Кибернетический сборник

НОВАЯ СЕРИЯ

ВЫПУСК

19

Сборник переводов
под редакцией
О. Б. ЛУПАНОВА

МОСКВА «МИР»
1983

ББК 32.81
К38
УДК 519.95

Научный совет по кибернетике
Академии наук СССР

Кибернетический сборник. Новая серия. Вып. 19. Сб. статей: Пер. с англ. — М.: Мир, 1983 — 200 с., ил.

Продолжение серии, начатой издательством «Мир» в 1965 г. В выпуске содержатся обзорные статьи и оригинальные работы известных зарубежных ученых по наиболее актуальным проблемам теоретической кибернетики. Большой интерес представляют статьи Р. Харалика по теории распознавания и Д. Кнута и Э. Яо о сложности моделирования неравномерных распределений.

Для научных работников, инженеров-исследователей, аспирантов и студентов, специализирующихся по теоретической кибернетике и ее приложениям.

Редакция литературы по математическим наукам

К $\frac{1702070000\text{-}050}{041\ (01)\text{-}83}$ 34-82, ч. 1

© «Мир», 1983

О шенноновской емкости графа[1])

Ласло Ловас

Доказано, что шенноновская емкость пятиугольника равна $\sqrt{5}$. Затем метод обобщается с целью получения верхних границ для емкости произвольного графа. Вводится детально исследованная и в известном смысле просто вычисляемая функция, которая является верхней границей емкости и во многих случаях равна ей. Получены некоторые результаты о емкости специальных графов; например, емкость графа Петерсена равна 4, а самодополнительный граф с n вершинами, группа автоморфизмов которого транзитивно действует на вершинах, имеет емкость $\sqrt{n}$.

1. ВВЕДЕНИЕ

Рассмотрим граф G, вершинами которого являются буквы алфавита, а смежность вершин означает, что буквы могут быть перепутаны. Тогда максимальное число однобуквенных сообщений, которые можно послать, не опасаясь перепутать их, очевидно, равно $\alpha(G)$ — максимальному числу независимых вершин в графе G. Обозначим через $\alpha(G^k)$ максимальное число k-буквенных сообщений, которые можно послать, не опасаясь их перепутать (два k-буквенных слова перепутываемы, если для каждого $1 \leqslant i \leqslant k$ их i-е буквы перепутываемы или равны). Ясно, что существует по крайней мере $\alpha(G^k)$ таких слов (образованных из максимального множества неперепутываемых букв), но их может быть и больше. Например, для пятиугольника C_5 справедливо $\alpha(C_5^2) = 5$. Действительно, если $v_1, \ldots, v_5$ — вершин в графе G. Обозначим через $\alpha(G)_k$ максимальное число v_2v_3, v_3v_5, v_4v_2 и v_5v_4 неперепутываемы.

Легко видеть, что

$$\Theta(G) = \sup_k \sqrt[k]{\alpha(G^k)} = \lim_{k\to\infty} \sqrt[k]{\alpha(G^k)}.$$

Эта величина была введена Шенноном [6] и называется *шенноновской емкостью* графа G. Предыдущее рассмотрение пока-

[1]) László Lovász. On the Shannon Capacity of a Graph, IEEE Transactions on Information Theory, IT-25, 1, 1979, 1—7.

© 1979 IEEE. Reprinted, with permission, from IEEE Transaction on Information Theory, vol. IT-25, No. 1, pp. 1—7, January 1979
© Перевод на русский язык, «Мир», 1983

зывает, что $\Theta(G) \geqslant \alpha(G)$ и что в общем случае равенство не выполняется.

Вычисление шенноновской емкости является очень трудной задачей даже для очень простых графов с небольшим числом вершин. Шеннон доказал, что $\alpha(G) = \Theta(G)$ для графов, которые могут быть покрыты $\alpha(G)$ кликами (наиболее известными такими графами являются так называемые совершенные графы; см. [1]). Однако уже для простейшего графа, не подпадающего под этот результат, — пятиугольника — шенноновская емкость была до настоящего времени неизвестна.

В [6] была также дана общая верхняя граница для $\Theta(G)$ (эта граница подробно обсуждена Розенфельдом [5]). Припишем вершинам x графа G неотрицательные веса $w(x)$ так, чтобы

$$\sum_{x \in C} w(x) \leqslant 1,$$

для любого полного подграфа C графа G такой набор весов называется *дробной упаковкой вершин*. Максимальное значение $\Sigma_x w(x)$, взятое по всем дробным упаковкам вершин, обозначается через $\alpha^*(G)$. Из теоремы двойственности линейного программирования легко получить, что величина $\alpha^*(G)$ может быть определена двойственным образом как минимальное значение $\Sigma_C q(C)$, взятое по всем распределениям неотрицательных весов $q(C)$ на кликах C графа G таким, что для любой вершины x

$$\sum_{C \ni x} q(C) \geqslant 1.$$

В этих обозначениях теорема Шеннона утверждает, что

$$\Theta(G) \leqslant \alpha^*(G).$$

В случае пятиугольника этот результат и замечание, сделанное выше, приводят к следующим границам:

$$\sqrt{5} \leqslant \Theta(C_5) \leqslant 5/2.$$

Мы докажем, что нижняя граница является точной. Это будет достигнуто путем получения общей верхней границы для $\Theta(G)$. Эта верхняя граница детально исследована и в известном смысле ее просто вычислять. Наши методы дадут также возможность вычислить или оценить емкость других графов, например показать, что емкость графа Петерсена равна 4.

2. ЕМКОСТЬ ПЯТИУГОЛЬНИКА

Пусть G — конечный неориентированный граф без петель. Будем говорить, что две вершины графа G *смежны*, если они соединены ребром или равны.

Множество вершин графа G обозначается через $V(G)$. *Дополнительный граф* к графу G определяется как граф $\bar{G}$ с множеством вершин $V(\bar{G}) = V(G)$, в котором две вершины соединены ребром тогда и только тогда, когда они не соединены ребром в графе G. Разбиение $V(G)$ на k подмножеств, каждое из которых независимо, называется *k-раскрашиванием* графа G. Отметим, что это соответствует покрытию k кликами вершин дополнительного графа. Наименьшее k, для которого граф G допускает k-раскрашивание, называется его *хроматическим числом*.

Подстановка на множестве $V(G)$ есть *автоморфизм*, если она сохраняет смежность вершин. Автоморфизмы графа G образуют группу подстановок, называемую *группой автоморфизмов графа G*. Если для любой пары вершин x, $y \in V(G)$ существует автоморфизм, переводящий x в y, то группа автоморфизмов называется *вершинно-транзитивной*. Аналогичным образом определяется *реберно-транзитивная* группа. Граф называется *регулярным* степени d, если каждая вершина инцидентна d ребрам. Заметим, что граф с вершинно-транзитивной группой автоморфизмов является регулярным. Это не обязательно справедливо при реберно-транзитивной группе (как, например, в случае звезды).

Для графов G и H их сильное произведение $G \cdot H$ определяется как граф с множеством вершин $V(G \cdot H) = V(G) \times V(H)$, в котором (x, y) смежна с (x', y') тогда и только тогда, когда x смежна с x' в графе G, и y смежна с y' в графе H. Если обозначить через G^k сильное произведение k копий графа G, то $\alpha(G^k)$ есть несомненно максимальное число независимых вершин в графе G^k.

Мы будем интенсивно использовать линейную алгебру. Относительно различных свойств матриц (в основном положительно определенных) см., например, книгу [4]. Все векторы рассматриваются как вектор-столбцы. Мы обозначаем через I единичную матрицу, через J квадратную матрицу, состоящую из одних единиц, и через j вектор, состоящий из одних единиц (размерность этих матриц и векторов будет ясна из контекста).

Кроме скалярного произведения векторов $\mathbf{v}$, $\mathbf{w}$ (обозначаемого через $\mathbf{v}^T\mathbf{w}$, где T означает транспонирование), мы будем использовать *тензорное произведение*, определяемое следующим образом. Для векторов $\mathbf{v} = (v_1, \ldots, v_n)$ и $\mathbf{w} = (w_1, \ldots, w_m)$ через $\mathbf{v} \circ \mathbf{w}$ обозначается nm-мерный вектор $(v_1w_1, \ldots, v_1w_m, v_2w_1, \ldots, v_nw_m)^T$. Простое вычисление показывает, что два этих вида произведения векторов связаны соотношением

$$(x \circ y)^T (v \circ w) = (x^T v)(y^T w). \tag{1}$$

Рассмотрим граф G. Для простоты мы всегда будем предполагать, что его вершинами являются $1, \ldots, n$. *Ортонормированное представление* графа G есть система $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ единичных (т. е. длины 1) векторов в евклидовом пространстве, такая, что если вершины i и j несмежны, то векторы $\mathbf{v}_i$ и $\mathbf{v}_j$ ортогональны. Очевидно, что каждый граф имеет ортонормированное представление, например систему попарно ортогональных единичных векторов.

Лемма 1. *Пусть* $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ *и* $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ *— ортонормированные представления графов* G *и* H *соответственно. Тогда векторы* $\mathbf{u}_i \circ \mathbf{v}_j$ *образуют ортонормированное представление графа* $G \cdot H$.

Доказательство немедленно следует из (1).

Определим *значение* ортонормированного представления $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ как

$$\min_{\mathbf{c}} \max_{1 \leqslant i \leqslant n} \frac{1}{(\mathbf{c}^T \mathbf{u}_i)^2},$$

где $\mathbf{c}$ пробегает множество всех единичных векторов. Вектор $\mathbf{c}$, на котором этот минимум достигается, назовем *ручкой* представления. Пусть $\vartheta(G)$ обозначает минимум значений всех представлений графа G. Легко видеть, что этот минимум достигается. Назовем представление *оптимальным*, если его значение минимально.

Лемма 2. $\vartheta(G \cdot H) \leqslant \vartheta(G)\vartheta(H)$.

Доказательство. Пусть $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ и $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ — оптимальные представления графов G и H с ручками $\mathbf{c}$ и $\mathbf{d}$ соответственно. Тогда, согласно (1), $\mathbf{c} \circ \mathbf{d}$ есть единичный вектор и, следовательно,

$$\vartheta(G \cdot H) \leqslant \max_{i,j} \frac{1}{((\mathbf{c} \circ \mathbf{d})^T (\mathbf{u}_i \circ \mathbf{v}_j))^2} = \max_{i,j} \frac{1}{(\mathbf{c}^T \mathbf{u}_i)^2} \frac{1}{(\mathbf{d}^T \mathbf{v}_j)^2} = \vartheta(G)\vartheta(H).$$

Замечание. Ниже мы увидим, что в лемме 2 имеет место равенство.

Лемма 3. $\alpha(G) \leqslant \vartheta(G)$.

Доказательство. Пусть $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ — оптимальное ортонормированное представление графа G с ручкой $\mathbf{c}$. И пусть $\{1, \ldots, k\}$, например, есть максимальное независимое множество в графе G. Тогда векторы $\mathbf{u}_1, \ldots, \mathbf{u}_k$ попарно ортогональны, и

таким образом справедливо

$$1 = \mathbf{c}^2 \geqslant \sum_{i=1}^{k} (\mathbf{c}^T \mathbf{u}_i)^2 \geqslant \alpha(G)/\vartheta(G).$$

Теорема 1. $\Theta(G) \leqslant \vartheta(G)$.

Доказательство. Согласно леммам 2 и 3, $\alpha(G^k) \leqslant \vartheta(G^k) \leqslant \vartheta(G)^k$.

Теорема 2. $\Theta(C_5) = \sqrt{5}$.

Доказательство. Рассмотрим зонтик, ручка и пять ребер которого имеют длину 1. Откроем зонтик так, чтобы максимальный угол между ребрами был равен $\pi/2$. Пусть $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$, $\mathbf{u}_4$, $\mathbf{u}_5$ будут ребра и $\mathbf{c}$ — ручка, рассматриваемые как векторы, направленные от их общей точки. Тогда $\mathbf{u}_1, \ldots, \mathbf{u}_5$ образуют ортонормированное представление графа G. Кроме того, легко подсчитать с помощью теоремы о сферических косинусах, что $\mathbf{c}^T \mathbf{u}_i = 5^{-1/4}$ и, следовательно,

$$\Theta(C_5) \leqslant \vartheta(C_5) \leqslant \max_i \frac{1}{(\mathbf{c}^T \mathbf{u}_i)^2} = \sqrt{5}.$$

Противоположное неравенство является известным, и тем самым теорема доказана.

3. ФОРМУЛЫ ДЛЯ $\vartheta(G)$

Чтобы иметь возможность применить теорему 1 для вычисления или оценки шенноновской емкости других графов, мы должны исследовать величину $\vartheta(G)$ более подробно.

Теорема 3. *Пусть G — граф с вершинами $\{1, \ldots, n\}$. Тогда величина $\vartheta(G)$ равна минимальному среди наибольших собственных значений произвольных симметрических матриц $(a_{ij})_{i,j=1}^n$, таких, что*

$$a_{ij} = 1, \quad \text{если } i = j \text{ или } i \text{ и } j \text{ несмежны.} \tag{2}$$

Доказательство.

1) Пусть $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ — ортонормированное представление графа G с ручкой $\mathbf{c}$. Определим матрицу $(A) = (a_{ij})_{i,j=1}^n$ следующим образом:

$$a_{ij} = 1 - \frac{\mathbf{u}_i^T \mathbf{u}_j}{(\mathbf{c}^T \mathbf{u}_i)(\mathbf{c}^T \mathbf{u}_j)}, \qquad i \neq j,$$
$$a_{ii} = 1,$$

Тогда (2) выполнено. Кроме того,

$$-a_{ij}=\left(\mathbf{c}-\frac{\mathbf{u}_i}{(\mathbf{c}^T\mathbf{u}_i)}\right)^T\left(\mathbf{c}-\frac{\mathbf{u}_j}{(\mathbf{c}^T\mathbf{u}_j)}\right),\qquad i\neq j,$$

и

$$\vartheta(G)-a_{ii}=\left(\mathbf{c}-\frac{\mathbf{u}_i}{\mathbf{c}^T\mathbf{u}_i}\right)^2+\left(\vartheta(G)-\frac{1}{(\mathbf{c}^T\mathbf{u}_i)^2}\right).$$

Эти уравнения означают, что матрица $\vartheta(G)I-A$ положительно полуопределенная и, следовательно, наибольшее собственное значение матрицы A не превосходит $\vartheta(G)$.

2) Обратно, пусть $A=(a_{ij})$ — произвольная матрица, удовлетворяющая условиям (2), а λ — ее наибольшее собственное значение. Тогда матрица $\lambda I-A$ положительно полуопределенная и, следовательно, существуют векторы $\mathbf{x}_1, \ldots, \mathbf{x}_n$, такие, что

$$\lambda\delta_{ij}-a_{ij}=\mathbf{x}_i^T\mathbf{x}_j.$$

Пусть $\mathbf{c}$ — единичный вектор, ортогональный векторам $\mathbf{x}_1, \ldots, \mathbf{x}_n$. Положим

$$\mathbf{u}_i=\frac{1}{\sqrt{\lambda}}(\mathbf{c}+\mathbf{x}_i).$$

Тогда

$$\mathbf{u}_i^2=\frac{1}{\lambda}(1+\mathbf{x}_i^2)=1,\qquad i=1, \ldots, n,$$

а для несмежных i и j

$$\mathbf{u}_i^T\mathbf{u}_j=\frac{1}{\lambda}(1+\mathbf{x}_i^T\mathbf{x}_j)=0.$$

Таким образом $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ есть ортонормированное представление графа G. Кроме того,

$$\frac{1}{(\mathbf{c}^T\mathbf{u}_i)^2}=\lambda,\qquad i=1, \ldots, n,$$

и отсюда $\vartheta(G)\leqslant\lambda$. Это и завершает доказательство теоремы.

Отметим, что, как следует из доказательства, среди оптимальных представлений существует такое, что

$$\vartheta(G)=\frac{1}{(\mathbf{c}^T\mathbf{u}_1)^2}=\ldots=\frac{1}{(\mathbf{c}^T\mathbf{u}_n)^2}.$$

Следующая теорема дает хорошее описание величины $\vartheta(G)$.

Теорема 4. *Пусть G — граф с множеством вершин $\{1, \ldots, n\}$, и пусть $B=(b_{ij})_{i,j=1}^n$ пробегает множество всех положительно полуопределенных симметрических матриц, таких, что*

(3) $$b_{ij}=0$$

для любой пары i, j различных смежных вершин и

$$\operatorname{Tr} B = 1. \tag{4}$$

Тогда

$$\vartheta(G) = \max_{B} \operatorname{Tr} BJ.$$

Заметим, что $\operatorname{Tr} BJ$ *есть сумма всех элементов матрицы* B.

Доказательство:

1) Пусть $A = (a_{ij})_{i,\,j=1}^{n}$ — матрица, удовлетворяющая (2), и ее наибольшее собственное значение равно $\vartheta(G)$, и пусть B — произвольная симметрическая матрица, удовлетворяющая условиям (3) и (4). Тогда, используя (2) и (3), имеем

$$\operatorname{Tr} BJ = \sum_{i=1}^{n}\sum_{j=1}^{n} b_{ij} = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}b_{ij} = \operatorname{Tr} AB$$

и, таким образом

$$\vartheta(G) - \operatorname{Tr} BJ = \operatorname{Tr}(\vartheta(G)I - A)B.$$

Здесь обе матрицы: $\vartheta(G)I - A$ и B являются положительно полуопределенными. Пусть $\mathbf{e}_1, \ldots, \mathbf{e}_n$ — множество ортогональных собственных векторов матрицы B с соответствующими собственными значениями $\lambda_1, \ldots, \lambda_n \geqslant 0$. Тогда

$$\operatorname{Tr}(\vartheta(G)I - A)B = \sum_{i=1}^{n} \mathbf{e}_i^T(\vartheta(G)I - A)B\mathbf{e}_i =$$

$$= \sum_{i=1}^{n} \lambda_i \mathbf{e}_i^T(\vartheta(G)I - A)\,\mathbf{e}_i \geqslant 0.$$

2) Построим матрицу B, для которой предыдущее неравенство превратится в равенство. Пусть $(i_1, j_1), \ldots, (i_m, j_m)$ $(i_k < j_k)$ — ребра графа G. Рассмотрим $(m+1)$-мерные векторы

$$\hat{\mathbf{h}} = \left(h_{i_1}h_{j_1}, \ldots, h_{i_m}h_{j_m}, \left(\sum h_i\right)^2\right)^T,$$

где $\mathbf{h} = (h_1, \ldots, h_n)$ пробегает множество всех единичных векторов, и вектор

$$\mathbf{z} = (0, 0, \ldots, 0, \vartheta(G))^T.$$

Утверждение: вектор $\mathbf{z}$ лежит в выпуклой оболочке векторов $\hat{\mathbf{h}}$. Предположим, что это не так. Так как векторы $\mathbf{h}$ образуют компактное множество, то существует гиперплоскость, отделяющая $\mathbf{z}$ от всех $\hat{\mathbf{h}}$, т. е. существует вектор $\mathbf{a}$ и действительное число α, такие, что $\mathbf{a}^T\hat{\mathbf{h}} < \alpha$ при всех единичных векторах $\mathbf{h}$, но $\mathbf{a}^T\mathbf{z} > \alpha$.

Положим

$$\mathbf{a} = (a_1, \ldots, a_m, y)^T.$$

Тогда, в частности, $\mathbf{a}^T\hat{\mathbf{h}} \leqslant \alpha$ при $\mathbf{h} = (1, 0, \ldots, 0)$ и, следовательно, $y \leqslant \alpha$. С другой стороны, $\mathbf{a}^T\mathbf{z} > \alpha$ влечет $\vartheta(G)y > \alpha$. Отсюда имеем $y > 0$ и $\alpha > 0$. Мы можем предположить, что $y = 1$ и, таким образом, $\alpha < \vartheta(G)$.

Определим теперь матрицу

$$a_{ij} = \begin{cases} \frac{1}{2}a_k + 1, & \text{если } \{i, j\} = \{i_k, j_k\}, \\ 1 & \text{в противном случае.} \end{cases}$$

Тогда условие $\mathbf{a}^T\hat{\mathbf{h}} \leqslant \alpha$ может быть переписано в виде

$$\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}h_ih_j \leqslant \alpha.$$

Поскольку наибольшее собственное значение матрицы $A = (a_{ij})$ равно

$$\max\{\mathbf{x}^TA\mathbf{x}: |\mathbf{x}| = 1\},$$

это означает, что наибольшее собственное значение матрицы (a_{ij}) не превосходит α. Так как матрица (a_{ij}) удовлетворяет условию (2), это влечет $\vartheta(G) \leqslant \alpha$. Полученное противоречие доказывает утверждение.

Согласно утверждению, существует конечное число единичных векторов $\mathbf{h}_1, \ldots, \mathbf{h}_N$ и неотрицательных целых чисел $\alpha_1, \ldots, \alpha_N$, таких, что

$$\alpha_1 + \alpha_2 + \ldots \alpha_N = 1, \tag{5}$$

$$\alpha_1\hat{\mathbf{h}}_1 + \ldots + \alpha_N\hat{\mathbf{h}}_N = \mathbf{z}. \tag{6}$$

Положим

$$\mathbf{h}_p = (h_{p,1}, \ldots, h_{p,n})^T,$$

$$b_{ij} = \sum_{p=1}^{N} \alpha_p h_{pi}h_{pj},$$

$$B = (b_{ij}).$$

Очевидно, что матрица B симметрическая и положительно полуопределенная. Кроме того, условие (6) означает, что

$$b_{i_kj_k} = 0, \qquad k = 1, \ldots, m,$$

и

$$\operatorname{Tr} BJ = \vartheta(G),$$

в то время как (5) означает, что

$$\operatorname{Tr} B = 1.$$

Это завершает доказательство.

Лемма 4. *Пусть* $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ — *ортонормированное представление графа* G *и* $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ — *ортонормированное представление дополнительного графа* $\bar{G}$. *Тогда для произвольных векторов* $\mathbf{c}$ *и* $\mathbf{d}$ *имеет место неравенство*

$$\sum_{i=1}^{n} (\mathbf{u}_i^T \mathbf{c})^2 (\mathbf{v}_i^T \mathbf{d})^2 \leqslant \mathbf{c}^2 \mathbf{d}^2.$$

Доказательство. Согласно (1), векторы $\mathbf{u}_i \circ \mathbf{v}_i$ удовлетворяют соотношению

$$(\mathbf{u}_i \circ \mathbf{v}_i)^T (\mathbf{u}_j \circ \mathbf{v}_j) = (\mathbf{u}_i^T \mathbf{u}_j)(\mathbf{v}_i^T \mathbf{v}_j) = \delta_{ij}.$$

Таким образом, они образуют ортонормированную систему векторов, и мы имеем неравенство

$$(\mathbf{c} \circ \mathbf{d})^2 \geqslant \sum_{i=1}^{n} ((\mathbf{c} \circ \mathbf{d})^T (\mathbf{u}_i \circ \mathbf{v}_i))^2,$$

которое есть в точности неравенство леммы.

Следствие 1. *Если* $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ — *ортонормированное представление* $\bar{G}$ *и* $\mathbf{d}$ — *произвольный единичный вектор, то*

$$\vartheta(G) \geqslant \sum_{i=1}^{n} (\mathbf{v}_i^T \mathbf{d})^2.$$

Следствие 2. $\vartheta(G)\vartheta(\bar{G}) \geqslant n$.

Дадим теперь другую минимаксную формулу для величины $\vartheta(G)$, которая показывает весьма неожиданную двойственность между графом G и дополнительным к нему графом $\bar{G}$.

Теорема 5. *Пусть* $(\mathbf{v}_1, \ldots, \mathbf{v}_m)$ *пробегает множество всех ортонормированных представлений графа* $\bar{G}$, *а* $\mathbf{d}$ — *множество всех единичных векторов. Тогда*

$$\vartheta(G) = \max \sum_{i=1}^{n} (\mathbf{d}^T \mathbf{v}_i)^2.$$

Доказательство. Согласно следствию 1, мы уже знаем, что неравенство $\geqslant$ справедливо. Построим теперь представление графа $\bar{G}$ и единичный вектор $\mathbf{d}$, для которых выполняется равенство. Пусть $B = (b_{ij})$ — положительно полуопределенная матрица, удовлетворяющая условиям (3) и (4) и такая, что $\operatorname{Tr} BJ =$

$= \vartheta(G)$. Поскольку матрица B положительно полуопределенная, то существуют векторы $\mathbf{w}_1, \ldots, \mathbf{w}_n$, такие, что

$$b_{ij} = \mathbf{w}_i^T \mathbf{w}_j. \tag{7}$$

Заметим, что

$$\sum_{i=1}^{n} \mathbf{w}_i^2 = 1, \qquad \left(\sum_{i=1}^{n} \mathbf{w}_i\right)^2 = \vartheta(G).$$

Положим

$$\mathbf{v}_i = \mathbf{w}_i / |\mathbf{w}_i|, \qquad \mathbf{d} = \left(\sum_{i=1}^{n} \mathbf{w}_i\right) \Big/ \left|\sum_{i=1}^{n} \mathbf{w}_i\right|.$$

Тогда, согласно условиям (7) и (3), векторы $\mathbf{v}_i$ образуют ортонормированное представление графа $\bar{G}$. Кроме того, используя неравенство Коши — Шварца, получим

$$\sum_{i=1}^{n} (\mathbf{d}^T \mathbf{v}_i)^2 = \left(\sum_{i=1}^{n} \mathbf{w}_i^2\right)\left(\sum_{i=1}^{n} (\mathbf{d}^T \mathbf{v}_i)^2\right) \geqslant \left(\sum_{i=1}^{n} |\mathbf{w}_i| (\mathbf{d}^T \mathbf{v}_i)\right)^2 =$$

$$= \left(\sum_{i=1}^{n} \mathbf{d}^T \mathbf{w}_i\right)^2 = \left(\mathbf{d}^T \sum_{i=1}^{n} \mathbf{w}_i\right)^2 = \left(\sum_{i=1}^{n} \mathbf{w}_i\right)^2 = \vartheta(G).$$

Этим доказательство завершено.

Отметим, что так как в неравенстве Коши — Шварца имеет место равенство, то отсюда также следует

$$(\mathbf{d}\mathbf{v}_i)^2 = \vartheta(G)\, \mathbf{w}_i^2 = \vartheta(G)\, b_{ii}. \tag{8}$$

Теорема 6. *Пусть A пробегает множество всех матриц, для которых $a_{ij} = 0$, если i и j смежны в графе G, и пусть $\lambda_1(A) \geqslant \geqslant \ldots \geqslant \lambda_n(A)$ обозначают собственные значения матрицы A. Тогда*

$$\vartheta(G) = \max_A \left\{1 - \frac{\lambda_1(A)}{\lambda_n(A)}\right\}.$$

Доказательство.

1) Пусть A произвольная матрица, такая, что $a_{ij} = 0$, если вершины i и j смежны в графе G. И пусть $\mathbf{f} = (f_1, \ldots, f_n)^T$ собственный вектор, отвечающий значению $\lambda_1(A)$ и такой, что $\mathbf{f}^2 = -1/\lambda_n(A)$ (заметим, что так как $\operatorname{Tr} A = 0$, то наименьшее собственное значение матрицы A *отрицательно*). Рассмотрим диагональную матрицу $F = \operatorname{diag}(f_1, \ldots, f_n)$ и матрицу

$$B = F(A - \lambda_n(A) I) F.$$

Очевидно, что матрица B положительно полуопределенная. Кроме того, $b_{ij} = 0$, если i и j — различные смежные вершины в графе G, и имеет место

$$\operatorname{Tr} B = -\lambda_n(A) \operatorname{Tr} F^2 = 1.$$

Таким образом, по теореме 4

$$\vartheta(G) \geqslant \operatorname{Tr} BJ = \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} f_i f_j - \lambda_n(A)\sum_{i=1}^{n} f_i^2 =$$

$$= \sum_{i=1}^{n} \{\lambda_1(A) f_i^2 - \lambda_n(A) f_i^2\} = 1 - \frac{\lambda_1(A)}{\lambda_n(A)}.$$

2) Факт достижения равенства может быть получен более или менее прямым обращением рассмотренных выше рассуждений, что здесь опущено.

Следствие 3 (см. Хоффман [3]). *Пусть* $\lambda_1 \geqslant \ldots \geqslant \lambda_n$ *— собственные значения матрицы смежности графа* G. *Тогда хроматическое число графа* G *не меньше чем* $1 - \lambda_1/\lambda_n$.

Доказательство. Покажем, что хроматическое число графа G не меньше чем $\vartheta(\bar{G})$. Действительно, если $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ — ортонормированное представление графа G, $\mathbf{c}$ — произвольный единичный вектор и $J_1, \ldots, J_k$ — одноцветные классы в произвольном k-раскрашивании графа G, то имеет место неравенство

$$\sum_{i=1}^{n} (\mathbf{c}^T \mathbf{u}_i)^2 = \sum_{m=1}^{k} \sum_{i \in J_m} (\mathbf{c}^T \mathbf{u}_i)^2 \leqslant \sum_{m=1}^{k} 1 = k,$$

из которого, согласно теореме 5, и следует утверждение. Теперь матрица смежности графа G удовлетворяет условиям теоремы (с заменой G на $\bar{G}$), из которой и вытекает искомое неравенство.

4. НЕКОТОРЫЕ ДАЛЬНЕЙШИЕ СВОЙСТВА $\vartheta(G)$

Результаты предыдущего раздела позволяют довольно просто вычислять $\vartheta(G)$. Теперь получим некоторые следствия.

Теорема 7. $\vartheta(G \cdot H) = \vartheta(G)\vartheta(H)$.

Доказательство. Мы уже знаем, что

$$\vartheta(G \cdot H) \leqslant \vartheta(G)\vartheta(H).$$

Для того чтобы доказать противоположное неравенство, рассмотрим ортонормированное представление $(\mathbf{v}_1, \ldots, \mathbf{v}_n)$ графа $\bar{G}$, ортонормированное представление $(\mathbf{w}_1, \ldots, \mathbf{w}_m)$ графа $\bar{H}$ и единые векторы $\mathbf{c}$, $\mathbf{d}$, такие, что

$$\sum_{i=1}^{n} (\mathbf{v}_i^T \mathbf{c})^2 = \vartheta(G), \qquad \sum_{i=1}^{m} (\mathbf{w}_i^T \mathbf{d})^2 = \vartheta(H).$$

Тогда векторы $\mathbf{v}_i \circ \mathbf{w}_j$ образуют ортонормированное представление графа $\overline{G \circ H}$ (это следует из включения $\overline{G \cdot H} \supseteq \bar{G} \cdot \bar{H}$ и того,

что эти векторы являются ортонормированным представлением графа $\bar{G}\cdot\bar{H}$). Кроме того, $\mathbf{c}\circ\mathbf{d}$ — единичный вектор. Таким образом,

$$\vartheta(G\cdot H)\geqslant\sum_{i=1}^{n}\sum_{j=1}^{m}((\mathbf{v}_i\circ\mathbf{w}_j)^T(\mathbf{c}\circ\mathbf{d}))^2=\sum_{i=1}^{n}\sum_{j=1}^{m}(\mathbf{v}_i^T\mathbf{c})^2(\mathbf{w}_j\mathbf{d})^2=$$

$$=\sum_{i=1}^{n}(\mathbf{v}_i^T\mathbf{c})^2\sum_{j=1}^{m}(\mathbf{w}_j^T\mathbf{d})^2=\vartheta(G)\vartheta(H).$$

Теорема 8. *Если группа автоморфизмов графа G вершинно-транзитивная, то $\vartheta(G)\vartheta(\bar{G})=n$.*

Следствие 4. *Если группа автоморфизмов графа G вершинно-транзитивная, то $\Theta(G)\Theta(\bar{G})\leqslant n$.*

Отметим, что теорема 8 и ее следствие справедливы не для всех графов, потому что существуют графы, для которых $\alpha(G)\alpha(\bar{G})>n$ (например, звезда).

Доказательство. Пусть Γ — группа автоморфизмов графа G. Элементы Γ можно рассматривать как $n\times n$-матрицы подстановок. Пусть матрица $B=(b_{ij})$ удовлетворяет условиям (3) и (4) и $\operatorname{Tr}BJ=\vartheta(\bar{G})$. Рассмотрим матрицу

$$\bar{B}=(\bar{b}_{ij})=\frac{1}{|\Gamma|}\left(\sum_{P\in\Gamma}P^{-1}BP\right).$$

Очевидно, что матрица $\bar{B}$ также удовлетворяет (3) и

$$\operatorname{Tr}\bar{B}=1,\quad \operatorname{Tr}\bar{B}J=\vartheta(G)$$

(здесь используется, что $PJ=JP=J$). Также очевидно, что матрица $\bar{B}$ симметрическая, положительно полуопределенная и удовлетворяет соотношению $P^{-1}\bar{B}P=\bar{B}$ для всех $P\in\Gamma$. Так как группа Γ транзитивно действует на вершинах, из этого следует, что $\bar{b}_{ii}=1/n$ для всех i. Построив, как при доказательстве теоремы 5, ортонормальное представление $(\mathbf{v}_1,\ldots,\mathbf{v}_n)$ и единичный вектор $\mathbf{d}$, получим, согласно (8),

$$(\mathbf{d}^T\mathbf{v}_i)^2=\frac{\vartheta(G)}{n}.$$

Таким образом, из определения $\vartheta(\bar{G})$ имеем

$$\vartheta(\bar{G})\leqslant\max_{1\leqslant i\leqslant n}\frac{1}{(\mathbf{d}^T\mathbf{v}_i)^2}=\frac{n}{\vartheta(G)}$$

и, следовательно,

$$\vartheta(G)\vartheta(\bar{G})\leqslant n.$$

Поскольку, как мы уже знаем, справедливо и обратное неравенство (следствие 2), то теорема 8 доказана.

Теорема 9. *Пусть G — регулярный граф, а $\lambda_1 \geqslant \lambda_2 \geqslant \ldots \geqslant \lambda_n$ — собственные значения его матрицы смежности. Тогда справедливо неравенство*

$$\vartheta(G) \leqslant \frac{-n\lambda_n}{\lambda_1 - \lambda_n}.$$

Если группа автоморфизмов графа $\bar{G}$ реберно-транзитивная, то имеет место равенство.

Следствие 5. *Для нечетных n*

$$\vartheta(C_n) = \frac{n\cos(\pi/n)}{1+\cos(\pi/n)}.$$

Доказательство. Рассмотрим матрицу $J - xA$, где значение x будет выбрано позже. Она удовлетворяет условию (2) теоремы 3, и, следовательно, ее наибольшее собственное значение не меньше $\vartheta(G)$. Обозначим через $\mathbf{v}_i$ собственный вектор матрицы A, соответствующий собственному значению λ_i. Тогда поскольку граф G регулярен, то $\mathbf{v}_1 = \mathbf{j}$, и поэтому векторы $\mathbf{j}$, $\mathbf{v}_2, \ldots, \mathbf{v}_n$ являются также собственными векторами матрицы J. Таким образом, собственные значения матрицы $J - xA$ равны $n - x\lambda_1, -x\lambda_2, \ldots, -x\lambda_n$. Наибольшее из них либо первое, либо последнее, и оптимальный выбор есть $x = n/(\lambda_1 - \lambda_n)$, когда они оба равны $-n\lambda_n/(\lambda_1 - \lambda_n)$. Это доказывает первое предложение.

Предположим теперь, что группа Γ автоморфизмов графа G транзитивно действует на ребрах. Пусть $C = (c_{ij})$ — произвольная симметрическая матрица, такая, что $c_{ij} = 1$, если i и j равны или несмежны, и ее максимальное собственное значение равно $\vartheta(G)$. Как и в доказательстве теоремы 8, рассмотрим матрицу

$$\bar{C} = \frac{1}{|\Gamma|}\sum_{P \in \Gamma} P^{-1}CP.$$

Тогда матрица $\bar{C}$ также удовлетворяет условиям (2) и, кроме того, ее максимальное собственное значение не больше $\vartheta(G)$. По теореме 3 оно равно $\vartheta(G)$. Кроме того, матрица $\bar{C}$, очевидно, имеет вид $J - xA$. Отсюда вытекает справедливость второго предложения.

5. СРАВНЕНИЕ С ДРУГИМИ ГРАНИЦАМИ ДЛЯ ЕМКОСТИ

Теорема 10. $\vartheta(G) \leqslant \alpha^*(G)$.

Доказательство. Мы воспользуемся теоремой 4. Пусть $(\mathbf{u}_i)$ — ортонормированное представление графа $\bar{G}$ и $\mathbf{c}$ — единичный

вектор, такие, что

$$\vartheta(G) = \sum_{i=1}^{n} (\mathbf{c}^T \mathbf{u}_i)^2.$$

Пусть C — произвольная клика в графе G. Тогда множество $\{\mathbf{u}_i : i \in C\}$ есть ортонормированная система векторов и, следовательно,

$$\sum_{i \in C} (\mathbf{c}^T \mathbf{u}_i)^2 \leqslant \mathbf{c}^2 = 1.$$

Отсюда веса $(\mathbf{c}^T \mathbf{u}_i)^2$ образуют дробную упаковку вершин, и, таким образом,

$$\vartheta(G) = \sum_{i=1}^{n} (\mathbf{c}^T \mathbf{u}_i)^2 \leqslant \alpha^*(G).$$

Очень простой верхней границей величины $\Theta(G)$ является размерность ортонормированного представления графа G.

Теорема 11. *Предположим, что у графа G имеется ортонормированное представление размерности d. Тогда*

$$\vartheta(G) \leqslant d.$$

Доказательство. Пусть $(\mathbf{u}_1, \ldots, \mathbf{u}_n)$ — ортонормированное представление графа G в d-мерном пространстве. Тогда $(\mathbf{u}_1 \circ \mathbf{u}_1, \mathbf{u}_2 \circ \mathbf{u}_2, \ldots, \mathbf{u}_n \circ \mathbf{u}_n)$ есть другое ортонормированное представление графа G. Пусть $\{\mathbf{e}_1, \ldots, \mathbf{e}_d\}$ — ортонормированный базис и вектор

$$\mathbf{b} = \frac{1}{\sqrt{d}} (\mathbf{e}_1 \circ \mathbf{e}_1 + \mathbf{e}_2 \circ \mathbf{e}_2 + \ldots + \mathbf{e}_d \circ \mathbf{e}_d).$$

Тогда $\mathbf{b}^2 = 1$ и

$$(\mathbf{u}_i \circ \mathbf{u}_i)^T \mathbf{b} = \frac{1}{\sqrt{d}} \sum_{k=1}^{d} (\mathbf{e}_k \circ \mathbf{e}_k)^T (\mathbf{u}_i \circ \mathbf{u}_i) =$$

$$= \frac{1}{\sqrt{d}} \sum_{k=1}^{d} (\mathbf{e}_k^T \mathbf{u}_i)^2 = \frac{1}{\sqrt{d}}.$$

Поэтому $\vartheta(G) \leqslant d$.

6. ПРИЛОЖЕНИЯ

Мы можем применить наши методы для вычисления шенноновской емкости графов, отличных от пятиугольника. Мы, конечно, будем иметь дело с такими графами G, для которых $\alpha(G) < \alpha^*(G)$, поскольку, если $\alpha(G) = \alpha^*(G)$, то $\Theta(G) = \alpha(G)$ согласно теореме Шеннона.

Теорема 12. *Если группа автоморфизмов графа G вершинно-транзитивная, то* $\Theta(G\cdot\bar{G})=|V(G)|$. *Если, кроме того, граф G самодополнительный, то* $\Theta(G)=\sqrt{|V(G)|}$.

Доказательство. «Диагональ» в графе $G\cdot\bar{G}$ является независимым множеством, следовательно,

$$\Theta(G\cdot\bar{G})\geqslant\alpha(G\cdot\bar{G})\geqslant|V(G)|.$$

С другой стороны, согласно теоремам 1, 6 и 7, имеем

$$\Theta(G\cdot\bar{G})\leqslant\vartheta(G\cdot\bar{G})=\vartheta(G)\vartheta(\bar{G})=|V(G)|.$$

Если граф самодополнительный, то

$$\Theta(G\cdot\bar{G})=\Theta(G^2)=\Theta(G)^2.$$

Это доказывает теорему. Из доказательства также видно, что в этих случаях $\Theta=\vartheta$.

Теорема 13. *Пусть $n\geqslant 2r$ и $K(n,r)$ обозначает граф, множеством вершин которого являются r-элементные подмножества n-элементного множества S, и два подмножества смежны, если и только если они не пересекаются. Тогда*

$$\Theta(K(n,r))=\binom{n-1}{r-1}.$$

Следствие 6. *Емкость графа Петерсена, изоморфного графу $K(5,2)$, равна* 4.

Следствие 7 (см. Эрдеш, Ко и Радо [2]).

$$\alpha(K(n,r))=\binom{n-1}{r-1}.$$

Заметим, что

$$\alpha^*(K(n,r))=\binom{n}{r}\Big/\left[\frac{n}{r}\right]$$

и это больше, чем $\binom{n-1}{r-1}$, *если r не делит n.*

Доказательство теоремы 13. Подмножества из r элементов, содержащие фиксированный элемент из S, образуют независимое множество в графе $K(n,r)$ и, следовательно,

$$\Theta(K(n,r))\geqslant\alpha(K(n,r))\geqslant\binom{n-1}{r-1}.$$

Чтобы доказать противоположное неравенство, мы вычислим $\vartheta(K(n,r))$. Так как группа автоморфизмов графа $K(n,r)$, очевидно, транзитивна и на вершинах и на ребрах, то можно

воспользоваться теоремой 9. Таким образом, следует вычислить собственные числа, соответствующие графу $K(n, r)$. Очевидно, что вектор j является собственным с собственным значением $\binom{n-r}{r}$.

Пусть $1 \leqslant t \leqslant r$ и для каждого подмножества $T \subset S$, такого, что $|T| = t$, пусть x_T обозначает действительное число, удовлетворяющее следующему соотношению: для любого подмножества $U \subset S$ мощности $|U| = t-1$ выполнено

$$\sum_{T \supseteq U} x_T = 0. \tag{9}$$

Имеется $\binom{n}{t} - \binom{n}{t-1}$ линейно независимых векторов (x_T) такого типа. Для каждого такого вектора определим число

$$\bar{x}_A = \sum_{\substack{T \subseteq A \\ |T|=t}} x_T$$

для любого подмножества $A \subset S$, $|A| = r$. Нетрудно видеть (и это хорошо известно), что числа x_T могут быть вычислены через числа $\bar{x}_A$, следовательно, имеется $\binom{n}{t} - \binom{n}{t-1}$ линейно независимых векторов вида $(\bar{x}_A)$.

Утверждение. Каждый вектор $(\bar{x}_A)$ является собственным для матрицы смежности графа $K(n, r)$ с собственным значением $(-1)^t \binom{n-r-t}{r-t}$. В самом деле, для любого $A_0 \subset S$, такого, что $|A_0| = r$, имеем

$$\sum_{A \cap A_0 = \varnothing} \bar{x}_A = \sum_{T \cap A_0 = \varnothing} \binom{n-r-t}{r-t} x_T = \binom{n-r-t}{r-t} \beta_0.$$

Чтобы определить эту величину, положим

$$\beta_i = \sum_{|T \cap A_0| = i} x_T.$$

Тогда, суммируя (9) по всем $U \subset S$, таким, что $|U| = t-1$ и $|U \cap A_0| = i$, получим

$$(i+1)\beta_{i+1} + (t-i)\beta_i = 0.$$

Это можно рассматривать как рекуррентное соотношение для величин β_i, что дает

$$\beta_i = (-1)^i \binom{t}{i} \beta_0,$$

откуда

$$\beta_0 = (-1)^t \beta_t = (-1)^t \bar{x}_{A_0},$$

что и доказывает утверждение.

С помощью этой конструкции мы нашли

$$1 + \sum_{t=1}^{r} \left(\binom{n}{t} - \binom{n}{t-1} \right) = \binom{n}{r}$$

линейно независимых собственных векторов (с векторами, соответствующими разным значениям t, нет сложностей, так как их собственные значения различны). Поэтому мы имеем все собственные векторы, и, следовательно, собственными значениями матрицы смежности графа $K(n, r)$ являются числа

$$(-1)^t \binom{n-r-t}{r-t}, \qquad t = 0, 1, \ldots, r.$$

Итак, наибольшее и наименьшее собственные значения равны $\binom{n-r}{r}$ и $-\binom{n-r-1}{r-1}$ соответственно, и теорема 9 дает соотношение

$$\vartheta(K(n, r)) = \frac{\binom{n-r-1}{r-1}\binom{n}{r}}{\binom{n-r}{r} + \binom{n-r-1}{r-1}} = \binom{n-1}{r-1}.$$

ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ

Величина $\vartheta(G)$ была введена с целью оценить $\Theta(G)$. Поэтому естественными являются следующие вопросы.

Задача 1. Справедливо ли равенство[1]) $\vartheta = \Theta$? Более скромно — найти другие графы, для которых $\vartheta(G) = \Theta(G)$. В частности, верно ли для циклического графа с нечетным числом вершин равенство $\vartheta(G) = \Theta(G)$?

Последний вопрос указывает на трудности, которые кажутся решающими. Во всех случаях точного определения величины $\Theta(G)$, которые известны автору, существует некоторое k (на самом деле $k = 1$ или 2), такое, что $\alpha(G^k) = \Theta(G)^k$. Но если $\Theta(G) = \vartheta(G)$, например, для циклического графа с 7 вершинами, то такого k не существует, поскольку никакая степень $\vartheta(C_7)$ не является целым числом.

[1]) Ответ на этот вопрос, а также вопросы задач 2 и 3 отрицательный, см. W. Haemers. On some problems of Lovasz concerning the Shannon capacity of a graph, IEEE Trans. Jnform. Theory, IT-25, 2, 1979, 231—232. — *Прим. перев.*

Различные свойства величины $\vartheta(G)$, установленные в этой статье, приводят к новым задачам, которые могут быть решены при положительном ответе на первый вопрос задачи 1.

Задача 2. Справедливо ли равенство $\Theta(G\cdot H)=\Theta(G)\Theta(H)$? (Заметим, что неравенство $\Theta(G\cdot H)\geqslant\Theta(G)\Theta(H)$ очевидно.)

Задача 3. Верно ли, что $\Theta(G)\cdot\Theta(\bar{G})\geqslant|V(G)|$?

Отметим, что положительный ответ на вопрос задачи 2 означает и положительный ответ на вопрос задачи 3, поскольку

$$\Theta(G)\Theta(\bar{G})=\Theta(G\cdot\bar{G})\geqslant\alpha(G\cdot\bar{G})\geqslant|V(G)|.$$

Это в свою очередь повлекло бы положительный ответ на последний вопрос задачи 1:

$$n\leqslant\Theta(C_n)\Theta(\bar{C}_n)\leqslant\vartheta(C_n)\vartheta(\bar{C}_n)=n,$$

откуда $\Theta(C_n)=\vartheta(C_n)$ и $\Theta(\bar{C}_n)=\vartheta(\bar{C}_n)$.

Следствие 7 показывает пример, когда вычисление $\vartheta(G)$ помогает нетривиальным образом вычислить $\alpha(G)$. Есть ли еще такие примеры?

Благодарности

Я искренне благодарен К. Вестергомби и М. Розенфельду за неоднократные обсуждения предмета этой статьи. Я также благодарен Т. Немецу, А. Схрейверу и рецензентам, которые обратили мое внимание на некоторые ошибки и предложили ряд улучшений первоначального текста. Я хотел бы особо отметить, что А. Дж. Хоффман и М. Розенфельд также распространили основную идею (разд. II) на другие графы. В частности, Хоффман получил теорему 9, а Розенфельд вычислил $\vartheta(\bar{C}_n)$.

ЛИТЕРАТУРА

1. Berge C., Graphs and Hypergraphs. Amsterdam and London: North-Holland; New York: American Elsevier, 1973.
2. Erdös P., Ko C. and Rado R. Intersection theorems for systems of finite sets, Quart. J. Math. Oxford, vol. 12, pp. 313—320, 1961.
3. Hoffman A. J. On eigenvalues and colorings of graphs, in B. Harris, ed., Graph Theory and its Applications. New York and London: Academic, 1970, pp. 79—91.
4. Lancaster P. Theory of Matrices. New Yoork and London: Academic, 1969. [Имеется перевод: П. Ланкастер. Теория матриц. — М.: Наука, 1978.]
5. Rosenfeld M. On a problem of Shannon, Proc. Amer Math. Soc., vol. 18, pp. 315—319, 1967.
6. Shannon C. (Sept. 1956), The zero-error capacity of a noisy channel, IRE Trans. Inform. Theory IT-2, 8—19. [Имеется перевод: Шеннон К. Пропускная способность канала с шумом при нулевой ошибке. — В кн.: Работы по теории информации и кибернетике. — М.: ИЛ, 1963, стр. 464—487.]

Сравнение границ Дельсарта и Ловаса[1])

Александр Схрейвер

Граница линейного программирования Дельсарта (верхняя граница мощности клик в ассоциативных схемах) сравнивается с границей Ловаса (верхней границей $\theta(G)$ шенноновской емкости графа G). Обе границы могут трактоваться единым образом. Так, граница линейного программирования Дельсарта может быть обобщена до границы $\theta'(G)$ для числа независимости $\alpha(G)$ произвольного графа G, такой, что $\theta'(G) \leqslant \theta(G)$. С другой стороны, если множество ребер графа G является объединением классов симметрической ассоциативной схемы, то величина $\theta(G)$ может быть вычислена с помощью методов линейного программирования. Для таких графов произведение $\theta(G)\cdot\theta(\bar{G})$ равно числу вершин графа G.

1. ВВЕДЕНИЕ

Цель этой заметки — сравнить две функции, являющиеся верхними границами величин, изучение которых в определенной степени мотивировано теоретико-информационными задачами: границу линейного программирования Дельсарта для мощности клик в ассоциативных схемах и границ Ловаса $\theta(G)$ шенноновской емкости графа G. Первая граница может рассматриваться как граница числа независимости $\alpha(G)$ для некоторых графов G, тогда как граница Ловаса ограничивает сверху $\alpha(G^k)$ — число независимости графа, являющегося сильным произведением k копий графа G.

Вначале мы кратко изложим эти границы и историю вопроса (*граф* — это неориентированный граф без петель и кратных ребер).

Ассоциативные схемы и граница линейного программирования Дельсарта [2], [5]

Пара $(X, \mathcal{R})$, где $\mathcal{R}=(R_0, \ldots, R_n)$ есть разбиение $X\times X$, называется (*симметрической*) *ассоциативной схемой с числами*

---

[1]) Schrijver A. A Comparison of the Delsarte and Lovász Bounds. IEEE Trans. on Inform. Theory, IT-25, No. 4, July 1979.

© 1979 IEEE. Reprinted, with permission, from IEEE Transactions on Information Theory, vol. IT-25, pp. 425—429, July 1979
© Перевод на русский язык, «Мир», 1983

пересечений p_{ij}^k $(i, j, k = 0, \ldots, n)$, если

(1) $$R_0 = \{(x, x) \mid x \in X\};$$

(2) $$R_k^{-1} = \{(y, x) \mid (x, y) \in R_k\} = R_k \quad \text{для } k = 0, \ldots, n;$$

и для всех $i, j, k = 0, \ldots, n$ и $(x, y) \in R_k$

(3) $$|\{z \mid (x, z) \in R_i \text{ и } (z, y) \in R_j\}| = p_{ij}^k.$$

Таким образом, $p_{ij}^k = p_{ji}^k$. Мы можем рассматривать пару (X, R_i) как граф $(i = 1, \ldots, n)$. Граф (X, R_i) является регулярным степени $v_i = p_{ii}^0$ $(v_0 = 1)$. Следовательно, $p_{ij}^0 = \delta_{ij} v_i$. Пусть D_i — матрица смежности графа (X, R_i); D_0 — единичная матрица. Поскольку в силу (3) симметрические матрицы $D_0, \ldots, D_n$ перестановочны, то существует матрица $P = (P_k^u)_{k,\, u=0}^n$, такая, что числа $P_k^0, \ldots, P_k^n$ есть собственные значения матрицы D_k $(k = 0, \ldots, n)$ и для матриц $D_0, \ldots, D_n$ существует общий собственный вектор с собственными значениями $P_0^u, \ldots, P_n^u$ соответственно $(u = 0, \ldots, n)$. Можно считать, что $P_k^0 = v_k$ для всех k. Положим

(4) $$Q_k^u = \frac{\mu_u}{v_k} P_k^u,$$

где μ_u — размерность общего для матриц $D_0, \ldots, D_n$ собственного подпространства с собственными значениями $P_0^u, \ldots, P_n^u$ соответственно $(u = 0, \ldots, n)$. Можно показать, что

(5) $$\sum_{u=0}^{n} P_k^u Q_l^u = m \cdot \delta_{kl} \quad \text{и} \quad \sum_{k=0}^{n} P_k^u Q_k^v = m \cdot \delta_{uv},$$

где $m = |X|$. Таким образом, матрицы P и $m^{-1} \cdot Q^T$ являются взаимно обратными.

В теории кодирования интересуются двумя семействами ассоциативных схем: так называемыми схемами Хэмминга и Джонсона. Пусть n и q натуральные числа, и пусть X есть множество n-мерных векторов с координатами из множества $\{0, \ldots, q-1\}$. При $k = 0, \ldots, n$ положим

(6) $$R_k = \{(x, y) \in X \times X \mid d_H(x, y) = k\},$$

где $d_H(x, y)$ обозначает *расстояние Хэмминга* между векторами x и y, т. е. число координат, в которых x и y различаются. Пусть $\mathcal{R} = (R_0, \ldots, R_n)$. Как легко может быть проверено, пара $(X, \mathcal{R})$ является симметрической ассоциативной схемой; схемы, получаемые таким образом, называются *схемами Хэмминга*. Для схем Хэмминга величины v_k, μ_u и P_k^u задаются следую-

щими формулами:

$$v_k = \binom{n}{k}\cdot(q-1)^k \qquad \mu_u = \binom{n}{u}\cdot(q-1)^u$$

$$P_k^u = K_k(u) = \sum_{j=0}^{k}(-1)^j(q-1)^{k-j}\binom{u}{j}\binom{n-u}{k-j} \tag{7}$$

$$= \sum_{j=0}^{k}(-q)^j(q-1)^{k-j}\binom{n-j}{k-j}\binom{u}{j},$$

при $k, u = 0, \ldots, n$ ($K_k(u)$ — это *многочлен Кравчука* степени k от переменной u).

Второе семейство получается следующим образом. Пусть v и n — натуральные числа, и пусть X есть множество v-мерных $(0, 1)$-векторов, содержащих ровно n единиц ($n \leqslant 1/2\ v$). При $k = 0, \ldots, n$ положим

$$R_k = \{(x, y) \in X \times X \mid d_J(x, y) = k\}, \tag{8}$$

где $d_J(x, y) = 1/2 d_H(x, y)$ — *расстояние Джонсона* между векторами x и y. Пусть $\mathscr{R} = (R_0, \ldots, R_n)$. Тогда пара $(X, \mathscr{R})$ является симметрической ассоциативной схемой; построенные таким образом схемы называются *схемами Джонсона*. Их параметры таковы

$$v_k = \binom{n}{k}\binom{v-n}{n-k},$$

$$\mu_u = \binom{v}{u} - \binom{v}{u-1} = \frac{v-2u+1}{v-u+1}\cdot\binom{v}{u}, \tag{9}$$

$$P_k^u = E_k(u) = \sum_{j=0}^{k}(-1)^{k-j}\binom{n-j}{k-j}\binom{n-u}{j}\binom{v-n+j-u}{j}$$

$$= \sum_{j=0}^{k}(-1)^j\binom{u}{j}\binom{n-u}{k-j}\binom{v-n-u}{k-j},$$

для $k, u = 0, \ldots, n$ ($E_k(u)$ — это *многочлен Эберлейна* степени $2k$ от переменной u).

(*Сильно регулярные графы* дают третье семейство симметрических ассоциативных схем. Это в точности те графы (X, R_1), для которых пара $(X, \mathscr{R})$ является симметрической ассоциативной схемой, где $\mathscr{R} = (R_0, R_1, R_2)$, $R_2 = (X \times X) \setminus (R_0 \cup R_1)$. Отсюда следует, что дополнительный граф к сильно регулярному графу также является сильно регулярным.)

Основная задача комбинаторной теории кодирования состоит в том, чтобы оценить максимальный объем любого подмножества

(кода) в схеме Хэмминга или Джонсона, обладающего тем свойством, что в нем нет двух элементов, расстояние (Хэмминга или Джонсона) между которыми меньше заданной величины d. Чтобы перевести эту задачу на язык ассоциативных схем, нам необходимо понятие M-клики: для заданного множества M: $0 \in M \subset \{0, \ldots, n\}$ подмножество Y множества X является *M-кликой*, если $(x, y) \in \bigcup_{k \in M} R_k$ для всех $x, y \in Y$. Таким образом, задача теории кодирования состоит в определении максимальной мощности $\{0, d, d+1, \ldots, n\}$ — клики в схемах Хэмминга и Джонсона.

Чтобы получить верхнюю границу объема клики в симметрической ассоциативной схеме $(X, \mathcal{R})$, определим для $Y \subset X$ внутреннее распределение $(a_0, \ldots, a_n)$ множества Y следующим образом:

$$a_k = \frac{|R_k \cap (Y \times Y)|}{|Y|} \tag{10}$$

при $k = 0, \ldots, n$. Таким образом, $a_0 = 1$ и $\sum_{k=0}^{n} a_k = |Y|$. Кроме того, если Y есть M-клика, то $a_k = 0$ при $k \notin M$. Дельсарт показал, что для внутреннего распределения произвольного подмножества Y из X имеют место неравенства

$$\sum_{k=0}^{n} a_k Q_k^u \geqslant 0 \tag{11}$$

при $u = 0, \ldots, n$. Поэтому для любой M-клики Y имеет место неравенство

$$\begin{aligned} |Y| \leqslant \max\Bigg\{ \sum_{k=0}^{n} a_k \,|\, a_0, \ldots, a_n \geqslant 0;\ a_0 = 1;\ a_k = 0 \text{ для } k \notin M; \\ \sum_{k=0}^{n} a_k Q_k^u \geqslant 0 \Bigg\} \\ = \min\Bigg\{ \sum_{u=0}^{n} b_u \,|\, b_0, \ldots, b_n \geqslant 0;\ b_0 = 1;\ \sum_{u=0}^{n} b_u P_k^u \leqslant 0 \\ \text{для } k \notin M \setminus \{0\} \Bigg\}. \end{aligned} \tag{12}$$

Равенство в (12) следует из теоремы двойственности линейного программирования. Получаемая граница для объема клики называется *границей линейного программирования Дельсарта*. Методы линейного программирования можно применить для вычисления этой величины (относительно приложений в теории кодирования ср. [1]).

Следующий результат Дельсарта показывает, что граница линейного программирования является уточнением известной в теории кодирования границы Хэмминга. Пусть $(X, \mathcal{R})$ есть сим-

метрическая ассоциативная схема с $\mathscr{R}=(R_0, \ldots, R_n)$, и пусть $0 \in M \subset \{0, \ldots, n\}$ и $\bar{M} = \{0\} \cup (\{0, \ldots, n\} \setminus M)$. Тогда

(13) произведение значения границы линейного программирования для M-клик и значения границы линейного программирования для $\bar{M}$-клик не превосходит $|X|$.

Следовательно, $|Y| \cdot |Z| \leqslant |X|$ для M-клики Y и $\bar{M}$-клики Z. Взяв для схемы Хэмминга $M = \{0, d, d+1, \ldots, n\}$, получим границу Хэмминга.

Шенноновская емкость и граница Ловаса

Ловас [4] ввел для произвольного графа G число $\theta(G)$, которое является верхней границей «шенноновской емкости» $\Theta(G)$. Пусть $\alpha(G)$ есть максимальное число независимых (т. е. попарно несмежных) вершин графа G, и пусть $G \cdot H$ обозначает сильное произведение графов G и H, т. е. множеством вершин графа $G \cdot H$ является декартово произведение множества вершин графов G и H, и две различные вершины графа $G \cdot H$ смежны, если и только если в обеих координатах элементы равны или смежны. Произведение k копий графа G обозначается через G^k.

Шеннон [9] ввел число

$$\Theta(G) = \sup_k \sqrt[k]{\alpha(G^k)} = \lim_{k \to \infty} \sqrt[k]{\alpha(G^k)}, \tag{14}$$

называемое теперь шенноновской емкостью графа G. Если рассматривать вершины графа G как буквы алфавита, так что две вершины смежны, если и только если они «перепутываемы», то величину $\alpha(G^k)$ можно интерпретировать как максимальное число k-буквенных сообщений, таких, что любые два из них неперепутываемы по крайней мере в одной координате.

Поскольку $\alpha(G)^k \leqslant \alpha(G^k)$, отсюда следует, что $\alpha(G) \leqslant \Theta(G)$. В общем случае равенство не имеет места: например, $\alpha(C_5) = 2$, тогда как $\alpha(C_5^2) = 5 \leqslant \Theta(C_5)^2$. Ловас показал, что $\Theta(C_5) = \sqrt{5}$. В действительности Ловас дал следующую общую верхнюю границу для величины $\Theta(G)$.

Для графа $G = (V, E)$ с множеством вершин $V = \{1, \ldots, n\}$ положим

(15) $\theta(G) = \min \{\operatorname{lev} A \mid A = (a_{ij}) —$ симметрическая $n \times n$-матрица, такая, что $a_{ij} = 1$, если $\{i, j\} \notin E\}$,

где $\operatorname{lev} A$ обозначает наибольшее собственное значение матрицы A. Если $\alpha(G) = k$, то каждая матрица A, удовлетворяющая условиям, упомянутым в (15), содержит главный $k \times k$-минор

из одних единиц (с наибольшим собственным значением k) и, следовательно[1]), $\operatorname{lev} A \geqslant k$. Поэтому $\alpha(G) \leqslant \theta(G)$. Ловас доказал, что $\theta(G \cdot H) = \theta(G)\theta(H)$ для всех графов G и H. Отсюда следует $\alpha(G^k) \leqslant \theta(G)^k$, что приводит к более сильному неравенству $\Theta(G) \leqslant \theta(G)$ (Хэмерс [3] показал существование графов G, для которых $\Theta(G) < \theta(G)$). Кроме того, Ловас показал, что

(16) $\theta(G) = \max\{\sum_{ij} b_{ij} \mid B = (b_{ij}) —$ положительная полуопределенная матрица, такая, что $\operatorname{Tr} B = 1$ и $b_{ij} = 0$, если $\{i, j\} \in E\}$.

Итак, величину $\theta(G)$ можно рассматривать и как максимум, и как минимум, что делает функцию θ более простой для вычислений. Ловас установил в числе других результатов, что для графа G с n вершинами

(17) $\theta(G) \cdot \theta(\bar{G}) \geqslant n$ (где $\bar{G}$ — дополнительный граф), причем имеет место равенство, если граф вершинно-транзитивный,

(18) $\theta(G) \leqslant \frac{-n\lambda_n}{\lambda_1 - \lambda_n}$, если граф G регулярен (где λ_1 и λ_n — наибольшее и наименьшее собственные значения матрицы смежности графа G), причем имеет место равенство, если граф реберно-транзитивный.

В качестве следствия (18) получается следующий результат. Пусть $v \geqslant 2n$, и пусть $K(v, n)$ есть граф, множеством вершин которого являются n-подмножества некоторого фиксированного v-множества, и две вершины соединены, если и только если подмножества не пересекаются. Такие графы называются графами Кнезера. Тогда

(19) $$\theta(K(v, n)) = \binom{v-1}{n-1}$$

(согласно (18) достаточно вычислить собственные значения матрицы смежности графа $K(v, n)$), что обобщает теорему Эрдеша — Ко — Радо, согласно которой

$$\alpha(K(v, n)) = \binom{v-1}{n-1}.$$

Теории Дельсарта и Ловаса, как оказалось, имеют некоторые общие черты, такие, как оценивание клик или независимых

[1]) Здесь используется тот факт, что $\operatorname{lev} A = \max\{\mathbf{x}^T A \mathbf{x} \mid \mathbf{x}^T \mathbf{x} = 1\}$. — *Прим. перев.*

множеств в графах, использование техники собственных значений матриц, определяемых графами, установление отношений между графом и его дополнением, и возможность приложения к таким родственным структурам, как равновесные коды и графы Кнезера. Цель этой заметки — продвинуться дальше в изучении этой взаимосвязи.

Очевидно, что границу линейного программирования Дельсарта можно понимать как верхнюю границу величины $\alpha(G)$ для графов G, чье множество вершин является объединением классов R_i симметрической ассоциативной схемы $(X, \mathscr{R})$. Мы покажем, что границу Дельсарта можно расширить до границы $\theta'(G)$ величины $\alpha(G)$ для произвольных графов G; описание границы $\theta'(G)$ имеет много общих характерных черт с границей Ловаса $\theta(G)$. Будет показано, что $\theta'(G) \leqslant \theta(G)$ (в общем случае $\theta'(G) \neq \theta(G)$). С другой стороны, если множество ребер графа G есть объединение классов симметрической ассоциативной схемы $(X, \mathscr{R})$, то величина $\theta(G)$ может быть вычислена как решение задачи линейного программирования, получаемой из задачи (12) отбрасыванием ограничений неотрицательности для $a_0, \ldots, a_n$. Из этого следует, что для таких графов G также имеет место равенство $\theta(G) \cdot \theta(\bar{G}) = |X|$ (ср. (13) и (17)).

2. СРАВНЕНИЕ ГРАНИЦ ДЕЛЬСАРТА И ЛОВАСА

Вначале напомним следующую усиленную форму теоремы двойственности линейного программирования. Пусть C и D — замкнутые выпуклые конусы в пространствах $\boldsymbol{R}^k$ и $\boldsymbol{R}^m$ соответственно, а C^* и D^* — двойственные им конусы (т. е. C^* состоит из всех векторов $\boldsymbol{R}^k$, скалярное произведение которых со всеми векторами из C неотрицательно). Пусть M — действительная $m \times k$-матрица, и пусть $s \in \mathbf{R}^k$. Тогда.

$$(20)\quad \max\{cx \mid x \in C;\ d - Mx \in D\} = \min\{yd \mid y \in D^*;\ yM - c \in C^*\}$$

при условии, что эти множества непусты и замкнуты. Отметим, кроме того, что двойственным конусом к замкнутому выпуклому конусу действительных симметрических положительно полуопределенных $n \times n$-матриц, рассматриваемых как n^2-мерные векторы, является множество действительных $n \times n$-матриц U, таких, что $y^T U y \geqslant 0$ для всех вещественных n-мерных векторов y. (Таким образом, симметрические матрицы в двойственном конусе являются положительно полуопределенными.) Для удобства мы будем использовать следующее обозначение скалярного произведения $n \times n$-матриц $A = (a_{ij})$ и $B = (b_{ij})$:

$$(21)\quad A * B = \sum_{i,j=1}^{n} a_{ij} b_{ij},$$

т. е. $A*B = \operatorname{Tr}(A^TB)$. Таким образом, $A*I = \operatorname{Tr} A$ и $A*J = \sum_{i,j=1}^{n} a_{ij}$.

Пусть G — граф с множеством вершин $\{1, \ldots, n\}$. Ловас определил функцию

$$\theta(G) = \max\{\textstyle\sum_{i,j} b_{ij} \mid B = (b_{ij}) \text{ — симметрическая положительно полуопределенная } n \times n\text{-матрица, такая, что } \operatorname{Tr} B = 1 \text{ и } b_{ij} = 0, \text{ если } \{i, j\} \in E\} = \tag{22}$$
$$= \min\{\operatorname{lev} A \mid A = (a_{ij}) \text{ — симметрическая } n \times n\text{-матрица, такая, что } a_{ij} = 1, \text{ если } \{i, j\} \notin E\}.$$

Теперь определим функцию $\theta'(G)$ следующим образом:

$$\theta'(G) = \max\{\textstyle\sum_{i,j} b_{ij} \mid B = (b_{ij}) \text{ — неотрицательная симметрическая положительно полуопределенная } n \times n\text{-матрица, такая, что } \operatorname{Tr} B = 1 \text{ и } b_{ij} = 0, \text{ если } \{i, j\} \in E\}. \tag{23}$$

Это определение отличается от (22) тем, что рассматриваются только неотрицательные матрицы B.

Теорема 1. $\alpha(G) \leqslant \theta'(G) \leqslant \theta(G)$.

Доказательство. Очевидно, что $\theta'(G) \leqslant \theta(G)$. Пусть $Y (Y \subset \subset \{1, \ldots, n\})$ является независимым множеством из $\alpha(G) = k$ элементов. Положим $b_{ij} = 1/k$, если $i, j \in Y$, и $b_{ij} = 0$ в противном случае. Тогда матрица $B = (b_{ij})$ неотрицательная и положительно полуопределенная со следом 1, причем $b_{ij} = 0$, если $\{i, j\} \in E$. Кроме того, $\sum_{i,j} b_{ij} = k$. Следовательно, $\alpha(G) \leqslant k \leqslant \theta'(G)$. ■

Теорема 2. $\theta'(G) = \min\{\operatorname{lev} A \mid A = (a_{ij}) —$ *симметрическая* $n \times n$*-матрица, такая, что* $a_{ij} \geqslant 1$, *если* $\{i, j\} \notin E\}$.

Доказательство. Согласно определению,

$$\theta'(G) = \max\{B * J \mid B = (b_{ij}) \text{ — симметрическая положительно полуопределенная } n \times n\text{-матрица, такая, что } B * I = 1,\ B * F_{ij} = 0 \text{ для } \{i, j\} \in E, \text{ и } B * F_{ij} \geqslant \tag{24}$$
$$\geqslant 0 \text{ для } \{i, j\} \notin E\},$$

где F_{ij} это $(0, 1)$-матрица порядка n, в которой 1 стоит только на (i, j) и (j, i) местах. Из упомянутой выше формы теоремы двойственности следует, что этот максимум равен

$$\min\{\lambda \in R \mid M = (m_{ij}) \text{ — симметрическая } n \times n\text{-матрица;}$$
$$m_{ij} \leqslant 0, \text{ если } \{i, j\} \notin F; \text{ матрица } \lambda I + M - F \text{ является} \tag{25}$$
положительно полуопределенной}.

Полагая $A = J - M$ и используя тот факт, что для симметрической матрицы A ее наибольшее собственное значение равно минимальному значению λ, при котором матрица $\lambda I - A$ является положительно полуопределенной, получаем, что

$$\theta'(G) = \min\{\operatorname{lev} A \mid A = (a_{ij}) - \text{симметрическая } n \times n\text{-матрица, такая, что } a_{ij} \geqslant 1, \text{ если } \{i, j\} \notin E\}. \tag{26}$$ ■

Поскольку наибольшее собственное значение матрицы не возрастает при уменьшении диагональных элементов, можно считать, что минимум достигается на некоторой матрице A с единицами на диагонали.

Теперь мы докажем, что для графов, получаемых из симметрических ассоциативных схем, величина $\theta'(G)$ совпадает с границей линейного программирования Дельсарта.

Пусть $(X, \mathcal{R})$ — симметрическая ассоциативная схема с $\mathcal{R} = (R_0, \ldots, R_n)$, и пусть $0 \in M \subset \{0, \ldots, n\}$. Рассмотрим граф $G = (X, E)$ с множеством ребер $E = \bigcup_{i \notin M} R_i$. Очевидно, что понятия M-клики в ассоциативной схеме и независимого множества в графе G совпадают.

Теорема 3. *Величина $\theta'(G)$ равна значению границы линейного программирования для M-клик в схеме $(X, \mathcal{R})$.*

Доказательство. Согласно определению, значение границы линейного программирования равно (ср. (12))

$$\max\left\{\sum_{k=0}^{n} a_k \mid a_0, \ldots, a_n \geqslant 0;\ a_0 = 1;\ a_k = 0 \text{ для } k \notin M;\ \sum_{k=0}^{n} a_k Q_k^u \geqslant 0 \text{ для } u = 0, \ldots, n\right\}. \tag{27}$$

Пусть максимум достигается на $a_0, \ldots, a_n$. Положим

$$(b_{ij}) = B = \sum_{k=0}^{n} \frac{a_k}{m \cdot v_k} D_k, \tag{28}$$

где m, v_k и D_k определены, как в разделе 1. Тогда матрица B удовлетворяет условиям, упомянутым в (23); матрица B положительно полуопределенная, поскольку в силу перестановочности матриц $D_0, \ldots, D_n$ она имеет собственные значения

$$\sum_{k=0}^{n} \frac{a_k}{m \cdot v_k} P_k^u = \sum_{k=0}^{n} \frac{a_k}{m \cdot \mu_u} Q_k^u \tag{29}$$

при $u = 0, \ldots, n$. Так как $D_k * J = v_k \cdot m$, то отсюда следует, что $B * J = \sum_{i,j} b_{ij} = \sum_k a_k$. Поэтому значение границы линейного программирования не превышает $\theta'(G)$. Чтобы доказать обратное

утверждение, предположим, что минимум в (12) достигается на $b_0, \ldots, b_n$, и положим $\lambda = \sum_u b_u$. Определим матрицу

$$(30)\quad A = \lambda I - \sum_{k,\, u=0}^{n} \frac{b_u}{\mu_u} Q_k^u \cdot D_k + J = \lambda I - \sum_{k=0}^{n} \left(\sum_{u=0}^{n} \frac{b_u}{\mu_u} Q_k^u - 1 \right) \cdot D_k.$$

Так как матрица $\lambda I - A$ имеет собственные значения

$$(31)\quad \sum_{k=0}^{n} \left(\sum_{u=0}^{n} \frac{b_u}{\mu_u} Q_k^u - 1 \right) \cdot P_k^v = \sum_{u=0}^{n} \frac{b_u}{\mu_u} \cdot m \cdot \delta_{uv} - m\delta_{v0} \geqslant 0$$

$(v = 0, \ldots, n)$, то наибольшее собственное значение матрицы A не превышает λ. Кроме того, согласно (4) и (12), $a_{ij} \geqslant 1$, если $\{i, j\} \notin E$. Поэтому минимум в (12) не меньше, чем минимум, содержащийся в формулировке теоремы 3, или, другими словами, значение границы линейного программирования не меньше $\theta'(G)$. ■

Если $(X, \mathscr{R})$ — это схема Джонсона с n классами (см. Дельсарт [2]) и $M = \{0, \ldots, n-1\}$, то $G = K(v, n)$. Так как Ловас показал, что

$$\theta(K(v, n)) = \binom{v-1}{n-1},$$

то из границы линейного программирования Дельсарта также следует теорема Эрдеша — Ко — Радо.

Применяя рассуждения, подобные тем, что были использованы в доказательстве теоремы 3, можно доказать следующий результат относительно симметрических ассоциативных схем $(X, \mathscr{R})$ и графов G, связанных таким образом, как это указано перед теоремой 3.

Теорема 4. $\theta(G) = \max \{\sum_{k=0}^{n} a_k \mid a_0 = 1;\ a_k = 0$ *для* $k \notin M;$ $\sum_{k=0}^{n} a_k Q_k^u \geqslant 0$ *для* $u = 0, \ldots, n\} = \min \{\sum_{u=0}^{n} b_u \mid b_0, \ldots, b_n \geqslant 0;$ $b_0 = 1;\ \sum_{u=0}^{n} b_u P_k^u = 0$ *для* $k \in M \setminus \{0\}\}$.

Таким образом, для графов, полученных из симметрических ассоциативных схем, это дает более простой способ вычисления величины θ. В качестве обобщения результата Дельсарта (13) получается следующая

Теорема 5. *Пусть множество ребер E графа $G = (V, E)$ является объединением некоторых классов симметрической ассоциативной схемы. Тогда* $\theta(G) \cdot \theta(\bar{G}) = |X|$.

Доказательство. Ловас доказал, что для всех графов G справедливо неравенство $\theta(G)\theta(\bar{G}) \geqslant |X|$. Предположим теперь, что

множество E есть объединение некоторых классов ассоциативной схемы, как это описано перед теоремой 3. Тогда, согласно теореме 4, $\theta(G)=\sum_k a_k$ для некоторых $a_0, \ldots, a_n$, где $a_k=0$ для $k\in M$ и $\sum_k a_k Q_k^u \geqslant 0$ для $u=0, \ldots, n$. Положим

$$b_u=\frac{\sum_{k=0}^{n} a_k Q_k^u}{\theta(G)}. \tag{32}$$

Тогда $b_0, \ldots, b_n \geqslant 0$ и $b_0=1$. Кроме того, для $k\notin M$ (ср. (5)) имеем

$$\sum_{u=0}^{n} b_u P_k^u=\frac{1}{\theta(G)}\cdot\sum_{u,\,l} a_l P_l^u Q_k^u=\frac{1}{\theta(G)}\cdot\sum_l a_l\cdot m\cdot\delta_{kl}=\frac{a_k\cdot m}{\theta(G)}=0. \tag{33}$$

Следовательно, $b_0, \ldots, b_n$ удовлетворяют условиям, упомянутым в последней строке теоремы 4 с заменой G на $\bar{G}$. Кроме того,

$$\begin{aligned}\sum_{u=0}^{n} b_u &= \frac{1}{\theta(G)}\cdot\sum_{k,\,u} a_k Q_k^u=\frac{1}{\theta(G)}\sum_k a_k\cdot\sum_u Q_k^u\\ &=\frac{1}{\theta(G)}\cdot\sum_k a_k\cdot m\cdot\delta_{k0}=\frac{|X|}{\theta(G)}.\end{aligned} \tag{34}$$

Так как, согласно теореме 4, $\sum_u b_u \geqslant \theta(\bar{G})$, то мы показали, что $\theta(G)\cdot\theta(\bar{G})\leqslant|X|$.

Поскольку существует много сильно регулярных графов, которые не являются вершинно-транзитивными (ср. Зайдель [8]), то теорема 5 не следует из (17). М. Р. Бест нашел следующий пример графа G, для которого $\theta'(G)<\theta(G)$. Вершинами графа G являются векторы в $\{0, 1\}^6$, и два вектора считаются смежными, если и только если расстояние Хэмминга между ними не больше 3 (таким образом множество ребер является объединением некоторым классов схемы Хэмминга). Тогда $\theta'(G)=4$, в то время как $\theta(G)=16/3$.

После завершения этого исследования мы узнали, что частично похожие результаты были получены независимо Мак-Элисом, Родемичем и Рамсеем [7] (ср. [6]). Их функции $\alpha_L(G)$ и $\theta_L(G)$ равны $\theta'(G)$ и $\theta(G)$ соответственно.

ЛИТЕРАТУРА

1. Best M. R., Brouwer A. E., MacWilliams, Odlyzko A. M., Sloane N. J. A. Bounds for binary codes of length less than 25, IEEE Trans. Inform. Theory, vol. IT-24, pp. 81—93, Jan. 1978. [Имеется перевод: Бест М. Р., Брауэр А. Е., Мак-Вильямс Ф. Дж., Одлижко А. М., Слоэн Н. Дж. А. Границы

для двоичных кодов длины меньше 25. — В кн.: Кибернетический сборник, вып. 17. — М.: Мир, 1980.]
2. Delsarte P. An algebraic approach to the association schemes in coding theory. Philips Res. Reps. Suppl. 10, 1973. [Имеется перевод: Дельсарт Ф. Алгебраический подход к схемам отношений теории кодирования. — М.: Мир, 1976.]
3. Haemers W. On some problems of Lovász concerning the Shannon capacity of a graph, IEEE Trans. Inform. Theory, vol. IT-25, no. 2, pp. 231—232, March 1979.
4. Lovász L. On the Shannon capacity of a graph, IEEE Trans. Inform. Theory, vol. IT-25, no. 1, pp. 1—7, Jan. 1979. [Имеется перевод: Ловас Л. О шенноновской емкости графа. — См. наст. сб.]
5. MacWilliams F. J., Sloane N. J. A. The theory of error-correcting codes. Amsterdam: North-Holland and New York: Elsevier North Holland, 1977. [Имеется перевод: Мак-Вильямс Ф., Слоэн Н. Теория кодов, исправляющих ошибки. — М.: Связь, 1979.]
6. McEliece R. J. The bounds of Delsarte and Lovász, and their applications to coding theory, presented at the International Centre for Mechanical Sciences, Udine, Italy, Summer, 1978.
7. McEliece R. J., Rodemich E. R., Rumsey H. C., Jr. The Lovász bound and some generalizations, J. Combinatorics, Inform. Syst. Sci., vol. 3, 1978, pp. 134—152. (Имеется перевод: Мак-Элис Р. Д., Родемич Е. Р., Рамсей Г. С. Граница Ловаса и некоторые обобщения. — См. настоящий сборник.)
8. J. J. Seidel, Graphs and two-graphs, in Proc. 5th Southeastern Conf. Comb., Graph Theory and Computing. Winnipeg: Utilitas, 1974, pp. 125—143.
9. Shannon C. E. The zero-error capacity of a noisy channel, IRE Trans. Inform. Theory, vol. IT-3, pp. 3—15, 1956. (Имеется перевод: Шеннон К. Пропускная способность канала с шумом при нулевой ошибке. — В кн.: Шеннон К. Работы по теории информации и кибернетике. — М.: ИЛ, 1963, с. 464—487.)

Граница Ловаса и некоторые обобщения[1])

Р. Мак-Элис, Е. Родемич, Г. Рамсей

Недавно Ловас [5] предложил удивительно простое решение чрезвычайно трудной комбинаторной задачи («нахождение емкости пятиугольника»), поставленной Шенноном [8] в 1956 г. В этой работе мы покажем, что идеи Ловаса в сочетании с некоторыми нашими идеями приводят к очень мощному и общему методу решения комбинаторных задач упаковки, который мы изложим на языке теории графов В частности, граница линейного программирования Дельсарта [3] для клик в ассоциативных схемах окажется специальным случаем границы Ловаса.

ВВЕДЕНИЕ

Пусть $V = \{v_1, \ldots, v_N\}$ — конечное множество из N элементов, а E — совокупность двухэлементных подмножеств V. Множество G, состоящее из элементов $\{v_i\}$ из V и элементов E, называется *графом*[2]) на V. Элементы V называются *вершинами*, а элементы E — *ребрами* графа G. На рис. 1 изображен граф, представляющий для нас особый интерес:

$$G = \{\{0\}, \{1\}, \{2\}, \{3\}, \{4\}, \{0, 1\}, \{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 0\}\}.$$

Вершины здесь представлены точками на плоскости, а ребра — линиями, соединяющими соответствующие пары вершин. (Для дальнейших ссылок обозначим этот граф через C_5. Это и есть «пятиугольник», упомянутый выше.)

Подмножество $Y = \{y_1, \ldots, y_M\} \subseteq V$ называется *независимым множеством*, если никакая пара $\{y_i, y_j\}$, $i \neq j$ не является ребром графа G. Мощность наибольшего независимого множества в G обозначается $\alpha(G)$:

$$\alpha(G) = \max\{|Y|: Y \text{ — независимое множество в } G\}. \quad (1.1)$$

Например, $\alpha(C_5) = 2$, и множество $Y = \{0, 2\}$, обведенное кружками на рис. 1, является максимальным независимым множеством.

---

[1]) McEliece R. J., Rodemich E. R., Rumsey, Jr., H. C., The Lovász bound and some generalizations, Journal of Combinatorics, Information and System Sciences, 3, No. 3 (1978), 134—152.

[2]) Более точно — неориентированным графом

© Journal of Combinatorics, Information & System Sciences, Delhi, 1978
© Перевод на русский язык, «Мир», 1983

Для любого целого $n \geqslant 2$ определим n-ю *прямую степень* G, обозначаемую через G^n, следующим образом. Множество вершин G^n — это декартова степень V^n, т. е. множество всех N^n n-последовательностей $\mathbf{v} = (v_1, \ldots, v_n)$ элементов из V. Множество ребер G^n состоит из всех пар $\{\mathbf{v}, \mathbf{v}'\}$ из V^n, таких, что

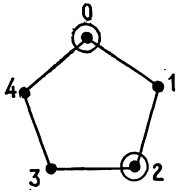


Рис. 1. Граф C_5; $\alpha(C_5) = 2$. (Это также граф Q_5 из примера 3 разд. 4.)

$\{v_i, v_i'\} \in G$ для всех i. Заметим, что если A — матрица смежности графа G, т. е. следующая $N \times N$-матрица, столбцы и строки которой занумерованы элементами множества V,

$$A(v, v') = \begin{cases} 1, & \text{если } \{v, v'\} \in G \text{ или } v = v', \\ 0 & \text{в противном случае,} \end{cases}$$

то матрицей смежности графа G^n является n-я прямая (кронекерова) степень матрицы A.

Емкость графа G, обозначаемая через $\theta(G)$, определяется следующим образом:

$$\theta(G) = \sup_n \alpha(G^n)^{1/n}. \tag{1.2}$$

Это понятие[1]) было введено Шенноном [8] в связи с проблемой нахождения пропускной способности дискретного канала без памяти при нулевой вероятности ошибки. Шеннон разработал методы, которые позволили ему вычислить емкость многих, но не всех графов.

Например, он показал, что если существует отображение ϕ множества V в независимое множество графа G, такое, что из $\{v, v'\} \notin G$ следует $\{\phi(v), \phi(v')\} \notin G$, то $\alpha(G^n) = \alpha(G^n)$ для всех n и, следовательно, $\theta(G) = \alpha(G)$.

Он также получил с помощью *линейного программирования* следующую верхнюю границу для $\theta(G)$. Пусть P — распределение вероятностей на V, т. е. $P(v) \geqslant 0$, $\sum \{P(v): v \in V\} = 1$. На подмножества $X \subseteq Y$ распределение P доопределяется по аддитивности: $P(X) = \sum \{P(x): x \in X\}$. Подмножество X называется *кликой* в G, если $\{x, x'\} \in G$ для всех $x, x' \in X$. Шеннон доказал, что $\theta(G) \leqslant \lambda^{-1}$, где

$$\lambda = \min_P \max \{P(X): X \text{ — клика}\}, \tag{1.3}$$

[1]) На самом деле Шеннон определял емкость как логарифм введенной нами величины $\theta(G)$.

и минимизация в (1.3) берется по всем возможным вероятностным распределениям.

Эти два результата позволили Шеннону вычислить емкости всех графов с пятью или меньшим числом вершин, за исключением графа C_5, изображенного на рис. 1. Для C_5 его результаты [1]) давали только

$$\sqrt{5} \leqslant \theta(C_5) \leqslant 5/2. \tag{1.4}$$

21 год спустя Ловас установил, что $\theta(C_5) \leqslant \sqrt{5}$; это в сочетании с нижней границей (1.4) показывает, что $\theta(C_5) = \sqrt{5}$. Кратко опишем метод Ловаса нахождения верхней границы для $\theta(G)$.

Ловас определяет *ортонормированное представление* G как реализацию G, в которой множество вершин V является множеством единичных (т. е. длины 1) векторов в евклидовом векторном пространстве, обладающих следующим свойством: если $v \cdot v' \neq 0$, то $\{v, v'\} \in G$.

Если G имеет такое ортонормированное представление и b — произвольный единичный вектор, то граница Ловаса имеет вид

$$\theta(G) \leqslant (\min_{v \in V} (v \cdot b)^2)^{-1}. \tag{1.5}$$

Ловас применил (1.5) к графу C_5, рассмотрев «зонтик» с пятью ребрами $\{v_1, \ldots, v_5\}$ единичной длины. Если зонтик открыт так, что угол между несмежными ребрами равен 90°, то $(v_1, \ldots, v_5)$ — ортонормированное представление G в трехмерном евклидовом пространстве. Если ручка зонтика b — также единичный вектор, то легко показать, что $b \cdot v_i = 5^{-1/4}$ для всех i и, следовательно, согласно (1.5), $\theta(C_5) \leqslant \sqrt{5}$.

Ловас также указал много других следствий из (1.5), которые мы сейчас не приводим. Однако заметим, что некоторые примеры разд. 4, приведенные ниже, фигурируют также в работе Ловаса и были получены им ранее. (Мы дадим ссылки на теоремы работы Ловаса в соответствующих местах разд. 4).

Настоящая статья возникла из попыток перенести результаты Ловаса на более общий случай. Думаем, что нам это удалось, но в нашем решении ортонормированные представления полностью исчезли. Тем не менее все границы, которые мы получили (по крайней мере границы для $\theta(G)$), могут быть получены из ортонормированного представления, поэтому мы и называем их *границами Ловаса*. (В приложении А содержатся доказательства эквивалентности нашего метода и метода Ловаса.)

---

[1]) Нижняя граница в (1.4) является следствием того факта (также установленного Шенноном), что $\alpha(C_5^2) = 5$.

В разд. 2 мы приведем наш вывод границ Ловаса. Как будет видно, вычисление этих границ для фиксированного графа G требует решения определенной задачи нелинейного программирования.

В разд. 3 мы покажем, что для достаточно симметричного графа G (точный смысл этого выражения дан в разд. 3) эта задача нелинейного программирования превращается в задачу *линейного* программирования.

В разд. 4 мы применяем наши более общие результаты к некоторым примерам. Сначала мы приводим очень простую границу для $\theta(G)$, которая применима к любым регулярным графам. Затем вычисляем границу Ловаса для двух бесконечных семейств графов: циклических графов C_N и квадратично-вычетных графов Q_p, которые являются двумя различными обобщениями графа C_5. Нам не удалось вычислить емкости графов C_N для нечетных N, $N \geqslant 7$, но для любого простого p, $p \equiv 1 \bmod 4$, мы покажем, что $\theta(Q_p) = \sqrt{p}$. Далее рассматриваются три специальных графа: граф Петерсена, графы икосаэдра и додекаэдра. Мы также рассматриваем очень интересный регулярный граф на 7 вершинах.

В разд. 4 в качестве последнего примера будет показано, что принадлежащая Дельсарту [3] граница «линейного программирования» для клик в ассоциативных схемах есть частный случай наших результатов.

2. ВЕРХНЯЯ ГРАНИЦА ЛОВАСА

В этом и следующем разделах G будет обозначать фиксированный граф. Мы постоянно будем иметь дело с векторами и матрицами, компоненты которых помечены множеством вершин V графа G. Если x — такой вектор и $v \in V$, то v-я компонента x будет обозначаться через $x(v)$; если A — такая матрица, то ее (v, v')-я компонента будет обозначаться через $A(v, v')$.

Будем также рассматривать *квадратичные формы*, ассоциированные с такими векторами и матрицами. Если x есть вектор-столбец и A — симметрическая матрица, то квадратичная форма x^TAx определяется как

$$x^TAx = \sum_{(v,\, v') \in V^2} x(v)\, x(v')\, A(v,\, v'). \tag{2.1}$$

Мы также будем рассматривать x^TAx как функцию от компонент вектора x.

Следующая теорема без сомнения выглядит совсем тривиальной, и все же это наш главный результат. Остальная часть статьи будет посвящена главным образом изучению его следствий.

Теорема 1. *Пусть A — симметрическая действительная матрица, такая, что*

$$A(v, v') = 1, \quad \text{если} \quad v = v',$$
$$A(v, v') \leqslant 0, \quad \text{если} \quad \{v, v'\} \notin G.$$

Тогда если $u = (1, \ldots, 1)$ обозначает вектор из одних единиц, то

$$\inf \{x^T A x\colon\ x \cdot u = 1\} \leqslant \alpha(G)^{-1}. \tag{2.2}$$

Доказательство. Пусть $Y \subseteq V$ — максимальное независимое множество в G, т. е. $|Y| = \alpha(G)$. Определим вектор y как $y(v) = \alpha(G)^{-1}$, если $v \in Y$, и $y(v) = 0$ в противном случае. Тогда ясно, что $y \cdot u = 1$, а в силу (2.1) $y^T A y \leqslant \alpha(G)^{-1}$. Это и доказывает (2.2). ■

Обозначим через $\lambda(A)$ значение левой части (2.2):

$$\lambda(A) = \inf \{x^T A x\colon\ x \cdot u = 1\} \tag{2.3}$$

Теорема 1 дает верхнюю границу для $\alpha(G)$, а именно $\alpha(G) \leqslant \lambda(A)^{-1}$ при условии, что $\lambda(A) > 0$. Ясно, что, для того чтобы применять эту границу, необходимо иметь больше сведений о функции $\lambda(A)$. Для дальнейшего мы перечислим теперь некоторые из наиболее важных ее свойств. Везде предполагаем, что матрица A — действительная и симметрическая (Доказательства этих свойств можно найти в приложении B.)

Прежде всего отметим, что $\lambda(A)$ отрицательно, если матрица A не является *положительной полуопределенной* (сокращенно п. п. о.), т. е. если не выполняется $x^T A x \geqslant 0$ для всех x:

$$\lambda(A) \geqslant 0, \quad \text{если и только если } A \text{ — п. п. о. матрица.} \tag{2.4}$$

Предполагая, что A — п. п. о. матрица, обозначим через $\xi_1, \ldots, \xi_N$ полное ортонормальное множество собственных векторов A, т. е. $\xi_i \cdot \xi_j = \delta_{ij}$ и $A\xi_j = \lambda_j \xi_j$. Так как A — п. п. о. матрица, то все собственные значения λ_j неотрицательны. Пусть $u = u_1\xi_1 + \ldots + u_N\xi_N$ — разложение u в этом базисе.

Тогда

$$\lambda(A) = 0, \quad \text{если} \quad u_j \neq 0,\ \lambda_j = 0 \text{ для некоторого } j, \tag{2.5}$$

$$\lambda(A) = \left(\sum u_j^2/\lambda_j\right)^{-1} \quad \text{в противном случае,} \tag{2.6}$$

где суммирование в (2.6) ведется только по индексам j, таким, что $\lambda_j > 0$. (Можно считать, что (2.6) включает (2.5) как особый случай, если мы расширим суммирование на все j и условимся, что $\infty^{-1} = 0$, $u^2/0$ равно 0, если $u = 0$, или равно ∞, если $u \neq 0$.)

Если u сам является собственным вектором A с собственным значением σ, то $\lambda(A)$ вычисляется намного проще:

$$\lambda(A)=\sigma/N, \quad \text{если} \quad Au=\sigma u \text{ и } A \text{ — п. п. о. матрица.} \tag{2.7}$$

Существует полезная двойственная формулировка определения (2.3) для п. п. о. матриц:

$$\lambda(A)=\max\{\lambda\colon A-\lambda J \text{ — п. п. о. матрица}\}, \tag{2.8}$$

где J обозначает матрицу из одних единиц.

Наш последний вспомогательный результат состоит в том, что функция $\lambda(A)$ *мультипликативна* на п. п. о. матрицах:

$$\lambda(A\times B)=\lambda(A)\lambda(B), \quad \text{если} \quad A \text{ и } B \text{ — п. п. о. матрицы}\}, \tag{2.9}$$

где $A\times B$ обозначает прямое произведение матриц A и B.

Согласно (2.4), теорема 1 дает нетривиальную информацию о $\alpha(G)$, если только A — п. п. о. матрица. Это приводит нас к определению следующих двух множеств матриц.

Определение 1. Множество $\Omega(G)$ определяется как множество всех матриц $A=(A(v,v'))$, занумерованных вершинами графа G и удовлетворяющих условиям:

$$A \text{ является п. п. о. матрицей,} \tag{2.10}$$

$$A(v,v)=1 \quad \text{для всех} \quad v\in V, \tag{2.11}$$

$$A(v,v')\leqslant 0, \quad \text{если} \quad \{v,v'\}\notin G. \tag{2.12}$$

Определение 2. Аналогично, $\Omega_0(G)$ определяется как множество матриц, удовлетворяющих (2.10) и (2.11), а условие (2.12) заменяется более сильным

$$A(v,v')=0, \quad \text{если} \quad \{v,v'\}\notin G. \tag{2.12'}$$

Значение класса $\Omega(G)$ очевидно из теоремы 1 и (2.4).

Теорема 2.

$$\alpha(G)\leqslant\lambda(A)^{-1} \quad \textit{для всех} \quad A\in\Omega(G).$$

Важность класса $\Omega_0(G)$ определяется следующей теоремой, которая по существу эквивалентна границе Ловаса (1.5). (Доказательство этой эквивалентности см. в приложении А.) ■

Теорема 3. $\theta(G)\leqslant\lambda(A)^{-1}$ *для всех* $A\in\Omega_0(G)$.

Доказательство. Ключом к доказательству является тот факт, что если $A\in\Omega_0(G)$, то n-я прямая степень $A^{[n]}=A\times A\times\ldots\times A$ (n сомножителей) принадлежит $\Omega_0(G^n)$.

Чтобы показать это, заметим вначале, что матрица $A^{[n]}$, будучи прямым произведением п. п. о матриц, является п. п. о.

матрицей. Далее, если $\mathbf{v}=(v_1,\ldots,v_n)$ и $\mathbf{v}'=(v'_1,\ \ldots,\ v'_n)$ — вершины графа G^n, то по определению прямого произведения

$$A^{[n]}(\mathbf{v},\ \mathbf{v}')=\prod_{j=1}^{n} A(v_j,\ v'_j). \tag{2.13}$$

Отсюда непосредственно следует, что $A^{[n]}(\mathbf{v},\mathbf{v})=1$ для всех $\mathbf{v}\in G^n$, так как каждый из сомножителей в правой части (2.13) равен 1. Если $\{\mathbf{v},\ \mathbf{v}'\}\notin G^n$, то существует по крайней мере один индекс j, такой, что $\{v_j,\ v'_j\}\notin G$. Поскольку $A\in\Omega_0(G)$, из (2.12′) следует, что $A(v_j,\ v'_j)=0$ и, значит, $A^{[n]}(\mathbf{v},\mathbf{v}')=0$. Следовательно, $A^{[n]}$ удовлетворяет (2.10), (2.11) и (2.12′) и, таким образом, принадлежит $\Omega_0(G^n)$.

Так как $\Omega_0(G^n)\subseteq\Omega(G^n)$, то можно применить теорему 2 к матрице $A^{[n]}$ и сделать вывод, что $\alpha(G^n)\leqslant\lambda(A^{[n]})^{-1}$. Но в силу (2.9) $\lambda(A^{[n]})=\lambda(A)^n$. Следовательно, $\alpha(G^n)\leqslant\lambda(A)^{-n}$ для всех n, и, таким образом, из определения (1.2) величины $\theta(G)$, получаем, что $\theta(G)\leqslant\lambda(A)^{-1}$. ■

Исходя из теорем 2 и 3, введем теперь *границы Ловаса* $\alpha_L(G)$ и $\theta_L(G)$:

$$\alpha_L(G)=\min\{\lambda(A)^{-1}:\ A\in\Omega(G)\}, \tag{2.14}$$

$$\theta_L(G)=\min\{\lambda(A)^{-1}:\ A\in\Omega_0(G)\}. \tag{2.15}$$

В теоремах 2 и 3 мы показали, что $\alpha(G)\leqslant\alpha_L(G)$, $\theta(G)\leqslant\theta_L(G)$. К сожалению, мы не знаем ни одного эффективного алгоритма вычисления $\alpha_L(G)$ и $\theta_L(G)$ для произвольного графа. Однако мы сейчас покажем, что можно использовать симметрии графа G для некоторого упрощения вычислений. В разд. 3 мы расширим идеи и покажем, что если граф G весьма симметричен, то границы $\alpha_L(G)$ и $\theta_L(G)$ могут быть вычислены с помощью линейного программирования.

Симметрия графа G — это подстановка на множестве вершин V, сохраняющая множество ребер E инвариантным. Таким образом, подстановка π на множестве V является симметрией графа G, если и только если $\{\pi(v),\ \pi(v')\}\in E$ для $\{v,v'\}\in E$. Заметим, что симметрия графа G является также симметрией *дополнительного графа* G', имеющего множество вершин V и множество ребер E', состоящее из пар, не принадлежащих E.

Пусть P — группа симметрий G, а $E_1,\ \ldots,\ E_s$ — орбиты E под действием P. Аналогично, пусть $E'_1,\ \ldots,\ E'_t$ — орбиты E'. Два ребра, лежащие в одной и той же орбите, назовем *эквивалентными*.

Предположим теперь, что $A \in \Omega(G)$ или $\Omega_0(G)$. Тогда легко видеть, что матрица $\bar{A}$, определенная следующим образом:

$$\bar{A}(v, v') = \frac{1}{|P|} \sum_{\pi \in P} A(\pi(v), \pi(v')), \tag{2.16}$$

также лежит в том же множестве. Кроме того, матрица $\bar{A}$ обладает тем свойством, что $\bar{A}(v_1, v_1') = \bar{A}(v_2, v_2')$, если $\{v_1, v_1'\}$ и $\{v_2, v_2'\}$ — эквивалентные ребра. Покажем дополнительно, что $\lambda(\bar{A}) \geqslant \lambda(A)$. Обозначив через $\pi(A)$ матрицу с элементами $A(\pi(v), \pi(v'))$, для любого значения λ имеем

$$\bar{A} - \lambda J = \frac{1}{|P|} \sum_{\pi \in P} (\pi(A) - \lambda I) = \frac{1}{|P|} \sum_{\pi \in P} \pi(A - \lambda J).$$

Если положить $\lambda = \lambda(A)$, то в силу (2.8) матрица $A - \lambda J$ является п. п. о. матрицей и, следовательно, таковой является каждая из матриц $\pi(A - \lambda J)$. Таким образом, $\bar{A} - \lambda J$ — п. п. о. матрица и из (2.8) следует, что $\lambda(\bar{A}) \geqslant \lambda(A)$.

Обозначим через B_j, $j = 1, 2, \ldots, s$, матрицы смежности орбит E_j

$$B_j(v, v') = \begin{cases} 1, & \text{если } \{v, v'\} \in E_j, \\ 0 & \text{в противном случае.} \end{cases}$$

Аналогично определим матрицы B_k', $k = 1, 2, \ldots, t$, как матрицы смежности орбит E'. Тогда, согласно предыдущим рассуждениям, матрица $\bar{A}$ может быть представлена в виде линейной комбинации этих матриц и единичной $N \times N$ матрицы I:

$$\bar{A} = I + \sum_{j=1}^{s} \mu_j B_j + \sum_{k=1}^{t} \mu_k' B_k'. \tag{2.17}$$

Мы показали таким образом, что для произвольной матрицы $A \in \Omega(G)$ (соответственно $\Omega_0(G)$) можно построить матрицу $\bar{A}$ вида (2.17), лежащую в том же классе и такую, что $\lambda(\bar{A}) \geqslant \lambda(A)$. Это означает, что при вычислении границ $\alpha_L(G)$ и $\theta_L(G)$ можно без ограничения общности рассматривать только матрицы вида (2.17). Более формально определим

$\bar{\Omega}(G) = \{$п. п. о. матрицы вида (2.17) с $u_k \leqslant 0$, $k = 1, 2, \ldots, t\}$,

$\bar{\Omega}_0(G) = \{$п. п. о. матрицы вида (2.17) с $u_k' = 0$, $k = 1, 2, \ldots, t\}$.

Тогда имеем более простое для вычислений определение границ Ловаша

$$\alpha_L(G) = \min\{\lambda(A)^{-1}: A \in \bar{\Omega}(G)\},$$

$$\theta_L(G) = \min\{\lambda(A)^{-1}: A \in \Omega_0(G)\}.$$

3. ГРАНИЦА ЛИНЕЙНОГО ПРОГРАММИРОВАНИЯ ДЛЯ $\alpha(G)$ и $\Theta(G)$

В этом разделе будет показано, что для достаточно симметричного графа G вычисление границ $\alpha_L(G)$ и $\theta_L(G)$ может быть значительно упрощено.

Требуемая степень симметричности будет заключаться в том, чтобы *матрицы смежности* $\{B_j\}\{B'_k\}$ в (2.17) *были перестановочны* друг с другом. То, что это фактически является утверждением о группе симметрий графа G, можно увидеть следующим образом.

Пусть P — группа симметрий G. Каждому $\pi \in P$ сопоставим соответствующую матрицу подстановки π^*:

$$\pi^*(v, v') = \begin{cases} 1, & \text{если } \pi(v) = v', \\ 0 & \text{в противном случае.} \end{cases}$$

Естественно, орбиты $\{E_j\}$, $\{E'_k\}$ левоинвариантны по отношению к симметриям $\pi \in P$; в терминах соответствующих матриц смежности это может быть выражено следующим образом:

$$\begin{aligned} &\pi^* B_j = B_j \pi^* \quad \text{для всех} \quad j = 1, 2, \ldots, s;\ \pi \in P, \\ &\pi^* B'_k = B'_k \pi^* \quad \text{для всех} \quad k = 1, 2, \ldots, t;\ \pi \in P. \end{aligned} \tag{3.1}$$

Пусть P^* обозначает группу всех матриц подстановок, соответствующую подстановкам группы P, и пусть $Z(P^*)$ — централизаторное кольцо группы P^*, т. е. множество всех матриц, которые перестановочны со всеми $\pi^* \in P^*$. В соответствии с (3.1) матрицы $\{B_j\}$, $\{B'_k\}$ принадлежат $Z(P^*)$.

Если известно, что кольцо $Z(P^*)$ коммутативно, то из этого немедленно следует, что матрицы B_j, B'_k перестановочны друг с другом. К счастью, этот случай часто осуществляется. Действительно, можно показать ([10], гл. 5), что для транзитивной группы P кольцо $Z(P^*)$ коммутативно тогда и только тогда, когда комплексное представление группы P, порожденное группой матриц P^*, разлагается в сумму неэквивалентных неприводимых представлений. В частности, если P содержит транзитивную абелеву подгруппу или если для любой пары (v, v') различных вершин в группе P существует элемент, переставляющий их, то это условие выполняется.

Исходя из предыдущих рассуждений, представим теперь наши результаты в следующем обобщенном виде.

Пусть V — конечное множество, содержащее N элементов, и пусть $\{E_1, E_2, \ldots, E_n\}$ — разбиение совокупности E всех двухэлементных подмножеств V. Для каждого $j = 1, 2, \ldots, n$ обозначим через A_j матрицу смежности E_j:

$$A_j(v, v') = \begin{cases} 1, & \text{если } \{v, v'\} \in E_j, \\ 0 & \text{в противном случае.} \end{cases}$$

Пусть A_0 обозначает единичную $N \times N$ матрицу. Предположим, что матрицы $\{A_j\colon j = 0, 1, \ldots, n\}$ перестановочны друг с другом. В итоге наши предположения состоят в том, что A_j суть матрицы из 0 и 1, удовлетворяющие следующим условиям:

$$A_0 = I, \qquad \sum_{j=0}^{n} A_j = J. \tag{3.2}$$

Все матрицы A_j симметрические. (3.3)

$$A_j A_k = A_k A_j \quad \text{для всех} \quad j, k = 0, 1, \ldots, n. \tag{3.4}$$

Для фиксированного подмножества C множества $\{1, 2, \ldots, n\}$ через G_G будем обозначать граф с множеством вершин V и множеством ребер

$$E_C = \bigcup_{j \in C} E_j. \tag{3.5}$$

Наша цель состоит в том, чтобы дать верхнюю границу «линейного программирования» для $\alpha(G_c)$ и $\theta(G_c)$ (теоремы 4 и 5 ниже). Однако, чтобы сформулировать эти результаты, необходимо некоторое предварительное обсуждение.

Заметим, что в силу (3.2) и (3.4) любая матрица A_j перестановочна с матрицей J из одних единиц и, следовательно, все $n+2$ матрицы J, A_0, A_1, ..., A_n попарно перестановочны. Поскольку эти матрицы, кроме того, симметрические, следовательно, они одновременно диагонализуемы, так как по известной теореме линейной алгебры (см. [4], гл. 6, теорема 4) для этих $n+2$ матриц существует множество $\{\xi_m\}_{m=1}^{N}$ линейно независимых *общих собственных векторов*.

В частности, ξ_m являются собственными векторами для J:

$$J\xi_m = \lambda_m \xi_m, \qquad m = 1, 2, \ldots, N, \tag{3.6}$$

где $\{\lambda_m\}$ — множество собственных значений J. Но собственными значениями J являются только значения $\{0, N\}$, и простое вычисление показывает, что если $J\xi = N\xi$, то вектор ξ должен быть кратен вектору u из одних единиц. Таким образом, можно считать, что $\xi_1 = u$.

Теперь для всех j, m определим собственные значения $\lambda_{j,m}$ следующим образом:

$$A_j \xi_m = \lambda_{j,m} \xi_m, \qquad j = 0, 1, \ldots, n; \quad m = 1, 2, \ldots, n. \tag{3.7}$$

Мы подошли к нашим границам «линейного программирования» для $\alpha(G)$ и $\theta(G)$.

Теорема 4. *Пусть* $\mu_1, \mu_2, \ldots, \mu_n$ — *действительные числа, такие, что*

$$\mu_j \leqslant 0, \quad \text{если} \quad j \in C \tag{3.8}$$

и

$$1+\sum_{j=1}^{n}\mu_j\lambda_{j,m}\geqslant 0 \qquad m=1,\ 2,\ \ldots,\ N. \tag{3.9}$$

Тогда

$$\alpha(G_C)\leqslant N\Big/\Big(1+\sum_{j=1}^{n}\mu_j\lambda_{j,1}\Big).$$

Теорема 5. *Пусть* $\mu_1, \ldots, \mu_n$ *удовлетворяют* (3.9) *и*

$$\mu_j=0 \quad \text{для} \quad j\notin C. \tag{3.8'}$$

Тогда

$$\theta(G_C)\leqslant N\Big/\Big(1+\sum_{j=1}^{n}\mu_j\lambda_{j,1}\Big).$$

Доказательство. Для данных констант $\{\mu_j\}$ определим матрицу

$$A=I+\sum_{j=1}^{n}\mu_jA_j.$$

Ясно, что векторы $\{\xi_m\}$ являются собственными для A, так как

$$A\xi_m=\xi_m+\sum_{j=1}^{n}\mu_jA_j\xi_m=\Big(1+\sum_{j=1}^{n}\mu_j\lambda_{j,m}\Big)\xi_m.$$

Более того, предположение (3.9) гарантирует, что собственные значения $\{1+\sum\mu_j\lambda_{j,m}\}$ матрицы A неотрицательны, и, следовательно, A — п. п. о. матрица. Условия (3.8) и (3.8′) теперь означают, что матрица A принадлежит $\Omega(G_C)$ или $\Omega_0(G_C)$. Таким образом, из теорем 2 и 3 получаем, что $\alpha(G_C)<\lambda(A)^{-1}$, $\theta(G_C)\leqslant\lambda(A)^{-1}$.

Для вычисления $\lambda(A)$ заметим, что $\xi_1=u$ является собственным вектором A с соответствующим собственным значением $1+\sum_{j=1}^{n}\mu_j\lambda_{j,1}$, а, согласно (2.7), $\lambda(A)=(1+\sum\mu_j\lambda_{j,1})/N$, откуда и следуют теоремы 4 и 5. ■

Для того чтобы получить наилучшие возможные границы среди границ, приведённых в теоремах 4 и 5, нам по существу требуется максимизировать линейную функцию $\sum_{1}^{n}\mu_j\lambda_{j,1}$ при линейных ограничениях (3.8) или (3.8′) и (3.9). Это задача линейного программирования (если известны собственные векторы и собственные значения матриц A_j), и, следовательно, нам удалось показать, что *границы Ловаса* $\alpha_L(G)$ и $\theta_L(G)$ *могут быть вычислены с помощью линейного программирования при условии, что матрицы смежности* B_j, B'_k *орбит* E_j, E'_k *перестановочны.*

4. НЕКОТОРЫЕ ПРИМЕНЕНИЯ

В этом разделе мы опишем некоторые из многих возможных применений предыдущих результатов. В частности, мы получим первоначальный результат Ловаса о емкости графа C_5 и границу «линейного программирования» Дельсарта для клик в ассоциативных схемах.

Пример 1. Регулярные графы. Говорят, что G — *регулярный* граф, если число ребер, содержащих данную вершину v, есть константа r, не зависящая от v. Число r называется *валентностью* графа G.

Следствием регулярности графа G является то, что матрица смежности B, соответствующая множеству ребер E, перестановочна с J: $JB = BJ$. Очевидно, что матрица B, перестановочна также с единичной матрицей I и, следовательно, матрицей $B' = J - I - B$, которая является матрицей смежности ребер дополнительного графа. Таким образом, можно применить теорему 5. Опуская очевидные подробности, получаем

$$\theta(G) \leqslant \frac{N}{1 + r/|\lambda_{\min}|}, \tag{4.1}$$

где $\lambda_{\min}$ — наименьшее собственное значение B (которое с необходимостью отрицательно, если E не пусто).

Более того, если к тому же группа симметрий графа G переставляет ребра транзитивно, то из наших результатов следует, что $\theta_L(G)$ равна правой части (4.1), т. е. (4.1) является наилучшей возможной границей такого типа. (Граница (4.1) эквивалентна теореме 9 Ловаса.)

Пример 2. Графы C_N. Обозначим через C_N циклический граф на N вершинах, т. е. $V = \{0, 1, \ldots, N-1\}$, $E = \{\{i, i+1\}: i = 0, 1, \ldots, N-1\}$, где индексы берутся по модулю N. Все эти графы регулярны, а циклическая группа порядка N переставляет ребра транзитивно, таким образом, мы можем вычислить $\theta_L(C_N)$ по формуле (4.1).

Для нахождения собственных значений матрицы инцидентности B заметим, что векторы $x(\zeta) = (1, \zeta, \ldots, \zeta^{N-1})$, где ζ — произвольный комплексный корень N-й степени из единицы, образуют независимое множество собственных векторов для B. Действительно,

$$Bx(\zeta) = (\zeta + \zeta^{-1})\, x(\zeta).$$

Отсюда собственные значения B равны $\{(\zeta + \zeta^{-1})\} = \{2\cos(2\pi k/N): k = 0, 1, \ldots, \lfloor N/2 \rfloor\}$. Очевидно, что наименьшее собственное значение равно -2, если N четное, и $2\cos(\pi - \pi/N) = -2\cos(\pi/N)$, если N нечетное и $N \geqslant 3$. Таким

образом, ввиду (4.1) имеем

$$\theta(C_N)\leqslant\begin{cases} N/2, & N\text{ — четное,}\\ N/(1+\cos(\pi/N)^{-1}, & N\text{ — нечетное,}\quad N\geqslant 3.\end{cases}\tag{4.2}$$

Для четных N или $N=3$ эта граница является точной, но эти результаты совсем элементарные и были известны уже Шеннону.

Однако для нечетных $N\geqslant 5$ эти границы нетривиальны (они содержатся в работе [5], следствие 5). Например, при $N=5$ имеем $\theta(C_5)\leqslant\sqrt{5}$. Для нечетных $N\geqslant 7$ верхние и нижние границы для $\theta(C_N)$ не совпадают. Ниже приводится таблица верхних границ (4.2) и наилучших известных нижних границ [1] для нечетных N, $7\leqslant N\leqslant 19$:

$$7^{3/5}=3.21410\leqslant\theta(C_7)\leqslant 3.31767$$
$$81^{1/3}=4.32675\leqslant\theta(C_9)\leqslant 4.36009$$
$$148^{1/3}=5.28957\leqslant\theta(C_{11})\leqslant 5.38630$$
$$247^{1/3}=6.27431\leqslant\theta(C_{13})\leqslant 6.40417$$
$$380^{1/3}=7.24316\leqslant\theta(C_{15})\leqslant 7.41715$$
$$4513^{1/4}=8.37214\leqslant\theta(C_{17})\leqslant 8.42701$$
$$7666^{1/4}=9.35712\leqslant\theta(C_{19})\leqslant 9.43477$$

Пример 3. Квадратично-вычетные графы. Пусть p — простое число, такое, что $p\equiv 1\,(\mathrm{mod}\,4)$. Граф Q_p имеет множество вершин $V=\{0, 1, \ldots, p-1\}$ и множество ребер $E\{\{v, v'\}: v-v'$ является квадратичным вычетом по модулю $p\}$. (Заметим, что Q_5 изоморфен графу «пятиугольника» C_5.) Граф Q_p регулярен с валентностью $(p-1)/2$. Матрица смежности задается следующим образом:

$$B_p(v, v')=\begin{cases} 1, & \text{если } v-v' \text{ является квадратичным вычетом,}\\ 0 & \text{в противном случае.}\end{cases}$$

Легко проверяется, что p-векторы $x(\xi)=(1, \xi, \ldots, \xi^{p-1})$, где ξ — произвольный комплексный корень p-й степени из 1, являются собственными векторами для B_p и что собственное значение вектора $x(\xi)$ равно $\sum\{\xi^a: a$ — квадратичный вычет$\}$. Хорошо известно (например, ср. [9] гл. 3 § 11), что эти суммы принимают только три различных значения $(p-1)/2$, $(-1\pm\sqrt{p})/2$. Отсюда наименьшим собственным значением матрицы B_p является $(-1-\sqrt{p})/2$, и из (4.1) следует, что $\theta(Q_p)\leqslant\sqrt{p}$. С другой стороны, если b — фиксированный квадратичный невычет по модулю p, то p упорядоченных пар

(v, bv), $v \in V$, образуют независимое множество в Q_p^2 и, следовательно, $\alpha(Q_p^2) \geqslant p$. Эти два неравенства устанавливают тот факт, что $\theta(Q_p) = \sqrt{p}$ для всех p, $p \equiv 1 \pmod 4$. Ввиду этого результата ясно, что графы Q_p представляют собой более подходящее обобщение пятиугольника, чем графы C_N. (Эти графы не фигурируют в работе Ловаса, но в теореме 12 он показы-

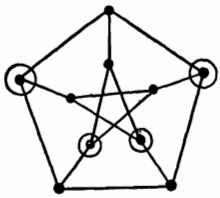


Рис. 2. Граф Петерсена.

вает, что если граф G самодополнительный и группа симметрии графа G транзитивна на вершинах, то $\theta(G) = \sqrt{|V|}$. Этот пример, таким образом, есть частный случай теоремы 12 Ловаса.)

Пример 4. Разные реберно-транзитивные графы. Здесь мы применим границу (4.1) к трем очень интересным графам. В каждом случае существует только один класс эквивалентности

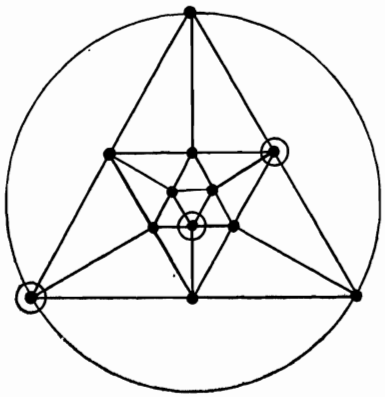


Рис. 3. Граф икосаэдра.

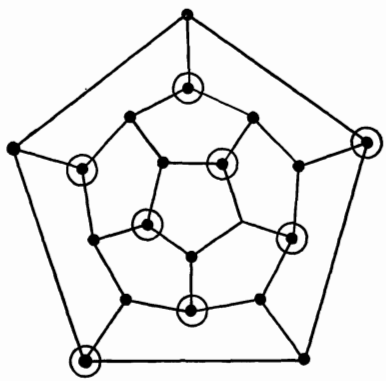


Рис. 4. Граф додекаэдра.

ребер, так что полученные границы равны $\theta_L(G)$. В каждом случае $\theta_L(G)$ строго меньше, чем любая граница, которую можно получить методами Шеннона.

Граф Петерсена. Это регулярный граф с $N = 10$, $r = 3$ (рис. 2).

Минимальное собственное значение здесь оказывается равным -2, и в силу (4.1) имеет место $\theta(G) \leqslant 4$. С другой сто-

роны, $\alpha(G)=4$ (см. четыре обведенные кружками вершины на рис. 2) и следовательно, $\theta(G)=4$. (Этот результат является частным случаем теоремы 13 Ловаса.)

Граф икосаэдра. Этот граф имеет $N=12$, $r=5$; его вершины и ребра образованы из вершин и ребер правильного икосаэдра (рис. 3).

Здесь минимальное собственное значение равно $-\sqrt{5}$. Поэтому из (4.1) следует, что $\theta(G)\leqslant 3(\sqrt{5}-1)=3{,}7802$. С другой стороны $\alpha(G)=3$, так что имеет место [1]) $3\leqslant\theta(G)\leqslant 3{,}7082$.

Граф додекаэдра. Это граф правильного додекаэдра с $N=20$, $r=3$ (рис. 4).

Здесь также $\lambda_{\min}=-\sqrt{5}$. Поэтому из (4.1) следует, что $\theta(G)\leqslant 15\sqrt{5}-25=8{,}5410$. С другой стороны, как показано на рисунке, $\alpha(G)=8$. Таким образом, $8\leqslant\theta(G)\leqslant 8{,}5410$.

Пример 5. Специальный граф на 7 вершинах. Рассмотрим граф, изображенный на рис. 5.

Это регулярный граф с $N=7$, $r=4$. Минимальное собственное значение равно $4\cos\frac{6\pi}{7}\cos\frac{2\pi}{7}=-2{,}2470$, и, следовательно, в силу (4.1) $\theta(G)\leqslant 2{,}5178$. Однако, под действием симметрий графа, его ребра разбиваются на два класса эквивалентности: ребра вида $\{i, i+1\}$ и ребра вида $\{i, i+2\}$ (сложение по модулю 7). В этом случае граница из (4.1) строго больше, чем граница Ловаса $\theta_L(G)$, и для вычисления $\theta_L(G)$ необходимо непосредственно воспользоваться теоремой 5.

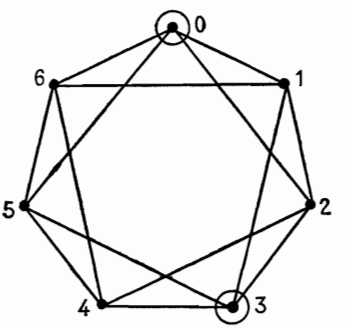


Рис. 5. Регулярный граф на семи вершинах.

Пусть A_0 — единичная 7×7 матрица, A_1 — матрица смежности для ребер типа $\{i, i+1\}$, A_2 — для ребер типа $\{i, i+2\}$ и A_3 — для ребер типа $\{i, i+3\}$ (которые не являются ребрами G).

Легко проверить, что матрицы $\{A_0, A_1, A_2, A_3\}$ удовлетворяют условиям (3.2)—(3.4). Для $C=\{1, 2\}$ граф G_C является графом

[1]) Д. Хикерсон показал, что $\alpha(G^2)\geqslant 12$, из чего вытекает, что $\theta(G)\geqslant\sqrt{12}=3{,}4641$. Его доказательство состоит в следующем. Пусть H будет графом с тем же множеством вершин V, что и G, но в котором вершины v и v' смежные, если и только если расстояние между ними в графе G равно 2. Заметим, что H изоморфен G. Пусть ϕ — перестановка V, отображающая G на H, и пусть $S=\{(v, \phi(v))\colon v\in V\}$. Тогда S — независимое множество в G^2, и поэтому $\alpha(G^2)\geqslant 12$.

на рис. 5. Семь векторов вида $x(\zeta)=(1, \zeta, \ldots, \zeta^6)$, где ζ — комплексный корень седьмой степени из 1, образуют множество общих собственных векторов матриц A:

$$A_0x(\zeta)=x(\zeta),$$
$$A_1x(\zeta)=(\zeta+\zeta^{-1})x(\zeta),$$
$$A_2x(\zeta)=(\zeta^2+\zeta^{-2})x(\zeta),$$
$$A_3x(\zeta)=(\zeta^3+\zeta^{-3})x(\zeta).$$

Таким образом, согласно теореме 5, если μ_1 и μ_2 удовлетворяют условию $1+\mu_1(\zeta+\zeta^{-1})+\mu_2(\zeta^2+\zeta^{-2})\geqslant 0$ для всех корней седьмой степени из 1, то $\theta(G)\leqslant 7/(1+2\mu_1+2\mu_2)$. Для получения наилучшей такой границы необходимо максимизировать функцию $\mu_1+\mu_2$ при условии, что указанные выше неравенства выполнены. Это легко делается вручную[1]), и наибольшее возможное значение достигается при $\mu_1=0{,}8020$, $\mu_2=0{,}3569$. В результате получаем границу $\theta(G)\leqslant 2{,}1098$. (Этот граф является дополнительным к циклическому графу C_7, мы запишем это как $G=C_7'$.) Теорема 8 Ловаса утверждает, что если G — произвольный граф с группой симметрий, транзитивно действующей на вершинах, то $\theta_L(G)\cdot\theta_L(G')=N$. Таким образом, из примера 2 следует, что $\theta_L(C_N')=1+\left(\cos\frac{\pi}{N}\right)^{-1}$ при нечетных N. Мы включили этот альтернативный вывод только для того, чтобы дать нетривиальный пример нашего подхода с позиций линейного программирования.

Пример 6. Границы «линейного программирования» Дельсарта. В [3] Дельсарт получил границы «линейного программирования» для клик в ассоциативных схемах. Здесь мы дадим набросок доказательства того, что эти границы являются частным случаем нашей теоремы 4. (Недавно Схрейвер [7] получил аналогичные результаты.)

Пусть V — конечное множество и $R=\{R_0, R_1, \ldots, R_n\}$ — семейство из $n+1$ подмножеств его декартова квадрата V^2. Множество R_j является отношением на V и может быть описано своей матрицей смежности

$$A_j(v, v')=\begin{cases}1, & \text{если } (v, v')\in R_j,\\ 0 & \text{в противном случае.}\end{cases}$$

Пара (V, R) называется (симметричной) *ассоциативной схемой*, если выполнены следующие условия:

[1]) В действительности необходимо рассматривать только три неравенства, а именно неравенства с $\zeta=\exp(2\pi ik/7)$, $k=1, 2, 3$.

A1: R является разбиением V^2 и R_0 — диагональ, т. е.

$$R_0 = \{(v, v): v \in V\}.$$

A2: Отношения $\{R_j\}$ симметричны, т. е. из $(v, v') \in R_j$ следует $(v', v) \in R_j$.

A3: Существуют числа $p_{i,j}^{(k)} = p_{j,i}^{(k)}$, такие, что для всех $i, j = 0, 1, \ldots, n$

$$A_i A_j = \sum_{k=0}^{n} p_{i,j}^{(k)} A_k.$$

Если M — подмножество множества $\{0, 1, \ldots, n\}$, причем $0 \in M$, то непустое подмножество $Y \subseteq V$ называется *M-кликой относительно R*, если

$$R_j \cap Y^2 = \phi \quad \text{для всех} \quad j \notin M.$$

Дельсарт [3] получил верхнюю границу для числа точек в M-клике, которая является решением определенной задачи линейного программирования. Но мы в равной мере можем применить теорему 4 к той же проблеме, поскольку для матрицы $\{A_j\}$ ассоциативной схемы условия (3.2)—(3.4), конечно, выполняются (заметим, что так как $p_{i,j}^{(k)} = p_{j,i}^{(k)}$, то условие (A.3) значительно сильнее, чем (3.4)). Если мы положим $C = \{j: j \notin M\}$, то определенная выше M-клика является независимым множеством в G_C, и поэтому верхняя граница теоремы 4 является верхней границей мощности любой M-клики. На самом деле можно показать, что эта граница совпадает с границей Дельсарта. Следовательно, теорема 4 является более общей, чем граница Дельсарта, так как она применима ко многим объектам, не являющимся ассоциативными схемами.

Приложение А. Эквивалентность теоремы 3 и границы Ловаса

Для данного ортонормированного представления графа G определим матрицу A следующим образом:

$$A(v, v') = v \cdot v'.$$

Ясно, что $A \in \Omega_0(G)$. Если b — вектор длины 1 и $\lambda = \min\{(v \cdot b)^2: v \in V\}$, то $A - \lambda J$ является п. п. о. матрицей [1]). Это так, потому

[1]) Матрица $A - \lambda J$ в случае произвольного ортонормированного представления, вообще говоря, не является п. п. о. матрицей. Например, для двумерного представления графа $C_4\{v_1 = (1, 0), v_2 = (-1, 0), v_3 = (0, 1), v_4 = (0, -1)\}$ и вектора $b = (1/\sqrt{2}, 1/\sqrt{2})$ величина λ равна $1/2$, но матрица $A - 1/2J$ не является п. п. о. матрицей, так как $\sum (a_{vv'} - 1/2)\, x(v)\, x(v') = -8 < 0$ для $x = (1, 1, 1, 1)$. Однако матрица $A - \lambda J$ будет п. п. о. матрицей, если ограничиться «равноугольными» представлениями, т. е. представ-

что $A - \lambda J = B + C$, где матрицы B и C определены следующим образом:

$$B(v, v') = (v - (v \cdot b)\, b) \cdot (v' - (v' \cdot b)\, b),$$
$$C(v, v') = (v \cdot b)(v' \cdot b) - \lambda.$$

Матрица B является п. п. о. матрицей, так как она является матрицей скалярных произведений некоторого множества векторов. Матрица C также п. п. о. матрица, так как для любого множества действительных чисел $\{x(v)\colon v \in V\}$

$$x^T C x = \Big(\sum_v x(v)(v \cdot b)\Big)^2 - \lambda \Big(\sum x(v)\Big)^2 \geqslant 0,$$

поскольку $(v \cdot b) \geqslant \sqrt{\lambda}$ для всех v. Таким образом, матрица $A - \lambda J$, будучи суммой двух п. п. о. матриц, также является п. п. о. матрицей.

Так как $A \in \Omega_0(G)$ и $A - \lambda J$ — п. п. о. матрицы, то из (2.8) следует, что $\lambda(A) \geqslant \lambda$ и, следовательно, по теореме 3 $\theta(G) \leqslant \leqslant \lambda^{-1} = (\min\{(v \cdot b)^2\colon v \in V\})^{-1}$. Таким образом, из теоремы 3 вытекает граница Ловаса (1.5).

Обратно, для $A \in \Omega_0(G)$ положим $\lambda = \lambda(A)$. Тогда матрица $A - \lambda J$ является п. п. о. матрицей, и по известной теореме ([2], гл. 9) существует матрица B, такая, что $B^T B = A - \lambda J$. Обозначив через $\{\mathbf{w}(v)\colon v \in B\}$ вектор-столбцы матрицы B, имеем

$$\mathbf{w}(v) \cdot \mathbf{w}(v') = A(v, v') - \lambda = \begin{cases} 1 - \lambda, & \text{если } v = v', \\ -\lambda, & \text{если } \{v, v'\} \notin G. \end{cases}$$

Пусть теперь $\mathbf{t}$ — вектор, ортогональный ко всем векторам $\mathbf{w}(v)$, причем $|\mathbf{t}|^2 = \lambda$ (если необходимо, увеличиваем размерность основного пространства). Положим

$$\mathbf{x}(v) = \mathbf{w}(v) + \mathbf{t}.$$

Длина векторов $\mathbf{x}(v)$ равна 1, так как

$$\mathbf{x}(v) \cdot \mathbf{x}(v') = A(v, v').$$

лениями, для которых $(b, v_i) = \sqrt{\lambda}$ при всех i. С другой стороны, для произвольного ортонормированного представления $(v_1, \ldots, v_n)$ и единичного вектора b существует «равноугольное» представление с тем же λ (достаточно увеличить размерность представления, добавив ортонормированные векторы $f_1, \ldots, f_n$, и рассмотреть новое представление

$$v_i' = \frac{\sqrt{\lambda}}{(b, v_i)}\, v_i + \sqrt{1 - \frac{\lambda}{(b, v_i)^2}}\, f_i\Big).$$

Поэтому граница Ловаса (1.5) действительно следует из границы теоремы 3. — *Прим. перев.*

Граф ортогональности, определяемый векторами x, является, таким образом, подграфом (то же множество вершин, но подмножество ребер) H' графа G.

Более того, определив вектор $\mathbf{b}$ как $\mathbf{b}=\frac{\mathbf{t}}{|\mathbf{t}|}$, имеем $\mathbf{x}(v)\times\mathbf{b}=|\mathbf{t}|=\sqrt{\lambda}$ для всех $v\in V$. Следовательно, по границе Ловаса (1.5) $\theta(H')\leqslant\lambda^{-1}=\lambda(A)^{-1}$. Но ясно, что $\theta(G)\leqslant\theta(H')$, и поэтому из результатов Ловаса следует теорема 3.

Приложение B. Доказательство утверждений (2.4)—(2.9)

Напомним, что A является действительной симметрической $N\times N$-матрицей и что

$$\lambda(A)=\inf\{x^TAx:\ x\cdot u=1\}, \tag{B.1}$$

где $u=(1, 1, \ldots, 1)$ — вектор из одних единиц. По теореме о главных осях ([2], гл. 9 и 10) существует множество $\{\xi_1, \ldots, \xi_N\}$ из N ортонормированных собственных векторов матрицы A:

$$\begin{aligned}\xi_i\cdot\xi_j&=1, \quad \text{если} \quad i=j,\\ &=0, \quad \text{если} \quad i\neq j;\end{aligned} \tag{B.2}$$

$$A\xi_j=\lambda_j\xi_j, \quad j=1, 2, \ldots, N. \tag{B.3}$$

Таким образом, для $x=x_1\xi_1+\ldots+x_N\xi_N$ и $y=y_1\xi_1+\ldots+y_N\xi_N$ имеет место

$$x\cdot y=\sum_{j=1}^{N}x_jy_j, \tag{B.4}$$

$$x^TAy=\sum_{j=1}^{N}\lambda_jx_jy_j. \tag{B.5}$$

Если одно из собственных значений λ_j отрицательно, то можно следующим образом построить вектор x, удовлетворяющий условиям $x\cdot u=1$ и $x^TAx<0$. Пусть $v=v_1\xi_1+\ldots+v_N\xi_N$ — фиксированный вектор, такой, что $v\cdot u=1$. Для любого действительного β определим вектор

$$x=\frac{\beta\xi_j+v}{\beta u_j+1}, \tag{B.6}$$

где $u=u_1\xi_1+\ldots+u_N\xi_N$ — разложение вектора u. Ясно, что $x\cdot u=1$, а из (B.5) вычисляем

$$x^TAx=\frac{1}{(\beta u_j+1)^2}\left\{\lambda_j\beta^2+2\beta\lambda_jv_j+\sum_{j=1}^{N}\lambda_jv_j^2\right\}.$$

Очевидно, что это выражение будет отрицательно, если β достаточно велико, так как выражение в скобках доминируется членом $-|\lambda_j|\beta^2$. Утверждение (2.4) доказано.

Таким образом, мы предполагаем, что матрица A является п. п. о. матрицей, т. е. все собственные значения $\{\lambda_j\}$ неотрицательны. Если для некоторого j имеем $\lambda_j = 0$, но $u_j \neq 0$, то, положив $x = u_j^{-1}\xi_j$, получим $x \cdot u = 1$, $x^T Ax = 0$. Этим утверждение (2.5) доказано.

С другой стороны, если $u_j = 0$ всякий раз, когда $\lambda_j = 0$, то, воспользовавшись неравенством Шварца, получим

$$\left(\sum x_j u_j\right)^2 \leqslant \left(\sum \lambda_j x_j^2\right)\left(\sum \lambda_j^{-1} u_j^2\right), \tag{B.7}$$

где суммирование ведется только по индексам, для которых $\lambda_j > 0$. Так как, согласно (B.4) и (B.5), $\sum x_j u_j = x \cdot u$ и $\sum \lambda_j x_j^2 = x^T Ax$, то из (B.7) сразу же следует, что если $x \cdot u = 1$, то $x^T Ax \geqslant \left(\sum \lambda_j^{-1} u_j^2\right)^{-1} = \lambda$. С другой стороны, выбирая $x_i = u_i \lambda_i^{-1} \lambda$ для всех i, получаем $x \cdot u = 1$ и $x^T Ax = \lambda$. Это доказывает (2.6).

Если u является собственным вектором матрицы A с собственным значением σ, то в разложении $u = u_1\xi_1 + \ldots + u_N\xi_N$ коэффициент u_j должен равняться нулю, за исключением $\lambda_j = \sigma$. Следовательно, $\lambda(A) = \left(\sigma^{-1}\sum u_j^2\right)^{-1} = \sigma/N$, так как $\sum u_j^2 = u \cdot u = N$. Это доказывает (2.7).

Чтобы доказать (2.8), заметим, что

$$x^T(A - \lambda J)x = x^T Ax - \lambda(x \cdot u)^2 = \sum \lambda_j x_j^2 - \lambda\left(\sum x_j u_j\right)^2. \tag{B.8}$$

Сравнивая это с (B.7), видим, что это выражение будет неотрицательно для всех x тогда и только тогда, когда $\lambda \leqslant \left(\sum \lambda_j^{-1} u_j^2\right)^{-1}$, т. е. $\lambda \leqslant \lambda(A)$. Это доказывает (2.8).

Наконец, обратимся к (2.9). Предположим, что $\xi_1, \ldots, \xi_N$ — главные оси матрицы A с соответствующими собственными значениями $\{\lambda_j\}$ и что $\eta_1, \ldots, \eta_M$ — главные оси матрицы B с собственными значениями $\{\mu_k\}$. Предположим далее, что $u^{(A)} = u_1^{(A)}\xi_1 + \ldots + u_N^{(A)}\xi_N$, $u^{(B)} = u_1^{(B)}\eta_1 + \ldots + u_M^{(B)}\eta_M$ есть разложение векторов из одних единиц по этим двум базисам.

Из известных результатов (см. [6], гл. VII) следует, что MN векторов $\xi_j \times \eta_k$ являются главными осями матрицы $A \times B$ с соответствующими собственными значениями $\lambda_j\mu_k$. Более того, очевидно, что разложение MN-мерного вектора из одних единиц по базису $\{\xi_j \times \eta_k\}$ имеет вид

$$u = \sum_{j,k} u_j^{(A)} u_k^{(B)} (\xi_j \times \eta_k).$$

Таким образом, согласно (2.6),

$$\lambda(A \times B) = \left(\sum_{j,k} (\lambda_j \mu_k)^{-1} \left[u_j^{(A)} u_k^{(B)}\right]^2\right)^{-1} = \lambda(A)\lambda(B),$$

что доказывает (2.9).

ЛИТЕРАТУРА

1. Baumert L. D. et al. A combinatorial packing problem, pp. 97—108, in Computers in Algebra and Number Theory, SIAM-AMS Proceedings, Vol. IV, American Mathematical Society, Providence, 1971.
2. Birkhoff G., MacLane S. A. Survey of Modern Algebra, MacMillan, New York, 1960.
3. Delsarte P. An algebraic approach to the association schemes of coding theory, Philips Research Reports Supplement, No. 10, 1973. [Имеется перевод: Дельсарт Ф. Алгебраический подход к схемам отношений теории кодирования. — М.: Мир, 1976.]
4. Hoffman K., Kunze R. Linear Algebra, Prentice-Hall, Inc., Englewood Cliffs, 1961.
5. Lovász L. On the Shannon capacity of a graph, IEEE Trans. Inform. Theory, T-25, No. 1, 1979, 1—7. [Имеется перевод: Ловас Л. О шенноновской емкости графа. — См. наст. сб.]
6. MacDuffee C. The Theory of Matrices, Chelsca, New York, 1956.
7. Schrijver A. A comparison of the bounds of Delsarte and Lovász, IEEE Trans. Inform. Theory, IT-25, No. 4, 1979, 425—429. [Имеется перевод: Схрейвер А. Сравнение границ Дельсарта и Ловаса. — См. наст. сб.]
8. Shannon C. The zero error capacity of a noisy channel, IRE Trans. Inform. Theory, IT-2, 8—19, 1956. [Имеется перевод: Шеннон К. Пропускная способность канала с шумом при нулевой вероятности ошибки. — В кн.: Работы по теории информации и кибернетике. — М.: ИЛ, 1963, стр. 464—487.]
9. Weyl H. Algebraic Theory of Numbers (Annals of Math, Studies No. 1), Princeton University Press, Princeton, 1940. [Имеется перевод: Вейль Г. Алгебраическая теория чисел. — М.: ИЛ, 1947.]
10. Wielandt H. Finite Permutation Groups, Academic Press, New York, 1964.

Последовательности комплексных чисел с низкой периодической корреляционной функцией[1])

В. Оллтоп

В работе указаны три семейства последовательностей комплексных чисел, имеющих почти минимально возможную величину корреляции, причем эти последовательности не содержат нулевых элементов. Для некоторых пар последовательностей периода $N = (p-1)/2$, p — простое число, все непиковые корреляционные коэффициенты имеют величину, близкую к $(2N)^{-1/2}$.

I. ВВЕДЕНИЕ

Периодические последовательности используются при построении систем сигналов для радаров с широким спектром и для систем связи. Главной целью при построении таких последовательностей является минимизация величины боковых лепестков автокорреляционной и взаимно-корреляционной функций. В настоящей работе рассматривается только случай периодических корреляционных функций.

Велч [1] показал, что для семейства M различных комплекснозначных последовательностей, каждая из которых имеет период N и единичную норму на периоде, величина

$$B(M, N) \triangleq \left(\frac{M-1}{MN-1}\right)^{1/2} \tag{1}$$

является нижней границей максимального непикового значения корреляционной функции. Это означает, что хотя бы один из $MN-M$ автокорреляционных боковых лепестков или один из $(M^2-M)N$ взаимно-корреляционных коэффициентов имеет величину, не меньшую, чем $B(M, N)$. Таким образом, величина $(2N)^{-1/2}$ служит нижней границей для пары последовательностей, тогда как в случае, когда число последовательностей M растет вместе с N, нижняя граница равна примерно $N^{-1/2}$.

Используя характеры группы единиц кольца Z_N вычетов по модулю N, Шольц и Велч [2] построили семейство последовательностей периода N, которые обладают желательными корре-

[1]) Alltop W. O. Complex Sequences with Low Periodic Correlations, IEEE Trans. Inform. Theory, vol. 26; No. 3; pp. 350—354, May 1980.

© 1980 IEEE. Reprinted, with permission, from IEEE Transactions on Information Theory, vol. IT-26, No. 3, pp. 350—354, May 1980
© Перевод на русский язык, «Мир», 1983

ляционными свойствами. Каждая из этих последовательностей содержит $\varphi(N)$ ненулевых элементов на периоде, где $\varphi(N)$ — число единиц в Z_N; $\varphi(N)$ чаще называют φ-функцией Эйлера [3]. При простом N эти последовательности имеют автокорреляционные боковые лепестки, равные по модулю $1/(N-1)$ и взаимно-корреляционные коэффициенты, равные либо 0, либо $N^{1/2}/(N-1)$. При составном N автокорреляционные боковые лепестки могут достигать $\varphi(d)^{-1}$ для некоторого делителя d числа N. (Следует заметить, что последовательности в [2] имеют норму $\varphi(N)$, а не 1, вследствие чего величины корреляции приобретают другой вид.) Голд [4] построил семейства двоичных последовательностей периода $N = 2^m - 1$, которые имеют максимальную величину корреляции, близкую к $2N^{-1/2}$ и $2^{1/2}N^{-1/2}$ при $m \equiv 2 \pmod 4$ и нечетном m соответственно.

В разд. III и IV приведены три типа семейств последовательностей, которые почти достигают границы (1). Все элементы этих последовательностей равны по модулю $N^{-1/2}$, тогда как непиковые значения корреляционной функции по модулю не превышают $N^{-1/2}$. Семейства, приведенные в разд. III, содержат $p-1$ или p последовательностей, где p — наименьший простой делитель числа N. Семейства, приведенные в разд. IV, содержат M последовательностей, где $MN+1$ — простое число. Пары (M, N), порождающие эти семейства, связаны с одним частным типом циклических разностных множеств в аддитивной группе Z_p — кольца вычетов по модулю p, $p = MN + 1$. Непиковые корреляционные коэффициенты этих последовательностей равны по модулю примерно $cN^{-1/2}$, где $c = ((M-1)/M)^{1/2}$, $M = 2, 4, 8$. Чакрабарти и Томлинсон [5] использовали простые числа p вида $MN+1$ для построения последовательностей комплексных чисел периода p, а не N. Некоторые сдвиги этих последовательностей дают последовательности с малой апериодической корреляцией. Гауссовы суммы типа той, которая вычислена в лемме 1, рассматриваются в [6].

II. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ

Пусть $\mathscr{H}$ означает семейство $\{h_\lambda: 1 \leqslant \lambda \leqslant M\}$ последовательностей комплексных чисел периода N. k-й элемент h_λ есть $h_\lambda(k)$, а один период h_λ есть $(h_\lambda(0), \ldots, h_\lambda(N-1))$. Z_N обозначает кольцо вычетов по модулю N, так что Z_p изоморфно конечному полю $\mathrm{GF}(p)$, когда p — простое число. Корреляционный коэффициент с номером τ между h_λ и h_μ определяется следующим образом:

$$H_{\lambda\mu}(\tau) \overset{\Delta}{=} \sum_{k=0}^{N-1} h_\lambda(k+\tau)\, h_\mu(k)^*,$$

где $h_\mu(k)^*$ — число, комплексно сопряженное с $h_\mu(k)$. Аргументы $k+\tau$ и k в h_λ и h_μ должны быть приведены по модулю N к одному из чисел 0, 1, ..., $N-1$. Если рассматривать бесконечные в обе стороны последовательности с периодом N, то приведение по модулю N необязательно. Предполагается, что все последовательности имеют единичную норму на периоде, т. е.

$$\| h_\lambda \|^2 = \sum_{k=0}^{N-1} | h_\lambda(k) |^2 = 1.$$

Отсюда вытекает, что все корреляционные пики $H_{\lambda\lambda}(0)$ равны единице. Мерой величины корреляции, представляющей наибольший интерес, является

$$\max(\mathscr{K}) = \max \{ | H_{\lambda\mu}(\tau) | : \lambda \neq \mu \text{ или } \tau \neq 0 \}.$$

Таким образом, $\max(\mathscr{K})$ есть максимум модуля всех $M^2N - M$ непиковых корреляционных коэффициентов. Граница Велча (1) показывает, что

$$\max(\mathscr{K}) \geqslant \left(\frac{M-1}{MN-1} \right)^{1/2}.$$

Каждое h_λ будет представлять собой последовательность корней из единицы, умноженных на $N^{-1/2}$, т. е.

$$h_\lambda(k) = N^{-1/2} \omega_n^{f(\lambda, k)},$$

где $\omega_n = \exp(2\pi i/n)$, а f — функция из $Z_M \times Z_N$ в Z_n. Для последовательностей разд. III $n = N$, тогда как в разд. IV предполагается $n = MN + 1 = p$.

Например, положим $M = N = n \geqslant 2$ и пусть $f(\lambda, k) = \lambda k$. Период h_λ есть

$$N^{-1/2} (\omega_N^0, \omega_N^\lambda, \omega_N^{2\lambda}, \ldots, \omega_N^{(N-1)\lambda})$$

и сопряжен с λ-й строкой матрицы $N \times N$ дискретного преобразования Фурье. В этом случае

$$| H_{\lambda\mu}(\tau) | = \begin{cases} 0, & \text{если } \lambda \neq \mu, \\ 1, & \text{если } \lambda = \mu. \end{cases}$$

Все взаимно-корреляционные коэффициенты равны нулю, тогда как автокорреляционные коэффициенты равны по модулю единице и, следовательно, $\max(\mathscr{K}) = 1$.

III. ПОСЛЕДОВАТЕЛЬНОСТИ С КВАДРАТИЧНОЙ И КУБИЧЕСКОЙ ФАЗОЙ

Для нечетного целого числа N ($N > 2$) λ-я последовательность с квадратичной фазой определяется как

$$a_\lambda(k) \stackrel{\Delta}{=} N^{-1/2} \omega_N^{\lambda k^2},$$

а λ-я последовательность с кубической фазой как

$$b_\lambda(k) \stackrel{\Delta}{=} N^{-1/2}\omega_N^{k^3+\lambda k}.$$

Последовательности с квадратичной фазой аналогичны последовательностям Чу [7].

Пример 1. Для $N = 5$

$$a_1 = 5^{-1/2}(1, \omega_5, \omega_5^4, \omega_5^4, \omega_5),$$
$$a_2 = 5^{-1/12}(1, \omega_5^2, \omega_5^3, \omega_5^3, \omega_5^2),$$
$$a_3 = 5^{-1/12}(1, \omega_5^3, \omega_5^2, \omega_5^2, \omega_5^3),$$
$$a_4 = 5^{-1/12}(1, \omega_5^4, \omega_5, \omega_5, \omega_5^4),$$
$$a_5 = 5^{-1/2}(1, 1, 1, 1, 1),$$
$$b_1 = 5^{-1/2}(1, \omega_5^2, 1, 1, \omega_5^3),$$
$$b_2 = 5^{-1/2}(1, \omega_5^3, \omega_5^2, \omega_5^3, \omega_5^2),$$
$$b_3 = 5^{-1/2}(1, \omega_5^4, \omega_5^4, \omega_5, \omega_5),$$
$$b_4 = 5^{-1/2}(1, 1, \omega_5, \omega_5^4, 1),$$
$$b_5 = 5^{-1/2}(1, \omega_5, \omega_5^3, \omega_5^2, \omega_5^4).$$

$\max(\mathscr{A}_5) = \max(\mathscr{B}_5) = 5^{-1/2}$, где

$$\mathscr{A}_5 = \{a_1, a_2, a_3, a_4\},$$
$$\mathscr{B}_5 = \{b_1, b_2, b_3, b_4, b_5\}.$$

Последовательность a_5 не используется из-за своей постоянной автокорреляции; в действительности период a_5 равен единице, а не пяти. Все боковые лепестки $A_{\lambda\lambda}(\tau)$, $\tau \neq 0$, семейства $\mathscr{A}_5$ равны нулю, а все взаимно-корреляционные коэффициенты $A_{\lambda\mu}(\tau)$, $\lambda \neq \mu$, этого семейства равны по модулю $5^{-1/2}$. Обозначив через $B_{\lambda\mu}(\tau)$ корреляцию между b_λ и b_μ, получаем

$$|B_{\lambda\mu}(\tau)| = \begin{cases} 1, & \text{если } \lambda = \mu, \quad \tau = 0, \\ 0, & \text{если } \lambda \neq \mu, \quad \tau = 0, \\ 5^{-1/2} & \text{в противном случае.} \end{cases}$$

Семейство $\mathscr{A}_5$ обобщается на все нечетные целые числа, тогда как $\mathscr{B}_5$ обобщается на все простые $p \geqslant 5$, в результате чего получаются семейства последовательностей, которые достигают границы $N^{-1/2}$ для величины корреляции. Для нечетного $N \geqslant 3$

определим $\mathscr{A}_N$ следующим образом:

$$\mathscr{A}_N \overset{\Delta}{=} \{a_1, \ldots, a_{p-1}\},$$

где p есть наименьший простой делитель N.

Теорема 1. *Для нечетного $N \geqslant 3$ $\mathscr{A}_N$ есть семейство, состоящее из $p-1$ последовательностей с квадратичной фазой, причем $\max(\mathscr{A}_N) = N^{-1/2}$, где p — наименьший простой делитель N. Точнее, корреляционные коэффициенты $A_{\lambda\mu}(\tau)$, $0 \leqslant \tau \leqslant N-1$, $1 \leqslant \lambda, \mu \leqslant p-1$, удовлетворяют следующему соотношению:*

$$|A_{\lambda\mu}(\tau)| = \begin{cases} 1 & \text{при } \lambda = \mu, \quad \tau = 0, \\ 0 & \text{при } \lambda = \mu, \quad \tau \neq 0, \\ N^{-1/2} & \text{в остальных случаях.} \end{cases}$$

Для нечетного простого p положим

$$\mathscr{B}_p \overset{\Delta}{=} \{b_1, b_2, \ldots, b_p\}.$$

Теорема 2. *Для каждого простого числа $p \geqslant 5$ $\mathscr{B}_p$ есть семейство из p последовательностей с кубической фазой, причем $\max(\mathscr{B}_p) = p^{-1/2}$. Точнее, корреляционные коэффициенты $B_{\lambda\mu}(\tau)$, $0 \leqslant \tau \leqslant p-1$, $1 \leqslant \lambda, \mu \leqslant p$, удовлетворяют следующему соотношению:*

$$|B_{\lambda\mu}(\tau)| = \begin{cases} 1 & \text{при } \lambda = \mu, \ \tau = 0, \\ 0 & \text{при } \lambda \neq \mu, \ \tau = 0, \\ p^{-1/2} & \text{в остальных случаях.} \end{cases}$$

Теоремы 1 и 2 доказываются с помощью преобразования и вычисления некоторых тригонометрических сумм. Используя определение $A_{\lambda\mu}(\tau)$, a_λ и a_μ, получаем

$$A_{\lambda\mu}(\tau) = \sum_{k=0}^{N-1} a_\lambda(k+\tau)\, a_\mu(k)^* =$$

$$= \frac{1}{N} \sum_{k=0}^{N-1} \omega_N^{\lambda(k+\tau)^2} \omega_N^{-\mu k^2} =$$

$$= \frac{1}{N}\, \omega_N^{\lambda\tau^2} \sum_{k=0}^{N-1} \omega_N^{(\lambda-\mu)k^2 + 2\lambda\tau k}.$$

Аналогично

$$B_{\lambda\mu}(\tau) = \frac{1}{N}\, \omega_N^{\tau^3+\lambda\tau} \sum_{k=0}^{N-1} \omega_N^{3\tau k^2 + (\lambda-\mu+3\tau^2)k}.$$

Отсюда вытекает, что

$$|A_{\lambda\mu}(\tau)| = \frac{1}{N}|Q_N(\lambda - \mu, 2\lambda\tau)|, \tag{2}$$

$$|B_{\lambda\mu}(\tau)| = \frac{1}{N}|Q_N(3\tau, \lambda - \mu + 3\tau^2)|, \tag{3}$$

где

$$Q_N(\eta, \sigma) \overset{\Delta}{=} \sum_{k=0}^{N-1} \omega_N^{\eta k^2 + \sigma k}. \tag{4}$$

Следовательно, модули корреляционных коэффициентов могут быть определены с помощью величин $|Q_N(\eta, \sigma)|$.

Лемма 1. *Предположим, что N — нечетное число и $\delta = (\eta, N)$ — наибольший общий делитель η и N. Тогда*

$$|Q_N(\eta, \sigma)|^2 = \begin{cases} N\delta, & \text{если } \delta \text{ делит } \sigma, \\ 0, & \text{если } \delta \text{ не делит } \sigma. \end{cases} \tag{5}$$

Доказательство. По определению

$$|Q_N(\eta, \sigma)|^2 = Q_N(\eta, \sigma)(Q_N(\eta, \sigma))^* =$$

$$= \left(\sum_{k=0}^{N-1} \omega_N^{\eta k^2 + \sigma k}\right)\left(\sum_{r=0}^{N-1} \omega_N^{-\eta r^2 - \sigma r}\right) = \sum_{k,\, r} \omega_N^{\eta(k^2 - r^2) + \sigma(k - r)},$$

где пара (k, r) пробегает множество $Z_M \times Z_N$. Тогда пара (k, s), где $s = k - r$, также пробегает все множество $Z_N \times Z_N$. Используя эту замену в знаке суммирования, получаем

$$|Q_N(\eta, \sigma)|^2 = \sum_{(k,\, r)} \omega_N^{(k-r)(\eta(k+r)+\sigma)} =$$

$$= \sum_{k,\, s} \omega_N^{s(\eta(2k-s)+\sigma)} = \sum_{s=0}^{N-1} \omega_N^{-\eta s^2 + \sigma s} \sum_{k=0}^{N-1} \omega_N^{2\eta s k}.$$

Так как ω_N — примитивный корень N-й степени из единицы, то

$$\sum_{k=0}^{N-1} \omega_N^{2\eta s k} = \begin{cases} N, & \text{если } 2\eta s \equiv 0 \pmod N, \\ 0 & \text{в противном случае.} \end{cases}$$

Следовательно,

$$|Q_N(\eta, \sigma)|^2 = N \sum_{s} \omega_N^{-\eta s^2 + \sigma s},$$

где суммирование распространено на те значения s, для которых $2\eta s \equiv 0 \pmod N$ и $0 \leqslant s \leqslant N - 1$. Используем теперь тот факт, что N — нечетное число. Так как $(2, N) = 1$, то сравнение $2\eta s \equiv 0 \pmod N$ эквивалентно сравнению $\eta s \equiv 0 \pmod N$. Таким образом,

$$\omega_N^{-\eta s^2 + \sigma s} = \omega_N^{\sigma s}, \quad \text{если} \quad 2\eta s \equiv 0 \pmod N.$$

Отсюда следует, что

$$|Q_N(\eta, \sigma)|^2 = N \sum_s \omega_N^{\sigma s},$$

где суммирование проводится по тем значениям s, для которых $\eta s \equiv 0 (\mathrm{mod}\, N)$, $0 \leqslant s \leqslant N-1$. Множество тех s в Z_N, для которых $\eta s \equiv 0 (\mathrm{mod}\, N)$, есть $\{0, m, 2m, \ldots, (\delta-1)m\}$, где $\delta = (\eta, N)$, $m = N/\delta$. Следовательно,

$$|Q_N(\eta, \sigma)| = N\left[1 + \omega_N^{\sigma m} + \omega_N^{2\sigma m} + \ldots + \omega_N^{(\delta-1)\sigma m}\right] = \begin{cases} N\delta, & \text{если } \sigma m \equiv 0 (\mathrm{mod}\, N), \\ 0, & \text{если } \sigma m \not\equiv 0 (\mathrm{mod}\, N). \end{cases}$$

Поскольку $\sigma m \equiv 0 (\mathrm{mod}\, N)$ тогда и только тогда, когда δ делит σ, то лемма доказана.

Доказательство теорем 1 и 2. Для a_λ и a_μ из $\mathscr{A}_N$ λ и μ являются взаимно простыми с N. Более того, $\lambda - \mu$ также взаимно просто с N, за исключением случая $\lambda = \mu$. Следовательно, значения $|A_{\lambda\mu}(\tau)|$, указанные в теореме 1, получаются в результате применения соотношения (5) к равенству (2).

При $\tau \neq 0$ число 3τ взаимно просто с p, $p \geqslant 5$. Следовательно, $|B_{\lambda\mu}(\tau)|^2 = N^{-1}$, если только $\tau \neq 0$. Из (3) и (5) следует, что $B_{\lambda\mu}(0) = 0$, если $\lambda \neq \mu$, что и доказывает теорему 2.

Возможность увеличения семейства $\mathscr{A}_N$ или использования последовательностей с кубической фазой при составных N заслуживает дальнейшего обсуждения. Как и ранее, обозначим через p наименьший простой делитель N. Если $(\lambda, N) = d$, то $|A_{\lambda\lambda}(m)| = 1$, где $m = N/d$. Следовательно, λ должно быть взаимно просто с N для каждой последовательности a_λ из $\mathscr{A}_N$. Если $\mathscr{A}_N$ содержит p различных последовательностей a_λ, причем каждое λ взаимно просто с N, то в нем должны быть две последовательности a_λ и a_μ, для которых $\lambda \equiv \mu (\mathrm{mod}\, p)$. В этом случае $|A_{\lambda\mu}(d)| = (d/N)^{1/2}$, где $d = (\lambda - \mu, N) \geqslant p$. Следовательно, для того чтобы $\max(\mathscr{A}_N)$ не превосходили $N^{-1/2}$, семейство $\mathscr{A}_N$ должно содержать не более $p-1$ последовательностей a_λ, для которых все λ и $\lambda - \mu$ взаимно просты с N. Конечно, такие семейства из $p-1$ последовательностей существуют, причем отличные от $\{a_1, a_2, \ldots, a_{p-1}\}$. Например, $\max(\mathscr{A}) = 35^{-1/2}$, когда $\mathscr{A} = \{a_4, a_8, a_{12}, a_{16}\}$, $N = 35$.

Для любой последовательности b_λ с кубической фазой имеет место $|B_{\lambda\lambda}(p)| = (p/N)^{1/2}$, когда N делится на p и не делится на 3. Таким образом, автокорреляция превосходит $N^{-1/2}$ для всех b_λ. Аналогично ведут себя последовательности характеров [2].

IV. ПОСЛЕДОВАТЕЛЬНОСТИ СТЕПЕННЫХ ВЫЧЕТОВ

Пусть γ обозначает примитивный корень в конечном поле Z_p, где $p = MN + 1$. Положим $\beta = \gamma^M$, так что β и γ есть примитивные корни из единицы в Z_p степени N и $p - 1$ соответственно. N различных степеней $\beta \pmod p$ образуют подгруппу в Z_p^* — мультипликативной группе ненулевых элементов поля Z_p. Обозначив эту подгруппу через $\mathscr{C}_N$, получаем разбиение Z_p^* на M смежных классов:

$$Z_p^* = \mathscr{C}_N \cup \gamma\mathscr{C}_N \cup \gamma^2\mathscr{C}_N \cup \ \dots \ \cup \gamma^{M-1}\mathscr{C}_N$$

с представителями $1, \gamma, \gamma^2, \dots, \gamma^{M-1}$.

Определим λ-ю последовательность (M, N) степенных вычетов следующим образом:

$$y_\lambda(k) = N^{-1/2}\omega_p^{\gamma^\lambda \beta^k}.$$

Пример 2. Для $p = 19$, $M = 3$, $N = 6$, $\gamma = 2$ получаются три различные последовательности (3.6) степенных вычетов:

$$y_0 = 6^{-1/2}(\omega_{19}, \omega_{19}^8, \omega_{19}^7, \omega_{19}^{18}, \omega_{19}^{11}, \omega_{19}^{12}),$$
$$y_1 = 6^{-1/2}(\omega_{19}^2, \omega_{19}^{16}, \omega_{19}^{14}, \omega_{19}^{17}, \omega_{19}^3, \omega_{19}^5),$$
$$y_2 = 6^{-1/2}(\omega_{19}^4, \omega_{19}^{13}, \omega_{19}^9, \omega_{19}^{15}, \omega_{19}^6, \omega_{19}^{10}).$$

Для $\mathscr{Y}_{3,6} = \{y_0, y_1, y_2\}$ имеет место $\max(\mathscr{Y}_{3,6}) = 0{,}418$. В действительности все 51 значение непиковых коэффициентов корреляции $Y_{\lambda\mu}(\tau)$ заключено среди

$$s_\lambda = 6^{-1/2}\sum_{k=0}^{5} y_\lambda(k), \quad 0 \leqslant \lambda \leqslant 2,$$
$$s_0 = -0{,}204, \quad s_1 = 0{,}418, \quad s_2 = -0{,}381.$$

В общем случае пусть $\mathscr{Y}_{M,N} = \{y_0, y_1, \dots, y_{M-1}\}$ и

$$Y_{\lambda\mu}(\tau) \stackrel{\Delta}{=} \sum_{k=0}^{N-1} y_\lambda(k + \tau)\, y_\mu(k)^*$$

для $0 \leqslant \lambda, \mu \leqslant M - 1$ и $0 \leqslant \tau \leqslant N - 1$, где $p = MN + 1$ — простое число.

Лемма 2. *Все непиковые коэффициенты корреляции семейства $\mathscr{Y}_{M,N}$ степенных вычетов принимают одно из M значений $s_0, s_1, \dots, s_{M-1}$, где*

$$s_\lambda \stackrel{\Delta}{=} N^{-1/2}\sum_{k=0}^{N-1} y_\lambda(k).$$

Точнее,

$$Y_{\lambda\mu}(\tau)=\begin{cases}1, & \text{если } \lambda=\mu \text{ и } \tau=0,\\ s_\eta & \text{в противном случае,}\end{cases}$$

где η *— такой вычет по модулю* M, $0\leqslant\eta\leqslant M-1$, *что* $\gamma^{-\eta}(\gamma^\lambda\beta^\tau-\gamma^\mu)=\beta^r$ *для некоторого* r, *т. е.* γ^n *и* $\gamma^\lambda\beta^\tau-\gamma^\mu$ *лежат в одном смежном классе группы* Z_p^* *по подгруппе* $\mathscr{C}_N$, *порожденной* $\beta=\gamma^M$.

Доказательство. По определению $Y_{\lambda\mu}(\tau)$, y_λ и y_μ имеем

$$Y_{\lambda\mu}(\tau)=\sum_{k=0}^{N-1}y_\lambda(k+\tau)\,y_\lambda(k)^*=$$

$$=\frac{1}{N}\sum_{k=0}^{N-1}\omega_p^{\gamma^\lambda\beta^k+\tau}\omega_p^{-\gamma^\mu\beta^k}=\frac{1}{N}\sum_{k=0}^{N-1}\omega_p^{\alpha\beta^k},$$

где $\alpha=\gamma^\lambda\beta^\tau-\gamma^\mu$. Показатель в сумме пробегает весь класс $\alpha\mathscr{C}_N$. Если $\alpha\not\equiv 0\pmod p$, то $\alpha\mathscr{C}_N=\gamma^\eta\mathscr{C}_N$ для единственного η, удовлетворяющего условию $0\leqslant\eta\leqslant M-1$. Следовательно,

$$Y_{\lambda\mu}(\tau)=\frac{1}{N}\sum_{k=0}^{N-1}\omega_p^{\gamma^\eta\beta^k}=N^{-1/2}\sum_{k=0}^{N-1}y_\eta(k)=s_\eta,$$

если только $\alpha\in Z_p^*$. Остается рассмотреть только случай $\alpha\equiv 0\pmod p$. Здесь $\gamma^\lambda\beta^\tau\equiv\gamma^\mu$ и $\beta^\tau=\gamma^{\mu-\lambda}$. Так как γ^0, γ^1, ..., γ^{M-1} суть представители различных смежных классов по $\mathscr{C}_N$, то $\mu=\lambda$ и $\beta^\tau\equiv 1\pmod p$. Таким образом, при $\alpha\equiv 0\pmod p$ рассматриваемая корреляция совпадает с пиковым значением, что и доказывает лемму.

$\max(\mathscr{Y}_{M,N})$ есть наибольшее из следующих M чисел $|s_0|$, $|s_1|$, ..., $|s_{M-1}|$. Из определения s_λ и $y_\lambda(k)$ вытекает, что

$$|s_\lambda|^2=\frac{1}{N}\left(\sum_{k=0}^{N-1}y_\lambda(k)\right)\left(\sum_{r=0}^{N-1}y_\lambda(r)\right)^*=\frac{1}{N}\sum_{k,\,r}y_\lambda(k)\,(y_\lambda(r))^*=$$

$$=N^{-2}\sum_{k,\,r}\omega_p^{\gamma^\lambda\beta^k}\omega_p^{-\gamma^\lambda\beta^r}=N^{-2}\sum_{k,\,r}\theta_\lambda^{\beta^k-\beta^r},$$

где $\theta_\lambda=\exp(2\pi j\gamma^\lambda/p)$, а пары (k,y) пробегают все N^2 пар в $Z_N\times Z_N$. Последнее эквивалентно тому, что

$$|s_\lambda|=N^{-2}\sum\{\theta_\lambda^{\mu-\zeta}\colon(\mu,\zeta)\in\mathscr{C}_N\times\mathscr{C}_N\},\qquad(6)$$

т. е. показатель $\beta^k-\beta^r$ пробегает все N^2 разностей $\mu-\zeta$, $\mu\in\mathscr{C}_N$, $\zeta\in\mathscr{C}_N$, когда (k,r) пробегает $Z_N\times Z_N$. В исключитель-

ных случаях, когда все N^2-N ненулевых разностей равномерно распределены по $p-1$ классам вычетов в Z_p^*, все $|s_\lambda|$ имеют одинаковую величину и $\max(\mathscr{Y}_{M,N})$ почти оптимален. Это происходит, когда $\mathscr{C}_N$ является разностным множеством в аддитивной группе Z_p.

Циклическое разностное множество (V, K, Λ) есть подмножество $\mathscr{D}$ циклической группы Z_V, удовлетворяющее следующим условиям: $\mathscr{D}$ содержит K элементов, каждое ненулевое z из Z_V допускает ровно Λ представлений $z=\mu-\zeta$ парами (μ, ζ) из $\mathscr{D}\times\mathscr{D}$.

Пример 3. Множество $\mathscr{D}=\{1, 2, 4\}$ есть $(7, 3, 1)$-циклическое разностное множество в Z_7

$$1=2-1, \quad 4=1-4,$$
$$2=4-2, \quad 5=2-4,$$
$$3=4-1, \quad 6=1-2.$$

Пример 4. Множество $\mathscr{D}=\{0, 1, 3, 9\}$ есть $(13, 4, 1)$-циклическое разностное множество в Z_{13}.

Основы теории циклических разностных множеств можно найти в [8] и [9].

Элементарный подсчет показывает, что необходимым условием существования циклического (V, K, Λ)-разностного множества является выполнение равенства

$$K^2-K=\Lambda(V-1). \tag{7}$$

Дальнейшие необходимые условия на параметры V, K, Λ могут быть установлены с помощью более глубоких теоретико-числовых и алгебраических методов [8], [9].

Основной интерес представляют здесь циклические (p, N, Λ)-разностные множества $\mathscr{D}$ при простом $p=MN+1$ и $\mathscr{D}=\mathscr{C}_N$, являющемся мультипликативной группой вычетов M-й степени в Z_p^*. $(7, 3, 1)$-разностное множество $\{1, 3, 4\}$ примера 3 есть $\mathscr{C}_3$, т. е. множество ненулевых квадратичных вычетов в Z_7^*. Для соответствующей пары последовательностей $(2, 3)$ степенных вычетов

$$s_0=(-1+j\sqrt{7})/6, \quad s_1=(-1-j\sqrt{7})/6,$$
$$\max(\mathscr{Y}_{2,3})=|s_0|=|s_1|=\sqrt{2}/3.$$

Множество $\{0, 1, 3, 9\}$ примера 4 есть $\mathscr{C}_3$, к которому добавлено число 0. Для $\gamma=2$ в Z_{13} параметры $(4, 3)$-последовательностей

степенных вычетов есть

$$s_0 = s_0^* = 0.217 + 0.174j,$$
$$s_1 = s_3^* = 0.384 + 0.575j,$$
$$|s_0| = |s_2| = 0.278 = ((5 - \sqrt{13})/18)^{1/2},$$
$$\max(\mathscr{Y}_{4,3}) = |s_1| = |s_3| = 0.691 = ((5 + \sqrt{13})/18)^{1/2}.$$

Величины s_λ, $0 \leqslant \lambda \leqslant 3$, не все равны, так как $\mathscr{C}_3$ не является циклическим разностным множеством в Z_{13}.

Предположим, что $\mathscr{C}_N$ есть циклическое (p, N, Λ)-разностное множество в Z_p, $p = MN + 1$. N^2 показателей в (6) содержат 0 точно N раз и каждый ненулевой элемент в Z_p точно Λ раз. Таким образом, (6) принимает вид

$$|s_\lambda|^2 = N^{-2}\left(N + \Lambda \sum_{k=1}^{p-1} \theta_\lambda^k\right) = N^2(N - \Lambda), \tag{8}$$

так как θ_λ есть комплексный (примитивный) корень p-й степени из единицы. Подставляя параметры (p, N, Λ) вместо (V, K, Λ) в (7), получаем

$$N^2 - N = \Lambda(p - 1) = \Lambda MN$$

и

$$\Lambda = (N - 1)/M. \tag{9}$$

Из (8) и (9) следует, что

$$|s_\lambda| = N^{-1/2}\left(1 - \frac{1}{M} + \frac{1}{MN}\right)^{1/2}.$$

Тем самым доказана следующая теорема.

Теорема 3. *Предположим, что $p = MN + 1$ есть простое число и множество $\mathscr{C}_N$ ненулевых вычетов M-й степени образует циклическое (p, N, Λ)-разностное множество в аддитивной группе Z_p. Тогда множество последовательностей $\mathscr{Y}_{M,N}$ (M, N)-степенных вычетов состоит из M различных последовательностей y_λ периода N, для которых*

$$|Y_{\lambda\mu}(\tau)| = \begin{cases} 1, & \text{если } \lambda = \mu \text{ и } \tau = 0, \\ N^{-1/2}\left(1 - \frac{1}{M} + \frac{1}{MN}\right) & \text{в противном случае.} \end{cases} \tag{10}$$

Известен только один бесконечный класс пар (M, N), к которому теорема 3 может быть применена. Это класс

$$\mathscr{P} = \{(2, N): N \text{ — нечетное, } 2N + 1 \text{ — простое}\}.$$

Каждый представитель $\mathscr{P}$ дает пару $\mathscr{Y}_{2,N}$ последовательностей степенных вычетов периода N, для которых

$$\max(\mathscr{Y}_{2,N}) = N^{-1/2}\left(\frac{1}{2} + \frac{1}{2N}\right)^{1/2} \approx (2N)^{-1/2}.$$

Первый нетривиальный случай возникает из (7,3,1)-циклического разностного множества примера 3:

$$y_0 = 3^{-1/2}(\omega_7, \omega_7^2, \omega_7^4),$$
$$y_1 = 3^{-1/2}(\omega_7^3, \omega_7^6, \omega_7^5).$$

Разностные множества, ассоциированные с $\mathscr{P}$, являются множествами Пэли — Адамара квадратичных вычетов в конечном поле Z_p, $p \equiv 3 \pmod 4$.

Другие известные разностные множества степенных вычетов, к которым может быть применена теорема 3, существуют для $M = 4$ или 8. Если $p = 4n^2 + 1$, n — нечетное, то $\mathscr{C}_N$ — циклическое разностное множество в Z_p при $N = n^2$. Каждое из таких множеств производит семейство из четырех последовательностей периода N. Подходящими значениями для N, меньшими 10 000, являются 3^2, 5^2, 7^2, 13^2, 27^2, 33^2, 37^2, 45^2, 47^2, 55^2, 63^2, 65^2, 67^2, 73^2, 75^2, 85^2. При $M = 8$ известны лишь случаи, когда $p = 8n^2 + 1 = 64m^2 + 9$, где n и m — нечетные числа. Первые два таких простых числа есть 73 и 140 411 704 393 (см. [8], стр. 124).

Сравнение (1) и (10) показывает, что $\max(\mathscr{Y}_{M,N})$ близок к минимуму всякий раз, когда $\mathscr{C}_N$ — разностное множество, в частности

$$\max(\mathscr{Y}_{M,N}) < \left[1 + \frac{1}{MN(M-1)}\right] B(M, N).$$

Отметим, что $\max(\mathscr{Y}_{M,N})$ может быть меньше $N^{-1/2}$ и в случае, когда $\mathscr{C}_N$ не является разностным множеством. Если $p \equiv 1 \pmod 4$, то $\mathscr{C}_N$ не является разностным множеством при $N = (p-1)/2$, но в этом случае

$$s_0 = (-1 + \sqrt{p})/(p-1), \quad s_1 = (-1 - \sqrt{p})/(p-1),$$
$$\max(\mathscr{Y}_{2,N}) = |s_1| = (1 + \sqrt{p})/(p-1) < \left(\frac{2}{p-1}\right)^{1/2} = N^{-1/2},$$

если только $p \geqslant 13$, $p \equiv 1 \pmod 4$. Другими примерами с $\mathscr{C}_N$, не являющимся разностным множеством, и $M > 2$ являются $\max(\mathscr{Y}_{4,27}) = 0.1878$; $\max(\mathscr{Y}_{4,69}) = 0.1201$; $\max(\mathscr{Y}_{4,87}) = 0.0994$; $\max(\mathscr{Y}_{4,93}) = 0.1016$; $\max(\mathscr{Y}_{4,127}) = 0.0814$.

V. РЕЗЮМЕ

В работе рассматривались три типа семейств комплексных периодических последовательностей, обладающих желательными корреляционными свойствами: последовательности с квадратичной фазой, с кубической фазой и степенных вычетов. В каждой нормированной последовательности периода N все элементы равны по абсолютной величине $N^{-1/2}$.

Последовательности с квадратичной и кубической фазой почти достигают границы Велча (1). Для каждого нечетного N

существует семейство из $(p-1)$ последовательностей с квадратичной фазой, которые имеют максимальную непиковую корреляцию, равную $N^{-1/2}$, где p — наименьший делитель N. Более того, все автокорреляционные лепестки каждой последовательности с квадратичной фазой равны нулю. Последовательности с кубической фазой имеют ненулевые автокорреляционные лепестки и достигают границы $N^{-1/2}$, только когда N простое.

Семейство из M последовательностей степенных вычетов периода N было построено при $MN+1=p$, где p — простое число. Эти последовательности порождаются мультипликативной группой $\mathscr{C}_N$ ненулевых M-х степеней в конечном поле Z_p. Когда $\mathscr{C}_N$ — циклическое разностное множество в аддитивной группе Z_p, получающееся семейство последовательностей степенных вычетов почти достигает границы Велча. Для каждого простого числа $p \equiv 3 \pmod 4$ группа $\mathscr{C}_N$, являющаяся разностным множеством Пэли — Адамара, порождает пару последовательностей периода N с максимальной величиной непиковой корреляции, лишь несколько превосходящей $(2N)^{-1/2}$, $N=(p-1)/2$. Для больших $p \equiv 1 \pmod 4$ справедлив аналогичный результат.

В этой работе основное внимание было сконцентрировано на изучении достижения границы Велча небольшими (т. е. $M \leqslant N$) семействами периодических последовательностей. Обобщение предложенной здесь конструкции приводит к другим интересным семействам последовательностей. Например, к семействам последовательностей степенных вычетов, имеющим величину корреляции, близкую к $N^{-1/2}$, для которых $\mathscr{C}_N$ не является циклическим множеством. Последовательности с квадратичной и кубической фазой определяются некоторыми полиномами второй и третьей степени соответственно. Полиномы более высокой степени будут порождать бо́льшие семейства последовательностей. Например, p^2 полиномов четвертой степени $k^4+\lambda k^2+\mu k$, $0 \leqslant \lambda, \mu \leqslant p-1$, определяют p^2 циклически отличных друг от друга последовательностей периода p. Расчеты на ЭВМ показали, что при $p \equiv 2 \pmod 3$ эти последовательности имеют величину корреляции, не превосходящую $2p^{-1/2}$.

Благодарности

Автор благодарен Брюсу Мак-Клангу за его помощь при проведении тестов на ЭВМ для последовательностей с биквадратичной фазой, отмеченных в разд. V.

Добавление

Недавние результаты Сарвэйта [10] появились после подготовки настоящей работы к публикации. Наша теорема 1 по

существу является теоремой 3 работы [10]. Для нечетного N последовательность $u^{(2\lambda)}$ Франка — Цадова — Чу из [10] отличается от последовательности a_λ с квадратичной фазой только нормирующим множителем $N^{1/2}$.

ЛИТЕРАТУРА

1. Welch L. R. Lower bounds on the maximum cross correlation of signals. IEEE Trans. Inform. Theory, 1974, IT-20, 397—399.
2. Scholtz R. A., Welch L. R. Group characters: Sequences with good correlation properties. IEEE Trans. Inform. Theory 1978, IT-24, 537—545.
3. Le Veque W. J. Topics in Number Theory. Addison-Wesley, 1956.
4. Gold R. Optimal binary sequences for spread spectrum multiplexing. IEEE Trans. Inform. Theory, 1967, IT-13, 619—621.
5. Chakrabarti N. B., Tomlinson M. Design of sequences with specified autocorrelation and cross correlation. IEEE Trans. Commun., 1976, COM-24, 1246—1252.
6. Landau E. Elementary Number Theory, New York, 1958.
7. Chu D. C. Polyphase codes with good periodic correlation properties. IEEE Trans. Inform. Theory, 1972, IT-18, 531—532.
8. Baumert L. D. Cyclic Difference Sets. Lecture Notes in Mathematics, No. 182. Berlin, Heidelberg, New York, Springer-Verlag, 1971.
9. Hall M., Jr. Combinatorial Theory. Blaisdell, 1967.
10. Sarwate D. V. Bounds on crosscorrelation and autocorrelation of sequences. IEEE Trans. Inform. Theory, 1979, IT-25, 720—724.
11*. В. И. Левенштейн. Границы максимальной мощности кода с ограниченным модулем скалярного произведения, ДАН СССР, 1982, 263, № 6, 1303—1308.
12*. Sarwate D. V., Pursley M. B. Cross-correlation properties of pseudo-random and related sequences. Proc. IEEE, 1980, v. 68, № 5, p. 593—619.
13*. Hardy G. H., Littlewood J. E. The trigonometrical series associated with elliptic θ-functions. Acta Math., 1914, v. 37, p. 193—238.
14*. Potter H. S. A. On diophantine approximation and generalized elliptic theta function. Quart. J. Math. Oxford Ser., 1938, p. 161—175.
15*. Fiedler H., Jurkat W., Körner O. Asymptotoc expansions of finite theta series. Acta Arithm., 1977, v. 32, p. 129—140.

ДОБАВЛЕНИЕ ПЕРЕВОДЧИКА

Приведем некоторые дополнительные свойства построенных в статье последовательностей с кубической фазой.

Рассмотрим множество E_q^n векторов $x=(x_1, \ldots, x_n)$, где

$$x_j \doteq \frac{1}{\sqrt{n}} \exp\left(2\pi i \frac{a_j}{q}\right), \quad a_j = 0, \ldots, q-1; \quad j = 1, \ldots, n.$$

Обозначим через $M(E_q^n, \rho)$ максимальную мощность множества $W \subseteq E_q^n$, для которого

$$\max_{x,\, y \in W;\, x \neq y} \left| \sum_{j=1}^{n} x_j y_j^* \right| \leqslant \rho.$$

[1]) Литература, отмеченная звездочкой, добавлена переводчиком.

В работе [11] было отмечено, что на множестве, состоящем из p последовательности $b_\lambda(k)$, $\lambda = 1, \ldots, p$, с кубической фазой и их циклических сдвигов, достигается граница

$$M\left(E_q^n, \frac{1}{\sqrt{n}}\right) \leqslant n^2$$

при $n = q = p \geqslant 5$.

Отсюда, в частности, следует, что построенное в статье множество из p последовательностей с кубической фазой имеет максимально возможную мощность в классе множеств последовательностей, для которых максимум модуля всех непиковых корреляционных коэффициентов не превышает $p^{-1/2}$.

Отметим также, что апериодическая функция взаимной корреляции каждой последовательности с кубической фазой имеет порядок $O(p^{-1/2})$, что на логарифмический множитель лучше, чем для всех известных ранее последовательностей (см., например, [12]). В самом деле для неполной суммы Гаусса известна оценка

$$\max_{1 \leqslant N \leqslant p-1} \left| \sum_{k=0}^{N} \omega_p^{Ak^2+Bk} \right| \leqslant cp^{1/2}, \quad 1 \leqslant A \leqslant p-1, \quad 0 \leqslant B \leqslant p-1,$$

где c некоторая абсолютная постоянная (см. [13] или более доступные работы [14—15]). Отсюда вытекает, что при $1 \leqslant \tau \leqslant p-1$

$$\left| \sum_{k=0}^{p-1-\tau} b_\lambda(k+\tau)\, b_\lambda^*(k) \right| = p^{-1} \left| \sum_{k=0}^{p-1-\tau} \omega_p^{3\tau k^2+3\tau^2 k} \right| \leqslant cp^{-1/2}.$$

О глубине формул[1])

Э. Шамир, М. Снир

Institute of Mathematics, Hebrew University of Jerusalem and Computer Science Department, Edinburgh University

Задача минимизации глубины формул с помощью эквивалентных преобразований формализуется здесь в общеалгебраическом плане. Для одной специфической алгебраической системы Σ_0 разработаны в духе динамического программирования специальные методы получения нижних оценок глубины. Из этих нижних оценок для Σ_0 автоматически вытекают точно такие же нижние оценки для двух систем: (i) класса арифметических выражений, содержащих только сложение и умножение, и (ii) класса выражений над конечными языками, содержащих только операции объединения и конкатенации. Получены следующие нижние оценки глубины: (i) для перманента — $2n - o(n)$, (ii) для симметрических многочленов — $(0{,}25 + o(1))\log^2 n$, (iii) для формул специального вида сложности n — $1{,}16 \log n$.

1. ВВЕДЕНИЕ

Нахождение нижних оценок сложности вычисления конкретных функций и предикатов — это одно из наиболее интересных направлений теории вычисления, в котором многие гипотезы и задачи все еще противостоят продолжающимся попыткам найти решение. Нехватка методов и слабость результатов бросается в глаза, даже когда класс рассматриваемых вычислений ограничен, например в случае монотонных булевых вычислений.

В связи с этим важно разрабатывать методы получения нижних оценок и для других ограниченных классов. Здесь рассматриваются «монотонные» вычисления, использующие только «сложение» и «умножение», в пределах системы аксиом Σ_0. Система Σ_0 сильнее системы ΣR — обычной арифметики над положительными числами с операциями сложения и умножения, а также сильнее (т. е. имеет больше тождеств) системы ΣL — алгебры над словами с операциями объединения и конкатенации. Поэтому нижние оценки, полученные для функций в системе Σ_0, остаются справедливыми для функций того же вида в системах ΣR и ΣL.

В системе Σ_0 можно получать экспоненциальные нижние оценки для числа операций, как это было сделано авторами в

[1]) Shamir E. and Snir M. On the depth complexity of formulas. Mathematical Systems Theory, 13, No. 4, 301—322, 1980.

© 1980 by Springer-Verlag
© Перевод на русский язык, «Мир», 1983

[9, 17] (см. также [4] для случая ΣL). В настоящей статье мы интересуемся глубиной (т. е. временной задержкой при параллельных вычислениях), которая лучше поддается точной оценке с помощью более тонких методов. При монотонных вычислениях (т. е. в системе Σ_0) окончательная формула накладывает сильные ограничения на промежуточные вычисления (на возможные подформулы). Устанавливая подходящие ограничения и рассматривая процесс построения формулы как процесс выбора операндов и операций ($+$ или $*$), мы можем свести поиск оптимальной по глубине формулы к одной из задач динамического программирования, решение которой дает оптимальную глубину, а иногда и оптимальные формулы.

Приводится несколько применений метода. В системе Σ_0 даются весьма точные нижние оценки глубины для перманента, симметрических функций и некоторых других специальных функций, которые естественным образом возникают в арифметических и комбинаторных вычислениях.

Хотя монотонные вычисления и имеют некоторые преимущества, например такое, как абсолютная устойчивость численного решения [5, 6], они обычно непрактичны для функций, которые можно значительно эффективнее вычислить, используя вычитание (или отрицание в булевом случае). Тем не менее вычисления в пределах ограниченной монотонной системы, такой, как Σ_0, оказываются весьма практичными при перестроении (или упрощении) выражений *общего вида*, которое проводится с целью минимизации глубины (или другой характеристики формулы). В действительности, по-видимому, нереально внедрение систем, которые вводят дополнительные члены с целью их поглощения, так как это расширяет область, в которой ведется поиск оптимальной формулы, до необъятных размеров.

Доказывать любые нижние оценки, в том числе NP — трудность проблемы упрощения, обычно труднее в системах, содержащих больше тождеств. В нашей ограниченной системе Σ_0 нам удалось доказать [11], что проблема упрощения (в отношении числа умножений или сложности неветвящегося алгоритма или глубины) является NP-трудной.

В разд. 9 мы применяем нашу технику к другой проблеме упрощения выражений в системе Σ_0, а именно мы получаем нижнюю оценку для перехода от оптимальной сложности формулы к оптимальной глубине. Имеется несколько статей, посвященных соответствующим верхним оценкам [1, 7], а некоторые нижние оценки в булевых системах с сильными ограничениями можно найти в [2, 3]. Нижние оценки для соотношения между сложностью и глубиной, затрагивающие члены низших порядков, в монотонных и немонотонных булевых системах были

недавно получены в [21, 22] (см. последний абзац в разд. 9).

Мы хотели бы поблагодарить рецензентов за их ценные замечания, поправки и рекомендации.

2. КОНГРУЭНТНОСТЬ ФОРМУЛ

Цель разд. 2 и разд. 3 дать четкое и строгое определение формул, отношения конгруэнтности и семантики формул (которая, как обычно, дается посредством интерпретаций). Однако читатель может положиться на свои собственные формальные или даже интуитивные представления об этих понятиях и перейти прямо к разд. 4.

Пусть $\Omega = \{0, 1\}^*$ — множество конечных слов из 0 и 1. Обозначим через Λ пустое слово, а через rs конкатенацию слов r и s. Множество Ω является частично упорядоченным с отношением порядка $\prec$, которое определяется следующим образом: $r \prec s$ тогда и только тогда, когда существует такое слово t, что $rt = s$. Если не выполняется ни соотношение $r \prec s$, ни соотношение $s \prec r$, то слова r и s называются *независимыми*. Множество $T \subseteq \Omega$ называется *двоичным деревом* (для краткости деревом), если $ri \in T \Rightarrow r \in T$ и $r0 \in T$ тогда и только тогда, когда $r1 \in T$. Назовем элементы множества T *вершинами*. Вершина $r \in T$ называется *концевой*, если $r0 \notin T$, и *внутренней*, если $r0 \in T$. *Длиной* дерева T (обозначение $l(T)$) называется число его концевых вершин. Известно, что $|T| = 2l(T) - 1$. *Глубиной* дерева T (обозначение $d(T)$) называется максимум длин слов из T.

Пусть C — некоторое множество *констант*, X — бесконечное множество *переменных*, $+$ (сложение) и $*$ (умножение) — символы двух двуместных операций. *Формулой* (над $\langle C, X, +, * \rangle$) называется функция F: $\Omega \to C \cup X \cup \{+, *\}$, у которой областью определения (обозначение $\mathrm{Dom}(F)$) является дерево, такая, что $Fr \in \{+, *\}$, если r — внутренняя вершина, и $Fr \in C \cup X$, если r — концевая вершина. Глубина и длина формулы это соответственно глубина и длина ее области определения. Пусть F_1 и F_2 — формулы, $\theta \in \{+, *\}$. Тогда

$$(F_1 \theta F_2) = \lambda, \quad \text{где} \quad \lambda r = \begin{cases} \theta, & \text{если } r = \Lambda, \\ F_i s, & \text{если } r = is. \end{cases}$$

Кроме того, когда это не вызывает путаницы, мы будем обозначать через α, где $\alpha \in C \cup X$, формулу λ, которая равна α, если $r = \Lambda$. *Подформула* формулы F, соответствующая слову r, определяется равенством $F/r = \lambda$, где $\lambda s = F(rs)$. Множество подформул формулы F обозначим через $Sb(F)$. Пусть F и G — формулы; формула, получающаяся из F подстановкой G вместо

подформулы, соответствующей слову r, имеет следующий вид:

$$F(r \leftarrow G) = \lambda, \quad \text{где} \quad \lambda s = \begin{cases} Gt, & \text{если } s = rt, \\ Fs & \text{в остальных случаях.} \end{cases}$$

Это определение можно распространить на множество взаимно независимых вершин следующим образом:

$$F(r \leftarrow G_r\colon\ r \in \phi) = F,$$
$$F(r \leftarrow G_r\colon\ r \in R \cup \{s\}) = F(r \leftarrow G_r\colon\ r \in R)(s \leftarrow G_s).$$

Пусть F, G и K — формулы; формула, получающаяся из F заменой G на K, имеет следующий вид:

$$F(G \leftarrow K) = F(r \leftarrow K\colon\ r \in \{s\colon\ F/s = G\}).$$

Пусть $\mathcal{G}$ — семейство взаимно независимых формул (ни одна из них не является подформулой другой). Тогда по определению

$$F(G \leftarrow K_G\colon\ G \in \mathcal{G}) = F(r \leftarrow K_{F/r}\colon\ r \in \{s\colon\ F/s \in \mathcal{G}\}).$$

Отношение $\equiv$ между формулами называется *конгруэнтностью*, если оно является отношением эквивалентности и из соотношений $F_1 \equiv F_2$, $G_1 \equiv G_2$ следует, что для любого $x \in X$ имеет место $F_1(x \leftarrow G_1) \equiv F_2(x \leftarrow G_2)$. Мы будем интересоваться главным образом конгруэнтностями, порожденными применением последовательности базисных правил преобразования. Каждое базисное правило заменяет одну подформулу другой. Формально *система преобразований* состоит из множества Σ пар формул (называемых *схемами*). Подстановка любых формул вместо переменных, выполненная одновременно в паре формул из схемы, дает *аксиому*. Таким образом, $\langle F_1', F_2' \rangle$ является аксиомой тогда и только тогда, когда $F_i' = F_i(x \leftarrow G_x\colon\ x \in X)$, $i = 1, 2$, а $\langle F_1, F_2 \rangle$ является схемой.

Будем писать $F \to F'$, если $F/u = G$, $\langle G, G' \rangle$ является аксиомой и $F' = F(u \leftarrow G')$. Обозначим через $\xrightarrow{*}$ рефлексивное и транзитивное замыкание отношения $\to$, тогда отношение $\xrightarrow{*}$ будет минимальным по включению отношением конгруэнтности, порожденным системой Σ. Это отношение будем обозначать через $\equiv_\Sigma$.

Упрощение формулы F относительно отношения конгруэнтности $\equiv_\Sigma$ это поиск либо более простой формулы $G \equiv_\Sigma F$, либо самой простой формулы при некотором критерии оптимальности. В этой статье таким критерием будет глубина. Обозначим через $d_\Sigma(F)$ минимальную глубину такой формулы G, что $G \equiv_\Sigma F$.

3. ИСТИННЫЕ И ПОЛНЫЕ ИНТЕРПРЕТАЦИИ

Интепретации формул над $\langle C, X, +, * \rangle$ введем обычным образом. Пусть D — множество с двумя двуместными операциями: сложением $(+)$ и умножением $(*)$. Функция ν: $C \to D$ называется *интерпретацией* совокупности $\langle C, X, +, * \rangle$ с областью определения D. С каждой формулой F ассоциируется функция F_ν: $D^X \to D$, определяемая индуктивно:

$$c_\nu = \nu(c) \quad \text{при} \quad c \in C,$$
$$x_\nu = x \quad \text{при} \quad x \in X,$$
$$(F + G)_\nu = F_\nu + G_\nu,$$
$$(F * G)_\nu = F_\nu * G_\nu.$$

Система Σ называется *истинной* для интерпретации ν, если $F \equiv_\Sigma G \Rightarrow F_\nu = G_\nu$; система Σ называется *полной* для интерпретации ν, если $F_\nu = G_\nu \Leftrightarrow F \equiv_\Sigma G$.

Если система Σ полна, то $d_\Sigma(F)$ это в точности минимальная глубина формулы, реализующей функцию F_ν, которая совпадает с минимальной задержкой параллельного вычисления функции F_ν при условии, что используются только сложение и умножение и каждая операция выполняется за единицу времени, причем нет конфликтных ситуаций в памяти и нет программных задержек.

Мы приведем две интерпретации с полными системами.

А. Обозначим через R^+ множество неотрицательных действительных чисел с обычными сложением и умножением. Тогда ΣR будет состоять из следующих схем (здесь x, y, z — переменные, 0 и 1 — константы):

1. $(x + y) + z \equiv x + (y + z)$
2. $x + y \equiv y + x$
3. $(x * y) * z \equiv x * (y * z)$
4. $x * y \equiv y * x$
5. $x * (y + z) \equiv x * y + x * z$
6. $(x + y) * z \equiv x * z + y * z$
7. $0 + x \equiv x$
8. $0 * x \equiv 0$
9. $x * 0 \equiv 0$
10. $x * 1 \equiv x$
11. $1 * x \equiv x$
12. При любых a, $b \in C$ и $\theta \in \{+, *\}$ имеет место $a \theta b \equiv c$.

Обозначим через ν_R интерпретацию с областью определения R^+, согласующуюся со всеми схемами вида 12 (т. е. если $a\theta b \equiv c$, то $\nu(a)\theta\nu(b) = \nu(c)$).

Лемма 3.1. *Система* ΣR *полна для интерпретации* ν_R.

В. Пусть Ξ — конечный алфавит. Обозначим через L множество конечных языков над алфавитом Ξ с операцией объединения в роли сложения и операцией конкатенации в роли умножения. Положим $C = \Xi \cup \{0, 1\}$. Тогда ΣL будет состоять из схем 1—3, 5—11 и схемы

13. $x + x \equiv x$.

Обозначим через ν_L интерпретацию, при которой каждая буква $\xi \in \Xi$ означает язык $\{\xi\}$, состоящий из одного слова, символ 0 означает пустой язык $\varnothing$, а символ 1 — язык $\{\Lambda\}$, состоящий из одного пустого слова Λ.

Лемма 3.2. *Система* ΣL *полна для интерпретации* ν_L.

Обе леммы устанавливаются путем доказательства того, что каждый класс конгруэнтности содержит единственную «каноническую» формулу (ср. с теоремой 4.1) и что неэквивалентные канонические формулы реализуют различные функции.

4. СИСТЕМА Σ_0

Начиная с этого момента, сосредоточим наше внимание на системе Σ_0, состоящей из следующих схем («аксиом») (x, y, z — переменные, 0 и 1 — константы):

1. $(x + y) + z \equiv x + (y + z)$
2. $x + y \equiv y + x$
3. $(x * y) * z \equiv x * (y * z)$
4. $x * y \equiv y * x$
5. $x * (y + z) \equiv x * y + x * z$
6. $x + x \equiv x$
7. $x + 0 \equiv x$
8. $x * 0 \equiv 0$
9. $x * 1 \equiv x$

10. Для любой константы $a \neq 0$, имеет место равенство $a \equiv 1$.

Будем обозначать через $\equiv$ (без индекса Σ_0) конгруэнтность, определяемую системой Σ_0, и положим $D(F) = d_{\Sigma_0}(F)$ (минимальная глубина формулы G, конгруэнтной формуле F).

Если система Σ *слабее* системы Σ_0 (Σ_0 *сильнее* Σ) в том смысле, что из соотношения $F \equiv_\Sigma G$ вытекает соотношение $F \equiv G$, то тогда у формулы F меньше конгруэнтных ей формул относительно системы Σ. При этом $d_\Sigma(F) \geqslant D(F)$ (аналогичное неравенство имеет место и для других характеристик формул). Заметим, что обе приведенные в § 3 системы: ΣR, которая полна для R^+ со сложением и умножением, и ΣL, которая полна для семейства конечных языков L с операциями объединения и конкатенации, слабее системы Σ_0. Поэтому все нижние оценки, которые в дальнейшем будут доказаны в пределах системы Σ_0, останутся также справедливыми и для параллельных вычислений в R^+ и в L.

Заметим еще, что подстановка константы 1 вместо всякой отличной от 0 константы делает Σ_0 наиболее сильной из возможных систем в том, что касается тождеств, содержащих константы. Единственная интерпретация U, для которой система Σ_0 истинна, это булево кольцо $\{0, 1\}$ с операциями И и ИЛИ. Тем не менее для интерпретации U система Σ_0 не полна: $x^2 \neq x$ в Σ_0. Таким образом, для нижних оценок в системе Σ_0 имеются только аналоги с содержательным смыслом.

Пусть $F_1, \ldots, F_k$ — формулы. Обозначим через $\prod_{i=1}^{k} F_i$ формулу $(\ldots(F_1 * F_2) * \ldots F_k)$ (аналогичным образом введем $\sum_{i=1}^{k} F_i$). Примем естественный порядок для символов, встречающихся в формулах. Формулы упорядочим лексикографически. Условимся считать F (стандартным) одночленом, если $F = 1$ или $F = \prod_{i=1}^{k} x_i$, где $x_1, \ldots, x_k$ — переменные, расположенные по порядку. Стандартный многочлен имеет вид: $F = \sum_{i=1}^{k} F_i$, где $F_1, \ldots, F_k$ — различные одночлены, расположенные по порядку. Справедлива следующая теорема о стандартном представлении.

Теорема 4.1. А. *Каждой формуле F соответствует единственный многочлен* (*полином*) $\mathrm{poly}(F)$, *такой, что* $F \equiv \mathrm{poly}(F)$.

В. $F \equiv G$ *тогда и только тогда, когда* $\mathrm{poly}(F) = \mathrm{poly}(G)$.

Мы опускаем несложное доказательство этой теоремы.

Обозначим через $\mathrm{Mon}(F)$ множество одночленов (monomials) многочлена $\mathrm{poly}(F)$, через $w(F)$ — *вес* формулы F, т. е. мощность множества $\mathrm{Mon}(F)$, и через $\deg(F)$ — *степень* формулы F,

т. е. максимальную длину одночлена из $\mathrm{Mon}(F)$. Очевидно, что

$$w(F_1 + F_2) \leqslant w(F_1) + w(F_2); \tag{4.1}$$

$$w(F_1 * F_2) \leqslant w(F_1) \cdot w(F_2); \tag{4.2}$$

$$\deg(F_1 + F_2) = \max_{i=1,2} \deg(F_i); \tag{4.3}$$

$$\deg(F_1 * F_2) = \deg(F_1) + \deg(F_2). \tag{4.4}$$

Функция g, определенная на формулах, называется Σ_0-*инвариантной*, если из $F_1 \equiv F_2$ вытекает, что $g(F_1) = g(F_2)$ (т. е. если одна из частей равенства определена, то определена и другая и равенство выполняется). По теореме 4.1.B степень и вес являются Σ_0-инвариантными функциями.

Формула F называется *однородной*, если все ее одночлены имеют одинаковую степень. Если формула F является однородной, то однородными являются и все ее подформулы.

5. ФУНКЦИИ РОСТА

Вес формулы, т. е. число одночленов в соответствующем многочлене, дает грубую нижнюю оценку для глубины. Из (4.1) и (4.2) сразу следует, что $D(F) \geqslant \lceil \log w(F) \rceil + 1$. В этой статье логарифмы берутся по основанию 2, $\lceil x \rceil$ означает наименьшее целое, не меньшее чем x, $\lfloor x \rfloor$ означает наибольшее целое не большее чем x. Пример формулы $\prod_{i=1}^{n} (x_{2i-1} + x_{2i})$ показывает, что в общем случае эту оценку нельзя улучшить. Однако мы приведем ряд формул, главным образом формул, встречающихся в комбинаторных вычислениях, для которых эту оценку можно значительно улучшить, и получим для них достаточно точные оценки величины $D(F)$. Нижние оценки для $D(F)$ в этом случае основаны на том, что в процессе вычисления F в пределах системы Σ_0 вес не может расти слишком быстро, поскольку это привело бы к появлению нежелательных одночленов. Ключевая идея связана с понятием *функции роста* формулы, которая должна быть Σ_0-инвариантной.

Пусть F — некоторая формула, а p — некоторый одночлен. *Дополнением* p в пределах F называется множество одночленов $\mathrm{Comp}_F(p) = \{q\colon p \cdot q \in \mathrm{Mon}(F)\}$ (через $p \cdot q$ обозначен одночлен $\mathrm{poly}(p * q)$). *Функция роста* формулы F определяется соотношением

$$W_F(k) = \max_{\deg(p)=k} \max_{q \in \mathrm{Comp}_F(p)} |\mathrm{Comp}_F(q)|,$$

где p и q — одночлены. Если формула F однородна, то функцию W_F можно определить так:

$$W_F(k) = \max\{|\operatorname{Comp}_F(q)|\colon \deg(q) = \deg(F) - k\}.$$

Пусть G — подформула формулы F; тогда $F \equiv G * H + K$. Пусть, далее, $g \in \operatorname{Mon}(G)$ и $h \in \operatorname{Mon}(H)$; тогда $h \in \operatorname{Comp}_F(g)$ и $\operatorname{Mon}(G) \subseteq \operatorname{Comp}_F(h)$. Отсюда следует, что $W_F(k)$ является верхней оценкой для числа одночленов в любой подформуле формулы F, содержащей одночлен степени k. Заметим, что $W_F(0) = 1$ [1]) и что если формула F содержит одночлен степени k, то тогда $W_F(k) = w(F)$. Функции $w(\ )$ и $\deg(\ )$, а также функция роста $W_F(\ )$ являются Σ_0-инвариантными. Эту инвариантность в пределах системы Σ_0 можно использовать для получения нижних оценок для величины $D(F)$.

Обозначим через $W_F(k, j)$ максимум веса подформул степени k и глубины j в формуле F (если таких подформул нет, то $W_F(k, j) = 0$). Из (4.1) и (4.2) легко увидеть, что

$$W_F(k, 0) = \begin{cases} 1, & \text{если } k \leqslant 1, \\ 0 & \text{в противном случае} \end{cases}$$

и

$$W_F(k, j+1) \leqslant \max[2W_F(k, j), \max_{0<i<k} W_F(i, j) \cdot W_F(k-i, j)]. \quad (5.1)$$

Кроме того,

$$W_F(k, j) \leqslant W_F(k) \quad (5.2)$$

и

$$W_F(k, n) = W_F(k), \quad (5.3)$$

где $k = \deg(F)$, $n = d(F)$.

Мы можем теперь сформулировать основную лемму 5.1 и более общую теорему 5.2 о мажоризации.

Лемма 5.1. *Пусть $T(k, j)$ — функция, определяемая следующим образом:*

$$T(k, 0) = \begin{cases} 1, & \textit{если } k \leqslant 1, \\ 0 & \textit{в противном случае}, \end{cases} \quad (5.4)$$

$$T(k, j+1) = \min(W_F(k), \max[2T(k, j), \max_{0<i<k} T(i, j) \cdot T(k-i, j)]). \quad (5.5)$$

Положим

$$t(k) = \min\{j\colon T(k, j) = W_F(k)\}.$$

Тогда

A. $W_F(k, j) \leqslant T(k, j)$,
B. $D(F) \geqslant t(\deg(F))$.

[1]) Это верно, если ни один из одночленов не является частью другого, например если формула F бесповторна. — *Прим. перев.*

Доказательство. A. Доказывается индукцией по j: утверждение справедливо при $j = 0$, переход от j к $j + 1$ получается применением (5.1) и (5.2).

B. Непосредственно вытекает из инвариантности функции $W_F(\ )$, утверждения A и соотношения (5.3). ■

В некоторых приложениях было бы полезно использовать другие функции, связанные с весом и степенью. К тому же удобнее обращаться с логарифмом веса. Общая теорема о мажоризации, в которой налагаются естественные условия на ω — логарифм веса и на δ — степень, показывает, как могла бы в процессе вычисления при изменении степени k и глубины j возникать максимально возможная величина логарифма веса $W(k, j)$, и утверждает, что эта величина действительно является мажорирующей для любого конкретного вычисления формулы F. Точная формулировка дается ниже.

Теорема 5.2. *Пусть F — некоторая формула. Пусть ω: $Sb(F) \to R$ такая функция, что*

$$\omega(G) \leqslant 0, \quad \text{если} \quad G \in C \cup X, \tag{5.6}$$

$$\omega(G_1 + G_2) \leqslant \max_{i=1,2} \omega(G_i) + 1 \tag{5.7}$$

и

$$\omega(G_1 * G_2) \leqslant \omega(G_1) + \omega(G_2). \tag{5.8}$$

Пусть δ $Sb(F) \to N$ такая функция, что

$$\delta(G_1 + G_2) \geqslant \max_{i=1,2} \delta(G_i) \tag{5.9}$$

и

$$\delta(G_1 * G_2) \geqslant \delta(G_1) + \delta(G_2). \tag{5.10}$$

Пусть W: $N \to R$ такая неубывающая, принимающая лишь неотрицательные значения функция (функция роста), что

$$\omega(G) \leqslant W(\delta(G)), \quad \text{если} \quad \omega(G) \text{ определено}, \tag{5.11}$$

и

$$\omega(F) = W(\delta(F)). \tag{5.12}$$

Определим функцию T: $N \times N \to R$ соотношениями

$$T(k, 0) = 0, \tag{5.13}$$

$$T(k, j+1) = \min\{W(k), \max[T(k, j) + 1, \max_{0<i<k} (T(i, j) + T(k - i, j))]\}. \tag{5.14}$$

Определим пороговую функцию t: $N \to N$ соотношением

$$t(k) = \min\{j:\ T(k, j) = W(k)\}. \tag{5.15}$$

Тогда

A. *Функция T не убывает по каждой переменной.*

B. *Для любой подформулы $G \in Sb(F)$ выполняется соотношение*: $\omega(G) \leqslant T(\delta(G), d(G))$.

C. $d(F) \geqslant t(\delta(F))$.

D. *Если функции ω и δ являются Σ_0-инвариантными, то тогда* $D(F) \geqslant t(\delta(F))$.

Отметим, что при $\omega(\ ) = \log w(\ )$, $\delta(\ ) = \deg(\ )$ и $W(\ ) = \max(0, \log W_F(\ ))$ из этой теоремы извлекается лемма 5.1.

Доказательство. A. Поскольку $T(k, j) \leqslant W(k)$, справедливо соотношение

$$T(k, j+1) \geqslant \min(W(k), T(k, j)+1) \geqslant T(k, j).$$

Итак, функция T не убывает по второй переменной. Если $j = 0$, то $T(k+1, 0) = 0 = T(k, 0)$. При $j > 0$ неравенство $T(k+1, j) \geqslant T(k, j)$ устанавливается индукцией по $k + j$. Действительно,

$$\begin{aligned} T(k+1, j+1) = \min\{&W(k+1), \\ &\max[T(k+1, j)+1, \max_{0<i<k+1}(T(i, j)+T(k+1-i, j))]\} \geqslant \\ \geqslant \min\{&W(k), \max[T(k, j)+1, \\ &\max_{0<i<k}(T(i, j)+T(k-i, j))]\} = T(k, j+1) \end{aligned}$$

(неравенство здесь следует из условия неубывания функции $W(k)$ и из предположения индукции).

B. Доказывается индукцией по построению формулы. Рассмотрим следующие случаи:

1. $G \in C \cup X$. Тогда $\omega(G) \leqslant 0$, $\delta(G) \geqslant 0$ и $d(G) = 0$. Поэтому

$$T(\delta(G), d(G)) \geqslant T(0, 0) = 0 \geqslant \omega(G).$$

2. $G = G_1 \vartheta G_2$. Пусть $k = \delta(G)$, $j = d(G)$ и $k_i = \delta(G_i)$, $j_i = d(G_i)$, $i = 1, 2$. При этом $j = \max j_i + 1$.

2a. $G = G_1 + G_2$. Тогда $k \geqslant \max k_i$ и

$$\begin{aligned} T(k, j) = \max_i T(k, j+1) \geqslant & \qquad \text{(в силу A)} \\ \geqslant \max_i \min\{W(k), T(k, j_i)+1\} \geqslant & \qquad \text{(в силу (5.14))} \\ \geqslant \min\{W(k), \max_i T(k_i, j_i)+1\} \geqslant & \qquad \text{(в силу A)} \\ \geqslant \min\{W(k), \max_i \omega(G_i)+1\} \geqslant & \text{(по предположению индукции)} \\ \geqslant \omega(G). & \end{aligned}$$

2b. $G = G_1 * G_2$. Тогда $k \geqslant k_1 + k_2$ и

$T(k, j) \geqslant T(k_1 + k_2, j) \geqslant$ (в силу A)

$\geqslant \min\{W(k), T(k_1, j) + T(k_2, j)\} \geqslant$ (в силу (5.14))

$\geqslant \min\{W(k), T(k_1, j_1) + T(k_2, j_2)\} \geqslant$ (в силу A)

$\geqslant \min\{W(k), \omega(G_1) + \omega(G_2)\} \geqslant$ (по предположению индукции)

$\geqslant \omega(G)$.

C. Из (5.14), утверждения B и (5.12) следует, что

$$W(\delta(F)) \geqslant T(\delta(F), d(F)) \geqslant \omega(F) = W(\delta(F)).$$

Теперь утверждение следует из определения функции t.

D. Непосредственно следует из утверждения C. ■

Теперь перед нами задача дискретного динамического программирования: дана неубывающая функция $W(k)$, требуется вычислить функцию $t(k)$, определяемую соотношениями (5.13)—(5.15). Другая точка зрения — это точка зрения с позиций оптимального управления: какая комбинация сложений и умножений, приводящая к построению множества $\mathrm{Mon}(F)$, при ограничивающей функции $W(k)$ дает формулу F минимальной глубины? На самом деле функция $W(k)$ представляет собой лишь часть действительных ограничений. Однако в ряде случаев она весьма точно определяет оптимальную глубину и в значительной мере оптимальную формулу в конгруэнтном классе. В плане динамического программирования задача становится интересной, когда начинает играть роль ограничение $T(k, j) \leqslant W(k)$, из-за которого функция T на некоторых шагах должна увеличиваться лишь на «+1», что означает для исходной задачи неизбежность чередования умножений со сложениями. Это имеет место в том случае, когда функция W *субаддитивна*, т. е.

$$W(i) + W(j) \leqslant W(i) + W(j).$$

Теорема 5.3. *Пусть W — субаддитивная, неубывающая функция, принимающая лишь неотрицательные значения. Определим функции T и t соотношениями* (5.13)—(5.15). *Тогда если* $W(\lfloor k/2 \rfloor) = 0$, *то*

$$t(k) = \lceil W(k) \rceil, \tag{5.21}$$

в противном случае

$$t(k) \leqslant \min_{0<i<k} \{\max[t(i), t(k-i)] + \lceil \Delta W(k, i) \rceil\} \tag{5.22}$$

(*здесь* $\Delta W(k, i) = W(k) - W(k-i) - W(i)$).

Доказательство. Если $W(\lfloor k/2 \rfloor) = 0$, то тогда

$$T(k, j+1) = \min\{W(k), T(k, j) + 1\}$$

и очевидно, что (5.21) выполняется. Если $W(\lfloor k/2 \rfloor) > 0$, то тогда

$$T(\lfloor k/2 \rfloor, 1) + T(\lceil k/2 \rceil, 1) = 2 \geqslant T(k, 2)$$

и существует такое максимальное $j \leqslant t(k)$, что при некотором i

$$T(k, j) \leqslant T(i, j-1) + T(k-i, j-1).$$

Рассмотрим два случая:

А. Пусть $j = t(k)$. Тогда $T(k, j) = W(k)$ и

$$W(k) = T(k, j) \leqslant T(i, j-1) + T(k-i, j-1) \leqslant$$
$$\leqslant W(i) + W(k-i) \leqslant W(k),$$

а значит, $\Delta W(k, i) = 0$, $T(i, j-1) = W(i)$ и $T(k-i, j-1) = W(k-i)$.

Таким образом, $j - 1 = \max[t(i), t(k-i)]$ и

$$t(k) = \max[t(i), t(k-i)] + \Delta W(k, i) + 1.$$

В. Пусть $j < t(k)$. Тогда $T(k, j+1) \leqslant T(k, j) + 1$. Если бы оказалось, что $j < t(i)$ или $j < t(k-i)$, то тогда выполнялось бы неравенство

$$T(i, j) \geqslant T(i, j-1) + 1$$

или неравенство

$$T(k-i, j) \geqslant T(k-i, j-1) + 1,$$

а значит, и неравенства

$$T(i, j) + T(k-i, j) \geqslant T(i, j-1) + T(k-i, j-1) + 1 \geqslant$$
$$\geqslant T(k, j) + 1 \geqslant T(k, j+1).$$

Это противоречит определению числа j. Следовательно,

$$j \geqslant \max[t(i), t(k-i)]. \tag{5.23}$$

Поскольку $T(k, j) \leqslant W(i) + W(k-i)$ и $T(k, j+m) \leqslant T(k, j) + m$, отсюда следует, что

$$t(k) \geqslant j + \Delta W(k, i),$$

а это вместе с (5.23) дает (5.22). ■

Отметим, что если W, а значит, и T принимают лишь целые значения, то тогда из неравенства $j \leqslant t(i)$ вытекает неравенство $T(i, j) \geqslant T(i, j-1) + 1$. В этом случае утверждение В можно усилить, доказав вместо (5.23) соотношение

$$j = \max[t(i), t(k-i)] + 1$$

и соответственно неравенство

$$t(k) \geqslant \max[t(i), t(k-i)] + 1 + \Delta W(k, i).$$

Следствие 5.4. *Если к условиям теоремы* 5.3 *добавить условие о целочисленности значений функции* W, *то будут выполняться следующие соотношения*:

$$t(k) = W(k), \quad \text{если} \quad W(\lfloor k/2 \rfloor) = 0, \tag{5.24}$$

$$t(k) = \min_{0<i<k} \{\max[t(i), t(k-i)] + \Delta W(k, i) + 1\}. \tag{5.25}$$

Из (5.25) легко увидеть, что t — неубывающая функция, так что (5.25) можно упростить до

$$t(k) = \min_{k/2 \leqslant i < k} (t(i) + \Delta W(k, i) + 1). \tag{5.26}$$

Теорема 5.3 дает прямое определение функции t. Из нее следует, что оптимальная стратегия для исходной задачи состоит в том, чтобы для каждого k строить подформулы степени k с максимальным весом посредством повторяющихся сложений произведений подформул степени i и $k-i$ при некотором конкретном i.

В ряде случаев можно упростить определение функции t еще дальше, показав, что дихотомическое разбиение является оптимальным.

Рассмотрим разность первого порядка $\Delta W(k) = W(k+1) - W(k)$ функции W. Будем называть функцию W *выпуклой*, если $\Delta W(k)$ является неубывающей функцией. Легко проверить, что если W — выпуклая функция и $W(0) \leqslant 0$, то функция W субаддитивна.

Лемма 5.5. *Пусть* W — *выпуклая функция, принимающая целые неотрицательные значения, такая, что* $W(0) = 0$. *Пусть* t — *функция, определяемая соотношениями* (5.24), (5.26). *Тогда если* $W(\lfloor k/2 \rfloor) > 0$, то

$$t(k) = t(\lceil k/2 \rceil) + \Delta W(k, \lceil k/2 \rceil) + 1.$$

Доказательство. Пусть i — максимальное такое число, что

$$0 < i \leqslant k/2$$

и

$$t(k) = t(k-i) + \Delta W(k, i) + 1.$$

Если $t(k-i) = W(k-i)$, то тогда также и $t(\lceil k/2 \rceil) = W(\lceil k/2 \rceil)$, так что

$$t(\lceil k/2 \rceil) + \Delta W(k, \lceil k/2 \rceil) + 1 = W(k) - W(\lfloor k/2 \rfloor) \leqslant$$
$$\leqslant W(k) - W(i) + 1 = t(k),$$

и таким образом $t = \lceil k/2 \rceil$. В противном случае рассмотрим такое число j, что

$$0 < j \leqslant (k-i)/2$$

и

$$t(k-i)=t(k-i-j)+\Delta W(k-i,\ j)+1.$$

Положим

$$f(x,\ y)=W(k)-W(x)-W(y)-W(k-x-y)+t(k-x-y)+2.$$

Тогда $f(i,\ j)=t(k)$ и по определению функции t

$$f(i,\ j)=\min\{f(x,\ y):\ 0<x\leqslant k/2,\ 0<y\leqslant(k-x)/2\}. \quad (5.27)$$

Предположим, что $i<j$. Легко проверить, что $0<j\leqslant k/2$ и $0<i\leqslant(k-j)/2$. Кроме того, $f(j,i)=f(i,j)$, а это противоречит максимальности числа i. Следовательно, $j\leqslant i$. Допустим, что $i<\lceil k/2\rceil$. Тогда

$$f(i+1,\ j-1)-f(i,\ j)=\Delta W(j-1)-\Delta W(i)\leqslant 0.$$

Если $j>1$, то тогда пара чисел $(i+1,\ j-1)$ удовлетворяет неравенству (5.27), а это противоречит определению числа i. Если же $j=1$, то тогда

$$f(i,\ j)\geqslant f(i+1,\ 0)=W(k)-W(i+1)-W(0)-W(k-i-1)+$$
$$+t(k-i-1)+2\geqslant t(k-i-1)+\Delta W(k,\ i-1)+1\geqslant t(k).$$

Это противоречит определению числа i. ■

Отсюда следует, что $i=\lceil k/2\rceil$, как и утверждает лемма.

Если функция W может принимать нецелые значения, то функцию t, удовлетворяющую соотношениям (5.21), (5.22), мы можем оценить снизу функцией $\bar{t}$, определяемой аналогичными соотношениями

$$\bar{t}(1)=W(1), \quad (5.28)$$

$$\bar{t}(k)=\min_{0<i<k}\{\max[\bar{t}(i),\ \bar{t}(k-i)]+\Delta W(k,\ i)\} \quad (5.29)$$

или эквивалентным образом (поскольку функция $\bar{t}$ не убывает)

$$\bar{t}(k)=\min_{k/2\leqslant i<k}[\bar{t}(i)+\Delta W(k,\ i)]. \quad (5.30)$$

Если W — выпуклая функция, то те же рассуждения, что при доказательстве леммы 5.5, показывают, что

$$\bar{t}(k)=\bar{t}(\lceil k/2\rceil)+\Delta W(k,\ \lceil k/2\rceil). \quad (5.31)$$

6. МОНОТОННЫЕ ВЫЧИСЛЕНИЯ ПЕРМАНЕНТА

Мы можем теперь использовать технику, развитую в разд. 5, для доказательства нижних оценок для конкретных формул.

Первой функцией, которую мы исследуем, будет перманент

$$P_n(x_{ij}) = \sum_{\sigma \in S_n} \prod_{i=1}^{n} x_{i\,\sigma(i)},$$

где S_n — группа перестановок на множестве $\{1, \ldots, n\}$. Перманент $P_n(x_{ij})$ является однородной формулой степени n веса $n!$ с n^2 переменными.

Теорема 6.1. $D(P_n) = 2n - 0{,}25 \log^2 n + O(\log n)$.

Доказательство. Функцией роста для P_n является $W_P(k) = k!$. Действительно, пусть p, q — одночлены, $\deg(p) = n - k$, $\deg(q) = k$ и $p \cdot q \in \mathrm{Mon}(P_n)$. Тогда q является членом перманента некоторого минора порядка $k \times k$, определяемого одночленом p. Таким образом, $|\mathrm{Comp}_{P_n}(p)| = k!$.

Положим $W(k) = \log W_P(k) = \log k!$. Тогда $W(0) = 0$ и разность первого порядка $\Delta W(k) = W(k+1) - W(k) = \log(k+1)$ функции W будет возрастающей функцией.

Определим $\bar{t}$ соотношениями (5.28), (5.30). При этом $\bar{t}(1) = W(1) = 0$,

$$\bar{t}(k) = \bar{t}(\lceil k/2 \rceil) + \Delta W(k, \lceil k/2 \rceil) = \bar{t}(\lceil k/2 \rceil) + \log C_k^{\lceil k/2 \rceil}.$$

Пользуясь формулой Стирлинга, получим

$$\log C_k^{\lceil k/2 \rceil} = k - 0{,}5 \log k + 0{,}5\,(1 - \log \pi) + O(1/k).$$

Решением рекуррентного соотношения будет

$$\bar{t}(k) = 2k - 0{,}25 \log^2 k - O(\log k).$$

В силу леммы 5.1 и теоремы 5.3

$$D(P_n) \geqslant \bar{t}(n) = 2n - 0{,}25 \log^2 n + O(\log n).$$

Верхняя оценка дается формулой Лапласа для перманента, которая определяется индуктивно. Пусть $X = (x_{ij})$ — квадратная матрица порядка $2n$. Пусть $N = 1, \ldots, 2n$ и $I \subseteq N$, $|I| = n$. Обозначим через $\bar{P}_I$ перманент минора порядка $n \times n$ матрицы X, построенного на первых n строках и множестве I столбцов, а через $\underline{P}_I$ — аналогично определяемый перманент на последних n строках. Тогда

$$P_{2n}(X) = \sum_{I \subseteq N,\ |I| = n} \bar{P}_I * \underline{P}_{N \setminus I}.$$

Случай матрицы нечетного порядка рассматривается аналогично.

Обозначим через $s(n)$ глубину определенной выше формулы для перманента P_n. Имеем

$$s(1)=0,$$

$$s(n)=s(\lceil n/2\rceil)+\lceil \log C_n^{\lceil n/2\rceil}\rceil+1.$$

Из сравнения индуктивного определения функции s с индуктивным определением функции $\bar{t}$ ясно, что

$$s(n)=\bar{t}(n)+O(\log n).$$

Это завершает доказательство. ■

7. УМНОЖЕНИЕ НЕКОТОРЫХ МАТРИЦ

Следующая задача, которую мы рассмотрим, это умножение p матриц порядка $n\times n$: $M=X^1\cdot\ \ldots\ X^p$ $(X^i=(x^i_{jk}))$. Эта задача часто встречается при вычислении линейных рекуррентных соотношений. Имеем

$$m_{rs}=\sum_{1\leqslant i_1<\ldots<i_{p-1}\leqslant n} x^1_{ri_1}\cdot x^2_{i_1i_2}\cdot\ \ldots\ \cdot x^p_{i_{p-1}s},$$

что представляет собой многочлен степени p веса n^{p-1} с $n^2(p-2)+2n$ переменными. Будут получены близкие верхняя и нижняя оценки для $D(m_{rs})$.

Теорема 7.1. $D(m_{rs})\geqslant \lceil \log p\rceil \log n$.

Доказательство. Легко показать, что функцией роста для m_{rs} является $W_m(k)=n^{k-1}$. Поэтому

$$W(k)=\log W_m(k)=(k-1)\log n,$$

$$W(1)=0$$

и

$$\Delta W(k)=W(k+1)-W(k)=\log n,$$

так что W — выпуклая функция. Определим $\bar{t}$ соотношениями

$$\bar{t}(1)=W(1)=0,$$

$$\bar{t}(k)=\bar{t}(\lceil k/2\rceil)+\Delta W(k,\lceil k/2\rceil)=\bar{t}(\lceil k/2\rceil)+\log n.$$

Отсюда следует, что

$$\bar{t}(k)=\lceil \log k\rceil \log n$$

и, таким образом,

$$D(m_{rs})\geqslant \bar{t}(p)=\lceil \log p\rceil \log n.$$ ■

Замечание. Прямым вычислением функции $\bar{t}$ можно улучшить полученную выше оценку и показать, что

$$D(m_{rs})\geqslant \lceil \log p\rceil \lceil \log n\rceil.$$

С другой стороны, очевидно, что

$$D(m_{rs}) \leqslant \lceil \log p \rceil (\lceil \log n \rceil + 1).$$

Следствие 7.2. *Если* $n = 2^q$, *то* $D(m_{rs}) = \lceil \log p \rceil (\log n + 1)$.

Доказательство. Верхняя оценка очевидна. Нижняя оценка доказывается с помощью следствия 5.4.

8. СИММЕТРИЧЕСКИЕ МНОГОЧЛЕНЫ

Рассмотрим теперь формулы, определяющие симметрические многочлены от n переменных $x_1, \ldots, x_n$:

$$\sigma_k^n(x_1, \ldots, x_n) = \sum_{I \subseteq N,\ |I|=k} \prod_{i \in I} x_i,$$

где $N = \{1, \ldots, n\}$. Это пример задачи, когда обычная функция роста не помогает. Необходимо воспользоваться теоремой 5.2 во всей ее общности.

Многочлен, определяемый формулой σ_k^n, обладает линейностью по каждой переменной. Это свойство по определению является Σ_0-инвариантным. Кроме того, если этим свойством обладает формула F, то им обладает и любая ее подформула. Формула σ_k^n является к тому же однородной.

Теорема 8.1. $D(\sigma_n^{2n}) = 0{,}25 \log^2 n + O(\log n)$.

Доказательство. Обозначим через $v(G)$ число различных переменных, встречающихся в формуле G. Функции δ и ω с областью определения $Sb(\sigma_n^{2n})$ определим с помощью соотношений:

$$\delta(G) = \deg(G),$$
$$\omega(G) = \log w(G) - v(G) + \deg(G).$$

Если $G \in C$, то тогда $w(G) = 1$, $v(G) = 0$ и $\deg(G) = 0$, так что $\omega(G) = 0$. Если $G \in X$, то тогда $w(G) = 1$, $v(G) = 1$ и $\deg(G) = 1$, так что снова $\omega(G) = 0$. Таким образом, если $G \in C \cup X$, то $\omega(G) = 0$.

Если $G = G_1 + G_2$, то тогда $w(G) \leqslant w(G_1) + w(G_2)$, $v(G) \geqslant \max\limits_{i=1,2} v(G_i)$ и $\deg(G) = \deg(G_i)$, $i = 1, 2$, так что $\omega(G) \leqslant \max\limits_{i=1,2} \omega(G_i) + 1$.

Если $G = G_1 * G_2$, то тогда $w(G) \leqslant w(G_1) \cdot w(G_2)$, $v(G) = v(G_1) + v(G_2)$ (G_1 и G_2 не имеют общих переменных, так как иначе формула G не была бы линейной) и $\deg(G) = \deg(G_1) + \deg(G_2)$. Таким образом, $\omega(G) \leqslant \omega(G_1) + \omega(G_2)$.

Функции δ и ω удовлетворяют условиям теоремы 5.2. Обе эти функции являются Σ_0-инвариантными.

Положим $W(k) = \log C_{2k}^k - k$. Докажем, что для любой подформулы $G \in Sb(F)$, где $F \equiv \sigma_n^{2n}$, выполняется соотношение: $\omega(G) \leqslant W(\delta(G))$. В самом деле, пусть $d = \deg(G)$, $v = v(G)$. Очевидно, что $0 \leqslant d \leqslant n$ и $d \leqslant v \leqslant 2n$. Кроме того, $w(G) \leqslant C_v^d$. Требуемое соотношение будет доказано, если мы покажем, что при любом v, $d \leqslant v \leqslant 2n$, выполняется неравенство: $\log C_v^d - v + d \leqslant \log C_{2d}^d - d$. Но ведь функция $f(v) = \log C_v^d - v + d$ принимает максимальное значение при $v = 2d$ и это значение равно $\log C_{2d}^d - d$.

Заметим, что $W(1) = 0$ и что функция $\Delta W(k) = W(k+1) - W(k) = \log(2k+1) - \log(k+1)$ является возрастающей. Теперь можно применить теоремы 5.2, 5.3. Как обычно, определим

$$\bar{t}(1) = W(1) = 0,$$

$$\bar{t}(k) = \bar{t}(\lceil k/2 \rceil) + \Delta W(k, \lceil k/2 \rceil) = \bar{t}(\lceil k/2 \rceil) + \log C_{2k}^k - \log C_{2\lceil k/2 \rceil}^{\lceil k/2 \rceil} - \log C_{2\lfloor k/2 \rfloor}^{\lfloor k/2 \rfloor} = \bar{t}(\lceil k/2 \rceil) + 0{,}5 \log k - 1 + 0{,}5 \log \pi + O(1/k).$$

(Окончательный результат получается с помощью формулы Стирлинга. Заметим, что линейные члены взаимно уничтожают друг друга, а добавок $0{,}5 \log k$ происходит от членов второго порядка.)

Решением этого рекуррентного соотношения является

$$\bar{t}(k) = 0{,}25 \log^2 k + c \log k + O(1),$$

где $c = -0{,}75 + 0{,}5 \log \pi = 0{,}07575 \ldots$. Поскольку $D(\sigma_n^{2n}) \geqslant \bar{t}(n)$, мы получаем

$$D(\sigma_n^{2n}) \geqslant 0{,}25 \log^2 n + c \log n + O(1).$$

Верхнюю оценку $0{,}25 \log^2 n + O(\log n)$ для $D(\sigma_n^{2n})$ легко вывести из работы Пиппенджера [8], в которой для σ_n^{2n} построены формулы длины $n^{0{,}25 \log n + O(1)}$ [1]). В этой конструкции в полной мере используется идемпотентность сложения. Если такой возможности нет, то обычное построение по индукции

$$\sigma_k^{2n}(x_1, \ldots, x_{2n}) = \sum_{j=0}^{k} \sigma_j^n(x_1, \ldots, x_n) \cdot \sigma_{k-j}^n(x_{n+1}, \ldots, x_{2n})$$

дает для σ_n^{2n} формулу глубины $0{,}5 \log^2 n + O(\log n)$. ■

[1]) Этот результат раньше был получен Л. С. Хасиным в работе [23]. — *Прим. перев.*

Приводим без доказательства следующее усиление предыдущей теоремы.

Теорема 8.2. $D(\sigma_n^{kn}) \geqslant 0{,}25 \log^2 n + (c - 0{,}5 \log(k/k-1)) \cdot \log n + \log k + O(1)$, *где* $c = 0{,}5(\log \pi - 0{,}5) = 0{,}5757 \ldots$.

9. 2-3-ДЕРЕВЬЯ С ЧЕРЕДУЮЩИМИСЯ ЯРУСАМИ

Хорошо известно, что любую формулу F длины n, пользуясь ассоциативностью, коммутативностью и дистрибутивностью, можно преобразовать в эквивалентную формулу, имеющую глубину $O(\log n)$. Лучший результат в этом направлении принадлежит Мюллеру и Препарате [7] и состоит в следующем: если $l(F) = n$, то тогда $D(F) \leqslant c \log n$, где $c = 2{,}08, \ldots$. Манро [14] поставил проблему — доказать, хотя бы в ограниченной системе, нижнюю оценку, которая была бы выше, чем $\log n + O(1)$ (тривиальной нижней оценкой является $\log n$). Мы покажем, что это можно сделать по крайней мере в пределах ограниченной аксиоматической системы Σ_0 или системы более слабой. Рассмотрим семейство формул T_n, $n = 1, 2, \ldots$. Формулу T_n можно представить равномерным 2-3-деревом, в котором ярусы трехместного сложения чередуются с ярусами двуместного умножения. Индуктивно T_n определяется следующим образом:

$$A^{\varepsilon_1, \ldots, \varepsilon_n}_{\delta_1, \ldots, \delta_n} = x^{\varepsilon_1, \ldots, \varepsilon_n}_{\delta_1, \ldots, \delta_n},$$

$$A^{\varepsilon_1, \ldots, \varepsilon_k}_{\delta_1, \ldots, \delta_k, \delta_{k+1}} = \prod_{\varepsilon=0}^{1} A^{\varepsilon_1, \ldots, \varepsilon_k, \varepsilon}_{\delta_1, \ldots, \delta_k, \delta_{k+1}} \qquad (*\text{— ярус}),$$

$$A^{\varepsilon_1, \ldots, \varepsilon_k}_{\delta_1, \ldots, \delta_k} = \sum_{\delta=0}^{2} A^{\varepsilon_1, \ldots, \varepsilon_{k-1}, \varepsilon_k}_{\delta_1, \ldots, \delta_{k-1}, \delta_k, \delta} \qquad (+\text{— ярус}).$$

$T_n = A_\Lambda$ (Λ — пустое слово, T_n приписано корню дерева). Формулу T_n можно также получить из T_{n-1} (или из T_{n-k}), подставляя вместо переменных копии формулы T_1 (или соответственно формулы T_k), не имеющие общих переменных.

Формула T_n содержит 6^n переменных $x^{\varepsilon_1, \ldots, \varepsilon_n}_{\delta_1, \ldots, \delta_n}$, где $\varepsilon_i \in \{0, 1\}$, $\delta_i \in \{0, 1, 2\}$, $1 \leqslant i \leqslant n$, каждая из которых встречается в T_n один раз. Поскольку для каждого трехместного сложения требуется два шага, глубина формулы T_n равна $3n$. Каждый одночлен $p \in \mathrm{Mon}(T_n)$ имеет вид

$$p = \prod_{\varepsilon_j \in \{0, 1\}} x^{\varepsilon_1, \ldots, \varepsilon_n}_{\delta_1, \ldots, \delta_n},$$

где $\delta_i \in \{0, 1, 2\}$, причем δ_i удовлетворяет определенным ограничениям. Действительно, если посмотреть на $*$-ярус в формуле

T_n, то можно заметить следующее: если две переменные в одночлене p имеют одинаковые значения индексов $\varepsilon_1, \ldots, \varepsilon_k$, то тогда они имеют одно и то же значение индекса δ_{k+1} и тем более индексов $\delta_k, \ldots, \delta_1$; значение δ_1 одно и то же для всех переменных из одночлена p. Таким образом,

$$T_n \equiv \sum_{\varepsilon_1, \ldots, \varepsilon_n} \prod x^{\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n}_{\delta_1(\Lambda),\, \delta_2(\varepsilon_1),\, \ldots,\, \delta_n(\varepsilon_1, \ldots, \varepsilon_{n-1})}.$$

Сумма берется по всем возможным наборам значений $\delta_1(\Lambda)$, $\delta_2(\varepsilon_1), \ldots, \delta_n(\varepsilon_1, \ldots, \varepsilon_{n-1})$, каждое из которых пробегает множество $\{0, 1, 2\}$. Общее число свободных индексов вида δ в одном одночлене $p \in \mathrm{Mon}(T_n)$ равно $1 + 2 + \ldots + 2^{n-1} = 2^n - 1$, и, таким образом, $w(T_n) = 3^{2^n - 1}$. (По-другому это равенство можно получить, заметив, что при переходе от T_n к T_{n+1} вес умножается на 3^{2^n}. Формула T_n однородна и $\deg(T_n) = 2^n$.)

Лемма 9.1. *Функция роста для формулы T_n имеет вид*

$$W_{T_n}(k) = 3^{\sum_{i=1}^{n} \lfloor k/2^i \rfloor}. \tag{9.1}$$

Доказательство. Индекс $\delta(\varepsilon_1, \ldots, \varepsilon_j)$ встречается у 2^{n-j} переменных одночлена $p \in \mathrm{Mon}(T_n)$. Выбирая одночлен q степени $m = 2^n - k$ так, чтобы он был частью одночлена p, мы фиксируем m переменных и, следовательно, не менее $\lceil m/2 \rceil$ индексов вида $\delta(\varepsilon_1, \ldots, \varepsilon_{n-1})$, не менее $\lceil m/4 \rceil$ индексов вида $\delta(\varepsilon_1, \ldots, \varepsilon_{n-2})$ и т. д. При этом общее число индексов, оставшихся свободными, не превосходит

$$2^n - 1 - \sum_{i=1}^{n} \lceil m/2^i \rceil = 2^n - 1 - \sum_{i=1}^{n} \lceil (2^n - k)/2^i \rceil =$$

$$= 2^n - 1 - \sum_{i=1}^{n} 2^{n-i} + \sum_{i=1}^{n} \lfloor k/2^i \rfloor = \sum_{i=1}^{n} \lfloor k/2^i \rfloor,$$

причем эта оценка достигается.

Каждый из оставшихся свободных индексов можно выбрать тремя способами из множества $\{0, 1, 2\}$. Поэтому максимальное число одночленов степени k, дополняющих q, дается соотношением (9.1), как и утверждается в лемме. ■

Теорема 9.2. $D(T_n) = 3n$.

Доказательство. Формула T_n имеет глубину $3n$. Для оценки $D(T_n)$ снизу мы пользуемся почти теми же рассуждениями, что в теореме 5.3, только мы должны более аккуратно проанализировать случаи, возникающие при различных значениях $\Delta W(k, i)$,

где $W(k)$, как обычно, определяется соотношением

$$W(k) = \log W_{T_n}(k) = c \sum_{m=1}^{n} \lfloor k/2^m \rfloor,$$

в котором $c = \log_2 3 = 1{,}58, \ldots$. Для величины $\Delta W(k, i) = W(k) - W(i) - W(k-i)$ имеем

$$\Delta W(k, i) = c \sum_{m=1}^{n} (\lfloor k/2^m \rfloor - \lfloor i/2^m \rfloor - \lfloor (k-i)/2^m \rfloor) \geqslant 0, \quad (9.2)$$

$$\Delta W(k, i) \geqslant c, \quad \text{если} \quad i,\ k-i < 2^r, \quad k \geqslant 2^r \quad (9.3)$$

(благодаря члену при $m = r$),

$$\Delta W(k, i) \geqslant 2c, \quad \text{если} \quad i < 2^r, \quad k-i < 2^{r-1}, \quad k \geqslant 2^r \quad (9.4)$$

(благодаря членам при $m = r,\ r-1$).

Определим $T(i, j)$ и $t(i)$, как в теореме 5.2.

Лемма 9.3. А. *Если* $2^r \leqslant k$, *то* $t(k) \geqslant 3r$.

В. *Если* $k < 2^{r+1}$ *и* $t(k) = 3r$, *то* $W(k) - T(k, t(k)-1) \geqslant 0{,}5$ (*т. е. последнее приращение величины* $T(k,)$ *не меньше чем* 0,5).

Заметим, что утверждение А при $k = 2^n$ содержит в себе утверждение теоремы, которую мы доказываем. Дополнительное утверждение В требуется для индуктивного перехода.

Доказательство ведется индукцией по k. В случае $k \leqslant 2$ все очевидно. Пусть $2^r \leqslant k < 2^{r+1}$. Так же как в доказательстве теоремы 5.3, выбираем такое максимальное число $j \leqslant t(k)$, что при некотором $i \geqslant k/2$ выполняется соотношение

$$T(k, j) \leqslant T(i, j-1) + T(k-i, j-1). \quad (9.5)$$

Имеем $i \geqslant 2^{r-1}$, так что

$$t(i) \geqslant 3r - 3.$$

Теми же рассуждениями, которыми мы доказали (5.23), доказываем, что

$$j \geqslant m = \max[t(i),\ t(k-i)],$$

так что

$$T(i, j) = W(i), \qquad T(k-i, j) = W(k-i)$$

и

$$t(k) \geqslant j + \lceil \Delta W(k, i) \rceil.$$

1. Если $\Delta W(k, i) = 0$, то тогда в силу (9.3) $i \geqslant 2^r$, так что по предположению индукции $t(i) \geqslant 3r$. При этом $j \geqslant 3r$. Если $t(k) > j$, то тогда $t(k) \geqslant 3r + 1$ и утверждения А и В доказаны. Если же $t(k) = j$, то тогда из (9.5) следует, что $T(i, j-1) =$

$= W(i)$ и $T(k-i,\ j-1) = W(k-i)$, так что $j-1 \geqslant t(i) \geqslant 3r$ и снова $t(k) \geqslant 3r+1$.

2. Допустим, что $\Delta W(k, i) = c$. Рассмотрим несколько случаев.

2a. Пусть $2^r \leqslant i$. Тогда по предположению индукции $t(i) \geqslant 3r$ и $t(k) \geqslant 3r+2$.

2b. Пусть $2^{r-1} \leqslant i < 2^r$. Тогда в силу (9.4) $k-i \geqslant 2^{r-1}$ и $t(k-i) \geqslant 3r-3$.

2b 1. Если либо $m \geqslant 3r-2$, либо $j \geqslant m+1$, то тогда $j \geqslant 3r-2$, $t(k) \geqslant j+2 \geqslant 3r$, что доказывает утверждение A.

Если при этом $t(k) > 3r$, то справедливо и утверждение B. Если же $t(k) = 3r$, то предыдущие неравенства обращаются в равенства: $j = 3r-2 = t(k)-2$. Отсюда следует, что

$$T(k,\ t(k)-2) = T(k,\ j) \leqslant T(i,\ j-1) + T(k-i,\ j-1) \leqslant$$
$$\leqslant W(i) + W(k-i) \leqslant W(k) - c,$$
$$T(k,\ t(k)-1) \leqslant W(k) - c + 1 \leqslant W(k) - 0{,}5,$$

что и доказывает утверждение B.

Нам осталось рассмотреть случай

2b 2. $j = m = 3r-3 = t(i) = t(k-i)$.

В силу части B индуктивного предположения

$$T(i,\ j) = W(i) \geqslant T(i,\ j-1) + 0{,}5$$

и

$$T(k-i,\ j) = W(k-i) \geqslant T(k-i,\ j-1) + 0{,}5.$$

Таким образом,

$$T(i,\ j) + T(k-i,\ j) \geqslant T(k,\ j) + 1 = T(k,\ j+1).$$

Это противоречит определению числа j. Значит, этот случай невозможен.

3. Если $\Delta W(k, i) \geqslant 2c$, то тогда $t(k) \geqslant j+4 \geqslant 3r+1$, что и завершает доказательство. ■

Следствие 9.4. *Для любого n существует формула F длины $l = 6^n$, для которой не существует эквивалентной формулы глубины меньше чем $c \log l$, где $c = 1{,}16\ \ldots$.*

Доказательство. Такой формулой является T_n, поскольку $l(T_n) = 6^n$, $D(T_n) = 3n$ и $3n/\log 6^n = 3/\log_2 6 = 1{,}16\ \ldots$. ■

Мы предполагаем, что многочлен, определяемый формулой T_n, нельзя вычислить меньше чем за $c \log l$ шагов, даже если допустить использование вычитания.

Для соотношения между сложностью и глубиной формул были бы интересны даже значительно более слабые нижние

оценки глубины, затрагивающие лишь члены низших порядков, такие, например, как

$$\log n + O(\log \log n),$$

где n — сложность формулы. Подобные результаты были недавно установлены [21] для схемы Горнера в случае монотонных булевых операций (И, ИЛИ) и даже в более слабой форме для немонотонных базисов [22].

ЗАКЛЮЧЕНИЕ И НЕКОТОРЫЕ ГИПОТЕЗЫ

Мы получили весьма близкие верхние и нижние оценки глубины для ряда формул в случае, когда при их упрощении используется только ограниченное множество преобразований. Тривиальным образом эти результаты дают оценки сложности формул при тех же ограничениях. Пользуясь простым рассуждением о разветвлении на входе операции и результатом из [7], мы устанавливаем, что $L(F)$ — минимальная сложность формулы, эквивалентной формуле F, удовлетворяет неравенствам

$$2^{D(F)/C} \leqslant L(F) \leqslant 2^{D(F)},$$

где $C = 2{,}08\ldots$. Таким образом для перманента сложность формулы экспоненциальна, а для симметрических функций сверхполиномиальна. В сущности близкие верхние и нижние оценки сложности формулы можно получить, если тот подход, основанный на идеях динамического программирования, который использовался здесь, применить непосредственно к сложности [15].

Связь между $D(F)$ — глубиной и $C(F)$ — минимальным числом операций (сложений и умножений) [1]), требующихся для вычисления формулы, эквивалентной формуле F при тех же самых ограничениях, слабее. Очевидно, что $D(F) \leqslant C(F)$, если же формула F представляет собой однородный многочлен степени n с v переменными, то тогда $D(F) \leqslant c \log n \cdot \log C(F) + \log v$, где c — некоторая константа (Хияфил [13]), причем эти неравенства нельзя существенно улучшить. Действительно, поскольку симметрические функции можно вычислять с помощью $O(n^2)$ операций, второе неравенство бывает точным (с точностью до постоянного множителя), а вычисление x^n оказывается тем случаем, когда точным является первое неравенство. Известно, что для вычисления перманента требуется

[1]) Подразумевается, что результаты операций можно использовать многократно. — *Прим. перев.*

$O(n2^n)$ операций [17]. Мы предполагаем, что для умножения p матриц порядка $m \times m$ требуется $O(pm^2)$ операций.

Может показаться удивительным, что рассмотренные функции можно вычислять быстрее, если допускать взаимное уничтожение членов. Однако при использовании вычитания перманент можно вычислить за $n + 2\log n + O(1)$ параллельных шагов (Райзер [16]). Неизвестно, дает ли выигрыш использование вычитания при последовательных вычислениях. Результат Валианта [12] о P-полноте задачи вычисления перманента внушает мысль о том, что надежда на существенный выигрыш невелика (и о том, что доказательство нетривиальной нижней оценки в общем случае — это очень трудное дело).

Еще более впечатляющая ситуация сложилась для определителя (формула для которого в системе Σ_0 формально эквивалентна формуле для перманента). В то время, как определитель можно вычислить за $O(\log^2 n)$ параллельных шагов при условии, что допускается взаимное уничтожение членов (Ксанки [18]), к нему применима нижняя оценка, полученная для перманента, и, таким образом, в случае когда взаимное уничтожение членов не допускается, для его вычисления требуется $2n - O(n)$ параллельных шагов.

Хотя вычитание можно использовать для построения более быстрых неветвящихся алгоритмов, вычисляющих симметрические функции или произведение матриц, авторам неизвестно, можно ли с его помощью построить более быстрые параллельные вычисления. Авторам неизвестен также ни один метод уменьшения числа шагов, требующихся для вычисления 2-3-дерева T_n. Авторы предполагают, что установленные здесь результаты останутся справедливыми, если допустить и взаимное уничтожение членов, причем задача о 2-3-деревьях должна быть одной из наиболее легких.

Другой интересный вопрос — это глубина рассмотренных формул, если их интерпретировать как булевы формулы (т. е. сложение интерпретировать как ИЛИ, а умножение как И). В полном базисе для формулы сложности n можно построить эквивалентную ей формулу глубины $O(\log n)$ [19] и формулу, эквивалентную перманенту глубины $O(\log^3 n)$ [20]. До сих пор авторам неизвестно ни одного результата, показывающего, что полученные здесь нижние оценки не сохраняются, если ограничиться монотонными булевыми формулами (хотя имеются примеры, когда нижние оценки глубины, справедливые для системы Σ_0, не переносятся на монотонный булев случай).

Наконец, было бы интересно получить близкие верхние и нижние оценки для относительной силы системы Σ_0 по сравнению с полной системой, использующей для упрощения бесповторных формул взаимное уничтожение членов.

ЛИТЕРАТУРА

1. Brent R. P. The parallel evaluaton of arithmetic expressions in logarithmic time, Complexity of Sequential and Parallel Numerical Algorithms, Academic Press, New York, 83—102, 1973.
2. McColl W. F. Some results on circuit depth, Ph. D. Thesis, University of Warwick, 1977.
3. McColl W. F. The circuit depth of symmetric Boolean functions, JCSS 17, 108—115, 1978.
4. Goodrich G. B., Ladner R. E. and Fischer M. J. Straight-line programs to compute finite languages, Proc. of a conference on Theoretical Computer Science, University of Waterloo, 221—229, 1977.
5. Miller W. Computational complexity and numerical stability, SIAM J. Computing 4, 97—107, 1975.
6. Miller W. Computer search for numerical stability, J. ACM 22, 512—521, 1975.
7. Miller D. E., Preparata F. P. Restructuring of arithmetic expressions for parallel evaluation, J. ACM 23, 534—543, 1976.
8. Pippenger N. Short monotone formulas for threshold functions, IBM Tech. Rep. RC 5405, 1975.
9. Shamir E., Snir M. Lower bounds on the number of multiplications and the number of additions in monotone computations, IBM Tech. Rep. RC 6757, 1977.
10. Shamir E., Snir M. Lower bounds on depth complexity in monotone arithmetic computations, IBM Tech. Rep. RC 7055, 1978.
11. Snir M. On the complexity of formula simplification (готовится к печати).
12. Valiant L. G. The complexity of computing the permanent, Theor. Comp. Sci. 8, 189—201, 1979.
13. Hyafil L. On the parallel evaluation of multivariate polynomials, SIAM J. Computing 8, 120—123, 1979.
14. Munro J. I. The parallel complexity of arithmetic computations, Proc. of the 1977 Internatinal FCT-conference on Fundamentals of Computations Theory, Springer, Berlin, 466—475, 1977.
15. Snir M. On the size of monotone formulas, University of Edinburgh Tech. Rep. CSR-46-79, 1979.
16. Райзер Г. Дж. Комбинаторная математика. — М.: Мир, 1966.
17. Jerrum M., Snir M. Some exact complexity results for straight-line computations over semirings, University of Edinburgh Tech. Rep. CSR-58-80, 1980.
18. Csansky L. Fast parallel matrix inversion algorithms, SIAM J. Computing 5, 618—623, 1976.
19. Muller D. E., Preparata F. P. Bounds to complexities of networks for sorting and switching, J. ACM 22, 195—201, 1975.
20. Valiant L. G., личное сообщение.
21. Commentz-Walter B. Size-depth tradeoff in monotone Boolean formulae, Acta Inform. 12, 227—243, 1979.
22. Commentz-Walter B., Sattler J. Size-depth tradeoff in non-monotone Boolean formulae, Acta Inform. (в печати).
23. Хасин Л. С. Оценки сложности реализации монотонных симметрических функций формулами в базисе $\vee$, &, $\neg$, ДАН СССР 189, № 4, 752—755, 1969.

Сложность моделирования неравномерных распределений[1])

Д. Кнут, Э. Яо

Computer Science Department, Stanford University, Stanford, Calif. 94305, U. S. A.

Целью настоящей работы является введение в теорию сложности моделирования случайных величин с неравномерными распределениями, исходя из имеющегося источника равнораспределенных случайных битов[2]). Мы рассмотрим процедуры, которые минимизируют среднее число случайных битов, необходимых для моделирования случайных величин с произвольными распределениями в произвольных системах счисления.

В разд. 1 представлены некоторые неформальные примеры алгоритмов и наброски теорем, которые более детально будут рассматриваться в последующих разделах. Оптимальные алгоритмы моделирования дискретных распределений изложены в разд. 2. В разд. 3 введено понятие общей системы представления вещественных чисел, а в разд. 4 на основе этого понятия конструируется оптимальная процедура моделирования произвольного распределения. Количественные аспекты подобного построения анализируются в разд. 5.

Раздел 6 содержит некоторые теоретические результаты, не имеющие непосредственного отношения к задаче моделирования случайных величин; в нем показано, что для вычисления первых k двоичных знаков значения произвольной монотонной функции, отображающей $[0, 1]$ в $[0, 1]$, в среднем требуется знание менее чем $k + 4$ первых двоичных знаков аргумента, когда усреднение проводится по всем значениям аргумента из $[0, 1]$.

Поскольку оптимальные алгоритмы моделирования случайных величин, построенные в рамках неограниченной модели в разд. 1—6, как правило, являются непрактичными, в разд. 7 обсуждается важный специальный класс алгоритмов с конечным числом состояний. Несколько более сильная модель, чем

---

[1]) Knuth D. E., Yao A. C. The Complexity of Nonuniform Random Number Generation. Algorithms and Complexity. Academic Press, New York, 1976, 357—428.

[2]) То есть независимых случайных величин, принимающих значения 0 и 1 с равными вероятностями. — *Прим. перев.*

© 1976 by Academic Press, Inc.
© Перевод на русский язык, «Мир», 1983

устройство с конечным числом состояний, используется при моделировании экспоненциального распределения в разд. 8.

В разд. 9 обсуждаются некоторые возможные направления дальнейшего развития этого нового раздела теории сложности, изложение которого в силу его новизны авторы старались сделать замкнутым [1]).

1. ПРЕДВАРИТЕЛЬНЫЕ ПРИМЕРЫ

Предположим, что вы хотите сыграть в кости, но единственное, чем вы располагаете, — это монета, подбрасывая которую можно получить «случайный бит», равный 1, если монета падает гербом вниз, или 0 в противном случае, так что результат серии подбрасываний монеты может быть представлен в виде цепочки нулей и единиц. Если ваша монета идеальная, то каждое ее подбрасывание обеспечивает получение 0 или 1 с одинаковой вероятностью; кроме того, результаты подбрасываний будут независимы друг от друга. Задача состоит в получении результатов, эквивалентных бросанию двух правильных костей, так чтобы суммарное количество выпавших очков (2, 3, ..., 6, 7, 8, ..., 12) встречалось соответственно с вероятностями (1/36, 2/36, ..., 5/36, 6/36, 5/36, ..., 1/36).

Один из возможных путей решения поставленной задачи может состоять в трех последовательных подбрасываниях монеты, так что исходы (001, 010, 011, 100, 101, 110) можно интерпретировать как представленное в двоичной системе счисления число очков (1, 2, 3, 4, 5, 6), выпавших при бросании кости. Если результатом трех последовательных подбрасываний монеты является 000 или 111, то процесс подбрасывания должен быть повторен до получения трех неидентичных исходов. Затем аналогичным образом осуществляется имитация процесса бросания второй кости.

Подобная процедура имитации бросания кости посредством троекратного подбрасывания монеты может быть представлена в графической форме (рис. 1). Каждый кружок обозначает одно подбрасывание монеты, которое имеет два исхода, показанные стрелками справа от кружка; каждый квадратик обозначает «концевой узел» с приписанным ему числом выпавших очков.

Легко видеть, что в данном случае среднее число T подбрасываний монеты в точности равно 4, поскольку эта величина удовлетворяет уравнению $T = 3 + \frac{2}{8}T$. Можно вычислить это

[1]) При переводе специальных терминов мы руководствовались терминологией, принятой в отечественных изданиях многотомной монографии Д. Кнута [5, 6, 8]. — *Прим. перев.*

среднее значение менее специальным способом, который будет использоваться позже при решении других задач. Пусть p_m — вероятность того, что необходимо ровно m подбрасываний, и

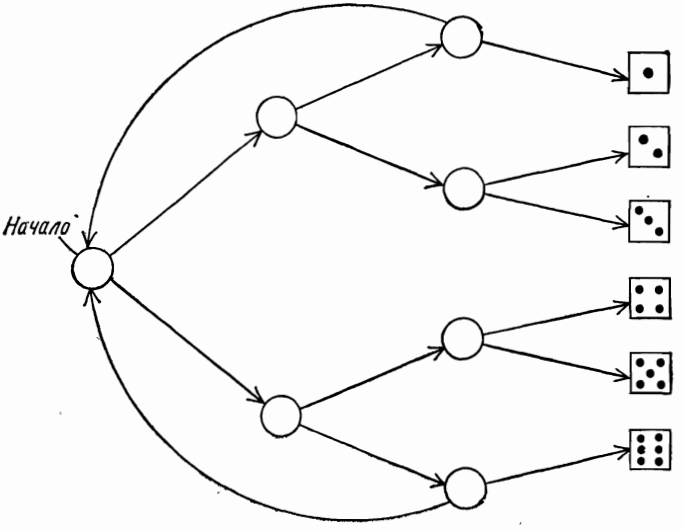


Рис. 1. Имитация бросания одной кости путем троекратного подбрасывания монеты.

пусть $q_m = p_{m+1} + p_{m+2} + \ldots$ — вероятность того, что потребуется более чем m подбрасываний монеты. Тогда среднее значение

$$\sum_{m\geqslant 1} mp_m = \sum_{m\geqslant 1} \sum_{0\leqslant l<m} p_m = \sum_{l\geqslant 0} \sum_{m>l} p_m = \sum_{l\geqslant 0} q_l. \tag{1.1}$$

(Изменения порядка суммирования законно, поскольку каждое $p_m \geqslant 0$.) Следовательно, для процедуры, изображенной на рис. 1, среднее число подбрасываний равно

$$\frac{1}{1} + \frac{2}{2} + \frac{4}{4} + \frac{2}{8} + \frac{4}{16} + \frac{8}{32} + \frac{4}{64} + \ldots$$

и сумма этого ряда равна 4. В общем случае легко видеть, что для *каждого* алгоритма, основанного на подбрасывании монеты, справедливо равенство

$$q_m = \frac{l(m)}{2^m}, \tag{1.2}$$

где целое $l(m) \geqslant 0$ равно числу «живых» (неокончившихся) ветвей в алгоритме после того, как произведено m подбрасываний. Можно представить себе расширение диаграммы, изображенной на рис. 1, в бесконечное бинарное дерево с корневым узлом на уровне 0 и $l(m)$ узлами разветвления на уровне m.

Заметим, что существует возможность продвижения по дереву, изображенному на рис. 1, без достижения когда-нибудь концевого узла. Это обстоятельство должно иметь место для *любой* корректной процедуры имитации бросания правильной кости посредством подбрасывания идеальной монеты, поскольку 6 ни при каком m не является делителем 2^m. Если не существует бесконечных путей в бинарном дереве, то, согласно «лемме о бесконечном дереве» (см., например, [5], разд. 2.3.4.3), существует такое m, что $q_m = 0$; следовательно, вероятность

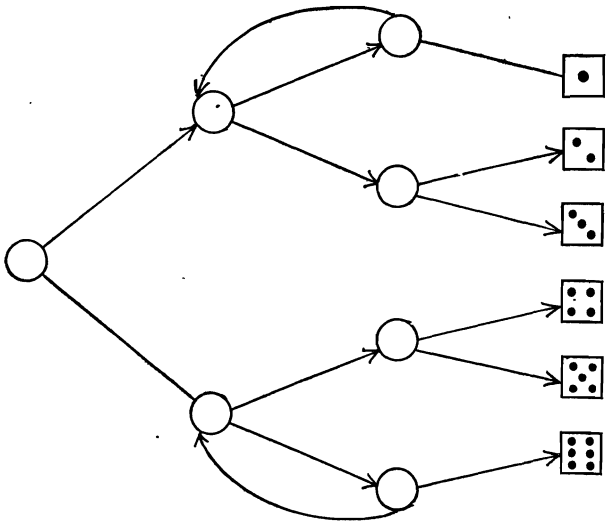
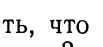
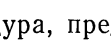


Рис. 2. Более эффективный способ имитации бросания одной кости.

каждого исхода в таком дереве должна быть кратна $1/2^m$. Поэтому анализ сложности алгоритмов, связанных с подбрасыванием монеты, не должен основываться просто на времени выполнения[1]) алгоритма в наихудшем случае. К счастью, необходимость в осуществлении большого числа подбрасываний на рис. 1 чрезвычайно мала — вероятность того, что потребуется более чем $3m$ подбрасываний, равна только $(2/8)^m$; на практике подобная процедура всегда завершается, за исключением, возможно, в книгах и пьесах [10]. В данной статье подобные процедуры, в том числе и бесконечные, мы называем «алгоритмами», хотя выражаясь более точно, их следовало бы называть «вычислительными методами», поскольку алгоритмами традиционно называются конечные процедуры.

[1]) Здесь и далее время выполнения алгоритма является синонимом необходимого числа подбрасываний монеты (случайных битов). — *Прим. перев.*

Читатель, возможно, уже заметил источник неэффективности на рис. 1: когда получаются три идентичных исхода подбрасываний монеты, мы могли бы использовать этот общий результат в качестве исхода следующего подбрасывания, поскольку исходы 000 и 111 равновероятны. Таким образом, мы получаем процедуру, представленную в графической форме на рис. 2, которая требует в среднем только $3\frac{2}{3}$ подбрасывания монеты [1]).

Если же использовать эту процедуру для имитации каждого из двух бросаний кости, то в среднем потребуется $7\frac{1}{3}$ подбрасывания монеты. На самом деле возможно значительное улучшение этой процедуры, поскольку нам не нужна полная информация об индивидуальном количестве очков, выпавшем в результате бросания каждой отдельной кости; все, что нам требуется, — это суммарное количество выпавших очков вне зависимости от того, является ли, например, результат, равный 7, суммой  или  .

Можно показать, что процедура, представленная в графической форме на рис. 3, обеспечивает правильные вероятности исходов бросания двух костей и для ее выполнения требуется в среднем только $4\frac{7}{18}$ подбрасывания монеты.

В разд. 2, в котором исследуются общие вопросы получения выборочных реализаций из произвольного дискретного распределения, будет показано, что рис. 3 соответствует *оптимальному* алгоритму моделирования исходов бросания двух костей среди всех алгоритмов, основанных на подбрасывании монеты, причем «оптимальность» понимается в самом сильном смысле. Пусть q_m — вероятность того, что для выполнения процедуры, изображенной на рис. 3, потребуется более чем m подбрасываний монеты, и пусть Q_m — соответствующая вероятность для любого пригодного для этой цели алгоритма; тогда $Q_m \geqslant q_m$. В частности, алгоритм, соответствующий рис. 3, минимизирует каждый член ряда (1.1), поэтому он имеет минимальное среднее время выполнения среди всех пригодных для этой цели алгоритмов.

Если мы захотим имитировать бросание одной кости, то рис. 2 предоставляет нам оптимальную процедуру для этой цели. Однако если мы захотим собрать полную информацию о каждом из 36 равновероятных исходов двух последовательных бросаний кости, то существует *лучший* способ, чем просто

[1]) Поскольку эта величина является решением уравнения $T = 3 + \frac{2}{8}(T - 1)$. — *Прим. перев.*

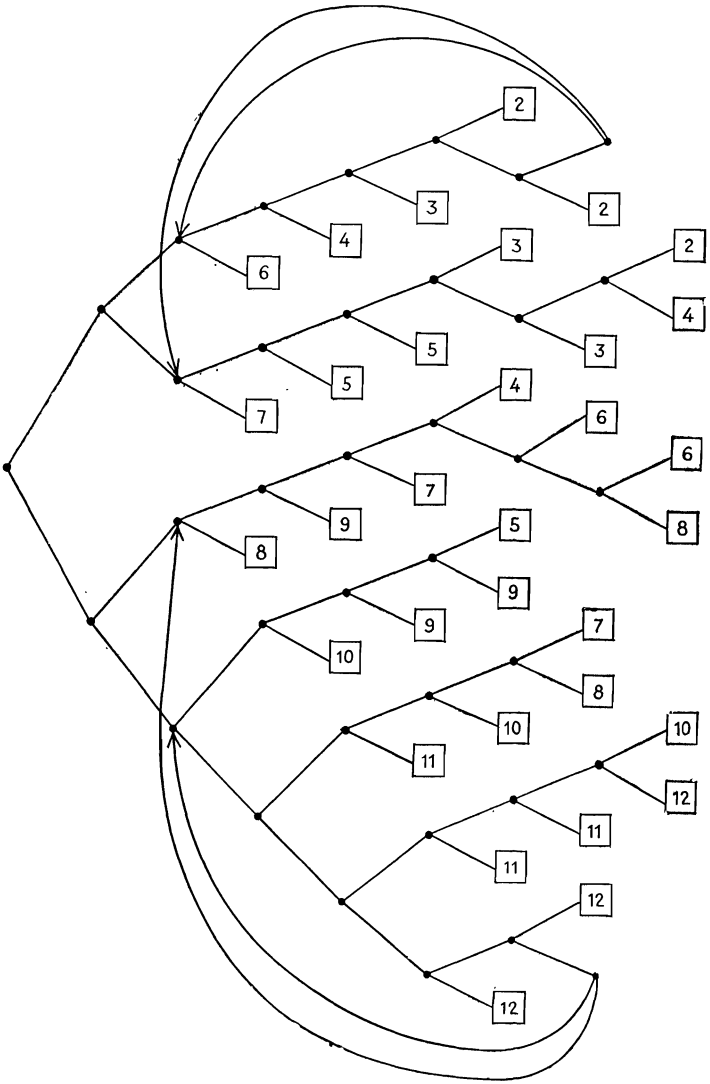


Рис. 3. Оптимальная процедура моделирования распределения суммы очков, выпавших в результате бросаний двух костей.

применение процедуры, представленной на рис. 2, дважды! Из результатов разд. 2 будет следовать, что два бросания кости можно имитировать путем подбрасывания монеты в среднем $6\frac{2}{3}$ раза. Более того, можно имитировать процесс k бросаний кости в среднем с помощью менее чем $2 + k \log_2 6$ подбрасываний мо-

неты; например, при $k = 3$ оптимальное среднее число подбрасываний равно $8\frac{35}{57}$.

Обратимся теперь к следующей задаче, которая, несмотря на свое кажущееся различие, в действительности тесно связана с предыдущей: в отличие от дискретного случая рассмотрим случай моделирования непрерывной *вещественнозначной* случайной величины X, имеющей функцию распределения [1])

$$F(x) = P(X \leqslant x), \quad -\infty < x < +\infty. \tag{1.3}$$

Поскольку мы не имеем возможности получать бесконечно много цифр значения случайной величины X в конечное время на ЭВМ, будем иметь дело с алгоритмами, которые доставляют последовательные двоичные знаки случайной величины X до тех пор, пока в этом есть необходимость; таким образом, если алгоритм выполняется достаточно долго, то в результате можно получить искомое значение со сколь угодно большой степенью точности. Задача состоит в моделировании подобных случайных величин, используя только источник равнораспределенных случайных битов, например подбрасывая идеальную монету достаточно много раз.

Предположим, например, что случайная величина X имеет функцию распределения

$$F(x) = \begin{cases} 0, & x < 0, \\ x^2, & 0 \leqslant x \leqslant 1, \\ 1, & x > 1. \end{cases} \tag{1.4}$$

Один из способов моделирования этой случайной величины состоит в присвоении $X \leftarrow \sqrt{U}$, где U — случайная величина, равномерно распределенная между 0 и 1, поскольку

$$P(X \leqslant x) = P(\sqrt{U} \leqslant x) = P(U \leqslant x^2) = x^2, \quad 0 \leqslant x \leqslant 1.$$

Следующая процедура, написанная на специально приспособленном для наших целей диалекте Алгола, доставляет последовательные двоичные знаки результата, используя вариант

---

[1]) В отечественной литературе функция распределения обычно определяется как $F(x) = P(X < x)$; по поводу соответствия между обоими определениями см. разд. 4. — *Прим. перев.*

обычного «школьного» правила извлечения квадратного корня.

```
A: begin integer n, r;
   n ← r ← 0;
   while true do
     begin r ← 4 × r + 2 × flip + flip; n ← 2 × n;
         if r > n then begin r ← r − n − 1; n ← n + 2;
           output (1) end
         else output (0);
       end;
   end;
```

Здесь *flip* означает входной случайный бит, равный 0 или 1, а результатом операции *output* является очередной знак двоичной дроби X. Корректность этого алгоритма вытекает из рассмотрения следующих соотношений, связывающих вход и выход алгоритма с величинами n и r в начале цикла **while**:

если входом алгоритма является цепочка двоичных знаков, состоящая из $2k$ нулей и единиц, которая определяет целое число m в двоичной записи, то выходом алгоритма является двоичная цепочка длины k, которая определяет целое число $\lfloor\sqrt{m}\rfloor$; кроме того $n = 2\lfloor\sqrt{m}\rfloor$ и $r = m - n^2/4$.

В качестве примера выполнения этого алгоритма предположим, что исходы первых 10 подбрасываний монеты образуют цепочку 1010011100; тогда первые пять битов результата равны 11001 и действительно квадратный корень из $(.1010011100\ \ldots)_2$ равен $(.11001\ \ldots)_2$. Таким образом, алгоритм A использует два входных бита для получения каждого выходного бита.

Изучающие моделирование случайных величин знают, что существует другой способ моделирования распределения (1.4), который состоит в выборе наибольшей из двух независимых равномерно распределенных между 0 и 1 случайных величин U и V, поскольку

$$P(\max(U, V) \leqslant x) = P(U \leqslant x,\ V \leqslant x) = x^2, \qquad 0 \leqslant x \leqslant 1.$$

Этот факт позволяет в большинстве случаев использовать менее чем $2k$ входных битов для получения k битов результата; действительно, достаточно нескольких первых битов U и V, чтобы выяснить, что $U \neq V$, и затем можно использовать оставшиеся биты наибольшего из чисел в качестве остальных битов результата. Таким образом, можно применять следующий ал-

горитм:

```
B: begin integer u, v;
       u ← v ← 0;
       while u = v do
          begin u ← flip; v ← flip;
             output (max (u, v));
          end;
       while true do output (flip);
   end;
```

Среднее число T_k входных битов, необходимых для получения k битов результата с помощью этого алгоритма, удовлетворяет рекуррентному соотношению

$$T_k = 2 + \frac{1}{2} T_{k-1} + \frac{1}{2}(k-1), \quad k \geqslant 1, \quad T_0 = 0,$$

откуда следует, что $T_k = k + 2 - 2^{1-k}$.

Более детальный анализ алгоритма B показывает, что он может быть улучшен, если использовать первое из двух подбрасываний монеты для решения вопроса о том, равны или нет u и v; если $u \neq v$, то второе подбрасывание является излишним. В результате мы получаем более простой алгоритм

```
C: begin integer u;
       u ← 0;
       while u = 0 do
          begin u ← flip;
             if u = 1 then output (1) else output (flip)
          end;
       while true do output (flip);
   end;
```

Среднее число T_k подбрасываний монеты, необходимых для получения с помощью этого алгоритма k битов результата, является решением уравнения

$$T_k = \frac{1}{2} k + \frac{1}{2}(2 + T_{k-1}), \quad k \geqslant 1, \quad T_0 = 0,$$

а именно

$$T_k = k + 1 - 2^{-k}. \tag{1.5}$$

Ниже будет показано, что алгоритм C реализует *оптимальный* способ моделирования случайной величины X с распределением (1.4), где оптимальность понимается в самом сильном смысле: при любых целых k и m среди всех пригодных для моделирования этого распределения алгоритмов алгоритм C минимизирует вероятность того, что потребуется m или более подбрасываний монеты, прежде чем будет получено k битов результата. Кроме того, в дальнейшем будет конструктивно

показано, что оптимальные в указанном смысле алгоритмы существуют для *любого* распределения F.

Задача имитации k бросаний кости, $k = 1, 2, 3, \ldots$, может рассматриваться как задача моделирования вещественнозначной случайной величины между 0 и 1 в системе счисления по основанию 6, исходя из последовательности случайных битов. В действительности существует общий подход, который включает в себя оба типа задач в качестве частных случаев. В разд. 3 будет обсуждаться идея общей *системы представления* для вещественных чисел, которая включает в себя в качестве специальных случаев как двоичную, так и системы счисления по смешанным основаниям и даже такие формы представления чисел, как система с «плавающей точкой» и представление в виде непрерывных дробей. В результате общая задача моделирования неравномерных распределений превращается в задачу моделирования случайной величины X с данным распределением F в конкретной системе представления, и в дальнейшем мы будем строить оптимальные методы для всех вариантов этой общей задачи. Задача моделирования дискретных распределений, рассматриваемая в разд. 2, представляет собой просто получение всех значащих цифр представления соответствующей случайной величины.

Прежде чем перейти непосредственно к рассмотрению дискретного случая, заметим, что Аренсом в [1] предложен несколько отличный от изложенных выше метод моделирования случайной величины X с распределением (1.4). В принятой нами форме записи алгоритмов его так называемый метод «втискивания бита» выглядит следующим образом:

```
D: begin integer i, j;
     j ← 0;
   while flip = 0 do j ← j + 1;
    for i ← 1 step 1 until j do output (flip);
    output (1);
    while true do output (flip);
  end;
```

Эта процедура весьма похожа на алгоритм C, за исключением того, что она удерживает первые j битов результата; среднее число подбрасываний монеты, необходимых для получения k битов результата, равно

$$k + 1 + 2^{-k}, \tag{1.6}$$

что несколько больше, чем (1.5) [1]). Таким образом, метод Аренса не является оптимальным с нашей точки зрения. Однако с

[1]) Поскольку в алгоритме D величина j может превышать k. — *Прим. перев.*

практической точки зрения j, как правило, очень мало, величина k обычно не меньше 10 и алгоритм D может быть очень эффективно реализован на двоичной ЭВМ с использованием операций сдвига. Поэтому алгоритм D в действительности является более практичным, чем оптимальный алгоритм C!

Как всегда мы должны быть осторожными по отношению к нашим теоретическим результатам по сложности алгоритмов. Теория очень полезна для локализации узких мест в задаче, для обнаружения новых алгоритмов и установления границ их применимости, однако не надо забывать при этом, что мы имеем дело лишь с упрощенной идеализацией действительности (см. [7] для дальнейшего обсуждения последствий злоупотребления теорией).

2. МОДЕЛИРОВАНИЕ ДИСКРЕТНЫХ СЛУЧАЙНЫХ ВЕЛИЧИН

Предположим, что случайная величина X принимает различные значения $x_1, x_2, \ldots, x_n$ соответственно с вероятностями $p_1, p_2, \ldots, p_n$, где

$$p_1 + p_2 + \ldots p_n = 1, \tag{2.1}$$

причем n может быть бесконечно. Мы будем изучать класс таких алгоритмов моделирования случайной величины X с этим распределением вероятностей, которые реализуют процесс подбрасывания монеты так, как это обсуждалось в разд. 1. Каждый такой алгоритм может быть представлен в виде бинарного (как правило, бесконечного) дерева, содержащего узлы двух типов:

(а) узлы разветвления, которые имеют двух потомков и означают «подбросить монету и перейти к одному из потомков в зависимости от исхода бросания»;

(б) концевые узлы, которые не имеют потомков; они помечены одним из значений x_i и означают «получить в качестве результата x_i и закончить выполнение алгоритма».

Диаграммы на рис. 1—3 представляют собой примеры подобных деревьев, если их очевидным образом расширить до бесконечных деревьев. Будем считать, что корень подобного бинарного дерева имеет уровень 0, а потомки узлов уровня m имеют уровень $m+1$. Если $t_i(m)$ — число концевых узлов уровня m, которые помечены значением x_i, то для каждого i должны выполняться равенства

$$\sum_{m \geqslant 0} t_i(m)/2^m = p_i. \tag{2.2}$$

Если эти условия выполняются, то мы будем говорить о *ПДР-дереве* (дереве, порождающем дискретное распределение) [1]), а соответствующие алгоритмы будем называть *ПДР-алгоритмами*. Целью настоящего раздела является изучение таких ПДР-деревьев, которым соответствуют наиболее быстродействующие алгоритмы.

Пусть

$$t(m) = t_1(m) + t_2(m) + \ldots + t_n(m) \tag{2.3}$$

— общее число концевых узлов ПДР-дерева уровня m. Из равенств (2.1), (2.2) и (2.3) следует, что

$$\sum_{m \geqslant 0} t(m)/2^m = 1, \tag{2.4}$$

следовательно, ПДР-алгоритм завершается с вероятностью 1.

Как отмечалось в разд. 1, среднее время выполнения подобного алгоритма (среднее число подбрасываний монеты) равно

$$\sum_{m \geqslant 0} q_m = \sum_{m \geqslant 0} \frac{l(m)}{2^m}, \tag{2.5}$$

где q_m — вероятность того, что потребуется более чем m подбрасываний монеты и $l(m)$ — количество узлов разветвления уровня m. Мы будем искать нижнюю границу для среднего времени выполнения алгоритма путем определения нижней границы для каждого $l(m)$.

Решение этой задачи очень простое. Во-первых,

$$\begin{aligned} &l(0) + t(0) = 1, \\ &l(m) + t(m) = 2l(m-1) \quad \text{при} \quad m \geqslant 1, \end{aligned} \tag{2.6}$$

так как каждый узел разветвления уровня $m-1$ имеет двух потомков. Кроме того,

$$l(m) = \sum_{\mu > m} t(\mu)/2^{\mu - m} \tag{2.7}$$

(это соотношение справедливо при $m = 0$ на основании (2.4), а при $m \geqslant 1$ следует по индукции из (2.6)). Далее

$$\sum_{\mu > m} t_i(\mu)/2^\mu = p_i - \sum_{0 \leqslant \mu \leqslant m} t_i(\mu)/2^\mu \geqslant \{2^m p_i\}/2^m, \tag{2.8}$$

где $\{x\} = x - \lfloor x \rfloor = x \pmod 1$ означает дробную часть x, так как левая сумма неотрицательна, а правая кратна 2^{-m}. Комбинируя (2.7) и (2.8), получаем искомую нижнюю границу

$$l(m) \geqslant \{2^m p_1\} + \{2^m p_2\} + \ldots + \{2^m p_n\}. \tag{2.9}$$

[1]) В оригинале DDG-tree (discrete distrubution generating tree). — *Прим. перев.*

Подстановка (2.9) в (2.5) показывает, что среднее время выполнения любого ПДР-алгоритма по меньшей мере равно

$$\nu(p_1)+\nu(p_2)+\ \dots\ +\nu(p_n), \tag{2.10}$$

где «новая» функция

$$\nu(x)=\sum_{m\geqslant 0}\{2^m x\}/2^m, \qquad 0\leqslant x\leqslant 1. \tag{2.11}$$

Если $n=\infty$, то (как мы увидим далее) нижняя граница (2.10) может быть бесконечной, и в таких случаях все ПДР-алгоритмы будут в среднем выполняться бесконечно долго.

Если же нижняя граница конечна, то ПДР-алгоритм будет достигать ее тогда и только тогда, когда в (2.8) имеет место равенство при всех i и m; это означает, что

$$\sum_{0\leqslant\mu\leqslant m} t_i(\mu)\,2^{m-\mu}=\lfloor 2^m p_i\rfloor, \tag{2.12}$$

т. е. $t_i(m)$ при $m=0,\ 1,\ 2,\ \dots$ просто является значением m-го бита в двоичном представлении p_i.

Поскольку нам часто придется иметь дело с двоичными разложениями, воспользуемся функциями Бореля

$$\varepsilon_m(x)=\lfloor 2^m x\rfloor \pmod 2, \tag{2.13}$$

которые определяют коэффициент при 2^{-m} в обычном двоичном представлении числа x, так что

$$x=\sum_m \varepsilon_m(x)/2^m \tag{2.14}$$

при $-\infty<x<\infty$, где суммирование осуществляется по всем целым значениям m. Предыдущее обсуждение привело нас, таким образом, к заключению, что ПДР-алгоритм имеет минимальное среднее время (2.10) тогда и только тогда, когда

$$t_i(m)=\varepsilon_m(p_i) \tag{2.15}$$

при всех i и m.

Обратно, мы должны показать, что для любых распределений вероятностей $(p_1,\ p_2,\ \dots,\ p_n)$ существуют ПДР-деревья, удовлетворяющие (2.15). Определим величины $t_i(m)$ и $t(m)$, как в (2.15) и (2.3), так что имеют место (2.2) и (2.4). Ясно, что ПДР-дерево с подобным определением $t_i(m)$ существует тогда и только тогда, когда целые $l(m)$, определенные в (2.6), неотрицательны. Но из (2.4) и (2.6) следует (2.7); следовательно, $l(m)\geqslant 0$ и оптимальные ПДР-деревья, удовлетворяющие (2.15), существуют.

В качестве примера рассмотрим распределение, определяемое трансцендентными вероятностями

$$
\begin{aligned}
p_1 &= 1/\pi = (0.010100010111110\ldots)_2,\\
p_2 &= 1/e = (0.010111100010110\ldots)_2,\\
p_3 &= 1 - p_1 - p_2 = (0.010100000101010\ldots)_2.
\end{aligned}
\tag{2.16}
$$

Одно из оптимальных ПДР-деревьев для этого распределения показано на рис. 4. Данное дерево существенно бесконечно — оно не может быть получено путем расширения конечной диаграммы, как в случаях, которым соответствуют рис. 1—3. Ясно, что алгоритмы с конечным числом состояний имеют место тогда и только тогда, когда все вероятности рациональны [1]). С другой

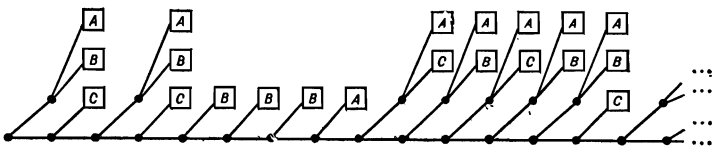


Рис. 4. Оптимальное ПДР-дерево, порождающее события (A, B, C) в соответствии с вероятностями (2.16).

стороны, алгоритм оптимального моделирования любого дискретного распределения может быть легко построен в виде функции (бесконечных) двоичных представлений данных вероятностей.

Следующая теорема резюмирует полученные результаты по моделированию дискретных распределений.

Теорема 2.1. *Пусть* $(p_1, p_2, \ldots, p_n)$ *— распределение вероятностей, причем* n *может быть бесконечно, и пусть* $\mathcal{A}$ *есть ПДР-алгоритм для этого распределения. Тогда среднее время* $T(\mathcal{A})$ *выполнения этого алгоритма (т. е. среднее число случайных битов, необходимых для выполнения* $\mathcal{A}$*) удовлетворяет неравенству*

$$T(\mathcal{A}) \geqslant v(p_1) + v(p_2) + \ldots + v(p_n), \tag{2.17}$$

где функция $v(x)$ *определена в* (2.11). *Следующие утверждения эквивалентны:*

(a) *при любом* $m \geqslant 0$ *среди всех ПДР-алгоритмов для данного распределения алгоритм* $\mathcal{A}$ *минимизирует вероятность того, что для их выполнения потребуется более чем* m *случайных битов;*

[1]) И, очевидно, их число конечно. — *Прим. перев.*

(б) *при каждом* $m \geqslant 0$ *и* $1 \leqslant i \leqslant n$ *ПДР-дерево, соответствующее алгоритму* $\mathscr{A}$, *имеет в точности* $\varepsilon_m(p_i)$ *концевых узлов уровня* m, *помеченных значением* x_i, *где* $\varepsilon_m(x)$ *обозначает коэффициент при* 2^{-m} *в двоичном представлении* x.

Кроме того, если $T(\mathscr{A})$ *конечно, то утверждение*

(в) $T(\mathscr{A}) = \nu(p_1) + \nu(p_2) + \ldots + \nu(p_n)$

также эквивалентно утверждениям (а) *и* (б). ■

Перейдем теперь к изучению функции $\nu(x)$, которая фигурирует в формуле для оптимального среднего времени. Имеем

$$\nu(x) = \sum_{m \geqslant 0} \Big(\sum_{\mu > m} \varepsilon_\mu(x)/2^{\mu-m} \Big) / 2^m = \sum_{\mu \geqslant 0} \sum_{0 \leqslant m < \mu} \varepsilon_\mu(x)/2^\mu$$
$$= \sum_{\mu \geqslant 0} \mu \varepsilon_\mu(x)/2^\mu \quad \text{при} \quad 0 \leqslant x \leqslant 1. \tag{2.18}$$

Удобнее доопределить $\nu(x)$ для произвольных вещественных неотрицательных чисел x так, что

$$\nu(x) = \sum_m m \varepsilon_m(x)/2^m \quad \text{при} \quad x \geqslant 0, \tag{2.19}$$

где суммирование производится по всем целым значениям m от $-\infty$ до $+\infty$. Таким образом, эта функция оказывается определенной для любого фиксированного x, поскольку $\varepsilon_m(x)$ будет равна 0 при всех достаточно больших отрицательных значениях m.

Когда x является степенью 2, то $\nu(x) = H(x)$, где

$$H(x) = x \log_2(1/x), \quad x > 0; \quad H(0) = 0. \tag{2.20}$$

Кроме того, легко видеть, что равенство

$$\nu(x/2^k) = kx/2^k + \nu(x)/2^k \tag{2.21}$$

справедливо также и для $H(x)$. Отсюда следует, что, по-видимому, $\nu(x)$ приближенно равна $H(x)$, и это действительно так.

Теорема 2.2. $H(x) \leqslant \nu(x) < H(x) + 2x$ *для всех* $x \geqslant 0$, *и эти границы являются неулучшаемыми.*

Доказательство. Ясно, что $\nu(0) = H(0)$. Пусть $x > 0$ такое, что $2^{-m} \leqslant x < 2^{1-m}$ [1]). Тогда

$$\nu(x) = \sum_{\mu \geqslant m} \mu \varepsilon_\mu(x)/2^\mu \geqslant \sum_{\mu \geqslant m} \log_2(1/x)\, \varepsilon_\mu(x)/2^\mu = H(x)$$

[1]) При остальных x утверждение теоремы справедливо на основании (2.21). — *Прим. перев.*

и

$$H(x)+2x-\nu(x)=\sum_{\mu\geqslant m}(\log_2(1/x)+2-\mu)\,\varepsilon_\mu(x)/2^\mu$$

$$>\sum_{\mu\geqslant m}(m+1-\mu)\,\varepsilon_\mu(x)/2^\mu=2^{-m}-\sum_{\mu\geqslant m+2}(\mu-m-1)\,\varepsilon_\mu(x)/2^\mu$$

$$>2^{-m}-\sum_{\mu\geqslant 1}\mu/2^{\mu+m+1}=0,$$

поскольку $\varepsilon_m(x)=1$.

Утверждение, что эти границы неулучшаемы, вытекает из рассмотрения значений $x=2^{-m}$ и $x=2^{1-m}-2^{1-m-k}$. ■

Следствие. *Пусть* $H(p_1, p_2, \ldots, p_n)=H(p_1)+H(p_2)+\ldots+H(p_n)$ *— энтропия распределения вероятностей* $(p_1, p_2, \ldots, p_n)$. *Тогда среднее время выполнения оптимального ПДР-алгоритма для этого распределения заключено между* $H(p_1, p_2, \ldots, p_n)$ *и* $H(p_1, p_2, \ldots, p_n)+2$. ■

Вероятно, наиболее известным свойством $H(x)$ является фундаментальное неравенство

$$H(x)+H(y)\leqslant H(x+y)+x+y, \tag{2.22}$$

которое превращается в равенство тогда и только тогда, когда $x=y$. Функция $\nu(x)$ не удовлетворяет этому соотношению (например, при $x=1/3$ и $y=2/3$ $\nu(x)=8/9$, $\nu(y)=10/9$ и $\nu(x+y)=0$), однако $\nu(x)$ удовлетворяет родственному *равенству*

$$\nu(x)+\nu(y)=\nu(x+y)+(x\oplus y), \tag{2.23}$$

где $x\oplus y$ является таким двоичным числом, в котором единицы встречаются ровно там, где происходят *переносы* при сложении x с y:

$$x\oplus y=\sum_m(\varepsilon_m(x+y)+\varepsilon_m(x)+\varepsilon_m(y))\,(\mathrm{mod}\,2)/2^m. \tag{2.24}$$

Равенство (2.23) следует из того факта, что величина ν сокращается на 2^{-m}, когда происходит перенос в позицию с номером m, так как

$$(m+1)/2^{m+1}+(m+1)/2^{m+1}=m/2^m+1/2^m.$$

Заметим, что

$$x\oplus y\leqslant 2(x+y), \tag{2.25}$$

поскольку, если $2^m\leqslant x+y<2^{m+1}$, то $x\oplus y\leqslant 2^m+2^{m-1}+2^{m-2}+\ldots=2^{m+1}$.

Закончим настоящий раздел определением экстремальных значений среднего времени выполнения оптимального алгоритма.

Теорема 2.3. *Пусть* $\nu(p_1, p_2, \ldots, p_n) = \nu(p_1) + \nu(p_2) + \ldots + \nu(p_n)$ *будет «новой энтропией» распределения вероятностей* $(p_1, p_2, \ldots, p_n)$. *Тогда*

$$\nu(p_1, p_2, \ldots, p_n) \leqslant \lceil \log_2 \rceil + (n-1)/2^{\lceil \log_2 n \rceil - 1}. \tag{2.26}$$

Кроме того, если все p_i *строго положительны, то*

$$\nu(p_1, p_2, \ldots, p_n) \geqslant 2 - 2^{2-n}, \tag{2.27}$$

и эти границы являются неулучшаемыми.

Доказательство. Справедливость верхней границы следует из того, что

$$\{2^m p_1\} + \{2^m p_2\} + \ldots + \{2^m p_n\} \leqslant \min(2^m, n-1) \tag{2.28}$$

при всех $m \geqslant 0$ (далее используется (2.10)); левая часть неравенства (2.28) равна $2^m - \lfloor 2^m p_1 \rfloor - \ldots - \lfloor 2^m p_n \rfloor$, т. е. целому числу, меньшему чем n. Когда $n = 6$, эта граница может быть достигнута, если, например, в двоичной системе счисления

$$\begin{aligned}
p_1 &= .001\ 111110\ 111110\ 111110\ \ldots \\
p_2 &= .001\ 111101\ 111101\ 111101\ \ldots \\
p_3 &= .001\ 111011\ 111011\ 111011\ \ldots \\
p_4 &= .000\ 110111\ 110111\ 110111\ \ldots \\
p_5 &= .000\ 101111\ 101111\ 101111\ \ldots \\
p_6 &= .000\ 011111\ 011111\ 011111\ \ldots .
\end{aligned}$$

Аналогично, в случае произвольного n можно положить

$$\begin{aligned}
p_j &= \frac{1}{2^q}\left(\frac{2^n - 2^{j-1} - 1}{2^n - 1}\right) + \frac{1}{2^q}, \quad 1 \leqslant j \leqslant 2^q + 1 - n, \\
p_j &= \frac{1}{2^q}\left(\frac{2^n - 2^{j-1} - 1}{2^n - 1}\right), \quad 2^q + 1 - n < j \leqslant n,
\end{aligned} \tag{2.29}$$

где $q = \lceil \log_2 n \rceil$. Заметим, что если $\theta = \lceil \log_2 n \rceil - \log_2(n-1)$, то $0 < \theta \leqslant 1$ и

$$\lceil \log_2 n \rceil + (n-1)/2^{\lceil \log_2 n \rceil - 1} = \log_2(n-1) + \theta + 2^{1-\theta}.$$

Так как функция $\theta + 2^{1-\theta} \leqslant 2$ при $0 \leqslant \theta \leqslant 1$, то

$$\nu(p_1, p_2, \ldots, p_n) \leqslant \log_2(n-1) + 2, \tag{2.30}$$

причем равенство возможно, когда $n - 1$ является степенью 2.

Перейдем теперь к рассмотрению нижней границы. Ясно, что нижняя граница может быть достигнута только в том случае, когда ПДР-дерево имеет наименьшее число концевых узлов, а именно n. В этом случае существует $n-1$ узлов разветвления и известно, что $\nu(p_1, p_2, \ldots, p_n)$ имеет минимальное значение, когда дерево вырождено (см. [8], упр. 6.3—37, где эквивалентный факт приводится без доказательства). Используя обозначения, принятые при доказательстве теоремы 2.1, из (2.15) и (2.18) следует, что для оптимального дерева всегда выполняется равенство

$$\nu(p_1, p_2, \ldots, p_n) = \sum_{m \geqslant 1} mt(m)/2^m, \tag{2.31}$$

что в свою очередь [1]) равно

$$\sum_{m \geqslant 1} m(2l(m-1) - l(m))/2^m = \sum_{m \geqslant 0} l(m)/2^m. \tag{2.32}$$

Поскольку в рассматриваемом случае $\sum_{m \geqslant 0} l(m) = n-1$, ясно, что (2.32) достигает минимума, когда $l(0) = l(1) = \ldots = (l(n-2) = 1$. Эта нижняя граница достигается, например, для распределения

$$p_j = \begin{cases} 2^{-j}, & 1 \leqslant j < n, \\ 2^{1-n}, & j = n. \end{cases} \tag{2.33}$$

■

Когда n бесконечно, мы можем, используя теорему 2.2, подобрать такое распределение вероятностей, что $\nu(p_1, p_2, \ldots) = \infty$. Вероятно, наиболее интересный пример подобного рода может быть получен, если положить $p_1 = 0$ и

$$p_j = \frac{1}{2^{\lfloor \log_2 j \rfloor + 2 \lfloor \log_2 \log_2 j \rfloor + 1}}, \qquad j \geqslant 2. \tag{2.34}$$

Любопытно проверить, что в этом случае $\Sigma p_j = 1$ и $\Sigma \nu(p_j) = \infty$. Поскольку каждое p_j обратно степени 2, то оптимальное ПДР-дерево может быть довольно легко построено.

3. ОБЩИЕ СИСТЕМЫ ПРЕДСТАВЛЕНИЯ

Прежде чем приступить к моделированию непрерывных распределений, рассмотрим довольно общую модель для представления вещественных чисел.

Пусть (возможно, неограниченный) интервал R вещественной оси разбит на один или более непересекающихся интервалов, скажем

$$R = R[a] \cup R[a+1] \cup \ldots \cup R[b], \tag{3.1}$$

[1]) Согласно (2.6). — *Прим. перев.*

где a и b либо целые, либо равны $\pm\infty$, $a \leqslant b$ и $R[i]$ находится левее $R[j]$ при $i < j$. Пусть далее каждый интервал $R[j]$ разбит аналогичным образом на один или более непересекающихся интервалов так, что

$$R[j] = R[j, a_j] \cup R[j, a_j + 1] \cup \ldots \cup R[j, b_j]. \tag{3.2}$$

Затем каждый интервал $R[j_1, j_2]$ разбивается подобным же образом, и этот процесс продолжается счетное число раз. При этом требуется единственно, чтобы длины интервалов в конечном счете стремились к нулю, т. е., выражаясь формально, предполагается, что бесконечное пересечение

$$R \cap R[j_1] \cap R[j_1, j_2] \cap R[j_1, j_2, j_3] \cap \ldots \tag{3.3}$$

содержит самое большее одну точку для каждой бесконечной последовательности целых чисел $j_1, j_2, j_3, \ldots$.

Подобное дерево интервальных разбиений будем называть *системой представления* для R; согласно определению, каждое вещественное число $x \in R$ имеет единственное представление

$$x = \langle j_1, j_2, j_3, \ldots \rangle, \tag{3.4}$$

которое определяется условием

$$x \in R[j_1, j_2, j_3, \ldots, j_r] \text{ при любом } r \geqslant 1. \tag{3.5}$$

Если $x < x'$, то представление для x всегда будет лексикографически меньше представления для x', т. е. если $x = \langle j_1, j_2, j_3, \ldots \rangle$ и $x' = \langle j'_1, j'_2, j'_3, \ldots \rangle$, то найдется некоторое $m \geqslant 1$, такое, что $j_i = j'_i$ при $1 \leqslant i < m$, но $j_m < j'_m$.

Существует много известных примеров таких систем представления.

(а) Пусть $R = [0, 1)$, и пусть при $0 \leqslant j_1, j_2, \ldots, j_r < d$

$$R[j_1, j_2, \ldots, j_r] = [(j_1 d^{r-1} + j_2 d^{r-2} + \ldots + j_r)/d^r, \\ (j_1 d^{r-1} + j_2 d^{r-2} + \ldots + j_r + 1)/d^r), \tag{3.6}$$

что, разумеется, представляет собой стандартную систему счисления по основанию d для дробей.

(б) Пусть $R = (-\infty, \infty)$, и пусть

$$R[-1] = (-\infty, 0),\ R[0] = [0, 0],\ R[1] = (0, \infty), \\ R[1, i] = [2^i, 2^{i+1}),\ R[-1, -i] = [-2^{i+1}, -2^i] \tag{3.7}$$

при всех целых i. Далее эти интервалы разбиваются следующим образом: полагается $R[0, \ldots, 0] = R[0]$ и

$$R[+1, i, j_1, \ldots, j_r] = [2^i \times (1.j_1 \ldots j_r)_2, \\ 2^i \times ((1.j_1 \ldots j_r)_2 + 2^{-r})), \\ R[-1, i, j_1, \ldots, j_r] = -R[1, -i, 1 - j_1, \ldots, 1 - j_r] \tag{3.8}$$

для всех $0 \leqslant j_1, \ldots, j_r \leqslant 1$. Это *двоичная система с плавающей точкой.*

(в) Пусть $R = (0, 1]$, и пусть при $j_1, j_2, \ldots, j_r \geqslant 1$ $R[-j_1, j_2, \ldots, (-1)^r j_r]$ — множество целых чисел, лежащих между

(3.9)
$$\cfrac{1}{j_1 + \cfrac{1}{j_2 + \quad \cdot \quad \cdot \quad \cdot \quad + \cfrac{1}{j_r}}}$$

(исключая эту границу) и

$$\cfrac{1}{j_1 + \cfrac{1}{j_2 + \quad \cdot \quad \cdot \quad \cdot \quad + \cfrac{1}{j_r + 1}}}$$

(включая эту границу). Пусть, кроме того, $R[-j_1, j_2, \ldots, (-1)^r j_r, 0, \ldots, 0]$ с одним или более заключительными нулями представляет собой единственное число $1/(j_1 + 1/(j_2 + 1/(\ldots + 1/(j_r + 1)) \ldots))$. Например, $R[-3] = [1/4, 1/3)$; $R[-3, 1] = (1/4, 2/7]$; $R[-3, 0] = [1/4, 1/4]$. Подобное представление является стандартным *представлением в виде непрерывной дроби* с небольшими изменениями, сделанными для согласования лексикографического порядка представления с обычным порядком.

(г) Пусть $R = (-\infty, \infty)$ и $x = n + \theta$ для любого $x \in R$, где n целое и $0 \leqslant \theta < 1$. Существует способ последовательного разбиения R на два подынтервала таким образом, чтобы на каждом шаге x было представлено $n + 1$ единицами и двоичным представлением θ, разделенными нулем, если $n \geqslant 0$, или $(-n)$ нулями и двоичным представлением θ, разделенными единицей в противном случае. Таким образом,

(3.10)
$$\begin{aligned} &R[0] = (-\infty, 0), \quad R[1] = [0, \infty); \\ &R[0, 0] = (-\infty, -1), \quad R[0, 1] = [-1, 0); \\ &R[1, 0] = [0, 1), \quad R[1, 1] = [1, \infty) \end{aligned}$$

и т. д. Это представление может быть названо «единично-двоичным» представлением x. (Заметим, что хотя мы проводили двоичные разбиения, так что каждое из них содержит бесконечные интервалы, тем не менее каждое бесконечное пересечение (3.3) будет иметь самое большее, одну общую точку. Существует бо-

лее «эффективный» способ кодирования произвольного целого n последовательностью нулей и единиц; этот вопрос изучался в [3].)

В соответствии с этими правилами числа 7/22 и $1/\pi$ имеют следующие представления:

	7/22	$1/\pi$
Десятичная система:	$\langle 3, 1, 8, 1, 8, 1, 8, 1, \ldots\rangle$	$\langle 3, 1, 8, 3, 0, 9, 8, 8, \ldots\rangle$
Двоичная система с плавающей точкой:	$\langle 1, -2, 0, 1, 0, 0, 0, 1, \ldots\rangle$	$\langle 1, -2, 0, 1, 0, 0, 0, 1, \ldots\rangle$
Непрерывная дробь:	$\langle -3, 6, 0, 0, 0, 0, 0, 0, \ldots\rangle$	$\langle -3, 7, -15, 1, -292, 1, -1, \ldots\rangle$
Единично-двоичная система:	$\langle 1, 0, 0, 1, 0, 1, 0, 0, 0, \ldots\rangle$	$\langle 1, 0, 0, 1, 0, 1, 0, 0, 0, \ldots\rangle$.

Заметим, что представления (а), (б) и (г) имеют такие последовательности $\langle j_1, j_2, j_3, \ldots\rangle$, которые соответствуют допустимым интервалам, но не представляют никакого вещественного числа; это случается, когда пересечение (3.3) пусто. Например, представление $\langle 9, 9, 9, 9, 9, \ldots\rangle$ никогда не встречается в десятичной системе счисления.

Всякий раз, когда шаг разбиения производит более двух подинтервалов, мы можем заменить его последовательностью двоичных разбиений. Например, $R = R_1 \cup R_2 \cup R_3 \cup R_4$ может быть заменено на $R = R' \cup R''$, $R' = R_1 \cup R_2$, $R'' = R_3 \cup R_4$ или на $R = R_1 \cup R'$, $R' = R_2 \cup R''$, $R'' = R_3 \cup R_4$. Единично-двоичная система (3.10) иллюстрирует, как бесконечное разбиение $R = \ldots \cup R_{-1} \cup R_0 \cup R_1 \cup \ldots$ может быть представлено в виде последовательности двоичных разбиений так, чтобы каждое R_j получалось после конечного числа делений.

Если разбиения производят самое большее две части, так что все индексы j_i ограничены значениями 0 и 1, то будем называть результирующую систему *двоично-кодированным представлением*. Из предыдущего следует, что каждая система представления может быть «расширена» до двоично-кодированной в следующем смысле: для всех конечных последовательностей целых $\langle j_1, \ldots, j_r\rangle$ найдется конечная последовательность $\langle j'_1, \ldots, j'_{r'}\rangle$ нулей и единиц, такая, что интервал $R[j_1, \ldots, j_r]$ в данной системе представления равен интервалу $R[j'_1, \ldots, j'_{r'}]$ в двоично-кодированной системе, которая расширяет данную; кроме того, отображение $\langle j_1, \ldots, j_r\rangle \to \langle j'_1, \ldots, j'_{r'}\rangle$ сохраняет лексикографический порядок элементов последовательностей.

В заключение настоящего раздела приведем несколько определений, касающихся так называемых «граней» представления; полезность этого понятия станет ясна в разд. 4. *Гранью* $\mathscr{F}$ представления является любое множество непересекающихся интервалов этого представления, объединение которых есть R.

В частности, множество всех $R[j_1, \ldots, j_r]$, таких, что r имеет фиксированное значение k, всегда является гранью, которую мы будем называть *k-гранью*. Когда общее представление отображается в двоично-кодированное представление, ее k-грани отображаются в грани двоично-кодированного представления. Образ $(k+1)$-грани, вообще говоря, является лучшим разбиением, чем образ k-грани.

Говорят, что грань $\mathscr{F}'$ в двоично-кодированном представлении покрывается другой гранью $\mathscr{F}$, если $\mathscr{F}'$ совпадает с $\mathscr{F}$, за исключением одного из интервалов $I \in \mathscr{F}'$, который разбит на два интервала $I = I' \cup I''$, где I', $I'' \in \mathscr{F}$. Каждая грань $\mathscr{F}$ двоично-кодированного представления может быть получена из последовательности граней $\mathscr{F}_0$, $\mathscr{F}_1$, $\mathscr{F}_2$, ..., где $\mathscr{F}_0 = \{R\}$ и $\mathscr{F}_{j+1}$ покрывает $\mathscr{F}_j$; если $\mathscr{F}$ конечна, то эта последовательность также конечна и заканчивается $\mathscr{F}$, в противном случае счетная грань $\mathscr{F}$ получается естественным путем как предел счетной последовательности покрывающих граней.

4. МОДЕЛИРОВАНИЕ ПРОИЗВОЛЬНЫХ РАСПРЕДЕЛЕНИЙ

Пусть $F(x)$ — монотонная неубывающая функция, заданная на всей вещественной оси, такая, что $F(-\infty) = 0$ и $F(+\infty) = 1$. Тогда F определяет распределение вещественнозначной случайной величины X, если условиться, что

$$(4.1) \qquad P(X \leqslant x) = F(x+0) \quad \text{и} \quad P(X < x) = F(x-0).$$

Как обычно, будем писать

$$(4.2) \qquad F(+\infty) = \lim_{n\to\infty} F(n), \quad F(-\infty) = \lim_{n\to\infty} F(-n),$$

$$F(x+0) = \lim_{n\to\infty} F\left(x+\frac{1}{n}\right), \quad F(x-0) = \lim_{n\to\infty} F\left(x-\frac{1}{n}\right).$$

Если $F(x+0) \neq F(x-0)$, то вероятность того, что $X = x$, отлична от нуля (и равна $F(x+0) - F(x-0)$).

Таким образом, как дискретные, так и непрерывные распределения соответствуют этой конструкции и нет необходимости требовать, чтобы F была непрерывной либо слева, либо справа.

Пусть I — некоторый интервал и $\Delta_F(I)$ — вероятность того, что $X \in I$; тогда

$$(4.3) \qquad \begin{aligned} \Delta_F((a, b)) &= F(b-0) - F(a+0), \\ \Delta_F((a, b]) &= F(b+0) - F(a+0), \\ \Delta_F([a, b)) &= F(b-0) - F(a-0), \\ \Delta_F([a, b]) &= F(b+0) - F(a-0). \end{aligned}$$

(Будем считать, что в этих формулах $F(+\infty - 0) = 1$, $F(-\infty + 0) = 0$, если a или b бесконечны.)

В данном разделе мы будем заниматься алгоритмами, которые сколь угодно точно порождают случайную величину X (в том смысле, как это объяснялось в разд. 1) в некоторой заданной системе представления для R так, как это было определено в разд. 3. В частности, для любого непустого интервала $R[j_1, \ldots, j_r]$ системы представления мы будем получать представление, начинающееся с $\langle j_1, \ldots, j_r\rangle$, с вероятностью $\Delta_F(R[j_1, \ldots, j_r])$. ПДР-деревья из разд. 2 показывают, как моделировать вероятностное распределение, соответствующее r-грани для любого отдельного r; в настоящем разделе мы рассмотрим решение этой задачи для всех r-граней одновременно.

Основная идея заключается в *уточнении* (refinement) ПДР-дерева в соответствии с процедурой интервальных разбиений. Это можно пояснить следующим образом. Предположим, что имеется оптимальное ПДР-дерево для некоторого распределения вероятностей $p_1, p_2, \ldots, p_n$, и мы хотим заменить значение x_1 двумя различными значениями x' и x'', которым соответствуют вероятности p' и p'', такие, что

$$p_1 = p' + p''. \tag{4.4}$$

В дальнейшем будет видно, что любое оптимальное ПДР-дерево для $(p_1, p_2, \ldots, p_n)$ можно «уточнить» путем замещения всех концевых узлов $\boxed{1}$ соответствующими поддеревьями, которые порождают x' и x'', так, что уточненное ПДР-дерево будет оптимальным для нового распределения вероятностей $(p', p'', p_2, \ldots, p_n)$. Если подобное оптимальное уточнение последовательно сделано для всех разбиений двоично-кодированного представления, то мы получаем дерево, которое является оптимальным одновременно для всех шагов процедуры моделирования.

Прежде чем формализовать эту идею, рассмотрим пример, иллюстрирующий процесс уточнения. Предположим, что $p = p_1$ в (2.16) должно быть заменено на $p' + p''$, где

$$\begin{aligned} p' &= 1/10 = (.000110011001100\ldots)_2, \\ p'' &= \quad\ = (.001101111110001\ldots)_2, \\ p &= 1/\pi\ = (.010100010111110\ldots)_2. \end{aligned} \tag{4.5}$$

На рис. 4 было изображено оптимальное ПДР-дерево, в котором узлы $\boxed{A}$ соответствуют единицам в двоичном представлении p; требуется заменить эти узлы поддеревьями, в которых узлы $\boxed{A'}$ и $\boxed{A''}$ соответствуют единицам в двоичных представлениях p'

и p''. Тот факт, что $p' + p'' = p$ делает это возможным: мы должны, просматривая двоичные представления слева направо, удалять узел $\boxed{A}$ всякий раз, когда в p встречается 1, и добавлять узлы $\boxed{A'}$ и (или) $\boxed{A''}$ всякий раз, когда 1 встречается в p' и (или) p''; после того как подобная процедура проделана с одним уровнем, мы переходим к следующему уровню, удваивая свободные узлы. В результате может быть получено уточ-

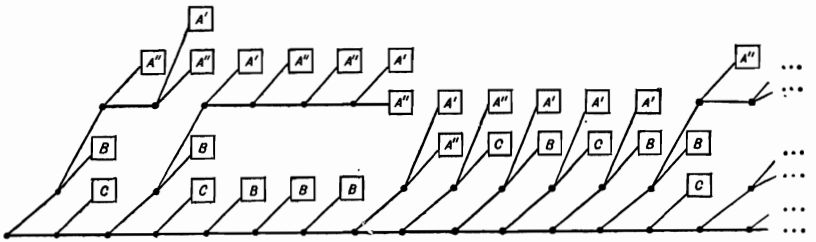
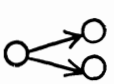


Рис. 5. Уточнение дерева, изображенного на рис. 4, с использованием вместо p_1 вероятностей p' и p'', определенных в (4.5).

ненное оптимальное ПДР-дерево, подобное тому, которое изображено на рис. 5.

При заданных $p = p' + p''$ алгоритм оптимального уточнения может быть записан следующим образом.

begin set of nodes S, T; **integer** m; **node** x;
  $m \leftarrow 0$; $S \leftarrow$ empty;
  while true do
  begin if $\varepsilon_m(p) = 1$ **then** включить узел $\boxed{A}$ уровня m в S;
    if $\varepsilon_m(p') = 1$ **then** исключить один узел из S и заменить его на $\boxed{A'}$;
    if $\varepsilon_m(p'') = 1$ **then** исключить один узел из S и заменить его на $\boxed{A''}$;
    $T \leftarrow$ empty;
    for all $x \in S$ **do**
    begin заменить x поддеревом  ; исключить x из S; включить два новых потомка x в T;
    end;
    $S \leftarrow T$; $m \leftarrow m + 1$;
  end;
end;

Поскольку S не может содержать одновременно более трех узлов, то следовало бы заменить цикл «**for all** $x \in S$» более простым правилом превращения S из одноэлементного множества в двухэлементное. (Число элементов в S во время выполнения этого цикла равно числу переносов в m-ю позицию, когда p' прибавляется к p'', следовательно, оно всегда равно 0 или 1.) Однако алгоритм представлен в таком виде, чтобы показать, что он может быть легко обобщен на случай многократного уточнения, когда p заменяется суммой произвольного числа вероятностей $p' + p'' + p''' + \ldots = p$. Результат многократного уточнения, по существу, аналогичен последовательным однократным уточнениям, однако мы можем предпочесть их одновременное выполнение.

Приведенный выше алгоритм уточнения является недетерминированным, так как он не указывает, какой узел множества S должен быть удален и заменен на $\boxed{A'}$ или $\boxed{A''}$. Таким образом, в общем случае можно найти несколько способов уточнения. Наоборот, легко видеть, что *каждое* оптимальное уточнение может быть осуществлено с помощью некоторой детерминированной версии приведенного выше алгоритма.

Частным случаем процесса уточнения является построение из разд. 2: можно положить $1 = p_1 + (1 - p_1)$, затем $1 - p_1 = p_2 + (1 - p_1 - p_2)$ и т. д. Даже когда существует бесконечно много вероятностей, подобные последовательные уточнения порождают естественным образом «в пределе» оптимальное ПДР-дерево.

Применим теперь изложенную идею уточнения к моделированию случайных величин. Предположим, что задано двоично-кодированное представление некоторого интервала R и распределение вероятностей F, сосредоточенное на R, т. е. $\Delta_F(R)$ должно быть равно 1. Определим для этого распределения и двоично-кодированного представления понятие ПР-дерева (дерева, порождающего распределение) [1]). *ПР-деревом* называется (обычно бесконечное) бинарное дерево, содержащее узлы следующих двух типов.

(а) Узлы разветвления, имеющие двух потомков и помеченные конечной (возможно, пустой) последовательностью нулей и единиц. Смысл узла разветвления таков: «получить эту последовательность нулей и единиц в качестве очередных битов представления X; затем подбросить монету и перейти к одному из потомков в зависимости от исхода бросания, равного 0 или 1».

(б) Концевые узлы, не имеющие потомков, которые помечены бесконечной последовательностью нулей и единиц. Смысл

[1]) В оригинале DG-tree (distribution generating tree). — *Прим. перев.*

концевого узла таков: «получить эту последовательность нулей и единиц в качестве остальных битов представления X».

Алгоритм, определенный ПР-деревом в соответствии с тем смыслом, который приписан его узлам, будем называть *ПР-алгоритмом.*

ПР-дерево должно гарантировать, что X имеет распределение F. Это означает, что *если* $R[j_1, \ldots, j_k]$ *— любой непустой интервал представления и если* $t(j_1, \ldots, j_k; m)$ *— число таких узлов уровня* m *дерева, что соответствующий ПР-алгоритм порождает последовательность* $\langle j_1, \ldots, j_k\rangle$ *и, возможно, дальнейшие биты, когда достигается этот узел, причем* k*-й бит порождается именно при достижении этого узла* (*а не раньше*), *то*

$$\sum_{m\geqslant 0} t(j_1, \ldots, j_k; m)/2^m = \Delta_F(R[j_1, \ldots, j_k]). \tag{4.6}$$

В содержательном плане это громоздкое утверждение сводится к следующему простому условию: при всех $j_1, \ldots, j_k$ вероятность того, что моделируемая случайная величина X имеет представление, начинающееся с $j_1, \ldots, j_k$, должна быть равна вероятности того, что случайная величина с распределением F принадлежит интервалу $R[j_1 \ldots, j_k]$.

На рис. 6, *б* приведен пример ПР-дерева, соответствующего алгоритму C из разд. 1 моделирования в двоичной системе счисления распределения $F(x) = x^2$, сосредоточенного на интервале $[0, 1)$. В этом дереве отсутствуют концевые узлы и последовательность битов, приписанная каждому узлу, состоит самое большее из одного бита.

Существует возможность получения в качестве результата выполнения ПР-алгоритма последовательности типа $(0.1111\ldots)_2$, но это происходит с вероятностью 0. С практической точки зрения мы, разумеется, не знаем, возникает ли такая аномалия, так как не видим всю выходную последовательность.

Пусть $t_k(m)$ — сумма $t(j_1, \ldots, j_k; m)$ по всем $0 \leqslant j_1, \ldots, j_k \leqslant 1$. Тогда $t_k(m)/2^m$ равно вероятности того, что ПР-алгоритм вычислит k-й бит представления X после того, как произведено ровно m подбрасываний монеты. Поскольку сумма всех $\Delta_F(R[j_1, \ldots, j_k])$ равна 1, то из (4.6) следует, что

$$\sum_{m\geqslant 0} t_k(m)/2^m = 1; \tag{4.7}$$

другими словами, при каждом фиксированном k ПР-алгоритм будет порождать k и более битов X с вероятностью 1.

Среднее время выполнения ПР-алгоритма может быть определено так же, как и в разд. 2. Пусть

$$q_k(m) = \sum_{\mu>m} t_k(\mu)/2^\mu \tag{4.8}$$

— вероятность того, что потребуется более чем m подбрасываний монеты для получения k-го бита X. Среднее число подбрасываний монеты тогда равно $\sum_{m \geqslant 0} q_k(m)$, и мы можем распространить теорему 2.1 на ПР-алгоритмы.

Теорема 4.1. *Пусть F — некоторое распределение, и пусть задана двоично-кодированная система представления над R, где*

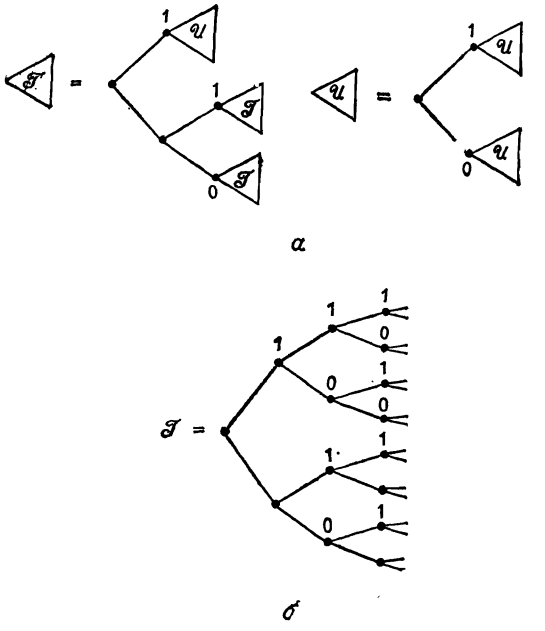


Рис. 6. ПР-дерево $\mathscr{T}$, соответствующее алгоритму C для распределения $F(x) = x^2$ в двоичной записи. В части (*а*) рисунка показана схема порождения полного дерева, начальные уровни которого изображены в части (б).

$\Delta_F(R) = 1$. *Тогда каждый ПР-алгоритм для этого представления и распределения F удовлетворяет неравенству*

$$(4.9) \qquad q_k(m) \geqslant \sum_{0 \leqslant j_1, \ldots, j_k \leqslant 1} \{2^m \Delta_F(R[j_1, \ldots, j_k])\}/2^m.$$

Кроме того, существует ПР-алгоритм, для которого при всех k и m в (4.9) имеет место равенство.

Доказательство. Неравенство (4.9) является просто частным случаем (2.9), поскольку $\Delta_F(R[j_1, \ldots, j_k])$ определяет дискретное распределение вероятностей. Остается показать, что существует ПР-дерево, для которого эта граница достигается.

Согласно теореме 2.1, это эквивалентно нахождению такого ПР-дерева, что

$$t(j_1, \ldots, j_k;\ m) = \varepsilon_m(\Delta_F(R[j_1, \ldots, j_k])) \tag{4.10}$$

при всех $j_1, \ldots, j_k$ и m.

Это ПР-дерево может быть построено индукцией по k, используя алгоритм уточнения ПДР-деревьев. Начнем с ПДР-дерева для простого распределения вероятностей $(\Delta_F(R[0]), \Delta_F(R[1]))$; это является ПДР-деревом, порождающим один бит. В ПДР-дереве, порождающем k битов, представим каждое $\Delta_F(R[j_1, \ldots, j_k])$ в виде суммы $\Delta_F(R[j_1, \ldots, j_k, 0]) + \Delta_F(R[j_1, \ldots, j_k, 1])$, получая, таким образом, ПДР-дерево, порождающее $k+1$ битов. Концевым узлам для $\Delta_F(R[j_1, \ldots, j_k, 0])$ приписан нуль, а концевым узлам для $\Delta_F(R[j_1, \ldots, j_k, 1])$ — единица.

ПР-дерево является предельным деревом, полученным с помощью подобного построения при $k \to \infty$; последовательность нулей и единиц [1]), связанная с каждым узлом ПР-дерева, является последовательностью знаков, приписанных каждому узлу в процессе уточнения. Пример построения ПР-дерева будет дан ниже. Ясно, что если условие (4.10) выполняется, то в (4.9) имеет место равенство. ■

Будем говорить, что ПР-алгоритм является *оптимальным*, если в (4.9) имеет место равенство при всех k и m. Далее, как и в теореме 2.1, мы можем охарактеризовать оптимальные деревья на языке условия, накладываемого только на k.

Следствие. *Пусть $\mathscr{A}$ есть ПР-алгоритм для распределения F и некоторого представления над R, и пусть $T_k(\mathscr{A})$ — среднее число случайных битов, необходимых для того, чтобы $\mathscr{A}$ породил по меньшей мере k результирующих битов. Тогда*

$$T_k(\mathscr{A}) \geqslant \sum_{0 \leqslant j_1, \ldots, j_k \leqslant 1} \nu(\Delta_F(R[j_1, \ldots, j_k])) \text{ [2])}. \tag{4.11}$$

Если $\mathscr{A}$ является оптимальным ПР-алгоритмом, то для всех k имеет место равенство. Обратно, если при всех k в (4.11) имеет место равенство, то $\mathscr{A}$ — оптимальный ПР-алгоритм. ■

В качестве примера этого следствия снова рассмотрим алгоритм C из разд. 1, который соответствует ПР-дереву, изображенному на рис. 6. Мы знаем из (1.5), что $T_k(C) = k + 1 - 2^{-k}$. Нижняя граница (4.11) для данного распределения в двоичном

[1]) Возможно, пустая. — *Прим. перев.*

[2]) Функция $\nu(x)$ определена в (2.11). — *Прим. перев.*

представлении может быть установлена следующим образом:

$$
\begin{aligned}
(4.12)\qquad & \sum_{0\leqslant j_1,\ \ldots,\ j_k\leqslant 1} \nu\left(\Delta_F\left(R\left[j_1,\ \ldots,\ j_k\right]\right)\right) \\
&= \sum_{0\leqslant j<2^k} \nu\left(\Delta_F\left(\left[j/2^k,\ (j+1)/2^k\right)\right)\right) \\
&= \sum_{0\leqslant j<2^k} \nu\left(((j+1)/2^k)^2-(j/2^k)^2\right) \\
&= \sum_{0\leqslant j<2^k} \nu\left((2j+1)/2^{2k}\right) \\
&= \sum_{0\leqslant j_1,\ \ldots,\ j_k\leqslant 1} \nu\left(\frac{j_1}{2^k}+\frac{j_2}{2^{k+1}}+\ \ldots\ +\frac{j_k}{2^{2k-1}}+\frac{1}{2^{2k}}\right) \\
&= \sum_{0\leqslant j_1,\ \ldots,\ j_k\leqslant 1} \left(\frac{kj_1}{2^k}+\frac{(k+1)\,j_2}{2^{k+1}}+\ \ldots\ +\frac{(2k-1)\,j_k}{2^{2k-1}}+\frac{2k}{2^{2k}}\right) \\
&= \frac{k}{2}+\frac{k+1}{4}+\ \ldots\ +\frac{2k-1}{2^k}+\frac{2k}{2^k}=k+1-2^{-k}.
\end{aligned}
$$

Таким образом, утверждение об оптимальности алгоритма C доказано.

Поучительно рассмотреть процесс построения оптимального ПР-дерева для более сложного распределения $F(x)=x^3$, $0\leqslant x\leqslant 1$, которое соответствует максимальной из *трех* равномерно распределенных случайных величин. Для этой цели мы будем использовать подход, намеченный в общих чертах при доказательстве теоремы 4.1, строя ПР-дерево путем суперпозиции соответствующих ПДР-деревьев для $k=1, 2, \ldots$. На k-м шаге существует 2^k интервалов $[j/2^k, (j+1)/2^k)$, $0\leqslant j<2^k$, и значение Δ_F для каждого интервала равно

$$
(4.13)\qquad \left(\frac{j+1}{2^k}\right)^3-\left(\frac{j}{2^k}\right)^3=\frac{3j^2+3j+1}{2^{3k}},\quad 0\leqslant j<2^k.
$$

Мы будем строить ПДР-дерево для этого распределения путем уточнения распределения, полученного на $(k-1)$-м шаге, а именно будем использовать тот факт, что

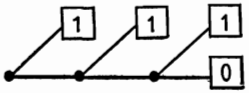
$$
(4.14)\qquad \left[\frac{j}{2^{k-1}},\ \frac{j+1}{2^{k-1}}\right)=\left[\frac{2j}{2^k},\ \frac{2j+1}{2^k}\right)\cup\left[\frac{2j+1}{2^k},\ \frac{2j+2}{2^k}\right),
$$

$$
\frac{3j^2+3j+1}{2^{3(k-1)}}=\frac{3(2j)^2+3(2j)+1}{2^{3k}}+\frac{3(2j+1)^2+3(2j+1)+1}{2^{3k}},
$$

$$
0\leqslant j<2^{k-1}.
$$

При $k = 1$ имеем простое распределение вероятностей (1/8, 7/8), так что первое ПДР-дерево выглядит следующим образом:

(4.15)

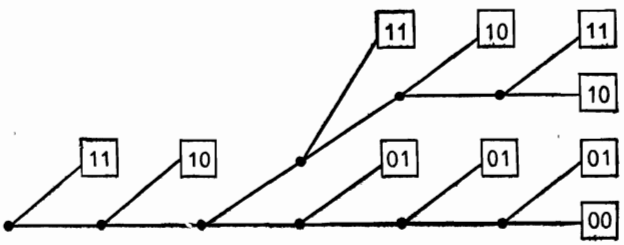


При $k = 2$ (4.14) позволяет использовать уточнения $1/8 = 1/64 + 7/64$, $7/8 = 19/64 + 37/64$, или

$$.001 = .000001 + .000111,$$
$$.111 = .010011 + .100101;$$

следовательно, (4.15) преобразуется в

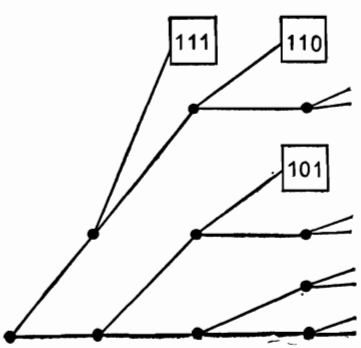
(4.16)



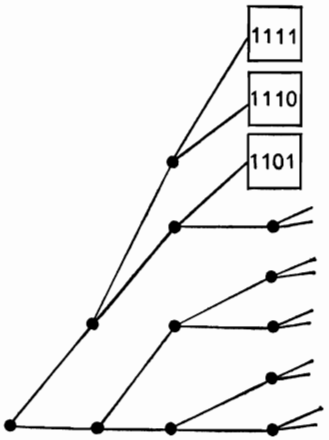
применяя алгоритм уточнения дважды: один раз к узлу [0] в (4.15), другой раз к узлам [1].

Поскольку при $k \geqslant 3$ дерево становится чрезмерно громоздким, будем изучать его структуру только до третьего уровня. ПДР-деревья при $k = 3$ и при $k = 4$ выглядят соответственно следующим образом:

(4.17)



(4.18)



при $k \geqslant 5$ все концевые узлы на первых трех уровнях пропадают. Путем суперпозиции этих ПДР-деревьев получается оптимальное ПР-дерево с размещением цепочек битов, так как это показано на рис. 7. Заметим, что на некоторых узлах встречаются по два бита, поскольку эти узлы были концевыми узлами двух ПДР-деревьев.

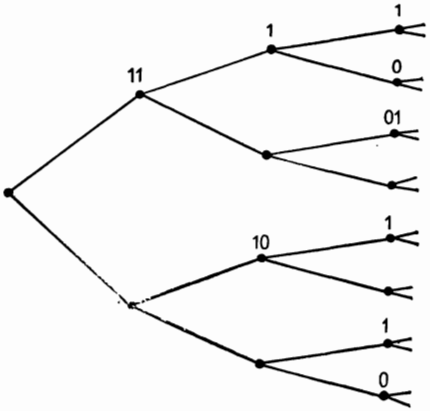


Рис. 7. Первые уровни оптимального ПР-дерева для распределения $F(x) = x^3$ в двоичном представлении.

В этом примере так же, как и в предыдущем, ПР-дерево не содержит концевых узлов. Такие узлы с приписанными им бесконечными последовательностями нулей и единиц встречаются, когда распределение F является разрывным.

Завершим настоящий раздел выяснением того, почему оптимальные ПР-деревья являются оптимальными в самом сильном смысле. Для любого заданного ПР-дерева и любой грани двоично-кодированного представления можно получить ПДР-дерево для вероятностного распределения, сосредоточенного на этой грани, непосредственной заменой узла разветвления и его преемников концевым узлом, помеченным $\langle j_1, \ldots, j_k \rangle$, всякий раз, когда $R[j_1, \ldots, j_k]$ является интервалом грани и узел разветвления завершает получение $j_1, \ldots, j_k$. В этих терминах мы

определили оптимальное ПР-дерево как такое, для которого соответствующие ПДР-деревья в каждой k-грани являются оптимальными для соответствующих им дискретных распределений. Но k-грани включают все узлы, следовательно, каждый концевой узел, помеченный $\langle j_1, \ldots, j_k \rangle$, встречается надлежащее число раз (а именно, $\varepsilon_m(\Delta_F(R[j_1, \ldots, j_k]))$ раз) на каждом уровне m, и отсюда следует, что ПДР-деревья, полученные из оптимального ПР-дерева в *любой* грани двоично-кодированного представления, являются оптимальными. В частности, оптимальному ПР-дереву будет соответствовать оптимальный алгоритм в *любой* системе представления, которая преобразуется в двоично-кодированное представление так, как это обсуждалось в разд. 3.

Другим следствием подобной сильной оптимальности является тот факт, что процедура построения из теоремы 4.1 порождает *все* оптимальные ПР-деревья.

Теорема 4.2. *ПР-дерево для данного распределения и двоично-кодированного представления оптимально тогда и только тогда, когда ПДР-дерево, соответствующее каждой грани $\mathscr{F}$, является оптимальным уточнением ПДР-дерева, соответствующего каждой грани, покрываемой $\mathscr{F}$.*

Доказательство. Предположим, что $\mathscr{F}$ покрывает $\mathscr{F}'$, где интервал I для $\mathscr{F}'$ разбит на интервалы I' и I'' для $\mathscr{F}$. ПДР-дерево, соответствующее $\mathscr{F}$ (обозначим его через $\mathscr{T}(\mathscr{F})$), может быть получено из $\mathscr{T}(\mathscr{F}')$ заменой всех концевых узлов, соответствующих I, поддеревьями с концевыми узлами, соответствующими I' и I''. Если $\mathscr{T}(\mathscr{F})$ не является оптимальным уточнением $\mathscr{T}(\mathscr{F}')$, то как $\mathscr{T}(\mathscr{F})$, так и $\mathscr{T}(\mathscr{F}')$ не могут быть оптимальными ПДР-деревьями. ■

5. ОПТИМАЛЬНОЕ СРЕДНЕЕ ВРЕМЯ ВЫПОЛНЕНИЯ АЛГОРИТМА

После того как мы установили существование (возможно, бесконечных[1])) алгоритмов, которые являются оптимальными в сильном смысле, выясним, насколько они хороши на самом деле. В относительном смысле они минимизируют среднее значение r-й степени времени выполнения алгоритма при всех r. Оказывается, что они всегда хороши и в абсолютном смысле, по крайней мере в среднем: какое бы распределение мы не рассматривали, после небольшой начальной задержки результат производится с такой же скоростью, с какой данные поступают на вход алгоритма.

---

[1]) То есть с бесконечным средним временем выполнения. — *Прим. перев.*

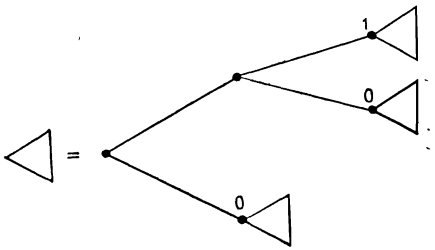
Теорема 5.1. *Для любого $k \geqslant 1$, любых распределений и двоично-кодированных представлений оптимальный ПР-алгоритм доставит k битов X после ввода в среднем менее чем $k + 2$ случайных битов.*

Доказательство. На самом деле среднее число случайных битов не превышает $k + 2 - 2^{1-k}$, поскольку в теореме 2.3 можно положить $n = 2^k$. ■

Эта теорема с очевидными изменениями справедлива для других представлений; например, в системе счисления по основанию d мы сможем получить k результирующих цифр в среднем после менее чем $k \log_2 d + 2$ шагов выполнения алгоритма. В частности, мы можем имитировать k бросаний кости в среднем с помощью менее чем $k \log_2 6 + 2$ подбрасываний монеты.

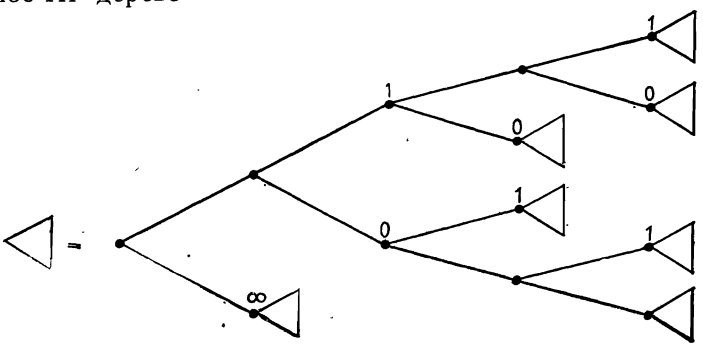
В известном смысле теорема 5.1 представляет собой удивительный результат. Например, предположим, что мы хотим получить случайную последовательность нулей и единиц таким образом, чтобы каждый бит принимал значение 1 с вероятностью 1/4 независимо от других битов этой последовательности. Интуитивно кажется, что наилучшим способом решения этой задачи является использование следующего ПР-дерева:

(5.1)



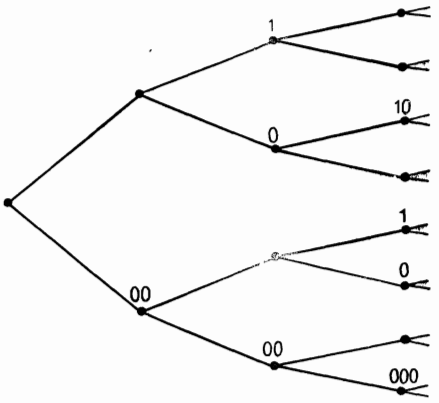
Однако подобная процедура требует в среднем $\frac{3}{2}k$ шагов для получения k результирующих битов, что в соответствии с теоремой, весьма далево от оптимального значения. Другое более сложное ПР-дерево

(5.2)



требует приблизительно $\frac{9}{8}k$ шагов, а начальная часть оптимального ПР-дерева для этой задачи выглядит следующим образом:

(5.3)



Выше мы вычислили среднее число подбрасываний монеты, необходимых для моделирования распределения $F(x)=x^2$ на $[0, 1)$ в двоичной системе счисления, и в результате получили $k+1-2^{-k}$ (см. (4.12)). Если же мы попытаемся воспользоваться аналогичной техникой для распределения $F(x)=x^3$, то очень скоро столкнемся с серьезными трудностями: сумма

$$\sum_{0\leqslant j<2^k} \nu\left(\frac{3j^2+3j+1}{2^{3k}}\right) \tag{5.4}$$

непредставима в какой-нибудь простой замкнутой форме как функция k. К счастью, существует выход из этого положения; следующая теорема показывает, как можно вычислить асимптотическое значение во многих представляющих интерес случаях.

Теорема 5.2. *Пусть $F(x)$ — функция распределения на $[0, 1)$, и предположим, что существует производная $F'(x)=f(x)$, где $f(x)$ ограничена и уравнение $f(x)=y$ имеет ограниченное число решений x для каждого $y>0$* [1]. *Пусть $T_F(k)$ — среднее время (в терминах случайных подбрасываний монеты) выполнения оптимального ПР-алгоритма для распределения F с использованием двоичной системы счисления. Тогда*

$$T_k(k)=k+d_F+O(k^2/2^k), \tag{5.5}$$

[1]) Впрочем, из приведенного ниже доказательства следует, что достаточно требовать, чтобы $f(x)$ была функцией ограниченной вариации на $[0, 1)$. — *Прим. перев.*

где асимптотическая задержка

$$d_F = \sum_m \frac{m}{2^m} \int_0^1 \varepsilon_m(f(x))\,dx. \tag{5.6}$$

(*Функция* $\varepsilon_m(x)$ *определена в* (2.13), *интеграл* $\int_0^1 \varepsilon_m(f(x))\,dx$ *представляет собой меру множества* $\{x \mid m$*-й бит в двоичном представлении* $f(x)$ *равен* $1\}$, *где под «*m*-м битом» понимается коэффициент при* 2^{-m}. *Константа, заключенная в символ* O *в* (5.5). *зависит только от ограничений на* $f(x)$.)

Доказательство. Имеем

$$T_F(k) = \sum_{0 \leqslant j < 2^k} v(A_{kj}),$$

где

$$A_{kj} = \Delta_F\left(\left[\frac{j}{2^k}, \frac{j+1}{2^k}\right)\right) = \int_{j/2^k}^{(j+1)/2^k} f(x)\,dx. \tag{5.7}$$

Следовательно,

$$T_F(k) = \sum_{m \geqslant 1} \sum_{0 \leqslant j < 2^k} \frac{m}{2^m} \varepsilon_m(A_{kj}). \tag{5.8}$$

Так как

$$\sum_{m \geqslant 1} \varepsilon_m(A_{kj})/2^m = A_{kj},$$

то

$$T_F(k) - k = \sum_{m \geqslant 1} \frac{m-k}{2^m} \sum_{0 \leqslant j < 2^k} \varepsilon_m(A_{kj}) = \sum_{m \geqslant 1-k} \frac{m}{2^m} J_m(k), \tag{5.9}$$

где

$$J_m(k) = \frac{1}{2^k} \sum_{0 \leqslant j < 2^k} \varepsilon_{m+k}(A_{kj}). \tag{5.10}$$

Пусть $\sup_{x \in [0,\ 1)} f(x) = 2^\mu$, тогда на основании (5.7) $A_{kj} \leqslant 2^{\mu-k}$ и, следовательно, $J_m(k) = 0$ для любого k, когда $m < -\mu$.

Рассмотрим теперь некоторое фиксированное m. Из нашего предположения относительно f следует, что следующее условие имеет место для всех, за исключением $O(2^m)$, значений j из промежутка $0 \leqslant j < 2^k$:

$$\lfloor 2^m f(x) \rfloor = \lfloor 2^m f(y) \rfloor \quad \text{для всех} \quad x, y \in \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right]. \tag{5.11}$$

Для всех таких значений j имеем

$$\varepsilon_{m+k}(A_{kj}) = \varepsilon_{m+k}\left(\int_{j/2^k}^{(j+1)/2^k} f(x)\,dx\right) = \varepsilon_m(f(x))$$

при любом $x \in \left[\frac{j}{2^k}, \frac{j+1}{2^k}\right]$;

$$\frac{1}{2^k}\varepsilon_{m+k}(A_{kj}) = \int_{j/2^k}^{(j+1)/2^k} \varepsilon_m(f(x))\,dx.$$

Следовательно,

$$J_m(k) = \int_0^1 \varepsilon_m(f(x))\,dx + O\left(2^{m-k}\right),$$

а поскольку $0 \leqslant J_m(k) \leqslant 1$, то на самом деле

$$J_m(k) = \int_0^1 \varepsilon_m(f(x))\,dx + O\left(\min\left(1, 2^{m-k}\right)\right). \tag{5.12}$$

Подставляя это значение в (5.9), имеем

$$T_F(k) = k + d_F + \sum_{-\mu \leqslant m \leqslant k} \frac{m}{2^m} O\left(2^{m-k}\right) + \sum_{m>k} \frac{m}{2^m} O(1),$$

откуда немедленно следует (5.5). ■

Из проведенного доказательства ясно, что (5.5) справедливо также и в других случаях, например когда существует конечное число точек $0 = x_0 < x_1 < \ldots < x_n = 1$, таких, что $F'(x) = f(x)$ существует и $f(x)$ является ограниченной и монотонной на $[x_{i-1}, x_i]$ функцией при $1 \leqslant i \leqslant n$; распределение F может иметь скачки в x_i, однако достаточно, чтобы оно было «кусочно-дифференцируемым» в указанном выше смысле. С другой стороны, как мы увидим позже, (5.5) не справедливо в общем случае, когда существует бесконечно много точек, в которых F не дифференцируемо.

Применим вначале теорему 5.2 к распределению $F(x) = x^3$. Интеграл в (5.6)

$$\int_0^1 \varepsilon_m(3x^2)\,dx = \int_0^3 \varepsilon_m(y)\,dy/\sqrt{12y}\,.$$

При $m = -1$ он равен $\int_2^3 dy/\sqrt{12y} = 1 - \sqrt{2/3}$; при $m \geqslant 0$ он равен

$$\sum_{1 \leqslant j \leqslant 3\cdot 2^{m-1}} \int_{(2j-1)/2^m}^{2j/2^m} dy/\sqrt{12y} = \frac{1}{2^{m/2}\sqrt{3}} \sum_{1 \leqslant j \leqslant 3\cdot 2^{m-1}} (\sqrt{2j} - \sqrt{2j-1}). \tag{5.13}$$

Пусть $g(n)=\sqrt{1}+\sqrt{2}+\ \dots\ +\sqrt{n}$; тогда

$$(5.14)\qquad 2\sqrt{2}\,g(n)-g(2n)+\sqrt{2n}=$$
$$=(\sqrt{2}-\sqrt{1})+(\sqrt{4}-\sqrt{3})+\ \dots\ +(\sqrt{2n}-\sqrt{2n-1}).$$

Используя формулу суммирования Эйлера,

$$(5.15)\qquad g(n)=\tfrac{2}{3}n^{3/2}-\tfrac{1}{2}\sqrt{n}+c+O(1/\sqrt{n}),$$

когда $n\to\infty$, с некоторой константой c (которая оказывается равной $\zeta(-1/2)=-\zeta(3/2)/4\pi=-.207886224977354566$; см. [8], упр. 6.1—8), следовательно,

$$(5.16)$$
$$h(n)=\sum_{1\leqslant j\leqslant n}(\sqrt{2j}-\sqrt{2j-1})=\sqrt{n/2}+(2\sqrt{2}-1)\,c+O(1/\sqrt{n}).$$

Для больших m мы, таким образом, доказали, что $\int_0^1 \varepsilon_m(3x^2)\,dx=$

$$=\tfrac{1}{2}+(2\sqrt{2}-1)\ \text{с}\ 2^{-m/2}/\sqrt{3}+O(2^{-m}).$$

Используя это асимптотическое представление, можно вычислить значение

$$(5.17)\qquad d_F=\sum_m \frac{m}{2^m}\int_0^1 \varepsilon_m(f(x))\,dx=$$
$$=-2(1-\sqrt{2/3})+\frac{1}{\sqrt{3}}\sum_{m\geqslant 1}\frac{m}{2^{3m/2}}\,h(3\cdot 2^{m-1})$$

весьма точно; оно равно 0.465825512311. Заметим, что поэтому распределение $F(x)=x^3$ можно моделировать быстрее, чем кажущееся более простым распределение $F(x)=x^2$.

Рассмотрим далее до некоторой степени патологические распределения [1]). Пусть

$$(5.18)\qquad F(x)=\begin{cases}\tfrac{1}{2}F(4x), & 0\leqslant x<\tfrac{1}{4},\\ \tfrac{1}{2}, & \tfrac{1}{4}\leqslant x<\tfrac{3}{4},\\ \tfrac{1}{2}+\tfrac{1}{2}F(4x-3), & \tfrac{3}{4}\leqslant x<1.\end{cases}$$

[1]) Рассматриваемые примеры распределений носят название сингулярных распределений канторовского типа; см., например, Феллер В. Введение в теорию вероятностей и ее приложения, т. 2. — М.: Мир, 1967, 54—55. — *Прим. перев.*

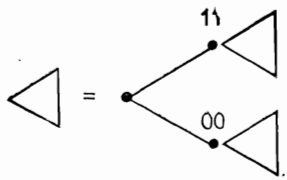
Подобные функциональные значения становятся более понятными с использованием двоичной записи: если через α обозначить бесконечную цепочку нулей и единиц, не содержащую на конце бесконечно много нулей, то

$$
\begin{aligned}
&F(0\ 0\ \alpha) = 0\ F(\alpha),\\
&F(0\ 1\ \alpha) = 1\ 0\ 0\ 0\ \ldots,\\
&F(1\ 0\ \alpha) = 1\ 0\ 0\ 0\ \ldots,\\
&F(1\ 1\ \alpha) = 1\ F(\alpha).
\end{aligned}
\tag{5.19}
$$

Легко видеть, что $F(x)$ является непрерывной неубывающей функцией, которая дифференцируема почти везде в $[0, 1]$. Производная не существует только в таких точках, двоичные разложения которых порождаются с помощью обобщенного регулярного выражения $(00+11)^*010^\infty+(00+11)^\infty$, причем всюду, где производная существует, это выражение равно нулю. Тем не менее эта функция непрерывно возрастает от 0 до 1.

Оптимальное ПР-дерево для распределения (5.18) в действительности очень простое,

(5.20)



и время, необходимое для получения k выходных битов, в точности равно $\lceil k/2 \rceil$ с нулевой дисперсией.

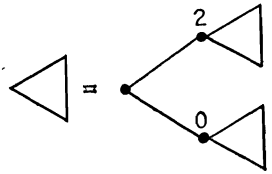
Можно также рассмотреть функцию, связанную с «тернарным множеством Кантора», которая часто приводится в учебниках в качестве примера монотонной непрерывной функции, производная которой равна нулю почти везде:

$$
F(x) = \begin{cases} \frac{1}{2}F(3x), & 0 \leqslant x < \frac{1}{3},\\ \frac{1}{2}, & \frac{1}{3} \leqslant x < \frac{2}{3},\\ \frac{1}{2} + \frac{1}{2}F(3x-2), & \frac{2}{3} \leqslant x < 1. \end{cases}
\tag{5.21}
$$

В этом случае оптимальное ПР-дерево для моделирования *тернарного* (по основанию 3) представления случайной величины X

имеет вид

(5.22)



Интуитивно ожидается, что k битов двоичного представления X будут порождаться с помощью приблизительно $k/\log_2 3$ подбрасываний монеты на один выходной бит. Это может быть доказано следующим образом. Пусть $l = \lfloor k/\log_2 3 \rfloor$, так что $-3^{-l-1} < 2^{-k} < 3^{-l}$. Разобьем множество целых j, таких, что $0 \leqslant j < 2^k$, на $3^{l-1}+1$ классов с номерами от 0 до 3^{l-1}, где класс с номером n состоит из таких j, для которых $(3n-2)/3^l < j/2^k < (3n+1)/3^l$. Каждому классу принадлежит самое большее девять значений j. Сумма вероятностей в классе с номером n равна $F((3n+1)/3^l) - F((3n-2)/3^l)$, и нетрудно убедиться, что эта сумма равна либо 2^{-l}, либо 0 для каждого n. Следовательно, она отлична от нуля в точности для 2^l значений n. Для таких n на основании (2.21)

$$\sum_{p \in \text{класс } n} \nu(p) = \frac{1}{2^l} \sum_{p \in \text{класс } n} ((2^l p)\, l + \nu(2^l p))$$

$$= \frac{l}{2^l} + \frac{1}{2^l} \sum_{p \in \text{классу } n} \nu(2^l p)$$

и последняя сумма, согласно теореме 2.3, не превосходит 5. Поэтому для распределения (5.21)

$$\lfloor k/\log_2 k \rfloor \leqslant T_F(k) \leqslant \lfloor k/\log_2 3 \rfloor + 5. \tag{5.23}$$

Другим примером распределения подобного общего типа является $F(x) = F_0(x)$, где

(5.24)

$$F_j(x) = \begin{cases} \frac{1}{2} F_{j+1}(2^{2j+1}x), & 0 \leqslant x < 2^{-2j-1}, \\ \frac{1}{2}, & 2^{-2j-1} \leqslant x < 1 - 2^{-2j-1}, \\ \frac{1}{2} + F_{j+1}(2^{2j+1}x - 2^{2j+1} + 1), & 1 - 2^{-2j-1} \leqslant x < 1. \end{cases}$$

В этом случае оптимальный ПР-алгоритм требует $\lceil \sqrt{k} \rceil$ шагов для получения k-го бита, он порождает $2j-1$ идентичных

выходных битов после j-го подбрасывания монеты. Подобным же образом можно построить непрерывные функции распределения, для которых время выполнения алгоритма является функцией от k, возрастающей сколь угодно медленно.

Все приведенные примеры ПР-деревьев предполагали непрерывность $F(x)$, поэтому интересно рассмотреть разрывное распределение

$$F(x)=\begin{cases}\frac{1}{4}F(2x), & 0\leqslant x<\frac{1}{2},\\ \frac{3}{4}+\frac{1}{4}F(2x-1), & \frac{1}{2}\leqslant x<1.\end{cases} \tag{5.25}$$

Эта функция распределения может быть названа «заикающейся» функцией, поскольку двоичное представление $F(x)$ совпадает с двоичным представлением x, в котором каждый бит повторяется дважды. Например, $F(.010100010111\ldots)=.00\,11\,00\,11\,00\,00\,00\,11\,00\,11\,11\,11\ldots$ в двоичной форме. Эта функция имеет разрывы, равные 2^{-2j+1} в каждой двоично-рациональной точке вида $x=q/2^j$, где q нечетно. Таким образом, функция $F(x)$ непрерывна, за исключением двоично-рациональных точек, и имеет нулевую производную почти везде. Кроме того, существуют иррациональные точки, в которых производная не существует; в качестве примера одной из таких точек укажем точку

$$x_0=1-2^{-1}-2^{-2}-2^{-4}-2^{-8}-2^{-16}-\ldots=1-\sum_{j\geqslant 0}2^{-2^j}. \tag{5.26}$$

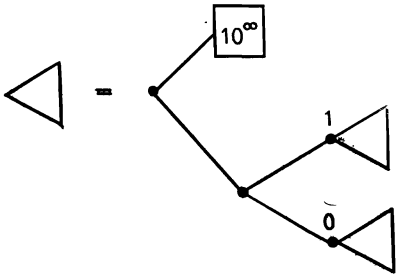
Если $\varepsilon=2^{-2^j}$ для любого $j\geqslant 1$, то $F(x_0+\varepsilon)-F(x_0)=3\varepsilon^2$, но $F(x_0+2\varepsilon)-F(x_0)=2\varepsilon+4\varepsilon^2$; таким образом, в точке x_0 производная не существует. (Между прочим, в этой точке имеет место курьезное равенство $F(x_0)=3x_0-\frac{1}{2}$.)

Величины $A_{jk}=\Delta_F([j/2^k,\ (j+1)/2^k))$, $0\leqslant j<2^k$, включают в себя 2^{k-i} вхождений $(2^{2i-1}+1)/2^{2k}$, $1\leqslant i\leqslant k$, и одно вхождение $1/2^{2k}$. (Напомним, что значение $F(j/2^k-0)$ должно быть использовано в этом вычислении, а $F(j/2^k)$ — нет.) Таким образом, среднее время, необходимое для моделирования k битов, может быть вычислено точно и оно равно

$$\sum_{0\leqslant j<2^k}\nu(A_{jk})=3-3/2^k. \tag{5.27}$$

Оптимальное ПР-дерево для распределения (5.25) оказывается очень простым:

(5.28)



Заметим, что значение x_0 в (5.26) имеет отмеченные свойства по отношению к F, но не к данному дереву.

Приведенные примеры показывают необходимость по крайней мере некоторых ограничительных предположений относительно $F(x)$ в теореме 5.2.

6. МОНОТОННЫЕ АЛГОРИТМЫ МОДЕЛИРОВАНИЯ

В разд. 1 мы рассмотрели метод «квадратного корня» для моделирования случайной величины X с функцией распределения $F(x) = x^2$. Этот частный метод требовал $2k$ входных битов для получения k битов результата. Можно, однако, указать более сложное правило извлечения квадратного корня, которое требует ровно столько битов, сколько действительно необходимо для получения k-го бита результата при любом k. Например, если $U = (.010\alpha)_2$, то всегда $\sqrt{U} = (.100\ldots)_2$ независимо от последующих битов α.

Подобная идея приводит к ПР-дереву, которое начинается так, как это показано на рис. 8. С первого взгляда может показаться, что это дерево должно быть оптимальным, однако среднее число входных битов, необходимых для получения третьего бита результата оказывается равным 4, в то время как алгоритм C требует для этой цели только $3^7/_8$ бита. Причина этого состоит в том, что нашей целью является не вычисление квадратного корня из равномерно распределенной случайной величины — задача заключается в моделировании случайной величины X, которая имеет такое же *распределение*, как и квадратный корень из равномерно распределенной случайной величины.

Из сравнения рис. 8 с рис. 6 становится ясно, в чем рис. 8 уступает рис. 6: результат 101 встречается только один раз на 3-м уровне дерева на рис. 6, в то время как он встречается

дважды на 4-м уровне дерева, изображенного на рис. 8. После небольшого размышления становится понятным, почему рис. 6 является более выигрышным: рис. 8 требует *монотонности* ПР-дерева в том смысле, что все цепочки, порождаемые путями на левых (т. е. нижних) ветвях, должны быть лексикографически меньшими или равными по отношению ко всем цепочкам, порождаемым путями на правых (т. е. верхних) ветвях. Для дерева, изображенного на рис. 6, этого не требуется.

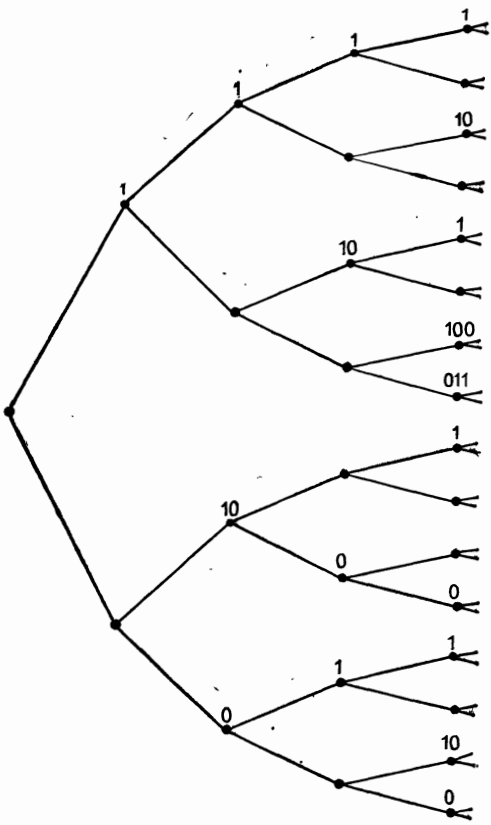


Рис. 8. Монотонное ПР-дерево для извлечения квадратного корня из двоичного числа между 0 и 1.

Существует единственное оптимальное монотонное ПР-дерево для каждого распределения и представления. Заметим, что подобные оптимальные деревья являются оптимальными для вычисления $F^{-1}(U)$, исходя из двоичного представления U; они представляют интерес с точки зрения сложности вычислений вне зависимости от того, используем ли мы их для моделирования случайных величин. Определение точного среднего времени выполнения оптимального алгоритма «квадратного корня», когда усреднение проводится по всем возможным последовательностям случайных битов, представляется трудным, однако можно показать, что среднее число входных битов, необходимых для получения k-го бита результата вычисления квадратного корня, меньше, чем $k+3$. В действительности подобная граница справедлива и в общем случае.

Теорема 6.1. *Для любого распределения и двоично-кодированного представления оптимальное монотонное ПР-дерево тре-*

бует для порождения k битов результата в среднем менее чем на 2 входных бита больше по сравнению с оптимальным ПР-деревом.

Доказательство. Среднее время получения k-го бита результата равно $\sum L(v)/2^{L(v)}$, где суммирование осуществляется по всем узлам v, для которых k-й бит является результирующим и $L(v)$ означает уровень узла v. Каждый узел v связан с некоторым интервалом I k-го уровня представления. Ниже мы докажем, что для каждого I существует не более чем 2 связанных с ним узла v любого данного уровня. Этого достаточно для доказательства теоремы, поскольку расход времени, связанный с I, может быть записан как

$$\nu(\Delta_I) + \nu(\Delta_I'), \quad \text{где} \quad \Delta_I + \Delta_I' = \Delta_F(I);$$

отсюда на основании (2.23) и (2.25) имеем

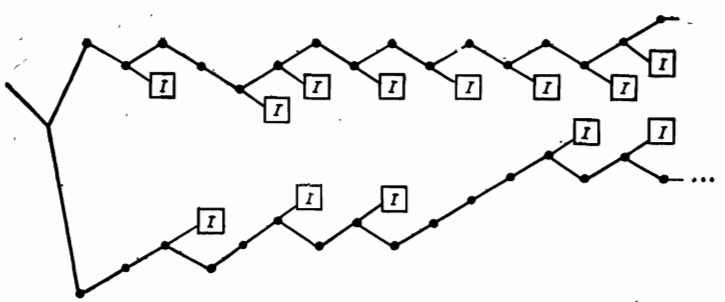
$$\nu(\Delta_F(I)) \leqslant \nu(\Delta_I) + \nu(\Delta_I') \leqslant \nu(\Delta_F(I)) + 2\Delta_F(I). \tag{6.1}$$

Суммируя неравенство по всем I и замечая, что для некоторых I выполняется строгое неравенство, получаем требуемый результат.

Рассмотрим узлы v, связанные с интервалом I; пусть $\Delta_F(I) = x - y$, где $x = F(b \pm 0)$ и $y = F(a \pm 0)$, как в (4.3). Предположим, что $x > y$. Мы собираемся отобразить все входы, которые $> y$ и $< x$, в интервал I. (Входы, *равные* x или y несущественны, так как они имеют бесконечную точность и, следовательно, встречаются с вероятностью нуль.) Рассмотрим, например, значения

$$x = .01101001010011010101011 \ldots,$$
$$y = .01101000110110101111010 \ldots.$$

Тогда узлы $\boxed{I}$ оптимального монотонного ПР-дерева, продолжающего путь для 0110100, будут располагаться примерно так:



В общем случае после первого бита, в котором различаются x и y, найдется узел $\boxed{I}$ для 0, соответствующий каждому последующему, равному 1 биту x, и узел $\boxed{\ }^{\neg}$ для 1, соответствующий каждому последующему, равному 0 биту y. (Если y — конечная двоичная дробь, то должно быть внесено небольшое изменение: следует заменить «10^∞» в конце y на «01^∞», чтобы сделать сформулированное правило справедливым в общем случае.) Ясно, что на один уровень приходится не более двух узлов $\boxed{I}$ или не более одного, если $y = 0$, как и требовалось. ■

Следствие. *Пусть $G(x)$ — любая монотонная возрастающая функция из $[0, 1]$ в $[0, 1]$. Тогда в предположении равномерного распределения значений x k-й бит $G(x)$ может быть вычислен с использованием в среднем менее чем $k + 4$ битов x.*

Доказательство. Пусть $F(x) = G^{-1}(x)$ — функция с произвольными значениями, сохраняющая монотонность, когда $G^{-1}(x)$ не определено. Примените теоремы 5.1 и 6.1. ■

В частности, существует способ вычисления с помощью достаточно искусного алгоритма таких функций, как $1/(2 - \exp \times \times (-\pi^2 \operatorname{tg} \sqrt{x}))$ почти в «реальное время».

Рассмотрим теперь оптимальное монотонное ПР-дерево в простом случае, для которого время выполнения соответствующего алгоритма может быть вычислено точно. Предположим, что θ — вещественное число между 0 и 1, и допустим, что мы хотим вычислить $\theta + x$ в двоичной форме, исходя из заданного двоичного значения x. Для того чтобы оставаться в интервале $[0, 1]$, мы будем вычислять $\frac{1}{2}(\theta + x)$, так как это значение имеет столько же двоичных цифр (только сдвинутых). Следующий алгоритм, реализующий правило сложения «слева направо», доставляет результат настолько быстро, насколько это возможно, в предположении, что $x = (.x_1x_2x_3\ldots)_2$ и что $\theta = (.\theta_1\theta_2\theta_3\ldots)_2$ имеет бесконечное двоичное разложение.

```
begin integer j, m;
    m ← 1;  j ← 0;
    while true do
    begin
        if θm + xm = 1 then j ← j + 1
        else begin if θm + xm = 0 then output (01^j)
                                  else output (10^j);
                  j ← 0;
             end;
```

```
        m ← m + 1;
    end;
end;
```

Среднее число требуемых входных битов, прежде чем этот алгоритм произведет k битов результата, равно T_{k0}, где

$$T_{kj} = \begin{cases} 0, & k \leqslant 0, \\ 1 + \frac{1}{2} T_{k-j-1,0} + \frac{1}{2} T_{k,\,j+1}, & k > 0. \end{cases} \tag{6.2}$$

Решением, которое удовлетворяет неравенству $0 \leqslant T_{kj} \leqslant T_{k0}$, является $T_{kj} = \min(k + 1 - j,\ 2)$ при $k \geqslant 1$; таким образом, среднее значение равно $k + 1$.

Можно показать, что для оптимального ПР-дерева, соответствующего функции распределения $F(x) = 2x - \theta$, $\frac{1}{2}\theta \leqslant x \leqslant \frac{1}{2}(\theta + 1)$, $T_F(k) = k - 1 + 2^{2-k}$. (Доказательство: Ненулевые числа A_{kj} для $0 \leqslant j < 2^k$ включают значение 2^{1-k} точно $2^{k-1} - 1$ раз плюс две бесконечные двоичные дроби, которые складываются с 2^{1-k}; таким образом, $\sum \nu(A_{kj}) = 2^{1-k}((2^{k-1} - 1)(k - 1) + k + 1)$). Поэтому константа 2 в теореме 6.1 является наилучшей даже в этом простом случае.

7. ГЕНЕРАТОРЫ С КОНЕЧНЫМ ЧИСЛОМ СОСТОЯНИЙ

Мы убедились, что существуют очень эффективные алгоритмы для произвольных распределений в произвольных представлениях, где «эффективность» понимается в том смысле, что при этом требуется сравнительно немного случайных битов. Однако большинство алгоритмов, которые достигают оптимальных границ, являются очень сложными и требуют чрезмерного (вплоть до бесконечного) объема памяти для хранения ПР-деревьев. Поэтому хотелось бы рассмотреть ограниченные классы алгоритмов, которые допускают относительно простую программную реализацию и требуют сравнительно мало памяти.

В настоящем разделе мы изучим алгоритмы, которые могут быть реализованы с помощью автоматов с конечным числом состояний; алгоритм C из разд. 1 представляет собой наилучший пример подобного алгоритма (см. рис. 6). Мы частично получим ответы на следующие вопросы: (а) какие распределения можно моделировать с помощью подобных алгоритмов? (б) насколько близко к оптимальному времени $T_F(k)$ моделирования

распределения F мы можем приблизиться, используя подобные алгоритмы?

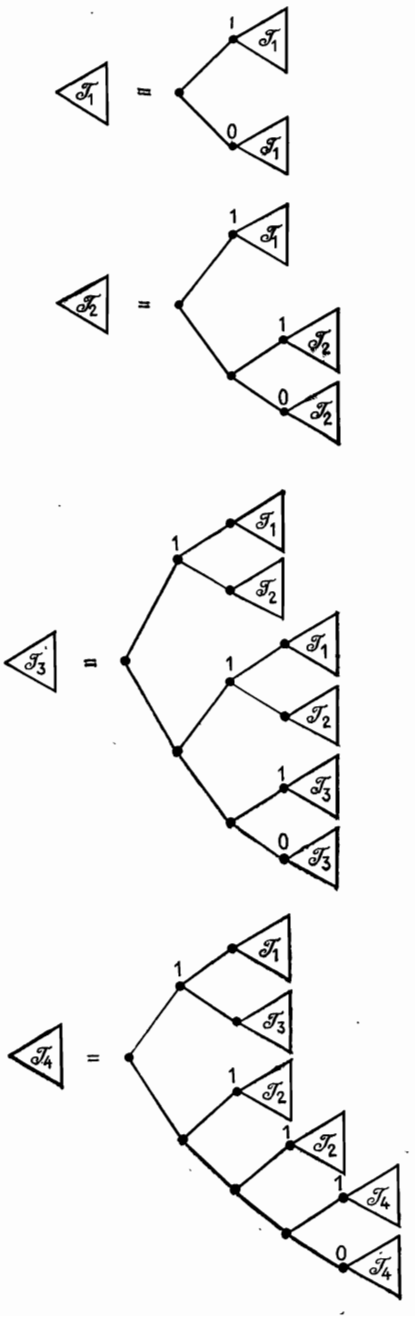
На протяжении всего раздела мы ограничимся рассмотрением распределений $F(x)$, сосредоточенных на $[0, 1)$, в двоичной системе счисления. *Генератором с конечным числом состояний* (сокр. ГКС) для F называется ПР-алгоритм моделирования F, соответствующее которому дерево имеет конечное число неизоморфных поддеревьев, причем два дерева считаются изоморфными, если они отличаются друг от друга разве что перестановкой левых и правых ветвей в узлах разветвления. Например, бесконечное ПР-дерево, изображенное на рис. 6, *б*, имеет только шесть неизоморфных поддеревьев, как это показано на рис. 6, *а*, а именно поддерево $\mathscr{T}$ с (или без) 0 или 1 в его корне, поддерево $\mathscr{U}$ с 0 или 1 в его корне и другие потомки $\mathscr{T}$. (На практике только три из этих шести типов поддеревьев соответствовали бы различным состояниям машинной программы, так как выходные цепочки могли бы быть связаны с переходами между состояниями.)

ГКС-модель вычислений тесно связана с хорошо разработанными теориями вероятностных конечных автоматов и с проблемой реализуемости ввода-вывода в цепях Маркова (см., например, [9]). Связь ГКС с генерированием случайных чисел представляет собой новое направление этой теории.

В настоящем разделе мы покажем, что каждое из распределений $F(x) = x^n$ при $n \geqslant 3$ может быть порождено с помощью подходящего ГКС, однако ни один из них не будет являться оптимальным ПР-алгоритмом. С другой стороны, для любого заданного $\varepsilon > 0$ мы укажем ГКС для полиномиального распределения $F(x)$, среднее время получения с помощью которого k битов результата в пределах ε будет совпадать с оптимальным средним временем $T_F(k)$. В заключение мы покажем, что некоторые важные распределения (такие, как экспоненциальное и нормальное) не могут быть порождены с помощью никакого ГКС.

Вначале для моделирования $F(x) = x^n$ мы воспользуемся техникой, по существу подобной той, которая в [1], гл. 2, приписывается Г. Гашюцу. Идея этого подхода состоит в моделировании максимальной из n независимых равномерно распределенных случайных величин. Дерево $\mathscr{T}_n$ начинается с использования ПДР-дерева для вероятностей $\binom{n}{j} 2^{-n}$ того, что старшие биты n моделируемых равномерно распределенных случайных величин содержит ровно j единиц. Если $j = 0$, то выводится 0 и продолжается дерево $\mathscr{T}_n$, в противном случае выводится 1 (как только обнаружено, что $j \neq 0$) и продолжается дерево $\mathscr{T}_j$.

Например,


$\mathcal{T}_1$ = 1 $\mathcal{T}_1$ 0 $\mathcal{T}_1$
$\mathcal{T}_2$ = 1 $\mathcal{T}_1$ 1 $\mathcal{T}_2$ 0 $\mathcal{T}_2$
$\mathcal{T}_3$ = 1 $\mathcal{T}_1$ $\mathcal{T}_2$ 1 $\mathcal{T}_1$ $\mathcal{T}_2$ 1 $\mathcal{T}_3$ 0 $\mathcal{T}_3$
$\mathcal{T}_4$ = 1 $\mathcal{T}_1$ $\mathcal{T}_3$ 1 $\mathcal{T}_2$ 1 $\mathcal{T}_2$ 1 $\mathcal{T}_4$ 0 $\mathcal{T}_4$

Среднее время выполнения алгоритма

$$T_n(k) = 2 - 2^{1-n}, \quad k = 1,$$
$$(7.1)\quad T_n(k) =$$
$$= \sum_j y\left(2^{-n}\binom{n}{j}\right) + \frac{2}{2^n} T_n(k-1) + \frac{1}{2^n} \sum_{0<j<n} \binom{n}{j} T_j(k-1), \quad k>1.$$

В частности, $T_1(k) = k$, $T_2(k) = k + 1 - 2^{-k}$, $T_3(k) = k + \frac{5}{2} - \frac{3}{2}2^{-k} - 4^{1-k}$; в общем случае $T_n(k)$ имеет вид $k + d_n + O(2^{-k})$, где $d_n = O(\log_2 n)$[2].[1]) (Напомним, что при $n = 3$ оптимальное время выполнения в соответствии с (5.17) равно $k + .446 + O(k^2 2^{-k})$.)

Теорема 7.1. *При $n \geqslant 3$ для $F(x) = x^n$ не существует ГКС, являющегося оптимальным ПР-алгоритмом.*

Доказательство. Покажем, что каждое оптимальное ПР-дерево имеет бесконечное число неизоморфных поддеревьев. Если v — узел любого ПР-дерева, пусть $P(v)$ — длина наибольшего «свободного пути», начинающегося в v, т. е. наибольшее расстояние, которое может быть пройдено из v до узлов, являющихся его потомками, прежде, чем встретится выход[2]). Если $P(v_1) \neq P(v_2)$, то поддеревья с корневыми узлами v_1 и v_2 не могут быть изоморфными; следовательно, достаточно показать, что для любого N существует узел v с $N < P(v) < \infty$. (Может показаться странным, что доказательство опирается на длинные свободные пути, так как свободные пути интуитивно представляются неэффективными и рассмотренные выше генераторы с конечным числом состояний для $F(x) = x^n$ не содержали длинных свободных путей. Тем не менее именно это отсутствие длинных свободных путей делает их неоптимальными в соответствии с данным доказательством!)

Напомним, что, согласно теореме 4.2, каждое оптимальное ПР-дерево получается с помощью последовательных уточнений. Когда $\boxed{v}$ — концевой узел ПДР-дерева уровня k, то уточнение, примененное к уровню $k+1$, заменяет $\boxed{v}$ поддеревом $\mathscr{T}(v)$ с концевыми узлами $\boxed{v'}$ или $\boxed{v''}$, наиболее длинный свободный путь из v в результирующем ПР-дереве является са-

[1]) В выходных данных к статье отмечается, что ряд вычислений и формальных преобразований выполнены с использованием системы MACSYMA. — *Прим. перев.*

[2]) То есть очередной бит результата. — *Прим. перев.*

мым длинным путем из v в $\mathscr{T}(v)$. Если $\mathscr{T}(v)$ содержит конечное число, скажем $M(v)$, концевых узлов $\boxed{v'}$ и $\boxed{v''}$, отсюда следует, что $\log_2 M(v) \leqslant P(v) > M(v)$. Для завершения доказательства необходимо только указать такие узлы v, что $M(v)$ является сколь угодно большим конечным числом.

Рассмотрим вначале случай $n = 3$, поскольку это упрощает выкладки. Будем использовать целые

$$b = \frac{4^{2t+1} - 4}{10}, \quad j = 5 \cdot 2^m + b, \quad k \geqslant m + 3, \tag{7.2}$$

где t — очень большое число, однако m много больше, скажем $m \geqslant 100t$. Тогда узлы v, v' и v'' будут соответствовать интервалам $[j/2^k, (j+1)/2^k) = [2j/2^{k+1}, (2j+1)/2^{k+1}) \cup [(2j+1)/2^{k+1}, (2j+2)/2^{k+1})$. Соответствующие $\Delta_F(I)$ для $F(x) = x^3$ есть $p = p' + p''$, где

$$\begin{aligned} p' &= (300 \cdot 2^{2m} + (12 \cdot 4^{2t+1} - 18) \cdot 2^m + 12b^2 + 6b + 1)/2^{3k+3}, \\ p'' &= (300 \cdot 2^{2m} + (12 \cdot 4^{2t+1} + 42) \cdot 2^m + 12b^2 + 18b + 7)/2^{3k+3}, \\ p &= (600 \cdot 2^{2m} + (24 \cdot 4^{2t+1} + 24) \cdot 2^m + 24b^2 + 24b + 8)/2^{3k+3}. \end{aligned} \tag{7.3}$$

Когда p' прибавляется к p'', в этих трех двоичных дробях не происходит переносов между заключенными в скобки коэффициентами при $2^{2m-3k-3}$, 2^{m-3k-3} и 2^{-3k-3}, поскольку m достаточно велико; следовательно, потомки v' и v'' каждого узла v, соответствующего биту в одной из этих трех групп, также будут принадлежать к той же самой группе. Вторая группа содержит ровно четыре узла v вместе с пятью узлами v'' и $4t+3$ узлами v', следовательно, один из четырех узлов v будет иметь в качестве потомков по крайней мере $(5 + 4t + 3)/4 = t + 2$ узлов v' и v''. Так как t может быть сколь угодно большим, теорема доказана полностью для случая $n = 3$.

Доказательство в общем случае $n \geqslant 3$ производится аналогично: выбираются b, j и k, удовлетворяющие (7.2), и величина m, достаточно большая по сравнению с t. Соответствующие коэффициенты при $2^{m(n-2)-nk-n}$ будут равны $n(n-1)10^{n-3}(2 \cdot 4^{2t+1} - 3) + n(n-1)10^{n-3} \cdot (2 \cdot 4^{2t+1} + 7) = n(n-1)10^{n-3}(4 \cdot 4^{2t+1} + 4)$, и снова мы можем указать узел v со сколь угодно большим числом потомков, когда t достаточно велико. ■

Аналогичным образом может быть показано, что распределение $F(x) = q^2x^2$, $0 \leqslant x \leqslant 1/q$, не может быть порождено оптимальным образом никаким ГКС, когда $q > 1$ нечетно: разбиение интервала $[j/2^k, (j+1)/2^k]$ в случае, когда $j = (2^t - a)/q^2$, является целым, а в случае, когда $q^2/4 \leqslant a < q^2/2$,

соответствует двоичной сумме

$$(2^{t+2}+3q^2-4a)+(2^{t+2}-(4a-q^2))=4(2^{t+1}+q^2-2a);$$

отсюда следует существование узла v, для которого $P(v)\approx t$. Подобные целые t и a существуют для сколь угодно большого t, поскольку можно положить $a=2^s$ и $t=rb+s$, где $2^r=1 \pmod{q^2}$.

Следующая теорема показывает, что многие распределения $F(x)$, которые являются полиномами с рациональными коэффициентами, могут быть порождены с помощью ГКС.

Теорема 7.2. *Пусть $F(x)$ — функция распределения на $[0,1]$, производная которой может быть представлена в виде конечной суммы полиномов*

$$f_{ij}(x)=x^i(1-x)^j, \quad 0\leqslant x\leqslant 1, \tag{7.4}$$

с положительными рациональными коэффициентами. Тогда для распределения $F(x)$ существует ГКС, среднее время получения с помощью которого k битов результата равно $k+0(1)$.

Доказательство. Рассмотрим вначале сравнительно легкий случай, когда

$$F(x)=p_1x+p_2x^2+\ldots+p_nx^n, \quad p_1+p_2+\ldots+p_n=1, \tag{7.5}$$

и все p_j неотрицательны и рациональны. План доказательства состоит в построении вначале ПДР-дерева для дискретного распределения $(p_1, p_2, \ldots, p_n)$: оптимальное ПДР-дерево с конечным числом неизоморфных поддеревьев может быть найдено, поскольку каждое p_j имеет периодическое двоичное разложение и в качестве их общей длины можно использовать наименьшее общее кратное длин периодов. Далее, где бы это ПДР-дерево не имело концевой узел, скажем для p_j, подставим вместо него ГКС $\mathscr{T}_j$, определенный выше для распределения x^j. Среднее время, необходимое для получения с помощью этого метода k битов результата, не превосходит $k+\max(d_1, \ldots, d_n)$ плюс не более чем $\log_2 n+2$ битов для начального определения j.

В случае когда $F(x)$ имеет отрицательные коэффициенты, но $F'(x)$ является положительной рациональной комбинацией функций $f_{ij}(x)$ из (7.4), справедливы аналогичные рассуждения: $F(x)$ должна быть выпуклой комбинацией бета-распределений

$$F_{ij}(x)=(i+j+1)\binom{i+j}{j}\int_0^1 t^i(1-t)^j\,dt, \quad 0\leqslant x\leqslant 1, \tag{7.6}$$

с рациональными коэффициентами p_{ij} и этого достаточно для того, чтобы доказать, что каждое $F_{ij}(x)$ может быть порождено с помощью соответствующего ГКС. Хорошо известно (см., на-

пример [8], упр. 5—7), что $F_{ij}(x)$ является распределением $(i+1)$-й наименьшей среди $i+j+1$ независимых равномерно распределенных случайных величин; таким образом, требуется только обобщить предыдущее построение для определения максимальной из n равномерно распределенных случайных величин.

ГКС $\mathscr{T}_{nt}$ для t-й наибольшей среди равномерно распределенных случайных величин может быть построен следующим образом: вначале используется $\mathscr{T}_n$ для порождения случайного целого j, имеющего биномиальное распределение $\binom{n}{j}2^{-n}$; в этом случае j равняется числу равномерно распределенных случайных величин, старший бит которых равен 1. Если $j \geqslant t$, то выводится 1 и далее используется ГКС $\mathscr{T}_{jt}$; в противном случае выводится 0 и используется $\mathscr{T}_{n-j,\,t-j}$. (Подобное построение сводится к предыдущему, когда $t=1$, т. е. $\mathscr{T}_{n1}=\mathscr{T}_n$.)

Поскольку $F_{ij}(x)$ порождаемо с помощью $\mathscr{T}_{i+j+1,\,j+1}$ за $k+O(1)$ единиц времени, то теорема доказана полностью. ■

Заметим, что проведенное в этой теореме построение доставляет ГКС, время выполнения которого не превосходит $k+D_n$, где задержка D_n зависит только от максимальной степени n полиномов $F_{ij}(x)$ в принятом разложении $F(x)$, а не от коэффициентов этой выпуклой комбинации. Однако иногда значение n может быть большим, чем собственно степень $F(x)$, например когда $F(x)$ имеет отрицательные коэффициенты. Действительно, если $F'(x)$ пропорциональна $1-3x+3x^2$, то мы не можем получить $F'(x)$ в виде неотрицательной линейной комбинации 1, x, $1-x$, x^2, $x(1-x)$ и $(1-x)^2$, однако $1-3x+3x^2=(1-x)^3+x^3$.

Можно распространить построение из теоремы 7.2 на значительно более общие полиномиальные распределения. Например, можно построить ГКС, который моделируют t-ю наибольшую среди n независимых случайных величин, равномерно распределенных в $[p, q]$, где p и q — рациональные числа, $0 \leqslant p < q \leqslant 1$. Это делает возможным моделирование $F(x)$ в случае, когда $F'(x)$ пропорциональна $(3x-1)^2$, например путем раздельного моделирования в интервалах $[0, 1/3]$ и $[1/3, 1]$. (Подобное распределение $F(x)$ не удовлетворяет условиям теоремы 7.2, так как $F_{ij}(1/3) > 0$ для всех i и j, но $(3x-1)^2=0$ при $x=1/3$.) С другой стороны, когда $F'(x)$ имеет иррациональные корни (т. е. она пропорциональна $(2x^2-1)^2$), то не существует очевидного способа построения ГКС для $F(x)$.

Теперь мы сможем использовать теорему 7.2 для получения более сильного результата, представляющего теоретический интерес,

Теорема 7.3. *Пусть $F(x)$ — полиномиальная функция распределения на $[0, 1]$, имеющая вид (7.5) с неотрицательными рациональными коэффициентами, и пусть $\varepsilon > 0$. Тогда существует ГКС, который порождает $F(x)$, и среднее время получения с помощью этого ГКС любого числа k битов результата менее чем на ε больше среднего времени выполнения оптимального ПР-алгоритма.*

Доказательство. Идея доказательства состоит в использовании оптимального ПР-алгоритма для моделирования большого числа результирующих битов вследствие чего основная часть распределения будет порождена с помощью «равномерных» [1]) поддеревьев (подобных $\mathcal{T}_1$), в то время как оставшаяся часть распределения является полиномом, и, следовательно, может быть довольно быстро порождена с помощью ГКС, используя теорему 7.2. Этой неопределенной идее может быть придан точный смысл, используя некоторые идеи из доказательства теоремы 5.2.

Пусть m — большое число. Так как $F(x)$ удовлетворяет условиям теоремы 5.2, мы можем заключить, что все значения j из диапазона $0 \leqslant j < 2^{2m}$, за исключением $O(2^m)$, будут удовлетворять (5.11); другими словами, найдется целое t_j, такое, что

$$(7.7)\qquad \frac{t_j}{2^m} \leqslant f(x) < \frac{t_j+1}{2^m} \quad \text{для всех} \quad x \in \left[\frac{j}{2^{2m}}, \frac{j+1}{2^{2m}}\right].$$

Будем говорить, что j является «хорошим», когда имеет место это неравенство. Поскольку вероятность $\Delta_F([j/2^m,\ (j+1)/2^m))$ равна

$$A_{2m,j} = \int_{j/2^{2m}}^{(j+1)/2^{2m}} f(x)\,dx,$$

то для любых $k \geqslant 0$ и $0 \leqslant l \leqslant 2^k - 1$ имеем

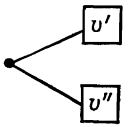
$$(7.8)\qquad \frac{t_j}{2^{3m+k}} \leqslant A_{2m+k,\,2^k j+l} < \frac{t_j+1}{2^{3m+k}}.$$

Другими словами, вероятности для всех разбиений «хорошего» интервала $[j/2^{2m},\ (j+1)/2^{2m})$ будут иметь одинаковые старшие биты, определяемые значениями t_j.

Пусть $p = p' + p''$, где p, p' и p'' имеют одинаковые старшие биты, т. е. мы предполагаем, что $\lfloor 2^s p \rfloor = \lfloor 2^{s+1} p' \rfloor + \lfloor 2^{s+1} p'' \rfloor$ для некоторого целого s. Тогда легко видеть, что алгоритм уточнения из разд. 4 может быть применен к первым s уровням де-

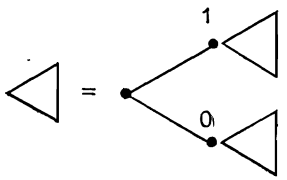
[1]) То есть порождающих равномерное распределение. — *Прим. перев.*

рева, если множество S этого алгоритма рассматривается как очередь («первым вошел — первым вышел»): каждый узел $\boxed{v}$ на уровне s или меньшем будет заменен поддеревом



В этом случае *каждое* уточнение разбиений «хорошего» интервала будет обладать свойством равенства старших битов; следовательно, существует оптимальное ПР-дерево, в котором *каждый* узел v на уровне, не превосходящем $3m$, принадлежащий «хорошему» интервалу, будет иметь «равномерное» поддерево

(7.9)



Другими словами, когда мы встречаемся с подобным узлом v, то все последовательные выходы попросту совпадают с входами. Таким образом, бóльшая часть оптимального ПР-дерева уже является ГКС в том смысле, что подавляющее большинство его поддеревьев к тому времени, когда мы порождаем $2m$-й бит результата, изоморфны «равномерному» поддереву. Для завершения доказательства мы должны разобраться с оставшейся частью дерева.

Рассмотрим вначале «плохое» значение j. Мы можем предположить, что в процессе построения оптимального ПР-дерева ПДР-дерево для результата на уровне $2m$ является конечным ПДР-деревом, поскольку все фигурирующие при этом вероятности являются рациональными числами. (Действительно, алгоритм уточнения преобразует конечное ПДР-дерево в конечное ПДР-дерево, когда p, p' и p'' являются рациональными, если S рассматривается как очередь.) Каждый концевой узел v этого ПДР-дерева, соответствующий «плохому» значению j, теперь будет замещен ГКС для полиномиального распределения, определенным следующим образом:

$$A_{2m,\,j}F_j(x) = F\left(\frac{j+x}{2^{2m}}\right) - F\left(\frac{j}{2^{2m}}\right), \quad 0 \leqslant x \leqslant 1. \tag{7.10}$$

Снова рассмотрим «хорошее» значение j. Заменим каждый концевой узел v ПДР-дерева, соответствующий такому j, на «равномерное» поддерево (7.9), если v имеет уровень $\leqslant 3m$; в противном случае заменим v на ГКС для полиномиального распределения $G_j(x)$ с неотрицательными рациональными коэффициентами, определенным следующим образом:

$$\frac{t_j}{2^{3m}}\,x+\left(A_{2m,j}-\frac{t_j}{2^{3m}}\right)G_j(x)=A_{2m,j}F_j(x). \tag{7.11}$$

Результирующее дерево, дополненное спецификациями вывода для узлов уровня $\leqslant 2$ так, как это делалось при построении оптимальных ПР-деревьев в теореме 4.1, является ГКС для $F(x)$. Остается проверить, что время получения результатов достаточно близко к оптимальному.

Среднее время, необходимое для получения k-го бита при $k\leqslant 2m$, является оптимальным, исходя из способа построения. Среднее время, необходимое для получения $(2m+k)$-го бита, превышает оптимальное самое большее на величину, ограниченную произведением D_n (см. замечание к теореме 7.2) на вероятность того, что используется ГКС для полиномиального распределения $F_j(x)$ или $G_j(x)$. Но вероятность того, что используется F_j, равна вероятности того, что j «плохое», а именно $O(2^{-m})$, и вероятность того, что использовано G_j, равна

$$\sum\left(A_{2m,j}-\frac{t_j}{2^{3m}}\right)<\sum_{0\leqslant j<2^{2m}}1/2^{3m}=2^{-m},$$

где первая сумма вычисляется по всем «хорошим» j. Следовательно, среднее время получения результата превышает оптимальное самое большее на величину $O(D_n/2^m)$, что для достаточно больших m меньше, чем ε. ■

В заключение настоящего раздела покажем, что класс распределений, реализуемых с помощью ГКС, очень ограничен. Как было показано в разд. 5, только некоторые редко встречающиеся непрерывные распределения можно моделировать с помощью ГКС. Однако оказывается, что если $F(x)$ является кусочно-аналитической функцией, реализуемой с помощью ГКС, то $F(x)$ должно быть кусочно-полиномиальной функцией.

Будем говорить, что $F(x)$ является *аналитической* на $[c, d]$, если для любого $x_0\in[c, d]$ существуют $\delta>0$ и последовательность коэффициентов $\langle a_0, a_1, a_2, \ldots\rangle$, ..., такая, что

$$F(x)=\sum_{j\geqslant 0}a_j(x-x_0)^j \tag{7.12}$$

для любых $x\in[c, d]$, таких, что $|x-x_0|\leqslant\delta$.

Теорема 7.4. *Если распределение* $F(x)$ *на* $[0, 1]$ *порождаемо с помощью ГКС и* $F(x)$ *является аналитической функцией на* $[c, d] \subseteq [0, 1]$ *для некоторых* $c < d$, *то существует полином* $p(x)$ *с рациональными коэффициентами, такой, что* $F(x) = p(x)$ *при всех* $x \in [c, d]$.

Доказательство. В доказательстве мы будем опираться на лемму, которая по существу является следствием теоремы Арбиба [2] и Хеллера [4], о реализуемости стохастических систем. Поскольку наша вычислительная модель несколько отличается от модели Арбиба и Хеллера, то мы докажем данную лемму применительно к нашей модели.

Лемма. *Пусть* $F(x)$ *порождаемо с помощью ГКС. Тогда существует конечное множество* $\{G_1(x), G_2(x), \ldots, G_s(x)\}$ *функций на* $[0, 1]$, *со следующим свойством: для каждого* $k \geqslant 0$ *и* $0 \leqslant j < 2^k$, *таких, что* F *является отличной от константы непрерывной функцией на* $[j/2^k, (j+1)/2^k)$, *функция*

$$F_{kj}(x) = \frac{F\left(\frac{j+x}{2^k} - 0\right) - F\left(\frac{j}{2^k} - 0\right)}{F\left(\frac{j+1}{2^k} - 0\right) - F\left(\frac{j}{2^k} - 0\right)}, \quad 0 \leqslant x \leqslant 1, \tag{7.13}$$

является выпуклой комбинацией G_i. *В частности, любое семейство, содержащее более чем* s *функций* $F_{kj}(x)$, *является линейно зависимым над* $[0, 1]$.

Доказательство. Поскольку ГКС-дерево имеет конечное число неизоморфных поддеревьев, существует конечное множество функций распределения $\{G_1(x), G_2(x), \ldots, G_s(x)\}$, соответствующих любому биту, которым помечен любой узел разветвления. (Так как цепочка битов, которой помечен каждый узел разветвления, является конечной, то ее длина ограниченна. После того как нами получен некоторый бит из этой цепочки, найдется распределение $G_i(x)$, которому будут соответствовать остальные биты из этой цепочки и биты, которыми помечены потомки данного узла. Согласно введенным выше определениям, существует ГКС для разрывного распределения $F(x) = 0$ при $x < 1/\pi$ и $F(x) = 1$ при $x \geqslant 1/\pi$, а именно тривиальное дерево, состоящее из единственного концевого узла, доставляющего в качестве результата бесконечную цепочку нулей и единиц, которая соответствует двоичному представлению $1/\pi$. Для такого ГКС существует бесконечно много линейно независимых распределений $F_{ij}(x)$. В действительности лемма верна также и для разрывных $F_{kj}(x)$ при условии, что $F_{kj}(x)$ не является «одноступенчатой» функцией, для которой $F_{kj}(x + 0) = F_{kj}(x - 0) + 1$ при некотором x.)

Для данных k и j рассмотрим все узлы $v \in V_{kj}$, которым был приписан k-й бит результата, а первые k битов определяли двоичное представление $j/2^k$. Поскольку каждый такой узел v встречается на некотором уровне $L(v)$ и определяет некоторое распределение $G^{(v,\,k)}(x) \in \{G_1(x), \ldots, G_s(x)\}$, то

$$(7.14)\quad F_{kj}(x) = \sum_{v \in V_{kj}} \frac{1}{2^{L(v)}} G^{(v,\,k)}(x) = \sum_{1 \leqslant i \leqslant s} \lambda_i G_i(x), \quad 0 \leqslant x \leqslant 1,$$

для некоторого λ_i, что и доказывает лемму. ■

Приступая к доказательству теоремы, выберем точку $x_0 = j/2^k$, такую, что $c < x_0 < d$. Тогда существуют δ и $\langle a_0, a_1, \ldots\rangle$, удовлетворяющие (7.12). Заменяя в случае необходимости j на $2^t j$ и k на $k + t$, можно предположить, что $1/2^k < \delta$ и $x_0 + 1/2^k < d$. Согласно лемме, существует номер s, такой, что при $0 \leqslant i \leqslant s$ функции $F_{k+i,\,2^i j}(x)$ являются линейно зависимыми, т. е. существуют $b_i \neq 0$, такие, что

$$(7.15)\quad \sum_{0 \leqslant i \leqslant s} b_i A_{k+i,\,2^i j} F_{k+i,\,2^i j}(x) = 0, \quad 0 \leqslant x \leqslant 1.$$

Поскольку на основании (7.12) и (7.13)

$$(7.16)\quad A_{k+1,\,2^i j} F_{k+1,\,2^i j}(x) = \sum_{l \geqslant 1} a_l \left(\frac{x}{2^{k+i}}\right)^l, \quad 0 \leqslant x \leqslant 1,$$

то равенство (7.15) означает, что

$$(7.17)\quad \sum_{0 \leqslant i \leqslant s} b_i / (2^{k+i})^l = 0$$

для всех l, таких, что $a_l \neq 0$. Если по меньшей мере $s + 1$ значений a_l отличны от нуля, то (7.17) определяет систему линейных уравнений, определитель которой кратен отличному от нуля определителю Вандермонда; следовательно, все $b_i = 0$. Отсюда следует, что $a_l = 0$ для всех больших l, т. е. $F(x)$ является полиномиальной функцией $p(x)$ в интервале $[x_0, x_0 + 1/2^k)$. Из хорошо известной теории аналитических функций следует, что $F(x) = p(x)$ для *всех* $x \in [c, d]$.

Таким образом, осталось показать, что коэффициенты $p(x)$ рациональны. В силу равенства (7.16) при $x = 1$ достаточно доказать, что для любого ГКС A_{kj} являются рациональными числами. Поскольку A_{kj} является вероятностью получения отдельной строки, состоящей из k нулей и единиц, A_{kj} может быть представлено в виде суммы элементов строки матрицы, являющейся результатом произведения k матриц перехода между «состояниями» ГКС. Здесь под «состоянием» понимается устройство, имеющее ровно один выход. Рассматриваемый нами ГКС

может быть легко представлен как имеющий конечное множество состояний с *не более* чем одним выходом в каждом состоянии, если разделить каждый узел, который порождает два и более результатов, на отдельные узлы, переход между которыми возможен с вероятностью 1. В то же время мы можем исключить состояния, которые *не* имеют ни одного выхода, поодиночке следующим стандартным способом. Если v является таким состоянием, которому соответствует вероятность p_{vv} перехода из v в v, то мы можем исключить v, если для каждой другой пары состояний u и w заменим вероятность p_{uw} перехода из u в w на вероятность

$$p_{uw} + \frac{p_{uv}p_{vw}}{1 - p_{vv}}. \tag{7.18}$$

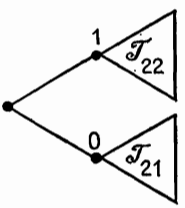
После применения конечного числа подобных преобразований мы получим устройство с конечным числом состояний, каждое из которых имеет ровно один выход и переходы между каждой парой состояний определяются рациональными вероятностями. Как установлено выше, отсюда следует, что каждое A_{kj} является рациональным числом. ■

Представляется интересным указать такие распределения, для которых ГКС является оптимальным. Одним из важных примеров может служить распределение $(U + V)/2$, где U и V являются независимыми равномерно распределенными случайными величинами, а именно

$$F(x) = \begin{cases} 2x^2, & 0 \leqslant x \leqslant 1/2, \\ 1 - 2(1-x)^2, & 1/2 \leqslant x \leqslant 1. \end{cases} \tag{7.19}$$

Оптимальный ГКС для этого распределения получается из дерева

(7.20)



где $\mathscr{T}_{22}$ было определено при доказательстве теоремы 7.2 (и соответствует $\min(U, V)$). Кроме того, читатель может проверить, что распределение

$$F(x) = (x - p)/q, \quad p \leqslant x \leqslant p + q, \tag{7.21}$$

также может быть оптимальным образом получено с помощью ГКС, когда p и q являются рациональными числами, такими, что $0 \leqslant p < p + q \leqslant 1$.

8. МОДЕЛИРОВАНИЕ ЭКСПОНЕНЦИАЛЬНОГО РАСПРЕДЕЛЕНИЯ

Выше было показано, что экспоненциальное распределение

$$F(x) = 1 - e^{-x}, \quad x \geqslant 0, \tag{8.1}$$

которое является одним из важнейших в практических приложениях, не может быть порождено с помощью ГКС. Отсюда следует, что ГКС представляют собой слишком ограниченную модель, в то время как общие ПР-алгоритмы являются слишком неограниченными моделями. В настоящем разделе мы вкратце обсудим простой алгоритм моделирования экспоненциального распределения в надежде на то, что это несколько прояснит главный вопрос о том, какая же промежуточная модель в действительности наиболее соответствует практическим задачам.

Следующий элегантный метод был предложен Джоном фон Нейманом [11] [1]). В настоящее время известно много других методов моделирования экспоненциального распределения (см., например, [1], гл. 7), однако мы будем иметь дело с методом фон Неймана, поскольку он допускает очень простую программную реализацию. Этот метод заключается в следующем.

1. Положить $c \leftarrow 0$.
2. Получить независимые реализации U_0, U_1, ... из равномерного распределения до тех пор, пока не найдется наименьшее $j \geqslant 1$, такое, что $U_{j-1} < U_j$.
3. Если j четно, то положить $c \leftarrow c + 1$ и вернуться к шагу 2; в противном случае получить в качестве результата $c + U_0$.

Рассмотренные до сих пор примеры распределений были ограничены интервалом [0, 1], однако экспоненциально распределенные случайные величины могут принимать сколь угодно большие значения. Будем использовать единично-двоичное представление для вещественных чисел так, как это обсуждалось в разд. 3, но без начальной единицы, поскольку все числа будут положительными; таким образом, число $c + t$, где c — целое и t — дробь, представляется в виде нуля и c единиц, за которыми следует двоичное представление t. Тогда экспоненциальное распределение может быть порождено адаптированным методом фон Неймана следующим образом:

```
begin integer f, p, s, t; integer array a, b [1 : ∞];
    comment В данной реализации метода Неймана переменные
            имеют следующий смысл:
            f = 1 означает, что j = 1, f = 0 означает, что j > 1,
            p представляет собой j (mod 2),
```

[1]) См. также Феллер В. Введение в теорию вероятностей и ее приложения. Т. 2. — М.: Мир, 1967, с. 62—63. — *Прим. перев.*

```
            a_1, ..., a_s представляют собой старшие биты U_0,
            b_1, ..., b_t представляют собой старшие биты U_{j-1};
  while true do
  begin f ← 1; p ← 1; t ← 0; comment j ← 1;
    advance: t ← t + 1; b_t ← flip;
             if f = 1 then begin s ← t; a_t ← b_t end;
             if flip = 0 then go to advance;
    inequality found: if b_t = 1 then
             begin comment U_{j-1} > U_j;
               f ← 0; p ← 1 − p; b_t ← 0; comment j ← j + 1;
               for i ← 1 step 1 until t do
               if flip = 1 then begin t ← i;
               goto inequality found end;
               goto advance;
             end else
             begin comment U_{j-1} < U_j;
               if p = 1 then
               begin output (0); comment переход от единичного представления к двоичному;
                 for i ← 1 step 1 until s do output (a_i);
                 while true do output (flip)
               end else output (1); comment c ← c + 1;
             end
  end
end;
```

Ясно, что этот алгоритм может быть выполнен стохастической машиной Тьюринга, которая работает в реальном времени (т. е. одна машинная операция соответствует одной инструкции *flip*), при условии, что в состав операций, выполняемых этой машиной, входят операции печати символа и смещения головки вправо на одну позицию или мгновенного возвращения ее к началу ленты. Вряд ли возможна реализация этого алгоритма на стохастическом магазинном автомате даже не в реальное время.

Среднее время выполнения оптимального ПР-алгоритма для экспоненциального распределения в единично-двоичном представлении может быть получено из доказательства теоремы 5.2. Согласно (5.9),

$$T_F(k) - k = \sum_{m \geqslant 1-k} \frac{m}{2^m} J_m(k), \tag{8.2}$$

но в случае единично-двоичного представления

$$J_m(k) = \frac{1}{2^k}\left(\sum_{0 \leqslant i \leqslant k} \sum_{0 \leqslant j < 2^{k-i-1}} \varepsilon_{m+k}\left(A_{kj}^{(i)}\right)\right) + \frac{1}{2^k}\varepsilon_{m+k}(1 - F(k)), \tag{8.3}$$

где

$$A_{kj}^{(i)} = F\left(i + \frac{j+1}{2^{k-i-1}}\right) - F\left(i + \frac{j}{2^{k-i-1}}\right). \tag{8.4}$$

Пусть m фиксировано и $k \to \infty$. Тогда для всех x из интервала, соответствующего $A_{kj}^{(i)}$, и для всех i и j, за исключением $O(2^m)$ «плохих» значений, функция плотности $f(x) = F'(x)$ будет иметь постоянное значение $t_{ij} = \lfloor 2^m f(x) \rfloor$; для «хороших» значений $2^{-k}\varepsilon_{m+k}(A_{kj}^{(i)})$ оказывается равным интегралу от $\varepsilon_m(2^{-1-i}f(x))$ по этому интервалу, умноженному на 2^{-1-i}. Следовательно, в единично-двоичном представлении аналогом (5.12) является

$$J_m(k) = \sum_{i \geqslant 0} \frac{1}{2^{i+1}} \int_i^{i+1} \varepsilon_m\left(\frac{f(x)}{2^{i+1}}\right) dx + O(\min(1, 2^{m-k})). \tag{8.5}$$

Пусть $f(x) = e^{-x}$ и J_m — предельное значение (8.5) при $k \to \infty$. Тогда $J_1 = 0$, $J_2 = \frac{1}{2}\ln 2$, $J_3 = \frac{1}{2}(1 + \ln(2/3))$, и дальнейшие вычисления показывают, что среднее число случайных битов, необходимых для получения с помощью оптимального ПР-алгоритма k битов результата, равно $k + 0.5385 + O(k^2/2^k)$.

Интересно проанализировать быстродействие приведенного выше простого алгоритма моделирования экспоненциального распределения, однако мы ограничимся только эмпирическим результатом. В результате 1000 испытаний этого алгоритма на ЭВМ оказалось, что для получения k битов результата требуется около $k + 5.4 \pm 0.2$ случайных битов при больших k.

9. ЗАКЛЮЧИТЕЛЬНЫЕ ЗАМЕЧАНИЯ И ОТКРЫТЫЕ ПРОБЛЕМЫ

Мы обсудили начальные понятия нового раздела теории сложности, который объединяет теоретический и численный анализ с целью изучения одного интересного приложения алгоритмов и автоматов и проливает некоторый свет на задачу моделирования неравномерно распределенных случайных величин. Хотя мы получили ответы на ряд вопросов, которые естественным образом возникают в рамках этой теории, многие проблемы остались открытыми.

Вероятно наиболее интересным направлением дальнейших исследований является изучение моделей, промежуточных между алгоритмами с конечным числом состояний и общими алгоритмами на деревьях, поскольку эти модели можно естественным образом связать с задачами моделирования на ЭВМ экспоненциального, нормального, гамма- и тому подобных распределений. Насколько эффективно мы можем моделировать нормально распределенные случайные величины, исходя из последователь-

ности случайных битов и используя сравнительно простой алгоритм?

В этой связи могут быть рассмотрены автоматы или алгоритмы, которым «сообщены» двоичные представления некоторых констант. Например, вероятность того, что экспоненциально распределенная случайная величина превышает значение $k \ln 2$, равна $1/2^k$ и один из способов моделирования экспоненциально распределенных случайных величин может состоять в использовании этого факта для сведения задачи к моделированию случайных величин в промежутке $[0, \ln)$. В этом случае для выполнения алгоритма потребуется двоичное представление $\ln 2$ и, возможно, кратные ему значения.

Другой интересной темой, не затронутой в настоящей статье, является моделирование важнейших дискретных распределений, таких, как распределение Пуассона с параметром λ или биномиальное распределение с параметрами n и p [1]). Заслуживающий особого внимания метод «подсчета единиц» для моделирования биномиального распределения предложен Аренсом ([1], гл. 12); нельзя ли его существенно улучшить?

Ряд следующих задач мы также оставили нерешенным. (1) Можно ли распространить теорему 5.2 на случаи, когда $F(x)$ имеет бесконечно много точек разрыва или недифференцируемости? Как связан коэффициент при k со свойствами F? (2) Определить среднее время выполнения простого алгоритма моделирования экспоненциального распределения, рассмотренного в разд. 8. (3) Охарактеризовать все $F(x)$, для которых ГКС является оптимальным. (4) Насколько большим должно быть число состояний ГКС, порождающего распределение $F(x) = x^3$, который является ε-оптимальным? (5) Какие полиномиальные распределения можно моделировать с помощью ГКС? (6) Можно ли с помощью магазинного автомата моделировать любое распределение? (7) Что если источник независимых случайных битов является смещенным (так что вероятность единицы равна p)?

ЛИТЕРАТУРА

1. Ahrens J. H., Dieter U. Non-uniform random numbers — Graz: Institüt für Math. Statistik, 1974.
2. Arbib M. A. Realization of stochastic systems, IEEE Conf. record, Symp. on switching and automata theory, 7 (1966), 262—265.
3. Bentley J. L., Yao A. C. An almost optimal algorithm for unbounded searching, Inform. Processing Letters, 5 (1976), 82—87.
4. Heller A. On stochastic processes derived from Markov chains, Ann. Math. Stat., 36 (1965), 1286—1291.

[1]) По-видимому, имеются в виду алгоритмы, являющиеся функциями параметров указанных распределений. — *Прим. перев.*

5. Knuth D. E. The art of computer programming, vol. 1: Fundamental algorithms, Addison-Wesley, Reading, 1968. [Имеется перевод: Кнут Д. Искусство программирования для ЭВМ. Т. 1. Основные алгоритмы. — М.: Мир, 1976.]
6. Knuth D. E. The art of computer programming, vol. 2: Seminumerical algorithms, Addison-Wesley, Reading, 1969. [Имеется перевод: Кнут Д. Искусство программирования для ЭВМ. Т. 2. Получисленные алгоритмы. — М.: Мир, 1977.]
7. Knuth D. E. The dangers of computer science theory, Proc. 4th Internat. Congress for logic, methodology and philosophy of science (Bucharest, 1971), ed. by P. Suppes et al., North-Holland, Amsterdam, 1973, 367—371.
8. Knuth D. E. The art of computer programming, vol. 3: Sorting and searching, Addison-Wesley, Reading, 1973. [Имеется перевод: Кнут Д., Искусство программирования для ЭВМ. Т. 3. Сортировка и поиск. — М.: Мир, 1978.]
9. Paz A. Introduction to probabilistic automata, Academic Press, New York, 1971.
10. Stoppard T. Rosencrantz and Guildenstern are dead, Faber and Faber, London, 1967.
11. Neumann von J. Various techniques used in connection with random digits, Collected works, vol. 5, Pergamon Press, London, 1963, 768—770.

Теоремы существования для семейств Шпернера[1])

Д. Дейкин, Дж. Годфри, А. Хилтон

Reading University, England

1. ВВЕДЕНИЕ

Пусть $\mathcal{S}$ — конечное семейство конечных множеств. Будем называть $\mathcal{S}$ семейством Шпернера, если для любых двух различных множеств X, Y из $\mathcal{S}$ выполняется соотношение $X \not\subseteq Y$. Параметрами семейства $\mathcal{S}$ являются числа $p_0, p_1, \ldots$, где p_i это число множеств мощности i в семействе $\mathcal{S}$. Последовательность $p_0, p_1, \ldots$, разумеется, является конечной. Если $p_0 > 0$, то семейство Шпернера состоит только из пустого множества $\varnothing$. Поэтому мы обычно считаем $p_0 = 0$. Через $[i, j]$ будем обозначать множество целых чисел $i, i+1, \ldots, j$, которое считаем пустым при $i > j$.

Эта статья была стимулирована гипотезой [2] Д. Клейтмана и Э. Милнера, которую мы доказываем в виде следующей теоремы.

Теорема 1. *Пусть $\mathcal{S}$ — семейство Шпернера, состоящее из подмножеств множества $[1, s]$, и пусть $p_0, p_1, \ldots, p_s$ — параметры семейства $\mathcal{S}$. Тогда существует семейство Шпернера $\mathcal{U}$, снова состоящее из подмножеств множества $[1, s]$, с параметрами $q_0, q_1, \ldots, q_s$, где $q_i = 0$ при $0 \leqslant i < \frac{1}{2}s$ и $q_i = p_{s-i} + p_i$ при $\frac{1}{2}s < i \leqslant s$, а когда s четно $q_{\frac{1}{2}s} = p_{\frac{1}{2}s}$.*

Для доказательства этой теоремы нам нужен специальный класс семейств Шпернера. Мы их смогли извлечь из работы Д. Краскала [3], описываем мы их в разд. 5. Они позволили нам также получить следующую теорему.

Теорема 2. *Пусть $p_0, p_1, \ldots, p_g$ — неотрицательные целые числа, удовлетворяющие условию $p_0 = 0 < p_g$. Тогда наименьшее такое целое s, для которого существует семейство Шпернера из подмножеств множества $[1, s]$ с параметрами $p_0, p_1, \ldots, p_g$,*

---

[1]) Daykin D. E., Godfrey J., Hilton A. J. W. Existence Theorems for Sperner Families, J. Comb. Theory (A) 17, 245—251, 1974.

© 1974 by Academic Press, Inc.
© Перевод на русский язык, «Мир», 1983

определяется соотношением

$$s = p_1 + K_2(p_2 + K_3(p_3 + \ldots + K_g(p_g)\ldots)), \qquad (1)$$

где K *обозначает функцию Краскала, которая определяется ниже соотношениями* (2) *и* (3).

Вывести теорему 1 из теоремы 2 нам не удалось. О приложениях теоремы 1 упоминается в [2].

2. УПОРЯДОЧИВАЮЩЕЕ ОТНОШЕНИЕ

Ради удобства мы будем рассматривать наши множества как подмножества множества $\{1, 2, \ldots\}$. Для двух различных множеств X и Y будем использовать запись $X < Y$, если наибольшее целое число, входящее в симметрическую разность $(X \setminus Y) \cup (Y \setminus X)$, принадлежит Y. Пусть $l \geqslant 1$, и пусть $\mathscr{L}$ обозначает семейство всех множеств мощности l. Упорядочивающее отношение превращает $\mathscr{L}$ в последовательность $L_1, L_2, \ldots$, которая, например, при $l = 3$ начинается так:

$$\begin{aligned} L_1 &= 1,\ 2,\ 3 \\ L_2 &= 1,\ 2,\ \ \ 4 \\ L_3 &= 1,\ \ \ 3,\ 4 \\ L_4 &= \ \ \ 2,\ 3,\ 4 \\ L_5 &= 1,\ 2,\ \ \ \ 5 \\ L_6 &= 1,\ \ \ 3,\ \ 5 \end{aligned}$$

В общем случае L_{i+1} можно получить из L_i следующим образом.

Правило, определяющее следующее множество. *Пусть даны целые числа* m *и* l, *удовлетворяющие условию*: $1 \leqslant m \leqslant l$, *и множество* L *мощности* l. *Пусть* α — *наименьшее такое целое число, что*

(i) $\alpha \in L$,

(ii) $\alpha + 1 \notin L$,

(iii) $|L \cap [1, \alpha]| \geqslant 1 + l - m$.

Тогда множество M *мощности* m, *следующее за множеством* L, *определяется так*:

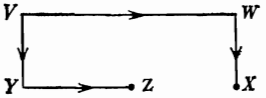
$$M = [1, \beta] \cup \{\alpha + 1\} \cup (L \setminus [1, \alpha]),$$

где

$$\beta = |L \cap [1, \alpha]| - 1 - l + m.$$

Заметим, что $L < M$, а когда $m = l$, условие (iii) становится лишним, поскольку оно следует из условия (i).

Приведем теперь три примера применения этого правила. Каким образом мы его применяем, показано на следующей диаграмме:



Множества V, W имеют мощность 3, а множества X, Y, Z — мощность 2.

	Пример 1	*Пример 2*	*Пример 3*
V	1, 2, 4	1, 3, 4	2, 3, 4
W	1, 3, 4	2, 3, 4	1, 2, 5
X	1, 5	1, 5	3, 5
Y	3, 4	1, 5	1, 5
Z	1, 5	2, 5	2, 5

В примерах соответственно $X = Z$, $X < Z$, $Z > X$, что показывает некоммутативность диаграммы. Эти соотношения показывают также, что за различными множествами одинаковой мощности может идти одно и то же множество другой мощности.

3. ТЕОРЕМА КРАСКАЛА — КАТОНЫ

В этом разделе мы приведем несколько результатов, которые первым получил Краскал [3], а затем независимо переоткрыл Катона [1].

При заданных положительных числах l и n существует единственное представление числа n в виде

$$n = C_{a_l}^{l} + C_{a_{l-1}}^{l-1} + \ldots + C_{a_t}^{t}, \tag{2}$$

где $l \geqslant t \geqslant 1$ и $a_l > a_{l-1} > \ldots > a_t \geqslant t$. Пусть $\mathscr{S}$ — некоторое семейство множеств, и пусть Δ — оператор, удаляющий всеми возможными способами один элемент из множеств семейства $\mathscr{S}$, давая тем самым другое семейство множеств $\Delta\mathscr{S}$.

Теорема 3. *Пусть $\mathscr{S}$ — семейство из n множеств мощности l. Тогда*

$$|\Delta\mathscr{S}| \geqslant K_l(n),$$

где

$$K_l(n) = C_{a_l}^{l-1} + C_{a_{l-1}}^{l-2} + \ldots + C_{a_t}^{t-1}. \tag{3}$$

В следующем разделе будет приведен пример семейства $\mathscr{S}$, которое Краскал назвал каскадом. Для каскада $|\Delta\mathscr{S}| = K_l(n)$, и это показывает, что теорему 3 усилить нельзя. Повторное применение теоремы 3 дает оценку для $|\Delta^r\mathscr{S}|$. Если рассмотреть

случай $\mathscr{S} = \mathscr{T} \cup \mathscr{U}$, где $\mathscr{T}$ и $\mathscr{U}$ — непересекающиеся семейства, состоящие соответственно из n_1 и n_2 множеств, то мы увидим, что

$$K_l(n_1 + n_2) \leqslant K_l(n_1) + K_l(n_2), \tag{4}$$

где n_1, n_2 — любые неотрицательные целые числа.

4. КАСКАД КРАСКАЛА

Для нас важно то, что мы можем пример каскада выписать в явном виде:

$$\mathscr{S} = \mathscr{S}_l \cup \mathscr{S}_{l-1} \cup \ldots \cup \mathscr{S}_t, \tag{5}$$

где при $l \geqslant i \geqslant t$

$$\mathscr{S}_i = \{X \cup Y_i : |X| = i,\ X \subseteq [1,\ a_i]\},$$

а

$$Y_i = \bigcup_{i<j\leqslant l} \{1 + a_j\}.$$

Заметим, что $Y_l = \varnothing$, $Y_{l-1} = \{1 + a_l\}$, $Y_{l-2} = \{1 + a_{l-1}\} \cup \{1 + a_l\}$ и т. д. Выразив $\mathscr{S}$ таким способом, мы можем сразу сделать ряд наблюдений.

Во-первых, если при $l \geqslant i \geqslant t$ имеет место $Z_i \in \mathscr{S}_i$, то тогда $Z_l < Z_{l-1} < \ldots < Z_t$. Таким образом, семейства вида $\mathscr{S}_i$ попарно не пересекаются и $|\mathscr{S}_i| = C^i_{a_i}$, так что в силу (2) и (5) $|\mathscr{S}| = n$. При упорядочении семейства $\mathscr{S}$ последним множеством в нем является

$$\Omega = [1 - t + a_t,\ a_t] \cup Y_t,$$

ибо Ω является последним множеством в $\mathscr{S}_t$. Если $|Z| = l$ и $Z < \Omega$, то тогда $Z \in \mathscr{S}$. Другими словами, $\mathscr{S}$ состоит из первых n множеств мощности l, поэтому в обозначениях разд. 2 $\mathscr{S} = \{L_1, L_2, \ldots, L_n\}$. Кроме того, семейство $\mathscr{S}$ можно породить правилом, определяющим следующее множество, при $l = m$.

Далее мы заметим, что

$$\Delta\mathscr{S}_l = \{X : |X| = l - 1,\ X \subseteq [1,\ a_l]\},$$

$$(\Delta\mathscr{S}_{l-1}) \setminus \Delta\mathscr{S}_l = \{X \cup Y_{l-1} : |X| = l - 2,\ X \subseteq [1,\ a_{l-1}]\},$$

$$(\Delta\mathscr{S}_{l-2}) \setminus \Delta(\mathscr{S}_l \cup \mathscr{S}_{l-1}) = \{X \cup Y_{l-2} : |X| = l - 3,\ X \subseteq [1,\ a_{l-2}]\},$$

и т. д., причем числа элементов в этих множествах это соответственно $C^{l-1}_{a_l}$, $C^{l-2}_{a_{l-1}}$, $C^{l-3}_{a_{l-2}}$ и т. д. Таким путем с помощью (3) мы устанавливаем, что $|\Delta\mathscr{S}| = K_l(n)$. Последнее множество семейства $\Delta\mathscr{S}$ получается удалением первого элемента из последнего множества Ω семейства $\mathscr{S}$, поэтому оно имеет вид

$$\Psi = [2 - t + a_t,\ a_t] \cup Y_t.$$

Если $|Z| = l - 1$ и $Z < \Psi$, то тогда $Z \in \Delta\mathscr{S}$. Следовательно, $\Delta\mathscr{S}$ состоит из первых $K_l(n)$ множеств мощности $l - 1$.

5. СПЕЦИАЛЬНЫЕ СЕМЕЙСТВА ШПЕРНЕРА

Теперь мы можем раскрыть, что мы понимаем под специальным семейством Шпернера $\mathscr{F}$, имеющим заданные параметры $p_0, p_1, \ldots, p_g$, которые удовлетворяют условию: $p_0 = 0 < p_g$. Пусть $\mathscr{F}_i^*$ обозначает совокупность из p_i множеств мощности i, принадлежащих $\mathscr{F}$. Тогда

$$\mathscr{F} = \mathscr{F}_1^* \cup \mathscr{F}_2^* \cup \ldots \cup \mathscr{F}_g^*.$$

Для упрощения записи положим $q_g = p_g$ и при $i = g-1, g-2, \ldots, 1$

$$q_i = p_i + K_{i+1}(q_{i+1}).$$

В частности, q_1 — это число из (1). Пусть теперь семейство $\mathscr{F}_g^*$ состоит из первых q_g множеств мощности g. Как показано в последнем разделе, $\Delta\mathscr{F}_g^*$ состоит из первых $K_g(p_g) = q_{g-1} - p_{g-1}$ множеств мощности $g-1$, а следующие p_{g-1} множеств мы берем в качестве второй части семейства $\mathscr{F}_{g-1}^*$. Таким образом, семейство $\mathscr{F}_{g-1}^* \cup \Delta\mathscr{F}_g^*$ состоит из первых q_{g-1} множеств мощности $g-1$. Аналогично, $\Delta(\mathscr{F}_{g-1}^* \cup \Delta\mathscr{F}_g^*)$ состоит из первых $q_{g-2} - p_{g-2}$ множеств мощности $g-2$, а следующие p_{g-2} множеств мы берем для $\mathscr{F}_{g-2}^*$ и т. д. для $\mathscr{F}_{g-3}^*, \ldots, \mathscr{F}_1^*$. Очевидно, что

$$\bigcup_{X \in \mathscr{F}} X = \mathscr{F}_1^* \cup \Delta(\mathscr{F}_2^* \cup \Delta(\ldots(\mathscr{F}_{g-1}^* \cup \Delta\mathscr{F}_g^*)\ldots)) = [1, q_1].$$

Другой тривиальный факт, который будет важен в следующем разделе, формулируется в виде леммы.

Лемма 1. *Специальным семейством Шпернера с параметрами* $p_0, p_1, \ldots, p_{g-2}, q_{g-1}$ *является* $\mathscr{F}_1^* \cup \mathscr{F}_2^* \cup \ldots \cup \mathscr{F}_{g-1}^* \cup \Delta\mathscr{F}_g^*$.

Чтобы доказать теорему 2, рассмотрим произвольное семейство Шпернера $\mathscr{S}$ с заданными параметрами $p_1, \ldots, p_g$. Обозначим через $\mathscr{S}_i^*$, $1 \leqslant i \leqslant g$, подсемейство семейства $\mathscr{S}$, состоящее из p_i множеств мощности i. Тогда по теореме 3

$$|\Delta\mathscr{S}_g^*| \geqslant K_g(p_g)$$

и, поскольку $\mathscr{S}_{g-1}^*$ и $\Delta\mathscr{S}_g^*$ не пересекаются,

$$|\mathscr{S}_{g-1}^* \cup \Delta\mathscr{S}_g^*| = |\mathscr{S}_{g-1}^*| + |\Delta\mathscr{S}_g^*| \geqslant q_{g-1}.$$

Очевидно, что при фиксированном l $K_l(n)$ не убывает с ростом n. Поэтому, применив снова теорему 3, мы получим

$$|\mathscr{S}_{g-2}^* \cup \Delta(\mathscr{S}_{g-1}^* \cup \Delta\mathscr{S}_g^*)| = |\mathscr{S}_{g-2}^*| + |\Delta(\mathscr{S}_{g-1}^* \cup \Delta\mathscr{S}_g^*)| \geqslant$$
$$\geqslant p_{g-2} + K_{g-1}(|\mathscr{S}_{g-1}^* \cup \Delta\mathscr{S}_g^*|) \geqslant p_{g-2} + K_{g-1}(q_{g-1}) = q_{g-2}.$$

Продолжая действовать таким способом, мы установим в конце концов, что

$$\Big|\bigcup_{X\in\mathscr{S}} X\Big| = |\mathscr{S}_1^* \cup \Delta(\mathscr{S}_2^* \cup \Delta(\ldots(\mathscr{S}_{g-1}^* \cup \Delta\mathscr{S}_g^*)\ldots))| \geqslant q_{g-1},$$

и теорема 2 доказана. Таким образом, справедлива также

Теорема 4. *Для каждого семейства Шпернера, состоящего из подмножеств множества* $[1, s]$, *существует специальное семейство Шпернера с теми же параметрами, состоящее также из подмножеств множества* $[1, s]$.

6. ДОКАЗАТЕЛЬСТВО ТЕОРЕМЫ 1

Пусть s — фиксированное положительное число, и пусть заданы параметры $p_0, p_1, \ldots, p_s$. Обозначим через h наибольшее такое целое, что $p_{s-h} + p_h > 0$. Если s четно и $h = \frac{1}{2}s$, утверждение теоремы 1 тривиально при $\mathscr{Y} = \mathscr{S}$. Если s нечетно и $h = \frac{1}{2}(s+1)$, то доказываемое утверждение также справедливо, ибо, согласно теореме Шпернера [2], $C_s^h \geqslant p_0 + p_1 + \ldots + p_s = p_{h-1} + p_h$, так что в качестве $\mathscr{Y}$ можно взять любые $p_{h-1} + p_h$ множеств мощности h. Сделаем индуктивное предположение о том, что доказываемое утверждение справедливо для всех таких семейств $\mathscr{S}$, что h удовлетворяет соотношению: $\frac{1}{2}s \leqslant h < k$, и рассмотрим случай $h = k$.

Итак, мы предполагаем, что имеется семейство $\mathscr{S}$ с параметрами $p_0, p_1, \ldots, p_s$, где $p_i = 0$ при $i < s-k$ и $i > k$, но $p_{s-k} + p_k > 0$. Допустим временно, что s четно и что $k \geqslant 3 + \frac{1}{2}s$. Отправляясь от $\mathscr{S}$, определим 6 семейств Шпернера: $\mathscr{R}_1, \ldots, \mathscr{R}_6$. Параметры всех этих семейств сведены в таблицу, где для удобства мы положили

$$a = p_k, \quad b = p_{k-1}, \quad c = p_{s-k+1}, \quad d = p_{s-k},$$
$$K^a = K_k(a), \quad K^d = K_k(d), \quad K^e = K^a + K^d,$$
$$K^f = K_k(a+d) \leqslant K^e.$$

Неравенство $K^f \leqslant K^e$ вытекает из (4).

Отправляясь от $\mathscr{S}$, мы применим оператор Δ к подсемейству $\mathscr{S}_k^*$ множеств из $\mathscr{S}$ мощности k с целью получить не менее K^a новых множеств мощности $k-1$. Пусть $\mathscr{W}$ — какие-нибудь K^a из этих множеств. Положим $\mathscr{R}_1 = (\mathscr{S} \setminus \mathscr{S}_k^*) \cup \mathscr{W}$. Семейство $\mathscr{R}_2$ получается заменой каждого множества из $\mathscr{R}_1$ его дополнением.

Таблица значений параметров

	$\mathscr{S}$	$\mathscr{R}_1$	$\mathscr{R}_2$	$\mathscr{R}_3$	$\mathscr{R}_4$	$\mathscr{R}_5$	$\mathscr{R}_6$
$i > k$	0	0	0	0	0	0	0
$i = k$	a	0	d	0	0	0	$a+d$
$i = k-1$	b	$b+K^a$	c	$c+K^d$	$b+c+K^e$	$b+c+K^f$	$b+c$
$k-1 > i > \frac{1}{2}s$	p_i	p_i	p_{s-i}	p_{s-i}	p_i+p_{s-i}	p_i+p_{s-i}	p_i+p_{s-i}
$i = \frac{1}{2}s$	p_i	p_i	p_i	p_i	p_i	p_i	p_i
$\frac{1}{2}s > i > s-k+1$	p_i	p_i	p_{s-i}	p_{s-i}	0	0	0
$i = s-k+1$	c	c	$b+K^a$	$b+K^a$	0	0	0
$i = s-k$	d	d	0	0	0	0	0
$i < s-k$	0	0	0	0	0	0	0

Затем $\mathscr{R}_3$ получается из $\mathscr{R}_2$ точно так же, как $\mathscr{R}_1$ было получено из $\mathscr{S}$. Поскольку параметры семейства $\mathscr{R}_3$ удовлетворяют равенству $p_k = p_{s-k} = 0$, по предположению индукции существует семейство $\mathscr{R}_4$ с параметрами, указанными в таблице. В свою очередь, поскольку $K^e \geqslant K^f$ и имеется семейство $\mathscr{R}_4$, в силу теоремы 4 существует специальное семейство $\mathscr{R}_5$, а значит, в силу леммы 1 и специальное семейство $\mathscr{R}_6$, у которых параметры имеют значения, указанные в таблице.

Приведенное рассуждение сохраняется, даже когда нарушено какое-нибудь из условий: s четно или $k \geqslant 3 + \frac{1}{2}s$, если в зависимости от случая удалять очевидным способом из таблицы 1, 2, 3 или 4 строки. Доказательство теоремы 1 получается по индукции.

Добавление в корректуре (21 марта 1974 г.). Авторы хотели бы поблагодарить Д. Катону, сообщившего им, что обобщение теоремы 2 было доказано Д. Ф. Клементсом (Clements G. F. A minimization problem concerning subsets of a finite set, Discrete Math., 4, 1973, 123—128).

ЛИТЕРАТУРА

1. Katona G. A theorem for finite sets, Theory of Graphs, Proc. of Colloquium, Tihany, Hungary, 1966, 187—207.
2. Kleitman D., Milner E. C. On the average size of the sets in a Sperner family (в печати).
3. Kruskal J. B. The number of simplices in a complex, Math. Optimisation Techniques, University of California, Berkeley and Los Angeles, 1963,

Простое доказательство теоремы Краскала — Катоны[1])

Д. Дейкин

Reading University, England

Заметка является продолжением предыдущей статьи [1] из этого же журнала [2]).

Пусть n и l — положительные числа, а $\mathscr{S}$ — совокупность первых n множеств мощности l. Кроме того, пусть $\mathscr{A}$ — любое другое семейство из n множеств мощности l. Тогда $\mathscr{S}$ представляет собой каскад Краскала и теорема, которую мы собираемся доказать ([1, теорема 3]), утверждает, что $|\Delta\mathscr{S}| \leqslant |\Delta\mathscr{A}|$.

Пусть m — минимум величины $|W \setminus X|$, взятый по всем множествам W мощности l, не принадлежащим $\mathscr{A}$, и всем множествам X, принадлежащим $\mathscr{A}$, для которых выполняется условие $W < X$. Поскольку $\mathscr{A} \neq \mathscr{S}$, число m определено и $m \geqslant 1$. Среди пар множеств W, X выберем произвольную такую пару W_0, X_0, что $m = |W_0 \setminus X_0|$. Положим

$$M = W_0 \setminus X_0, \quad N = X_0 \setminus W_0,$$

так что $M \cap N = \varnothing$, $m = |M| = |N|$ и $M < N$. Далее, положим

$$\mathscr{B} = \{X : X \in \mathscr{A}, \quad M \cap X = \varnothing, \quad N \subseteq X, \quad (X \setminus N) \cup M \notin \mathscr{A}\},$$

$$\mathscr{C} = \{(X \setminus N) \cup M : X \in \mathscr{B}\},$$

$$\mathscr{D} = \mathscr{C} \cup (\mathscr{A} \setminus \mathscr{B}),$$

так что $|\mathscr{A}| = |\mathscr{D}|$. Грубо говоря, $\mathscr{D}$ формируется из $\mathscr{A}$ путем замены везде, где это возможно, подмножества N множества X из $\mathscr{A}$ на подмножество M. Множество, получающееся из X в результате такой замены, удовлетворяет условию: $(X \setminus N) \cup M < X$. В частности, W_0, получающееся из X_0, удовлетворяет условию: $W_0 < X_0$. Таким образом $\mathscr{D}$ ближе к $\mathscr{S}$, чем $\mathscr{A}$, и повторение этой процедуры постепенно преобразует $\mathscr{A}$ в $\mathscr{S}$. Поэтому для доказательства теоремы достаточно показать, что $|\Delta\mathscr{D}| \leqslant |\Delta\mathscr{A}|$.

Пусть $Z \in (\Delta\mathscr{D}) \setminus (\Delta\mathscr{A})$. Тогда существует такое множество $Y \in \mathscr{D} \setminus \mathscr{A} = \mathscr{C}$, что $Z \subseteq Y$. По определению семейства $\mathscr{C}$ су-

---

[1]) Daykin D. E. A Simple Proof of the Kruskal — Katona Theorem. J. Comb. Theory (A) 17, 252—253, 1974.

[2]) См. настоящий сборник, стр. 159—166.

© 1974 by Academic Press, Inc.
© Перевод на русский язык, «Мир», 1983

ществует такое множество $X \in \mathscr{B}$, что $Y = (X \setminus N) \cup M$ и, таким образом, $N \cap Z = \varnothing$. Предположим, что $M \not\subseteq Z$. Поскольку Z получается удалением одного элемента из Y и $M \subseteq Y$, из соотношения $P = M \cap Z$ теперь следует, что $|P| = m - 1$. Далее, пусть Q — множество, полученное удалением из N наименьшего его элемента β, тогда $|Q| = m - 1$, $P \cap Q = \varnothing$ и либо $P = Q = \varnothing$, либо $P < Q$. Поскольку $X \in \mathscr{B}$, выполняются соотношения $M \cap X = \varnothing$ и $N \subseteq X$, так что $P \cap X = \varnothing$ и $Q \subseteq X$. Положим $W_1 = (X \setminus Q) \cup P$, тогда $|W_1| = l$, $|W_1 \setminus X| = |P| = m - 1$ и $W_1 \leqslant X \in \mathscr{A}$. Пользуясь определением числа m, мы приходим к выводу, что $W_1 \in \mathscr{A}$. Однако $Q \cap Z = N \cap Z = \varnothing$ и $M \cap Z = P$, так что $Z \subseteq W_1$, а это дает соотношение $Z \in \Delta\mathscr{A}$, противоречащее определению множества $\mathscr{A}$. Таким образом, мы показали, что $N \cap Z = \varnothing$ и $M \subseteq Z$.

Введем

$$\psi Z = (Z \setminus M) \cup N.$$

Тогда $\psi Z \subseteq X$ и, таким образом, ψ определяет отображение $(\Delta\mathscr{D}) \setminus (\Delta\mathscr{A})$ в $\Delta\mathscr{A}$. Покажем, что $\psi Z \notin \Delta\mathscr{D}$. Действительно, предположим, что $\psi Z \subseteq V \in \mathscr{D}$. Тогда $N \subseteq V$, так что $V \notin \mathscr{C}$, но $V \in \mathscr{A} \setminus \mathscr{B}$. Поскольку $N \subseteq V \notin \mathscr{B}$, имеет место один из двух случаев: (i) $M \cap V \neq \varnothing$, (ii) $M \cap V = \varnothing$ и $(V \setminus N) \cup M \in \mathscr{A}$. В случае (ii) выполняется соотношение: $Z \subseteq (V \setminus N) \cup M$, так что $Z \in \Delta\mathscr{A}$, а это противоречит определению множества Z. Поэтому должен иметь место случай (i) $M \cap V \neq \varnothing$. В этом случае существует такое целое α, что $V = \alpha \cup \psi Z$ и $M \cap \psi Z = \varnothing$, так что на самом деле $M \cap V = \alpha$. Положим $P = M \setminus \alpha$ и $Q = N \setminus \beta$, как и раньше. Тогда $Q \subseteq N \subseteq V$, $P \cap V = \varnothing$, и мы полагаем $W_2 = (V \setminus Q) \cup P$. Теперь $|W_2| = l$, $|W_2 \setminus V| = |P| = m - 1$ и $W_2 \leqslant V \in \mathscr{A}$, так что, пользуясь определением числа m, мы подобно тому, как раньше, делаем вывод, что $W_2 \in \mathscr{A}$. Кроме того, ввиду того что $\alpha \in V$, имеет место соотношение $M \subseteq W_2$ и, следовательно, $Z \subseteq W_2$. Однако отсюда следует, что $Z \in \Delta\mathscr{A}$, а это противоречит определению множества Z. Таким образом, мы доказали, что ψ отображает $(\Delta\mathscr{D}) \setminus (\Delta\mathscr{A})$ в $(\Delta\mathscr{A}) \setminus (\Delta\mathscr{D})$ и, следовательно, $|\Delta\mathscr{D}| \leqslant |\Delta\mathscr{A}|$, как и требовалось.

Автор благодарит Джина Годфри за любезное указание ошибки в первом варианте этой заметки.

ЛИТЕРАТУРА

1. Daykin D. E., Godfrey J., Hilton A. J W. Existence theorems for Sperner families, J. Comb. Theory (A) 17, 1974, 245—251. [Имеется перевод: см. настоящий сб., стр. 159—166.

Теорема Эрдёша — Ко — Радо из теоремы Краскала — Катоны[1])

Д. Дейкин

Reading University, England

Теорема Эрдеша — Ко — Радо непосредственно следует из теоремы Краскала — Катоны.

Теорема, о которой будет идти речь, была опубликована в [3], в [4] было доказано, что ее экстремальный случай является единственным, изящное доказательство теоремы было дано в [6]. Здесь будет показано, что эта теорема является непосредственным следствием теоремы Краскала — Катоны, которая впервые была опубликована в [7], переоткрыта в [5] и получила короткое доказательство в [1].

Теорема. *Пусть k, r — целые, причем $1 \leqslant k \leqslant r \leqslant 2k$. Пусть F — такое семейство различных собственных подмножеств множества $R = \{1, 2, \ldots, r\}$, каждое из которых имеет мощность не менее k, что ни одно подмножество из F не содержит в себе другое и R не является объединением никаких двух подмножеств из F. Тогда*

$$|F| \leqslant C_{r-1}^{k}.$$

Доказательство. Будем считать, что все множества из F имеют мощность k. Иначе мы бы удалили всеми возможными способами из множеств семейства F, имеющих максимальную мощность, скажем мощность h, по одному элементу. В результате получилось бы новое семейство G множеств мощности не более $h - 1$ с теми же свойствами, что и семейство F, и в силу леммы Шпернера имело бы место неравенство $|G| \geqslant |F|$ (подробности см. в [3, стр. 316]).

Обозначим через C семейство множеств, дополняющих множества из F до множества R. Таким образом, каждое множество из C имеет мощность $r - k$. Далее, обозначим через S семейство множеств мощности $r - k$, являющихся подмножествами множеств из F.

В случае $r = k$ теорема тривиальна. Поэтому пусть $r > k$.

---

[1]) Daykin D. E. Erdös — Ko — Rado from Kruskal — Katona. J. Comb. Theory (A) 17, 254—255, 1974.

© 1974 by Academic Press, Inc.
© Перевод на русский язык, «Мир», 1983.

Предположим, что

$$|F| \geqslant C_{r-1}^{k} + C_{k-1}^{k-1}.$$

Поскольку $k \geqslant r - k$, теорема Краскала — Катоны утверждает, что

$$|S| \geqslant C_{r-1}^{r-k} + C_{k-1}^{r-k-1}.$$

Следовательно,

$$|C| + |S| = |F| + |S| > C_{r-1}^{k} + C_{r-1}^{r-k} = C_{r}^{r-k}$$

и, таким образом, существует множество X, принадлежащее одновременно C и S. Поскольку X принадлежит C, $R \setminus X$ принадлежит F, а поскольку X принадлежит S, существует множество Y из F, содержащее X. Объединение множеств $R \setminus X$ и Y есть R. Это противоречие и доказывает теорему.

ЛИТЕРАТУРА

1. Daykin D. E. A simple proof of the Kruskal — Katona theorem, J. Comb. Theory, Ser. A17, 1974, 252—253. [Имеется перевод: см. настоящий сб., стр. 167—168.]
2. Daykin D. E., Godfrey J., Hilton A. J. W. Existence theorems for Sperner families, J. Comb. Theory, Ser. A17, 1974, 245—251. [Имеется перевод: см. настоящий сб., стр. 159—166.]
3. Erdös P., Ko C., Rado R. Intersection theorems for systems of finite sets, Quart. J. Math. Oxford 12, 1961, 313—318.
4. Hilton A. J. W., Milner E. C. Some intersection theorems for systems of finite sets, Quart. J. Math. Oxford 18, 1967, 369—384.
5. Katona G. A theorem for finite sets, Theory of Graphs, Proc. of Colloquium, Tihany, Hungary, 1966, 187—207.
6. KatonaG. A. simple proof of the Erdös — Chao Ko — Rado theorem, J. Comb. Theory 13, 1972, 183—184.
7. Kruskal J. B. The number of simplices in a complex, Math. Optimisation Technlques, U. of California, Berkeley and Los Angeles, 1963, 251—278.

Структурное распознавание образов, гомоморфизмы и размещения[1])

Р. Харалик

Department of Electrical Engineering, Department of Computer Science, University of Kansas, Lawrence, KS 66045, U. S. A.

В работе обсуждается общая задача распознавания образов с обучением со структурной точки зрения. Показано, что как задача определения решающего правила, так и задача применения решающего правила являются задачами нахождения гомоморфизмов независимо от того, совпадает ли структура данных объектов с N-мерным вектором в случае статического распознавания образов или же она представляет собой цепочку или ее обобщение в случае синтаксического распознавания образов. Затем вводится понятие размещения в качестве более сложной структуры данных объектов, которое является размеченным N-арным отношением, и демонстрируется способ построения и применения решающих правил к размещениям исходя из понятия гомоморфизма. Подход, предложенный в этой работе, дает возможность получать методы структурного распознавания образов, служащие обобщением синтаксического распознавания образов, заданных в виде фраз.

1. ВВЕДЕНИЕ

Эта статья преследует три цели: 1) обеспечить единообразное рассмотрение схемы распознавания N-мерных векторов и схемы распознавания цепочек, что проясняет их общую структурную основу; 2) ввести понятие размещения в качестве структуры данных объекта, 3) показать, как именно использовать размещения аналогично использованию N-мерных векторов и цепочек в структурном распознавании образов.

В статистическом распознавании образов N-мерные векторы используются в качестве основной структуры данных объекта. Каждый объект должен быть упорядоченным набором чисел, и все наборы должны иметь одинаковую длину. Если этими значениями являются действительные числа, то исходя из понятия расстояния можно получить линейные или квадратичные решающие правила для определения класса, к которому следует отнести N-мерный вектор [11]. Если указанные значения —

---

[1]) Haralick Robert M. Structural Pattern Recognition, Homomorphisms, and Arrangements, Pattern Recognition, vol. 10, 1978, 223—236.

© Pattern Recognition Society, 1978
© Перевод на русский язык, «Мир», 1982

нечисловые символы, то можно использовать решающие правила с помощью покрытий цилиндрических множеств [14, 19].

В синтаксическом распознавании образов в качестве основной структуры данных объекта используются цепочки. Любой объект называется предложением и представляет собой конкатенацию символов, принадлежащих заданному множеству символов. Если множество предложений (которые порождаются регулярной грамматикой) сопоставленных с классом, является автоматным языком, то решающее правило, определяющее, к каким именно классам они относятся, может осуществляться конечным автоматом.

Много усилий было затрачено для обобщения структуры данных объекта в синтаксическом распознавании образов, представленной в виде цепочки. Уже описаны структуры данных в виде деревьев [2, 5], матриц [3, 15, 21], плексов [1]) и вебов [2]). Однако в данной работе при обсуждении порождающих грамматик мы ограничимся лишь цепочками.

В структурном распознавании образов каждый класс объектов характеризуется при помощи определенного множества взаимосвязей между частями объектов из этого класса. Главное внимание уделяется комбинаторной сложности и однозначности, которые присущи этой совокупности связей. В структурном распознавании образов предполагается, что комбинаторная сложность объекта достаточно велика, чтобы обеспечить почти 100%-ную точность правильного распознавания. Поэтому вместо методов анализа ошибок, типичных для статического распознавания образов, в структурном распознавании образов выясняются возможности методов построения покрытий для порождения решающих правил. Метод покрытия множества Михальского [19] или процедуры восстановления грамматик, обзор которых дан Фу и Бутом [10], Бирманом и Фельдманом [1], Пао [23], являются частными случаями общей схемы построения покрытий, которая, по нашему мнению, характеризует структурное распознавание образов.

В структурном распознавании имеются четыре основные задачи: 1) выбор структуры данных для объекта (задача его представления); 2) преобразование физических замеров в структуру данных объекта (задача выделения признаков); 3) построе-

---

[1]) В языках цепочек каждый символ имеет две «точки примыкания». Если допускать символы с произвольным конечным числом точек примыкания для связи с другими символами, то можно строить структуры, образуемые взаимосвязанными символами указанного вида, которые называются плексами. — *Прим. перев.*

[2]) Вебами называются ориентированные графы с символами на вершинах. Термины «plexus» и «web» имеют сходные значения — сплетение, паутина. — *Прим. перев.*

ние решающего правила (задача обобщения); 4) применение решающего правила (задача его реализации).

В данной работе показано, что с позиций структурного распознавания образов задача 3 (построение решающего правила) фактически является задачей, выразимой в терминах определения покрытий посредством гомоморфизмов и задания решающих правил, использующих эти покрытия с помощью гомоморфизмов. Если читателю стало ясно, что это один и тот же общий тип задачи независимо от того, является ли структура данных объекта N-мерным вектором, цепочкой или обобщением цепочки, как это упоминалось ранее, то естественно ставить вопрос о возможностях структурного распознавания образов относительно других типов структур данных. С этой целью мы вводим в качестве структуры данных размеченное N-арное отношение, которое является более сложным по сравнению с цепочкой или N-мерным вектором и включает в себя N-мерный вектор, цепочку и ее обобщения как частные случаи. Мы назовем размеченное N-арное отношение размещением и покажем, как именно может осуществляться построение решающего правила, исходя из размещения, выбранного в качестве структуры данных.

В разд. 2 обсуждается вопрос о сведении задачи структурного распознавания образов к задаче нахождения гомоморфизмов. В разд. 3 иллюстрируется пример покрытия множества с помощью методологии структурного распознавания образов, в то время как разд. 4 посвящен интерпретации синтаксического примера восстановления грамматики, которая укладывается в рамки методологии структурного распознавания образов. В разд. 5 описано размещение, предлагаемое в качестве структуры данных, и дан пример его применения в структурном подходе к распознаванию образов.

2. ГОМОМОРФИЗМЫ И СТРУКТУРНОЕ РАСПОЗНАВАНИЕ ОБРАЗОВ

В этом разделе мы опишем способ сведения задачи построения решающего правила к задаче нахождения гомоморфизмов. Приведенное здесь математическое описание является достаточно общим для всего структурного распознавания образов и соответствует цели достижения единообразия всех рассматриваемых процессов на теоретическом уровне. Некоторые из ранних идей, способствовавших появлению этого раздела, могут быть найдены в работе Харалика [13]. Начнем с введения некоторых определений.

Пусть P — множество объектов и C — множество классов. Предположим, что $P' \subseteq P$ есть множество наблюдаемых объектов. Множество обучающих данных представляет собой бинарное отношение $T \subseteq P' \times C$, которое сопоставляет наблюденным

объектам их истинные классы. Обычно доля наблюдаемых объектов из P весьма мала. Каждому классу также сопоставляется по крайней мере один наблюденный объект: следовательно, $T(P')=C$, так что T является отображением P' на C. Поскольку в структурном распознавании образов предполагается, что ошибки отсутствуют, любой наблюденный объект, относящийся к некоторому классу, соответствует не более чем одному классу. Таким образом, T однозначно; из $(p, c)\in T$ и $(p, c')\in T$ следует, что $c=c'$. При задании некоторого типа структуры на множестве объектов P задача построения решающего правила состоит в нахождении бинарного отношения D (решающего правила), которое относит каждый объект из P к классам из C и которое значительно шире, чем T; итак, мы должны иметь $T\subseteq D\subseteq \subseteq P\times C$. Заметим, что мы не требуем однозначности D.

Структура на множестве объектов P представляет собой набор подмножеств p, который его покрывает. Подобный набор обозначим через $\mathscr{L}$; $\mathscr{L}$ является некоторым заданным подмножеством множества всех подмножеств P. Структура $(P, \mathscr{L})$ может быть пространством окрестностей или топологическим пространством. Частичный порядок на P, индуцированный посредством $\mathscr{L}$, может оказаться решеткой или, как это подчеркивается в данной работе, $\mathscr{L}$ может быть набором гомоморфных образов элементов P.

Задача построения решающего правила состоит в нахождении способа отнесения к определенному классу объектов, которые ранее не наблюдались. Иными словами, она заключается в определении способа обобщения, исходя из обучающего множества. В структурном распознавании образов обобщение осуществляется посредством покрытия $\mathscr{L}$. Решающее правило относит объект $p\in P$ к классу $c\in C$ тогда, когда имеется подмножество $L\in\mathscr{L}$, удовлетворяющее условиям обобщения:

1) $L\cap P'\neq\varnothing$,
2) из того, что $p\in L\cap P'$, следует $(p, c)\in T$.

Первое условие констатирует, что обобщение может иметь место лишь посредством подмножества L, содержащего один из наблюденных объектов. Можно обобщать исходя лишь из объектов, для которых имеется некоторая информация об идентификации их классов. Во втором условии требуется однозначность: все наблюденные объекты в L должны идентифицироваться одинаково по T. Подмножество $L\in\mathscr{L}$, удовлетворяющее условиям 1 и 2, называется обобщающим множеством. В последующей части этого раздела мы будем анализировать обобщающие множества и иллюстрировать роль, которую играют гомоморфизмы.

Сужением набора $\mathscr{L}$ назовем набор $\mathscr{L}'$, который содержит лишь те члены L набора $\mathscr{L}$, которые удовлетворяют условиям

обобщения 1 и 2. Тем самым

$\mathscr{L}' = \{L \in \mathscr{L} \mid$ 1) $L \cap P' \neq \varnothing$ и
2) существует $c \in C$, удовлетворяющее условию $(p, c) \in T$ для каждого $p \in L \cap P'\}$.

С набором $\mathscr{L}$ связывается естественное бинарное отношение $F \subseteq P \times \mathscr{L}$, которое сопоставляет каждый p с элементом набора $\mathscr{L}$, которому этот объект принадлежит:

$$F = \{(p,\ L) \in P \times \mathscr{L} \mid p \in L\}.$$

Пусть F' — сужение F к $\mathscr{L}'$; $F' = F \cap (P \times \mathscr{L}')$. Элементов $\mathscr{L}$ должно быть достаточно, чтобы допустить обобщение с помощью всей информации, содержащейся в обучающем множестве. Это означает, что для каждого наблюденного объекта должно существовать совместимое подмножество в $\mathscr{L}'$, содержащее этот наблюденный объект. Таким образом, мы полагаем, что для каждого наблюденного объекта p существует $L \in \mathscr{L}'$, удовлетворяющее условию $(p, L) \in F'$; следовательно, F' определено на всем P'.

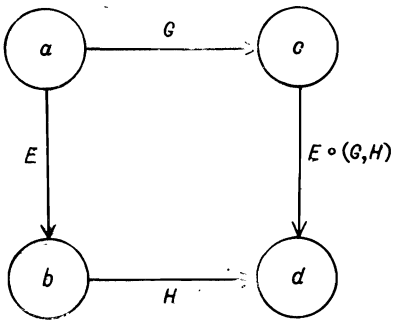


Рис. 1. Иллюстрация понятия композиции отношений. Если $(a, b) \in E$, $(a, c) \in G$ и $(b, d) \in H$, то пара (a, b) переводится в пару (c, d) при помощи композиции E с (G, H).

Чтобы показать, как именно может быть построено решающее правило, использующее покрытие $\mathscr{L}'$, необходимо определить конкретный вид композиции отношений и гомоморфизм, основанный на этой композиции. Пусть $E \subseteq A \times B$, $G \subseteq A \times C$ и $H \subseteq B \times D$. Данный тип композиции переводит пары из E через G и H в пары из $C \times D$. Это преобразование выполняется покомпонентно. Первая компонента переводится через G, вторая компонента — через H.

Определение 1

Обозначим композицию E с (G, H) через $E \cdot (G, H)$ и по определению положим $E \cdot (G, H) = \{(c, d) \in C \times D \mid$ для некоторых $(a, b) \in E$, $(a, c) \in G$ и $(b, d) \in H\}$.

Рис. 1 иллюстрирует идею композиции с помощью коммутативной диаграммы. Заметим, что если $A = C$ и G — отношение тождества на A, то результирующая композиция соответствует обычной композиции функций.

Наше естественное понятие гомоморфизма определяется как отображение одного множества в другое, сохраняющее его структуру. Отображение структуры достигается посредством только что введенной композиции.

Определение 2

(G, H) называется гомоморфизмом из $E \subseteq A \times B$ в $W \subseteq C \times D$ тогда и только тогда, когда $E \cdot (G, H) \subseteq W$.

Иными словами, каждая пара из E должна переводиться в некоторую пару из W. Однако в W могут быть пары, в которые не переводится никакая пара из E.

Мы используем понятия композиции отношений и гомоморфизма следующим образом. Определим отношение включения $I_{p'} = \{(p, q) \in P \times P' \mid p = q\}$. Связь элементов $\mathscr{L}'$ с классами из C может быть задана в виде композиции $I_{p'}$ с (F', T): $Q = I_{p'} \cdot (F', T)$. Q — гомоморфный образ отношения включения $I_{p'}$. Как это показано на рис. 2, обобщающее множество L связывается по Q с классом c, если существует наблюденный объект $p \in P$, который входит в L, $(p, L) \in F'$, и его истинным классом является c, $(p, c) \in T$.

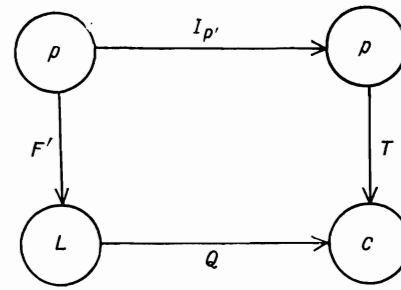


Рис. 2. Коммутативная диаграмма, определяющая отношение Q посредством равенства $Q = I_p \cdot (F', T)$. Также должно быть ясно, что при некоторых разумных условиях T является гомоморфным образом; $F'\,T = F' \cdot (I_p, Q)$.

Обобщение производится через множества из $\mathscr{L}'$, и можно ожидать, что взаимосвязь между объектом p и обобщающим множеством L, как это определено посредством F', должно иметь сходство со взаимосвязью между объектом p и меткой его класса (если она существует), как она задается по T.

В некотором смысле F' должно содержать по крайней мере столько же информации, сколько и T. По данному F' мы должны быть в состоянии преобразовать его и получить T. Действительно, в этом случае T — гомоморфный образ F', полученный по гомоморфизму $(I_{p'}, Q)$. В предложении 1 доказывается, что из $Q \subseteq I_{p'} \cdot (F', T)$ следует $F' \cdot (I_{p'}, Q) \subseteq T$. В предложении 2 доказывается обратное утверждение, а теорема 1 устанавливает, что из соотношения $Q = I_{p'} \cdot (F', T)$ следует $T = F' \cdot (I_{p'}, Q)$.

Предложение 1. Пусть $Q \subseteq I_{p'} \cdot (F', T)$. Тогда $F' \cdot (I_{p'}, Q) \subseteq T$.

Доказательство. Пусть $(p, c) \in F' \cdot (I_{p'}, Q)$. Тогда существует $(p', L) \in F'$, такая, что $(p', p) \in I_{p'}$ и $(L, c) \in Q$. Но из того, что

$(p', p) \in I_{p'}$, следует, что $p = p'$. Поскольку $(L, c) \in Q$ и $Q \subseteq I_{p'} \cdot (F', T)$, существует пара $(q, q') \in I_{p'}$, такая, что $(q, L) \in F'$ и $(q, c) \in T$. Так как $L \in \mathscr{L}'$, то существует $c' \in C$, удовлетворяющий условию $(p^*, c') \in T$ для каждого $p^* \in L \cap P'$. Но $q \in L \cap P'$, так что $(q, c') \in T$. Однако T однозначно, поэтому из того, что $(q, c') \in T$ и $(q, c) \in T$, следует $c = c'$. Окончательно, из условий $(p', L) = (p, L) \in F'$ и $(p, p) \in I_{p'}$ вытекает, что $(p, c) \in T$.

Предложение 2. Пусть $I_{p'} \cdot (F', T) \subseteq Q$. Тогда $T \subseteq F' \cdot (I_{p'}, Q)$.

Доказательство. Пусть $(p, c) \in T$. Тогда $(p, p) \in I_{p'}$. Поскольку F' всюду определена, то существует $L \in \mathscr{L}'$ такое, что $(p, L) \in F'$. Следовательно, $(L, c) \in I_{p'} \cdot (F', T) \subseteq Q$. Теперь из условий $(p, L) \in F'$, $(p, p) \in I_{p'}$ и $(L, c) \in Q$ следует, что $(p, c) \in F' \cdot (I_{p'}, Q)$. Таким образом, $T \subseteq F' \cdot (I_{p'}, Q)$.

Теорема 1. *Если* $Q = I_{p'} \cdot (F', T)$, *то* $T = F' \cdot (I_{p'}, Q)$.

Доказательство. Оно следует из предложений 1 и 2.

Основной смысл теоремы состоит в том, что по заданным отношению обучающих данных T и покрытию $\mathscr{L}$ схема структурного распознавания образов направлена на определение отношения F', которое является обратным гомоморфным образом T. Обобщение означает нахождение обратных гомоморфных образов обучающих данных.

После того как обобщение проведено, легко построить решающее правило. Для этого лишь требуется преобразовать отношение F' в отношение, которое сопоставляет объектам классы. Мы сделаем это следующим образом. Пусть I_P — отношение тождества на P. Определим решающее правило D в виде бинарного отношения из P в C посредством равенства $D = F' \cdot (I_P, Q)$. В то время как отношение обучающих данных T — гомоморфный образ F' при гомоморфизме (I_P, Q), решающее правило является гомоморфным образом F' при гомоморфизме (I_P, Q). Решающее правило относит объект к классу c, если объект принадлежит к определенному обобщающему множеству L, содержащему наблюденный объект, который принадлежит к классу c по T.

Резюмируя, можно сказать, что схема структурного распознавания образов исходит из покрытия $\mathscr{L}$ множества объектов P. Затем с помощью отношения обучающих данных T определяется отношение F', которое связывает объекты с обобщающими множествами из покрытия $\mathscr{L}$. Это отношение F' является обратным гомоморфным образом отношения обучающих данных. Решающее правило D строится как гомоморфный образ отношения F'.

В разд. 3 и 4 мы описываем построение структурных решающих правил для N-мерных векторов в терминах указанных типов гомоморфизмов. В разд. 5 мы вводим размещение как структуру данных объекта и показываем, что и для него решающие правила вписываются в ту же схему.

3. РАСПОЗНАВАНИЕ ОБЪЕКТОВ, ЗАДАННЫХ В ВИДЕ N-МЕРНЫХ ВЕКТОРОВ, С ПОМОЩЬЮ ПОКРЫТИЙ

В этом разделе мы приводим пример построения решающего правила с помощью метода структурного распознавания образов, изложенного в разд. 2. Мы показываем, как именно определить $\mathscr{L}$ в виде гомоморфного образа элементов P, когда сохранение структуры при помощи гомоморфизма связано с псевдометрикой на множестве всех подмножеств P. Затем мы сопоставляем этот метод с методом покрытия Михальского [17—20], который является исходящим вариантом метода Квайна — Маккласки минимизации булевых переменных [16, 25] в случае небулевых переменных. Харалик [14] дал описание некоторых из обсуждаемых в данном разделе идей, касающихся покрытий.

Мы начнем с определений декартовых произведений, метрик, цилиндрических операторов и гомоморфизмов.

Определение 3

Пусть $J = \{j_1, j_2, \ldots, j_N\}$ — линейно упорядоченное множество, элементы которого удовлетворяют условию $j_n < j_{n+1}$, $n = 1, \ldots, N-1$. *Декартовым произведением* относительно множества индексов J из отобранной группы множеств $D_1, D_2, \ldots, D_K$ назовем

$$\underset{j \in J}{\times} D_j = D_{j_1} \times \ldots \times D_{j_N}.$$

Определение 4

Вещественнозначная функция ρ, заданная на $P \times P$, называется *метрикой* тогда и только тогда, когда

1) $\rho(x, y) \geqslant 0$, причем равенство достигается тогда и только тогда, когда $x = y$ (условие неотрицательности),
2) $\rho(x, y) = \rho(y, x)$ (условие симметричности),
3) $\rho(x, y) \leqslant \rho(x, y) + \rho(y, z)$ (неравенство треугольника).

Пара (P, ρ) называется *метрическим пространством.*

Функция ρ, которая удовлетворяет более слабым условиям:
1) $\rho(x, y) \geqslant 0$, и из $x = y$ следует $\rho(x, y) = 0$,
2) $\rho(x, y) = \rho(y, x)$,
называется *псевдометрикой.*

Когда множество объектов P представляется декартовым произведением $\underset{j \in J}{\times} D_j$, на P можно определить естественную мет-

рику ρ, которая вычисляет число несовпадающих компонент любых двух объектов.

Предложение 3. Пусть ρ — функция, определенная на $(\underset{i\in I}{\chi} D_i)\times(\underset{i\in I}{\chi} D_i)$, которая задается при помощи соотношения $\rho((x_1, \ldots, x_N), (y_1, \ldots, y_N)) = \#\{i\in I \mid x_i \neq y_i\}$. Тогда ρ является метрикой на $\underset{i\in I}{\chi} D_i$.

Доказательство. Ясно, что ρ удовлетворяет условиям неотрицательности и симметричности. Для доказательства неравенства треугольника рассмотрим $(z_1, \ldots, z_N)\in \underset{i\in I}{\chi} D_i$. Тогда

$$\begin{aligned}\rho((x_1, \ldots, x_N), (y_1, \ldots, y_N)) &= \#\{i\in I \mid x_i \neq y_i\} = \\ &= \#\{i\in I \mid x_i \neq y_i,\ y_i \neq z_i,\ x_i \neq z_i\} \cup \\ &\cup \{i\in I \mid x_i \neq y_i,\ y_i \neq z_i,\ x_i = z_i\} \cup \\ &\cup \{i\in I \mid x_i \neq y_i,\ y_i = z_i,\ x_i \neq z_i\} \cup \\ &\cup \{i\in I \mid x_i \neq y_i,\ y_i = z_i,\ x_i = z_i\}.\end{aligned}$$

Поскольку $\{i\in I \mid x_i \neq y_i,\ y_i = z_i,\ x_i = z_i\} = \varnothing$, то можно объединить первые два множества, а также первое и третье множества, чтобы получить

$$\begin{aligned}\#\{i\in I \mid x_i \neq y_i\} &= \#(\{i\in I \mid x_i \neq y_i,\ y_i \neq z_i\} \cup \\ &\cup \{i\in I \mid x_i \neq y_i,\ x_i \neq z_i\}) \leqslant \#(\{i\in I \mid y_i \neq z_i\} \cup \\ &\cup \{i\in I \mid x_i \neq z_i\}) \leqslant \#(\{i\in I \mid y_i \neq z_i\} + \{i\in I \mid x_i \neq z_i\}).\end{aligned}$$

Поэтому

$$\begin{aligned}\rho((x_1, \ldots, x_N), (y_1, \ldots, y_N)) \leqslant \rho((x_1, \ldots, x_N), (z_1, \ldots, z_N)) + \\ + \rho((z_1, \ldots, z_N), (y_1, \ldots, y_N)).\end{aligned}$$

Легко проверить, что функция, определяемая при помощи соотношения

$$\rho'(A;\ B) = \min_{x\in A}\min_{y\in B}\rho(x,\ y),$$

является псевдометрикой на множестве всех подмножеств $\mathscr{P}(\underset{i\in I}{\chi} D_i)$.

Функции, которые убывают по расстоянию на метрическом или псевдометрическом пространстве, называются сжимающими отображениями, и для нас они играют роль гомоморфизмов.

Определение 5

Пусть (P', ρ') — метрическое или псевдометрическое пространство. Допустим, что h: $P' \to P'$ удовлетворяет неравенству

$\rho'(x, y) \geqslant \rho'(h(x), h(y))$. Тогда h называется *сжимающим отображением* на P'. Для нашей цели h назовем *гомоморфизмом.*

Существует класс функций от множества всех подмножеств $\mathscr{P}(\underset{i\in I}{\chi} D_i)$ в $\mathscr{P}(\underset{i\in I}{\chi} D_i)$, соответствующий естественной метрике ρ на декартовом произведении $\underset{i\in I}{\chi} D_i$, которые назовем цилиндрическими операторами. Цилиндрический оператор расширяет любое множество до его окрестности, включающей его самого. Для расширения нужно взять каждый N-мерный вектор и считать произвольными значения указанных его координат.

Определение 6

Пусть $J \subseteq I$. Цилиндрический оператор Ψ_J из $\mathscr{P}(\underset{i\in I}{\chi} D_i)$ в $\mathscr{P}(\underset{i\in I}{\chi} D_i)$ определяется посредством соотношения

$$\Psi_J(A) = \{(y_1, \ldots, y_N) \in \underset{i\in I}{\chi} D_i \mid \text{для некоторого } (x_1, \ldots, x_N) \in A,$$

$$x_j = y_j \text{ при всех } j \in J\}.$$

Подмножество $\Psi_J(A)$ называется *J-цилиндрическим множеством A. Порядком* цилиндрического множества или порядком цилиндрического оператора называется число элементов в J. На рис. 3 приведены примеры некоторых цилиндрических множеств порядка 1.

Предложение 4. Пусть $J \subseteq I$. Цилиндрический оператор Ψ_J является гомоморфизмом на $\mathscr{P}(\underset{i\in I}{\chi} D_i)$.

Доказательство. Пусть $A, B \in \mathscr{P}(\underset{i\in I}{\chi} D_i)$. Так как $A \subseteq \Psi_J(A)$ и $B \subseteq \Psi_J(B)$, то $\min\limits_{x\in\Psi_J(A)} \min\limits_{y\in\Psi_J(A)} \rho(x, y) \leqslant \min\limits_{x\in A} \min\limits_{y\in B} \rho(x, y)$. Следовательно, по определению ρ', $\rho'(\Psi_J(A), \Psi_J(B)) \leqslant \rho'(A, B)$, и поэтому Ψ_J — гомоморфизм.

При структурном распознавании N-мерных векторов совокупность $\mathscr{L}$ может задаваться как множество всех образов элементов декартова произведения $P = \underset{i\in I}{\chi} D_i$ с помощью любого цилиндрического оператора, имеющего порядок, который не превышает подходящую константу k.

$$\mathscr{L} = \{L \subseteq P \mid L = \Psi_J(\{p\}) \text{ для некоторого } p \in P = \underset{i\in I}{\chi} D_i \text{ и}$$

$$J \subseteq I, \text{ такого, что } \# J \leqslant k\}.$$

После задания совокупности $\mathscr{L}$ можно определить решающее правило D посредством равенства $D = F' \cdot (I_{p'}, Q)$, где $Q = I_{p'} \cdot (F', T)$, T — отношение обучающих данных и F' — отношение,

которое связывает элементы с обобщающими подмножествами из $\mathscr{L}$.

Мы проиллюстрируем этот метод структурного распознавания образов следующим примером. Пусть пространство образов

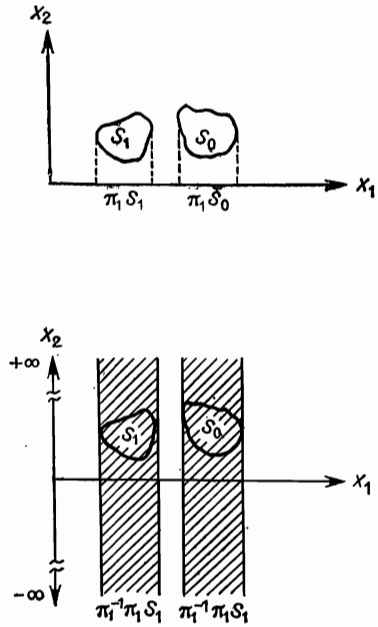


Рис. 3. Иллюстрация операторов проектирования и обратного проектирования для двух подмножеств S_0 и S_1 множества $D_1 \times D_2$.

или замеров P совпадает с $D_1 \times D_2 \times D_3 \times D_4$, где $D_1 = \{x, y, z\}$, $D_2 = \{1, 2, 3\}$, $D_3 = \{A, B, R, S\}$, $D_4 = \{\alpha, \beta, \gamma\}$.

Выберем в качестве $\mathscr{L}$ совокупность всех цилиндрических множеств первого порядка множества P.

$$\mathscr{L} = \{L \subseteq P \mid \text{для некоторого } p \in P,$$
$$L = \Psi_j(\{p\}),\ j = 1,\ 2,\ 3,\ 4\}.$$

Пусть множество классов $C = \{0, 1\}$, а множество наблюденных объектов

$$P' = \{(z,\ 3,\ R,\ \alpha),\ (y,\ 3,\ S,\ \alpha),\ (z,\ 3,\ R,\ \beta),\ (y,\ 3,\ B,\ \beta),$$
$$(z,\ 2,\ S,\ \alpha),\ (z,\ 3,\ B,\ \gamma),\ (z,\ 2,\ S,\ \beta),\ (z,\ 3,\ A,\ \gamma)\}.$$

Отношение обучающих данных $T \subseteq P' \times C$ задается следующей таблицей:

объект	T класс	объект	T класс
$(z, 3, R, \alpha)$	0	$(z, 2, S, \alpha)$	1
$(y, 3, S, \alpha)$	0	$(z, 3, B, \gamma)$	1
$(z, 3, R, \beta)$	0	$(z, 2, S, \beta)$	1
$(y, 3, B, \beta)$	0	$(z, 3, A, \gamma)$	1

Цилиндрические множества первого порядка, которые имеют непустое пересечение с множеством объектов P', приведены в следующем списке, где символ * обозначает любое допустимое значение:

$$z***,\ y***,\ *2**,\ *3**,\ **A*,\ **B*,$$

$$**R*,\ **S*,\ ***\alpha,\ ***\beta,\ ***\gamma.$$

Среди этих цилиндрических множеств первого порядка обобщающими множествами (включающими в себя объекты лишь из одного класса) являются множества $y***$, $*2**$, $**R*$, $**A*$, $***\gamma$.

Отметим, что в данном случае каждый объект из P' принадлежит по крайней мере одному из цилиндрических множеств. Следовательно, отношение $F' \subseteq P \times \mathscr{L}'$ действительно определено всюду на P'. F' задается следующей таблицей:

объект	F' цилиндрическое множество	объект	F' цилиндрическое множество
$(z, 3, R, \alpha)$	$**R*$	$(z, 2, S, \alpha)$	$*2**$
$(y, 3, S, \alpha)$	$y***$	$(z, 3, B, \gamma)$	$***\gamma$
$(z, 3, R, \beta)$	$**R*$	$(z, 2, S, \beta)$	$*2**$
$(y, 3, B, \beta)$	$y***$	$(z, 3, A, \gamma)$	$**A*$, $***\gamma$

Отношение Q между цилиндрическими множествами и классами определяется посредством равенства $Q = I_{P'} \cdot (F', T)$. Q задается следующей таблицей:

цилиндрическое множество	Q класс	цилиндрическое множество	Q класс
$y***$	0	$**A*$	1
$*2**$	1	$***\gamma$	1
$**R*$	0		

Решающее правило D, которое является бинарным отношением между множествами объектов и классов, определяется равен-

ством $D = F'.(I_{P'}, Q)$. Следовательно,

$D = \{(p, c) \in P \times C \mid$

если $c = 0$, то $p = (p_1, p_2, p_3, p_4)$, $p_1 = y$ или $p_3 = R$,

если $c = 1$, то $p_2 = 2$ или $p_3 = A$, или $p_4 = \gamma\}$.

На рис. 4 приведено изображение пространства замеров D на карте Карнау. Отметим, что могут существовать объекты, не принадлежащие никакому классу, объекты, отнесенные ровно к

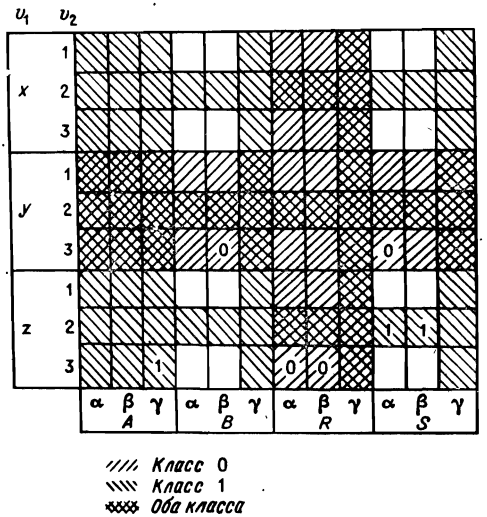


Рис. 4. Обобщенная карта Карнау для решающего правила. Клетки, заштрихованные в одном направлении, соответствуют объектам, отнесенным лишь к одному классу. Незаштрихованные клетки отвечают объектам, не отнесенным ни к какому классу. Клетки, заштрихованные в двух направлениях, соответствуют объектам, отнесенным к обоим классам.

одному классу и объекты, отнесенные более чем к одному классу. Конечно, обучающие объекты однозначно отнесены к одному классу, как это и требовалось.

Этот метод имеет прямую связь с методом покрытия Михальского. Сначала определим покрытие множества S_0 относительно множества S_1, затем покрытие n-го порядка множества S_0 относительно S_1.

Определение 7

Покрытием множества S_0 относительно множества S_1 называется произвольное семейство цилиндрических множеств $\mathscr{L}$, удовлетворяющих условию $S_0 \subseteq \bigcup_{L \subseteq \mathscr{L}} L \subseteq S_1^c$, где S_1^c — дополнение множества S_1.

Из этого определения следует, что если семейство цилиндрических множеств является покрытием, то теоретико-множественное объединение его членов должно полностью покрыть S_0 и не должно покрывать никакую часть множества S_1.

Определение 8

Покрытием n-го порядка множества S_0 относительно множества S_1 называется произвольная совокупность цилиндрических множеств, таких, что

1) $\bigcup_{L \in \mathscr{L}} L \supseteq S_0$,

2) $\bigcup_{L \in \mathscr{L}} L \subseteq S_1^c$,

3) из того, что $L \in \mathscr{L}$, следует, что порядок L не превосходит n.

Метод покрытия требует определения покрытия почти минимального порядка множества S_0 относительно S_1. Как это указывается в теореме 2, для построения покрытия $\mathscr{L}$ необходимо рассматривать лишь цилиндрические множества L со свойствами $L \cap S_1 = \varnothing$ и $L \cap S_0 \neq \varnothing$. Тогда среди цилиндрических множеств, удовлетворяющих этим условиям, должны отбираться последовательно те, которые по мере возрастания их порядков покрывают объекты из S_0, не покрытые предыдущими отобранными множествами [14].

Теорема 2. *Пусть* $\mathscr{L}_n = \{L \mid L$ — *цилиндрическое множество порядка* $\leqslant n$, $L \cap S_1 = \varnothing$ *и* $L \cap S_0 \neq \varnothing\}$. *Если* $S_0 \subseteq \bigcup_{L \in \mathscr{L}_n} L$, то $\mathscr{L}_n$ является покрытием S_0 относительно S_1.

Доказательство. По предположению $\bigcup_{L \in \mathscr{L}_n} L \supseteq S_0$ и порядок $L \in \mathscr{L}_n$ не превосходит n. Поскольку из $L \in \mathscr{L}_n$ следует, что $L \cap S_1 = \varnothing$, мы должны иметь $L \in S_1^c$. Но это верно для всех $L \in \mathscr{L}_n$. Следовательно, $\bigcup_{L \in \mathscr{L}_n} L \subseteq S_1^c$. Поэтому $\mathscr{L}_n$ — покрытие S_0 относительно S_1.

Наше требование $L \cap S_0 \neq \varnothing$ означает, что обобщение может быть осуществлено лишь посредством множеств, содержащих некоторые наблюденные объекты. Требование $L \cap S_1 = \varnothing$ означает, что множества из $\mathscr{L}'$ должны быть обобщающими множествами, так что все наблюденные объекты в $L \in \mathscr{L}'$ должны быть отнесены к общему классу. В задаче, рассмотренной в качестве примера, покрытие первого порядка S_0 относительно S_1 в точности задается при помощи пары цилиндрических

множеств $**R*$ и $y***$, которые сопоставлены с классом 0 посредством отношения Q.

Тот же результат отнесения $**R*$ и $y***$ к классу 0 может быть получен с помощью обобщенного варианта итеративного метода Маккласки. Однако для больших множеств объектов высокой размерности, указанный метод группировки, по-видимому, требует больше памяти и времени вычислений по сравнению с процедурой, исходящей из цилиндрических множеств первого порядка и последовательно переходящей к цилиндрическим множествам k-го порядка, где k — наименьшая константа, позволяющая покрыть обобщающими множествами данное множество объектов.

4. РАЗЛИЧЕНИЕ ОБЪЕКТОВ С ПОМОЩЬЮ ВОССТАНОВЛЕНИЯ ГРАММАТИК

Определение решающего правила в случае, когда объекты задаются цепочками, включает в себя алгебраическую структуру в большей степени чем в случае N-мерных векторов. В этом разделе мы рассмотрим синтаксические методы восстановления грамматик в свете структурного подхода, изложенного в разд. 2, для случая регулярных грамматик. Мы покажем, во-первых, как исходя из отношения обучающих данных можно построить частично заданный акцептор, а во-вторых, что совокупность $\mathscr{L}'$ может состоять из классов эквивалентностей входных моноидов произвольных акцепторов, которые являются гомоморфными образами акцептора, построенного исходя из отношения обучающих данных при помощи гомоморфизма заданного типа. Затем мы проиллюстрируем этот метод на простом примере синтаксического распознавания образов. Идеи этого раздела навеяны работами Фу [9], Фу и Бута [10], Бирмана и Фельдмана [1] и Пао [23].

Начнем с введения некоторых обозначений и определений. Пусть Σ — множество символов, из которых могут состоять цепочки, описывающие объекты, Σ^* — множество всех цепочек символов из Σ. Множеством объектов P считается Σ^*, которое бесконечно, даже если Σ конечно. Как известно, Σ^* — свободный моноид, порожденный из Σ по операции конкатенации цепочек. Пусть $\Sigma^{*'} \subseteq \Sigma^*$ — множество наблюденных объектов и $C = \{1, 2, \ldots, K\}$ — множество K классов. Предположим, что все объекты из $\Sigma^{*'}$ имеют конечную длину. Отношение обучающих данных T является подмножеством $\Sigma^{*'} \times C$: $T \subseteq \Sigma^{*'} \times C$.

Мы хотим определить совокупность $\mathscr{L}$ подмножеств Σ^*. Однако бесконечность Σ^* затрудняет прямое определение подобной совокупности. Наш метод будет заключаться в определении

совокупности $\mathscr{L}$ косвенным путем посредством конечных акцепторов для наблюденных цепочек из $\Sigma^{*\prime}$. Акцептор для K классов представляет собой автомат, который, начиная свою работу с заданного начального состояния, будет переходить по цепочке $\sigma \in \Sigma^*$ в заключительное состояние, соответствующее классу для σ.

Определение 9

Акцептором с конечным числом состояний для K классов называется $(K+4)$-мерный вектор $\mathscr{A} = (S, \Sigma, \delta, s_0, A_1, \ldots, A_K)$, где S — конечное множество состояний, $\delta \subseteq (S \times \Sigma) \times S$ — отношение переходов, $s_0 \in S$ — начальное состояние, $A_k \subseteq S$, $k = 1, \ldots, K$ суть взаимно исключающие друг друга подмножества заключительных состояний, сопоставленных соответственно классам $1, \ldots, K$. Следовательно, при $i \neq j$ имеет место $A_i \cap A_j = \varnothing$. Когда δ задано всюду на $S \times \Sigma \times S$, говорят, что $\mathscr{A}$ — *полностью определенный акцептор*. В противном случае $\mathscr{A}$ — *частично заданный акцептор*. Если δ однозначно, то $\mathscr{A}$ называется *детерминированным*, в противном случае — *недетерминированным*. Если существует класс k, такой, что по цепочке $\sigma \in \Sigma^*$ можно достичь состояние A_k посредством отношения переходов δ, то говорят, что σ *допускается* по классу k. В противном случае она *отвергается*.

По отношению обучающих данных T можно построить акцептор с конечным числом состояний. Если s_0 — начальное состояние и $(\sigma_1, \sigma_2, \ldots, \sigma_N, k)$ — первая пара вида цепочка-класс из T, то мы определим отношение переходов δ так, чтобы оно включало в себя (s_{i-1}, σ_i, s_i), $i = 1, \ldots, N$, и s_N включим в A_k. Затем возьмем следующую пару вида цепочка-класс из T и поступим также, исходя из состояния s_0 и переходя в следующее еще неиспользованное состояние вместо s_1. Результирующий акцептор с конечным числом состояний окажется частично определенным и в общем случае недетерминированным. Этот акцептор легко можно преобразовать в детерминированный путем итеративного слияния пар состояний r и t, если $(s, \sigma, r) \in \delta$ и $(s, \sigma, t) \in \delta$. Грамматика, соотнесенная с этим акцептором с конечным числом состояний, известна как каноническая определенная автоматная грамматика [10].

Акцептор, построенный описанным выше способом, будет правильно относить каждую наблюденную цепочку из $\Sigma^{*\prime}$ к ее истинному классу. Однако цепочки, не принадлежащие $\Sigma^{*\prime}$, будут отвергаться. Процесс восстановления грамматики предназначен для обобщения построенного акцептора с конечным числом состояний, который может допускать намного больше цепочек, чем первоначально наблюденные цепочки. Основой процесса восстановления грамматики для случая регулярных грамматик

служит слияние или комбинирование состояний в исходном акцепторе для K классов [23] или вспомогательных символов в правилах подстановки грамматики [8]. Эти два процесса по существу эквивалентны друг другу, и мы сконцентрируем наше внимание на методе слияния состояний. Правило слияния состояний заключается в том, что два состояния комбинируются, если для некоторого входа состояния, следующие за ними, могут быть скомбинированы и результирующий акцептор окажется детерминированным, причем состояния акцептора, соответствующие различным классам, никогда не комбинируются.

Мы можем считать, что комбинирование состояний множества S осуществляется при помощи функции h: $S \to W$. Если для некоторого $w \in W$ существуют s, $s' \in S$, удовлетворяющие условию $h(s) = W = h(s')$, то говорят, что состояния s и s' комбинируются посредством h. Конечно, не все пары состояний могут быть скомбинированы, если требуется детерминированность результирующего акцептора. Результирующий акцептор будет детерминированным тогда и только тогда, когда комбинирующая функция сохраняет переходы состояний. Функция сохраняет переходы тогда и только тогда, когда она комбинирует два состояния, которые имеют последующие состояния по цепочке $\sigma \in \Sigma^*$, причем она должна комбинировать и эти последующие состояния.

Определение 10

Пусть $\delta \subseteq (S \times \Sigma) \times S$ и h: $S \to W$. h называется функцией, сохраняющей переходы δ, тогда и только тогда, когда $(s_1, \sigma, s_2) \in \delta$, $(s_1', \sigma, s_2') \in \delta$, $w_1 = h(s_1)$, $w_2 = h(s_2)$ и из $w_1 = h(s_1')$ следует $w_2 = h(s_2')$.

Следующее предложение устанавливает, что результирующий акцептор является детерминированным тогда и только тогда, когда комбинирующая функция сохраняет переходы.

Предложение 5. Пусть $\delta \subseteq (S \times \Sigma) \times S$ и h: $S \to W$. Определим $\gamma \subseteq (W \times \Sigma) \times W$ посредством равенства $\gamma = \{(w_1, \sigma, w_2) \in W \times \Sigma \times W \mid$ существуют s_1, $s_2 \in S$, такие, что $(s_1, \sigma, s_2) \in \delta$ и $w_1 = h(s_1)$, $w_2 = h(s_2)\}$: h является функцией, сохраняющей переходы δ тогда и только тогда, когда γ однозначно.

Доказательство. Предположим, что h сохраняет переходы и $(w_1, \sigma, w_2) \in \gamma$, $(w_1, \sigma, w_2') \in \gamma$. Тогда существуют s_1, $s_2 \in S$, такие, что $(s_1, \sigma, s_2) \in \delta$, $w_1 = h(s_1)$, и существуют s_1', $s_2' \in S$, такие, что $(s_1', \sigma, s_2') \in \delta$, $w_1 = h(s_1')$ и $w_2' = h(s_2')$. Поскольку h сохраняет переходы, $(s_1, \sigma, s_2) \in \delta$, $(s_1', \sigma, s_2') \in \delta$ и из соотношений $w_1 = h(s_1) = h(s_1')$ и $w_2 = h(s_2)$ следует, что $w_2 = h(s_2')$. Следова-

тельно, $w_2 = w_2'$ и тем самым γ однозначно. Обратно, допустим, что γ однозначно и $(s_1, \sigma, s_2) \in \delta$, $(s_1', \sigma, s_2') \in \delta$, $w_1 = h(s_1) = h(s_1')$ и $w_2 = h(s_2)$. Пусть $w_2 = h(s_2')$. Тогда по определению $\gamma(w_1, \sigma, w_2) \in \gamma$, $(w_1, \sigma, w_2') \in \gamma$. Поскольку γ однозначно, $w_2 = w_2'$. Тем самым h оказывается функцией, сохраняющей переходы δ.

Функции, которые сохраняют переходы и не комбинируют заключительные состояния, соответствующие различным классам, не только гарантируют однозначность результирующего акцептора, но также гарантируют сохранение всей структуры переходов состояний. Подобные функции называются гомоморфизмами.

Определение 11

Пусть $\mathscr{A} = (S, \Sigma, \delta, s_0, A_1, \ldots, A_K)$ и $\mathscr{B} = (T, \Sigma, \gamma, t_0, B_1, \ldots, B_K)$ — два акцептора с конечным числом состояний для K классов. Функция h: $S \to T$ называется гомоморфизмом из $\mathscr{A}$ в $\mathscr{B}$ тогда и только тогда, когда

1) из того, что $(s_1, \sigma, s_2) \in \delta$ и $h(s_2) = w_2$, следует существование $w_1 \in W$, такого, что $(w_1, \sigma, w_2) \in \gamma$ и $w_1 = h(s_1)$,

2) $t_0 = h(s_0)$,

3) $B_k = h(A_k)$, $k = 1, \ldots, K$.

Предложение 6. Пусть $\mathscr{A} = (S, \Sigma, \delta, s_0, A_1, \ldots, A_K)$ — акцептор с конечным числом состояний для K классов, h — функция, сохраняющая переходы, отображающая множество S на множество W, которая удовлетворяет условию: если $i \neq j$, то $h(A_i) \cap h(A_j) = \varnothing$. Определим $\gamma \subseteq (W \times \Sigma) \times W$ посредством равенства $\gamma = \{(w_1, \sigma, w_2) \in W \times \Sigma \times W \mid$ существуют $s_1, s_2 \in S$, такие, что $w_1 = h(s_1)$, $w_2 = h(s_2)$ и $(s_1 \sigma, s_2) \in \delta\}$. Тогда $\mathscr{B} = (W, \Sigma, \gamma, h(s_0), h(A_1), \ldots, h(A_K))$ — детерминированный акцептор с конечным числом состояний, который является гомоморфным образом A.

Доказательство. По предложению 5 γ однозначно, что гарантирует детерминированность $\mathscr{B}$. Пусть $(s_1, \sigma, s_2) \in \delta$ и $h(s_2) = w_2$, $w_1 = h(s_1)$. Тогда из условий $(s_1, \sigma, s_2) \in \delta$, $h(s_2) = w_2$ и $h(s_1) = w_1$ по определению γ следует, что $(w_1, \sigma, w_2) \in \gamma$. Наконец, при $i \neq j$ имеет место $h(A_i) \cap h(A_j) = \varnothing$. Тем самым $\mathscr{B}$ — детерминированный акцептор с конечным числом состояний.

Чтобы показать, что $\mathscr{B}$ — гомоморфный образ $\mathscr{A}$, предположим, что $(s_1, \sigma, s_2) \in \delta$ и $h(s_2) = w_2$. Пусть $w_1 = h(s_1)$. Тогда по определению γ $(w_1, \sigma, w_2) \in \gamma$. Поэтому h удовлетворяет условиям 1, 2 и 3 определения и h является гомоморфизмом из $\mathscr{A}$ в $\mathscr{B}$. Так как h — отображение S на W, $\mathscr{B}$ — гомоморфный образ $\mathscr{A}$ при h.

Важность комбинирования состояний посредством гомоморфизма очевидна, когда мы принимаем во внимание то, что множество цепочек, допущенных в любой класс гомоморфного образа акцептора, должно включать в себя те цепочки, которые допускаются в соответствующий класс исходного акцептора. Тем самым для обобщения множества цепочек, допущенных при помощи любого акцептора с конечным числом состояний, нам нужно лишь найти гомоморфизм, который комбинирует состояния способом, аналогичным способу сохранения переходов. В следующем предложении доказывается, что множество цепочек, допущенных при помощи гомоморфного образа акцептора, по крайней мере не у́же множества цепочек, допущенных при помощи исходного акцептора.

Предложение 7. Пусть h — гомоморфизм акцептора с конечным числом состояний $\mathscr{A} = (S, \Sigma, \delta, s_0, A_1, \ldots, A_K)$ в $\mathscr{B} = (T, \Sigma, \gamma, t_0, B_1, \ldots, B_K)$. Допустим, что для некоторой $\sigma \in \Sigma^*$ имеет место включение $(s_0, \sigma, s_f) \in \delta$ и для некоторого $k \in \{1, \ldots, K\}$ — включение $s_f \in A_k$. Тогда существует $t_f \in T$, такой, что $(t_0, \sigma, t_f) \in \gamma$ и $t_f \in B_k$.

Доказательство. Пусть $t_f = h(s_f)$. Так как $t_0 = h(s_0)$ и $(s_0, \sigma, s_f) \in \delta$, h — гомоморфизм из $\mathscr{A}$ в $\mathscr{B}$, то $(t_0, \sigma, t_f) \in \gamma$. Когда h — гомоморфизм, то $B_k = h(A_k)$ и $s_f \in A_k$, отсюда мы должны иметь $t_f = h(s_f) \in B_k$.

Исходя из этой информации об акцепторах с конечным числом состояний и их гомоморфизмах, нетрудно понять, как можно построить покрытие $\mathscr{L}$ множества Σ^*. Известно, что с каждым акцептором с конечным числом состояний связано разбиение множества Σ^* со следующим свойством: две цепочки принадлежат общему слою разбиения, если их переходы состояний совпадают (предложение 8). Справедливо также то, что слои указанного разбиения являются гомоморфными образами Σ^* (предложения 9 и 10). Тем самым мы можем определить $\mathscr{L}$ как совокупность слоев, принадлежащих всем разбиениям множества Σ^*, определенным при помощи акцепторов с конечным числом состояний, которые заданы всюду гомоморфными образами акцепторов с конечным числом состояний, определенных посредством отношения обучающих данных. При желании мы можем ограничить число или тип гомоморфизмов. $\mathscr{L}$ может состоять из всех слоев разбиений множества Σ^*, определенных при помощи гомоморфизмов в заданном классе гомоморфизмов на исходном акцепторе с конечным числом состояний. Поскольку гомоморфизм акцептора с конечным числом состояний определялся не для комбинирования заключительных состояний, соответствующих различным классам, сужение $\mathscr{L}'$ совокупности $\mathscr{L}$ в данном случае равно $\mathscr{L}$.

Предложение 8. Пусть $\delta \subseteq (S \times \Sigma) \times S$. Определим $R = \{(a, b) \in \Sigma^* \times \Sigma^* | (s, a, t) \in \delta$ тогда и только тогда, когда $(s, b, t) \in \delta\}$. Тогда R — отношение эквивалентности, заданное на Σ^*.

Доказательство. R рефлексивно, симметрично и транзитивно.

Предложение 9. Пусть $\delta \subseteq (S \times \Sigma) \times S$. Определим $R = \{(a, b) \in \Sigma^* \times \Sigma^* | (s, a, t) \in z$ тогда и только тогда, когда $(s, b, t) \in \delta\}$. Пусть $[a]$ обозначает класс эквивалентности для элемента a. Тогда классы эквивалентности R образуют моноид по бинарной операции, заданной как $[a] \times [b] = [ab]$.

Доказательство. Во-первых, эта операция определена корректно, так как если $a' \in [a]$ и $b' \in [b]$, то мы должны иметь $a'b' \in [ab]$. Чтобы доказать это, рассмотрим $a' \in [a]$ и $b' \in [b]$. Тогда $(s, a, r) \in \delta$, если и только если $(s, b, r) \in \delta$, и $(r, a', t) \in \delta$, если и только если $(r, b', t) \in \delta$. Отметим, что из условия $(s, ab, t) \in \delta$ следует существование $r \in S$, удовлетворяющего условиям $(s, a, r) \in \delta$ и $(r, b, t) \in \delta$. Так как из того, что $(s, a, r) \in \delta$ следует $(s, a', r) \in \delta$ и из $(r, b, t) \in \delta$ вытекает $(r, b' t) \in \delta$, мы имеем $(s, a', r) \in \delta$ и $(r, b', t) \in \delta$. Однако из того, что $(s, a', r) \in \delta$ и $(r, b', t) \in \delta$, следует, что $(s, a', b', t) \in \delta$. Тем самым из $(s, ab, t) \in \delta$ следует $(s, a', b', t) \in \delta$. Аналогичными рассуждениями доказывается обратное утверждение. Поэтому $(s, ab, t) \in \delta$ тогда и только тогда, когда $(s, a', b', t) \in \delta$, так что $a'b' \in [a, b]$.

Во-вторых, эта операция ассоциативна, поскольку

$$([a] \cdot [b]) \cdot [c] = ([ab]) \cdot [c] = [(ab) \cdot c] = [a \cdot (bc)] = \\ = [a] \cdot ([bc]) = [a] \cdot ([b] \cdot [c]).$$

Наконец, тождество является единичным классом $[\lambda]$, так как $[a] \cdot [\lambda] = [a\lambda] = [a]$, $[\lambda] \cdot [a] = [\lambda a] = [a]$.

Предложение 10. Пусть $\delta \subseteq (S \times \Sigma) \times S$. Определим $R = \{(a, b) \in \Sigma^* \times \Sigma^* | (s, a, t) \in \delta$ тогда и только тогда, когда $(s, b, t) \in \delta\}$. Пусть $\mathscr{E}$ — совокупность классов эквивалентностей R и $[\sigma]$ обозначает класс эквивалентности для σ. Тогда функция h: $\Sigma^* \to \mathscr{E}$, заданная посредством соотношения $h[\sigma] = [\sigma]$, является гомоморфизмом из моноида Σ^* в моноид $\mathscr{E}$.

Доказательство. $h[ab] = [ab] = [a] \cdot [b] = h(a) \cdot h(b)$.

Проиллюстрируем этот метод структурного распознавания образов на примере. Пусть $\Sigma = \{A, B, \lambda\}$, где λ — пустая цепочка и пространством объектов служит Σ^*. Множество наблюденных объектов $\Sigma^{*\prime}$ задается так:

$$\Sigma^{*\prime} = \{A, AB, BA, BBA, AAA, ABA, ABAA, AA, B, \lambda, ABAB\}.$$

Допустим, что имеются два класса $C=\{0, 1\}$ и отношение обучающих данных $T \subseteq \Sigma^{*\prime} \times C$ задается следующей таблицей:

T		T	
объект	класс	объект	класс
A	0	$ABAA$	0
AB	0	AA	1
BA	0	B	1
BBA	0	λ	1
AAA	0	$ABAB$	1
ABA	0		

Диаграмма переходов для детерминированного акцептора с конечным числом состояний, которая может быть выведена прямо из отношения обучающих данных T, приведена на рис. 5. Мы

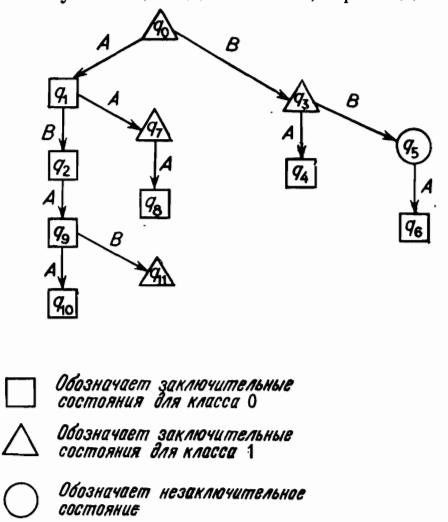


Рис. 5. Диаграмма переходов для акцептора с конечным числом состояний, определенного посредством отношения обучающих данных.

построим покрытие $\mathscr{L}$ совокупности Σ^*, включая в $\mathscr{L}$ классы эквивалентности входных моноидов двух из семнадцати всюду определенных акцепторов с конечным числом состояний, которые являются гомоморфными образами исходного акцептора.

Эти два гомоморфизма и акцепторы гомоморфных образов приведены на рис. 6. Классы эквивалентности входных моноидов для этих акцепторов гомоморфных образов перечислены на рис. 7. Квадратные скобки обозначают класс эквивалентности цепочки, заключенной в них. Нижний индекс у скобок указывает, относится ли данный класс эквивалентности к акцептору гомоморфного образа 1 или 2. Рис. 7 также дает для каждого

класса эквивалентности регулярное выражение, указывающее множество его цепочек. Для упрощения членов в регулярных выражениях метки классов эквивалентности, чьи регулярные выражения были определены ранее, используются вместо соответствующих регулярных выражений в тех случаях, когда это удобно.

За исключением классов эквивалентности $[BB]_1$, $[ABB]_1$, $[AAB]_1$, $[BAB]_2$, $[BABA]_2$, $[ABABA]_2$, $[BBAB]_2$, все классы

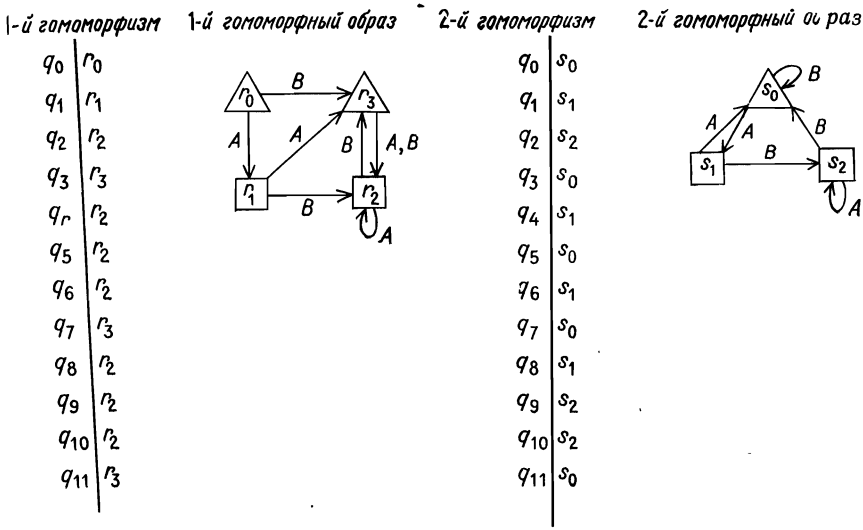


Рис. 6. Два гомоморфизма и их гомоморфные образы акцептора, определенного посредством отношений обучающих данных (см. рис. 5).

эквивалентности имеют непустые пересечения с множеством наблюденных цепочек. Более того, каждый класс эквивалентности, имеющий непустое пересечение с множеством наблюденных цепочек, является обобщающим множеством, так как он содержит наблюденные цепочки лишь из одного класса. Отношение $F' \subseteq \Sigma^{*\prime} \times \mathscr{L}'$, которое связывает наблюденные цепочки с обобщающими множествами, задается следующей таблицей:

объект F'	слой покрытия	объект Γ'	слой покрытия
A	$[A]_1$, $[A]_2$	AAA	$[BA]_1$, $[AB]_2$
AB	$[AB]_1$, $[AB]_2$	AA	$[AA]_1$, $[\lambda]_2$
BA	$[BA]_1$, $]BA]_2$	B	$[B]_1$, $[B]_2$
BBA	$[BA]_1$, $[BBA]_2$	λ	$[\lambda]_1$, $[\lambda]_2$
AAA	$[BA]_1$, $[A]_2$	$ABAB$	$[BAB]_1$, $[ABAB]_2$
ABA	$[BA]_1$ $[ABA]_2$		

Отношение Q, связывающее обобщающие множества и классы, определяется равенством $Q = I_{\Sigma^*/} \cdot (F', T)$. Отношение Q задается следующей таблицей:

Q слой покрытия	класс	Q слой покрытия	класс	Q слой покрытия	класс
$[\lambda]$	1	$[BA]_1$	0	$[AB]_2$	0
$[A]_1$	0	$[BAB]_1$	1	$[BA]_2$	0
$[B]_1$	1	$[\lambda]_2$	1	$[ABA]_2$	0
$[AB]_1$	0	$[A]_2$	0	$[ABAB]_2$	1
$[AA]_1$	1	$[B]_2$	1	$[BBA]_2$	0

Решающее правило D, которое является бинарным отношением, связывающим объекты с классами, определяется посредством равенства $D = F' \cdot (I_{\Sigma^*}, Q)$. Поэтому мы можем записать D в виде множества всех пар объект — класс, удовлетворяющих некоторым условиям

$$D = \{(\sigma, c) \in \Sigma^* \times c \mid$$
если $c = 0$, то $\sigma \in [A]_1, [AB]_1, [BA]_1,$
$$[A]_2, [AB]_2, [BA]_2, [ABA]_2, [BBA]_2,$$
если $c = 1$, то $\sigma \in [\lambda]_1, [B]_1, [AA]_1, [BAB]_1, [\lambda]_2, [B]_2, [ABAB]_2\}.$

Поскольку не все классы эквивалентности обязательно должны содержать наблюденные объекты, классы эквивалентности, не используемые при задании D, могут порождать некоторые цепочки из Σ^*, которые не могут быть отнесены ни к какому классу. Так как классы эквивалентности, определенные исходя из различных гомоморфизмов, могут перекрываться, для D возможен случай, когда некоторые цепочки могут относиться более чем к одному классу.

Имеется прямая связь между построением решающего правила и методами восстановления грамматик Пао [23] и Фельдмана [8]. Эти методы исходят либо из акцептора с конечным числом состояний, либо из грамматики, которая порождает лишь наблюденные цепочки. Восстановление грамматики осуществляется или путем влияния состояний акцепторов с конечным числом состояний, как это мы продемонстрировали на примере, или путем слияния вспомогательных символов грамматики. Поскольку существует естественный изоморфизм между акцепторами с конечным числом состояний и регулярными грамматиками, эти два метода слияния равносильны одному и тому же

$$[\lambda]_1 = \lambda$$

$$[A]_1 = A$$

$$[B]_1 = B(BB)^*$$

$$[BB]_1 = [B]_1[B]_1$$

$$[AB]_1 = A[B]_1$$

$$[ABB]_1 = AB[B]_1$$

$$[AA]_1 = AA(BB)^*$$

$$[AAB]_1 = [AA]_1 B$$

$$[BA]_1 = ([AB]_1 + [B]_1 + [BB]_1 + [AA]_1)\, A(A^* + B(A + B))^*$$

$$[BAB]_1 = [BA]_1\, [B]_1$$

$$[\lambda]_2 = (AA)^*$$

$$[A]_2 = A(AA)^*$$

$$[B]_2 = [\lambda]_2\, B((([A]_2\, B[A]_2\, B[\lambda]_2) + [\lambda]_2)^*$$

$$[AB]_2 = [A]_2\, [B]_2$$

$$[BA]_2 = [B]_2\, [A]_2$$

$$[BAB]_2 = [BA]_2\, [B]_2$$

$$[BABA]_2 = [BAB]_2\, [A]_2$$

$$[ABA]_2 = [AB]_2\, [A]_2$$

$$[ABAB]_2 = [ABA]_2\, [B]_2$$

$$[ABABA]_2 = [ABAB]_2\, [A]$$

$$[BB]_2 = [A]_2[B]_2\, ([B]_2 + [A]_2[B]_2[B]_2)\, ((B + (AA)^*) + [A]_2[B]_2 A^*[B]_2)^* + [B]_2\, ([B]_2 + [A]_2[B]_2[B]_2)$$

$$[BBA]_2 = [BB]_2\, [A]_2$$

$$[BBAB]_2 = [BBA]_2\, B\, A^*$$

Рис. 7. Классы эквивалентности и соответствующие им регулярные выражения входного моноида для двух акцепторов с конечным числом состояний из рис. 6.

основному процессу, даже если слияние вспомогательных символов может привести к недетерминированным акцепторам с конечным числом состояний.

5. РАЗЛИЧИЯ ОБЪЕКТОВ, ПРЕДСТАВЛЕННЫХ РАЗМЕЩЕНИЯМИ

В этом разделе мы вводим понятие размещения [14] в качестве структуры данных и показываем, как его можно использовать в структурном подходе к распознаванию образов подобно цепочке и N-мерному вектору.

5.1. Размещение

Пусть A — множество элементов, размещения которых описывается. Каждая группа связанных элементов из A размечается меткой из множества меток L. Пусть R — размеченное N-арное отношение, которое состоит из размеченных N-ок элементов из A, связанных между собой.

Определение 12

Простым размещением N-го порядка называется тройка (R, A, L), где $R \subseteq A^N \times L$. Мы можем использовать R для задания размещения (R, A, L), когда ясно, о каких именно множествах A и L идет речь.

N-мерный вектор является частным случаем размещения первого порядка. Чтобы увидеть это, предположим, что $(x_1, \ldots, x_N)$ — данный N-мерный вектор. Пусть в качестве L выбрано множество целых чисел, используемых для меток. Рассмотрим множество $A = \{x_1, x_2, \ldots, x_N\}$ и определим отношение $R \subseteq A \times L$ посредством равенства $R = \{(\alpha, l) \in A \times L \mid x_l = \alpha\}$. Цепочка на самом деле является N-кой переменной длины, так что и она представляет собой частный случай размещения первого порядка.

Мы определим общее размещение как множество связанных друг с другом размещений, имеющих общее множество меток. Это понятие позволяет совместно рассматривать взаимосвязи элементов общего множества, имеющих различные порядки.

Определение 13

Общим размещением называется множество простых размещений различных порядков, определенных на одном и том же множестве и имеющих общее множество меток. Если имеются K простых размещений в размещении A, то $A = \{R_1, R_2, \ldots, R_K; A, L\}$, где $R_k \subseteq A^{N_k} \times L$.

Поскольку размещения выбираются для представления наших объектов и мы уже показали, что структурное распознавание может рассматриваться как процесс обобщения посредством гомоморфных образов наблюденных обучающих объектов, мы должны определить операцию композиции и гомоморфизм. С этой целью мы будем использовать ту же идею, изложенную в общей форме в разд. 2. Композиция размещения при помощи бинарного отношения дает размещение, содержащее все размеченные N-ки данного размещения, покомпонентно преобразованного посредством этого бинарного отношения. Единственное отличие этой композиции от композиции, введенной в разд. 2, состоит в том, что в разд. 2 указанное преобразование осуществляется с помощью бинарного отношения для каждой компоненты упорядоченных пар в данном отношении, причем различным

компонентам соответствуют различные бинарные отношения. Здесь мы считаем, что данное отношение N-арное, и требуем, чтобы все компоненты были преобразованы операцией композиции посредством бинарного отношения одинаковым способом.

Определение 14

Пусть $A = \{R_1, \ldots, R_K; A, L\}$ — размещение и $H \subseteq A \times B$. *Композиция* размещения A с H определяется как размещение, задаваемое в виде

$$A \cdot H = B = \{S_1, S_2, \ldots, S_k; B, L\},$$

где

$$S_k = \{(b_1, b_2, \ldots, b_{N_k}, l) \mid (a_1, \ldots, a_{N_k}, l) \in R_k,$$
$$(a_n, b_n) \in H, \quad n = 1, \ldots, N_k\}.$$

Понятие гомоморфизма, введенное в разд. 2, выражало требование, чтобы образ композиции содержался в данном отношении. Мы используем то же самое понятие применительно к размещениям.

Определение 15

Размещение $A = \{R_1, \ldots, R_K; A, L\}$ *содержится* в размещении $D = \{T_1, \ldots, T_K; A, L\}$ тогда и только тогда, когда $R_k \subseteq T_k$, $k = 1, \ldots, K$. В этом случае мы записываем $A \subseteq D$.

Определение 16

Два размещения $A = \{R_1, \ldots, R_K; A, L\}$ и $B = \{S_1, \ldots, S_N; B, M\}$ *сравнимы*, если числа отношений в обоих размещениях равны ($K = N$), они имеют общее множество меток ($M = L$) и порядки отношений R_K и S_K равны

$$(R_k \subseteq A^{N_k} \times L \quad \text{и} \quad S_k \subseteq B^{N_k} \times L).$$

Определение 17

Пусть $A = \{R_1, \ldots, R_K; A, L\}$ и $B = \{S_1, \ldots, S_K; B, L\}$ — два сравнимых размещения, H: $A \to B$. Функция H называется гомоморфизмом из размещения A в размещение B тогда и только тогда, когда $A \cdot H \subseteq B$.

Мы рассмотрим структурное распознавание образов, заданных в виде простого размещения, а не в виде общего размещения. Можно взять в качестве множества объектов P совокупность всех размещений и считать семейство $\mathscr{L}$ состоящим из элементов, являющихся множествами размещений, каждое из которых характеризуется тем, что оно является прообразом гомоморфизма в признаковое размещение, сопоставленное этому множеству. Признаковые размещения являются гомоморфными образами наблюденных объектов, представленных размещениями. Сужение $\mathscr{L}'$ до $\mathscr{L}''$, отношение Q и решающее правило D могут быть определены так же, как и в разд. 2.

Для нашего примера предположим, что множество классов C равно $\{A, B\}$, множество объектов является совокупностью всех размещений второго порядка на множестве $S = \{a, b, c, d, e, f\}$ с метками $K = \{0, 1\}$, размещения наблюденных объектов — это $P' = \{T_0, T_1\}$ и отношение обучающих данных $T \subseteq P' \times C$.

Ниже приведены определения размещений T_0, T_1 и обучающего отношения T.

$T_0 \subseteq S \times S \times L$

$aa0$
$ba0$
$ca0$
$da0$
$ad1$
$bd1$
$cd1$
$dd1$

$T_1 \subseteq S \times S \times L$

$aa0$
$ba0$
$ce0$
$de0$
$ea0$
$fa0$
$ad1$
$bd1$
$cd1$
$dd1$
$ed1$
$fd1$

T	
размещения для наблюденных объектов	класс
T_0	A
T_1	B

На рис. 8 показано, как именно четыре функции, которые комбинируют элементы S, вместе образуют гомоморфные образы наблюденных обучающих размещений. Обозначим эти гомоморфные образы размещений через A_1, A_2, ..., A_8. Заметим, что размещения A_1, A_2, A_3 и A_6 либо совпадают, либо же изоморфны.

Определим семейство $\mathscr{L}$ как набор множеств, каждое из которых содержит все размещения, имеющие одно из признаковых размещений в качестве своего гомоморфного образа, $\mathscr{L} = \{L \subseteq P \mid$ для каждого размещения $R \in L$ существует функция f и признаковое размещение B, такие, что $R \cdot f = B\}$.

На основе выбранных четырех комбинирующих функций и восьми признаковых размещений, которые они задают, можно считать, что набор $\mathscr{L}$ может содержать до 8 множеств. Однако в силу того, что четыре признаковые размещения изоморфны, $\mathscr{L}$ содержит лишь 5 множеств размещений L_1—L_5, определенных следующим образом:

$L_1 = \{R \in P \mid$ существует функция f, такая, что $R \cdot f = A_1\}$,

$L_2 = \{R \in P \mid$ существует функция f, такая, что $R \cdot f = A_4\}$,

$L_3 = \{R \in P$ существует функция f, такая, что $R \cdot f = A_5\}$,

$L_4 = \{R \in P$ существует функция f, такая, что $R \cdot f = A_7\}$,

$L_5 = \{R \in P$ существует функция f, такая, что $R \cdot f = A_8,\}$

Поскольку L_1 содержит наблюденные обучающие размещения более чем из одного класса, сужение $\mathscr{L}'$ семейства $\mathscr{L}$ со-

Комбинирование функций F	Образ наблюденного размещения T_0 при функции F	Образ наблюденного размещения T_1 при функции F
F_1: $S \to E$	$A_1 \subseteq E \times E \times K$	$A_5 \subseteq E \times E \times K$
$a, b \to v$ $c, d \to w$ $e \to x$ $f \to y$	$vv0$ $wv0$ $vw1$ $ww1$	$vv0$ $xv0$ $yv0$ $vw1$ $ww1$ $xw1$ $yw1$
F_2: $S \to F$	$A_2 \subseteq F \times F \times K$	$A_6 \subseteq F \times F \times K$
$a, c, e \to v$ $b, d, f \to w$	$vv0$ $wv0$ $vw1$ $ww1$	$vv0$ $wv0$ $vw1$ $ww1$
F_3: $S \to H$	$A_3 \subseteq H \times H \times K$	$A_7 \subseteq H \times H \times K$
$a, b \to p$ $c, d \to q$ $e, f \to r$	$pp0$ $qp0$ $pq1$ $qq1$	$pp0$ $qp0$ $p0$ $pq1$ $qq1$ $rq1$
F_4: $S \to G$	$A_4 \subseteq G \times G \times K$	$A_8 \subseteq G \times G \times K$
$a, d, f \to v$ $b, c, e \to w$	$vv0$ $wv0$ $vv1$ $ww1$	$vv0$ $wv0$ $ww0$ $vw0$ $vv1$ $wv1$

Рис. 8. Иллюстрация способа комбинирования функциями F_1, F_2, F_3 и F_4 элементов S и формирования гомоморфных образов A_1—A_8 наблюденного обучающего размещения.

стоит лишь из L_2—L_5 служащих его обобщающими множествами. Отношение F', которое связывает размещения с обобщающими множествами, задается в виде

$$F' = \{(p,\ L) \in P \times \mathscr{L}' \mid p \in \mathscr{L}\}.$$

Отношение Q, которое сопоставляет обобщающие множества с классами, определяется соотношением

$$Q = I_{P'} \cdot (F',\ T) = \{(L_2,\ A),\ (L_3,\ B),\ (L_4,\ B),\ (L_5,\ B)\}.$$

Решающее правило D, которое относит объекты к классам, определяется соотношением $D = F' \cdot (I_P,\ Q)$ и может быть записано в виде $D = \{(p,\ c) \in P \times C \mid$ если $c = A$, то $p \in L_2$,

если $c = B$, то $p \in L_3 \cup L_4 \cup L_5\}$.

6. ЗАКЛЮЧЕНИЕ

В этой работе мы описали подход структурного распознавания образов, исходя из алгебры отношений. Мы высказали мнение о том, что обобщение наблюденных объектов осуществляется посредством явного или неявного использования ограниченного покрытия на пространстве объектов. Сужение покрытия производится с тем, чтобы гарантировать соотнесение лишь одного класса с каждым элементом покрытия. Вид покрытия зависит от структуры данных объектов: для N-мерных векторов — цилиндрические множества, для цепочек — классы эквивалентностей входных моноидов. Покрытие может быть задано до или после выявления обучающего отношения объектов.

В рамках схемы отношений, в которой мы описали этот процесс, естественное отношение F', связывающее объекты с множествами в ограниченном покрытии, к которым они принадлежат, является обратным гомоморфным образом обучающего отношения объектов. Решающее правило D, которое сопоставляет объекты классам, является гомоморфным образом отношения F'. Гомоморфизмы дают возможность не только построить D по заданному обучающему отношению объектов и покрытию пространства объектов, но, как это было сделано для N-мерных векторов и цепочек, само покрытие может быть определено в терминах гомоморфизмов на пространстве объектов. Тем самым гомоморфизмы играют существенную роль при построении решающих правил структурного распознавания образов.

Наконец, мы показали, что, поскольку структурное распознавание образов рассматривается с этих позиций, можно предложить и естественно использовать другие типы реляционных структур данных объектов. Для иллюстрации этого факта мы определили размещение в виде размеченного N-арного отношения и показали, как именно размещения могут использоваться в качестве структуры данных объекта. В этом случае пространством объектов является множество всех размеченных N-арных отношений. Множества из покрытия определяются посредством отобранных гомоморфных образов обучающих размещений объектов. Тогда решающее правило является гомоморфным образом отношения, которое связывает объекты с классами в ограниченном покрытии.

Благодарности. Хочу поблагодарить Линду Дж. Шапиро за помощь и идеи, которыми она делилась со мной на этапе подготовки этой статьи, и Линн Эртебати за перепечатку рукописи.

ЛИТЕРАТУРА

1. Biermann A. W., Feldman J. A. On the synthesis of finite state machines from samples of their behavior, IEEE Trans. Comput., **21**, 592—597 (1972).
2. Brainerd W. Free generating regular systems, Inf. Control, **14**, 217—231 (1969).

3. Dacey M. F. The syntax of a triangle and some other figures, Pattern Recognition, **2**, 11—31 (1970).
4. Davis R., Buchanan B., Shortliffe E. Production rules as a representation for a knowledge-based consultation program, Art. Intelligence, **8**, 15—45 (1977).
5. Donor J. E. Tree acceptors and some of their applications, J. Comput. Syst. Sci., **4**, 406—451 (1970).
6. Feder J. Plex languages, Inf. Sci., **3**, 225—241 (1971).
7. Feder J. Languages of encoded line patterns, Inf. Control, **13**, 230—244 (1968).
8. Feldman J. A. First thoughts on grammatical inference, Stanford Artificial Intelligence Project Memo 55, Stanford University, Stanford, California (1967).
9. Fu K. S. Synactic Methods in Pattern Recognition, Academic Press, New York, 295 pp. (1974).
10. Fu K. S., Booth T. L Grammatical inference: introduction and survey — I, IEEE Trans. Syst., Man. Cybern., **5**, 95—111, January (1975).
11. Fukanaga K. Introduction to Statistical Pattern Recognition, Academic Press, New York (1972).
12. Ginsburg S. The Mathematical Theory of Context-Free Language, McGraw-Hill, New York (1966).
13. Haralick R. M. The pattern recognition problem from the perspective of relation theory, Pattern Recognition, **7**, 67—79 (1975).
14. Haralick R. M. Structural pattern recognition, arrangements and theory of covers, IEEE Conf. on Pattern Recognition and Image Processing, Troy, New York, June (1977).
15. Kirsch R. A. Computer interpretation of English text and patterns, IEEE Trans. electron. Comput., **13**, 362—376 (1964).
16. McCluskey E. J., Jr. Minimization of boolean functions, Bell System Tech. J., **35**, 1417—1444, November (1956).
17. Michalski R. S. On the quasi-minimal solution of the general covering problem, Proc. 5th Intern. Symposium, Yugoslavia, Bled, 8—11 Okt.
18. Michalski R. S., McCormick B. H. Interval generalization of switching theory, Proc. 3rd Ann. Houston Conf. on Computer and System Science, Houston, Texas, pp. 231—246 (1971).
19. Michalski R. S., AQVAL/1 — Computer implemental of a variable-valued logic system VL_1 and example of its application to pattern recognition, Int. Jnt Conf. on Pattern Recognition, Washington, D. C., pp. 3—17 (1973).
20. Michalski R. S. Variable-valued logic: System VL_1, Proc. 1974 Int. Symp. on Multiple-Valued Logic, West Virginia State University, Morgantown, West Virginia, pp. 323—346 (1974).
21. Milgram D. M., Rosenfeld A. Array automata and array grammars, IFIP Congr., **71**, Booklet TA-2, North-Holland, Amsterdam, 166—173 (1971).
22. Narasimhan R. N. Syntax-directed interpretation of classes of pictures, Comm. ACM **9**, 166—173 (1966).
23. Pao T. W. A solution to the syntactical induction-inference problem for a non-trivial subset of context-free languages, Interim Technical Report 69-19, Moore School of Engineering, University of Pennsylvania, Philadelphia, Pennsylvania (1969).
24. Pfaltz J. L., Rosenfeld A. Web grammars, Proc. 1st Int. Jnt Conf. om Artificial Intelligence, Washington, D. C., May, 609—619 (1969).
25. Quine W. V. A way to simplify truth function, Am. Math. Monthly, **62**, 627—631, November (1955).
26. Shaw A. C. A formal picture description scheme as a basis for a picture processing system, Inf. Control, **14**, 9—52 (1969).
27. Shortliffe E. H. Computer-Based Medical Consultation: MYCIN, Elsevier, New York, 264 pp. (1976).

Содержание

Математические вопросы

Теория распознавания

КИБЕРНЕТИЧЕСКИЙ СБОРНИК

Выпуск 19

Научн. ред. Л. Н, Бабынина Младш. научн. ред. Э. Г. Иванова Художник Н. К. Сапожников. Художественный редактор В. И. Шаповалов. Технический редактор З. И. Резник Корректор Н. А. Гиря

ИБ № 3119

Сдано в набор 26.03.82. Подписано к печати 13.12.82. Формат 60×90/16. Бумага типографская № 2. Гарнитура латинская. Печать высокая. Объем 6,25 бум. л. Усл. печ. л. 12,5. Усл. кр. отт. 12,5. Уч.-изд. л. 10,68. Изд. № 1/2202. Тираж 3500 экз. Зак. 152. Цена 1 р. 90 к.

ИЗДАТЕЛЬСТВО «МИР» 129820, Москва, И-110, ГСП 1-й Рижский пер., 2.

Ленинградская типография № 2 головное предприятие ордена Трудового Красного Знамени Ленинградского объединения «Техническая книга» им. Евгении Соколовой Союзполиграфпрома при Государственном комитете СССР по делам издательств, полиграфии и книжной торговли. 198052, г. Ленинград, Л-52, Измайловский проспект, 29