2008 Теоретические основы прикладной дискретной математики № 1(1)

DOI 10.17223/20710410/1/6

УДК 519.728

ЭНТРОПИЯ НЕДООПРЕДЕЛЕННЫХ ПОСЛЕДОВАТЕЛЬНОСТЕЙ ПРИ ОГРАНИЧЕНИЯХ НА ДООПРЕДЕЛЕНИЯ

Л.А. Шоломов

Институт системного анализа РАН, г. Москва

E-mail: sholomov@isa.ru

Рассматриваются последовательности недоопределенных символов, каждому из которых соответствует некоторое множество полностью определенных символов, одним из которых он может быть замещен (доопределен). При заданных ограничениях на вид доопределний получены оценки минимальной мощности доопределяющего множества для класса последовательностей с заданными кратностями появления символов.

Ключевые слова: *недоопределенный символ*, *доопределение*, *комбинаторная энтропия*, *ограниченная комбинаторная энтропия*, *W-энтропия.*

Задано множество $M = \{0, 1, \ldots, m-1\}$ и каждому непустому $T \subseteq M$ сопоставлен символ a_T. Алфавит всех символов a_T обозначим через A, а его подалфавит $\{a_0, a_1, \ldots, a_{m-1}\}$, символы которого соответствуют элементам множества M, – через A_0. Символы из A_0 будем называть *основными*, из A – *недоопределенными*. *Доопределением* символа a_T назовем всякий основной символ a_i, $i \in T$, а доопределением последовательности в алфавите A – любую последовательность в алфавите A_0, полученную из исходной заменой всех символов некоторыми доопределениями. Будем рассматривать последовательности длины n в алфавите A. Для набора натуральных чисел $\boldsymbol{n} = (n_T, T \subseteq M)$, $\sum_T n_T = n$, обозначим через $K_n(\boldsymbol{n})$ множество всех последовательностей длины n, в которых символы a_T, $T \subseteq M$, встречаются n_T раз. Скажем, что некоторое множество последовательностей в алфавите A_0 *доопределяет* класс $K_n(\boldsymbol{n})$, если в нем найдется доопределение каждой последовательности из этого класса. Пусть $N_n(\boldsymbol{n})$ – минимальная мощность множества, доопределяющего $K_n(\boldsymbol{n})$. Величину $\log N_n(\boldsymbol{n})$ (здесь и дальше логарифмы двоичные) будем называть *комбинаторной энтропией* класса $K_n(\boldsymbol{n})$.

Введем обобщение энтропийной функции Шеннона на случай недоопределенных данных (подробнее см. в [1]), положив для всякого набора $P = (p_T, T \subseteq M)$, $p_T \geq 0$, $\sum_T p_T = 1$:

$$H(P) = \min_Q \left\{ -\sum_{T \subseteq M} p_T \log \sum_{i \in T} q_i \right\},$$

где минимум берется по наборам $Q = (q_i, i \in M)$, $q_i \geq 0$, $\sum_i q_i = 1$. Для полностью определенных данных, когда множествам T мощности $|T| \geq 2$ соответствуют $p_T = 0$, функция $H(P)$ совпадает с обычной энтропийной функцией $H(p_0, \ldots, p_{m-1}) = -\sum_{0 \leq i \leq m-1} p_i \log p_i$. Введем функционал $h_n(\boldsymbol{n}) = n\, H(\boldsymbol{n}/n)$, где $\boldsymbol{n}/n = (n_T/n, T \subseteq M)$. В [1] получены следующие оценки комбинаторной энтропии.

Утверждение 1. Существует константа $c = c(m)$, такая, что комбинаторная энтропия класса $K_n(\boldsymbol{n})$ заключена в пределах

$$h_n(\boldsymbol{n}) - c \log n \leq \log N_n(\boldsymbol{n}) \leq h_n(\boldsymbol{n}) + c \log n.$$

В ряде случаев возникают ограничения на вид используемых доопределений (см., например, [2]). Для набора натуральных чисел $\boldsymbol{s} = (s_0, \ldots, s_{m-1})$, $s_0 + \ldots + s_{m-1} = n$, обозначим через $K_n(\boldsymbol{s})$ класс всех последовательностей из $(A_0)^n$, в которых символы a_i, $i \in M$, встречаются s_i раз. Считаем, что набор $\boldsymbol{s}$ *совместим* с $\boldsymbol{n}$, если в $K_n(\boldsymbol{s})$ имеются доопределения всех (эквивалентно – хотя бы одной) последовательностей из $K_n(\boldsymbol{n})$. Будем говорить, что *множество* $\boldsymbol{S} \subseteq (A_0)^n$ *совместимо с* $\boldsymbol{n}$, если любой набор из $\boldsymbol{S}$ совместим с $\boldsymbol{n}$. Пусть $K_n(\boldsymbol{S}) = \bigcup_{s \in \boldsymbol{S}} K_n(\boldsymbol{s})$. Для $\boldsymbol{S}$, совместимого с $\boldsymbol{n}$, обозначим через $N_n(\boldsymbol{n})_{\boldsymbol{S}}$ минимальное число последовательностей из $K_n(\boldsymbol{S})$, доопределяющих класс $K_n(\boldsymbol{n})$, и величину $\log N_n(\boldsymbol{n})_{\boldsymbol{S}}$ назовем ***S*-*ограниченной комбинаторной энтропией*** класса $K_n(\boldsymbol{n})$.

В работе [2] приведены оценки $\boldsymbol{S}$-ограниченной энтропии для возникших там случаев – когда множество $\boldsymbol{S}$ одноэлементно (теорема 6) и еще в одной ситуации (теорема 7 без доказательства). В данной работе рассматривается произвольное множество $\boldsymbol{S}$, совместимое с $\boldsymbol{n}$. Чтобы сформулировать основной результат, введем ряд понятий.

Пусть $R = \|r_{Ti}\|$ – матрица, столбцы которой соответствуют множествам $T \subseteq M$, строки – элементам $i \in M$. Элементы матрицы удовлетворяют условиям $r_{Ti} \geq 0$, $r_{Ti} = 0$ для $i \notin T$, $\sum_{T,i} r_{Ti} = 1$. С матрицей R свяжем функцию

$$I(R) = \sum_{T,i} r_{Ti} \log \frac{r_{Ti}}{\sum_T r_{Ti} \sum_i r_{Ti}}.$$

Если рассматривать R как матрицу совместных вероятностей, то $I(R)$ – величина взаимной информации [3]. Пусть заданы наборы $P = (p_T, T \subseteq M)$, $p_T \geq 0$, $\sum_T p_T = 1$, и некоторое множество $\boldsymbol{Q}$ наборов $Q = (q_i, i \subseteq M)$, $q_i \geq 0$, $\sum_i q_i = 1$. Будем говорить, что *матрица R согласована с P при ограничениях $\boldsymbol{Q}$*, если $\sum_i r_{Ti} = p_T$ и набор $Q = (q_0, \ldots, q_{m-1})$, $q_i = \sum_T r_{Ti}$, содержится в $\boldsymbol{Q}$. Величину $\inf_R I(R)$ по всем матрицам R, согласованным с P при ограничениях $\boldsymbol{Q}$, обозначим через $H(P)_{\boldsymbol{Q}}$. В частности, если множество $\boldsymbol{Q}$ состоит из одного набора Q, будем использовать обозначение $H(P)_Q$. Значение $H(P)_Q$ достижимо, и вместо $\inf_R$ в этом случае можно использовать $\min_R$. Отметим, что функции $H(P)_W$ подобного типа для различного рода ограничений W на вид совместных распределений возникают в задачах кодирования с заданным критерием верности и носят название W-энтропии [4].

Для множества $\boldsymbol{S}$, согласованного с $\boldsymbol{n}$, введем функционал $h_n(\boldsymbol{n})_{\boldsymbol{S}} = n\, H(\boldsymbol{n}/n)_{\boldsymbol{S}/n}$, где $\boldsymbol{S}/n$ означает множество наборов $\boldsymbol{s}/n$, $\boldsymbol{s} \in \boldsymbol{S}$. Если $\boldsymbol{S}$ состоит из одного набора $\boldsymbol{s}$, будем использовать обозначение $h_n(\boldsymbol{n})_{\boldsymbol{s}}$. Так как число наборов длины n конечно и величина $h_n(\boldsymbol{n})_{\boldsymbol{s}}$ достижима, то $h_n(\boldsymbol{n})_{\boldsymbol{S}} = \min_{\boldsymbol{s} \in \boldsymbol{S}} h(\boldsymbol{n})_{\boldsymbol{s}}$. Сформулируем основной результат работы.

Теорема. Существует константа $c = c(m)$, такая, что для любых $\boldsymbol{n}$ и $\boldsymbol{S}$, где $\boldsymbol{S}$ согласовано с $\boldsymbol{n}$, $\boldsymbol{S}$-ограниченная комбинаторная энтропия класса $K_n(\boldsymbol{n})$ заключена в пределах

$$h_n(\boldsymbol{n})_{\boldsymbol{S}} - c \log n \leq \log N_n(\boldsymbol{n})_{\boldsymbol{S}} \leq h_n(\boldsymbol{n})_{\boldsymbol{S}} + c \log n.$$

Для доказательства понадобится следующий факт.

Лемма. Если для целочисленных наборов $\boldsymbol{n} = (n_T, T \subseteq M)$ и $\boldsymbol{s} = (s_i, i \in M)$, таких, что $\sum_T n_T = \sum_i s_i$, система уравнений относительно переменных w_{Ti}

$$\sum_i w_{Ti} = n_T \ (T \subseteq M), \qquad \sum_T w_{Ti} = s_i \ (i \in M), \qquad w_{Ti} = 0 \ (i \notin T) \tag{*}$$

имеет неотрицательное решение, то для всякого такого решения $\boldsymbol{w} = (w_{Ti}, T \subseteq M, i \in M)$ найдется целочисленное неотрицательное решение $\boldsymbol{w}^0 = (w_{Ti}^0, T \subseteq M, i \in M)$, для которого $\left|w_{Ti}^0 - w_{Ti}\right| \leq 1$ при всех T и i.

Доказательство. Рассмотрим решение $\boldsymbol{w}$ и положим $u_{Ti} = \lfloor w_{Ti} \rfloor$, $w'_{Ti} = w_{Ti} - u_{Ti}$, $n'_T = \sum_i w'_{Ti}$, $s'_i = \sum_T w'_{Ti}$. Очевидно, что числа n'_T и s'_i будут целыми и что выполнены уравнения, полученные из (*) заменой всех величин штрихованными, а также условие $\sum_T n'_T = \sum_i s'_i$. Построим ориентированную потоковую сеть [5] с полюсами s (источник) и t (сток), внутренними вершинами α_T, $T \subseteq M$, и β_i, $i \in M$, дугами (s, α_T) с пропускными способностями n'_T, дугами (α_T, β_i), $i \notin T$, с пропускными способностями 1 и (β_i, t) с пропускными способностями s'_i. Поток в этой сети не превосходит $\sum_T n'_T = \sum_i s'_i$, и эта величина достигается при назначении потоков в дугах (α_T, β_i) равными w'_{Ti}. По теореме Форда – Фалкерсона в этой сети существует максимальный целочисленный поток. Обозначив через w''_{Ti} $(w''_{Ti} \in \{0,1\})$ соответствующие потоки в дугах (α_T, β_i) и положив $w_{Ti}^0 = u_{Ti} + w''_{Ti}$, получим величины w_{Ti}^0, которые, как нетрудно видеть, удовлетворяют условиям леммы.

Доказательство теоремы. Нижняя оценка. Обозначим через $t_{\boldsymbol{S}}(\boldsymbol{n})$ максимальное число последовательностей из $K_n(\boldsymbol{n})$, доопределимых одной последовательностью из $K_n(\boldsymbol{S})$. Рассмотрим последовательности $\boldsymbol{x} \in K_n(\boldsymbol{n})$ и $\boldsymbol{y} \in K_n(\boldsymbol{S})$, такие, что $\boldsymbol{x}$ доопределяет $\boldsymbol{y}$. Пусть $\boldsymbol{y} \in K_n(\boldsymbol{S})$, $\boldsymbol{s} = (s_i, i \in M)$. Обозначим через w_{Ti} число символов a_T в $\boldsymbol{x}$, доопределенных в $\boldsymbol{y}$ символом a_i. Числа w_{Ti} удовлетворяют условиям (*). При заданных w_{Ti} последовательность $\boldsymbol{y}$ доопределяет

$$\frac{s_0!}{\prod_T w_{T0}!} \cdots \frac{s_{m-1}!}{\prod_T w_{T,m-1}!} = \frac{\prod_i s_i!}{\prod_{T,i} w_{Ti}!}$$

последовательностей из $K_n(\boldsymbol{n})$, а всего она доопределяет

$$t_{\boldsymbol{s}}(\boldsymbol{n}) = \sum_{\{w_{Ti}\},(*)} \frac{\prod_i s_i!}{\prod_{T,i} w_{Ti}!}$$

последовательностей из этого класса, где сумма берется по всем наборам неотрицательных чисел w_{Ti}, удовлетворяющих условиям (*). В силу $0 \le w_{Ti} \le n$ и того, что количества индексов i и множеств T ограничены константами (зависящими от m), имеем

$$t_{\boldsymbol{s}}(\boldsymbol{n}) \le n^{c_1} \max_{\{w_{Ti}\},(*)} \frac{\prod_i s_i!}{\prod_{T,i} w_{Ti}!},$$

где $c_1 = c_1(m)$ – константа. Отсюда

$$t_{\boldsymbol{S}}(\boldsymbol{n}) = \max_{\boldsymbol{s}\in\boldsymbol{S}} t_{\boldsymbol{s}}(\boldsymbol{n}) \le n^{c_1} \max_{\boldsymbol{s}\in\boldsymbol{S}} \max_{\{w_{Ti}\},(*)} \frac{\prod_i s_i!}{\prod_{T,i} w_{Ti}!}.$$

Класс $K_n(\boldsymbol{n})$ содержит $n!\big/\prod_T n_T!$ последовательностей. Отсюда и из предшествующего неравенства заключаем, что

$$N_n(\boldsymbol{n})_{\boldsymbol{S}} \ge \frac{|K_n(\boldsymbol{n})|}{t_{\boldsymbol{S}}(\boldsymbol{n})} \ge n^{-c_1} \min_{\boldsymbol{s}\in\boldsymbol{S}} \min_{\{w_{Ti}\},(*)} \frac{n!\prod_{T,i} w_{Ti}!}{\prod_T n_T!\prod_i s_i!}.$$

Из формулы Стирлинга следует, что для любых целых $z, z_1, \ldots, z_k$, $z \ge 2$, $z_1 + \ldots + z_k = z$, выполнено $\log\left(z!\big/\prod_j z_j!\right) = z\log z - \sum_j z_j \log z_j + \theta \log z$, $-c_2 \le \theta \le c_2$, $c_2 = c_2(k)$. С учетом этого получаем

$$\log N_n(\boldsymbol{n})_{\boldsymbol{S}} \ge \min_{\boldsymbol{s}\in\boldsymbol{S}} \min_{\{w_{Ti}\},(*)} \left(n\log n - \sum_T n_T \log n_T - \sum_i s_i \log s_i + \sum_{T,i} w_{Ti} \log w_{Ti} \right) - c\log n.$$

Используя равенства $\sum_T n_T = \sum_i s_i = \sum_{T,i} w_{Ti} = n$ и (*), преобразуем минимизируемое выражение к виду

$$n\left(-\sum_{T,i} \frac{w_{Ti}}{n} \log\frac{w_{Ti}}{n} + \sum_T \frac{n_T}{n}\log\frac{n_T}{n} + \sum_i \frac{s_i}{n}\log\frac{s_i}{n} \right) = n\,\mathrm{I}\left(\left\|\frac{w_{Ti}}{n}\right\|\right),$$

откуда

$$\log N_n(\boldsymbol{n})_{\boldsymbol{S}} \ge \min_{\boldsymbol{s}\in\boldsymbol{S}} \min_{\{w_{Ti}\},(*)} \left\{ n\,\mathrm{I}\left(\left\|\frac{w_{Ti}}{n}\right\|\right) \right\} - c\log n \ge$$
$$\ge \min_{\boldsymbol{s}\in\boldsymbol{S}} nH(\boldsymbol{n}/n)_{\boldsymbol{s}/n} - c\log n = \min_{\boldsymbol{s}\in\boldsymbol{S}} h_n(\boldsymbol{n})_{\boldsymbol{s}} - c\log n = h_n(\boldsymbol{n})_{\boldsymbol{S}} - c\log n.$$

В е р х н я я о ц е н к а. Рассмотрим значение $\boldsymbol{s}$, при котором $h_n(\boldsymbol{n})_{\boldsymbol{s}} = h_n(\boldsymbol{n})_{\boldsymbol{S}}$. Занумеруем последовательности $\boldsymbol{x} \in K_n(\boldsymbol{n})$ и $\boldsymbol{y} \in K_n(\boldsymbol{s})$ индексами $r = 1, 2, \ldots, n!\big/\prod_T n_T!$ и $q = 1, 2, \ldots, n!\big/\prod_i s_i!$ соответственно. Образуем таблицу $\|d_{qr}\|$, строки которой соответствуют последовательностям $\boldsymbol{y}_q$, столбцы – последовательностям $\boldsymbol{x}_r$, положив $d_{qr} = 1$ в случае, когда $\boldsymbol{y}_q$ доопределяет $\boldsymbol{x}_r$, и $d_{qr} = 0$, когда нет. Все столбцы таблицы содержат одинаковое число единиц; обозначим его $b_{\boldsymbol{s}}(\boldsymbol{n})$. Оценим $b_{\boldsymbol{s}}(\boldsymbol{n})$ снизу.

Пусть $R = \|r_{Ti}\|$ – матрица, на которой достигается значение $H(\boldsymbol{n}/n)_{\boldsymbol{s}/n}$. Числа $w_{Ti} = nr_{Ti}$ удовлетворяют уравнениям (*). По лемме найдется целочисленное решение $\left\{w_{Ti}^0\right\}$ уравнений (*), компоненты которого отличаются от w_{Ti} не более чем на 1. Для произвольного $\boldsymbol{x}_r = K_n(\boldsymbol{n})$ число доопределений $\boldsymbol{y}_q \in K_n(\boldsymbol{s})$, в которых число символов a_T, доопределяемых символом a_i, равно w_{Ti}^0, составляет $\prod_T n_T!\big/\prod_{T,i} w_{Ti}^0!$. Поэтому $b_{\boldsymbol{s}}(\boldsymbol{n}) \ge \prod_T n_T!\big/\prod_{T,i} w_{Ti}^0!$.

В матрице из 0 и 1 с u строками и v столбцами, содержащей не менее s единиц в каждом столбце, можно выделить множество из не более $\frac{u}{s}(\ln\frac{vs}{u} + 1) + 1$ строк, содержащих 1 в каждом столбце [6]. Из этого результата при

$$u = \frac{n!}{\prod_i s_i!}, \qquad v = \frac{n!}{\prod_T n_T!}, \qquad s = \frac{\prod_T n_T!}{\prod_{T,i} w_{Ti}^0!}$$

получаем оценку

$$N_n(\boldsymbol{n})_{\boldsymbol{S}} \le N_n(\boldsymbol{n})_{\boldsymbol{s}} \le c_1 n \frac{n! \prod_{T,i} w_{Ti}^0!}{\prod_T n_T! \prod_i s_i!},$$

где $c_1 = c_1(m)$ – некоторая константа. Применение формулы Стирлинга дает

$$\log N_n(\boldsymbol{n})_{\boldsymbol{S}} \le n \log n - \sum_T n_T \log n_T - \sum_i s_i \log s_i + \sum_{T,i} w_{Ti}^0 \log w_{Ti}^0 + c_2 \log n.$$

Нетрудно убедиться непосредственно, что при изменении числа x не более чем на 1 величина $x \log x$ изменяется не более чем на $O(\log x)$. С учетом этого и того, что количества индексов T и i ограничены константами (зависящими от m), перейдем от величин w_{Ti}^0 к $w_{Ti} = nr_{Ti}$:

$$\log N_n(\boldsymbol{n})_{\boldsymbol{S}} \le n \log n - \sum_T n_T \log n_T - \sum_i s_i \log s_i + \sum_{T,i} w_{Ti} \log w_{Ti} + c_3 \log n,$$

изменив константу перед $\log n$. Подобно случаю нижней оценки это неравенство может быть преобразовано к виду

$$\log N_n(\boldsymbol{n})_{\boldsymbol{S}} \le n \,\mathrm{I}\left(\left\| \frac{w_{Ti}}{n} \right\|\right) + c_3 \log n.$$

Принимая во внимание, что

$$n \,\mathrm{I}\left(\left\| \frac{w_{Ti}}{n} \right\|\right) = n \,\mathrm{I}(\|r_{Ti}\|) = nH(\boldsymbol{n}/n)_{\boldsymbol{s}/n} = h_{\boldsymbol{s}}(\boldsymbol{n}) = h_{\boldsymbol{S}}(\boldsymbol{n}),$$

приходим к нужной верхней оценке.

Теорема доказана.

Рассмотрим функцию

$$H(P,Q) = -\sum_T p_T \log \sum_i q_i,$$

участвующую в определении функции $H(P)$, и функцию $H(P)_Q$, введенную выше. Из [1, 2] следует, что $\min_Q H(P, Q) = \min_Q H(P)_Q$ $(= H(P))$ и, более того, минимумы достигаются в одной и той же точке Q_0. Функция $H(P, Q)$ задана явно, в то время как $H(P)_Q$ определена как минимум функции $I(R)$ по матрицам R, удовлетворяющим определенным условиям. В связи с этим возникает вопрос о взаимоотношении функций $H(P, Q)$ и $H(P)_Q$ и возможности замены в теореме функционала $h_n(\boldsymbol{n})_{\boldsymbol{S}}$ вычислительно более простым функционалом $h_n(\boldsymbol{n}, \boldsymbol{S})$, где $h_n(\boldsymbol{n}, \boldsymbol{S}) = \min_{\boldsymbol{s}\in\boldsymbol{S}} h_n(\boldsymbol{n}, \boldsymbol{s})$, $h_n(\boldsymbol{n}, \boldsymbol{s}) = n\, H(\boldsymbol{n}/n, \boldsymbol{s}/n)$.

Утверждение 2. Имеет место неравенство $H(P, Q) \le H(P)_Q$.

Доказательство. Пусть значение $H(P)_Q$ достигается на матрице $R = \|r_{Ti}\|$. Ее элементы удовлетворяют соотношениям $\sum_i r_{Ti} = p_T$, $\sum_T r_{Ti} = q_i$. При заданном T, воспользовавшись для выпуклой функции $f(x) = x \log x$ неравенством Иенсена $\sum_i \alpha_i f(x_i) \ge f\left(\sum_i \alpha_i x_i\right)$ при $\alpha_i = q_i \big/ \sum_{j\in T} q_j$, $x_i = r_{Ti} / p_T q_i$, получаем с учетом $r_{Ti} = 0$ для $i \notin T$

$$\sum_i r_{Ti} \log \frac{r_{Ti}}{p_T q_i} = p_T \left(\sum_{j\in T} q_j\right) \sum_i \frac{q_i}{\sum_{j\in T} q_j} \frac{r_{Ti}}{p_T q_i} \log \frac{r_{Ti}}{p_T q_i} \ge p_T \log \frac{1}{\sum_{j\in T} q_j}.$$

Суммирование этих неравенств по T дает $H(P)_Q \ge H(P, Q)$. Утверждение доказано.

Из него вытекает, что $h_n(\boldsymbol{n}, \boldsymbol{S}) \le h_n(\boldsymbol{n})_{\boldsymbol{S}}$, и в силу теоремы получаем следующий факт.

Следствие. Существует константа $c = c(m)$, такая, что

$$\log N_n(\boldsymbol{n})_{\boldsymbol{S}} \ge h_n(\boldsymbol{n}, \boldsymbol{S}) - c \log n.$$

Таким образом, более просто вычислимый функционал $h_n(\boldsymbol{n}, \boldsymbol{S})$ может быть использован в нижней оценке $\boldsymbol{S}$-ограниченной комбинаторной энтропии. Следующий пример показывает, что он не применим в качестве верхней оценки.

Пример. Рассмотрим некоторый класс $K_n(\boldsymbol{n})$ *частично определенных последовательностей*, т.е. последовательностей в алфавите $A = \{a_0, \ldots, a_{m-1}, \bullet\}$, где $\bullet$ – н*еопределенный символ*, допускающий доопределение любым символом алфавита $A_0 = \{a_0, \ldots, a_{m-1}\}$ (в прежних обозначениях он совпадает с a_M). Пусть

$\boldsymbol{n} = (n_0, \ldots, n_{m-1}, n_\bullet)$ и доопределения берутся из $K_n(\boldsymbol{s})$, $\boldsymbol{s} = (s_0, \ldots, s_{m-1})$, $s_0 \geq n_0, \ldots, s_{m-1} \geq n_{m-1}$. Удобнее иметь дело с частотами, поэтому положим $\boldsymbol{n}/n = P = (p_0, \ldots, p_{m-1}, p_\bullet)$, $P' = (p_0, \ldots, p_{m-1})$, $\boldsymbol{s}/n = Q = (q_0, \ldots, q_{m-1})$.

Поскольку $\log \sum_i q_i = 0$, функция $H(P, Q)$ приобретает вид $H(P,Q) = -\sum_i p_i \log q_i$. Отсюда, учитывая $\sum_i p_i = 1 - p_\bullet$, по свойству энтропийной функции получаем

$$H(P,Q) = -(1-p_\bullet)\sum \frac{p_i}{1-p_\bullet}\log q_i \geq (1-p_\bullet)H\left(\frac{p_0}{1-p_\bullet},\ldots,\frac{p_{m-1}}{1-p_\bullet}\right) = (1-p_\bullet)H\left(\frac{P'}{1-p_\bullet}\right).$$

Равенство здесь достигается лишь при $Q_0 = P'/(1 - p_\bullet)$, поэтому $H(P) = H(P, Q_0) = (1 - p_\bullet)H(P'/(1 - p_\bullet))$.

В данном случае имеется единственная матрица R, согласованная с P при ограничении Q. Ее ненулевыми элементами являются $r_{ii} = p_i$, $r_{\bullet i} = q_i - p_i$, $0 \leq i \leq m - 1$. Следовательно,

$$H(P)_Q = I(R) = \sum_i p_i \log\frac{p_i}{p_i q_i} + \sum_i (q_i - p_i)\log\frac{q_i - p_i}{p_\bullet q_i} =$$

$$= -\sum_i q_i \log q_i + p_\bullet \sum_i \frac{q_i - p_i}{p_\bullet}\log\frac{q_i - p_i}{p_\bullet} = H(Q) - p_\bullet H\left(\frac{Q-P'}{p_\bullet}\right).$$

С учетом $\sum_i (q_i - p_i) = p_\bullet$ и свойств энтропийной функции оценим разность

$$H(P)_Q - H(P,Q) = -p_\bullet \sum_i \frac{q_i - p_i}{p_\bullet}\log q_i - p_\bullet H\left(\frac{Q-P'}{p_\bullet}\right) \geq 0.$$

Равенство достигается лишь при условии $Q = (Q - P')/p_\bullet$, т.е. при $Q = P'/(1 - p_\bullet) = Q_0$. Вычислим

$$H(P)_{Q_0} = H(Q_0) - p_\bullet H\left(\frac{Q_0 - P'}{p_\bullet}\right) = H\left(\frac{P'}{1-p_\bullet}\right) - p_\bullet H\left(\frac{P'}{1-p_\bullet}\right) = H(P).$$

Таким образом, функция $H(P, Q)$ всюду меньше функции $H(P)_Q$, исключая точку Q_0, где $\mathrm{H}(P)_{Q_0} = \mathrm{H}(P, Q_0) = \mathrm{H}(P)$. Применительно к функционалам это означает, что $h_n(\boldsymbol{n}, \boldsymbol{s})$ всюду меньше $h_n(\boldsymbol{n})_s$, исключая точку минимума, где их значения совпадают и равны $h_n(\boldsymbol{n})$.

ЛИТЕРАТУРА

1. *Шоломов Л.А.* Сжатие частично определенной информации // Нелинейная динамика и управление. Вып. 4. М.: Физматлит, 2004. С. 385 – 399.
2. *Шоломов Л.А.* О сложности последовательной реализации частичных булевых функций схемами // Дискретный анализ и исследование операций. Сер. 1. 2007. Т. 14. № 1. С. 110 – 139.
3. *Галлагер Р.* Теория информации и надежная связь. М.: Сов. радио, 1974.
4. *Вероятность* и математическая статистика: Энциклопедия. М.: БРЭ, 1999.
5. *Форд Л.*, *Фалкерсон Д.* Потоки в сетях. М.: Мир, 1966.
6. *Сапоженко А.А.*, *Асратян А.С.*, *Кузюрин Н.Н.* Обзор некоторых результатов по задачам о покрытии // Методы дискретного анализа в решении комбинаторных задач. Вып. 30. Новосибирск: ИМ СО АН СССР, 1977. С. 46 – 75.