегия $Б\Phi_{\mathcal{T}\mathcal{R}}$ направит свою новую гипотезу через точку S_j в сторону от φ. Очевидно

$$\beta_{ij} = \frac{\beta_j}{\alpha + B_i} \qquad (I)$$

Вероятность исправления гипотезы в точке S_m выражается через числа β_{ij}:

$$P\{A_m\} = \sum \beta_{-1,i_1}\cdot\beta_{i_1+1,i_2}\cdot\ldots\cdot\beta_{i_k+1,m},$$

где суммирование идет по всем кортежам $(i_1 \dots i_k)$ таким, что $k\geqslant 0$ и $-1\leqslant i_1<i_2<\dots<i_k<m$.

Вероятность совместных исправлений гипотезы в точках $S_{m_1},\dots,S_{m_t}$ (где $m_1<m_2<\dots<m_t$) можно выразить аналогично

$$P\{A_{m_1}\cap\dots\cap A_{m_t}\} = \sum \beta_{-1,i_1}\cdot\beta_{i_1+1,i_2}\cdot\ldots\cdot\beta_{i_k+1,m_t},$$

где суммирование идет по всем кортежам $(i_1 \dots i_k)$ таким, что $k\geqslant 0$, $-1\leqslant i_1<i_2<\dots<i_k<m_t$ и среди чисел i_x встречаются все числа $m_1,\dots,m_{t-1}$. В силу (I), вероятность $P\{A_{m_1}\cap\dots\cap A_{m_t}\}$ зависит только от чисел $\alpha, \beta_{-1}, \beta_0, \dots, \beta_m$ и $B_{m+1}=\beta_{m+1}+\dots$, где $m=m_t=\max m_i$. Введем новые переменные γ_j $(-1\leqslant j\leqslant m+1)$:

$$\alpha + B_{m+1} = \alpha\gamma_{m+1}$$

$$\beta_m = \alpha(\gamma_m - 1)\gamma_{m+1}$$

$$\alpha + B_m = \alpha\gamma_m\gamma_{m+1}$$

$$\beta_{m-1} = \alpha(\gamma_{m-1} - 1)\gamma_m \gamma_{m+1}$$

$$\alpha + B_{m-1} = \alpha \gamma_{m-1} \gamma_m \gamma_{m+1}$$

$$\alpha + B_j = \alpha \prod_j^{m+1} \gamma \qquad \beta_j = \alpha(\gamma_j - 1) \prod_{j+1}^{m+1} \gamma$$

Здесь $j = -1, 0, \ldots, m$. Таким образом,

$$\beta_{ij} = \frac{\beta_j}{\alpha + B_i} = \frac{\gamma_j - 1}{\prod_i^j \gamma},$$

$$\beta_{-1, i_1} \beta_{i_1+1, i_2} \cdots \beta_{i_k+1, m} \quad \frac{(\gamma_{i_1} - 1) \ldots (\gamma_{i_k} - 1)(\gamma_m - 1)}{\prod_{-1}^m \gamma}$$

Вынесем за скобки произведение $(m_t = m)$:

$$(\gamma_{m_1} - 1) \ldots (\gamma_{m_t} - 1) \frac{1}{\prod_{-1}^m \gamma}.$$

В выражении для $P\{A_{m_1} \cap \ldots \cap A_{m_t}\}$ тогда суммируются произведения

$$(\gamma_{i_1} - 1) \cdots (\gamma_{i_k} - 1)$$

по всем кортежам $(i_1 \ldots i_k) \in K(L)$, где

$$L = \{-1, 0, 1, \ldots, m\} - \{m_1, \ldots, m_t\}$$

(определение $K(L)$ см. в § I, предложение 2). По предложению 2, сумма эта равна $\prod_{i \in L} \gamma_i$, т.е.

$$P\{A_{m_1} \cap \ldots \cap A_{m_t}\} = (\gamma_{m_1} - 1) \ldots (\gamma_{m_t} - 1) \frac{\prod_{i \in L} \gamma_i}{\prod_{-1}^m \gamma} =$$

$$= \frac{\gamma_{m_1} - 1}{\gamma_{m_1}} \cdot \ldots \cdot \frac{\gamma_{m_t} - 1}{\gamma_m}.$$

При $t = 1$ мы имели бы

$$P\{A_m\} = P_m(Б\phi_{\mathfrak{C}\mathfrak{P}}, \varphi) = \frac{\gamma_m - 1}{\gamma_m} = \frac{\beta_m}{\alpha + B_m}.$$

Таким образом,

$$P\{A_{m_1} \cap \ldots \cap A_{m_t}\} = P\{A_{m_1}\} \cdot P\{A_{m_2}\} \cdot \ldots \cdot P\{A_{m_t}\},$$

что и доказывает независимость событий $\{A_m\}$.

Лемма 2 доказана.

ДОКАЗАТЕЛЬСТВО ЛЕММЫ I. Мы уже знаем, что

$$P(Б\phi_{\mathfrak{C}\mathfrak{P}}, \varphi) = \frac{\beta_m}{\alpha + B_m} = 1 - \frac{\alpha + B_{m+1}}{\alpha + B_m}.$$

Суммируем по всем m , учитывая тривиальное неравенство

$$1 - x \leq \ln \frac{1}{x}.$$

$$\sum_m \frac{\beta_m}{\alpha + B_m} \leq \ln \prod_m \frac{\alpha + B_m}{\alpha + B_{m+1}}.$$

Частичное произведение

$$\prod_{m \leq s} \frac{\alpha + B_m}{\alpha + B_{m+1}} = \frac{\alpha + B_{-1}}{\alpha + B_{s+1}}.$$

Переходя к пределу при $s \to \infty$, получаем:

$$\sum_m P_m(Б\phi_{\mathfrak{C}\mathfrak{P}}, \varphi) \leq \ln \frac{\alpha + B_{-1}}{\alpha}.$$

Если $\varphi = \mathfrak{T}_n$, то $\alpha \geq \mathfrak{P}_n$ (см. начало доказательства леммы 2). Кроме того, $\alpha + B_{-1} \leq 1$ как вероятность в точке S_{-1} (см. там же). Таким образом,

$$\ln \frac{\alpha + B_{-1}}{\alpha} \leq \ln \frac{1}{\mathfrak{P}_n}.$$

Лемма I доказана.

Леммы I,2 дают нам всю информацию о В-стратегии $Б\Phi_{\mathfrak{T}\pi}$, необходимую для получения искомой оценки. По предложению I (см. § I) и лемме I,

$$P\{M, \varphi, \mathfrak{T}\} = 1$$

для всех $\varphi \in U$. Если $\mathfrak{T}$ -вычислимая нумерация и мы зафиксируем (считая, что $1/0 = 1$ и $\ln 0 = 1$):

$$\pi_n = \frac{c}{n(\ln n)^2},$$

то В-стратегия $Б\Phi_{\mathfrak{T}\pi}$ станет рекурсивной. По лемме I,

$$\sum_m P_m(Б\Phi_{\mathfrak{T}\pi}, \mathfrak{T}_n) \leqslant \ln n + O(\log \log n). \qquad (2)$$

Введем случайные величины

$$x_m = \begin{cases} 1, \text{ если } Б\Phi_{\mathfrak{T}\pi}(\mathfrak{T}_n, m) \neq Б\Phi_{\mathfrak{T}\pi}(\mathfrak{T}_n, m+1), \\ 0 \text{ иначе} \end{cases}$$

Согласно (2):

$$\sum_m P\{x_m = 1\} \leqslant \ln n + O(\log \log n),$$

кроме того, по лемме 2 случайные величины x_m независимы. Выполнены все условия предложения 3 (см. § I), поэтому при $n \to \infty$

$$P\{Б\Phi_{\mathfrak{T}\pi}, \mathfrak{T}_n, \leqslant \ln n + \sqrt{\log n}\, \log \log n\} \to 1. \qquad (3)$$

Теорема I доказана.

Остается показать, как из В-стратегии $Б\Phi_{\mathfrak{T}\pi}$ получить рекурсивную всюду финитную В-стратегию Q и рекурсивную финитную на U и $\mathfrak{T}$ -регулярную В-стратегию R сохраняя при этом оценку (3).

В обоих случаях берется рекурсивная последовательность рациональных чисел $\{\varepsilon_m\}$ такая, что $\prod (1+\varepsilon_m) < \infty$

Определение гипотезы $Q(\varphi, m)$ $(m \geqslant -1)$ при условии $Q(\varphi, m-1) = n$. Если $\mathfrak{T}_n(m) = \varphi(m)$, то $Q(\varphi, m) = n$ с вероятностью 1. Если же $\mathfrak{T}_n(m) = \varphi(m)$,

ищется $p \leqslant m$ такое, что $p \in \mathcal{F}$, где

$$\mathcal{F} = \{p \mid (\forall x \leqslant m)\, \mathfrak{T}_p(x) = \varphi(x)\}.$$

Если такого p не существует, $Q(\varphi, m) = n$ с вероятностью 1. Если же $p \leqslant m \wedge p \in \mathcal{F}$ найдено, начинаем вычисление рациональных приближений чисел π_n, $n \in \mathcal{F}$. Продолжая это до тех пор, пока не найдем двоично рациональное распределение $(\lambda_{n_1}, \dots, \lambda_{n_k})$ такое, что $\sum_i \lambda_{n_i} = 1$ и для всех i

$$n_i \in \mathcal{F} \wedge \lambda_{n_i} \leqslant (1+\varepsilon_m) \frac{\pi_{n_i}}{\sum_{n \in \mathcal{F}} \pi_n}.$$

Затем полагаем для $i = 1, 2, \dots, k$: $Q(\varphi, m) = n_i$ с вероятностью λ_{n_i}.

Так как нумерация $\mathfrak{T}$ вычислима, а распределение $\{\pi_n\}$ – рекурсивно, то рекурсивна и В-стратегия Q. Кроме того, Q, очевидно, всюду финитна.

Чтобы вместо всюду финитной стратегии Q получить рекурсивную, $\mathfrak{T}$ -регулярную и финитную на U стратегию R, в определении $Q(\varphi, m)$ поиск $p \leqslant m$ со свойством $p \in \mathcal{F}$ нужно заменить на безусловный поиск p с этом свойством.

ЛЕММА 3. Для всех $\varphi \in U$ и всех $k \in N$

$$P\{Q, \varphi, \geqslant k\} = P\{R, \varphi, \geqslant k\} \leqslant$$
$$\leqslant P\{Б\Phi_{\mathfrak{T}\pi}, \varphi, \geqslant k\} \cdot \prod_m (1+\varepsilon_m).$$

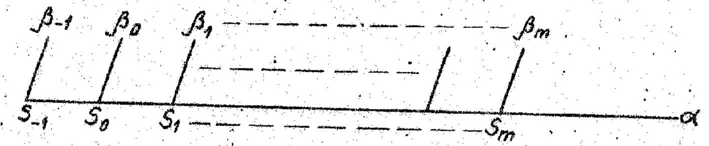
ДОКАЗАТЕЛЬСТВО. Для $\varphi \in U$ равенство

$$P\{Q, \varphi, \geqslant k\} = P\{R, \varphi, \geqslant k\}$$

очевидно, ибо стратегия Q – "растянутый" вариант стратегии R. Докажем неравенство

$$P\{R, \varphi, \leqslant k\} \leqslant P\{Б\Phi_{\mathfrak{T}\pi}, \varphi, \geqslant k\} \cdot \prod_m (1+\varepsilon_m).$$

Рассмотрим дерево всех функций класса $\mathcal{U}$, совпадающих до некоторого места с функцией φ.



Обозначения см. в доказательстве леммы 2. Вероятность β_{ij} $(i \leqslant j)$ того, что гипотеза стратегии $БФ_{\tau\mathfrak{T}}$, определяемая в точке S_i, проведет через точку S_j в сторону от функции φ, выражается так:

$$\beta_{ij} = \frac{\beta_j}{\alpha + B_i}$$

где $B_i = \beta_i + \beta_{i+1} + \cdots$. Для стратегии R соответствующие вероятности β'_{ij} обладают свойством

$$\beta'_{ij} \leqslant (1+\varepsilon_i)\,\beta_{ij}. \tag{4}$$

Вероятность $P\{R, \varphi, \geqslant k\}$ можно выразить через β'_{ij}:

$$P\{R, \varphi, \geqslant k\} = \sum \beta'_{-1,i_1}\,\beta'_{i_1+1,i_2}\cdots\beta'_{i_{k-1}+1,i_k},$$

где суммирование идет по всем кортежам $(i_1 \ldots i_k)$ таким, что

$$-1 \leqslant i_1 < i_2 < \cdots i_k.$$

Теперь, используя (4), получаем искомое неравенство

$$P\{R, \varphi, \geqslant k\} \leqslant \prod_m (1+\varepsilon_m)\cdot \sum \beta_{-1,i_1}\,\beta_{i_1+1,i_2}\cdots\beta_{i_{k-1}+1,i_k}$$

$$= \prod_m (1+\varepsilon_m)\cdot P\{БФ_{\tau\mathfrak{T}}, \varphi, \geqslant k\}.$$

Лемма 3 доказана.

Лемма 3 позволяет перенести без изменений свойство (3) стратегии $БФ_{\tau\mathfrak{T}}$ на стратегии Q, R. Сохраняется для Q, R также свойство

$$(\forall \varphi \in \mathcal{U})\; P\{БФ_{\tau\mathfrak{T}}, \varphi, \tau\} = 1.$$

Этим завершается доказательство теоремы I.

§ 4. Доказательство теоремы 2

Нумерованный класс $(\mathcal{U}, \tau)$, доказывающий нижнюю оценку $\ln n$, строится с помощью специальной диагональной процедуры, перебирающей одну за другой все вероятностные машины (в порядке их естественной нумерации $\{\mathfrak{M}_i\}$). Для каждой машины $\mathfrak{M}_i$ определяется серия блоков, каждый блок состоит из конечного числа функций. Если $\mathfrak{M}_i$ реализует В-стратегию, которая с вероятностью 1 синтезирует τ-номера всех функций некоторого блока, то на одной из функций блока эта машина с достаточно большой вероятностью достаточно много раз исправляет гипотезу. Из серий блоков, соответствующих всевозможным машинам $\mathfrak{M}_i$ складывается нумерация τ.

Такого рода диагональные конструкции впервые использовались Я.М.Барздинем [4,5,6].

Рассмотрим следующую последовательность независимых случайных величин $\{Z_i \mid i \geqslant 1\}$:

$$P\{Z_i = 1\} = \frac{1}{i}, \quad P\{Z_i = 0\} = 1 - \frac{1}{i}.$$

Таким образом,

$$\sum_{i=2}^{n} P\{Z_i = 1\} = \ln n + O(1).$$

По предложению 3 (см. § I) отсюда следует, что

$$P\Big\{\sum_{2}^{n} Z_i \geqslant \ln n - \sqrt{\log n}\,\log\cdot\log n\Big\} \to 1. \tag{5}$$

Имея в виду это свойство случайных величин $\{Z_i\}$, нетрудно понять значение следующей леммы.

ЛЕММА 4. Пусть даны: вероятностная машина $\mathfrak{M}$, натуральные числа k, n, $k < n$, рациональное число $\varepsilon > 0$. Тогда можно построить набор n функций

$$\sigma_1,\ \sigma_2,\ \ldots,\ \sigma_n$$

такой, что если машина $\mathfrak{M}$ с вероятностью 1 синтезирует в пределе $\mathfrak{S}$ -номер любой функции набора, то на одной из этих функций $\mathfrak{M}$ исправляет гипотезу $\geq k$ раз с вероятностью

$$\geq (1-\varepsilon)\, P\left\{ \sum_{1}^{n} z_i \geq k \right\}.$$

Доказательство см. ниже, § 5.

Если взять в лемме 4

$$\mathfrak{M}=\mathfrak{M}_i, \quad n=2^j, \quad k=j\ln 2-\sqrt{j}\,\log j, \quad \varepsilon=2^{-j},$$

то в силу (5), если машина $\mathfrak{M}_i$ с вероятностью 1 синтезирует $\mathfrak{S}$ -номера всех 2^j функций набора $\mathfrak{S}$, то на одной из этих функций с вероятностью, которая стремится к 1, $\mathfrak{M}_i$ исправляет гипотезу не менее $j\ln 2-\sqrt{j}\,\log j$ раз.

Блок $Б_{ij}$, соответствующий паре чисел (i,j), легко можно устроить таким образом, чтобы:

(а) все функции блока начинались словом $1^{\langle i,j\rangle}0$, где $\langle i,j\rangle$ - канторовский номер пары,

(б) "$\mathfrak{S}$ -номерами" считались не числа $1,2,\dots,n$, а любые другие, наперед выбранные n чисел.

Выбор этих чисел определяется положением функций блока $Б_{ij}$ в искомой нумерации $\mathfrak{T}$, которая складывается из всевозможных таких блоков.

Введем следующее кодирование всех троек (i,j,s) таких, что $0\leq s<2^j$ (это кодирование заимствовано из доказательства теоремы 4 работы [6]). Кодом тройки (i,j,s) будет число, имеющее двоичную запись

$$1\underbrace{0\dots0\alpha_k\dots\alpha_0}_{j}1\underbrace{0\dots0}_{i} \qquad (6)$$

где $\alpha_k\dots\alpha_0$ - двоичная запись числа s. В случае $j=0$ для $s=0$ берется "пустая запись".

Определение нумерованного класса $(U,\mathfrak{T})$, доказывающего теорему 2. Если число n не имеет двоичной записи вида (6), функцию $\mathfrak{T}_n$ полагаем тождественно равной нулю.

Если же n имеет вид (6) для тройки (i,j,s), где $0\leq s<2^j$, функцию $\mathfrak{T}_n$ полагаем равной s-ой функции блока $Б_{ij}$. При этом в процессе построения блока $Б_{ij}$ $\mathfrak{S}$ -номерами считаются все 2^j чисел вида (6), соответствующие паре (i,j).

Очевидно

$$2^{i+j+1}+2^i \leq \text{код}\ (i,j,s) < 2^{i+j+2}-2^i.$$

Пусть рекурсивная B-стратегия M синтезирует с вероятностью 1 $\mathfrak{T}$-номер любой функции полученного класса U. Пусть $\mathfrak{M}_i$ - машина, реализующая стратегию M. Для каждого $j\geq 0$ в блоке $Б_{ij}$ найдется функция, на которой M исправляет гипотезу не менее $j\ln 2-\sqrt{j}\,\log j$ раз с вероятностью

$$\geq (1-2^{-j})\, P\left\{ \sum_{1}^{2^j} z_i \geq j\ln 2-\sqrt{j}\,\log j \right\}. \qquad (7)$$

Обозначим $\mathfrak{T}$-номер этой функции через n_j. Очевидно

$$2^{i+j+1}<n_j<2^{i+j+2}.$$

Отсюда следует, что последовательность $\{n_j\}$ - возрастающая, и что

$$\log_2 n_j - i-2 < j < \log_2 n_j - i-1.$$

Таким образом, на функции $\mathfrak{T}_{n_j}$ B-стратегия M исправляет гипотезу не менее

$$j\ln 2-\sqrt{j}\,\log j \geq \ln 2\cdot\log_2 n_j - \sqrt{\log n_j}\,\log\log n_j$$

раз с вероятностью (7), которая при $j\to\infty$ стремится к 1.

§ 5. Доказательство леммы 4

Лемма 4 сформулирована в начале § 4. Итак, дана вероятностная машина $\mathfrak{M}$, даны натуральные числа n и k, $k < n$, и рациональное число $\varepsilon > 0$. Требуется построить набор n функций $\sigma_1, \sigma_2 \dots \sigma_n$ такой, что если $\mathfrak{M}$ реализует В-стратегию, которая с вероятностью 1 синтезирует (в пределе) σ -номер любой функции σ_j, то на одной из этих функций $\mathfrak{M}$ исправляет гипотезу не менее k раз с достаточно большой вероятностью, именно с вероятностью

$$\geqslant (1-\varepsilon)\, P\Big(\sum_{1}^{n} z_i \geqslant k \Big).$$

Определение случайных величин z_i см. в начале § 4.

Общий ход конструкции набора $(\sigma_1 \dots \sigma_n)$. Рассмотрим только случай $n = 4$, $k = 2$. Обобщение тривиально.

Процедура $P_{\mathfrak{M}}$. Параллельно вычисляются вероятности всевозможных событий вида

$$\mathfrak{M}(\Lambda) = s_{-1} \wedge \mathfrak{M}(\beta_0) = s_0 \wedge \dots \wedge \mathfrak{M}(\beta_0 \dots \beta_m) = s_m,$$

где $\bar{\beta}$ - двоичное слово, $\bar{s}$ - кортеж натуральных чисел. Делается это следующим образом. Параллельно для всех пар двоичных слов $(\bar{\alpha}, \bar{\beta})$ ведутся вычисления на машине $\mathfrak{M}$, при которых на ленте бернуллиевского диатчика записано слово $\bar{\alpha}$ (воспринимаемое как реализация работы датчика), а на ленте синтезируемой функции - слово $\bar{\beta}$ (воспринимаемое как начальный кусок функции, принимающей значения 0,1). Вначале каждой паре $(\bar{\beta}, \bar{s})$, где $\bar{\beta}$ - двоичное слово $\beta_0 \dots \beta_m$, $\bar{s}$ - кортеж натуральных чисел $s_{-1}, s_0, s_1, \dots, s_m$, приписывается пустое множество двоичных слов. Если в описанном процессе наступает ситуация, в которой машина $\mathfrak{M}$ для пары слов $(\bar{\alpha}, \bar{\beta})$ напечатала на входной ленте последовательность чисел $s_{-1}, s_0, s_1, \dots, s_m$ (соблюдая при этом правила реализации В-стратегий, см. [2],

§ 1), то к множеству слов, уже приписанному паре $(\bar{\beta}, \bar{s})$, прибавляется слово $\bar{\alpha}$. Конец определения процедуры $P_{\mathfrak{M}}$.

Параллельно работе процедуры $P_{\mathfrak{M}}$ функции $\sigma_1 \sigma_2 \sigma_3 \sigma_4$ получают в качестве значений все новые и новые нули:

$$\sigma_i(0) = 0, \quad \sigma_i(1) = 0, \dots$$

Только в некоторые исключительные моменты мы будем вмешиваться в этот процесс и давать функциям σ_i отдельные значения, равные 1.

Первый такой момент наступит, когда процедура $P_{\mathfrak{M}}$ вычислит информацию, достаточную для того, чтобы можно было получить следующее рациональное приближение для распределения некоторой гипотезы $\mathfrak{M}(0^{j_1})$:

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ p_1 & p_2 & p_3 & p_4 \end{pmatrix} \qquad (8)$$

где $\sum p_i = 1$ и для всех i:

$$p_i \leqslant \frac{1}{1-\delta} P\{ \mathfrak{M}(0^{j_1}) = i \}.$$

Ненаступление ситуации, в которой такое приближение возможно, означает, что для каждого $j \in N$ гипотеза $\mathfrak{M}(0^j)$ неопределена или $= 0$ или > 4 с вероятностью $\geqslant c\delta$. Но тогда с вероятностью $\geqslant c\delta$ это выполняется для бесконечно многих j одновременно. Получается, что

$$P\{ \mathfrak{M}, 0^{\infty}, \sigma \} \leqslant 1 - c\delta < 1,$$

где (в этом случае) $0^{\infty} = \sigma_1 = \sigma_2 = \sigma_3 = \sigma_4$. Этот случай с точки зрения леммы 4 неинтересен.

Рассмотрим теперь случай, когда распределение (8) получить удается. На основе (8), с помощью алгоритма, который описан ниже, один из номеров $1, 2, 3, 4$ исключается в следующем смысле. (Пусть это будет номер 1.) Функция σ_1 вместо очередного $\sigma_1(x_1) = 0$ получает значение

$\sigma_1(x_1)=1$, а дальше σ_1 полагается тождественно равной 0. Остальные функции $\sigma_2, \sigma_3, \sigma_4$ при $x=x_1$ получают значение 0. Таким образом, функция σ_1 становится отличной от $\sigma_2, \sigma_3, \sigma_4$ и в процессе σ-синтеза этих трех функций гипотезу 1 следует считать неверной.

Второе вмешательство в процесс становления функций σ_i (теперь уже без σ_1), происходит сразу за первым, если $p_1=0$, если же $p_1>0$, оно происходит, когда на основе информации, вычисленной процедурой $P_{\mathfrak{M}}$, удается получить следующее приближение распределения гипотезы $\mathfrak{M}(0^{j_2})$ (где $j_2>j_1$) при условии $\mathfrak{M}(0^{j_1})=1$:

$$\begin{pmatrix} 2 & 3 & 4 \\ q_2 & q_3 & q_4 \end{pmatrix} \tag{9}$$

где $\sum q_i=1$ и для всех i

$$q_i \leqslant \frac{1}{1-\delta}\, p\left\{ \mathfrak{M}(0^{j_2})=1 \mid \mathfrak{M}(0^{j_1})=1 \right\}.$$

Ненаступление ситуации, в которой такое приближение возможно, означает, что для всех $j>j_1$ при условии $\mathfrak{M}(0^{j_1})=1$ вероятность того, что $\mathfrak{M}(0^{j})$ неопределено, $=0$, $=1$ или >4, больше $c\delta$. Тогда $p\left\{\mathfrak{M}, 0^{\infty}, \sigma\right\}<1$, где $0^{\infty}=\sigma_2=\sigma_3=\sigma_4$. Этот случай можно не разбирать дальше.

В случае же, когда распределение (9) возможно, с помощью алгоритма, который описан ниже, исключается еще одна функция, пусть это σ_2. Мы полагаем $\sigma_2(x_2)=1$ и $\sigma_3(x_2)=\sigma_4(x_2)=0$, кроме того, $\sigma_2(x)=0$ для всех $x>x_2$. На этом второе вмешательство в определение набора кончается. Рассматривать распределение $\mathfrak{M}(0^{j_2})$ при условии $\mathfrak{M}(0^{j_1})=i$, где $i=2,3,4$ бесполезно, так как эти σ-номера пока еще являются верными гипотезами, и надеяться, что $\mathfrak{M}$ заменит их другими, не приходится.

Третье (и последнее при $n=4$) вмешательство в становление набора $\sigma_1\, \sigma_2\, \sigma_3\, \sigma_4$ происходит сразу за вторым,

если $q_2=0 \wedge p_2=0$. Если же $q_2>0$ и (или) $p_2>0$ то сначала ищется $j_3>j_2$ такое, что можно вычислить распределение (10) и (или) распределение (11). Первое распределение приближает распределение гипотезы $\mathfrak{M}(0^{j_3})$ при условии $\mathfrak{M}(0^{j_1})=1 \wedge \mathfrak{M}(0^{j_2})=2$:

$$\begin{pmatrix} 3 & 4 \\ r_3 & r_4 \end{pmatrix}, \tag{10}$$

здесь $r_3+r_4=1$ и для $i=2,3$

$$r_i \leqslant \frac{1}{1-\delta}\, p\left\{ \mathfrak{M}(0^{j_3})=i \mid \mathfrak{M}(0^{j_1})=1 \wedge \mathfrak{M}(0^{j_2})=2 \right\}.$$

Второе распределение приближает распределение $\mathfrak{M}(0^{j_3})$ при условии $\mathfrak{M}(0^{j_1})=2$:

$$\begin{pmatrix} 3 & 4 \\ s_3 & s_4 \end{pmatrix}, \tag{11}$$

здесь $s_3+s_4=1$ и для $i=3,4$:

$$s_i \leqslant \frac{1}{1-\delta}\, p\left\{ \mathfrak{M}(0^{j_3})=i \mid \mathfrak{M}(0^{j_1})=2 \right\}.$$

Ненаступление ситуации, в которой нужные приближения возможны, означает, как и прежде, что

$$p\left\{\mathfrak{M}, 0^{\infty}\ \sigma\right\}<1,$$

где $0^{\infty}=\sigma_3=\sigma_4$. Этот случай незачем разбирать дальше.

В случае же, когда (10) и (или) (11) получить удается, в соответствии с алгоритмом, описанным ниже, исключается еще одна функция, пусть это σ_3. Мы полагаем $\sigma_3(x_3)=1$, $\sigma_4(x_3)=0$, а при $x>x_3$ обе функции пусть равны тождественно нулю. Отметим, что условия

$$\mathfrak{M}(0^{j_1})=1 \wedge \mathfrak{M}(0^{j_2})=3 \vee 4,$$

$$\mathcal{M}(0^{j_1}) = 3 \vee 4$$

рассматривать не стоит. До исключения σ_3 гипотезы $3, 4$ одинаково верны и нет надежды, что $\mathcal{M}$ заменит их другими.

Таков общий ход конструкции набора функций, соответствующего машине $\mathcal{M}$ и натуральному числу n. Описанный ниже алгоритм исключения (мы рассмотрим случай $n=4$, $k=2$, обобщение тривиально) обеспечит, что на функции $\sigma_4 = 0^{\infty}$ стратегия машины $\mathcal{M}$ будет работать плохо с достаточно большой вероятностью.

Если машина $\mathcal{M}$ с вероятностью 1 синтезирует σ-номер любой функции набора $\sigma_1\ \sigma_2\ \sigma_3\ \sigma_4$, то в процессе построения этого набора произошли все три акта исключения, пусть это были $1\ 2\ 3\ (4)$. В результате получается следующая картина:

<table>
<tr><td colspan="4">1</td><td colspan="2">2</td><td>3</td><td>4</td></tr>
<tr><td colspan="2">2</td><td>3</td><td>4</td><td colspan="2">2</td><td>3</td><td>4</td></tr>
<tr><td>3</td><td>4</td><td>3</td><td>4</td><td>3</td><td>4</td><td>3</td><td>4</td></tr>
<tr><td colspan="8">4</td></tr>
</table>

В четвертой строке с вероятностью 1 фигурирует номер 4, это единственный номер функции $\sigma_4 = 0^{\infty}$. Разбиение первой строки соответствует распределению $(P_1\ P_2\ P_3\ P_4)$, см. (8). Вторая строка разбивает P_1 в соответствии с распределением $(q_2\ q_3\ q_4)$ (см. (9)). Третья строка разбивает $P_1\ q_2$ в соответствии с $(r_3\ r_4)$, см. (10), она разбивает также P_2 в соответствии с $(s_3\ s_4)$ (см. (11)).

Введем следующие обозначения:

$$|1| = P_1, \qquad |2| = |22| = P_2,$$
$$|12| = P_1 q_2, \qquad |133| = |13| = P_1 q_3,$$
$$|123| = P_1 q_2 r_3, \quad |224| = P_2\, s_4.$$

Допустим временно, что эти вероятности совпадают т о ч н о с вероятностями, характеризующими машину $\mathcal{M}$, например:

$$|1\,2\,3| = P\{\mathcal{M}(0^{j_1}) = 1 \wedge \mathcal{M}(0^{j_2}) = 2 \wedge \mathcal{M}(0^{j_3}) = 3\},$$
$$|2\,2\,3| = P\{\mathcal{M}(0^{j_1}) = 2 \wedge \mathcal{M}(0^{j_3}) = 3\}.$$

Тогда вероятность того, что $\mathcal{M}$ на функции $\sigma_4 = 0^{\infty}$ не менее $k=2$ раза исправляет гипотезу, будет

$$\geqslant |123| + |124| + |133| + |223| =$$
$$= |12| + |13| + |223|. \qquad \text{(a)}$$

Если бы мы на последнем шаге исключили не σ_3, а σ_4, то вместо (а) получилось бы

$$|12| + |14| + |224|. \qquad \text{(aa)}$$

Таким образом, если σ_1 и σ_2 уже исключены, следующим резонно исключить σ_3 или σ_4 в зависимости от того, что больше – (а) или (аа). И это разумное решение осуществимо: все слагаемые (а) и (аа) нам известны, как только мы достигли уровня j_3.

Рассмотрим теперь уровень j_2. Функция σ_1 уже исключена. Нам известны вероятности $|1|$, $|2|$, $|3|$, $|4|$, $|12|$, $|13|$, $|14|$. Какую из функций исключить – σ_2, σ_3 или σ_4? "Качество" решения "исключить σ_2" можно оценивать значением суммы (а)+(аа)

$$|12| + |13| + |223| + |12| + |14| + |224| =$$
$$= |1| + |22| + |12| = |1| + |2| + |12|. \qquad \text{(б)}$$

К счастью, эта сумма зависит только от вероятностей, известных на уровне j_2. Соответствующие суммы для σ_3 и σ_4

$$|1| + |3| + |13|, \qquad \text{(бб)}$$
$$|1| + |4| + |14|. \qquad \text{(ббб)}$$

Следовательно, большее из чисел (б), (бб), (ббб) и должно определить, какую из функций $\sigma_2, \sigma_3, \sigma_4$ исключить вслед за σ_1. Это решение осуществимо на уровне j_2.

Наконец, рассмотрим уровень j_1. "Качество" решения "исключить сначала σ_1" будем оценивать суммой (б)+(бб)+(ббб)

$$3\cdot|1| + |2| + |3| + |4| + |12| + |13| + |14| =$$
$$= 1+3\cdot|1|.$$

Аналогичные показатели для $\mathfrak{S}_2, \mathfrak{S}_3, \mathfrak{S}_4$ равны

$$1+3\cdot|2|,\quad 1+3\cdot|3|,\quad 1+3\cdot|4|.$$

Все они вычислимы на уровне j_1, где известны только вероятности $|i|$. Следовательно, сначала следует исключить $\mathfrak{S}_i$ с наибольшим $|i|$.

В этих соображениях и заключается сущность предлагаемого алгоритма, определяющего порядок исключения функций. Правда, здесь нужно еще доказать в общем виде, что все рассмотренные суммирования действительно всегда приводят к возвращению на уровень, где должно быть принято соответствующее решение. Этой проблемой "сокращения хвостов" мы займемся в § 6.

ЗАМЕЧАНИЕ. Рассмотренный простой пример может создать иллюзию, что предлагаемый алгоритм можно заменить более простым. Достаточно рассмотреть случай $n=5$, $k=3$, чтобы убедиться в обратном. Например, после исключения $\mathfrak{S}_1$, "качество" решения "исключить $\mathfrak{S}_i$" $(i=2,3,4,5)$ характеризуется суммой $|1|+|i|+3\cdot|1i|$. Вряд ли такой критерий можно угадать без вычислений, подобных проведенным выше.

Если бы имело место наше предположение о совпадении вероятностей, характеризующих машину $\mathfrak{M}$, и вероятностей $|i|, |ij|$ и т.д., фигурирующих выше, то в рассмотренном случае $(n=4,\ k=2)$ вероятность того, что на функции $\mathfrak{S}_4 = 0^\infty$ машина $\mathfrak{M}$ не менее 2 раз исправляет гипотезу, можно оценить следующим путем.

На уровне j_1 мы имели 4 числа:

$$1+3\cdot|1|,\quad 1+3\cdot|2|,\quad 1+3\cdot|3|,\quad 1+3\cdot|4|.$$

Их сумма равна $4+3\cdot 1=7$, т.е. зависит только от n, k и может быть вычислена, зная эти числа. Если мы, действуя согласно алгоритму, исключили $\mathfrak{S}_1$, то

$$1+3\cdot|1| \geqslant \frac{7}{4}.$$

Если затем исключается $\mathfrak{S}_2$, то

$$|1|+|2|+|12| \geqslant \frac{7}{4\cdot 3}.$$

Наконец, исключение $\mathfrak{S}_3$ означает

$$|12|+|13|+|223| \geqslant \frac{7}{4\cdot 3\cdot 2} = \frac{7}{24}.$$

Таким образом, если бы наше предположение было верным, то машина $\mathfrak{M}$ на функции $\mathfrak{S}_4 = 0^\infty$ исправляла бы гипотезу $\geqslant 2$ раз с вероятностью $\geqslant \frac{7}{24}$. Но наше предположение неверно – вероятности $|i|, |ij|$ и т.д. лишь приблизительно характеризуют машину $\mathfrak{M}$. Нетрудно, однако, подобрать число δ, определяющее степень приближения, так, чтобы, не получая для машины $\mathfrak{M}$ оценку $\geqslant \frac{7}{24}$, мы получили все же $\geqslant \frac{7}{24}(1-\varepsilon)$, где ε – наперед заданное число, фигурирующее в лемме 4.

В самом деле,

$$\begin{aligned}
P\{\mathfrak{M}, \mathfrak{S}_4, \geqslant 2\} &\geqslant P\{\mathfrak{M}(0^{d_1}) = 1 \wedge \mathfrak{M}(0^{d_2}) = 2\} + \\
&+ P\{\mathfrak{M}(0^{d_1}) = 1 \wedge \mathfrak{M}(0^{d_2}) = 3\} + \\
&+ P\{\mathfrak{M}(0^{d_1}) = 2 \wedge \mathfrak{M}(0^{d_3}) = 3\} \geqslant \\
&\geqslant (1-\delta)p_1[(1-\delta)q_2 + (1-\delta)q_3] + (1-\delta)p_2(1-\delta)s_3 = \\
&= (1-\delta)^2(p_1q_2 + p_1q_3 + p_1s_3) = \\
&= (1-\delta)^2(|12|+|13|+|223|) \geqslant (1-\delta)^2\frac{7}{24}.
\end{aligned}$$

Если δ выбрать таким, что $(1-\delta)^2 \geqslant 1-\varepsilon$, то

$$P\{\mathfrak{M}, \mathfrak{S}_4, \geqslant 2\} \geqslant (1-\varepsilon)\frac{7}{24}.$$

Обобщение этих соображений тривиально.

Для доказательства леммы 4 остается проверить, что

$$\frac{7}{24} = P\left\{ \sum_{2}^{4} z_i \geqslant 2 \right\}. \qquad (12)$$

где $\{z_i\}$ - независимые случайные величины такие, что

$$P\{z_i = 1\} = \frac{1}{i}, \qquad P\{z_i = 0\} = 1 - \frac{1}{i}.$$

В конкретном случае $n=4$, $\kappa=2$ равенство (I2) легко проверить непосредственно:

$$\frac{1}{2}\cdot\frac{1}{3}\cdot\left(1-\frac{1}{4}\right)+\frac{1}{2}\cdot\left(1-\frac{1}{3}\right)\cdot\frac{1}{4}+\left(1-\frac{1}{2}\right)\cdot\frac{1}{3}\cdot\frac{1}{4}+$$
$$+\frac{1}{2}\cdot\frac{1}{3}\cdot\frac{1}{4}=\frac{7}{24}$$

Общий случай (n, κ - произвольны, $\kappa < n$) доказывается с помощью леммы 3 (см. выше § 3). Заметим, что в оценке

$$|12| + |13| + |223| \geqslant \frac{7}{24}$$

равенство достигается на "симметричной" таблице:

<table>
<tr><td colspan="4">1</td><td colspan="2">2</td><td>3</td><td>4</td></tr>
<tr><td colspan="2">2</td><td>3</td><td>4</td><td colspan="2">2</td><td>3</td><td>4</td></tr>
<tr><td>3</td><td>4</td><td>3</td><td>4</td><td>3</td><td>4</td><td>3</td><td>4</td></tr>
<tr><td colspan="8">4</td></tr>
</table>

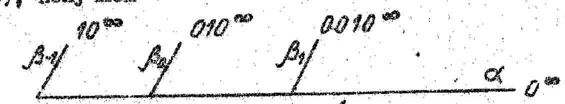
Здесь всякое деление (сверху вниз) является делением на равные части: четыре, три, две. Оказывается, что эта таблица в точности передает работу на функции $\sigma_4 = 0^\infty$ B-стратегии $Б\Phi_{\sigma\pi}$ (см. § 3), где

$$\sigma_1 = 10^\infty, \quad \sigma_2 = 010^\infty, \quad \sigma_3 = 0010^\infty, \quad \sigma_4 = 0^\infty,$$

$$\pi = \left(\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}\right).$$

В самом деле, $F\Phi_{\sigma\pi}(\Lambda)$ равно $1, 2, 3$ или 4 с одинаковой вероятностью $1/4$. Далее, при условии $Б\Phi_{\sigma\pi}(\Lambda) = 1$ гипотеза $Б\Phi_{\sigma\pi}(0)$ равна $2, 3$ или 4 с одинаковой вероятностью $1/3$. При условии $Б\Phi_{\sigma\pi}(\Lambda) = 1 \wedge Б\Phi_{\sigma\pi}(0) = 2$ гипотеза $Б\Phi_{\sigma\pi}(00)$ равна 3 или 4 с вероятностями $1/2$. При условии $Б\Phi_{\sigma\pi}(\Lambda) = 2$ гипотеза $Б\Phi_{\sigma\pi}(00)$ также равна 3 или 4 с вероятностями $1/2$. Наконец, гипотеза $Б\Phi_{\sigma\pi}(000) = 4$ с вероятностью 1.

Используя метод дерева (см. доказательство леммы 3 в § 3), получаем



$$\beta_{-1} = \beta_0 = \beta_1 = \alpha = \frac{1}{4},$$

$$P\{Б\Phi_{\sigma\pi}(\Lambda) \neq Б\Phi_{\sigma\pi}(0)\} = \frac{\beta_{-1}}{\alpha + \beta_{-1} + \beta_0 + \beta_1} = \frac{1}{4},$$

$$P\{Б\Phi_{\sigma\pi}(0) \neq Б\Phi_{\sigma\pi}(00)\} = \frac{\beta_0}{\alpha + \beta_0 + \beta_1} = \frac{1}{3},$$

$$P\{Б\Phi_{\sigma\pi}(00) \neq Б\Phi_{\sigma\pi}(000)\} = \frac{\beta_1}{\alpha + \beta_1} = \frac{1}{2}.$$

По лемме 3 независимы следующие случайные величины ($i = 0, 1, 2$):

$$x_i = \begin{cases} 1, & \text{если } Б\Phi_{\sigma\pi}(0^i) \neq Б\Phi_{\sigma\pi}(0^{i+1}), \\ 0, & \text{иначе.} \end{cases}$$

Таким образом, система случайных величин (x_0, x_1, x_2) тождественна системе (z_4, z_3, z_2). Так как сумма $\sum_0^2 x_i = \sum_2^4 z_i$ представляет собой число исправлений гипотезы стратегии $Б\Phi_{\sigma\pi}$ на функции $\sigma_4 = 0^\infty$, то

$$\frac{7}{24} = P\left\{ \sum_{2}^{4} z_i \geqslant 2 \right\},$$

где $\frac{7}{24}$ - нижняя оценка, вычисленная с помощью описанного выше алгоритма при $n = 4$, $\kappa = 2$.

Очевидно, это рассуждение проходит для произвольных n, κ, $\kappa < n$.

Итак, лемма 4 будет доказана, если удастся обосновать в общем случае использованное в построении алгоритма явление "сокращения хвостов".

§ 6. Доказательство "сокращения хвостов"

Напомним ситуацию. Дано натуральное число n. Пустому кортежу Λ соответствует некоторое распределение вероятностей на множестве $\{1, 2, \ldots, n\}$, соответствующие вероятности обозначим через:

$$|1|_\Lambda,\ |2|_\Lambda,\ \ldots,\ |n|_\Lambda.$$

Если теперь исключить некоторое число i, вероятность $|i|_\Lambda$ распадется следующим образом:

$$|i|_\Lambda = \sum_{y \neq i} |iy|_i.$$

Вероятности $|x|_\Lambda$, где $x \neq i$, не распадаются:

$$|x|_\Lambda = |xx|_i.$$

Полученные таким образом кортежи длины 2 назовем (i) - допустимыми.

Если на следующем шаге исключить число $j \neq i$, распадаются вероятности $|ij|_i$, $|jj|_i$:

$$|xj|_i = \sum_{y \neq i,j} |xjy|_{ij}, \qquad \text{если } x = i \vee x = j$$

Все остальные вероятности сохраняются:

$$|xy|_i = |xyy|_{ij},$$

где $y \neq j$. Полученные таким образом кортежи длины 3 будем называть (ij) -допустимыми.

В общем случае, если после исключения чисел $\bar{p} = (p_1 \ldots p_a)$ (тогда получено уже понятие $\bar{p}$ -допустимого кортежа $\bar{x}$ длины $a+1$ и соответствующие таким $\bar{x}$ вероятности $|\bar{x}|_{\bar{p}}$) исключается число $q \notin \bar{p}$, то вероятности вида $|\bar{x}q|_{\bar{p}}$ распадаются:

$$|\bar{x}q|_{\bar{p}} = \sum_{\substack{y \notin \bar{p} \\ y \neq q}} |\bar{x}qy|_{\bar{p}q}.$$

Остальные вероятности сохраняются:

$$|\bar{x}y|_{\bar{p}} = |\bar{x}yy|_{\bar{p}q},$$

где $y \neq q$ (и, конечно, $y \notin \bar{p}$, если кортеж $\bar{x}y$ $\bar{p}$-допустим).

В разряд $\bar{p}q$ -допустимых кортежей включаются все кортежи, входящие в правые части последних двух равенств.

Очевидно, кортеж $(x_1 \ldots x_{a+1}$ будет $\bar{p}$ -допустимым, где $\bar{p} = (p_1 \ldots p_a)$, если и только если выполняются условия: для всех $i \leqslant a$:

(I) $x_{i+1} \notin (p_1 \ldots p_i)$,

(2) $x_i \notin (p_1 \ldots p_i) \to x_{i+1} = x_i$.

В самом конце процесса исключения мы имеем некоторую перестановку $(p_1 \ldots p_n)$ множества $(1 2 \ldots n)$. Обозначим $\bar{p}(p_1 \ldots p_{n-1})$. Каждый $\bar{p}$ - допустимый кортеж $\bar{x}$ (длины n) получил вероятность $|\bar{x}|_{\bar{p}}$. В соответствии с нашей задачей, особое внимание обращается на кортежи $\bar{x} = (x_1 \ldots x_n)$, обладающие свойством

$$x_i = x_{i+1} \quad \text{для} \quad \geqslant k \text{ различных значений } i.$$

Здесь k - натуральное число, $k < n$. Множество всех таких кортежей обозначим через S. Мы воспользуемся только симметричностью множества S, именно, если π - любая перестановка $(1 2 \ldots n)$, то

$$(x_1 \ldots x_n) \in S \to (\pi x_1 \ldots \pi x_n) \in S.$$

"Качество" перестановки $(p_1 \ldots p_n)$ множества $(1 2 \ldots n)$ характеризуется вероятностью

$$H(\bar{p}) = \sum_{\bar{x}} |\bar{x}|_{\bar{p}}.$$

где суммирование идет по всем $\bar{p}$ -допустимым $\bar{x} \in S$, где $\bar{p} = (p_1 \ldots p_{n-1})$. Т.е. суммирование идет по некоторым кортежам длины n.

Свой показатель "качества" приписывается каждому кортежу $\bar{q} = (q_1 \ldots q_a)$, не содержащему повторений, $a \leqslant n-1$.

Во-первых, если $a = n-2$ (т.е. вне $\bar q$ остаются два числа множества $(1 2 \dots n)$, пусть это s, t), то

$$H(\bar q) = H(\bar q s) + H(\bar q t).$$

В общем случае $a \le n-2$, если вне $\bar q$ остаются числа $q_{a+1}, \dots, q_n$, определяем по индукции:

$$H(\bar q) = \sum_{i=a+1}^{n} H(\bar q q_i).$$

В частности:

$$H(\Lambda) = H(1) + H(2) + \dots + H(n).$$

Требуется показать, что для всех $\bar q$ и всех $q \notin \bar q$ значение $H(\bar q q)$ можно вычислить, зная только вероятности $|\bar x|_{\bar q}$ для всех $\bar q$-допустимых кортежей $\bar x$. Это будет сделано, если удастся показать, что $H(\bar q q)$ можно представить в виде

$$H(\bar q q) = \sum_{\bar x} c(\bar x)\,|\bar x|_{\bar q}$$

где $c(\bar x)$ – натуральные числа, которые можно вычислить, зная $\bar x, \bar q, q$ (суммирование идет по всем $\bar q$-допустимым кортежам $\bar x$).

Сначала рассмотрим случай $\bar q = (q_1 \dots q_{n-2})$. Если вне $\bar q$ остались числа q, t то для всякого $\bar q q$-допустимого $\bar x = (x_1 \dots x_n)$ имеем $x_n = t$ и $|\bar x|_{\bar q q} = |x_1 \dots x_{n-1}|_{\bar q}$, т.е.

$$H(\bar q q) = \sum_{\bar x} c(\bar x)\,|\bar x|_{\bar q}, \qquad (13)$$

где суммирование идет по всем $\bar q$-допустимым $\bar x$ (длины $n-1$), при этом

$$c(\bar x) = \begin{cases} 1, & \text{если } \bar x t \in S, \\ 0, & \text{иначе}. \end{cases}$$

Рассмотрим теперь следующий случай: $\bar q = (q_1 \dots q_{n-3})$, пусть вне $\bar q$ остались числа q, s, t, тогда, по определению,

$$H(\bar q q) = H(\bar q q s) + H(\bar q q t).$$

Согласно (13)

$$H(\bar q q s) = \sum_{\bar x} |\bar x|_{\bar q q}, \qquad (14)$$

где суммирование идет по всем $\bar q q$-допустимым $\bar x$ таким, что $\bar x t \in S$. Множество S симметрично, кроме того, свойство $\bar q q$-допустимости симметрично относительно перестановки s, t. Таким образом, если в (14) всюду заменить s на t и t на s, мы получим $H(\bar q q t)$ вместо $H(\bar q q s)$. Сумму $H(\bar q q s) + H(\bar q q t)$ можно упростить, различая следующие случаи ($\bar x$ – произвольный кортеж, входящий в (14)):

(1) $\bar x = \bar y s$, где $\bar y$ не содержит чисел s, t. Тогда, меняя s, t местами, получаем кортеж $\bar y t$, при этом

$$|\bar y s|_{\bar q q} + |\bar y t|_{\bar q q} = |\bar y|_{\bar q}$$

Кортеж $\bar y$ $\bar q$-допустим.

(2) $\bar x = \bar y t$, аналогично.

(3) $\bar x = \bar y s \dots s$, где $\bar y$ не содержит s, t. Тогда, меняя s и t местами, получаем $\bar y t \dots t$, при этом

$$|\bar y s \dots s|_{\bar q q} = |\bar y s \dots|_{\bar q},$$
$$|\bar y t \dots t|_{\bar q q} = |\bar y t \dots|_{\bar q}.$$

Кортежи $\bar y s \dots$ и $\bar y t \dots$ $\bar q$-допустимы.

(4) $\bar x = \bar y t \dots t$, аналогично.

Таким образом, для $H(\bar q q)$ мы получили представление

$$H(\bar q q) = \sum_{\bar x} c(\bar x)\,|\bar x|_{\bar q}, \qquad (15)$$

где суммирование идет по всем $\bar{q}$ -допустимым $\bar{x}$ (в данном случае длина таких $\bar{x}$ равна $n-2$), и $c(\bar{x})$ можно вычислить по $\bar{x}$, $\bar{q}$, q. Очевидна следующая особенность выражения (I5): оно симметрично относительно s, t, т.е. если $\bar{x}$ содержит s (тогда в этом случае, $\bar{x}$ не содержит t), то замена в $\bar{x}$ числа s на число t дает кортеж $\bar{y}$ такой, что $c(\bar{y}) = c(\bar{x})$.

Теперь можно перейти к шагу индукции. Пусть $\bar{q} = (q_1 \dots q_a)$ - кортеж без повторений и числа q, r_2, $r_3, \dots, r_{n-a}$ остаются вне $\bar{q}$. Предположим, что

$$H(\bar{q}\, q\, r_2) = \sum_{\bar{x}} c(\bar{x})\, |\bar{x}|_{\bar{q}q}, \tag{I6}$$

где суммирование идет по всем $\bar{q}q$ -допустимым $\bar{x}$. При этом выполняется следующее условие симметричности: если кортеж $\bar{x}$ $\bar{q}q$-допустим, а $(s_3 \dots s_{n-a})$ - любая перестановка чисел $(r_3 \dots r_{n-a})$, то действие $\bar{s}$ на $\bar{x}$ приводит к кортежу y такому, что $c(\bar{y}) = c(\bar{x})$.

По определению,

$$H(\bar{q}\, q) = \sum_{i=2}^{n-q} H(\bar{q}\, q\, r_i). \tag{I7}$$

Чтобы получить из (I6) выражение для $H(\bar{q}\, q\, r_i)$ $(i>2)$, нужно применить ко всем $\bar{q}\,q$ -допустимым кортежам $\bar{x}$ перестановку множества $(1\,2 \dots n)$, меняющую местами r_2 и r_i (остальные числа остаются на месте). Отсюда следует, что сумму (I7) можно упростить, различая следующие случаи:

(1) $\bar{x} = \bar{y}\, r_2$, где $\bar{y}$ не содержит r_2. В результате перестановок получаем кортежи $\bar{y}\, r_3, \dots, \bar{y}\, r_{n-a}$, при этом

$$|\bar{y} r_2|_{\bar{q}q} + |\bar{y} r_3|_{\bar{q}q} + \dots + |\bar{y}\, r_{n-a}|_{\bar{q}q} = |\bar{y}|_{\bar{q}}.$$

(2) $\bar{x} = \bar{y}\, r_i$, где $i>2$ и $\bar{y}$ не содержит r_i. Тогда (I6) содержит <u>все</u> кортежи этого рода, притом с одина-

ковым коэффициентом $c(\bar{x})$:

$$\bar{y}\, r_3,\ \bar{y}\, r_4, \dots, \bar{y}\, r_{n-a}.$$

В результате применения упомянутых $n-a-2$ перестановок каждый из этих $n-a-2$ кортежей порождает один кортеж $\bar{y} r_2$ плюс еще $n-a-3$ экземпляров самого себя. Все это войдет в сумму (I7)

$$(n-a-2)(|y r_2|_{\bar{q}q} + \dots + |\bar{y}\, r_{n-a}|_{\bar{q}q}) = (n-a-2) \cdot |\bar{y}|_{\bar{q}}.$$

(3) $\bar{x} = \bar{y}\, r_2 \dots r_2$, где $\bar{y}$ не содержит r_2. В результате перестановок возникают все аналогичные кортежи $\bar{y}\, r_i \dots r_i$ $(i>2)$. Нам достаточно заметить, что

$$|\bar{y}\, r_i \dots r_i|_{\bar{q}q} = |\bar{y} r_i \dots|_{\bar{q}} \tag{I8}$$

и что все r_i $(i \geqslant 2)$ войдут в сумму (I7) симметрично.

(4) $\bar{x} = \bar{y}\, r_i \dots r_i$, где $i>2$ и $\bar{y}$ не содержит r_i. Тогда в (I6) содержатся с тем же $c(\bar{x})$ все аналогичные кортежи:

$$(\bar{y}\, r_3 \dots r_3), \dots, (\bar{y}\, r_{n-a} \dots r_{n-a}).$$

Применение перестановок дает в итоге по $n-a-2$ экземпляра каждого $\bar{y}\, r_i \dots r_i$ $(i \geqslant 2)$. Нам достаточно заметить, что имеет место (I8) и что все r_i $(i \geqslant 2)$ войдут в сумму (I7) симметрично.

Таким образом, мы получили для $H(\bar{q}\, q)$ (где q - любое число $\notin \bar{q}$) представление

$$H(\bar{q}\, q) = \sum_{\bar{x}} c'(\bar{x})\, |\bar{x}|_{\bar{q}},$$

где суммирование идет по $\bar{q}$ -допустимым кортежам $\bar{x}$, а коэффициенты $c'(\bar{x})$ можно вычислить, зная $\bar{q}$, q, $\bar{x}$. Все числа $r \notin \bar{q}q$ входят в это выражение $H(\bar{q}\, q)$ симметрично.

Этим завершается шаг индукции.

Таким образом, доказано, что характеристику $H(\bar{q}\, q)$ любого кортежа $\bar{q}\, q$, не содержащего повторений, можно вычислить, зная только вероятности $|\bar{x}|_{\bar{q}}$ для всех $\bar{q}$ -допустимых кортежей $\bar{x}$. Этого достаточно для обоснования возможности своевременно принять требуемое в алгоритме леммы 4 решение (см. выше, § 5).

В заключение отметим, что в случае пустого кортежа полученный результат означает, что характеристика $H(\Lambda)$ зависит только от числа n и симметричного множества S кортежей длины n , которые суть единственные параметры описанной конструкции. В случае, когда

$$S = \{ (x_1 \dots x_n) \mid x_i \neq x_{i+1} \text{ для } \geqslant k \text{ значений } i \}$$

в § 5 показано, что

$$\frac{1}{n!} H(\Lambda) = P \left\{ \sum_{2}^{n} z_i \geqslant k \right\}.$$

Определение случайных величин z_i см. в § 5.

Результаты этой статьи изложены без доказательств в [7].

ЛИТЕРАТУРА

1. Колмогоров А.Н. Три подхода к определению понятия "количество информации". - "Проблемы передачи информации", 1965, т.1, № 1, с.1-7.
2. Подниекс К.М. Вероятностное прогнозирование вычислимых функций. - "Уч. зап. Латв. ун-та", 1975, т.233, с.57-76.
3. Ламперти Дж. Вероятность. М., 1973, 189 с.
4. Барздинь Я.М., Фрейвалд Р.В. О прогнозировании общерекурсивных функций. - "Доклады АН СССР", 1972, т.206, № 3, с.521-524.
5. Барздинь Я.М. Предельный синтез τ-номеров. - "Уч.зап. Латв. ун-та", 1974, т.210, с.112-116.
6. Барздинь Я.М., Фрейвалд Р.В. Прогнозирование и предельный синтез эффективно перечислимых классов функций. - "Уч. зап. Латв. ун-та", 1974, т.210, с.102-111.
7. Подниекс К.М. Вероятностный синтез нумерованных классов функций. - "Доклады АН СССР", 1975, т.223, № 5, с.1071-1074.