УДК 621.396.664

Е.Т. РАМАЗАНОВ

О ФАЗИФИКАЦИИ РАСПОЗНАЮЩИХ АЛГОРИТМОВ ТИПА ВЫЧИСЛЕНИЯ ОЦЕНОК И ПОСТРОЕНИИ КОРРЕКТНОГО АЛГОРИТМА

(Представлена академиком НАН РК К.А. Касымовым)

Предложен способ фазификации решающего правила алгоритма решения задачи сегментации программ для модели распознающих операторов в основе которых лежит операторная теория распознающих алгоритмов Журавлева. Фазификация исходной модели распознающих алгоритмов позволяет сформулировать и обосновать утверждение о корректности фаззи алгебры алгоритмов над множеством Ω - регулярных задач сегментации.

В работе Журавлёва [1] отмечена возможность фазификации класса алгоритмов вычисления оценок, а именно, интерпретацией оценок G_{ij}, элементов матрицы: $G = \left\| G_{ij} \right\|$, как значений функции принадлежности[2]. В работе [1] построена модель распознающих операторов (алгоритмов) над множеством алгоритмов вычисления оценок и определены условия корректности операторной алгебры над множеством задач распознавания. В [3,4] на основе работы [1] построена модель алгоритмов для решения задачи сегментации программ и определены условия корректности операторной алгебры над множеством Ω-регулярных задач сегментации, что позволяет дать аналитическое описание корректного алгоритма для каждой такой задачи. Напомним, что под задачей сегментации программ обычно принято понимать задачу разбивки программы на части, например, блоки и распределение этих блоков страницам(сегментам) виртуальной памяти программы с целью минимизировать критерий качества разбивки. В качестве такого критерия может выступать число страничных отказов, а если говорить более точно, то или теоретическое среднее, или экспериментальное среднее число страничных отказов, вычисленное по k прогонам программы. Часто считается, что начальная разбивка блоков программы по страницам (сегментам) виртуальной памяти задано заранее (естественная компоновка программы). Здесь будем считать, что число блоков программы равно n, а число сегментов (страниц) виртуальной памяти равно m, при этом каждый блок b_i программы и каждый сегмент S_j программы имеют свою длину- l_i и V_j соответственно, $i = 1,2,...,n$ $j = 1,2,...,m$; Согласно работе [1], а также работам [2,3] введём некоторые определения. Итак, пусть $P_j(x)$ - предикат $: P_j(x) \equiv "x \in S_j"$, где x-произвольный объект(блок программы), а S_j -сегмент(класс), $j = 1,2,...,m$; Пусть далее $\left\| \alpha_{ij} \right\|_{n \times m}$ матрица $: \alpha_{ij} = P_j(x_i)$, $j = 1,2,...,m$; $i = 1,2,...,n$.

Основываясь на работах[1, 3], рассмотрим модель M_0 алгоритмов типа вычисления оценок для сегментации блоков $b_1, b_2, ..., b_n$ по сегментам $S_1, S_2, ..., S_m$ виртуальной памяти, при этом считаем, что во время выполнения программы ни длина блоков $l_1, l_2, ..., l_n$, ни длина сегментов $V_1, V_2, ..., V_m$ не изменяются . Итак.

1) Зададим систему опорных множеств $\Omega = \{\omega\}$. Пусть $\{A_\alpha\}$-совокупность множеств активности программы[8], тогда в качестве $\omega \in \Omega$ будем использовать $A_\alpha \in \{A_\alpha\}$, т.е. $\Omega \subset \{A_\alpha\}$. Следуя работе [3], определим также функцию близости $B_\omega^\varepsilon(b_i, b_j)$. Пусть b_i, b_j - блоки, ω - опорное множество; $r_\omega(b_i, b_j)$ -число взаимных ссылок между блоками b_i, b_j в течение

суммарного интервала жизни множества ω; параметр ε выбирается из промежутка $0 \le \varepsilon \le 1$. Определим $B_\omega^\varepsilon(b_i, b_j)$:

$$B_\omega^\varepsilon(b_i, b_j) = \begin{cases} 1, \text{ если } r_\omega(b_i, b_j) \ge \varepsilon \cdot r_{ij} \\ 0, \text{ в противном случае} \end{cases} \quad (1)$$

где r_{ij} - число взаимных ссылок между блоками b_i и b_j, элемент матрицы близости, $r = [r_{ij}]_{n \times n}$.

Вес $p(\omega)$ опорного множества ω определим как функцию от τ_ω-времени жизни множества ω и длин сегментов $i_1, i_2, ..., i_{|\omega|}$, составляющих ω.

$$p(\omega) = \frac{\tau_\omega}{(V_{i_1} + V_{i_2} + ... + V_{i_{|\omega|}})} \quad (2)$$

Определим вес γ_j блока b_j:

$$\gamma_j = \frac{(l_1 + l_2 + ... + l_n)}{l_j} \text{ для всех}$$

$j = 1,2,...,n$ (3)

близость между блоками b_i, b_j по множеству ω:

$$\beta_{ij}^{(\omega)} = \frac{r_\omega(b_i, b_j)}{r_{ij}}, \quad r_{ij} \ne 0 \quad (4)$$

Пусть блоки $b_{r_1}, b_{r_2}, ..., b_{r_{|S_r|}}$ входят в сегмент S_r, тогда оценку $G_r(b_i)$ блока b_i за сегмент S_r согласно [3]:

$$G_r(b_i) = x_r \cdot \sum_{j \in S_r} \gamma_j \cdot \sum_{\omega \in \Omega} p(\omega) \cdot \beta_{ij}^{(\omega)} \cdot B_\omega^\varepsilon(b_i, b_j) \quad (5)$$

где $x_r \in \{0, 1\}$.

Следует заметить, что если длина сегментов $V_1 = V_2 = ... = V_m$, то имеем последовательность страниц $S_1, S_2, ..., S_m$.

Здесь интересно обратить внимание также и на более общую модель[3], алгоритмов вычисления оценок, чем модель M_0. Обобщение модели M_0, обозначим его через M_1, отличается от самой модели M_0 способом вычисления оценки блока b_i за сегмент S_r:

$$G_r(b_i) = x_r^1 \cdot G_r^1(b_i) + x_r^0 \cdot G_r^0(b_i) \quad (6)$$

где $x_r^0, x_r^1 \in \{-1, 0, 1\}$, $r = 1,2,...,m$, $i = 1,2,...,n$;

$$G_r^1(b_i) = \sum_{j \in S_r} \gamma_j \cdot \sum_{\omega \in \Omega} p(\omega) \cdot \beta_{ij}^{(\omega)} \cdot B_\omega^\varepsilon(b_i, b_j) \quad (7)$$

$$G_r^0(b_i) = \sum_{j \in CS_r} \gamma_j \cdot \sum_{\omega \in \Omega} p(\omega) \cdot \overline{\beta}_{ij}^{(\omega)} \cdot \overline{B}_\omega^\varepsilon(b_i, b_j),$$

$$CS_r = X^n \setminus S_r \quad (8)$$

$$\overline{\beta}_{ij}^{(\omega)} = \frac{r(\neg b_i, b_j) + r(b_i, \neg b_j)}{r_\omega(b_i, b_j)},$$

$$r_\omega(b_i, b_j) \ne 0 \quad (9)$$

где $r(\neg b_i, b_j)$-суммарное число взаимных ссылок между блоками не b_i и b_j. И соответственно $r(b_i, \neg b_j)$- суммарное число взаимных ссылок между b_i и не b_j.

$$\overline{B}_\omega^\varepsilon(b_i, b_j) = 1 - B_\omega^\varepsilon(b_i, b_j) \quad (10)$$

Отметим, что в приведенных моделях алгоритмов типа вычисления оценок в качестве решающего правила можно взять правило C^* (11).

$$C^*(Z) = \left\|\delta_{ij}\right\|, \quad \delta_{ij} = \begin{cases} \delta_{ij} = 1, \; if \;\; G_{ij} > c_{2j} \\ \delta_{ij} = 0, \; if \;\; G_{ij} < c_{1j} \\ \Delta, \; if \;\; c_{1j} \le G_{ij} \le c_{2j} \end{cases} \quad (11)$$

где G_{ij} - оценка отношения i объекта j классу. Пусть параметры $c_1 = \min_j c_{1j}, c_2 = \max_j c_{2j}$ В записи (11) учитываем, что решающее правило C^* по задаче Z может вычислять матрицу, обозначим её через $\delta = \left\|\delta_{ij}\right\|$, вообще говоря, не обязательно совпадающую с информационной матрицей задачи Z, т.е. $\left\|\alpha_{ij}\right\|_{n\times m}$. Кроме того, дополнительное значение Δ – означает неопределенность распознавания.

Напомним, что распознающий алгоритм А является корректным для задачи Z , если алгоритм по этой задаче вычисляет информационную матрицу $\left\|\alpha_{ij}\right\|_{n\times m}$ этой задачи. Исследование вопроса о корректности алгоритма А для задачи Z нужно начинать с построения алгебры $\vartheta(M)$ над исходным множеством параметрических алгоритмов $M=\{A_p \mid p=(p_1,p_2,...,p_m)\}$ распознавания. Такой операторный подход был предложен Журавлевым Ю. И. Формально, алгебру над множеством алгоритмов $M=\{A_p \mid p=(p_1,p_2,...,p_m)\}$ строят введением операций:

$$\lambda\cdot A(Z)=\lambda\cdot\varphi=\left\|\lambda\cdot\varphi_{ij}\right\|,\quad \lambda=const \tag{12}$$

$$(A'+A'')(Z)=A'(Z)+A''(Z)=$$
$$=\left\|\varphi'_{ij}\right\|+\left\|\varphi''_{ij}\right\|=\left\|\varphi'_{ij}+\varphi''_{ij}\right\| \tag{13}$$

$$(A'\cdot A'')(Z)=A'(Z)\cdot A''(Z)=$$
$$=\left\|\varphi'_{ij}\right\|\cdot\left\|\varphi''_{ij}\right\|=\left\|\varphi'_{ij}\cdot\varphi''_{ij}\right\| \tag{14}$$

где $\varphi=\left\|\varphi_{ij}\right\|$ - вещественная матрица оценок принадлежности блоков $b_1,b_2,...,b_n$ сегментам $S_1,S_2,...,S_m$ вычисленная алгоритмом A.

Будем считать, что в операторной алгебре, построенной над множеством распознающих алгоритмов типа вычисления оценок, каждый алгоритм А представляет собой суперпозицию $A\circ C$, т.е. $\mathsf{A}=A\circ C^*$, где A распознающий оператор, вычисляющий матрицу оценок $G=\left\|G_{ij}\right\|_{n\times m}$, а C^* корректное решающее правило, вычисляющее по матрице $G=\left\|G_{ij}\right\|_{n\times m}$, матрицу $\delta=\left\|\delta_{ij}\right\|$, задачи Z.

Далее будем считать, что для системы Ω подмножеств множества $\{1,2,...,m\}$ выполняется условие:

$$\forall\omega_\alpha(\omega_\alpha\in\Omega\Rightarrow\omega_\alpha\not\subset\bigcup_{\beta\neq\alpha}\omega_\beta) \tag{15}$$

Согласно определению, если выполняются условия (15), то система множеств Ω называется базой множества $\{1,2,...,m\}$. Заметим, что системы $\Omega_0=\{\{1\},\{2\},...,\{m\}\}$, $\Omega_1=\{1,2,...,m\}$ являются базами множества $\{1,2,...,m\}$.

Важно отметить, что задача сегментации $Z=<I,X^n>$, где I -начальная информация, а $X^n=\{b_1,b_2,...,b_n\}$ множество блоков называется Ω - регулярной, если :

1. $S_j\neq\varnothing$, $j=1,2,...,m$;
2. $S_i\cap S_j=\varnothing$, $i\neq j$;
3. для любых $b_i,b_j\in X^n$ и $\forall\omega\in\Omega$: $r_\omega(b_i,b_j)\neq\max\limits_{i,j} r_{ij}$;
4. I является Ω -существенной для задачи Z, т.е. для всякой пары блоков $b_i,b_j\in X^n$ найдется блок $b\in X^n$ ($b\neq b_i$, $b\neq b_j$) и множество $\omega\in\Omega$ такие, что $r_\omega(b_i,b)\neq r_\omega(b_j,b)$.

Корректный алгоритм A^* для задачи сегментации, точнее для задачи Z из множества Ω-регулярных задач сегментации представи'м[3]:

$$\mathsf{A}^*=\left((c_1+c_2)\sum_{i=1}^{q}\sum_{j=1}^{l}\alpha_{ij}\cdot B^k(i,j)\right)\circ C^* \tag{16}$$

где $\left\|\alpha_{ij}\right\|_{n\times m}$ - информационная матрица задачи Z, а C^* корректное решающее правило(11).

Параметры $c_1=\min\limits_j c_{1j}$, $c_2=\max\limits_j c_{2j}$, . В (16) величина k есть степень замыкания. Для степени k имеет место формула

$$k = \left[\frac{\ln q + \ln l + |\ln(c_1 + c_2)| - |\ln(c_1)|}{|\ln a|} \right] + 1 \quad ;$$

где $a < 1$, $a = \max\limits_{(r,t)\neq(i,j)} |G_{rt}(i,j)|$ (17)

Параметры c_1, c_2, таковы, что $c_2 > c_1 > \sqrt{a}$, $c_1 + c_2 > 1$, а k есть степень замыкания операторной алгебры, алгебры построенной, согласно правилам (12)-(14), над множеством исходных операторов. Напомним, что значение Δ – означает неопределенность распознавания.

Фазифицируем исходную модель M_1 алгоритмов типа вычисления оценок, заменив в M_1 решающее правило $C^*(Z)$. При этом новый вид решающего правила в модели $\tilde{M}_1$ ($\tilde{M}_1$-модификация модели M_1)алгоритмов для задачи сегментации программ, может быть и решающим правилом задачи распознавания. Следует отметить, что принципиальная возможность применения распознающего алгоритма для решения задачи сегментации программ была показана в работе [3]. Итак, вместо решающего правила (11) введем решающее правило $C_s^*(Z)$:

$$C_s^*(Z) = \delta_{ij} =$$

$$= \begin{cases} \delta_{ij} = 1, \ if \ \ G_{ij} > c_{2j} \\ \delta_{ij} = 0, \ if \ \ G_{ij} < c_{1j} \\ \delta_{ij} = \dfrac{1}{1+e^{-G_{ij}}} \ , if \ c_{1j} \le G_{ij} \le c_{2j} \end{cases} \tag{18}$$

которое в определенной степени позволяет более точно отразить ситуацию с распознаванием объектов, а именно детализируя «неопределенность»- Δ в формуле (11). Итак, как уже сказано, для фаззи- модели алгоритмов используем обозначение $\tilde{M}_1$.

В формуле (18) участвует сигмоидальная функция $f(x) = \dfrac{1}{1+e^{-x}}$. На основе сигмоидальной функции, кроме (18,) представляет интерес и решающее правило (19), которое на первый взгляд дает другую картину фазификации исходной модели:

$$\bar{C}_s^*(Z) = \left\| \delta_{ij} \right\|_{nxm} =$$

$$= \begin{cases} \delta_{ij} = 1, \quad if \ h_{2j} < \dfrac{1}{1+e^{-G_{ij}}} \\ \delta_{ij} = 0, \quad if \ h_{1j} > \dfrac{1}{1+e^{-G_{ij}}} \\ \dfrac{1}{1+e^{-G_{ij}}}, \ \ if \ h_{1j} \le G_{ij} \le h_{2j} \end{cases} \tag{19}$$

где h_{1j}, h_{2j} параметры решающего правила $\bar{C}_s^*$

Однако заметим, что, по-существу, решающее правило (19) сводится к правилу (18).

Пусть $\vartheta^k(\tilde{M}_1)$- операторная алгебра степени k, построенная по правилам (12)-(14) над фаззи-моделью алгоритмов $\tilde{M}_1$. Заметим, что опираясь на доказательства теоремы 5.1[3] и теоремы 1[6], нетрудно доказать следующее утверждение:

Теорема 1. Операторная алгебра $\vartheta^k(\tilde{M}_1)$ степени k, построенная над фаззи-моделью операторов $\tilde{M}_1$, является корректной над множеством Ω- регулярных задач сегментации. Если задача сегментации Z является Ω-регулярной задачей, то фаззи-алгоритм A_s^* представленный в виде

$$A_s^* = \left((c_1 + c_2) \sum_{i=1}^{q} \sum_{j=1}^{l} \alpha_{ij} \cdot B^k(i,j) \right) \circ \bar{C}_s^*$$

является корректным для Z, где величина k вычисляется по формуле (17).

ЛИТЕРАТУРА

1. *Журавлев Ю.И.* Корректные алгебры над множествами некорректных (эвристических) алгоритмов // Проблемы кибернетики. Вып.3.-М., 1978. С. 5-65.

2. *Заде Л.А.* Тени нечетких множеств// Проблемы передачи информации. 1966, том II, вып.1. С. 37 - 44.

3. *Дюсембаев А.Е.* Математические модели сегментации программ.- М.: Физматлит, серия «Библиотечка программиста»2001г.,208с.

4. *Дюсембаев А.Е.* Об одном подходе к задаче сегментации программ // Доклады РАН. 1993. Т. 329. №3. С. 712-714.

5. *Журавлев Ю.И.* Корректные алгебры над множествами некорректных(эвристических) алгоритмов // Кибернетика., 1978. №2. С. 35-46.

6. *Пападимитриу Х., Стайглиц К.* Комбинаторная оптимизация. М.: Мир, 1985.365с.

7. *Плотников А. Д.* Некоторые свойства экстремальных комбинаторных задач// Математические и программные методы проектирования информационных и управляющих систем. Тезисы докладов к зональной конференции 28-29 май1990. Пенза.С.14-15.

8. *Феррари Д.* Оценка производительности вычислительных систем. М.:Наука, 1974. 545с.

9. *Мэтьюз Джон. Г., Курис Д. Финк.* Численные методы.Киев.: Вильямс, 2001.713с.

Резюме

Программаларды сегменттеу есебінің шешу алгоритмының Журавлевтың операторлық тәсілін қолдану арқылы шешу ережесін фазификациялаудың жолы және бағаларды есептеу алгоритмының моделі қарастырылған. Алғашқы бейнелерді тану алгоритмдердің моделін фазификациялау қортындысында Ω - регулярлы сегменттеу есептер жиынында алгоритмдердің алгебрасы корректі болуы туралы тұжырымдар келтірілген.

Summary

In article the way is offered of fuzzyfication a solving rule of algorithm of the decision of a problem of segmentation of programs for model of distinguishing operators in which basis the theory of distinguishing algorithms of Zhuravlyov lies. Fuzzyfication initial model of distinguishing algorithms allows to formulate and prove the statement about a correctness fuzzy-algebras of algorithms over set - regular problems of segmentation.

Международная Академия Бизнеса *Поступила 16.05.09*