

Т.М. КОСОВСКАЯ
**ПОДХОД К РЕШЕНИЮ ЗАДАЧИ ПОСТРОЕНИЯ
МНОГОУРОВНЕВОГО ОПИСАНИЯ КЛАССОВ НА ЯЗЫКЕ
ИСЧИСЛЕНИЯ ПРЕДИКАТОВ**

Косовская Т.М. Подход к решению задачи построения многоуровневого описания классов на языке исчисления предикатов.

Аннотация. Рассматривается задача построения многоуровневого описания классов, объекты которых характеризуются свойствами своих элементов и отношениями между ними. Задачи распознавания и анализа таких объектов являются NP-трудными, но при наличии достаточно коротких и часто встречающихся подформул в описаниях классов можно построить многоуровневое описание классов, существенно понижающее значение показателя степени в оценках числа шагов алгоритмов, решающих эти задачи. До сих пор выделение таких подформул оставлялось на усмотрение разработчика системы распознавания. В работе предлагается подход к их автоматическому выделению.

Ключевые слова: искусственный интеллект, предикатные формулы исчисления, сложность алгоритма, NP-полнота, уровневое описание классов.

Kosovskaya T.M. An approach to the construction of a level description of classes by means of a predicate calculus language.

Abstract. A problem of construction of a level description of classes with objects characterized by properties of their elements and relations between them is under consideration in the paper. The problems of recognition and analysis of such objects are NP-hard, but if descriptions of classes contain short enough and frequently occurred sub-formulas then it is possible to build a level description of classes essentially decreasing an exponent in upper bounds of steps for an algorithm solving the problem. Usually an extracting of these sub-formulas is leaved to the investigator will. An approach to their automatic extraction is proposed in the paper.

Keywords: artificial intelligence, predicate calculus formulas, complexity of algorithm, NP-completeness, level description of classes.

1. Введение. В работе рассматриваются такие задачи искусственного интеллекта, как распознавание образов и прогнозирование, допускающие формализацию средствами исчисления предикатов при условии, что объект рассматривается как совокупность своих элементов. Такой подход к решению сформулированных задач назван логико-предметным. Для рассматриваемых задач в [3, 6] доказано, что они являются NP-трудными.

Для уменьшения числа шагов их решения в [4] предложено создание уровневого описания классов, заключающееся в выделении «часто» встречающихся среди исходных описаний классов подформул «небольшой сложности». Термины «часто» и «небольшой сложности» уточняются в зависимости от алгоритма, с помощью которого решается поставленная задача. В частности, в этих работах доказаны условия уменьшения числа шагов при использовании уровневого описания

классов для переборного алгоритма и для алгоритмов, основанных на построении вывода в исчислении предикатов. Примеры, приведённые в [4], иллюстрируют существенное уменьшение числа шагов распознавания контурного изображения на сложной сцене при введении двухуровневого описания, основанного на «здравом смысле».

Ниже предлагается подход к решению задачи построения уровневого описания классов, базирующийся на применении понятия неполной выводимости предикатной формулы, введённого в [5] для решения задач распознавания объектов с неполной информацией при логико-предметном подходе.

Рассматривается модельный пример построения трёхуровневого описания классов по имеющемуся исходному описанию.

2. Общая постановка задачи. Пусть исследуемый объект представлен как множество своих элементов $\omega = \{\omega_1, \dots, \omega_t\}$. На ω задан набор предикатов p_1, \dots, p_n , характеризующих свойства элементов ω и отношения между ними. Логическим описанием $S(\omega)$ объекта ω называется множество всех атомарных формул или их отрицаний, истинных на ω . Множество всех объектов разбито на классы $\Omega = \cup_{k=1}^K \Omega_k$. Логическим описанием класса называется формула $A_k(\bar{x})$, заданная в виде дизъюнкции элементарных конъюнкций, такая что если $A_k(\bar{\omega})$ истинна, то $\omega \in \Omega_k$.¹

С помощью построенных описаний объектов и классов предлагается решать следующие задачи.

Задача идентификации. Проверить, удовлетворяет ли объект ω или его часть описанию класса $A_k(\bar{x})$ и предъявить эту часть объекта.

Задача классификации. Найти все такие номера k , что верна формула $A_k(\bar{\omega})$.

Задача анализа. Найти и классифицировать все части τ объекта ω , для которых $A_k(\bar{\tau})$.

Решение задач идентификации, классификации и анализа для распознавания сложного объекта сведено в к доказательству соответ-

¹Здесь и далее посредством \bar{x} обозначается список элементов конечного множества x , соответствующий некоторой перестановке номеров его элементов. Тот факт, что элементами списка \bar{x} являются элементы множества y , будем записывать в виде $x \subseteq y$.

ственно логических следований²:

$$S(\omega) \Rightarrow \exists \bar{x}_{\neq} A_k(\bar{x}), \quad (1)$$

$$S(\omega) \Rightarrow \bigvee_{k=1}^M A_k(\bar{\omega}), \quad (2)$$

$$S(\omega) \Rightarrow \bigvee_{k=1}^M \exists \bar{x}_{\neq} A_k(\bar{x}). \quad (3)$$

Строго говоря, вместо формул (1), (2), (3) следовало бы писать соответственно:

$$S(\omega) \Rightarrow (? \bar{x}_{\neq}) A_k(\bar{x}), \quad (1')$$

$$S(\omega) \Rightarrow (? k_{k=1}^M) A_k(\bar{\omega}), \quad (2')$$

$$S(\omega) \Rightarrow (? k_{k=1}^M) (? \bar{x}_{\neq}) A_k(\bar{x}), \quad (3')$$

но рассматриваемые алгоритмы доказательства логических следований не только отвечают на вопрос «*существует ли ... ?*», но и предъявляют значения для переменных.

При решении задач медицинской диагностики в качестве множеств ω выступают пациенты, рассматриваемые как множество своих частей (органов), а исходные предикаты p_i — свойства этих частей и отношения между ними, обычно называемые симптомами. В этом случае классы объектов — это классы заболеваний или синдромы. Решение задач идентификации, классификации и анализа в этом случае сводится к доказательству тех же самых формул и соответствуют таким задачам как «обладает ли пациент заданным заболеванием (синдромом) и какие именно симптомы каких органов влекут это заболевание?», «какими заболеваниями (синдромами) страдает заданный орган?», «какие области пациента какими заболеваниями (синдромами) обладают?».

При решении задач анализа рынка в качестве множеств ω выступают множества субъектов и объектов рынка, исходные предикаты p_i — свойства этих объектов и субъектов, а также отношения между ними.

²Для того, чтобы записать, что значения для переменных списка \bar{x} , удовлетворяющие формуле $A(\bar{x})$, различны, вместо формулы:

$$\exists x_1 \dots \exists x_m (\&_{i=1}^m \&_{j=i+1}^m (x_i \neq x_j) \& A(x_1, \dots, x_m))$$

, будет использоваться обозначение:

$$\exists \bar{x}_{\neq} A(\bar{x}).$$

Описания классов могут задавать желаемые (или негативные) ситуации на рынке. Решение задач идентификации, классификации и анализа в этом случае сводится к доказательству тех же самых формул и соответствуют таким задачам как «какие объекты рынка влияют на то, что при имеющемся состоянии рынка будет получена заданная ситуация?», «какие ситуации следуют из заданных свойств объектов рынка?», «какие подмножества объектов рынка к каким ситуациям приводят?».

Можно привести ещё много задач искусственного интеллекта, решение которых является решением сформулированных задач и сводится к доказательству логических следований (1), (2), (3).

Заметим, что для того, чтобы уметь доказывать (1), (2), (3), достаточно уметь доказывать логическое следование:

$$S(\omega) \Rightarrow \exists \bar{x}_{\neq} A(\bar{x}), \quad (4)$$

где $A(\bar{x})$ — элементарная конъюнкция атомарных формул и их отрицаний. В [3, 6] доказаны оценки числа шагов алгоритмов, решающих задачу (4), а также задачи (1), (2), (3).

Так, например, для алгоритма полного перебора всех возможных подстановок констант из ω вместо переменных из $A(\bar{x})$, число шагов составляет $A_t^m \cdot \sum_{i=1}^{2^n} a_i \cdot (s_i + 1)$, где t — число элементов в ω , m — количество переменных в $A(\bar{x})$, s_i, a_i — количества вхождений атомарных формул с предикатом p_i в $S(\omega)$ и в $A(\bar{x})$ соответственно, A_t^m — число размещений из t по m . Эта оценка составляет $O(t^m \cdot \sum_{i=1}^n a_i \cdot (s_i + 1))$.

Для алгоритмов, основанных на построении вывода в исчислении предикатов [?]cl, оценки имеют вид $O(s_1^{a_1} \cdot \dots \cdot s_n^{a_n}) = O(s^a)$, где $s = \max_i(s_i)$, $a = \sum_{i=1}^n a_i$

Доказана NP-полнота задач (1), (2), (3) и, следовательно, NP-трудность задач (1'), (2'), (3').

3. Многоуровневое описание классов. Для уменьшения числа шагов работы алгоритмов, решающих описанные задачи, в [4] предложено иерархическое многоуровневое описание классов распознаваемых объектов.

Рассматриваются объекты, структура которых позволяет выделить достаточно простые их части и дать описание объекта в терминах свойств этих частей и отношений между ними. В частности, это можно сделать, выделяя «часто» встречающиеся подформулы $P_i^1(\bar{y})$, формул $A_k(\bar{x})$ «небольшой сложности». При этом записывается система эквивалентностей вида $p_i^1(y^1) \leftrightarrow P_i^1(\bar{y})$, где p_i^1 — новые предикаты, которые будем называть предикатами 1-го уровня, а переменные y^1 —

новые переменные для списков исходных переменных, которые будем называть переменными 1-го уровня.

Обозначим предикаты, задаваемые этими подформулами, посредством p_i^1 ($i = 1, \dots, n_1$) и назовем их предикатами первого уровня. Эти предикаты при $i = 1, \dots, n_1$ определяются равносильностями:

$$p_i^1(x_i^1) \leftrightarrow P_i^1(\bar{y}_i^1).$$

Обозначим формулы, полученные из $A_k(\bar{x})$ путем замены всех вхождений формул вида $P_i^1(\bar{y}_i^1)$ на атомарные формулы $p_i^1(x_i^1)$ (при $y_i^1 \subseteq x$) посредством $A_k^1(\bar{x}^1)$. Здесь \bar{x}^1 – список всех переменных формулы $A_k^1(\bar{x}^1)$, состоящий как из некоторых (быть может всех) исходных переменных формулы $A_k(\bar{x})$, так и из переменных первого уровня, появившихся в формуле $A_k^1(\bar{x}^1)$. Такие формулы $A_k^1(\bar{x}^1)$ можно рассматривать как описания классов в терминах предикатов исходного (нулевого) и первого уровней.

Описанием объекта $S^1(\omega)$ первого уровня назовем множество всех атомарных формул вида $p_i^1(\omega_{ij}^1)$, для которых истинна определяющая подформула $P_i^1(\bar{\tau}_{ij}^1)$ при $\tau_{ij}^1 \subset \omega$, а объект первого уровня ω_{ij}^1 представляет из себя список исходных объектов $\bar{\tau}_{ij}^1$.

Процедуру выделения «часто» встречающихся подформул «небольшой сложности» можно повторить с формулами $A_k^1(\bar{x}^1)$.

В результате построения составных предикатов (т.е. предикатов различных уровней) и многоуровневого описания классов исходная система описаний классов может быть записана с помощью равносильной ей многоуровневой системы описаний классов вида:

$$\left\{ \begin{array}{l} A_k^L(\bar{x}^L) \\ p_1^1(x_1^1) \Leftrightarrow P_1^1(\bar{y}_1^1) \\ \vdots \\ p_{n_1}^1(x_{n_1}^1) \Leftrightarrow P_{n_1}^1(\bar{y}_{n_1}^1) \\ \vdots \\ p_i^l(x_i^l) \Leftrightarrow P_i^l(\bar{y}_i^l) \\ \vdots \\ p_{n_L}^L(x_{n_L}^L) \Leftrightarrow P_{n_L}^L(\bar{y}_{n_L}^L) \end{array} \right. .$$

Решение рассмотренных задач распознавания может быть сведено к последовательной реализации $L + 1$ этапов.

1. Для каждого $j = 1, \dots, n_1$ осуществляется проверка выводимости из $S(\omega)$ формулы $\exists \bar{x}_{j \neq}^1 P_j^1(\bar{x}_j^1)$ и нахождение всех таких объектов 1-го уровня, существование которых утверждается в правой части логического следования и, следовательно, истинны атомарные формулы $p_j^1(\omega_j^1)$. При этом будет получено описание объекта 1-го уровня $S^1(\omega)$.

⋮

l . Для каждого $j = 1, \dots, n_l$ осуществляется проверка выводимости из $S(\omega) \cup S^1(\omega) \dots \cup S^{l-1}(\omega)$ формулы $\exists \bar{x}_{j \neq}^l P_j^l(\bar{x}_j^l)$ и нахождение всех таких объектов l -го уровня, существование которых утверждается в правой части логического следования, а следовательно истинны атомарные формулы $p_j^l(\omega_j^l)$. При этом будет получено описание объекта l -го уровня $S^l(\omega)$.

⋮

$L+1$. Проверка выводимости из $S(\omega) \cup S^1(\omega) \dots \cup S^L(\omega)$ формулы, соответствующей решаемой задаче, с заменой исходных описаний классов A_k на описания классов L -го уровня A_k^L .

Суммарное число шагов решения сформулированных задач (1), (2) и (3) складывается из числа шагов, затраченных на выполнение п.п. 1, ..., L .

Следует отметить, что при реализации каждого из п.п. 1, ..., L по сути дела проверяется логическое следование вида (4). В связи с этим можно уточнить понятие «небольшая сложность». Для переборного алгоритма это означает «небольшое» количество предметных переменных в формулах. Для алгоритмов же, основанных на проверке выводимости в исчислении предикатов, «небольшая сложность» означает «небольшое» количество атомарных формул в формулах, участвующих в выполнении п.п. 1, ..., L .

4. Понятие неполной выводимости формулы. Понятие неполной выводимости предикатной формулы было введено в [5] для распознавания объектов с неполной информацией. При этом рассматривается задача проверки того, что из истинности всех формул множества $S(\omega)$ следует истинность $A(\bar{x})$ или некоторой её максимальной подформулы $\tilde{A}(\bar{y})$ на наборе различных констант из ω , где список переменных \bar{y} является подписанием списка переменных \bar{x} .

Пусть a и \tilde{a} – количество атомарных формул в $A(\bar{x})$ и в $\tilde{A}(\bar{y})$ соответственно, m и \tilde{m} – количество предметных переменных в $A(\bar{x})$ и $\tilde{A}(\bar{y})$ соответственно.

Числа q и r вычисляются по формулам $q = \frac{\tilde{a}}{a}$, $r = \frac{\tilde{m}}{m}$ и характе-

ризуют степень совпадения формул $A(\bar{x})$ и $\tilde{A}(\bar{y})$. При этом $0 < q \leq 1$, $0 < r \leq 1$. Кроме того, $q = r = 1$ тогда и только тогда, когда $\tilde{A}(\bar{y})$ совпадает с $A(\bar{x})$.

Замечание. Возможен другой вариант определения чисел q и r . Каждому предикату и каждой предметной переменной формулы можно приписать их «вес», определяемые либо экспертами, либо из вероятностных соображений. Тогда $q = \frac{\tilde{w}}{w}$, $r = \frac{\tilde{v}}{v}$, где w и \tilde{w} – сумма «весов» предикатных формул в $A(\bar{x})$ и в $\tilde{A}(\bar{y})$ соответственно, v и \tilde{v} – сумма «весов» предметных переменных в $A(\bar{x})$ и $\tilde{A}(\bar{y})$ соответственно.

При таких обозначениях формулу $\tilde{A}(\bar{y})$ будем называть (q, r) -фрагментом формулы $A(\bar{x})$.

Задача нахождения (q, r) -фрагмента с максимально возможным значением параметра q называется задачей проверки неполной выводимости формулы. В [5] приведён один из возможных алгоритмов её решения.

5. Подход к построению многоуровневого описания классов. Понятие неполной выводимости формулы позволяет разработать подход к выделению подформул с требуемыми свойствами.

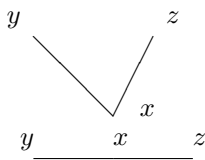
1. Для каждой пары элементарных конъюнкций $A_i(\bar{x}_i)$ и $A_j(\bar{x}_j)$, входящих в описания классов, посредством проверки неполной выводимости для $A_i(\bar{x}_i) \Rightarrow \exists \bar{x}_j \neq A_j(\bar{x}_j)$ и $A_j(\bar{x}_j) \Rightarrow \exists \bar{x}_i \neq A_i(\bar{x}_i)$ выделяем их максимальную (с точностью до имён предметных переменных) подформулу $Q_{ij}^1(\bar{x}_{ij})$.

2. Повторяем процесс выделения общих подформул для $Q_{i_1 \dots i_{2l-1}}^{l-1}(\bar{x}_{i_1 \dots i_{2l-1}})$ и $Q_{j_1 \dots j_{2l-1}}^{l-1}(\bar{x}_{j_1 \dots j_{2l-1}})$, получив их общие (с точностью до имён предметных переменных) подформулы $Q_{i_1 \dots i_{2l-1} j_1 \dots j_{2l-1}}^l(\bar{x}_{i_1 \dots i_{2l-1} j_1 \dots j_{2l-1}})$ ($l = 2, \dots, L$). Процесс завершится, так как на каждой итерации длины подформул уменьшаются.

3. Выберем среди подформул $Q_{i_1 \dots i_{2l-1} j_1 \dots j_{2l-1}}^l(\bar{x}_{i_1 \dots i_{2l-1} j_1 \dots j_{2l-1}})$ такие, которые удовлетворяют требуемым условиям и обозначим их посредством $P_i^1(\bar{y}_i^1)$ ($i = 1, \dots, n_1$).

4. Формулы $P_i^{l+1}(\bar{y}_i^{l+1})$ ($i = 1, \dots, n_{l+1}$, $l = 2, \dots, L$) строятся из выделенных ранее подформул $Q_{i_1 \dots i_{2l-1} j_1 \dots j_{2l-1}}^l(\bar{x}_{i_1 \dots i_{2l-1} j_1 \dots j_{2l-1}})$ с учётом требуемых условий.

6. Пример построения уровневого описания. Пример контурных изображений взят из [1]. Пусть имеется множество контурных изображений, составленных из отрезков прямых, задаваемых своими концами. Заданы два предиката V и L , определяемых следующим образом:



$$V(x, y, z) \iff (\angle yxz < \pi)$$

$$L(x, y, z,) \iff x \text{ лежит на отрезке с концами } y \text{ и } z$$

Задан класс контурных изображений ящика в различных ракурсах, представленных на рисунке 1.

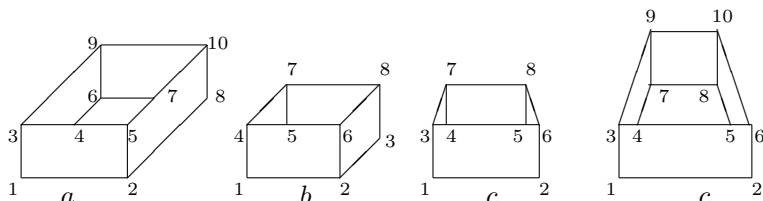


Рис. 1. Контурные изображения ящика в различных ракурсах

Эти эталонные изображения позволяют создать описание класса почти всех контурных изображений ящика (с точностью до зеркального отображения) путём замены различных констант в их описаниях на различные переменные и выписывания знака & между атомарными формулами. Это описание класса содержит 4 элементарные конъюнкции, содержащие соответственно 10, 8, 10, 8 переменных и 30, 22, 26, 32 атомарные формулы. Так, например, элементарная конъюнкция, соответствующая изображению *b*, имеет вид:

$$\begin{aligned} &V(x_1, x_4, x_2) \& V(x_2, x_1, x_6) \& V(x_2, x_6, x_3) \& V(x_2, x_1, x_3) \& \\ &V(x_3, x_2, x_8) \& V(x_4, x_5, x_1) \& V(x_4, x_6, x_1) \& V(x_4, x_7, x_5) \& \\ &V(x_5, x_4, x_7) \& V(x_5, x_7, x_6) \& V(x_6, x_2, x_5) \& V(x_6, x_2, x_4) \& \\ &V(x_6, x_5, x_8) \& V(x_6, x_4, x_8) \& V(x_6, x_8, x_2) \& V(x_7, x_5, x_4) \& \\ &V(x_7, x_8, x_5) \& V(x_7, x_8, x_4) \& V(x_8, x_3, x_6) \& V(x_8, x_6, x_7) \& \\ &V(x_8, x_3, x_7) \& L(x_5, x_4, x_6) \end{aligned}$$

Если требуется найти изображения всех коробок на сложной сцене, содержащей *t* вершин, то экспоненциальный сомножитель в оценке числа шагов для переборного алгоритма имеет вид t^{10} , а для алгоритма, основанного на построении вывода, s^{32} (здесь *s* – максимальное число вхождений одного и того же предиката в описание сцены и приблизительно равно, но меньше, чем $3t$).

Попарная проверка частичной выводимости между этими элементарными конъюнкциями позволяет выделить их общие подформулы, соответствующие изображениям на рисунке 2.

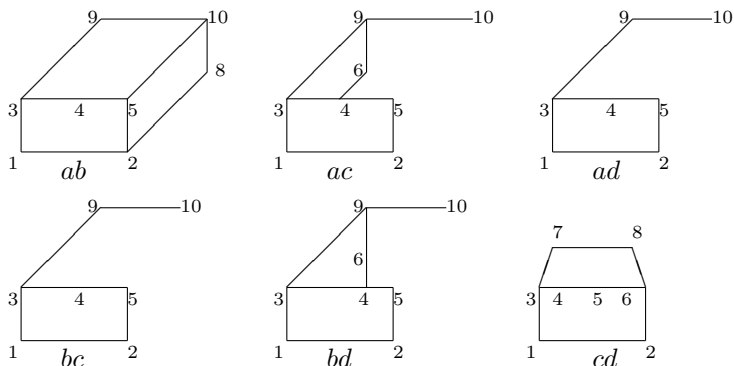


Рис. 2. Изображения, соответствующие выделенным подформулам

Эти подформулы содержат соответственно 8, 8, 7, 7, 7, 8 переменных и 18, 15, 11, 11, 15, 16 атомарных формул.

Последующее попарное выделение общих подформул даёт одну подформулу, содержащую 7 переменных и 11 атомарных формул, соответствующую изображению на рисунке 3.

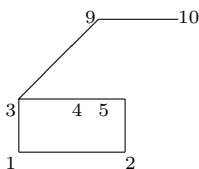


Рис. 3. Изображение, соответствующее повторному выделению подформул

Элементарная конъюнкция $P^1(x_1, x_2, x_3, x_4, x_5, x_9, x_{10}) = V(x_1, x_3, x_2) \& V(x_2, x_1, x_5) \& V(x_3, x_4, x_1) \& V(x_3, x_5, x_1) \& V(x_3, x_9, x_4) \& V(x_3, x_9, x_5) \& V(x_3, x_9, x_1) \& V(x_5, x_2, x_4) \& V(x_5, x_2, x_3) \& V(x_9, x_{10}, x_3) \& L(x_4, x_3, x_5)$, соответствующая этому изображению, определяет предикат 1-го уровня $p^1(x^1)$. Переменная первого уровня x^1 – это переменная для списка из 7 исходных переменных.

Элементарные конъюнкции $P_1^2(\overline{y_1^2})$, $P_2^2(\overline{y_2^2})$, $P_3^2(\overline{y_3^2})$, $P_4^2(\overline{y_4^2})$, соответствующие изображениям ab , ac , bd , cd и записанные с использо-

ванием предиката $p^1(x^1)$ определяют предикаты второго уровня $p_1^2(x_1^2)$, $p_2^2(x_2^2)$, $p_3^2(x_3^2)$, $p_4^2(x_4^2)$.

Например, подформула $P_1^2(\overline{y_1^2}) = p^1(x^1) \& V(x_2, x_5, x_8) \& V(x_2, x_1, x_8) \& V(x_5, x_4, x_{10}) \& V(x_5, x_3, x_{10}) \& V(x_8, x_2, x_{10}) \& V(x_{10}, x_8, x_5) \& V(x_{10}, x_5, x_9) \& V(x_{10}, x_8, x_9)$ соответствует изображению ab . Здесь y_1^2 – переменная для списка, состоящего из переменных $x^1, x_1, x_2, x_4, x_5, x_8, x_9, x_{10}$, причём переменная 1-го уровня x^1 – это переменная для списка исходных переменных $x_1, x_2, x_3, x_4, x_5, x_9, x_{10}$. Заметим, что при определении значения для переменной 1-го уровня x^1 только переменная x_8 не получила значения.

Если требуется найти изображения всех коробок на сложной сцене, содержащей t вершин, доказательство следования из $S(\omega)$ элементарной конъюнкции $P^1(x_1, x_2, x_3, x_4, x_5, x_9, x_{10})$, определяющей предикат 1-го уровня $p^1(x^1)$, и определение значений для переменных $x_1, x_2, x_3, x_4, x_5, x_9, x_{10}$ то экспоненциальный сомножитель в оценке числа шагов для переборного алгоритма имеет вид t^7 , а для алгоритма, основанного на построении вывода, s^{11} . При этом описание $S^1(\omega)$ содержит столько членов, сколько на сцене имеется объектов «подозрительных» на то, чтобы быть коробкой. Пусть их t^1 штук (t^1 много меньше, чем t).

Элементарные конъюнкции $P_1^2(\overline{y_1^2})$, $P_2^2(\overline{y_2^2})$, $P_3^2(\overline{y_3^2})$, $P_4^2(\overline{y_4^2})$ содержат соответственно 1, 1, 0, 1 «новых» переменных и 7, 4, 4, 5 «новых» атомарных формул. Доказательство следования из $S(\omega) \cup S^1(\omega)$ этих элементарных конъюнкций, определяющих предикаты 2-го уровня $p_1^2(x_1^2)$, $p_2^2(x_2^2)$, $p_3^2(x_3^2)$, $p_4^2(x_4^2)$ и определение значений для не более, чем одной переменной, экспоненциальный сомножитель в оценке числа шагов для переборного алгоритма имеет вид $t + t^1 = O(t)$, а для алгоритма, основанного на построении вывода, $(s + t^1)^7 = O(s^7)$. При этом описание $S^2(\omega)$ содержит не более членов, чем на сцене имеется объектов «подозрительных» на то, чтобы быть коробкой. Пусть их t^2 штук (t^2 много меньше, чем t).

Элементарные конъюнкции, полученные из описаний классов подстановкой предикатов 1-го и 2-го уровня вместо соответствующих подформул содержат соответственно 2, 0, 2, 2 «новых» переменных и 7, 4, 11, 16 «новых» атомарных формул. Доказательство следования из $S(\omega) \cup S^1(\omega) \cup S^2(\omega)$ элементарных конъюнкций, определяющих описание класса, и определение значений для не более, чем двух переменных, экспоненциальный сомножитель в оценке числа шагов для переборного алгоритма имеет вид $t + t^1 + t^2 = O(t)$, а для алгоритма, основанного на построении вывода, $(s + t^1 + t^2)^{16} = O(s^{16})$.

Так как $O(t^7) + O(t) + O(t^2) = O(t^7) < O(t^{10})$ и $O(s^{11}) + O(s^7) + O(s^{16}) = O(s^{16}) < O(s^{32})$, то как переборный алгоритм, так и алгоритм, основанный на построении вывода в исчислении предикатов, при использовании построенного двухуровневого описания классов совершают меньшее число шагов.

7. Обсуждение полученных результатов. В работе описан подход к построению уровневого описания классов, существенно уменьшающих число шагов алгоритмов, решающих NP-трудные задачи искусственного интеллекта. Этот подход основан на ранее введённом понятии неполной выводимости предикатных формул.

Несмотря на то, что алгоритмы проверки неполной выводимости имеют достаточно высокие (экспоненциальные) верхние оценки числа шагов, их применение производится однократно на предварительном этапе формирования описаний классов, до начала решения основных задач. При этом число шагов алгоритмов, многократно решающих основные задачи, существенно уменьшается.

Литература

1. Дуда Р., Харт П. Распознавание образов и анализ сцен // М.: Мир. 1976. 511 с.
2. Клини С. Математическая логика // М.: Мир. 1973. 480 с.
3. Косовская Т.М. Доказательства оценок числа шагов решения некоторых задач распознавания образов, имеющих логические описания // Вестник Санкт-Петербургского университета. 2007. Сер. 1. Вып. 4. С. 82–90.
4. Косовская Т.М. Многоуровневые описания классов для уменьшения числа шагов решения задач распознавания образов, описываемых формулами исчисления предикатов // Вестник Санкт-Петербургского университета. 2008. Сер. 10. Вып. 1. С. 64–72.
5. Косовская Т. М. Частичная выводимость предикатных формул как средство распознавания объектов с неполной информацией // Вестник Санкт-Петербургского университета. 2009. Сер. 10. Вып. 1. С. 74–84.
6. Косовская Т.М. Некоторые задачи искусственного интеллекта, допускающие формализацию на языке исчисления предикатов, и оценки числа шагов их решения // Труды СПИИРАН, 2010. Вып. 14. С. 58–75.

References

1. Duda R., Hart P. Raspoznavanie obrazov i analiz scen [Pattern recognition and scene analysis]. M., Mir. 1976. 511 p. (In Russ.).
2. Klini S. Matematicheskaja logika [Mathematical logic]. M., Mir. 1973. 480 p. (In Russ.).
3. Kosovskaja T.M. [Prove the estimates of the number of steps to solving some problems of pattern recognition with the logical description]. *Vestnik*

- of Saint Petersburg university – Vestnik Sankt-Peterburgskogo universiteta.* 2007. Series 1. vol. 4. pp. 82–90. (In Russ.).
4. Kosovskaja T.M. [Multi-level description of the classes to reduce the number of steps solving pattern recognition, described by formulas of the predicate calculus]. *Vestnik of Saint Petersburg university – Vestnik Sankt-Peterburgskogo universiteta.* 2008. Series 10. vol. 1. pp. 64–72. (In Russ.).
 5. Kosovskaja T.M. [Partial hatchability predicate formulas as a means of recognition of objects with incomplete information]. *Vestnik of Saint Petersburg university – Vestnik Sankt-Peterburgskogo universiteta.* 2009. Series 10. vol. 1. pp. 74–84. (In Russ.).
 6. Kosovskaja T.M. [Some problems of artificial intelligence, allowing the formalization of the language of the predicate calculus, and estimates of the number of steps to solve them]. *Trudy SPIIRAN – SPIIRAS Proceedings.* 2010. vol. 14. pp. 58–75. (In Russ.).

Косовская Татьяна Матвеевна – д-р физ.-мат. наук, доцент, профессор кафедры информатики математико-механического факультета С.-Петербургского государственного университета (СПбГУ); ст. научн. сотр. лаборатории информационных технологий в управлении и робототехнике СПИИРАН. Область научных интересов: логический подход к решению задач искусственного интеллекта, теория сложности алгоритмов. Число научных публикаций – 78. kosovtm@gmail.com; мат-мех СПбГУ, Университетский пр., д. 28, Старый Петергоф, С.-Петербург, 198504, РФ; р.т. +7(812)428-4233; СПИИРАН, 14-я линия В.О., д. 39, г. Санкт-Петербург, 199178, РФ; р.т. +7(812)328-0421.

Tatiana M. Kosovskaya – Ph.D., D.Sci., associate professor, professor of Computer Science Chair of St.Petersburg State University (SPbSU); senior researcher of Laboratory of Information Technologies in Control and Robototechniques, SPIIRAS. Research area: logical approach to the solving of Artificial Intelligence problems, theory of complexity of algorithms. Number of publications – 78. kosovtm@gmail.com; Faculty of Mathematics and Mechanics SPbSU, University av., 28, Stary Petergof, St.Petersburg, 198504, Russia; office phone +7(812)428-4233; SPIIRAS, 14-th line V.O., 39, St. Petersburg, 199178, Russia; office phone +7(812)328-0421.

Поддержка исследований. Работа выполнена при финансовой поддержке РФФИ, проект № 14-08-01276-а.

Acknowledgements. This research is supported by RFBR (grant 14-08-01276-a).

РЕФЕРАТ

Косовская Т.М. **Подход к решению задачи построения многоуровневого описания классов на языке исчисления предикатов.**

Рассматриваются задачи искусственного интеллекта, в которых исследуемый объект представлен как множество элементов со свойствами и отношениями из фиксированного множества, а также допускающие формализацию на языке исчисления предикатов. Исследуемый объект задаётся своим описанием, представленным как множество постоянных атомарных формул (или их отрицаний), истинных для этого объекта. Целевые формулы записываются в виде дизъюнкции элементарных конъюнкций атомарных формул.

К рассматриваемым задачам относятся:

- задача идентификации (проверить, удовлетворяет ли исследуемый объект или его часть заданной целевой формуле и предъявить эту часть объекта);
- задача классификации (найти все такие целевые формулы, которые верны для исследуемого объекта);
- задача анализа (Найти и классифицировать все части объекта, поддающиеся классификации).

Все эти задачи NP-трудны. Построение уровня описания целевых формул посредством выделения из них общих подформул позволяет существенно уменьшить число шагов алгоритмов, решающих рассматриваемые задачи. Такие уровневые описания соответствуют декомпозиции задачи большой размерности на несколько последовательно решаемых задач меньшей размерности. При этом в оценке числа шагов алгоритма вместо большого показателя степени экспоненты, соответствующего размерности исходной задачи, получается сумма, каждое слагаемое которой в показателе степени имеет размерность подзадачи.

В настоящей статье предлагается подход к построению такого многоуровневого описания, основанного на попарном выделении общих подформул (с точностью до имён переменных) целевых формул. Такое попарное выделение общих подформул, в свою очередь, базируется на ранее введённом понятии неполной выводимости элементарной конъюнкции из множества постоянных атомарных формул.

Многоуровневое описание, после попарного выделения общих подформул (с точностью до имён переменных) у элементарных конъюнкций (входящих в целевые формулы), общих подформул выделенных подформул, и т.д., строится, начиная с самых «простых» (например, с самых коротких или с зависящих от наименьшего числа переменных) подформул.

Приведён модельный пример построения двухуровневого описания контурных изображений.

SUMMARY

Kosovskaya T.M. **An approach to the construction of a level description of classes by means of a predicate calculus language.**

Artificial Intelligence problems, an investigated object of which is presented as a set of elements with properties and relations from a fixed set, and permitting their formalization in predicate calculus language, are under consideration. An approach to the construction of a level description of classes of such objects is suggested. An investigated object is represented by its description as a set of constant atomic formulas (or their negations) which are true for this object. A goal formula is written in the form of disjunction of elementary conjunctions of atomic formulas.

Problems under consideration are such as:

- identification problem (whether an object or its part satisfies the given goal formula and extract such a part of the object);
- classification problem (to find all such numbers goal formulas that are valid for the investigated object);
- analysis problem (to find and classify all parts of an object which may be classified).

All these problems are NP-hard. Construction of a level description for goal formulas, by means of extracting their common sub-formulas, allows essentially to decrease the number of steps for an algorithm solving a problem under consideration. These descriptions correspond to decomposition of a high-dimensional problem up to several less-dimensional sub-problems which are solved sequentially. In this case the upper bound of number of steps of an algorithm is the sum of terms with exponents equal to dimensions of sub-problems instead of one term which exponent equals to the dimension of the initial problem.

An approach, based on pairwise extraction of common sub-formulas (up to the names of variables) from all goal formulas, is suggested to the construction of such a level description. This pairwise extraction of common sub-formulas is based, in turn, on the earlier introduced notion of partial deducibility of an elementary conjunction from a set of constant atomic formulas.

After pairwise extraction of common sub-formulas (up to the names of variables) from elementary conjunctions (occurred in goal formulas), common sub-formulas of already extracted sub-formulas, ... etc. the level description is built beginning with the most «simple» (for example, the most short or containing the least number of variables) sub-formulas.

A model example of construction of two-level description for contour images is presented.