СИБИРСКИЙ ЖУРНАЛ ИССЛЕДОВАНИЯ ОПЕРАЦИЙ
Июль – сентябрь 1994. Том 1, № 3, 49–59

УДК 519.171

КАНОНИЧЕСКИЕ РАЗБИЕНИЯ ГРАФОВ И ИХ ПРИМЕНЕНИЕ ДЛЯ КОДИРОВАНИЯ ГРАФОВ*)

В. В. Лозин

Разбиение множества вершин графа G называется каноническим, если любые два класса разбиения порождают в G либо несвязный, либо дополнительный к несвязному граф. Таким образом, каждому каноническому разбиению графа G может быть поставлен в соответствие другой граф, вершины которого взаимно однозначно соответствуют классам разбиения, а ребра — парам классов, порождающих в G подграф, дополнительный к несвязному. Этот граф называется канонической основой графа G, если единственным его каноническим разбиением является тривиальное разбиение, т. е. разбиение на одновершинные классы. Показано, что каноническая основа произвольного графа единственна с точностью до изоморфизма. Предложен полиномиальный алгоритм поиска канонической основы и рассмотрено применение канонических разбиений для оптимального кодирования графов.

Пусть $G = (V, E)$ — обыкновенный граф с множеством вершин V и множеством ребер E, R_l — симметричное бинарное отношение на $\langle l \rangle = \{1, 2, \ldots, l\}$. Разбиение $\{V_i \mid i \in \langle l \rangle\}$ множества вершин V графа G называется R_l-каноническим, если для любых $(i,j) \in R_l$, $x_1, x_2 \in V_i$, $y_1, y_2 \in V_j$ выполняется следующее условие:

$$(x_1, y_1) \in E \text{ тогда и только тогда, когда } (x_2, y_2) \in E.$$

Например, если $R_l = \langle l \rangle^2$ — универсальное отношение на $\langle l \rangle$, то R_l-канонические разбиения — это канонические разбиения, описанные в [1]. Если $R_l = \{(i,i) \mid i \in \langle l \rangle\}$, то R_l-каноническими являются, в частности, разбиения вершин графа на клики, независимые множества. Заметим, что поиск таких разбиений при $l > 2$ составляет NP-полную задачу в классе всех графов [2].

В данной работе рассматриваются R_l-канонические разбиения (в дальнейшем называемые просто каноническими) для $R_l = \{(i,j) \mid i \neq$

*) Работа выполнена при финансовой поддержке Российского фонда фундаментальных исследований (код проекта 93-01-00491).

$j;\ i, j \in \langle l \rangle\}$. Таковыми являются, например, разбиения несвязных графов на компоненты связности. Очевидно, что каноническим для любого графа G является тривиальное разбиение, т. е. разбиение на классы, каждый из которых содержит одну вершину. Нетрудно видеть, что любые два класса канонического разбиения порождают либо несвязный, либо дополнительный к несвязному граф. Таким образом, каждому каноническому разбиению W графа G соответствует некоторый граф $\mathcal{T}_W$, вершинами которого являются классы разбиения W, а ребра образованы парами классов, порождающих в G подграф, дополнительный к несвязному. Очевидно, что всякое каноническое разбиение графа $\mathcal{T}_W$ дает соответствующее каноническое разбиение графа G. Граф $\mathcal{T}_W$, не имеющий нетривиальных канонических разбиений, называется канонической основой графа G (точные определения даны ниже).

В работе используются следующие обозначения:

VG — множество вершин графа G,
EG — множество ребер графа G,
$G\langle D \rangle$ — подграф графа G, порожденный множеством $D \subseteq VG$,
$\overline{D} = VG \backslash D$,
$N(v) = \{x \mid x \in VG,\ (v, x) \in EG\}$ — окрестность вершины v,
O_n — пустой граф с n вершинами,
K_n — полный граф с n вершинами.

§ 1. Канонические классы

Определение 1. Множество вершин $U \subset VG$ графа G называется *каноническим классом*, если $\overline{U} \neq \varnothing$ и для любой вершины $v \in \overline{U}$ имеем $N(v) \cap U = \varnothing$ либо $\overline{N(v)} \cap U = \varnothing$.

Множество канонических классов графа G обозначим через $K(G)$. Наиболее типичными примерами канонических классов являются компоненты несвязного графа. Действительно, если K — компонента связности G и $\overline{K} \neq \varnothing$, то $N(v) \cap K = \varnothing$ для любой вершины $v \in \overline{K}$.

Непосредственно из определения следует, что если граф G имеет больше одной вершины, то любая его вершина образует канонический класс. Канонические классы, состоящие более чем из одной вершины, будем называть *нетривиальными*. Граф, у которого каждый канонический класс тривиален, называется *неприводимым*. Очевидно, неприводимыми являются двухвершинные графы O_2 и K_2.

Установим некоторые свойства канонических классов.

Лемма 1. *Если $U_1, U_2 \in K(G)$ и $U_1 \cap U_2 \neq \varnothing$, то $U_1 \cap U_2 \in K(G)$.*

Доказательство. Поскольку $\overline{U}_1 \neq \varnothing$ и $\overline{U}_2 \neq \varnothing$, справедливо соотношение $\overline{U_1 \cap U_2} = \overline{U}_1 \cup \overline{U}_2 \neq \varnothing$. Если $v \in \overline{U}_1$, то ввиду включения $U_1 \in K(G)$ имеем $N(v) \cap U_1 \cap U_2 = \varnothing$ либо $\overline{N(v)} \cap U_1 \cap U_2 = \varnothing$. Аналогичный факт справедлив для $v \in \overline{U}_2$. Лемма 1 доказана.

Лемма 2. *Если* $U_1, U_2 \in K(G)$, $U_1 \cap U_2 \neq \varnothing$ *и* $\overline{U}_1 \cap \overline{U}_2 \neq \varnothing$, *то* $U_1 \cup U_2 \in K(G)$.

Доказательство. По условию $\overline{U_1 \cup U_2} = \overline{U}_1 \cap \overline{U}_2 \neq \varnothing$. Пусть $v \in \overline{U_1 \cup U_2} = \overline{U}_1 \cap \overline{U}_2$ и для определенности $N(v) \cap U_1 = \varnothing$. Тогда $N(v) \cap U_2 = \varnothing$. Действительно, если $N(v) \cap U_2 \neq \varnothing$, то с учетом $U_2 \in K(G)$ получаем $\overline{N(v)} \cap U_2 = \varnothing$, что равносильно $U_2 \subseteq N(v)$. По условию $U_1 \cap U_2 \neq \varnothing$. Поэтому для $x \in U_1 \cap U_2$ имеем $x \in U_1$, $x \in U_2 \subseteq N(v)$, т. е. $x \in U_1 \cap N(v)$, что противоречит условию $N(v) \cap U_1 = \varnothing$. Таким образом, $N(v) \cap U_2 = \varnothing$ при $N(v) \cap U_1 = \varnothing$. Следовательно, $N(v) \cap (U_1 \cup U_2) = \varnothing$. Аналогично $\overline{N(v)} \cap (U_1 \cup U_2) = \varnothing$ при $\overline{N(v)} \cap U_1 = \varnothing$. Лемма 2 доказана.

Лемма 3. *Если* $U_1, U_2 \in K(G)$, $U_1 \cap \overline{U}_2 \neq \varnothing$ *и* $\overline{U}_1 \cap U_2 \neq \varnothing$, *то* $U_1 \cap \overline{U}_2 \in K(G)$.

Доказательство. Обозначим $U = U_1 \cap \overline{U}_2$. Поскольку $\overline{U}_1 \neq \varnothing$, имеем $\overline{U} = \overline{U}_1 \cup U_2 \neq \varnothing$, а так как $U_1 \in K(G)$, для $v \in \overline{U}_1$ имеем $N(v) \cap U = N(v) \cap U_1 \cap \overline{U}_2 = \varnothing$ либо $\overline{N(v)} \cap U = \overline{N(v)} \cap U_1 \cap \overline{U}_2 = \varnothing$. Пусть теперь $v \in \overline{U} = \overline{U}_1 \cup U_2$, $v \notin \overline{U}_1$, т. е. $v \in U_1 \cap U_2$. По условию $\overline{U}_1 \cap U_2 \neq \varnothing$. В силу включения $U_1 \in K(G)$ для $x \in \overline{U}_1 \cap U_2$ выполняется $N(x) \cap U_1 = \varnothing$ либо $\overline{N(x)} \cap U_1 = \varnothing$. Пусть для определенности $N(x) \cap U_1 = \varnothing$. Тогда $N(v) \cap U = \varnothing$. Действительно, если $a \in N(v) \cap U = N(v) \cap U_1 \cap \overline{U}_2$, то $v \in N(a)$. Отсюда ввиду предположения $v \in U_2$ следует, что $N(a) \cap U_2 \neq \varnothing$. Учитывая условия $a \in \overline{U}_2$ и $U_2 \in K(G)$, получаем $\overline{N(a)} \cap U_2 = \varnothing$, т. е. $U_2 \subseteq N(a)$. Таким образом, $x \in U_2 \subseteq N(a)$. В этом случае $a \in N(x)$, и по предположению $a \in U_1$, что противоречит $N(x) \cap U_1 = \varnothing$. Следовательно, $N(v) \cap U = \varnothing$ при $N(x) \cap U_1 = \varnothing$. Аналогично, $\overline{N(v)} \cap U = \varnothing$ при $\overline{N(x)} \cap U_1 = \varnothing$. Лемма 3 доказана.

Лемма 4. *Если* $U_1, U_2 \in K(G)$ *и* $U_1 \cap U_2 = \varnothing$, *то* $(U_1 \times U_2) \cap EG = \varnothing$ *либо* $U_1 \times U_2 \subseteq EG$.

Доказательство. Пусть $(x_1, x_2) \in (U_1 \times U_2) \cap EG$. Покажем, что $(y_1, y_2) \in EG$ для любой пары $(y_1, y_2) \in (U_1 \times U_2)$. Действительно, имеем $x_1 \in U_1$, $x_2 \in U_2$, $x_1 \in N(x_2)$, $x_2 \in N(x_1)$. Таким образом, $N(x_2) \cap U_1 \neq \varnothing$. Отсюда с учетом $U_1 \in K(G)$ получаем $\overline{N(x_2)} \cap U_1 = \varnothing$, что равносильно $U_1 \subseteq N(x_2)$. В этом случае $x_2 \in N(y_1)$ для $y_1 \in U_1 \subseteq N(x_2)$. Таким образом, $N(y_1) \cap U_2 \neq \varnothing$. Поскольку $U_2 \in K(G)$, получаем $\overline{N(y_1)} \cap U_2 = \varnothing$, т. е. $U_2 \subseteq N(y_1)$. Следовательно, $y_2 \in U_2 \subseteq N(y_1)$, т. е. $(y_1, y_2) \in EG$. Лемма 4 доказана.

§ 2. Каноническая основа

Определение 2. Разбиение множества вершин графа G на канонические классы называется *каноническим разбиением*.

Пусть $W = \{V_1, \dots, V_k\}$ — каноническое разбиение графа G. По лемме 4 для любых классов $V_i, V_j \in W$ $(i \neq j)$ $(V_i \times V_j) \cap EG = \varnothing$ либо $(V_i \times V_j) \subseteq EG$. Обозначим через $\mathcal{T}_W$ граф, для которого $V\mathcal{T}_W = W$ и $(V_i, V_j) \in E\mathcal{T}_W$ тогда и только тогда, когда $(V_i \times V_j) \subseteq EG$. Будем говорить, что *граф $\mathcal{T}_W$ порождается разбиением W*.

Определение 3. Граф H называется *канонической основой* графа G, если существует каноническое разбиение W графа G такое, что H изоморфен $\mathcal{T}_W$ и неприводим.

Например, O_2 является канонической основой любого несвязного графа, а единственной основой неприводимого графа является сам граф.

Следующая теорема отвечает на вопрос о том, насколько разнообразными могут быть канонические основы данного графа.

Теорема 1. *Любые две канонические основы графа изоморфны.*

Доказательство. Обозначим через $K_m(G)$ множество всех максимальных по включению канонических классов графа G. Другими словами, $M \in K_m(G)$ тогда и только тогда, когда $M \in K(G)$ и $M \cap \overline{U} \neq \varnothing$ для любого класса $U \in K(G)$, отличного от M. Ясно, что $K_m(G)$ покрывает VG. Рассмотрим два случая.

Случай 1: $M_1 \cap M_2 = \varnothing$ для любых $M_1, M_2 \in K_m(G)$. Тогда $K_m(G)$ образует каноническое разбиение графа G, причем $\mathcal{T}_{K_m(G)}$ неприводим в силу максимальности входящих в него классов. Пусть W' — каноническое разбиение, отличное от $K_m(G)$ и $M \in K_m(G) \backslash W'$. Поскольку каждый класс $U \in W'$ содержится точно в одном из максимальных классов, минимальное подмножество $W'' \subset W'$, покрывающее M, является разбиением M. Следовательно, W'' — нетривиальный канонический класс графа $\mathcal{T}_{W'}$. Таким образом, если $K_m(G)$ является разбиением графа VG, то это единственное каноническое разбиение, порождающее неприводимый граф. В этом случае каноническая основа единственная.

Случай 2: $M_1 \cap M_2 \neq \varnothing$ для некоторых $M_1, M_2 \in K_m(G)$. Из условия максимальности M_1, M_2 имеем $M_1 \cap \overline{M}_2 \neq \varnothing$, $\overline{M}_1 \cap M_2 \neq \varnothing$. Поэтому ввиду леммы 3 $M_1 \cap \overline{M}_2 \in K(G)$. Кроме того, $\overline{M}_1 \cap \overline{M}_2 = \varnothing$, ибо в противном случае по лемме 2 получаем $M_1 \cup M_2 \in K(G)$, что противоречит максимальности M_1, M_2. Таким образом, $M_1 \cap \overline{M}_2 = \overline{M}_2 \in K(G)$ и $W = \{M_2, \overline{M}_2\}$ — каноническое разбиение графа G. Пусть для определенности $(M_2 \times \overline{M}_2) \cap EG = \varnothing$. Другими словами, $\mathcal{T}_W = O_2$.

Рассмотрим теперь произвольное каноническое разбиение W' такое, что $\mathcal{T}_{W'}$ неприводим. Если минимальное подмножество $W'' \subset$

W', покрывающее M_2, является разбиением M_2, то $W' \backslash W''$ — разбиение $\overline{M}_2$, следовательно, $|W''| = |W' \backslash W''| = 1$ в силу неприводимости $\mathcal{T}_{W'}$, т. е. $\mathcal{T}_{W'} = \mathcal{T}_W$. В противном случае имеется класс $U \in W'$ такой, что $U \cap M_2 \neq \varnothing$, $U \cap \overline{M}_2 \neq \varnothing$. Если при этом $\overline{U} \cap \overline{M}_2 \neq \varnothing$, то по лемме 2 получаем $U \cup M_2 \in K(G)$, что противоречит максимальности M_2. Следовательно, $\overline{U} \cap \overline{M}_2 = \varnothing$. Предположим, что для некоторого U' такого, что $U' \in W'$, $U' \subseteq \overline{U}$, выполняется $(U' \times U) \subseteq EG$. Тогда для вершин $x \in U' \subseteq \overline{U} \subseteq M_2$ и $y \in \overline{M}_2 \subseteq U$, с одной стороны, из $(M_1 \times \overline{M}_2) \cap EG = \varnothing$ имеем $(x, y) \notin EG$, а с другой стороны, из $(U' \times U) \subseteq EG$ имеем $(x, y) \in EG$. Таким образом, учитывая лемму 4, получаем $(U' \times U) \cap EG = \varnothing$ для каждого U' такого, что $U' \in W'$, $U' \subseteq \overline{U}$. Следовательно, $(\overline{U} \times U) \cap EG = \varnothing$, откуда вытекает $\overline{U} \in K(G)$. Поскольку $\mathcal{T}_{W'}$ неприводим, $\overline{U}$ содержит точно один класс из W', т. е. $|W'| = 2$, причем из условия $(U' \times U) \cap EG = \varnothing$ следует, что $\mathcal{T}_{W'} = O_2$. Теорема 1 доказана.

§ 3. Алгоритм поиска канонической основы

Из доказательства теоремы 1 вытекает, что для несвязных графов поиск канонического разбиения, порождающего основу, сводится к поиску произвольной компоненты связности. Выявление компонент связности легко осуществляется за полиномиальное время методом поиска в глубину или ширину [3]. Если же сам граф и его дополнение связны, то единственным разбиением, порождающим каноническую основу, является разбиение на максимальные канонические классы. Ниже приводится алгоритм, который для заданной вершины v графа G находит максимальный канонический класс, содержащий v.

АЛГОРИТМ ПОИСКА МАКСИМАЛЬНОГО КАНОНИЧЕСКОГО КЛАССА

Шаг 1. Полагаем $M = \{v\}$.

Шаг 2. Цикл по всем вершинам $x \in \overline{M}$. Если цикл закончился, то полагаем $M \in K_m(G)$. Конец.

Шаг 3. Полагаем $A = M \cup \{x\}$.

Шаг 4. Цикл по всем вершинам $y \in \overline{A}$. Окончание цикла означает, что $A \in K(G)$, т. е. «найдено каноническое расширение M» и $M = A$. Если $|\overline{M}| > 1$, переходим к шагу 2 для продолжения цикла, в противном случае полагаем $M \in K_m(G)$. Конец.

Шаг 5. Если $N(y) \cap A = \varnothing$ или $\overline{N(y)} \cap A = \varnothing$, то переходим к шагу 4 для продолжения цикла.

Шаг 6. Полагаем $A = A \cup \{y\}$. Если $\overline{A} \neq \varnothing$, то переходим к шагу 4 для продолжения цикла. Если $\overline{A} = \varnothing$, т. е. «не существует канонического расширения M, содержащего x», то переходим к шагу 2 для продолжения цикла.

Оценим трудоемкость данного алгоритма. Очевидно, циклы шагов 2 и 4 выполняются не более n раз, а все проверки внутри циклов осуществляются за время, линейным образом зависящее от числа вершин графа. Таким образом, поиск одного максимального канонического класса требует $O(n^3)$ операций. Поскольку в худшем случае $K_m(G)$ содержит n классов, поиск канонического разбиения, порождающего основу, осуществляется за время $O(n^4)$, где $n = |VG|$ — число вершин графа G.

§ 4. Канонические разбиения и кодирование графов

Определение 4. *Каноническим числом* $C(G)$ графа G называется число вершин его канонической основы.

Поскольку граф K_1 не имеет канонических разбиений, $C(G) \geqslant 2$ для любого графа G, у которого имеется не менее чем две вершины. Для K_1 положим $C(K_1) = 0$.

Определение 5. Граф G называется

— *l-каноническим*, если $C(G) \leqslant l$,

— *(l,m)-почти каноническим*, если G содержит l-канонический порожденный подграф не менее чем с $|VG| - m$ вершинами.

Обозначим через $F^{l,m}$ класс всех графов G таких, что каждый порожденный подграф графа G является (l,m)-почти каноническим.

Согласно определению $F^{l,m} \subseteq F^{l+1,m}$, $F^{l,m} \subseteq F^{l,m+1}$ для любых $l \geqslant 2$, $m \geqslant 0$. Кроме того, из определения вытекает наследственная замкнутость рассматриваемых классов (*наследственным* называется класс графов X, который вместе с каждым графом $G \in X$ содержит все порожденные подграфы).

Рассмотрим некоторые примеры.

Пример 1. Минимальный класс $F^{2,0}$ образует известный класс кографов [4]: это наследственный класс, замкнутый относительно операций объединения и соединения графов. Каждый граф в этом классе несвязный либо дополнительный к несвязному графу. В [5] показано, что такие графы имеют простое описание в терминах запрещенных подграфов: кографы не содержат простой цепи с четырьмя вершинами в качестве порожденного подграфа.

Пример 2. Легко видеть, что всякий ациклический граф является (2,1)-почти каноническим. Таким образом, класс лесов содержится в $F^{2,1}$. Аналогично класс планарных графов содержится в $F^{2,5}$. Последнее следует, в частности, из того, что любой планарный граф содержит вершину степени не более 5 и, кроме того, класс планарных графов является наследственным.

Для произвольного класса графов X через X_n обозначим множество n-вершинных графов из X, через B^* — множество всех слов в алфавите $B = \{0,1\}$, через α_j и $|\alpha|$ — j-й символ и длину слова α соответственно.

Определение 6. *Кодированием графов* из класса X называется семейство взаимно однозначных отображений $\Phi = \{\varphi_n \mid n = 1,2,\dots\}$, где $\varphi_n\colon X_n \to B^*$.

Очевидно, что для любого кодирования имеет место неравенство

$$\log|X_n| \leqslant \max_{G\in X_n} |\varphi_n(G)| \tag{1}$$

(здесь и далее логарифм берется по основанию 2).

Определение 7. Кодирование называется

— *асимптотически оптимальным*, если

$$\lim_{n\to\infty} \frac{\max_{G\in X_n}|\varphi_n(G)|}{\log|X_n|} = 1;$$

— *оптимальным по порядку*, если существует константа $M > 0$ такая, что

$$\frac{\max_{G\in X_n}|\varphi_n(G)|}{\log|X_n|} < M.$$

Ниже условие оптимальности по порядку мы будем записывать так:

$$\max_{G\in X_n} |\varphi_n(G)| = O(\log|X_n|).$$

В [5] показано, что для любого бесконечного наследственного класса графов x существует энтропия $h(X) = \lim_{n\to\infty}(\log|X_n|)/\binom{n}{2}$, и предложен универсальный метод кодирования графов, асимптотически оптимальный для наследственных классов с ненулевой энтропией. Однако, асимптотическая оптимальность данного метода не распространяется на классы графов с нулевой энтропией. Ниже доказано равенство $h(F^{l,m}) = 0$ для любых $l \geqslant 2$, $m \geqslant 0$ и описано семейство универсальных кодирований, содержащее оптимальное по порядку кодирование для любого из рассматриваемых классов $F^{l,m}$ $(l \geqslant 2, m \geqslant 0)$. Для описания этого семейства примем следующие соглашения. Все графы считаются помеченными, т. е. вершинам произвольного n-вершинного графа G приписаны натуральные числа от единицы до n. Упорядоченность множества вершин VG индуцирует упорядоченность произвольного подмножества $W \subseteq VG$ и тем самым устанавливает лексикографический порядок на множестве всех подмножеств VG, а следовательно, и на множестве порожденных подграфов графа G.

Пусть $s \geqslant 0$ — целое число. Опишем универсальное кодирование Φ^s, зависящее от параметра s. Для этого произвольному графу G сопоставим корневое упорядоченное дерево $T_s(G)$, каждой вершине которого приписана некоторая метка и поставлен в соответствие некоторый порожденный подграф графа G. Опишем $T_s(G)$ индуктивно.

Корню $T_s(G)$ соответствует граф G.

Пусть теперь вершине v дерева $T_s(G)$ соответствует подграф G_v графа G. Если $|VG_v| \leqslant s+1$, то v — висячая вершина в $T_s(G)$ и ей приписывается список номеров вершин графа G, порождающих G_v, и матрица смежности G_v. Если $|VG_v| > s+1$, то среди порожденных подграфов графа G_v с $|VG_v| - s$ вершинами находим подграф G'_v с наименьшим каноническим числом. Если таких подграфов несколько, то в качестве G'_v выбираем лексикографически наименьший из них. Пусть $t_v = C(G'_v)$, H_v — каноническая основа G'_v, $W_v = \{V_1, \ldots, V_{t_v}\}$ — лексикографически упорядоченное каноническое разбиение, порождающее основу H_v, $V_v = VG_v \backslash VG'_v$. Тогда вершина v имеет t_v сыновей. Вершине v приписываем матрицу смежности графа H_v, а ее сыновьям ставим в соответствие подграфы $G\langle V_1 \cup V_v\rangle, \ldots, G\langle V_{t_v} \cup V_v\rangle$. Нетрудно видеть, что граф G полностью определяется деревом $T_s(G)$. Закодируем дерево $T_s(G)$ с помощью двух слов: μ^0 — код дерева без меток и μ^1 — код меток. Для построения μ^0 используем метод кодирования корневых упорядоченных деревьев путем обхода [3]. Обход всегда начинается с корня и представляет собой вариант поиска в глубину, при котором направление движения из любой вершины определяется непройденным ребром с меньшим номером. Таким образом, каждое ребро дерева участвует в обходе дважды, и в μ^0 ему соответствуют два символа: 1 — при движении от корня к висячей вершине, 0 — при движении в обратном направлении. Поэтому длина слова μ^0 четна и

$$\sum_{i=1}^{k} \mu_i^0 \geqslant k/2 \tag{2}$$

для любого k, $1 \leqslant k \leqslant |\mu^0|$, причем при $k = |\mu^0|$ сумма в левой части (2) равна $k/2$. Слово μ^1 состоит из двоичных записей меток всех вершин дерева $T_s(G)$, упорядоченных в соответствии с последовательностью его обхода. Полагаем $\varphi_n^s(G) = \mu^0 0 \mu^1$, $\Phi^s = \{\varphi_n^s \mid n = 1, 2, \ldots\}$.

Теорема 2. *Кодирование Φ^s является оптимальным по порядку для любого класса $F^{l,m}$ при $l \geqslant 2$, $0 \leqslant m \leqslant s$.*

Доказательство. Нетрудно убедиться, что $\varphi_n^s(G)$ является кодом G. Действительно, ввиду (2) позиция нуля, разделяющего μ^0 и μ^1 в $\varphi_n^s(G)$, определяется наименьшим нечетным k таким, что $\sum\limits_{i=1}^{k} \varphi_n^s(G)_i <$

$k/2$. Восстановив дерево по коду μ^0 и зная значения n и s, нетрудно определить длину метки каждой вершины и тем самым ее положение в μ^1: метки вершин степени $k > 1$ имеют длину $(k-1)^2$, метки вершин степени 1 — длину $(s+1)\lceil\log n\rceil + (s+1)^2$. При $n \leqslant s+1$ дерево $T_s(G)$ состоит из одной вершины, имеющей метку длины $n\lceil\log n\rceil + n^2$. Таким образом, код $\varphi_n^s(G)$ полностью определяет дерево $T_s(G)$, а следовательно, и граф G.

Оценим сверху число вершин в дереве $T_s(G)$. Пусть M — множество висячих вершин $T_s(G)$, а $\overline{M} = VT_s(G)\backslash M$ — множество внутренних вершин. Очевидно, что $\sum\limits_{v\in\overline{M}} t_v = |ET_s(G)|$. Поскольку в любом дереве число ребер на единицу меньше числа вершин, справедливо равенство

$$\sum_{v\in\overline{M}}(t_v-1) = |ET_s(G)| - |\overline{M}| = |M| - 1.$$

Кроме того, из условия теоремы 2 имеем $t_v \geqslant 2$ для любой $v \in \overline{M}$, следовательно,

$$|\overline{M}| \leqslant \sum_{v\in\overline{M}}(t_v-1) = |M| - 1.$$

Пусть v — внутренняя вершина, $v_1, \ldots, v_{t_v}$ — ее сыновья. Тогда по построению $T_s(G)$

$$|VG_v| = \sum_{i=1}^{t_v}|VG_{v_i}| - (t_v-1)|V_v|.$$

Отсюда индукцией по числу внутренних вершин нетрудно установить соотношение

$$\sum_{v\in M}|VG_v| - \sum_{v\in\overline{M}}(t_v-1)|V_v| = |VG|. \tag{3}$$

Заметим, что по построению $|V_v| = s$ для всех $v \in \overline{M}$. Нетрудно видеть, что $|VG_v| = s+1$ для всех $v \in M$. Таким образом, соотношение (3) переписывается в виде $(s+1)|M| - s(|M|-1) = |VG|$, откуда $|M| = |VG| - s$. Учитывая, что $|\overline{M}| < |M|$, окончательно получаем

$$|T_s(G)| = |M| + |\overline{M}| < 2|M| = 2(|VG| - s).$$

Пусть теперь $G \in F_n^{l,m}$, где $0 \leqslant m \leqslant s$, $l \geqslant 2$. Тогда каждая внутренняя вершина в $T_s(G)$ помечена словом длины не более l^2, а каждая висячая вершина — словом длины не более $(s+1)\lceil\log n\rceil + (s+1)^2$. Следовательно, $|\mu^1| \leqslant (n-s)(l^2 + (s+1)\lceil\log n\rceil + (s+1)^2)$.

Для μ^0 очевидно неравенство $|\mu^0| \leqslant 4(n-s)$. Поэтому

$$\max_{G \in F_n^{l,m}} |\varphi_n^s(G)| = O(n \log n). \tag{4}$$

С другой стороны класс $F^{2,0}$ (минимальный из рассматриваемых) содержит класс графов $\mathcal{L}$ таких, что каждая компонента связности каждого графа состоит не более чем из двух вершин. Таким образом, число n-вершинных графов из $\mathcal{L}$ есть число разбиений n-элементного множества на подмножества, состоящие не более чем из двух элементов. Это число равно

$$|\mathcal{L}_n| = \sum_{i=0}^{\lfloor n/2 \rfloor} \binom{n}{2i} \frac{(2i)!}{i!\, 2^i}.$$

Отсюда видно, что $n \log n = O(\log |\mathcal{L}_n|)$. Следовательно, для любых $l \geqslant 2$, $m \geqslant 0$

$$n \log n = O(\log |F_n^{l,m}|). \tag{5}$$

Из (4) и (5) с учетом (1) получаем, что Φ^s — оптимальное по порядку кодирование для любого $F^{l,m}$ при $l \geqslant 2$, $0 \leqslant m \leqslant s$. Теорема 2 доказана.

Непосредственно из (1) и (4) вытекает

Следствие. *Для любых* $l \geqslant 2$, $m \geqslant 0$ *имеем* $h(F^{l,m}) = 0$.

Оценивая трудоемкость кодирования Φ_s в общем случае, заметим, что наибольшие затраты связаны с построением дерева $T_s(G)$. Очевидно, всякий n-вершинный граф G имеет не более n^s порожденных подграфов с $n-s$ вершинами. Следовательно, с учетом алгоритма поиска канонической основы, каждая внутренняя вершина при построении $T_s(G)$ требует $O(n^{s+4})$ операций. Поскольку $|T_s(G)| < 2n$, при построении $T_s(G)$ используется не более $O(n^{s+5})$ операций. Таким образом, кодирование Φ^s для произвольного n-вершинного графа G может быть получено за время, ограниченное полиномом степени не выше $s+5$.

Автор благодарен В. Е. Алексееву и рецензенту за внимательное отношение к работе и указанные замечания.

ЛИТЕРАТУРА

1. **Erdős P., Hajnal A.** On the number of distinct induced subgraphs of a graph // Discrete Math. 1989. V. 75, N 1–3. P. 145–154.
2. **Гэри М., Джонсон Д.** Вычислительные машины и труднорешаемые задачи. М.: Мир, 1982.

3. **Евстигнеев В. А., Касьянов В. Н.** Алгоритмы на деревьях. Новосибирск: ВЦ СО АН СССР, 1989.
4. **Corneil D. G., Lerchs H., Burlingham L. S.** Complement reducible graphs // Discrete Appl. Math. 1981. V. 3, N 3. P. 163–174.
5. **Алексеев В. Е.** Наследственные классы и кодирование графов // Проблемы кибернетики. М.: Наука, 1982. Вып. 39. С. 151–164.

Адрес автора:

РОССИЯ,
603600, Нижний Новгород, ГСП-20,
пр. Гагарина, 23, корп. 2
Нижегородский гос. ун-т

Статья поступила

19 апреля 1994 г.