Вестник КРАУНЦ. Физ.-мат. науки. 2016. № 4-1(16). С. 107-111. ISSN 2079-6641

DOI: 10.18454/2079-6641-2016-16-4-1-107-111

УДК 519.7

ПРОЦЕДУРА НАПРАВЛЕННОГО ПОИСКА КОРРЕКТНЫХ ОПЕРАЦИЙ НАД АЛГОРИТМАМИ *

Л. А. Лютикова, Е. В. Шматова

Институт прикладной математики и автоматизации, 36000, г. Нальчик, ул. Шортанова, 89а

E-mail: lylarisa@yandex.ru, lenavsh@yandex.ru

В данной работе рассматривается логический подход к теоретическому обоснованию построения корректных алгоритмов, расширяющих область получаемых решений на базе существующих алгоритмов. Предложенный метод позволяет на основе заданного множества алгоритмов распознавания выявить дополнительные знания заданной предметной области и построить минимальное правило, обеспечивающее дообучение данных алгоритмов.

Ключевые слова: алгоритм, корректная операция, база знаний, предметная область, решающие правила

MSC 68T27

THE PROCEDURE DIRECTED SEARCH FOR SPECIFIC OPERATIONS ON ALGORITHMS

L. A. Lyutikova, E. V. Shmatova

Institute of Applied Mathematics and Automation, 360000, KBR, Nalchik, st. Shortanova 89a, Russia

E-mail: lylarisa@yandex.ru, lenavsh@yandex.ru

This paper considers the logical approach to the theoretical justification of constructing correct algorithms that extend the area of the solutions obtained on the basis of existing algorithms. The proposed method makes it possible on the basis of a given set of pattern recognition algorithms to identify additional knowledge of a given subject area and to build a minimum rule, providing additional training of these algorithms.

Key words: algorithm, correct operation, knowledge base, subject area, decision rules.

*Работа выполнена при поддержке гранта РФФИ № 15-01-03381-а

Введение

Методы распознавания образов и восстановления зависимостей по неполным, неточным и разнородным данным используются при создании интеллектуальных информационных и аналитических систем в самых разных прикладных областях. Предпосылкой для их применения являются сбор и первичная обработка данных, в результате которых формируются описания объектов, ситуаций или явлений предметной области. Вслед за этим возникают задачи классификации объектов восстановления неизвестных значений некоторых их свойств прогнозирования их состояний, и т. д. Во всех этих случаях требуется строить алгоритмы, преобразующие исходную (начальную) информацию об объектах в выходную (финальную)информацию об этих же объектах. Как правило, в задачах такого класса наличие некоторой зависимости между начальными и финальными информациями представляется несомненным, однако она может оказаться сложной и неявной, а предметная область, недостаточно формализованной, чтобы построить её адекватную модель. В таких случаях зависимость восстанавливают по прецедентам [1, 2]. А алгоритмы обрабатывающие данную область представляют собой сложную многопараметрическую модель, которая не может быть снабжена четкими численными методам, что приводит к определенному числу некорректных решений. Однако в некоторых случаях подобные алгоритмы все же справляются с достаточным объемом данных и не могут быть проигнорированы как непригодные. В результате возникло развитие теории корректирующих операций [3, 4], синтеза корректных алгоритмов минимальной сложности, решение вопросов об их устойчивости с помощью математических методов. И появляется необходимость синтеза и корректировки с целью улучшения работы каждого алгоритма в отдельности и их совокупности.

Постановка задачи

Будем считать, что на приведенной предметной область, которая представлена в виде набора объектов и соответствующих ему признаков проработало ряд алгоритмов. Результат обработки данных каждым алгоритмом оценивается k-значной функцией $a_j(X_i,y_i)$. Можно предположить из опыта, что алгоритмы будут ошибаться на каких-то данных. Где-то ошибаться несколько алгоритмов, где-то только один, что является хорошим случаем, потому что его можно заменить на более успешный, если такой есть в нашем наборе. Задача синтеза алгоритмов состоит в построение на основе заданных наиболее успешного алгоритма, или построение более надежного алгоритма из менее надежных.

Пусть $\{A_1,A_2,\dots,A_n\}$ - множество алгоритмов,

$X=\{x_1,x_2,\dots,x_m\}$, $x_i\in\{0,1,\dots,k_r-1\}$, $k_r\in[2,\dots,N],N\in Z$ – обрабатываемые входные данные $X_i=\{x_1(y_i),x_2(y_i),\dots,x_m(y_i)\}$, $i=1,\dots,l$,

$y_i\in Y, Y=\{y_1,y_2,\dots,y_l\}$ – выходные данные;

$a_j(X_i,y_i)\in\{0,...,k-1\}$; $i=1,2,\dots,l$; $j=1,2,\dots,n$ – оценка работы алгоритма определяемое формулой

$$a_j(y_i)=\begin{cases}k-1, & A_j(X_i)=y_i\\ 0, & A_j(X_i)\neq y_i\end{cases},\ i=1,2,\dots,l,\ j=1,2,\dots,n$$

$k-1$ – алгоритм A_j распознал объект y_i по заданным признакам X_i,

$k-2$– алгоритм A_j распознал объект y_i вместе с другим объектам, 0 - алгоритм A_j не распознал объект y_i по заданным признакам X_i.

Входные, выходные данные, а также оценку работы алгоритмов представим следующим образом:

$$\begin{pmatrix} x_1(y_1) & x_2(y_1) & ... & x_m(y_1) \\ x_1(y_2) & x_2(y_2) & ... & x_m(y_2) \\ ... & .... & ... & ... \\ x_1(y_l) & x_2(y_l) & ... & x_m(y_l) \end{pmatrix} \to \begin{pmatrix} y_1 \\ y_2 \\ ... \\ y_l \end{pmatrix} \sim \begin{pmatrix} a_1(y_1) & a_2(y_1) & ... & a_m(y_1) \\ a_1(y_2) & a_2(y_2) & ... & a_m(y_2) \\ ... & .... & ... & ... \\ a_1(y_l) & a_2(y_l) & ... & a_m(y_l) \end{pmatrix}$$

$$A_i' = \{a_i(y_1), a_i(y_2)), \dots, a_i(y_l)\}, i = 1, 2, \dots, n$$

.

Будем считать, что может сложиться такая ситуация когда

$$\exists y_i \in Y | A_1(X_i) \neq y_i, A_2(X_i) \neq y_i, \dots, A_n(X_i) \neq y_i, i = 1, 2, \dots, l,$$

или ситуация когда

$$\exists y_i \in Y | A_1(X_i) = y_i \&_{j=1}^{r} y_i, A_2(X_i) = y_i \&_{j=1}^{r} y_i, ..., A_n = y_i \&_{j=1}^{r} y_i, i = 1, 2, ..., l; j \leqslant l - 1.$$

Необходимо создать алгоритм корректор такой, что

$$A_{n+1}(X_i) | A_{n+1}(X_i) = y_i \text{ и } A_{n+1}(X) | A_{n+1}(X) = Y.$$

Основные операции и определения

Пусть B — непустое множество, над элементами которого определены операции: отрицание (унарная операция), конъюнкция (бинарная), дизъюнкция (бинарная), а также константы $0, 1, ..., k-1$.

$$\begin{gathered} 0 \ \& \ X = 0, \\ (k-1) \vee \ X = (k-1), \\ 1 \ \& \ X = X, \\ 0 \vee X = X. \\ x^j \ \& \ x^k = \begin{cases} x^j, & j = k \\ 0, & j \neq k \end{cases}, \\ x_i^j = \begin{cases} j, & x_i = j \\ 0, & x_i \neq j \end{cases} \end{gathered}$$

Обобщенной инверсией является следующее выражение [5, 6]:

$$\overline{x^j} = x^0 \vee x^1 \vee \dots \vee x^{j-1} \vee x^{j+1} \vee \dots \vee x^{k-1}$$

Дизъюнкция

$$X \vee Y = \max\left[\frac{X}{k_i - 1}; \frac{Y}{k_i - 1}\right] * l, \text{ где } l = \begin{cases} k_i - 1, & \frac{X}{k_i - 1} < \frac{Y}{k_j - 1} \\ k_j - 1, & \text{иначе} \end{cases}$$

Конъюнкция

$$X \ \& \ Y = \min\left[\frac{X}{k_i - 1}; \frac{Y}{k_j - 1}\right] * l, \text{ где } l = \begin{cases} k_i - 1, & \frac{X}{k_i - 1} < \frac{Y}{k_j - 1} \\ k_j - 1, & \text{иначе} \end{cases}$$

Импликация: $X \to Y = \overline{X} \vee Y$

Процедура направленного поиска корректных операций над алгоритмами

Для осуществления логического анализа результатов работы данных алгоритмов и выявление скрытых закономерностей (неявных знаний) предлагается следующий подход:

Шаг 1: $A'_{n+1} = max(A'_1, \ldots, A'_n)$ - находим лучший алгоритм для распознавания данного образа на заданном наборе данных. В итоге получим алгоритм $_max$, который на каждом наборе данных выбирает лучший алгоритм из заданных для их обработки.

Шаг 2: Разобьём A_{max} на A^0_{max},A^1_{max},...,A^{k-1}_{max} соответственно множества нераспознанных объектов, распознанных в большом числе других, распознанным в малом числе других объектов, распознанных абсолютно точно. Если из множества всех объектов вычесть множество объектов, которое выдано A^0_{max}, затем A^1_{max} и так до A^{k-1}_{max}, т.е. $Y \setminus A^0_{max} \setminus A^1_{max} \setminus \ldots \setminus A^{k-1}_{max}$ получим нераспознанные алгоритмами объекты. Если после данных операций остается один объект его записываем во множество A^{k-1}_{max} и возвращаемся к второму шагу.

Шаг 3: Для разделения объектов в каждом из множеств проводим следующую процедуру: для каждой пара объектов из рассматриваемого множества заводим вектор значений: $Z_{i,j} = \{z_1, z_2, \ldots, z_m\}$, где $\forall$ пары y_i и y_j,

$$z_p = \begin{cases} 0, & x_p(y_i) = x_p(y_j) \\ 1, & x_p(y_i) \neq x_p(y_j) \end{cases}, p \in [1, \ldots, m],$$

из значения $x_p(y_i)$ или $x_p(y_j)$ на которых $z_p = 1$ будут составляется решающие правила для до обучения существующих алгоритмов.

Пример

x_1	x_2	x_3	Y	A'_1	A'_2	A'_3	A'_{max}	A_{max}
0	2	1	a	0	0	0	0	cd
1	0	1	b	1	2	1	2	b
2	1	2	c	2	2	1	2	c
0	1	1	d	1	1	1	1	deb
2	0	1	e	1	1	1	1	dec

Шаг 1:

$$A_{max} = \{A^0_{max} = \{c,d\}; A^1_{max} = \{d,e,b,d,e,c\}; A^2_{max} = \{b,c\}\}.$$

Шаг 2:

$$Y \setminus A^0_{max} \setminus A^1_{max} \setminus A^2_{max};$$
$$\{a,b,c,d,e\} \setminus \{c,d\} \setminus \{b,c\} \setminus \{d,e,b,d,e,c\} = a;$$
$$A^1_{max} \setminus A^2_{max} = \{d,e,b,d,e,c\} \setminus \{b,c\} = \{de : de\}.$$

Шаг 3: $d(0,1,1); e(2,0,1); z = (1,1,0)$ следовательно дообучающее правило будет состоять из двух переменных и одного объекта: $x^0_1 x^1_2 d$. То есть общий алгоритм будет выглядеть как:

$$A^2_{max} \vee a \vee x^0_1 x^1_2 d \vee A^1_{max} \setminus d.$$

Т.е для получения корректного алгоритма достаточно дообучить множество заданных алгоритмов одним правилом.

Заключение

В работе продемонстрирована возможность выявления дополнительных знаний заданной предметной области и построение системы минимальных правил, обеспечивающих дообучение данных алгоритмов.

Список литературы/References

[1] Журавлёв Ю. И., "Об алгебраическом подходе к решению задач распознавания или классификации", *Проблемы кибернетики*, **33** (1978), 5–68, [Zhuravlev Yu. I. Ob algebraicheskom podkhode k resheniyu zadach raspoznavaniya ili klassifikatsii. Problemy kibernetiki. 1978. no 33. pp. 5–68].

[2] Воронцов К. В., "Оптимизационные методы линейной и монотонной коррекции в алгебраическом подходе к проблеме распознавания", *Журнал вычислительной математики и математической физики*, **40**:1 (2000), 166–176, [Vorontsov K. V. Optimizatsionnye metody lineynoy i monotonnoy korrektsii v algebraicheskom podkhode k probleme raspoznavaniya. Zhurnal vychislitel'noy matematiki i matematicheskoy fiziki. 2000. vol. 40. no 1. pp. 166–176 (in Russian)].

[3] Шибзухов З.М., "О некоторых конструктивных и корректных классах алгебраических ΣΠ-алгоритмов", *Доклады РАН*, **432**:4 (2010), 43–55, [Shibzukhov Z. M. O nekotorykh konstruktivnykh i korrektnykh klassakh algebraicheskikh ΣΠ-algoritmov. Doklady RAN. 2010. vol. 432. no 4. pp.43–55 (in Russian)].

[4] Shibzukhov Z. M., "Correct Aggregation Operations with Algorithms", *Pattern Recognition and Image Analysis*, **24**:3 (2014), 377–382.

[5] Лютикова Л.А., *Моделирование и минимизация баз знаний в терминах многозначной логики предикатов*, Препринт–Нальчик, НИИ ПМА КБНЦ РАН, 2006, 33 с., [Lyutikova L. A. Modelirovanie i minimizatsiya baz znaniy v terminakh mnogoznachnoy logiki predikatov. Preprint–Nal'chik: NII PMA KBNTs RAN, 2006. 33 p. (in Russian)].

[6] Лютикова Л. А., "Логический подход к модели представления знаний", *Естественные и технические науки*, 2014, №6(74), 107–108, [Lyutikova L. A. Logicheskiy podkhod k modeli predstavleniya znaniy. Estestvennye i tekhnicheskie nauki. 2014. no 6(74). pp. 107–108 (in Russian)].

Список литературы (ГОСТ)

[1] Журавлёв Ю. И. Об алгебраическом подходе к решению задач распознавания или классификации // Проблемы кибернетики. 1978. № 33. С. 5–68

[2] Воронцов К. В. Оптимизационные методы линейной и монотонной коррекции в алгебраическом подходе к проблеме распознавания // Журнал вычислительной математики и математической физики. 2000. Т. 40. №1. С. 166–176

[3] Шибзухов З. М. О некоторых конструктивных и корректных классах алгебраических ΣΠ-алгоритмов // Доклады РАН. 2010. Т. 432. №4. С.43–55

[4] Shibzukhov Z. M. Correct Aggregation Operations with Algorithms // Pattern Recognition and Image Analysis. 2014. vol. 24. no 3. pp. 377–382

[5] Лютикова Л. А. Моделирование и минимизация баз знаний в терминах многозначной логики предикатов. Препринт–Нальчик: НИИ ПМА КБНЦ РАН,2006. 33 с.

[6] Лютикова Л. А. Логический подход к модели представления знаний //Естественные и технические науки. 2014. №6(74). С. 107–108

Для цитирования: Лютикова Л. А., Шматова Е. В. Процедура направленного поиска корректных операций над алгоритмами // *Вестник КРАУНЦ. Физ.-мат. науки*. 2016. № 4-1(16). C. 107-111. DOI: 10.18454/2079-6641-2016-16-4-1-107-111

For citation: Lyutikova L. A., Shmatova E. V. The procedure directed search for specific operations on algorithms, *Vestnik KRAUNC. Fiz.-mat. nauki*. 2016, **16**: 4-1, 107-111. DOI: 10.18454/2079-6641-2016-16-4-1-107-111

Поступила в редакцию / Original article submitted: 09.11.2016