ISSN 1683-4720 Труды ИПММ НАН Украины. 2009. Том 18

УДК 519.6

©2009. **И.С. Грунский, Е.А. Пряничникова**

ОБ АЛГЕБРЕ ЯЗЫКОВ, ПРЕДСТАВИМЫХ ГРАФАМИ С ОТМЕЧЕННЫМИ ВЕРШИНАМИ

Вводится алгебра языков, представимых конечными ориентированными графами с отмеченными вершинами, и рассматриваются свойства этой алгебры.

Введение. В настоящее время существует ряд актуальных задач, связанных с анализом графов с отмеченными вершинами [1, 2]. В большинстве работ, посвященных решению таких задач, для алгебраической интерпретации множества путей в графах используется алгебра Клини [4]. В данной работе вводится новая алгебра, которая, на наш взгляд, более адекватно описывает свойства языков, представимых в графах с отмеченными вершинами, и исследуются свойства этой алгебры.

Работа имеет следующую структуру. В разделе 1 вводятся основные определения. В разделе 2 формулируется теорема, аналогичная теореме Клини для алгебры $\langle 2^{X^*}, \cdot, \cup, *, \varnothing, \lambda \rangle$ и приводится доказательство ее первой части. В разделе 3 вводятся матричные методы представления графов с отмеченными вершинами, с помощью которых доказывается вторая часть теоремы.

1. Основные определения. Пусть X – конечный алфавит, X^+ – множество всех непустых слов конечной длины в алфавите X. Введем на языках L, $R \subseteq X^+$ следующие операции:

1) $L \cup R = \{w | w \in L$ или $w \in R\}$;

2) $L \circ R = \{w_1 \circ w_2 | w_1 \in L$ и $w_2 \in R\}$;

Операция $\circ$ на множестве X^+ определяется следующим образом: для всех $w_1, w_2 \in X^+$ и всех $x, y \in X$ $w_1x \circ yw_2 = w_1xw_2$, если $x = y$; $w_1x \circ yw_2$ – не определено в противном случае.

3)$L^+ = \bigcup_{i=1}^{\infty} L^i$, где $L^1 = L$; $L^{n+1} = L^n \circ L$ для всех $n \geq 1$;

4) $L^* = \bigcup_{i=0}^{\infty} L^i$, где $L^0 = X$; $L^{n+1} = L^n \circ L$ для всех $n \geq 0$;

5)$L^{\circledast} = L_{beg} \circ L^* \circ L_{end}$, где $L_{beg} = \{x | xw \in L, x \in X, w \in X^*\}$; $L_{end} = \{x | wx \in L, x \in X, w \in X^*\}$.

Рассмотрим алгебраическую систему $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$. Множество 2^{X^+} является полукольцом относительно операций $\circ$ и $\cup$, то есть $\langle 2^{X^+}, \cup, \varnothing \rangle$ является идемпотентным коммутативным моноидом, $\langle 2^{X^+}, \circ, X \rangle$ является моноидом, причем $\varnothing \circ R = R \circ \varnothing = \varnothing$ для всех $R \in 2^{X^+}$, и для операций $\circ$ и $\cup$ выполняются дистрибутивные законы: $R \circ (Q \cup P) = R \circ Q \cup R \circ P$, $(R \cup Q) \circ P = R \circ P \cup Q \circ P$.

Основные особенности рассматриваемой алгебраической системы связаны с тем, что, в отличие от конкатенации в алгебре Клини, операция $\circ$ частичная. Если в алгебре Клини $R \cdot Q = \varnothing$ только в том случае, когда $R = \varnothing$ или $Q = \varnothing$, то в алгебре $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$ легко подобрать примеры, когда это не верно: пусть $R = \{ab\}$, $Q = \{cd\}$, тогда $R \circ Q = \varnothing$. Еще одно отличие заключается в том, что в алгебре Клини множество R^* всегда содержит пустое слово и является бесконечным, а в алгебре $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X\rangle$ можно привести примеры множеств, для которых результат применения операции $\circledast$ будет бесконечным, конечным или пустым множеством. Операция $\circledast$ обладает следующими основными свойствами:

$$R^{\circledast} = R_{beg} \circ R_{end} \cup R^+ = (R_{beg} \circ R_{end} \cup R)^+ \tag{1}$$

$$R^{\circledast} = R_{beg} \circ (X \cup R^+) \circ R_{end} = R_{beg} \circ (X \cup R)^+ \circ R_{end} \tag{2}$$

$$R^+ = R^{\circledast} \circ R \cup R \tag{3}$$

$$R^{\circledast} = R^{\circledast} \circ R \cup R \cup R_{beg} \circ R_{end} \tag{4}$$

$$R^{\circledast} \circ R^{\circledast} \cup R = R^{\circledast} \tag{5}$$

$$R^{\circledast} \circ R^{\circledast} = R_{beg} \circ R_{end} \cup R^{\circledast} \circ R. \tag{6}$$

Доказательства этих соотношений непосредственно следуют из определения операций.

Множество 2^{X^+} упорядочено отношением включения, и $R \leq Q$ тогда, и только тогда, когда $Q \cup R = R$. Если $R \leq Q$, то $R \circ P \leq Q \circ P$, $R \cup P \leq Q \cup P$, $R^{\circledast} \leq Q^{\circledast}$, то есть отношение $\leq$ является естественным частичным порядком и монотонно по отношению к операциям алгебры $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X\rangle$.

По определению операции $\circledast$ для любого $R \in 2^{X^+}$ выполняется $R_{beg} \circ R_{end} \leq X$, следовательно, для любого $Q \in 2^{X^+}$

$$R_{beg} \circ R_{end} \circ Q \leq Q. \tag{7}$$

Регулярными выражениями в алгебре $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X\rangle$ будем называть формулы:
1)$\varnothing$ является регулярным выражением и представляет пустое множество.
2) Для каждого $x \in X$ символ x является регулярным выражением и представляет язык $\{x\}$.
3) Для каждого $x \in X$ и $y \in X$ xy является регулярным выражением, представляющим язык $\{xy\}$.
4) Если p и q – регулярные выражения, то выражения $(p \circ q)$,$(p \cup q)$,$(p^{\circledast})$ также являются регулярными.

Язык, который представляет регулярное выражение R, будем обозначать $L(R)$. Два регулярных выражения будем называть эквивалентными, если совпадают языки, которые они обозначают.

Назовем графом с отмеченными вершинами четверку $G = (Q, E, X, \mu)$, где Q – конечное множество вершин, $\mid Q \mid= n$, $E \subseteq Q \times Q$ – множество дуг, X – множество отметок вершин, $\mu : Q \to X$ – функция отметок вершин. Пусть $I \subseteq Q$ – множество начальных вершин графа, $F \subseteq Q$ – множество финальных вершин. Вершины q_1 и q_2 будем называть смежными, если дуга $(q_1, q_2) \in E$. Путем в графе G будем называть конечную последовательность смежных вершин $l = q1q2...qk$, где $(q_i, q_{i+1}) \in E$, число k будем называть длиной пути, q_1 – начальная вершина пути, q_k – конечная вершина. Элемент $x = \mu(q_1)\mu(q_2)...\mu(q_k) = x_1x_2..x_k \in X^+$ будем называть отметкой пути. Для каждой вершины q определим путь нулевой длины, начинающийся и заканчивающийся в этой вершине. Отметкой нулевого пути будет $x \in X$, такое, что $\mu(q) = x$. Множество отметок всех путей, начальной вершиной которых является вершина q_i, а конечная вершина $q_k \in F$, назовем языком, порожденным вершиной q_i. Отметки всех путей в графе G, начальные вершины которых принадлежат множеству I, а конечные – множеству F, назовем языком, порождаемым графом G, и обозначим $L(G)$. Для так называемых детерминированных графов с одним начальным состоянием свойства языков изучались в работе [3].

2. Решение уравнений в алгебре $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$. В теории конечных автоматов одним из важнейших результатов является теорема Клини, в которой утверждается, что класс языков, распознаваемых конечными автоматами, совпадает с классом рациональных языков, представимых регулярными выражениями алгебры Клини $\langle 2^{X^*}, \cdot, \cup, *, \varnothing, \lambda \rangle$ [1].

Основная цель данной работы – доказать аналогичную теорему для языков, порождаемых графами с отмеченными вершинами.

Теорема 1. *Для любого языка $L \subseteq X^+$ следующие утверждения эквивалентны:*
1) L является языком, представимым регулярным выражением алгебры
$\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$;
2) L является языком, порождаемым графом с отмеченными вершинами.

Доказательство. Покажем, что для любого графа G, порождающего язык L(G), можно найти такое регулярное выражение R алгебры $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$, что $L(R) = L(G)$.

Обозначим язык, порожденный вершиной q_i графа G, через L_i.

Для каждой дуги $(q_i, q_j) \in E$ в язык L_i входят все слова языка L_j с приписанным слева символом $\mu(q_i)$. Если вершина q_i в графе G является финальной, то символ $\mu(q_i)$ входит в язык L_i как отдельное слово. Таким образом, язык L_i можно описать уравнением:

$$L_i = \mu(q_i)\mu(q_1) \circ L_1 \cup \mu(q_i)\mu(q_2) \circ L_2 \cup ... \cup \mu(q_i)\mu(q_n) \circ L_n \cup \alpha \circ \mu(q_i), \tag{8}$$

где n – количество вершин в графе; если вершина q_i является финальной, то $\alpha = X$, иначе $\alpha = \varnothing$.

Язык, порождаемый графом G, можно определить как $L(G) = \bigcup_{q_i \in I} L_i$, следовательно, для определения L(G) нужно решить систему уравнений вида

$$\begin{aligned} x_1 &= r_{11} \circ x_1 \cup r_{12} \circ x_2 \cup ... \cup r_{1n} \circ x_n \cup p_1 \\ x_2 &= r_{21} \circ x_1 \cup r_{22} \circ x_2 \cup ... \cup r_{2n} \circ x_n \cup p_2 \\ &\dots\dots\dots\dots \\ x_n &= r_{n1} \circ x_1 \cup r_{n2} \circ x_2 \cup ... \cup r_{nn} \circ x_n \cup p_n, \end{aligned} \tag{9}$$

где коэффициенты $r_{ij}, p_i \in 2^{X^+}$, а все x_i – неизвестные языки. Так как все уравнения системы (9) имеют вид $Y = R \circ Y \cup Q$, для нахождения решения можно воспользоваться следующим результатом.

Лемма 1. *В алгебре* $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X\rangle$ *уравнение с одним неизвестным вида* $Y = R \circ Y \cup Q$ *имеет наименьшее решение* $Y = R^{\circledast} \circ Q \cup Q$, *которое является единственным, если* $X \cap R = \varnothing$.

Доказательство леммы. Для доказательства можно использовать известную теорему о неподвижной точке, согласно которой в любом замкнутом полукольце всякое уравнение вида $x = f(x)$, где f – непрерывное отображение носителя этого полукольца в себя, имеет наименьшее решение [1].

Идемпотентное полукольцо $\langle 2^{X^+}, \circ, \cup, \varnothing, X\rangle$ замкнутое, так как любая упорядоченная последовательность языков L_1, L_2,...L_n,... имеет наименьшую верхнюю грань, равную объединению всех языков в последовательности, а операция $\circ$ сохраняет наименьшие верхние грани последовательностей.

В силу непрерывности операций $\circ$ и $\cup$, а также того, что суперпозиция непрерывных функций непрерывна, правая часть уравнения $Y = R\circ Y\cup Q$ является непрерывным отображением $f : 2^{X^+} \to 2^{X^+}$. По теореме о неподвижной точке наименьшим решением уравнения $Y = R \circ Y \cup Q$ будет $Y = \bigcup_{n\geq 0}^{\infty} f^n(\varnothing)$, где $\varnothing$ – наименьший элемент множества 2^{X^+}, $f^0(x) = x$, $f^n(x) = f(f^{n-1}(x))$ для любого $n > 0$.

Учитывая, что $f^0(\varnothing) = \varnothing$, $f^1(\varnothing) = Q$, $f^2(\varnothing) = R \circ Q \cup Q$, $f^3(\varnothing) = R \circ R \circ Q \cup R \circ Q \cup Q$, ... $f^n(\varnothing) = (R^{n-1} \cup R^{n-2} \cup ... \cup R \circ R \cup R) \circ Q \cup Q$, получим, что

$$Y = \bigcup_{n\geq 0}^{\infty} f^n(\varnothing) = (\bigcup_{n\geq 0}^{\infty} (R^{n-1} \cup R^{n-2} \cup ... \cup R \circ R \cup R)) \circ Q \cup Q = R^+ \circ Q \cup Q = R^* \circ Q.$$

Воспользовавшись свойствами (7) и (1), получим

$$Y = R^+ \circ Q \cup R_{beg} \circ R_{end} \circ Q \cup Q = (R^+ \cup R_{beg} \circ R_{end}) \circ Q \cup Q = R^{\circledast} \circ Q \cup Q.$$

Покажем, что это решение единственно, когда $X \cap R = \varnothing$. Пусть Y' – любое решение уравнения $Y = R \circ Y \cup Q$, слово $p \in Y'$, длина p равна m. Подставляя многократно правую часть уравнения $Y = R \circ Y \cup Q$ вместо Y в себя, получим

$$\begin{gathered} Y = R \circ Y \cup Q = R \circ (R \circ Y \cup Q) \cup Q = R \circ R \circ Y \cup R \circ Q \cup Q = \\ R \circ R \circ (R \circ Y \cup Q) \cup R \circ Q \cup Q = R^m \circ Y \cup R^{m-1} \circ Q \cup ... \cup R \circ Q \cup Q. \end{gathered} \tag{10}$$

Если $X \cap R = \varnothing$, то в случае, когда $R^m \neq \varnothing$, все слова, принадлежащие языку R^m имеют длину не меньше $m+1$, откуда следует, что $p \in R^{m-1} \circ Q \cup ... \cup R \circ Q \cup Q$. Тогда $p \in R^{\circledast} \circ Q \cup Q$, и решение $Y = R^{\circledast} \circ Q \cup Q$ действительно является единственным решением уравнения $Y = R \circ Y \cup Q$, что и требовалось доказать. $\square$

Исходя из того, каким образом составлена система (9), можно отметить, что во всех уравнениях для всех коэффициентов выполняется $r_{ij_{beg}} = r_{ij_{end}} = p_{i beg}$ и $r_{ij} \cap X = \varnothing$. Из определения операции $\circledast$ следует, что $R^{\circledast} \circ Q = R_{beg} \circ R^* \circ R_{end} \circ Q$, поэтому, если $R_{beg} = R_{end} = Q_{beg}$, то $R^{\circledast} \circ Q = R_{beg} \circ R_{end} \circ Q \cup R \circ Q \cup R \circ R \circ Q.... = R^{\circledast} \circ Q \cup Q$.

Таким образом, если уравнение вида $Y = R \circ Y \cup Q$ описывает язык, порожденный вершиной Y в графе G, его решение может быть записано в виде

$$Y = R^{\circledast} \circ Q \tag{11}$$

и это решение является единственным.

Для решения системы (9) по аналогии с решением систем уравнений с регулярными коэффициентами в алгебре Клини $\langle 2^{X^*}, \cdot, \cup, *, \varnothing, \lambda \rangle$ можно использовать метод последовательного исключения неизвестных, подобный классическому методу Гаусса [1].

Используя (11) можно удалить L_n из правой части последнего уравнения в системе. Полученное значение для L_n можно подставить в предпоследнее уравнение, и, воспользовавшись тем же приемом, получить выражение для L_{n-1}, которое зависит только от $L_1 .... L_{n-2}$. Продолжая такие подстановки для всех оставшихся в системе уравнений, можно получить значения всех L_i, которые, в силу леммы 1, будут регулярными выражениями в алгебре $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$.

Таким образом, в результате решения системы (9) для любого графа G можно получить регулярное выражение R, для которого $L(R) = L(G)$, что и требовалось доказать.

3. Матричное представление графов с отмеченными вершинами. Докажем вторую часть теоремы, т.е. что для любого регулярного выражения R в алгебре $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$ существует граф G, порождающий язык $L(G) = L(R)$.

Для доказательства воспользуемся матричным представлением графов с отмеченными вершинами, по аналогии с тем, как задаются с помощью матриц конечные автоматы [5] .

Каждому графу G поставим в соответствие матрицу переходов M – квадратную матрицу порядка n, где n – количество вершин графа, элементы которой определяются следующим образом:

$$M_{ij} = \begin{cases} \mu(q_i)\mu(q_j), & \text{если } (q_i, q_j) \in E; \\ \varnothing, & \text{иначе.} \end{cases}$$

Для задания отметок вершин графа используем вектор Q длины n, такой, что $Q_i = \mu(q_i)$. Для задания начальных вершин графа используем вектор U длины n, такой, что

$$U_i = \begin{cases} \mu(q_i), & \text{если } q_i \in I; \\ \varnothing, & \text{иначе.} \end{cases}$$

Финальные состояния зададим с помощью вектора V длины n, для которого

$$V_i = \begin{cases} \mu(q_i), & \text{если } q_i \in F; \\ \varnothing, & \text{иначе.} \end{cases}$$

Пусть L – вектор длины n, элементы которого L_i – языки, порожденные вершинами q_i графа G. Рассмотрим семейство M_n квадратных матриц порядка n над алгеброй

$\langle 2^{X^+}$,$\circ$,$\cup$,$\circledast$,$\varnothing$,$X\rangle$. Определим на таких матрицах операции $\circ$ и $\cup$:
1) Если $A \cup B = C$,где $A, B, C \in M_n$, то $C_{ij} = A_{ij} \cup B_{ij}$ для всех $i, j \in \overline{1, n}$;
2) Если $A \circ B = C$, то $C_{ij} = \bigcup_{k=1}^{n} A_{ik} \circ B_{kj}$.

Таким образом, операции $\circ$ и $\cup$ на матрицах из M_n определяются аналогично обычным операциям сложения и умножения матриц, а операции сложения и умножения элементов этих матриц понимаются как операции алгебры $\langle 2^{X^+}$,$\circ$,$\cup$,$\circledast$,$\varnothing$,$X\rangle$.

Обозначим через Z_n квадратную матрицу порядка n, все элементы которой равны $\varnothing$; через E_n – квадратную матрицу порядка n, в которой все элементы, не лежащие на главной диагонали, равны $\varnothing$,а все диагональные элементы $E_{n_{ij}} = \bigcup_{a_i \in X} a_i$.

Определим для матриц из M_n операцию $+$, как $M^+ = M \cup M \circ M \cup M \circ M \circ M \ldots$ Если M – это матрица переходов графа G, а ее элемент M_{ij} – отметка пути длины 1 из вершины q_i в вершину q_j, то $M \circ M_{ij} = \bigcup_{k=1}^{n} M_{ik} \circ M_{kj}$ – сумма всех путей длины 2, ведущих из q_i в q_j; $M \circ M \circ M_{ij}$ – всех путей длины 3, и т.д. Таким образом, M^+ – это матрица, элементы которой $M^+{}_{ij}$ являются регулярными выражениями алгебры $\langle 2^{X^+}$,$\circ$,$\cup$,$\circledast$,$\varnothing$,$X\rangle$, описывающими все возможные пути в графе G из вершины q_i в вершину q_j, длина которых больше нуля.

Пусть $M^* = E_n \cup M^+$, $M^{\circledast} = M' \cup M^+$, где $M' = Q \circ Q^T$. Если M – это матрица переходов графа G, то M' – это матрица, описывающая все пути нулевой длины в графе G:

$$M' = \begin{cases} \mu(q_i), & \text{если } i = j; \\ \varnothing, & i \neq j, \end{cases}$$

а $M_{ij}^{\circledast}$ задает все пути в графе G из вершины q_i в вершину q_j, длина которых больше или равна нулю. Из определения операций следует, что $M^{\circledast} = (M' \cup M)^+$. Введем на множестве M_n отношение частичного порядка. Будем считать, что $A \leq B$ означает, что $A_{ij} \cup B_{ij} = B_{ij}$ для всех $i, j \in \overline{1, n}$.

Из определения M' следует, что $M' \circ P \leq P$ для любой матрицы M порядка n и вектора P длины n .

Используя матричное представление графов с отмеченными вершинами и свойства операций над матрицами, систему уравнений (9) можно записать в виде уравнения

$$L = M \circ L \cup V, \tag{12}$$

а язык, порожденный графом G, как $L(G) = U^T \circ L$.

Легко показать, что алгебра $\langle M_n$,$\circ$,$\cup$,$\circledast$,$Z_n, En\rangle$ является замкнутым идемпотентным полукольцом, поэтому по теореме о неподвижной точке уравнения вида $Y = R \circ Y \cup Q$, где Y – неизвестная матрица $R, Q \in M_n$, имеют наименьшее решение $Y = R^{\circledast} \circ Q \cup Q$.

В матрице M все элементы каждого столбца j – это пути в вершину q_j, поэтому $M_{ij_{end}} = M_{kj_{end}} = \mu(q_j)$, $V_j = \varnothing$ или $V_j = \mu(q_j)$, причем для всех диагональных элементов матрицы M выполняется $M_{iibeg} = M_{iiend} = V_i$, если $V_i \neq \varnothing$, откуда следует, что $M' \circ V = V$. Тогда решение уравнения (12) можно записать в виде

$$L = M^+ \circ V \cup V = M^+ \circ V \cup M' \circ V = (M^+ \cup M') \circ V = M^{\circledast} \circ V = (M \cup M')^{\circledast} \circ V, \tag{13}$$

а язык, порожденный любым графом G

$$L(G) = U^T \circ M^{\circledast} \circ V = U^T \circ (M \cup M')^{\circledast} \circ V. \tag{14}$$

Для вычисления значений матрицы $M^{\circledast}$ можно воспользоваться методом исключения неизвестных, который был применен для нахождения решений системы (9), но для использования в дальнейших вычислениях удобнее использовать рекурсивный способ, аналогичный применяемому в [5], для матриц переходов конечных автоматов.

Для произвольного графа G обозначим через R матрицу $M \cup M'$. Из (14) следует, что $L = R^{\circledast} \circ V$. Пусть n – количество вершин графа G. Выберем произвольное $n_1 < n$ и разобьем матрицу $R = \begin{pmatrix} R_{11} & \dots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n1} & \dots & R_{nn} \end{pmatrix}$ на блоки следующим образом:

R_1	R_3
R_4	R_2

$R_1 = \begin{pmatrix} R_{11} & \dots & R_{1n_1} \\ \vdots & \ddots & \vdots \\ R_{n_1 1} & \dots & R_{n_1 n_1} \end{pmatrix}$, $R_2 = \begin{pmatrix} R_{(n_1+1)(n_1+1)} & \dots & R_{(n_1+1)n} \\ \vdots & \ddots & \vdots \\ R_{n(n_1+1)} & \dots & R_{nn} \end{pmatrix}$,
$R_3 = \begin{pmatrix} R_{1(n_1+1)} & \dots & R_{1n} \\ \vdots & \ddots & \vdots \\ R_{n_1(n_1+1)} & \dots & R_{n_1 n} \end{pmatrix}$, $R_4 = \begin{pmatrix} R_{(n_1+1)1} & \dots & R_{(n_1+1)n_1} \\ \vdots & \ddots & \vdots \\ R_{n1} & \dots & R_{nn_1} \end{pmatrix}$.

Аналогично разобьем на блоки векторы L и V:

L_1	V_1
L_2	V_2

$L_1 = \begin{pmatrix} L_1 \\ \vdots \\ L_{n_1} \end{pmatrix}$, $L_2 = \begin{pmatrix} L_{(n_1+1)} \\ \vdots \\ L_n \end{pmatrix}$, $V_1 = \begin{pmatrix} V_1 \\ \vdots \\ V_{n_1} \end{pmatrix}$, $V_2 = \begin{pmatrix} V_{(n_1+1)} \\ \vdots \\ V_n \end{pmatrix}$.

С использованием блоков можно записать исходное уравнение в виде

$$\boxed{\begin{array}{c} L_1 \\ \hline L_2 \end{array}} = \boxed{\begin{array}{c|c} R_1 & R_3 \\ \hline R_4 & R_2 \end{array}} \circ \boxed{\begin{array}{c} L_1 \\ \hline L_2 \end{array}} \cup \boxed{\begin{array}{c} V_1 \\ \hline V_2 \end{array}}$$

или в виде системы уравнений

$$\begin{aligned} L_1 &= R_1 \circ L_1 \cup R_3 \circ L_2 \cup V_1 \\ L_2 &= R_4 \circ L_1 \cup R_2 \circ L_2 \cup V_2. \end{aligned} \tag{15}$$

Применяя к системе (15) лемму 1 и метод исключения неизвестных, получим:

$$\begin{aligned} L_1 &= (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ R_3 \circ R_2{}^{\circledast} \circ V_2 \cup \\ &\cup (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ V_1 \\ L_2 &= R_2{}^{\circledast} \circ R_4 \circ (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ R_3 \circ R_2{}^{\circledast} \circ V_2 \cup \\ &\cup R_2{}^{\circledast} \circ R_4 \circ (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ V_1 \cup R_2{}^{\circledast} \circ V_2. \end{aligned} \tag{16}$$

Рассмотрим квадратную матрицу X порядка n, построенную по следующим правилам: предположим, что R_1, R_2, R_3, R_4 – блоки матрицы R, и для матриц R_1 и R_2 уже известны матрицы $R_1^{\circledast}$ и $R_2^{\circledast}$, тогда матрица X имеет вид:

$(R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast}$	$(R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ R_3 \circ R_2{}^{\circledast}$
$R_2{}^{\circledast} \circ R_4 \circ (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast}$	$R_2{}^{\circledast} \circ R_4 \circ (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ R_3 \circ R_2{}^{\circledast} \cup R_2{}^{\circledast}$

(17)

Вычисляя значения элементов матрицы $X \circ V$, получим:

$(R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ V_1 \cup (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ R_3 \circ R_2{}^{\circledast} \circ V_2$	=	L_1
$R_2{}^{\circledast} \circ R_4 \circ (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ V_1 \cup R_2{}^{\circledast} \circ V_2 \cup R_2{}^{\circledast} \circ R_4 \circ (R_1 \cup R_3 \circ R_2{}^{\circledast} \circ R_4)^{\circledast} \circ R_3 \circ R_2{}^{\circledast} \circ V_2$		L_2

Таким образом, матрица, составленная по правилу (17) действительно совпадает с матрицей $R^{\circledast}$, и рекурсивный способ нахождения $R^{\circledast}$ заключается в следующем:
1) Если $n = 1$, то $R^{\circledast} = (R_{11})^{\circledast}$;
2) Если $n > 1$, то нужно выбрать $n_1 < n$, разбить матрицу R на блоки, вычислить $R_1{}^{\circledast}$ и $R_2{}^{\circledast}$, а затем получить $R^{\circledast}$ по правилу (17).

Для доказательства теоремы воспользуемся методом индукции по структуре формулы R.

Пусть R построена с использованием n операций. Если $n = 0$, то по определению регулярных выражений алгебры $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$ возможны следующие случаи: 1)$R = x$, где $x \in X$; 2)$R = xy$, где $x \in X$ и $y \in X$; 3)$R = \varnothing$.

В первом случае в графе G, для которого выполняется $L(G) = L(R)$, будет только одна вершина с отметкой x, которая одновременно и начальная, и финальная. Согласно (14), язык такого графа $L(G) = x = L(R)$. Во втором случае матрица переходов графа $M = \begin{pmatrix} \varnothing & xy \\ \varnothing & \varnothing \end{pmatrix}$, вектор начальных вершин $U = \begin{pmatrix} x \\ \varnothing \end{pmatrix}$, вектор финальных вершин $V = \begin{pmatrix} \varnothing \\ x \end{pmatrix}$. Согласно (14) $L(G) = U^T \circ M^{\circledast} \circ V = xy = R(L)$. В третьем случае положим, что в графе G нет ни одной вершины и $L(G) = \varnothing = L(R)$.

Если $n > 0$, то по определению регулярных выражений возможны следующие варианты: либо $R = R_1 \cup R_2$, либо $R = R_1 \circ R_2$, либо $R = R_1^{\circledast}$, причем R_1 и R_2 – формулы, построенные с использованием $n - 1$ операции.

Пусть для регулярных выражений R_1 и R_2 найдены такие графы G_1 и G_2, что $L(G_1) = L(R_1)$, $L(G_2) = L(R_2)$. Построим граф G, для которого $L(G) = L(R)$, где $R = R_1 \cup R_2$.

Обозначим через M_1 и M_2 матрицы переходов, U_1 и U_2 – векторы начальных вершин, V_1 и V_2 – векторы финальных вершин графов G_1 и G_2 соответственно. Искомый граф в матричной форме задается следующим образом. Пусть количество вершин в графах G_1 и G_2 равно n_1 и n_2, соответственно, M – квадратная матрица порядка $n = n_1 + n_2$, U и V – векторы длины n, которые с помощью блоков можно записать

$$M = \begin{array}{|l|l|}\hline M_1 & \varnothing \\ \hline \varnothing & M_2 \\ \hline \end{array}, \quad U = \begin{array}{|l|}\hline U_1 \\ \hline U_2 \\ \hline \end{array}, \quad V = \begin{array}{|l|}\hline V_1 \\ \hline V_2 \\ \hline \end{array}.$$

По правилу (17) вычислим $M^{\circledast}$:

$$M^{\circledast} = \begin{array}{|l|l|}\hline (M_1 \cup \varnothing \circ {M_2}^{\circledast} \circ \varnothing)^{\circledast} & {M_1}^{\circledast} \circ \varnothing \circ {M_2}^{\circledast} \\ \hline {M_2}^{\circledast} \circ \varnothing \circ {M_1}^{\circledast} & {M_2}^{\circledast} \cup M_2^{\circledast} \circ \varnothing \circ {M_1}^{\circledast} \circ \varnothing \circ {M_2}^{\circledast} \\ \hline \end{array} = \begin{array}{|l|l|}\hline {M_1}^{\circledast} & \varnothing \\ \hline \varnothing & {M_2}^{\circledast} \\ \hline \end{array},$$

Используя (14), получим, что
$L(G) = U^T \circ M^{\circledast} \circ V = U_1^T \circ {M_1}^{\circledast} \circ V_1 \cup U_2^T \circ {M_2}^{\circledast} \circ V_2 = L(R_1) \cup L(R_2) = L(R)$.

Покажем, что для G_1 и G_2, у которых $L(G_1) = L(R_1)$, $L(G_2) = L(R_2)$, и $R = R_1 \circ R_2$, можно построить граф G, для которого $L(G) = L(R)$.

Графу G соответствует следующее матричное представление:

$$M = \begin{array}{|l|l|}\hline M_1 & V_1 \circ {U_2}^T \circ M_2 \\ \hline \varnothing & M_2 \\ \hline \end{array}, \quad U = \begin{array}{|l|}\hline U_1 \\ \hline \varnothing \\ \hline \end{array}, \quad V = \begin{array}{|l|}\hline V_1 \circ {U_2}^T \circ V_2 \\ \hline V_2 \\ \hline \end{array}.$$

По правилу (17)

$$M^{\circledast} = \begin{array}{|l|l|}\hline {M_1}^{\circledast} & {M_1}^{\circledast} \circ V_1 \circ {U_2}^T \circ M_2 \circ {M_2}^{\circledast} \\ \hline \varnothing & {M_2}^{\circledast} \\ \hline \end{array}.$$

$L(G) = U^T \circ M^{\circledast} \circ V = {U_1}^T \circ {M_1}^{\circledast} \circ V_1 \circ {U_2}^T \circ V_2 \cup {U_1}^T \circ {M_1}^{\circledast} \circ V_1 \circ {U_2}^T \circ M_2 \circ {M_2}^{\circledast} \circ V_2 =$
$= {U_1}^T \circ {M_1}^{\circledast} \circ V_1 \circ ({U_2}^T \circ V_2 \cup {U_2}^T \circ M_2 \circ {M_2}^{\circledast} \circ V_2) = {U_1}^T \circ {M_1}^{\circledast} \circ V_1 \circ {U_2}^T \circ {M_2}^{\circledast} \circ V_2 =$
$= L(R_1) \circ L(R_2) = L(R)$.

Покажем, что для любых графа G_1 и регулярного выражения R_1, если $L(G_1) = L(R_1)$ и $R = {R_1}^{\circledast}$, то возможно построить такой граф G, что $L(G) = L(R)$.

Согласно (14) и (15) $L(G_1) = {U_1}^T \circ {M_1}^{\circledast} \circ V_1 = {U_1}^T \circ {M_1}^{+} \circ V_1 \cup {U_1}^T \circ V_1 =$
$= {U_1}^T \circ {M_1}^{*} \circ V_1$

Рассмотрим граф G, у которого матрица переходов $M = M_1 \cup V_1 \circ {U_1}^T \circ M_1$, вектор начальных вершин $U = U_1$, вектор финальных вершин $V = V_1$. Язык нового графа $L(G) = U^T \circ M^{\circledast} \circ V = U^T \circ (M \cup M')^{\circledast} \circ V$.

Так как $M' = {M_1}'$, можно записать, что $L(G) = {U_1}^T \circ ({M_1}' \cup M_1 \cup V_1 \circ {U_1}^T \circ (M_1 \cup$
$\cup {M_1}'))^{\circledast} \circ V_1 = {U_1}^T \circ ((M_1 \cup {M_1}') \cup V_1 \circ {U_1}^T \circ (M_1 \cup {M_1}'))^{*} \circ V_1$.

Из определения операции $*$ следует, что для любых R и Q

$$(R \cup Q)^* = R^* \circ (Q \circ R^*)^* \tag{17}$$

$$(R \circ Q)^* \circ R = R \circ (Q \circ R)^*. \tag{18}$$

Используя (17), можно записать
$L(G) = {U_1}^T \circ ((M_1 \cup {M_1}') \cup V_1 \circ {U_1}^T \circ (M_1 \cup {M_1}'))^* \circ V_1 =$
$= {U_1}^T \circ (M_1 \cup {M_1}')^* \circ (V_1 \circ {U_1}^T \circ (M_1 \cup {M_1}') \circ (M_1 \cup {M_1}')^*)^* \circ V_1$.

Применив (18), получим, что

$L(G) = {U_1}^T \circ (M_1 \cup {M_1}')^* \circ V_1 \circ ({U_1}^T \circ (M_1 \cup {M_1}') \circ (M_1 \cup {M_1}')^* \circ V_1)^*$.

Из определений матриц ${M_1}^*$ и ${M_1}'$ следует, что $(M_1 \cup {M_1}') \circ (M_1 \cup {M_1}')^* = {M_1}^{\circledast}$, а $(M_1 \cup {M_1}')^* = {M_1}^*$.

Таким образом, $L(G) = {U_1}^T \circ {M_1}^* \circ V_1 \circ ({U_1}^T \circ {M_1}^{\circledast} \circ V_1)^* = L(G_1) \circ (L(G_1))^* = (L(G_1))^+$. Согласно (1), $R^{\circledast} = R_{beg} \circ R_{end} \cup R^+$, поэтому, для того, чтобы из графа G, у которого $L(G) = (L(G_1))^+$, получить граф G, у которого $L(G) = (L(G_1))^{\circledast}$, нужно для всех вершин $q_i \in I$ и $q_j \in F$, у которых $\mu(q_i) = \mu(q_j)$, добавить в граф G новую вершину с отметкой $\mu(q_i)$ и сделать ее одновременно начальной и финальной.

Таким образом, во всех случаях для любого регулярного выражения R можно найти граф G, порождаемый которым язык $L(G) = L(R)$, что и требовалось доказать. □

Заключение.

В работе введена алгебра $\langle 2^{X^+}, \circ, \cup, \circledast, \varnothing, X \rangle$, и показано, что эта алгебра обладает некоторыми свойствами, аналогичными свойствам алгебры Клини $\langle 2^{X^*}, \cdot, \cup, *, \varnothing, \lambda \rangle$, что дает возможность использовать ее при анализе графов с отмеченными вершинами таким же образом, каким алгебра Клини используется при анализе графов с отмеченными переходами.

Авторы выражают благодарность В.А.Козловскому и А.Н.Курганскому за полезные замечания.

1. *Капитонова Ю.В., Летичевский А.А.* Математическая теория проектирования вычислительных систем. – М.: Наука, 1988. – 296с.
2. *Андон Ф.И., Дорошенко А.Е., Цейтлин Г.Е., Яценко Е.А.* Алгеброалгоритмические модели и методы параллельного программирования. – К.: Академпериодика, 2007. – 634с.
3. *Grunsky I., Kurganskyy O., Potapov I.* Languages Representable by Vertex-labeled Graphs. Proceedings of the 30th International Symposium on Mathematical Foundations of Computer Science, LNCS. – V.3618. – 2005. – P.435-446.
4. *Eilenberg S.* Automata, Languages, and Machines. – Academic Press, New York and London, 1974.
5. *ConwayJ. H.* Regular Algebra and Finite Machines. Chapman and Hall, London, U.K., 1971.

Ин-т прикл. математики и механики НАН Украины, Донецк
Гос. ун-т информатики и искусственного интеллекта
grunsky@iamm.ac.donetsk.ua

Получено 21.05.09