УДК 510.5+519.17

© *А. Ю. Сапаров, А. П. Бельтюков*

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ ИЗОБРАЖЕНИЙ ФОРМУЛ С ЦЕЛЬЮ ИХ РАСПОЗНАВАНИЯ

Работа посвящена использованию основных элементов теории графов в задаче распознавания математических формул. Вводятся понятия двухуровневых и двумерно ориентированных графов, которые позволяют описывать сложные изображения, состоящие из иерархии частей с особым взаимным расположением. Рассматривается специальное отображение, которое из математической формулы строит соответствующий двумерно ориентированный граф, называемый графом изображения формулы. Приводятся правила отображения для основных классов математических формул. Описывается метод решения задачи распознавания, основанного на обратной задаче получения графа изображения формулы.

Ключевые слова: двухуровневый граф, двумерно ориентированный граф, граф изображения формулы, распознавание математических формул.

Введение

Основная сложность в распознавании математических формул состоит в сложности структуры соответствующего изображения. В отличие от строк обычных текстов, в которых отдельные символы располагаются в строгой последовательности, в математических выражениях отдельные составляющие могут быть расположены в более сложной форме, удовлетворяющей некоторому набору правил. Если в обычном тексте соблюдается последовательность записи слева направо (или в отдельных случаях справа налево, как, например, в арабском тексте), то в формулах прослеживается некоторая иерархия частей. В таких текстах одни символы могут находиться над другими (например, дроби), правее и чуть выше (степени), правее и чуть ниже (нижние индексы) и т.д.

Дополнительная сложность заключается в сложности отдельных символов, а именно в разнообразии их форм. Для рукописных символов невозможно выделить базу шаблонов в растровом формате, как это обычно делается при распознавании печатных текстов, так как печатных шрифтов ограниченное число, чего нельзя сказать о рукописных. Но все же вне зависимости от почерка в них соблюдается определенная закономерность, заключающаяся в сходстве их скелетов, а отличия выражаются только в толщине линий, наклонах, поворотах. В этом случае выделение некоторого шаблона становится возможным, если символы рассматривать не как фрагменты растровых изображений, а как более сложные объекты, состоящие из отдельных кривых и точек, в которых эти кривые соединяются друг с другом.

При распознавания текстов с такими сложными структурами необходимо строго соблюдать взаимное расположение их отдельных частей, а также иерархию внутри этих частей. Для этого предлагается рассмотреть специальные математические объекты, которые описывают соответствующие изображения текстов. Так как исходное изображение состоит из двух основных уровней, первым из которых является изображение формулы как единого целого, а вторым изображения отдельных символов как независимых объектов, то для ее моделирования должна быть использована структура, также состоящая их двух уровней и учитывающая взаимосвязь между ними.

§ 1. Понятие двумерно ориентированного графа

Следующие три определения целиком взяты из работы [1].

Определение 1. *Обыкновенным графом* $L = (X, V)$ называется упорядоченная пара множеств: конечного непустого множества X, элементы которого называются *вершинами* графа L, и произвольного подмножества $V \subseteq \tilde{X}^{[2]}$, элементы которого называются *ребрами* этого графа. Вершины $x, y \in X$ смежны, если $\tilde{xy} \in V$, и *несмежны*, если $\tilde{xy} \notin V$. Ребро $\tilde{xy} \in V$ соединяет вершины x и y (или, что то же самое, y и x), а также *инцидентно* каждой из этих вершин (и наоборот, обе вершины инцидентны этому ребру); ребро можно обозначать и одной буквой (u, v, w и т.п.), если не требуется указывать, какие именно вершины оно соединяет.

Определение 2. Связный граф без циклов называется *деревом* и обозначается T.

Определение 3. Пусть $L = (X, V)$ — обыкновенный граф. Последовательность вида

$$x_0 \tilde{x_0 x_1} x_1 \tilde{x_1 x_2} x_2 \ldots x_{l-1} \tilde{x_{l-1} x_l} x_l, \tag{1.1}$$

где $x_0, x_1, x_2, \ldots, x_l \in X$, $\tilde{x_0 x_1}, \tilde{x_1 x_2}, \ldots, \tilde{x_{l-1} x_l} \in V$, называется *маршрутом длины l из вершины x_0 в вершину x_l*. При $x_l = x_0$ и $l \geqslant 1$ маршрут называется циклическим. В случае $l = 0$ маршрут состоит из единственной вершины, совсем не имеет ребер и циклическим не считается.

С понятиями связности и цикла также можно ознакомиться в работе [1]. Рассмотрим несколько новых классов графов.

Определение 4. Граф $F = (S, U)$ назовем *двухуровневым*, если множеством его вершин S является множество графов $L_i = (X_i, V_i) \in S$, $i = 1, 2, 3, \ldots$. При этом F будем считать графом *первого уровня*, а его вершины — графами *второго уровня*.

Определение 5. Граф $F = (S, U)$ назовем *двумерно ориентированным*, если для каждого ребра $u_i \in U$ задана последовательность векторов $V_i = \{v_{i,j} | v_{i,j} = (x_{i,j}; y_{i,j}) \in D \subset \mathbb{R}^2,\ j = 1, 2, \ldots, n;\ n \in \mathbb{N}^+\}$. При этом число n назовем *коэффициентом сегментации ребра*, а множество D — *базой ориентации*. Если $n = 1$, то ребро называется *несегментированным*. Такие ребра обозначаются указанием вершин графа, которые они соединяют, и в скобках указывается последовательность векторов: $\tilde{xy}(\{v_{xy,i}\}_{i=1}^n)$. Пара, обозначающая вершины графа, является упорядоченной, так как граф ориентированный ($\tilde{xy} \neq \tilde{yx}$). Если не требуется указывать направление, то ребро обозначается только указанием вершин. Если не требуется указания вершин, то ребро может обозначаться одной буквой.

Замечание 1 (о нормировке). Для вектора базы ориентации значимым является только его направление, а не его длина, соответственно, для базы ориентации определена норма $r : D \longrightarrow \mathbb{R}^+$ такая, что $\forall v_1, v_2 \in D(r(v_1) = r(v_2))$.

Определение 6. Двумерно ориентированный граф назовем *ориентированным по «сторонам света»*, если его базой ориентации является множество $(\{-1; 0; 1;\} \times \{-1; 0; 1;\}) \backslash \{(0, 0)\}$. «Северному» направлению соответствует вектор $(0, 1)$, «северо-восточному» — вектор $(1, 1)$ и т.д.

Замечание 2. В качестве нормы для базы ориентации ориентированного по «сторонам света» графа возьмем величину, равную максимуму по абсолютным значениям координат вектора: $r((x, y)) = \max\{|x|, |y|\}$.

Определение 7. Пусть $F = (S, U)$ — ориентированный по сторонам света граф. *Началом графа F* назовем вершину $s \in S$ такую, что $\neg\exists s_1 \in S((s_1, s) \in U)$. Иными словами, начало графа — это вершина, в которую не входит ни одно ребро.

Определение 8. Пусть $F = (S, U)$ — ориентированный по сторонам света граф. *Концом графа F* назовем вершину $e \in S$ такую, что $\neg\exists s_1 \in S((e, s_1) \in U)$. Другими словами, конец графа — это вершина, из которой не выходит ни одно ребро.

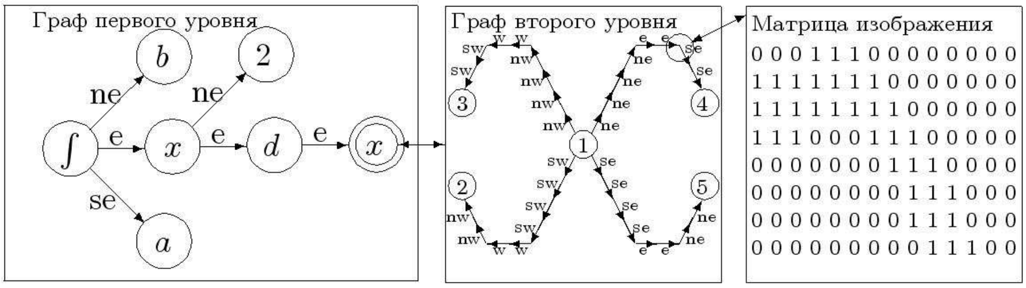

Рис. 1. Граф изображения формулы для интеграла

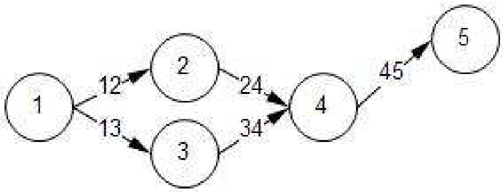

Рис. 2. Граф изображения формулы

Определение 9. *Графом изображения формулы* назовем двухуровневый ориентированный по «сторонам света» связный граф $F = (S, U)$ без циклов с несегментированными ребрами U и одним началом, где множество графов второго уровня $L_i = (X_i, V_i) \in S, i = 1, 2, \ldots, n$ также ориентировано по «сторонам света» и ребра V_i являются сегментированными. При этом направление каждого из ребер $u \in U$ обозначается через $\bar{u}$.

Приведем небольшой пример. Рассмотрим простую математическую формулу, а именно интеграл с верхним и нижним пределами:

$$\int_a^b x^2\,dx$$

Для данной формулы граф изображения приведен на рис. 1.

Определение 10. Пусть $F = (S, U)$ — граф изображения формулы, тогда через $s(F)$ будем обозначать вершину, являющуюся началом графа F, а через $E(F)$ — множество вершин, являющихся концами графа.

Определение 11. Пусть $F = (S, U)$ — граф изображения формулы. *Подграфом связности вершин $s_1 \in S$ и $s_2 \in S$ графа F* назовем граф $F' = (S', U')$ изображений формул, удовлетворяющих следующим условиям: $S' \subset S$, $U' \subset U$, s_1 — начало графа F', s_2 — конец графа F', а из вершины s_1 в вершину s_2 существует единственный маршрут, включающий все вершины и ребра графа F'. Например, для графа, изображенного на рис. 2, подграфами связности вершин 1 и 4 являются графы, изображенные на рис. 3.

Определение 12. Пусть $F = (S, U)$ — граф изображения формулы. *Множеством связности вершин $s_1 \in S$ и $s_2 \in S$ графа F (обозначается $W(F, s_1, s_2)$)* назовем множество всех подграфов связности вершин $s_1 \in S$ и $s_2 \in S$ графа F.

Замечание 3. Пусть $X = W(F, s_1, s_2)$ — множество связности вершин s_1 и s_2 графа F. Тогда выполнено следующее свойство: $\forall w \in X (W(w, s_1, s_2) = \{w\})$.

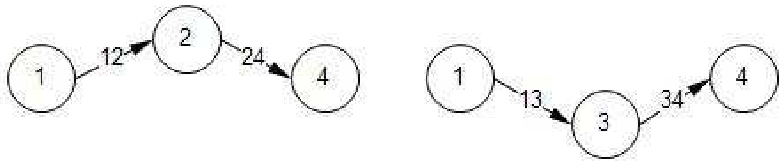

Рис. 3. Графы множества связности

Определение 13. Пусть $F = (S, U)$ — обыкновенный граф, тогда через $S(F)$ будем обозначать множество вершин графа F, а через $U(F)$ — множество его ребер.

Определение 14. Пусть $F = (S, U)$ — граф изображения формулы. Вершину $e \in S$ назовем *концом графа F по направлению v*, если $\exists w \in W(F, s(F), e)(\forall u \in U(w)\ (\bar{u} = v))$ и $\neg\exists s_1 \in S(es_1(v) \in U)$. Иными словами, конец графа по направлению v — это вершина, до которой можно добраться при движении только по v от начала графа и которая сама не содержит исходящих ребер с направлением v. Множество всех концов графа F по направлению v будем обозначать через $E_v(F)$. Например, для графа, изображенного на рис. 2, концом по направлению $(1,0)$ является вершина с номером 4, если учесть, что ребра $12, 13, 24, 34$ имеют направление $(1,0)$, а ребро 45 имеет направление $(1,1)$.

Определение 15. Пусть $F = (S, U)$ — граф изображения формулы. *Множеством правых концов графа* (обозначается $E_r(F)$) назовем множество концов графа F по направлению $(1,0)$. Например, для графа, изображенного на рис. 2, множество правых концов состоит из одной вершины с номером 4, если учесть, что ребра $12, 13, 24, 34$ имеют направление $(1,0)$, а ребро 45 имеет направление $(1,1)$.

§ 2. Отображения символов и классов формул

Рассмотрим отображения, значениями которых являются двумерно ориентированные графы.

Определение 16. Пусть L — множество классов формул. *Отображением класса формулы* назовем отображение $if : L \longrightarrow P$, где P — множество графов изображений формул.

Определение 17. Путь C — множество всех символов. *Отображением символа* назовем отображение $ic : C \longrightarrow P$, где P — множество ориентированных по «сторонам света» графов.

Замечание 4 (о несвязности графа второго уровня). Отметим, что значением ic не всегда является связный граф. Например, при применении отображения к символам ‘i’ и ‘=’ результатом будет несвязный граф.

Рассмотрим правила отображения для основных классов формул.

Пусть f_1, f_2, f_3 и f_4 формулы из произвольного класса, а $if(f_1)$, $if(f_2)$, $if(f_3)$ и $if(f_4)$ соответственно их отображения. Для данных отображений введем несколько обозначений, которые будем использовать в описываемых правилах для сокращения записи:

$$\begin{array}{llll} S_1 = S(if(f_1)), & U_1 = U(if(f_1)), & S_2 = S(if(f_2)), & U_2 = U(if(f_2)), \\ S_3 = S(if(f_3)), & U_3 = U(if(f_3)), & S_4 = S(if(f_4)), & U_4 = U(if(f_4)), \\ s_1 = s(if(f_1)), & E_{1r} = E_r(if(f_1)), & s_2 = s(if(f_2)), & E_{2r} = E_r(if(f_2)), \\ s_3 = s(if(f_3)), & E_{3r} = E_r(if(f_3)), & s_4 = s(if(f_4)), & E_{4r} = E_r(if(f_4)), \end{array}$$

$(E, s)_v = \{(e, s)v | e \in E\}$, где $F = (S, U)$ — граф изображения формулы, $s \in S$, $E \subset S$, $v \in D(F)$; $(e, s)_v = \{(e, s)v\}$, где $F = (S, U)$ — граф изображения формулы, $s, e \in S$, $v \in D(F)$.

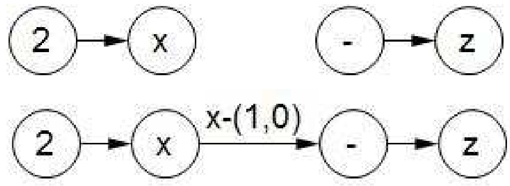

Рис. 4. Последовательность формул

I. Пусть f — формула, состоящая из одного символа c ($f = c$). Отображение формулы, состоящей из одного символа — это граф изображения формулы, состоящий из одной вершины. Этой вершиной является отображение соответствующего символа.

$$if(f) = (\{ic(c)\}, \emptyset). \tag{2.1}$$

Пример: $c =$ ‘ $\int$ ’. Тогда $if($‘ $\int$ ’$) = (\{ic($‘ $\int$ ’$)\}, \emptyset)$.

II. Пусть f — последовательность, состоящая из двух формул f_1 и f_2 ($f = f_1 f_2$). Отображение последовательности двух формул — это граф изображения формулы, содержащий все вершины и ребра отображения первой формулы, все вершины и ребра отображения второй формулы, а также дополнительные ребра, соединяющие правые концы отображения первой формулы с началом отображения второй формулы.

$$if(f) = (S_1 \bigcup S_2, U_1 \bigcup U_2 \bigcup (E_{1r}, s_2)_{(1,0)}). \tag{2.2}$$

Пример: $f_1 =$ “$2x$”, $f_2 =$ “$-z$”. $f =$ “$2x - z$”.

На рис. 4 приведены графы изображения формулы для $f_1 =$ “$2x$”, $f_2 =$ “$-z$” и $f =$ “$2x-z$”. Правым концом f_1 является вершина ‘x’, а началом f_2 — вершина ‘ $-$ ’. Следовательно, при построении графа изображения формулы для f кроме ребер из отображения f_1 и f_2 должно быть добавлено одно дополнительное ребро, которое на рисунке обозначено через $x-(1,0)$, что означает ребро, соединяющее вершину ‘x’ с вершиной ‘ $-$ ’ по направлению $(1,0)$.

$$if(\text{“}2x - z\text{”}) = (\{ic(\text{‘}2\text{’}), ic(\text{‘}x\text{’}), ic(\text{‘}-\text{’}), ic(\text{‘}z\text{’})\}, \{2x(1,0), -z(1,0), x-(1,0)\}).$$

III. Пусть f — функция возведения в степень ($f = {f_1}^{f_2}$). Отображение функции возведения в степень — это граф изображения формулы, содержащий все вершины и ребра исходного отображения (f_1), все вершины и ребра отображения показателя степени (f_2), а также дополнительное ребро, соединяющее правый конец исходного отображения с началом отображения показателя степени. Отметим, что рассматриваются только такие классы формул, которые имеют единственный правый конец. Также исключаются классы формул, содержащие в себе взятие индекса у степенной функции (это формулы вида ${X^2}_2$), и формулы, у которых показатель степени имеет сдвиг вправо по отношению к нижнему индексу (${X_2}^2$). Соответственно, рассматриваются только такие формулы, у которых нижний индекс и показатель степени находятся на одном уровне (X_2^2).

$$if(f) = (S_1 \bigcup S_2, U_1 \bigcup U_2 \bigcup (e_1, s_2)_{(1,1)}). \tag{2.3}$$

Пример: $f_1 =$ “2”, $f_2 =$ “x”. $f =$ “2^x”.

На рис. 5 приведены графы изображения формулы для $f_1 =$ “2”, $f_2 =$ “x” и $f =$ “2^x”. Правым концом f_1 является вершина ‘2’, а началом f_2 — вершина ‘x’. Следовательно, при построении графа изображения формулы для f, кроме ребер из отображения f_1 и f_2, должно быть добавлено одно дополнительное ребро, которое на рисунке обозначено через $2x(1,1)$, что означает ребро, соединяющее вершину ‘2’ с вершиной ‘x’ по направлению $(1,1)$.

$if(\text{“}2^x\text{”}) = (\{ic(\text{‘}2\text{’}), ic(\text{‘}x\text{’})\}, \{2x(1,1)\})$.

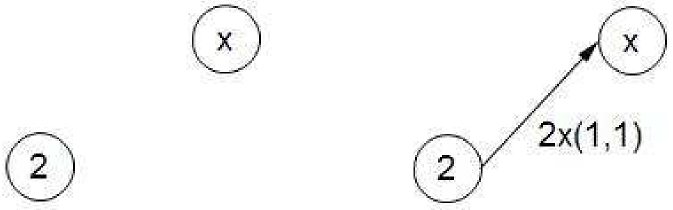

Рис. 5. Функция возведения в степень

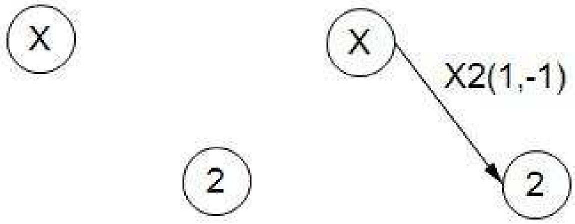

Рис. 6. Функция взятия индекса

IV. Пусть f — функция взятия индекса ($f = f_{1_{f_2}}$). Отображение функции взятия индекса — это граф изображения формулы, содержащий все вершины и ребра исходного отображения (f_1), все вершины и ребра отображения показателя индекса (f_2), а также дополнительное ребро, соединяющее правый конец исходного отображения с началом отображения показателя индекса. Отметим, что рассматриваются только такие классы формул, которые имеют единственный правый конец.

$$if(f) = (S_1 \bigcup S_2, U_1 \bigcup U_2 \bigcup (e_1, s_2)_{(1,-1)}). \tag{2.4}$$

Пример: $f_1 =$ “X”, $f_2 =$ “2”. $f =$ “X_2”.

На рис. 6 приведены графы изображения формулы для $f_1 =$ “X”, $f_2 =$ “2” и $f =$ “X_2”. Правым концом f_1 является вершина ‘X’, а началом f_2 — вершина ‘2’. Следовательно, при построении графа изображения формулы для f кроме ребер из отображения f_1 и f_2 должно быть добавлено одно дополнительное ребро, которое на рисунке обозначено через $X2(1, -1)$, что означает ребро, соединяющее вершину ‘X’ с вершиной ‘2’ по направлению $(1, -1)$.

$if(\text{“}X_2\text{”}) = (\{ic(\text{‘}X\text{’}), ic(\text{‘}2\text{’})\}, \{X2(1, -1)\})$.

V. Пусть f — интеграл с верхним и нижним пределами ($f = \displaystyle\int_{f_1}^{f_2} f_3$). Отображение интеграла — это граф изображения формулы, содержащий все вершины и ребра отображения подынтегрального выражения, все вершины и ребра отображений нижнего и верхнего пределов, отображение символа интегрирования, а также три дополнительных ребра. Эти ребра соединяют отображение символа интегрирования соответственно с началом отображения подынтегрального выражения, с началом отображения нижнего предела и с началом отображения верхнего предела.

$$\begin{aligned} if(f) = \Big(&S_1 \bigcup S_2 \bigcup S_3 \bigcup \{int\}, \\ &U_1 \bigcup U_2 \bigcup U_3 \bigcup (int, s_3)_{(1,0)} \bigcup (int, s_1)_{(0,-1)} \bigcup (int, s_2)_{(0,1)}\Big), \end{aligned} \tag{2.5}$$

где $int = ic(\text{‘}\int\text{’})$.

Пример: $f_1 =$ “0”, $f_2 =$ “1”, $f_3 =$ “$x\,dx$”. $f =$ “$\displaystyle\int_0^1 x\,dx$”.

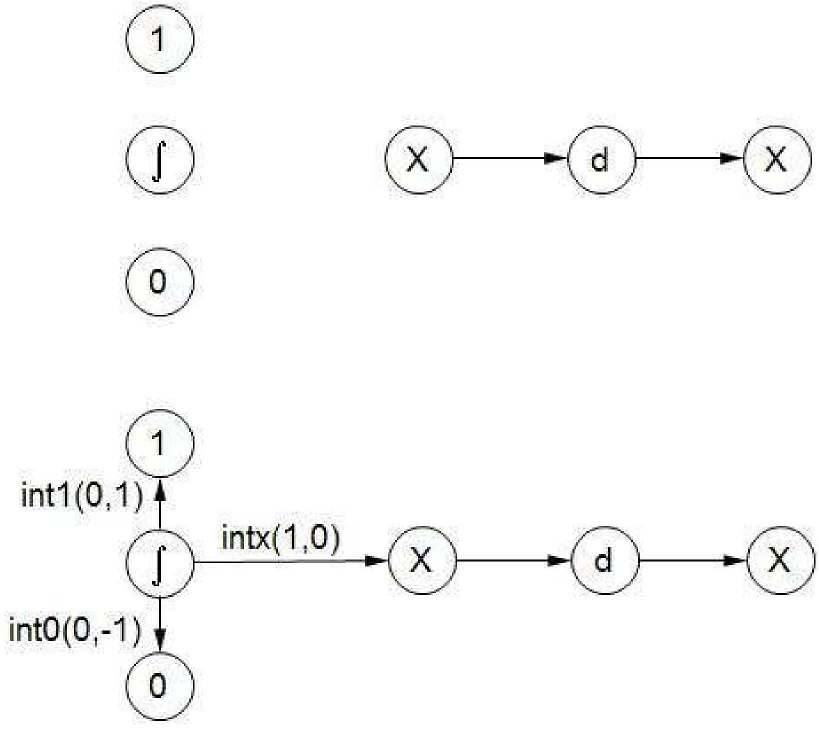

Рис. 7. Интеграл

На рис. 7 приведены графы изображения формулы для $f_1 =$ “0”, $f_2 =$ “1”, $f_3 =$ “$x\ dx$”, “$\int$” и $f =$ “$\displaystyle\int_0^1 x\,dx$”. Началом f_1 является вершина ‘0’, началом f_2 — вершина ‘1’, а началом f_3 — первая из вершин ‘x’ Следовательно, при построении графа изображения формулы для f кроме ребер из отображений f_1, f_2 и f_3 должны быть добавлены три дополнительных ребра, которые на рисунке обозначены через $intx(1,0)$ (ребро, соединяющее вершину ‘$\int$’ с вершиной ‘x’ по направлению $(1,0)$), $int0(0,-1)$ (ребро, соединяющее вершину ‘$\int$’ с вершиной ‘0’ по направлению $(0,-1)$) и $int1(0,1)$ (ребро, соединяющее вершину ‘$\int$’ с вершиной ‘1’ по направлению $(0,1)$).

$$if(\text{“}\int_0^1 x\,dx\text{”}) = (\{ic(\text{‘}0\text{’}), ic(\text{‘}1\text{’}), ic(\text{‘}\int\text{’}), ic(\text{‘}x\text{’}), ic(\text{‘}d\text{’}), ic(\text{‘}x\text{’})\},$$
$$\{int0(0,-1), int1(0,1), intx(1,0), xd(1,0), dx(1,0)\}).$$

VI. Пусть f — функция извлечения корня ($f = \sqrt[f_2]{f_1}$). Отображение функции извлечения корня — это граф изображения формулы, содержащий все вершины и ребра отображения подкоренного выражения, все вершины и ребра отображения показателя корня, отображение символа извлечения корня, а также двух дополнительных ребер, соединяющих отображение символа извлечения корня с началом отображения подкоренного выражения и с началом отображения показателя корня.

$$if(f) = (S_1 \bigcup S_2 \bigcup \{ic(\text{‘}\surd\text{’})\}, U_1 \bigcup U_2 \bigcup (ic(\text{‘}\surd\text{’}), s_1)_{(0,-1)}) \bigcup (ic(\text{‘}\surd\text{’}), s_2)_{(0,1)}) \tag{2.6}$$

Пример: $f_1 =$ “y”, $f_2 =$ “3”. $f =$ “$\sqrt[3]{y}$”.

На рис. 8 приведены графы изображения формулы для $f_1 =$ “y”, “$\surd$”, $f_2 =$ “3” и $f =$ “$\sqrt[3]{y}$”. Началом f_1 является вершина ‘y’, а началом f_2 — вершина ‘3’. Следовательно, при построении графа изображения формулы для f кроме ребер из отображений f_1 и f_2 должно быть добавлено два дополнительных ребра, которые на рисунке обозначены через $sqrty(0,-1)$ и $sqrt3(0,1)$ соответственно.

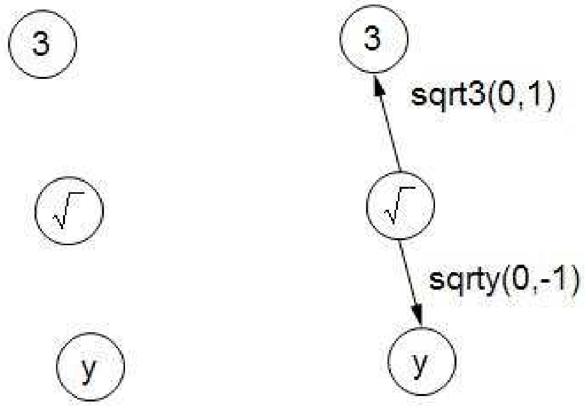

Рис. 8. Извлечение корня

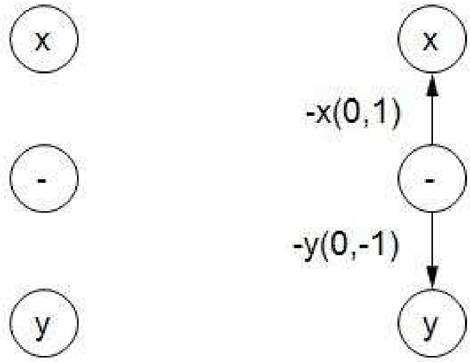

Рис. 9. Дробь

$if(``\sqrt[3]{y}") = (\{ic(`\surd'), ic(`y'), ic(`3')\}, \{sqrty(0,-1), sqrt3(0,1)\})$.

VII. Пусть f — дробь ($f = \frac{f_1}{f_2}$). Отображение дроби — это граф изображения формулы, содержащий все вершины и ребра отображения числителя, все вершины и ребра отображения знаменателя, отображение символа дробной черты, а также двух дополнительных ребер, соединяющих отображение символа дробной черты соответственно с началом отображения числителя и с началом отображения знаменателя.

$$if(f) = (S_1 \bigcup S_2 \bigcup \{frac\}, U_1 \bigcup U_2 \bigcup (frac, s_1)_{(0,1)} \bigcup (frac, s_2)_{(0,-1)}), \tag{2.7}$$

где $frac = ic(`-')$.

Пример: $f_1 = ``x"$, $f_2 = ``y"$. $f = ``\frac{x}{y}"$.

На рис. 9 приведены графы изображения формулы для $f_1 = ``x"$, $f_2 = ``y"$, $``-"$ и $f = ``\frac{x}{y}"$. Началом f_1 является вершина $`x'$, началом $f_2$ — вершина $`y'$. Следовательно, при построении графа изображения формулы для f кроме ребер из отображений f_1 и f_2 должны быть добавлены два дополнительных ребра, которые на рисунке обозначены через $-x(0,1)$ и $-y(0,-1)$.

$if(``\frac{x}{y}") = (\{ic(`-'), ic(`x'), ic(`y')\}, \{-x(0,1), -y(0,-1)\})$.

§ 3. Семантические сети

Определение 18. *Семантической сетью* (*деревом*) назовем информационную модель предметной области, представленную ориентированным графом $G = (S, V)$, вершины S которого соответствуют объектам предметной области P, а рёбра V задают отношения R между ними.

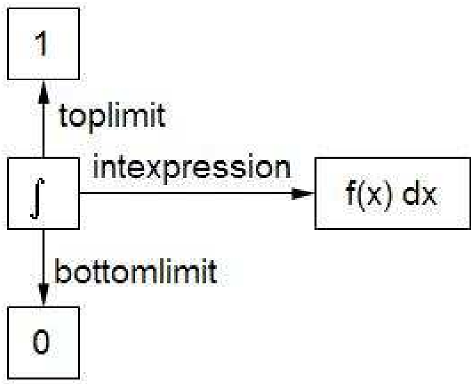

Рис. 10. Семантическое дерево для интеграла

Определение 19. *Однородной семантической сетью* назовем семантическую сеть, в которой все вершины имеют одинаковый набор типов отношений. *Неоднородной семантической сетью* назовем семантическую сеть, в которой вершины имеют различный набор типов отношений. Другими словами, можно сказать, что семантическая сеть является однородной, если все вершины имеют одинаковую структуру и неоднородной — в противном случае.

Рассмотрим небольшой пример. Пусть дана семантическая сеть, вершинами которой являются люди, находящиеся в университете. Если в качестве отношений взять отношение «быть знакомым», то данная сеть является однородной, так как каждому человеку доступно свойство «быть знакомым» с тем или иным человеком. Если же людей в университете рассматривать как студентов и преподавателей, то предложенная семантическая сеть становится неоднородной, так как в этом случае отношения могут быть двух типов («быть учеником» и «быть учителем»), и доступность этих свойств зависит от того, является ли человек студентом или преподавателем.

Более подробно с понятием семантическая сеть можно ознакомиться в работе [2].

Определение 20. *Семантическим деревом формулы* назовем связную семантическую сеть без циклов $F = (S, U)$, в котором предметной областью P является множество классов математических формул, а множеством отношений R — множество связей между формулами.

Рассмотрим пример. Для формулы $\int_0^1 f(x)\,dx$ семантическое дерево изображено на рис. 10. Как видно из рисунка, формула интеграла содержит исходящие ребра, описывающие отношения трех типов: нижний предел (bottomlimit), верхний предел (toplimit) и подынтегральное выражение (intexpression).

Замечание 5. Семантическое дерево формулы является неоднородным.

Определение 21. Пусть L — множество классов формул. *Семантическим отображением класса формулы* назовем отображение $sf : L \longrightarrow F$, где F — множество семантических деревьев формул.

§ 4. Задача распознавания математических формул

Рассмотрим задачу распознавания простой рукописной математической формулы. Пусть имеется сканированное изображение рукописного текста. Требуется распознать содержащуюся на изображении формулу.

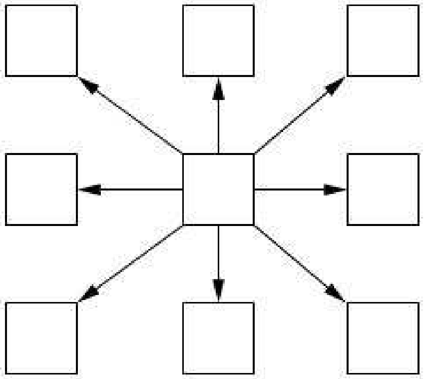

Рис. 11. Ограничение

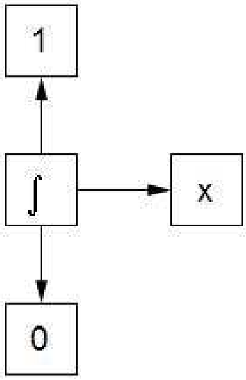

Рис. 12. Пример

Будем считать, что все содержащиеся в тексте символы уже распознаны по отдельности и уже построено дерево изображения для исходной формулы. Для простоты ограничим сложность структуры исходного дерева. Будем рассматривать только такие деревья, у которых вершины имеют не более одного исходящего ребра в каждом направлении базы ориентации. Схематически ограничивающая структура приведена на рис. 11.

Пример. Вершина дерева изображения формулы, содержащая символ интегрирования в выражении $\int_0^1 x\,dx$, может быть схематически представлена так, как изображена на рис. 12. В этом случае вершина имеет только 3 исходящих ребра с разными направлениями (вверх, вниз, вправо).

Требуется по имеющемуся дереву изображения формулы восстановить соответствующее семантическое дерево формулы.

Пусть имеется некоторое множество L классов формул и для него описаны некоторые правила для отображений $if : L \longrightarrow P$ и $sf : L \longrightarrow F$, где P — множество графов изображений формул, F — множество семантических деревьев формулы. Необходимо по заданному графу p изображения формулы получить соответствующее семантическое дерево формулы f. Иными словами, требуется найти отображение if^{-1}, так как из наличия двух отображений $if^{-1} : P \longrightarrow L$ и $sf : L \longrightarrow F$ исходная задача считается решенной.

§ 5. Описание решения

Пусть $p = (S, U) \in P$ — исходный граф изображения формулы. Рассмотрим все возможные структуры данного графа и приведем правила для отображения $if^{-1}(p)$.

Для упрощения записей будем использовать те же сокращения, которые были использованы в § 2.

При описании каждого правила отображения будем рассматривать примеры и для сокращения записи в примерах введем некоторые сокращения. С помощью символа, включенного в одинарные кавычки, будем обозначать отображение соответствующего символа (например, $`a' = ic(`a')$). Это же обозначение будем использовать для указания множества вершин графа, если множество вершин содержит только один элемент и этим элементом является отображение соответствующего символа (например, $(\{ic(`a')\}, \emptyset) = (`a', \emptyset)$); а также, если множество вершин представлено объединением одноэлементных множеств (например, $(\{ic(`a')\}\bigcup\{ic(`b')\}, S) = (`a'\bigcup`b', S)$). Для обозначения ребер будем использовать символы соответствующих вершин с указанием направления ребра в качестве нижнего индекса (например, $(`a', `b')_{(1,0)} = ab_{(1,0)}$). Аналогично с вершинами, это же обозначение будем использовать для указания множества ребер (например, $\big(`a'\bigcup`b', \{(`a', `b')_{(1,0)}\}\big) = \big(`a'\bigcup`b', ab_{(1,0)}\big)$ и $\big(`a'\bigcup`b'\bigcup`c', \{(`a', `b')_{(1,0)}, (`b', `c')_{(1,0)}\}\big) = \big(`a'\bigcup`b'\bigcup`c', ab_{(1,0)}\bigcup bc_{(1,0)}\big)$). С помощью c будем обозначать формулу, состоящую из одного символа; множество вершин графа, содержащего одну вершину соответствующую отображению c; а также сам граф, состоящий из одной вершины c. Отметим также, что через S_i и U_j будем обозначать множество вершин и множество ребер графа, причем (S_i, U_j) является графом изображения формулы тогда и только тогда, когда $i = j$.

1)

$$if^{-1}\left((ic(c), \emptyset)\right) = c, \tag{5.1}$$

где c — формула, состоящая из одного символа.

Пример. Пусть $p = (`x', \emptyset)$. Тогда $if^{-1}(p) = “x”$.

2)

$$if^{-1}\left((c\bigcup S_2, U_2\bigcup(c, s_2)_{(1,1)})\right) = c^{if^{-1}((S_2,U_2))}, \tag{5.2}$$

где c — начало графа p, а s_2 — начало графа (S_2, U_2).

Пример. Пусть $p = (`2'\bigcup`x', 2x_{(1,1)})$. По описанному правилу получаем соответствия: $c = `2'$, $S_2 = `x'$, $U_2 = \emptyset$, $(c, s_2)_{(1,1)} = 2x_{(1,1)}$. Тогда искомое отображение примет вид: $if^{-1}(p) = if^{-1}(`2', \emptyset)^{if^{-1}(`x',\emptyset)} = 2^x$.

3)

$$if^{-1}\left((c\bigcup S_2, U_2\bigcup(c, s_2)_{(1,-1)})\right) = c_{if^{-1}((S_2,U_2))}, \tag{5.3}$$

где c — начало графа p, а s_2 — начало графа (S_2, U_2).

Пример. Пусть $p = (`x'\bigcup`2', x2_{(1,-1)})$. По описанному правилу получаем соответствия: $c = `x'$, $S_2 = `2'$, $U_2 = \emptyset$, $(c, s_2)_{(1,-1)} = x2_{(1,-1)}$. Тогда искомое отображение примет вид: $if^{-1}(p) = if^{-1}(`x', \emptyset)^{if^{-1}(`2',\emptyset)} = x_2$.

4)

$$if^{-1}\left(S_1\bigcup S_2, U_1\bigcup U_2\bigcup(E_{1r}, s_2)_{(1,0)})\right) = if^{-1}((S_1, U_1))if^{-1}((S_2, U_2)), \tag{5.4}$$

где s_2 — начало графа (S_2, U_2), а E_{1r} — множество правых концов графа (S_1, U_1).

Пример. Пусть $p = (`2'\bigcup`e'\bigcup`x', 2e_{(1,0)}\bigcup ex_{(1,1)})$. По описанному правилу получаем соответствия: $S_1 = `2'$, $S_2 = `e'\bigcup`x'$, $U_1 = \emptyset$, $U_2 = ex_{(1,1)}$, $(E_{1r}, s_2)_{(1,0)} = 2e_{(1,0)}$. Тогда искомое отображение примет вид: $if^{-1}(p) = if^{-1}((`2', \emptyset))if^{-1}((`e'\bigcup`x', ex_{(1,1)}))$. Далее, применяя правила 2) и 1), получаем $if^{-1}(p) = if^{-1}((`2', \emptyset))if^{-1}((`e', \emptyset))^{if^{-1}((`x',\emptyset))} = 2e^x$.

5)

$$if^{-1}\Bigg(\Big(S_1\bigcup S_2\bigcup S_3\bigcup `\int', U_1\bigcup U_2\bigcup U_3\bigcup \Big(`\int', s_3\Big)_{(1,0)}\bigcup\Big(`\int', s_1\Big)_{(0,-1)}\bigcup\Big(`\int', s_2\Big)_{(0,1)}\Big)\Bigg) = \int_{if^{-1}((S_1,U_1))}^{if^{-1}((S_2,U_2))} if^{-1}((S_3,U_3)), \quad (5.5)$$

где '$\int$' — начало графа p, s_1 — начало графа (S_1,U_1), s_2 — начало графа (S_2,U_2), s_3 — начало графа (S_3,U_3).

Пример. Пусть $p = (`\int'\bigcup`0'\bigcup`1'\bigcup`d'\bigcup`x', \int d_{(1,0)}\bigcup\int 0_{(0,-1)})\bigcup\int 1_{(0,1)}\bigcup dx_{(1,0)})$. По описанному правилу получаем соответствия: $S_1 = `0'$, $S_2 = `1'$, $S_3 = `d'\bigcup`x'$, $U_1 = \emptyset$, $U_2 = \emptyset$, $U_3 = dx_{(1,0)}$, $\big(`\int', s_3\big)_{(1,0)} = \int d_{(1,0)}$, $\big(`\int', s_1\big)_{(0,-1)} = \int 0_{(0,-1)}$, $\big(`\int', s_2\big)_{(0,1)} = \int 1_{(0,1)}$. Тогда искомое отображение примет вид:

$$if^{-1}(p) = \int_{if^{-1}((`0',\emptyset))}^{if^{-1}((`1',\emptyset))} if^{-1}((`d'\bigcup`x', dx_{(1,0)})).$$

Далее, применяя правило 4), получаем $if^{-1}(p) = \int_0^1 if^{-1}((`d',\emptyset))if^{-1}((`x',\emptyset)) = \int_0^1 dx.$

6)

$$if^{-1}\Big((S_1\bigcup S_2\bigcup `\surd', U_1\bigcup U_2\bigcup(`\surd', s_1)_{(0,-1)}\bigcup(`\surd', s_2)_{(0,1)})\Big) = \\ = \sqrt[if^{-1}((S_2,U_2))]{if^{-1}((S_1,U_1))}, \quad (5.6)$$

где '$\surd$' — начало графа p, s_1 — начало графа (S_1,U_1), s_2 — начало графа (S_2,U_2).

Пример. Пусть $p = (`\surd'\bigcup`1'\bigcup`6', \surd 1_{(0,-1)}\bigcup 16_{(1,0)})$. По описанному правилу получаем соответствия: $S_1 = `1'\bigcup`6'$, $S_2 = \emptyset$, $U_1 = 16_{(1,0)}$, $U_2 = \emptyset$, $(`\surd', s_1)_{(0,-1)} = \surd 1_{(0,-1)}$, $(`\surd', s_2)_{(0,1)} = \emptyset$. Тогда искомое отображение примет вид: $if^{-1}(p) = \sqrt{if^{-1}((`1'\bigcup`6', 16_{(1,0)}))}$. Далее, применяя правило 4), получаем $if^{-1}(p) = \sqrt{if^{-1}((`1',\emptyset))if^{-1}((`6',\emptyset))} = \sqrt{16}$.

7)

$$if^{-1}\Big((S_1\bigcup S_2\bigcup `-', U_1\bigcup U_2\bigcup(`-', s_1)_{(0,1)}\bigcup(`-', s_2)_{(0,-1)})\Big) = \frac{if^{-1}((S_1,U_1))}{if^{-1}((S_2,U_2))}, \quad (5.7)$$

где '$-$' — начало графа p, s_1 — начало графа (S_1,U_1), s_2 — начало графа (S_2,U_2).

Пример. Пусть $p = (`-'\bigcup`1'\bigcup`2', -1_{(0,1)}\bigcup -2_{(0,-1)})$. По описанному правилу получаем соответствия: $S_1 = `1'$, $S_2 = `2'$, $U_1 = \emptyset$, $U_2 = \emptyset$, $(`-', s_1)_{(0,1)} = -1_{(0,1)}$, $(`-', s_2)_{(0,-1)} = -2_{(0,-1)}$. Тогда искомое отображение примет вид: $if^{-1}(p) = \dfrac{if^{-1}((`1',\emptyset))}{if^{-1}((`2',\emptyset))} = \dfrac{1}{2}$.

Докажем, что отображение if^{-1} является обратным к отображению if. Для этого докажем следующее равенство: $if(if^{-1}(p)) = p$.

1) Пусть $p = (ic(c),\emptyset)$. По правилу отображения if^{-1} получаем $if^{-1}((ic(c),\emptyset)) = c$. Далее, применяя правило I для отображения if, получаем $if(c) = (ic(c),\emptyset)$.

2) Пусть $p = (c\bigcup S_2, U_2\bigcup(c,s_2)_{(1,1)})$, c — начало графа p, а s_2 — начало графа (S_2,U_2). По правилу отображения if^{-1} получаем $if^{-1}(p) = c^{if^{-1}((S_2,U_2))}$. Далее, применяя правило III для отображения if, получаем

$$if(c^{if^{-1}((S_2,U_2))}) = (c\bigcup S_2, \emptyset\bigcup U_2\bigcup(c,s_2)_{(1,1)}) = (c\bigcup S_2, U_2\bigcup(c,s_2)_{(1,1)}).$$

3) Пусть $p = (c\bigcup S_2, U_2\bigcup(c,s_2)_{(1,-1)})$, c — начало графа p, а s_2 — начало графа (S_2,U_2). По правилу отображения if^{-1} получаем $if^{-1}(p) = c_{if^{-1}((S_2,U_2))}$. Далее, применяя правило IV для отображения if, получаем

$$if(c_{if^{-1}((S_2,U_2))}) = (c\bigcup S_2, \emptyset\bigcup U_2\bigcup(c,s_2)_{(1,-1)}) = (c\bigcup S_2, U_2\bigcup(c,s_2)_{(1,-1)}).$$

4) Пусть $p = (S_1 \bigcup S_2, U_1 \bigcup U_2 \bigcup (E_{1r}, s_2)_{(1,0)})$, s_2 — начало графа (S_2, U_2), а E_{1r} — множество правых концов графа (S_1, U_1). По правилу отображения if^{-1} получаем

$$if^{-1}(p) = if^{-1}((S_1, U_1))if^{-1}((S_2, U_2)).$$

Далее, применяя правило II для отображения if, получаем

$$if(if^{-1}((S_1, U_1))if^{-1}((S_2, U_2))) = (S_1 \bigcup S_2, U_1 \bigcup U_2 \bigcup (E_{1r}, s_2)_{(1,0)}).$$

5) Пусть

$$p = \Big(S_1 \bigcup S_2 \bigcup S_3 \bigcup `\int', U_1 \bigcup U_2 \bigcup U_3 \bigcup \Big(`\int', s_3\Big)_{(1,0)} \bigcup \Big(`\int', s_1\Big)_{(0,-1)} \bigcup \Big(`\int', s_2\Big)_{(0,1)}\Big),$$

'$\int$' — начало графа p, s_1 — начало графа (S_1, U_1), s_2 — начало графа (S_2, U_2), s_3 — начало графа (S_3, U_3). По правилу отображения if^{-1} получаем

$$if^{-1}(p) = \int_{if^{-1}((S_1,U_1))}^{if^{-1}((S_2,U_2))} if^{-1}((S_3, U_3)).$$

Далее, применяя правило V для отображения if, получаем

$$if\Big(\int_{if^{-1}((S_1,U_1))}^{if^{-1}((S_2,U_2))} if^{-1}((S_3, U_3))\Big) =$$
$$\Big(S_1 \bigcup S_2 \bigcup S_3 \bigcup `\int', U_1 \bigcup U_2 \bigcup U_3 \bigcup \Big(`\int', s_3\Big)_{(1,0)} \bigcup \Big(`\int', s_1\Big)_{(0,-1)} \bigcup \Big(`\int', s_2\Big)_{(0,1)}\Big).$$

6) Пусть $p = (S_1 \bigcup S_2 \bigcup `\surd', U_1 \bigcup U_2 \bigcup (`\surd', s_1)_{(0,-1)} \bigcup (`\surd', s_2)_{(0,1)})$, где '$\surd$' — начало графа p, s_1 — начало графа (S_1, U_1), s_2 — начало графа (S_2, U_2) . По правилу отображения if^{-1} получаем $if^{-1}(p) = \sqrt[if^{-1}((S_2,U_2))]{if^{-1}((S_1, U_1))}$. Далее, применяя правило VI для отображения if, получаем

$$if(\sqrt[if^{-1}((S_2,U_2))]{if^{-1}((S_1, U_1))}) = (S_1 \bigcup S_2 \bigcup `\surd', U_1 \bigcup U_2 \bigcup (`\surd', s_1)_{(0,-1)} \bigcup (`\surd', s_2)_{(0,1)}).$$

7) Пусть $p = (S_1 \bigcup S_2 \bigcup `-', U_1 \bigcup U_2 \bigcup (`-', s_1)_{(0,1)} \bigcup (`-', s_2)_{(0,-1)})$, '$-$' — начало графа p, s_1 — начало графа (S_1, U_1), s_2 — начало графа (S_2, U_2). По правилу отображения if^{-1} получаем $if^{-1}(p) = \frac{if^{-1}((S_1,U_1))}{if^{-1}((S_2,U_2))}$. Далее, применяя правило VII для отображения if, получаем

$$if(\frac{if^{-1}((S_1, U_1))}{if^{-1}((S_2, U_2))}) = (S_1 \bigcup S_2 \bigcup `-', U_1 \bigcup U_2 \bigcup (`-', s_1)_{(0,1)} \bigcup (`-', s_2)_{(0,-1)}).$$

Обратим внимание на дополнительные ограничения в пунктах 2), 3) и 5)–7). Если соответствующие условия не выполнены, то разбор графа изображения формулы выполняется по правилу 4); считаем, что в 4) дополнительные условия выполнены всегда, так как были введены ограничения на сложность формулы.

Доказано равенство $if(if^{-1}(p)) = p$ для любых рассматриваемых графов изображений формул p. Следовательно, if^{-1} является обратным к отображению if, что и требовалось доказать.

Заключение

Рассматривая такие понятия, как двумерно ориентированный граф и граф изображения формулы, мы можем сильно упростить анализ текстов со сложными структурами. Предложенные математические структуры дают полную информацию о взаимном расположении отдельных частей изображения, а также о связях внутри этих частей.

Также были рассмотрены специальные отображения, которые ставят в соответствие каждой математической формуле определенный граф изображения формулы и благодаря которым становится возможным сопоставление изображений с шаблонами формул и соответственно дальнейшее их распознавание.

Отметим, что в работе рассматривались математические формулы с ограниченной сложностью, а именно формулы, содержащие верхние и нижние индексы, дроби, интегралы, корни. Для расширения разнообразия формул достаточно добавить новые правила аналогично существующим. Не были рассмотрены такие конструкции, которые содержат элементы с недревовидной структурой. К таким элементам можно отнести системы, матрицы, различные диаграммы. Для их моделирования требуется рассматривать понятие двумерно ориентированных графов в более широком смысле для возможности описания более сложных связей.

Также были введены общие ограничения на число разнообразных связей между объектами формулы. Ограничение заключалось в том, что каждый символ имеет соседей не более чем в восьми направлениях на плоскости, что в большинстве случаев является достаточным. Для распознавания более сложных формул, помимо графов, ориентированных по сторонам света, необходимо вводить понятия специализированных графов с более точной двумерной ориентацией.

Предложенный механизм может быть использован не только в распознавании текстов с математическими формулами, но и в других задачах, требующих анализа структуры изображения, в которых основным критерием является взаимное расположение частей. По аналогии с графом изображения формулы в различных задачах могут быть использованы подобные структуры, основанные на базовых понятиях: двухуровневый граф и двумерно ориентированный граф.

СПИСОК ЛИТЕРАТУРЫ

1. Зыков А. А. Основы теории графов. М.: Наука, 1986. 381 с.
2. Рассел С., Норвиг П. Искусственный интеллект: современный подход. 2-е изд. / пер. с англ. М.: Вильямс, 2006. 1407 с.

Поступила в редакцию 19.12.2012

Сапаров Алексей Юрьевич, аспирант, кафедра теоретических основ информатики, Удмуртский государственный университет, 426034, Россия, г. Ижевск, ул. Университетская, 1.
E-mail: say.saplh@gmail.com

Бельтюков Анатолий Петрович, профессор, д.ф.-м.н., заведующий кафедрой теоретических основ информатики, Удмуртский государственный университет, 426034, Россия, г. Ижевск, ул. Университетская, 1. E-mail: belt@uni.udm.ru

A. Yu. Saparov, A. P. Beltyukov
Mathematical modeling of formula images for their recognition

Keywords: two-level graph, two-dimensional oriented graph, graph of mathematical expression image, mathematical formula recognition.

Mathematical Subject Classifications: 03D20, 05C20

The work is devoted to the use of the basic elements of graph theory to solve the mathematical formula recognition problem. We introduce the concepts of two-level graphs and two-dimensional oriented graphs that make it possible to describe complex images consisting of the hierarchy of parts with a particular relative position. We consider a special function that builds a two-dimensional oriented graph from a mathematical formula; the graph is called a graph of mathematical expression image. The mapping rules for basic classes of mathematical formulae are presented. We describe a problem-solving procedure for a recognition problem, which is based on the reverse problem of constructing of a mathematical-expression image graph.

REFERENCES

1. Zykov A.A. *Osnovy teorii grafov* (Foundations of the graphs theory), Moscow: Nauka, 1986, 381 p.
2. Russel S.J., Norvig P. *Iskusstvennyi intellekt: sovremennyi podkhod* (Artificial intelligence a modern approach. Second edition), Moscow: Williams, 2006, 1407 p.

Received 19.12.2012

Saparov Aleksei Yur'evich, post-graduate student, Department of Theoretical Foundations of Computer Science, Udmurt State University, ul. Universitetskaya, 1, Izhevsk, 426034, Russia.
E-mail: say.saplh@gmail.com

Beltyukov Anatolii Petrovich, Professor, Doctor of Physics and Mathematics, Head of the Department of Theoretical Foundations of Computer Science, Udmurt State University, ul. Universitetskaya, 1, Izhevsk, 426034, Russia.
E-mail: belt@uni.udm.ru