УДК 519.6

О ПОЛНОТЕ СИСТЕМЫ АКСИОМ ДЛЯ АЛГЕБРЫ ЯЗЫКОВ, ПРЕДСТАВИМЫХ КОНЕЧНЫМИ ОРИЕНТИРОВАННЫМИ ГРАФАМИ С ОТМЕЧЕННЫМИ ВЕРШИНАМИ

Е.А.Пряничникова,
Государственный университет информатики и искусственного интеллекта

Алгебра Клини является удобным средством для представления языков, задаваемых графами с отмеченными переходами[1]. В работе [2] было показано, что класс алгебр Клини образует конечно-базируемое эквациональное квазимногообразие, и доказана полнота системы аксиом, состоящей из тождеств и квазитождеств, для алгебры регулярных языков.

В настоящее время существует ряд актуальных задач, связанных с анализом графов с отмеченными вершинами. Свойства языков, представимых такими графами, изучались в работах [3,4]. В работе [5] вводится алгебра, которая может быть использована при работе с языками, представимыми графами с отмеченными вершинами, аналогично тому, как алгебра Клини используется для регулярных языков.

Основная цель данной работы – на основе системы аксиом алгебры Клини разработать систему аксиом для алгебры языков, представимых в графах с отмеченными вершинами, и доказать ее полноту.

Работа имеет следующую структуру. В разделе 1 вводятся основные определения и обозначения, дается формальная постановка задачи. В разделе 2 формулируются основные результаты и приводятся доказательства.

Основные определения и обозначения. Назовем графом с отмеченными вершинами четверку $G=(Q,E,X,\mu)$, где Q – конечное множество вершин, $|Q|=n$, $E\subseteq Q\times Q$ – множество дуг, X – конечный алфавит отметок вершин, $\mu: Q\to X$ - функция отметок вершин. Пусть $I\subseteq Q$ - множество начальных, а $F\subseteq Q$ - множество финальных вершин графа. Вершины q_1 и q_2 будем называть смежными, если $(q_1,q_2)\in E$. Путем в графе G будем называть конечную последовательность смежных вершин $l=q_1\,q_2...q_k$, где $(q_i,q_{i+1})\in E$, число $k-1$ будем называть длиной пути, q_1 - начальная вершина пути, q_k - конечная вершина. Пусть X^+ - множество всех непустых слов конечной длины в алфавите X. Элемент $x=\mu(q_1)\,\mu(q_2)...\mu(q_k)=x_1x_2...x_k\in X^+$ будем называть отметкой пути l в графе G. 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. Языком, порождаемым графом G, назовем отметки всех путей в графе G, начальные вершины которых принадлежат множеству I, а конечные – множеству F. Язык, порождаемый графом G, обозначим $L(G)$.

В работе [3] было предложено использовать для соединения слов из языков, представимых в графах с отмеченными вершинами, вместо операции конкатенации, применяемой в алгебре Клини, операцию $\circ$, которая на множестве слов в конечном алфавите X определяется следующим образом: для всех $w_1,w_2\in X^+$ и всех $x,y\in X$

$$w_1x\circ yw_2=\begin{cases} w_1xw_2, & \textit{если}\ \ x=y;\\ \textit{не определено}, & \textit{в противном случае.}\end{cases}$$

Рассмотрим алгебру $\left\langle 2^{X^+},\circ,\cup,\otimes,\varnothing,X\right\rangle$ со следующими операциями на языках $L,R\in 2^{X^+}$.

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

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

3) $L^*=\bigcup\limits_{i=0}^{\infty}L^i$, где $L^0=X;L^{n+1}=L^n\circ L,n\ge 0$;

4) $L^{\otimes} = L_{beg} \circ L^{*} \circ L_{end}$,

$$L_{beg} = \{x \mid xw \in L,\ x \in X, w \in X^{*}\},\ L_{end} = \{x \mid wx \in L,\ x \in X, w \in X^{*}\}.$$

Назовем множеством регулярных выражений алгебры $\left\langle 2^{X^{+}}, \circ, \cup, \otimes, \varnothing, X \right\rangle$ множество $T_{\sigma}(X)$ всех термов сигнатуры $\sigma = (\circ, \cup, \otimes, \varnothing, X)$, порожденных алфавитом $X \cup X^{2}$, где $X^{2} = \{xy \mid x \in X,\ y \in X\}$.

Пусть $\ell \subseteq 2^{X^{+}}$ – множество всех языков, представимых в конечных графах с отмеченными вершинами с алфавитом отметок X. Назовем стандартной интерпретацией регулярных выражений такой гомоморфизм $L : T_{\sigma}(X) \to \ell$, что $L(x) = \{x\}$ для всех $x \in X$, $L(xy) = \{xy\}$ для всех $xy \in X^{2}$, а для всех $p, q \in T_{\sigma}(X)$ выполняется $L(p \cup q) = L(p) \cup L(q)$, $L(p \circ q) = L(p) \circ L(q)$, $L(p^{\otimes}) = (L(p))^{\otimes}$.

В работе [5] доказана теорема, аналогичная теореме Клини для конечных автоматов: язык $L \subseteq X^{+}$ представим регулярным выражением алгебры $\left\langle 2^{X^{+}}, \circ, \cup, \otimes, \varnothing, X \right\rangle$ тогда и только тогда, когда он порождается графом с отмеченными вершинами.

Рассмотрим систему аксиом для алгебры языков, порождаемых графами с отмеченными вершинами, построенную на основе системы аксиом для алгебры Клини из работы [2] с учетом особенностей рассматриваемой алгебраической системы.

$p \cup (q \cup r) = (p \cup q) \cup r$	(1)	$(p \cup q) \circ r = p \circ r \cup q \circ r$	(9)
$p \cup q = q \cup p$	(2)	$\varnothing \circ p = \varnothing$	(10)
$p \cup \varnothing = p$	(3)	$p \circ \varnothing = \varnothing$	(11)
$p \cup p = p$	(4)	$p^{\otimes} \cup X \geq p^{\otimes} \circ p \cup p \cup X$	(12)
$p \circ (q \circ r) = (p \circ q) \circ r$	(5)	$p^{\otimes} \cup X \geq p \circ p^{\otimes} \cup p \cup X$	(13)
$X \circ p = p$	(6)	$q \cup p \circ y \leq y \Rightarrow p^{\otimes} \circ q \cup q \leq y$	(14)
$p \circ X = p$	(7)	$q \cup y \circ p \leq y \Rightarrow q \circ p^{\otimes} \cup q \leq y$	(15)
$p \circ (q \cup r) = p \circ q \cup p \circ r$	(8)	$p \leq X \Rightarrow p \circ p = p$	(16)

Отношение $\leq$ является естественным частичным порядком: $p \leq q$ тогда и только тогда, когда $p \cup q = q$, и монотонно по отношению к операциям алгебры $\left\langle 2^{X^{+}}, \circ, \cup, \otimes, \varnothing, X \right\rangle$.

Для любой алгебры $A = \langle A, \circ, \cup, \otimes, \varnothing, X \rangle$ из (1-4) следует, что $\langle A, \cup, \varnothing \rangle$ – идемпотентный коммутативный моноид, из (5-7) следует, что $\langle A, \circ, X \rangle$ – моноид, тождества (1-11) говорят о том, что $\langle A, \cup, \circ, \varnothing, X \rangle$ – идемпотентное полукольцо.

Основная цель данной работы – доказательство того, что дедуктивная система, состоящая из аксиом (1-16) полна для алгебры языков, представимых конечными ориентированными графами с отмеченными вершинами: два регулярных выражения p и q над конечным алфавитом X обозначают один и тот же язык, представимый графом, тогда и только тогда, когда формула $p = q$ является логическим следствием аксиом (1-16).

Доказательство полноты системы аксиом. Система аксиом, используемая в данной работе, является результатом приспособления системы аксиом алгебры Клини из работы [2] для адекватного представления алгебры языков, порождаемых конечными графами с отмеченными вершинами.

В работе [2] доказательство полноты построено следующим образом: для двух регулярных выражений, обозначающих один и тот же регулярный язык, строятся минимальные детерминированные автоматы. Так как распознаваемые этими автоматами языки равны, автоматы изоморфны. Показано, что поведение любого конечного автомата можно задать с помощью матриц над алгеброй Клини, причем класс таких матриц тоже образует алгебру Клини. Идея доказательства заключается в том, что если минимальные автоматы изоморфны, то, применяя аксиомы к матрицам, описывающим эти автоматы, можно получить равенство тех регулярных выражений, по которым эти автоматы строились.

Доказательство полноты системы аксиом (1-16), приведенное в данной работе, основано на других соображениях, поскольку авторам не хотелось привлекать для доказательства средства, лежащие за пределами алгебры языков – вводить дополнительно алгебру графов и алгебру матриц.

Вместо этого в данной работе показано, что любому терму p свободной алгебры квазимногообразия, описываемого тождествами и квазитождествами (1-16) можно поставить в соответствие систему уравнений таким образом, что линейная комбинация решений этой системы даст регулярное выражение p', представляющее тот же язык, что и исходный терм, причем из аксиом (1-16) следует, что $p = p'$. За счет введения порядка на множестве всех слов X^+ можно добиться того, что для всех регулярных выражений, представляющих один и тот же язык, соответствующие коэффициенты систем уравнений будут одинаковыми. Тогда можно показать, что для любых регулярных выражений p и q, представляющих один и тот же язык, из (1-16) следует равенство регулярных выражений, соответствующих решениям систем для p и q, то есть $p' = q'$, откуда следует, что $p = q$ также следует из (1-16).

Теорема 1. Система аксиом (1-16) полна для алгебры языков, представимых ориентированными конечными графами с отмеченными вершинами, то есть, если q и p – два регулярных выражения над алфавитом X, то $L(p) = L(q)$ тогда и только тогда, когда из (1-16) следует, что $p = q$.

Доказательство. Докажем первую часть теоремы: если из аксиом (1-16) логически следует, что $p = q$, то $L(p) = L(q)$. В силу того, что все аксиомы обладают тем свойством, что стандартные интерпретации регулярных выражений в левых частях тождеств и квазитождеств совпадают со стандартными интерпретациями их правых частей, если $p = q$ является следствием из аксиом, то $L(p) = L(q)$ для любых p и q. Что и требовалось доказать.

Докажем вторую часть теоремы. Пусть $L(p) = L(q)$, покажем, что $p = q$ является следствием аксиом (1-16). Построим для произвольного регулярного выражения p систему уравнений, решением которой будет набор регулярных выражений, из которых можно будет получить такое регулярное выражение p', что $p = p'$ будет следствием из аксиом (1-16).

Введем понятие производной по слову $x \in X^+$ для регулярных выражений алгебры $\left\langle 2^{X^+}, \circ, \cup, \otimes, \varnothing, X \right\rangle$ по аналогии с введенным в [7] понятием производной для регулярных выражений алгебры Клини.

Рекурсивно определим производную $D_w(p)$ регулярного выражения p по слову w следующим образом.

Если $p = \varnothing$, то $D_w(p) = \varnothing$ для любого $w \in X^+$. (17)

Если $w \in X$, то $D_w(p) = w \circ p$ для любого регулярного выражения p. (18)

Если $w \in X^2$, то

В случае, когда p состоит из одного слова u, причем $u \in X$, $D_w(p) = \varnothing$; (19)

В случае, когда p состоит из одного слова u, где $u \in X^2$, (20)

$$D_w(p) = \begin{cases} a \mid w \circ a = w,\ a \in X,\ \textit{если}\ u = w \\ \varnothing,\ \textit{в противном случае} \end{cases}$$

Если $p = p_1 \cup p_2$, то $D_w(p) = D_w(p_1) \cup D_w(p_2)$. (21)

Если $p = p_1 \circ p_2$, то $D_w(p) = D_w(p_1) \circ p_2 \cup D_w(\varepsilon(p_1) \circ p_2)$. (22)

Если $p = p_1^{\otimes}$, то $D_w(p) = D_w(p) \cup D_w(p) \circ p^{\otimes}$. (23)

Если $w = u \circ v$, где $u, v \in X^2$, то $D_w(p) = D_v(D_u(P))$. (24)

Здесь $\varepsilon(p)$ – это объединение всех букв алфавита X, которые входят в язык, являющийся стандартной интерпретацией p, как отдельные слова. Рекурсивно определим $\varepsilon(p)$ следующим образом:

$$\varepsilon(\varnothing)=\varnothing;\ \varepsilon(a)=a;\ \varepsilon(ab)=\varnothing \tag{25}$$

$$\varepsilon(p\cup q)=\varepsilon(p)\cup\varepsilon(q) \tag{26}$$

$$\varepsilon(p\circ q)=\varepsilon(p)\circ\varepsilon(q) \tag{27}$$

$$\varepsilon(p^{\otimes})=\varepsilon(p)+\alpha\text{, где } \alpha=\bigcup a \mid a\in X, a\circ p\neq\varnothing, p\circ a\neq\varnothing, \tag{28}$$

$$\varepsilon(\varepsilon(p))=\varepsilon(p) \tag{29}$$

Покажем, что из аксиом (1-16) следует, что любое регулярное выражение можно представить в виде комбинации его производных.

Лемма 1. Из аксиом (1-16) следует, что для любого регулярного выражения p

$$p=\bigcup_{w\in X^2} w\circ D_w(p)\cup\varepsilon(p) \tag{29}$$

Доказательство. Для доказательства воспользуемся методом индукции по структуре формулы p. Пусть p построена с использованием n операций. Если $n=0$, то по определению регулярных выражений возможны следующие случаи:

1) $p=\varnothing$. В этом случае $\varepsilon(p)=\varnothing$ по (25), $D_w(\varnothing)=\varnothing$ для всех w по (17).

Исходя из (11) и (3) $\bigcup_{w\in X^2} w\circ D_w(p)\cup\varepsilon(p)=\varnothing=p$.

2) $p=x,\ x\in X$. В этом случае $\varepsilon(p)=x$ по (25), $D_w(x)=\varnothing$ для всех w по (19).

По аксиомам (11) и (3) $\bigcup_{w\in X^2} w\circ D_w(p)\cup\varepsilon(p)=x=p$.

3) $p=xy,\ xy\in X^2$. В этом случае $\varepsilon(xy)=\varnothing$ по (25); $D_{xy}(xy)=z$,

где $z\in X,\ xy\circ z=xy$; $D_{ab}(xy)=\varnothing$ для всех $ab\in X^2$ если $ab\neq xy$ по (20). Тогда $\bigcup_{w\in X^2} w\circ D_w(p)\cup\varepsilon(p)=xy\circ z=xy=p$.

Если $n>0$, то по определению регулярных выражений возможны следующие варианты: либо $p=p_1\cup p_2$, либо $p=p_1\circ p_2$, либо $p=p_1^{\otimes}$, причем p_1 и p_2 – формулы, построенные с использованием $n-1$ операции.

Рассмотрим первый вариант.

Пусть $p=p_1\cup p_2$, причем для p_1 и p_2 выполняется утверждение леммы:

$$p_1=\bigcup_{w\in X^2} w\circ D_w(p_1)\cup\varepsilon(p_1);\ p_2=\bigcup_{w\in X^2} w\circ D_w(p_2)\cup\varepsilon(p_2).$$

Покажем, что $p=\bigcup_{w\in X^2} w\circ D_w(p)\cup\varepsilon(p)$.

По определениям (21) и (26) можно записать:

$$\bigcup_{w\in X^2} w\circ D_w(p)\cup\varepsilon(p)=\bigcup_{w\in X^2} w\circ D_w(p_1\cup p_2)\cup\varepsilon(p_1\cup p_2)=$$

$$=\bigcup_{w\in X^2} w\circ(D_w(p_1)\cup D_w(p_2))\cup\varepsilon(p_1)\cup\varepsilon(p_2).$$

Полученное выражение с помощью аксиом (8) и (2) можно записать в виде

$$\bigcup_{w\in X^2} w\circ D_w(p_1)\cup\varepsilon(p_1)\cup\bigcup_{w\in X^2} w\circ D_w(p_2)\cup\varepsilon(p_2)=p_1\cup p_2=p$$

Рассмотрим второй вариант. Пусть $p = p_1 \circ p_2$, и для p_1 и p_2 выполняется утверждение леммы. По определениям (22) и (27) и аксиомам (2) и (8) можно записать:

$$\bigcup_{w \in X^2} w \circ D_w(p) \cup \varepsilon(p) = \bigcup_{w \in X^2} w \circ D_w(p_1 \circ p_2) \cup \varepsilon(p_1 \circ p_2) =$$

$$= \bigcup_{w \in X^2} w \circ D_w(p_1) \circ p_2 \cup \bigcup_{w \in X^2} w \circ D_w(\varepsilon(p_1) \circ p_2) \cup \varepsilon(p_1) \circ \varepsilon(p_2).$$

По определению (29) $\varepsilon(\varepsilon(p_1) \circ p_2) = \varepsilon(\varepsilon(p_1)) \circ \varepsilon(p_2) = \varepsilon(p_1) \circ \varepsilon(p_2)$, поэтому можно записать:

$$\bigcup_{w \in X^2} w \circ D_w(p_1) \circ p_2 \cup \bigcup_{w \in X^2} w \circ D_w(\varepsilon(p_1) \circ p_2) \cup \varepsilon(p_1) \circ \varepsilon(p_2) =$$

$$= \bigcup_{w \in X^2} w \circ D_w(p_1) \circ p_2 \cup \bigcup_{w \in X^2} w \circ D_w(\varepsilon(p_1) \circ p_2) \cup \varepsilon(\varepsilon(p_1) \circ p_2) =$$

$$= \bigcup_{w \in X^2} w \circ D_w(p_1) \circ p_2 \cup \varepsilon(p_1) \circ p_2 = p_1 \circ p_2 = p$$ (по аксиоме (8)).

Наконец, в случае, когда $p = p_1^{\otimes}$, и $p_1 = \bigcup_{w \in X^2} w \circ D_w(p_1) \cup \varepsilon(p_1)$, по определению (23) можно записать: $\bigcup_{w \in X^2} w \circ D_w(p) \cup \varepsilon(p) = \bigcup_{w \in X^2} w \circ D_w(p_1^{\otimes}) \cup \varepsilon(p_1^{\otimes}) =$

$$= \bigcup_{w \in X^2} w \circ D_w(p_1) \cup \bigcup_{w \in X^2} w \circ D_w(p_1) \circ p_1^{\otimes} \cup \varepsilon(p_1^{\otimes}) = p_1^{\otimes} = p$$ (по аксиоме (13)).

Таким образом, во всех случаях для регулярного выражения p выполняется $p = \bigcup_{w \in X^2} w \circ D_w(p) \cup \varepsilon(p)$, что и требовалось доказать. □

Лемма 2. У любого регулярного выражения p количество различных производных конечно.

Доказательство. Докажем лемму по индукции. Базисом будут следующие три случая:

1) $p = \varnothing$, 2) $p = x$, где $x \in X$, 3) $p = xy$, где $xy \in X^2$. Во всех этих случаях конечность числа производных непосредственно следует из определений (17), (19), (20).

Предположим, что $p = p_1 \cup p_2$, где p_1 и p_2 – формулы, число различных производных которых конечно и равно n_1 и n_2 соответственно. Тогда из определения (21) следует, что для формулы p число различных производных $n = n_1 + n_2$, то есть тоже конечно.

Аналогичными рассуждениями можно показать, что число производных формулы p будет конечным и в тех случаях, когда $p = p_1 \circ p_2$ и $p = p_1^{\otimes}$.□

По лемме 1 для любого регулярного выражения p каждую из различных производных $D_i(p)$ (которых по лемме 2 конечное число) можно записать в виде

$$D_i(p) = \bigcup_{w \in X^2} w \circ D_j(p) \cup \varepsilon(D_i(p)), \tag{30}$$

где $D_j(p)$ – это такая из различных производных регулярного выражения p, что, если s_i – это любое слово, производная по которому совпадает с $D_i(p)$, то производная по слову $s_i \circ w$ совпадает с $D_j(p)$.

Лемма 3. Из аксиом (1-16) следует, что уравнение вида $p \circ x \cup q = x$ имеет минимальное решение $p^{\otimes} \circ q \cup q$.

Доказательство. Преобразуем аксиому $p^{\otimes} \cup X \geq p \circ p^{\otimes} \cup p \cup X$ (13) воспользовавшись аксиомами (9) и (6) в $(p^{\otimes} \cup X) \circ q \geq (p \circ p^{\otimes} \cup p \cup X) \circ q$, откуда получим

$p^{\otimes} \circ q \cup q \geq p \circ p^{\otimes} \circ q \cup p \circ q \cup q$ или $p^{\otimes} \circ q \cup q \geq p \circ (p^{\otimes} \circ q \cup q) \cup q$. Таким образом, $p^{\otimes} \circ q \cup q$ является решением уравнения $p \circ x \cup q = x$. Из (14) следует, что, если x является решением этого уравнения, то $x \geq p^{\otimes} \circ q \cup q$, то есть $p^{\otimes} \circ q \cup q$ – наименьшее решение уравнения $p \circ x \cup q = x$, что и требовалось доказать. □

Из леммы 3 следует, что систему, составленную из всех уравнений (30), можно решить способом, аналогичным методу Гаусса для систем линейных уравнений и получить ее наименьшее решение.

Лемма 4. Из аксиом (1-16) следует, что для любого регулярного выражения

$$p = \bigcup_{w \in X^+} w \circ D_w(p) \tag{31}$$

Доказательство. Для доказательства леммы воспользуемся методом структурной индукции. По определению производных

если $p = \varnothing$, то $\bigcup_{w \in X^+} w \circ D_w(p) = \varnothing = p$;

если $p = x$, $x \in X$, то $\bigcup_{w \in X^+} w \circ D_w(p) = \bigcup_{w \in X} w \circ D_w(p) = X \circ x = x = p$ (по аксиоме (8));

если $p = xy$, $xy \in X^2$, то $\bigcup_{w \in X^+} w \circ D_w(p) = X \circ xy \cup xy \circ D_{xy}(xy) = xy = p$ (по аксиомам (8) и (1) и (3));.

Предположим, что $p = p_1 \cup p_2$ и для p_1 и p_2 выполняется утверждение леммы: $p_1 = \bigcup_{w \in X^+} w \circ D_w(p_1)$ и $p_2 = \bigcup_{w \in X^+} w \circ D_w(p_2)$. Тогда

$$\bigcup_{w \in X^+} w \circ D_w(p) = \bigcup_{w \in X^+} w \circ D_w(p_1 \cup p_2) = \bigcup_{w \in X^+} w \circ D_w(p_1) \cup \bigcup_{w \in X^+} w \circ D_w(p_2) = p_1 \cup p_2 = p.$$

Пусть $p = p_1 \circ p_2$ и $p_1 = \bigcup_{w \in X^+} w \circ D_w(p_1)$, $p_2 = \bigcup_{w \in X^+} w \circ D_w(p_2)$. В этом случае

$$\bigcup_{w \in X^+} w \circ D_w(p) = \bigcup_{w \in X^+} w \circ D_w(p_1 \circ p_2) =$$

$$= \bigcup_{w \in X^+} w \circ D_w(p_1) \circ p_2 \cup \bigcup_{w \in X^+} w \circ D_w(\varepsilon(p_1) \circ p_2) = p_1 \circ p_2 \cup \varepsilon(p_1) \circ p_2 =$$

$= (p_1 \cup \varepsilon(p_1)) \circ p_2$. Так как по определению $\varepsilon(p_1) \leq p_1$, то $(p_1 \cup \varepsilon(p_1)) \circ p_2 = p_1 \circ p_2 = p$

Пусть $p = p_1^{\otimes}$ и $p_1 = \bigcup_{w \in X^+} w \circ D_w(p_1)$. В этом случае по определениям (18) и (23)

$$\bigcup_{w \in X} w \circ D_w(p_1^{\otimes}) = \bigcup_{w \in X} w \circ p_1^{\otimes} = X \circ p_1^{\otimes} = p_1^{\otimes},$$

$$\bigcup_{w \in X^2} w \circ D_w(p_1^{\otimes}) = \bigcup_{w \in X^2} w \circ D_w(p_1) \cup \bigcup_{w \in X^2} w \circ D_w(p_1) \circ p_1^{\otimes} = p_1 \cup p_1 \circ p_1^{\otimes},$$ то есть

$$\bigcup_{w \in X^+} w \circ D_w(p) = \bigcup_{w \in X^+} w \circ D_w(p_1^{\otimes}) = p_1^{\otimes} \cup p_1 \cup p_1 \circ p_1^{\otimes} = p_1^{\otimes} = p$$ (по аксиоме (13)

Таким образом, во всех случаях $p = \bigcup_{w \in X^+} w \circ D_w(p)$, что и требовалось доказать.

Из лемм 3 и 4 следует, что $p = \bigcup_{i=1}^{n} w_i \circ D_i(p)$ для любого регулярного выражения p, где n – число различных производных формулы p, $D_i(p)$ – одна из таких производных, w_i – любое слово, производная по которому совпадает с $D_i(p)$.

Таким образом, с использованием понятия производных регулярных выражений можно построить для произвольного регулярного выражения p систему уравнений, решение которой существует и дает возможность получить такое регулярное выражение p', что из аксиом (1-16) следует $p = p'$.

Покажем, что систему уравнений для регулярного выражения можно строить таким образом, чтобы для двух регулярных выражений с одинаковыми интерпретациями системы уравнений были одинаковыми, то есть количество неизвестных было одинаковым и соответствующие коэффициенты совпадали. Введем понятие производной языка $L \subseteq X^+$,представимого конечным ориентированным графом с отмеченными вершинами, по слову $w \in X^+ : D_w L = \{y | w \circ y \in L\}$.

Из определения производных регулярных выражений следует, что все производные также являются регулярными выражениями, значит для них тоже можно рассматривать стандартные интерпретации.

Лемма 5. Пусть язык $L(p)$ - стандартная интерпретация регулярного выражения p. Тогда для любого слова $w \in X^+$ производная регулярного выражения p по слову w описывает язык, совпадающий с производной языка $L(p)$ по слову w : $L(D_w(p)) = D_w(L(p))$.

Доказательство. Докажем лемму по индукции.

Если $p = x$, $x \in X$, то тогда $L(p) = \{x\}$. По определению производных регулярных выражений, у формулы p в этом случае есть только 2 различные производные: $D_w(x) = x$, если $w = x$, и $D_w(x) = \varnothing$, если $w \neq x$. У языка $L(p) = \{x\}$ также есть только две различные производные: $D_w(L(p)) = \begin{cases} \{x\}, \textit{если } w = x \\ \varnothing, \textit{если } w \neq x \end{cases}$.

Таким образом, в обоих случаях $L(D_w(p)) = D_w(L(p))$.

Если $p = xy$, $xy \in X^2$, то у формулы p есть три различные производные: $D_x xy = xy$; $D_{xy} xy = y$; $D_w x = \varnothing$, если $w \neq x$ и $w \neq xy$. У языка $L(p) = \{xy\}$ тоже есть три различные производные: $D_x\{xy\} = \{xy\}$; $D_{\{xy\}}\{xy\} = \{y\}$; $D_w\{xy\} = \varnothing$, если $w \neq x$ и $w \neq xy$. Таким образом, и в этом случае выполняется утверждение леммы $L(D_w(p)) = D_w(L(p))$.

Если $p = \varnothing$, то единственная производная этого регулярного выражения $D_w(p) = \varnothing$ для всех $w \in X^+$ совпадает с единственной производной языка $L(p) = \varnothing$.

Пусть $p = p_1 \cup p_2$, причем $L(D_w(p_1)) = D_w(L(p_1))$ и $L(D_w(p_2)) = D_w(L(p_2))$ для всех $w \in X^+$.

По определению производных регулярных выражений (21) $D_w(p_1 \cup p_2) = D_w(p_1) \cup D_w(p_2)$, откуда $L(D_w(p_1 \cup p_2)) = L(D_w(p_1) \cup D_w(p_2)) = L(D_w(p_1)) \cup L(D_w(p_2))$.

По определению производных языков $D_w(L(p_1)) = \{y | w \circ y \in L(p_1)\}$, $D_w(L(p_2)) = \{y | w \circ y \in L(p_2)\}$, поэтому $D_w(L(p_1)) \cup D_w(L(p_1)) = \{y | w \circ y \in L(p_1) \cup L(p_2)\} = \{y | w \circ y \in L(p)\} = D_w(L(p))$, значит в этом случае $L(D_w(p)) = D_w(L(p))$.

Пусть $p = p_1 \circ p_2$, $L(D_w(p_1)) = D_w(L(p_1))$ и $L(D_w(p_2)) = D_w(L(p_2))$ для всех $w \in X^+$. В этом случае $L(D_w(p)) = L(D_w(p_1 \circ p_2)) = L(D_w(p_1) \circ L(p_2) \cup L(\varepsilon(p_1)) \circ L(D_w(p_2)) =$
$= D_w(L(p_1)) \circ L(p_2) \cup L(\varepsilon(p_1)) \circ D_w(L(p_2))$.
$D_w(L(p)) = \{y | y \circ w \in L(p_1) \circ L(p_2)\} = \{u \circ v | w \circ u \in L(p_1), v \in L(p_2)\} \cup$

$\cup \{u \circ w \circ v | u \in L(p_1), u \in X, w \circ v \in L(p_2)\} = \{u | w \circ u \in L(p_1)\} \circ \{v | v \in L(p_2)\} \cup$
$\cup \{u | u \in L(p_1), u \in X\} \circ \{v | w \circ v \in L(p_2)\} = D_w(L(p_1)) \circ L(p_2) \cup L(\varepsilon(p_1)) \circ D_w(L(p_2)) =$
$= L(D_w(p))$, то есть в этом случае также $L(D_w(p)) = D_w(L(p))$.

Пусть $p = p_1^{\otimes}$, $L(D_w(p_1)) = D_w(L(p_1))$ и $L(D_w(p_2)) = D_w(L(p_2))$ для всех $w \in X^+$. В этом случае $L(D_w(p)) = L(D_w(p_1^{\otimes})) = L(D_w(p_1)) \cup L(D_w(p_1) \circ L(p_1^{\otimes})$.
$D_w(L(p)) = D_w(L(p_1^{\otimes})) = \{u | w \circ u \in L(p_1^{\otimes})\} = \{u | w \circ u \in L(p_1)\} \cup \{u | w \circ u \in L(p_1) \circ L(p_1^{\otimes})\} = L(D_w(p))$
Таким образом, во всех случаях $L(D_w(p)) = D_w(L(p))$, что и требовалось доказать. □

Из леммы 5 следует, что для регулярных выражений p и q, описывающих одинаковые языки $L(p) = L(q)$ выполняется $L(D_w(p)) = D_w(L(p)) = D_w(L(q)) = L(D_w(q))$ для любого слова $w \in X^+$. Таким образом, если $L(p) = L(q)$, то каждой из различных производных $D_i(p)$ регулярного выражения p соответствует одна из различных производных $D_j(q)$ регулярного выражения q: $L(D_i(p)) = L(D_j(q))$.

Пусть число различных производных регулярного выражения p равно n. Множество X^+ разбивается на n классов эквивалентности Y_i так, что слова u и v принадлежат одному классу, когда $L(D_u(p)) = L(D_v(p))$, то есть каждой из различных производных регулярного выражения p соответствует один из классов. Внутри каждого из классов слова можно упорядочить по длине, а слова одной длины – в лексикографическом порядке, и каждому классу сопоставить минимальное слово в этом классе. Эти минимальные слова можно в свою очередь упорядочить таким же образом, одновременно упорядочивая и классы. Пусть $L(p) = L(q)$. Тогда каждому классу соответствует одна из производных регулярного выражения q и одна из производных регулярного выражения p (не обязательно одинаковые), и различные производные каждого из этих регулярных выражений тоже можно упорядочить так, чтобы производным $D_i(p)$ и $D_i(q)$ соответствовал один и тот же класс Y_i и соответствующее этому классу минимальное слово s_i.

Тогда по лемме 1 каждую производную $D_i(q)$ можно записать в виде

$D_i(q) = \bigcup_{w \in X^2} w \circ D_j(q) \cup \varepsilon(D_i(q))$, где слово s_i – это минимальное слово в классе Y_i, который соответствует производной $D_i(q)$, а $D_j(q)$ – это такая из различных производных, что слово $s_i \circ w$ принадлежит классу Y_j.

Так как классы Y_i для производных $D_i(p)$ и $D_i(q)$ совпадают, то и системы уравнений, описывающие все $D_i(p)$ и $D_i(q)$ будут совпадать, а следовательно, будут совпадать и их решения. В этом случае регулярные выражения p' и q' можно записать в виде: $p' = \bigcup_{i=1}^{n} s_i \circ D_i(p)$ $q' = \bigcup_{i=1}^{n} s_i \circ D_i(q)$.

Таким образом, из аксиом (1-16) следует, что $p' = q'$. Ранее было показано, что из аксиом следует $p = p'$ и $q = q'$, поэтому $p = q$ также является следствием из аксиом (1-16). Что и требовалось доказать. □

Выводы. В работе показано, что алгебра языков, представимых в графах с отмеченными вершинами, является свободной алгеброй в конечно-базируемом квазимногообразии, задаваемом универсальной хорновской теорией (1-16).

РЕЗЮМЕ

Досліджуються властивості алгебри мов, що можуть бути представлені за допомогою скінченних орієнтованих графів з поміченими вершинами. Запропонована система аксіом для цієї алгебри та доведена її повнота.

SUMMARY

In this paper we study an algebra of languages that can be represented by vertex-labeled graphs, give a finitary axiomatization of this algebra and prove a completeness theorem.

СПИСОК ЛИТЕРАТУРЫ

1. Eilenberg S. Automata, Languages, and Machines / S. Eilenberg. – New York: Academic Press, 1974. – 451p.
2. Kozen D. A completeness theorem for Kleene algebras and the algebra of regular events/ D. Kozen // Information and Computation. – 1994. – №110(2) – P. 366–390.
3. Капитонова Ю.В. Математическая теория проектирования вычислительных систем / Ю.В. Капитонова, А.А. Летичевский. – М.: Наука, 1988. – 296 с.
4. Grunsky I. Languages Representable by Vertex-labeled Graphs / I. Grunsky, O. Kurganskyy, I. Potapov // Proc. of the 30th International Symposium on Mathematical Foundations of Computer Science. – 2005. – V. 3618 – P. 435-446.
5. Грунский И.С. Об алгебре языков, представимых графами с отмеченными вершинами / И.С. Грунский, Е.А. Пряничникова // Труды Ин-та прикл. математики и механики НАН Украины. – 2009. – т.18. – С. 37-46.
6. Пентус А. Е. Теория формальных языков: Учебное пособие / А. Е. Пентус, М. Р. Пентус. — М.: Изд-во ЦПИ при механико-математическом ф-те МГУ, 2004. — 80 с.
7. Brzozowski J. A. Derivatives of regular expressions / J.A.Brzozowski // Journal of the Association of Computing Machinery– 1964. – №11 – P. 481– 494.

Поступила в редакцию 10.03.2010 г.