CONDITIONS OF EFFECTIVENESS OF PATTERN RECOGNITION PROBLEM SOLUTION USING LOGICAL LEVEL DESCRIPTIONS OF CLASSES

Tatiana Kosovskaya

Abstract: Earlier the author have suggested a logical level description of classes which allows to reduce a solution of various pattern recognition problems to solution of a sequence of one-type problems with the less dimension. Here conditions of the effectiveness of the use of such a level descriptions are proposed.

Keywords: compound images, logical description of classes, effectiveness.

Introduction

Various pattern recognition problems wich may be described in the terms of predicates (which characterize the whole object or its parts) were reduced in [1] to the proof of deducibility of propositional and predicat calculus formulas from a set of atomic formulas.

Upper bounds of the number of steps of an algorithm solving pattern recognition problems with logical description were proved in [3]. For example, such an upper bound for an algorithm solving the problem of analysis of a compound object is polynomial but the degree of such a polynomial depends of the number of objective variables included to the class description. As a rule such a number is rather large.

A level description of classes offered in [2] allows to reduce the solution of various pattern recognition problems to the solution of a sequence of one-type problems with the less dimension. Here conditions of decreasing of the number of steps of algorithm solving the described pattern recognition problems with the use of many-level description are proposed.

1. Setting of a problem of compound objects logical recognition

Let Ω be a set of finite sets $\omega = \{\omega_1, \dots, \omega_t\}$. The set ω will be called a recognizable object. Let $p_1, \dots, p_n$ be a collection of predicates which characterize an object (global indication) or describe properties or relations between elements of ω (local indication). The set Ω is a union of K (may be intersected) classes Ω_k.

Logical description $S(\omega)$ of an object ω is a set of all true formulas in the form $p_i(x)$ or its negation written for all parts x of the object ω.

Logical description of a class (DC) Ω_k is such a formula $A_k(x)$ that $A_k(x)$ contains as an atomic only formulas of the form $p_i(y)$ where y is a subset of x; $A_k(x)$ has no quantifiers; if for some ordering ω' of the object ω the formula $A_k(\omega)$ is true then $\omega \in \Omega_k$

These descriptions may be used for solving the following problems.

Identification problem. To check weather object ω or its part belongs to the class Ω_k.

This problem was reduced in[1] to the proof of deducibility of the formula $\exists y\ (y \subset \omega\ \&\ A_k(y))$ from the description $S(\omega)$.

Classification problem. To find all such numbers k that $\omega \in \Omega_k$.

This problem was reduced in[1] to the proof of deducibility of disjunction of formulas $A_k(\omega')$ (for some ordering ω' of the object ω) from the description $S(\omega)$ and pointing out all such numbers k for which the corresponding disjunct is true for ω.

Problem of analysis of a compound object. To find and classificate all parts x of the object ω.

This problem was reduced in[1] to the proof of deducibility of disjunction of formulas $\exists y\ (y \subset \omega\ \&\ A_k(y))$ from the description $S(\omega)$ and pointing out all parts of ω wich mai be classified.

2. Level logical description of classes

Objects the structure of which allows to extract more simple fragments and to describe these objects in the terms of properties and relations between such fragments are regarded. In particular it may be done be means of selecting «frequently» appeared subformulas of formulas $A_k(x)$ with «small complexity». A system of equivalences in the form $p_j^1(x_j^1) \Leftrightarrow P_j^1(y_j^1)$ (where x_j^1 – new first-level variables, p_j^1 – new first-level predicates, $P_j^1(y_j^1)$ – subformulas of formulas $A_k(x)$) is written. The result of substitution of $p_j^1(x_j^1)$ instead of $P_j^1(y_j^1)$ into $A_k(x)$ is denoted by $A_k^1(x^1)$.

Such a procedure may be repeated with $A_k^1(x^1)$ but not later than $A_k^l(x^1)$ contains at least two subformulas in the same form.

3. Conditions of effectiveness of level description with the use of global indications

Let $p_1, \ldots, p_n$ be global indications (i.e. they are boolean variables). Then class descriptions are disjunctive normal forms (DNF) and any subformula of formulas $A, \ldots, A_K$ which appears at least two times is a simple conjunction.

Definition. Atom is variable or its negation.

Definition. Simple conjunctions $B_1, \ldots, B_m$ are called disjoint if there not exists such an atom that is included simultaneously into two different conjunctions.

Notifications.

a – a number of occurrences of boolean variables in formulas in DNF $A_1, \ldots, A_K$,

$P_1^1, \ldots, P_{n1}^1$ – subformulas of $A_1, \ldots, A_K$,

N_j^1 – a number of occurrences of subformula P_j^1,in $A_1, \ldots, A_K$,

v_j^1 – a number of occurrences of boolean variables in P_j^1,

$A_1^1, \ldots, A_K^1$ – the result of substitutions of atomic formulas p_j^1 instead of P_j^1 into $A_1, \ldots, A_K$,

Theorem 1. If formulas $P_1^1, \ldots, P_{n1}^1$ are disjoint then for the equality $a^1 = d\ a$ (for some $0 < d < 1$) it is necessary and sufficient

$$\Sigma_{j=1}^{n1}\ (v_j^1 - 1)\ N_j^1 = (1-d)\ a\ . \qquad (1)$$

Corollary 1.1. If formulas $P_1^1, \ldots, P_{n1}^1$ are disjoint then for decreasing the number of occurrences of boolean variables in formulas $A_1^1, \ldots, A_K^1$ in comparison with the number of occurrences of boolean variables in formulas $A_1, \ldots, A_K$ it is necessary and sufficient

$$\Sigma_{j=1}^{n1}\ (v_j^1 - 1)\ N_j^1 > \ a\ . \qquad (2)$$

Corollary 1.2. If formulas $P_1^1, \ldots, P_{n1}^1$ are disjoint and $N_j > N$ for some N then for decreasing the number of occurrences of boolean variables in formulas $A_1^1, \ldots, A_K^1$ in comparison with the number of occurrences of boolean variables in formulas $A_1, \ldots, A_K$ it is necessary and sufficient

$$\Sigma_{j=1}^{n1} \ (v_j^1 - 1) \ \geq \ a / N . \quad (3)$$

The next theorem gives a necessary condition for not disjoint formulas $P_1^1, \ldots , P_{n1}^1$.

Theorem 2. For the equality $a^1 = d\, a$ (for some $0 < d < 1$) it is necessary and sufficient

$$\Sigma_{j=1}^{n1} \ (v_j^1 - 1) \ N_j^1 \geq (1\text{-}d) \ a . \quad (4)$$

If $p_1, \ldots , p_n$ are boolean variables then both the identification problem and the classification problem may be solved with the use of resolution method or sequent propositional calculus the number of rule applications of which is not more then the number of occurrences of boolean variables in formula A_k (in formulas $A_1, \ldots , A_K$ for classification problem) [3]. Note that the upper bound of number of steps needed for calculation of p_j^1 equals to such a bound for classification problem if instead of A_k we take P_j^1.

Theorem 3. If $a^1 = d \ a$ then for decreasing of number of rule application steps while using the 2-level description it is sufficient

$$\Sigma_{j=1}^{n1} \ v_j^1 \ \leq \ (1\text{-}d) \ a . \quad (5)$$

4. Examples of two-level descriptions

Illustrate an application of the received conditions with a model example of 2-level description.

Let the set of recognizable objects is divided into 3 classes and objects may be described by means of 5 boolean variables *x, y, z, u, v*. Classes descriptions which allow to identify and classify an object have the form

$A_1 = \sim x \,\&\, \sim y \quad V \quad x \,\&\, \sim y \,\&\, z \quad V \quad x \,\&\, y \,\&\, z \,\&\, \sim v$

$A_2 = \sim x \,\&\, y \,\&\, \sim z \,\&\, \sim u \quad V \ \sim x \,\&\, y \,\&\, u \,\&\, \sim v \quad V \quad x \,\&\, \sim z \,\&\, \sim u \quad V \quad x \,\&\, z \,\&\, u \,\&\, \sim v$

$A_3 = \sim x \,\&\, y \,\&\, z \,\&\, \sim u \quad V \quad \sim x \,\&\, y \,\&\, u \,\&\, v \quad V \quad x \,\&\, \sim z \,\&\, u \,\&\, v \quad V \quad x \,\&\, y \,\&\, z \,\&\, v$

The number of occurrences of boolean variables in formulas A_1, A_2, A_3 is a=40.

Example 1. Let the following subformulas are extracted.

$P_1^1 = x \,\&\, y \,\&\, z$

$P_2^1 = \sim x \,\&\, y \,\&\, u$

$P_3^1 = \sim x \,\&\, y \,\&\, \sim u$

$P_4^1 = x \,\&\, z$

$P_5^1 = y \,\&\, z$

The number of occurrences of boolean variables in these subformulas is $\nu_1^1=3$, $\nu_2^1=3$, $\nu_3^1=3$, $\nu_4^1=2$, $\nu_5^1=2$. The number of occurrences of each of these subformulas is $N_j^1=2$.

The formulas are not disjoint and we can use the condition (4) $\Sigma_{j=1}^{n1} \ (v_j^1 - 1) N_j^1 \geq (1\text{-}d)\ a$. In this example $\Sigma_{j=1}^{n1} \ (v_j^1 - 1) \ N_j^1 = 16$. Hence for decreasing the length of description it is necessary $16 \leq 40$ (1-d), i.e. $d \geq 0.4$.

In fact

$A_1^1 = \sim x \,\&\, \sim y \quad V \quad \sim y \,\&\, p_4^1 \quad V \quad \sim v \,\&\, p_1^1 \,\&\, p_4^1 \,\&\, p_5^1$

$A_2^1 = \sim z \,\&\, p_3^1 \quad V \quad \sim v \,\&\, p_2^1 \quad V \quad x \,\&\, \sim z \,\&\, \sim u \quad V \quad u \,\&\, \sim v \,\&\, p_4^1$

$A_3^1 = p_3^1 \,\&\, p_5^1 \quad V \quad v \,\&\, p_2^1 \quad V \quad x \,\&\, \sim z \,\&\, u \,\&\, v \quad V \quad v \,\&\, p_1^1 \,\&\, p_4^1$

The number of occurrences of boolean variables in A_1^1, A_2^1, A_3^1 is a^1=29. As $a^1 = d\, a$ we have d= 29/40 = 0.725.

Verify, weather we may guarantee that such a 2-level description provides a decreasing of an upper bound of number of steps of a solution of classification problem, i.e. weather the condition (5) $\Sigma_{j=1}^{n1} \ v_j^1 \ \leq (1\text{-}d)\ a$ is fulfilled. In this example $\Sigma_{j=1}^{5} v_j^1$ =13, $(1\text{-}d)\ a$ = 11. Hence the condition (5) is not fulfilled and we can not guarantee a decreasing of an upper bound of number of steps of a solution of classification problem.

Example 1. Let the following subformulas are extracted.

$P_1^1 = \ {\sim}x\ \&\ y$

$P_2^1 = \ x\ \&\ z$

$P_3^1 = \ y\ \&\ z$

The number of occurrences of boolean variables in these subformulas is ν_1^1=2, ν_2^1=2, ν_3^1=2. The number of occurrences of each of these subformulas is N_1^1=4, N_2^1=4, N_3^1=3.

The formulas are not disjoint and we can use the condition (4) $\Sigma_{j=1}^{n1} (v_j^1 -1)N_j^1 \geq (1\text{-}d)\ a$. In this example $\Sigma_{j=1}^{n1} (v_j^1 -1)\ N_j^1 = 11$. Hence for decreasing the length of description it is necessary $11 \leq 40$ (1-d), i.e. $d \geq 0.725$.

In fact

$A_1^1 = {\sim}x\ \&\ {\sim}y \quad V \quad {\sim}y\ \&\ p_2 \quad V \quad {\sim}v\ \&\ p_2^1\ \&\ p_3^1$

$A_2^1 = {\sim}z\ \&\ {\sim}u\ \&\ p_1^1 \quad V \quad u\ \&\ {\sim}v\ \&\ p_1^1 \quad V \quad x\ \&\ {\sim}z\ \&\ {\sim}u \quad V \quad u\ \&\ {\sim}v\ \&\ p_2^1$

$A_3^1 = {\sim}u\ \&\ p_1^1\ \&\ p_3^1 \quad V \quad u\ \&\ v\ \&\ p_1^1 \quad V \quad x\ \&\ {\sim}z\ \&\ u\ \&\ v \quad V \quad v\ \&\ p_2^1\ \&\ p_3^1$

The number of occurrences of boolean variables in A_1^1, A_2^1, A_3^1 is a^1=32. As a^1=$d\ a$ we have d= 32/40 = 0.8.

Verify, weather we may guarantee that such a 2-level description provides a decreasing of an upper bound of number of steps of a solution of classification problem, i.e. weather the condition (5) $\Sigma_{j=1}^{n1} v_j^1 \leq (1\text{-}d)\ a$ is fulfilled. In this example $\Sigma_{j=1}^{3} v_j^1$ = 6, $(1\text{-}d)a$= 8. Hence the condition (5) is fulfilled and we can guarantee a decreasing of an upper bound of number of steps of a solution of classification problem.

Conditions of effectiveness of level description with the use of local indications

Let $p_1, \ldots ,p_n$ characterize properties and relations of a recognizable object elements. In such a case it was proved in [3] that the number of steps of an algorithm solving identification problem is bounded by the number of arrangement of m_k from t : A_t^{mk}. For classification problem and problem of analysis of a compound object such a bound is $\Sigma_{k=1}^{K} A_t^{mk}$. (Here m_k – the number of objective variables in the description of the k-th class.

Consequently the number of steps of an algorithm solving these problems is an exponent of the number of objective variables in the description of classes (and a polynomial of a high degree for any particular description). Moreover, if it is possible to construct an algorithm which in a polynomial (over the length of classes descriptions) number of steps solves such problems then one of the most difficult problems of XXI century P=NP will be solved. However with the use of level logical descriptions it is possible to decrease an exponent in the upper bound of the number of steps of such an algorithm.

Notifications.

$m_1, \ldots , m_K$ – number of objective variables in formulas $A_1, \ldots , A_K$,

r – a number which is greater than number of objective variables in every formula $P_1^1, \ldots , P_{n1}^1$,

x_k – the string of variables of the formula A_k,

x_j^1 – new variables of the 1st level defined be equivalences $p_j^1(x_j^1) \Leftrightarrow P_j^1(y_j^1)$,

$s1$ – the number of variables occurred in $A_1, \ldots , A_K$ but not occurred in $P_1^1, \ldots , P_{n1}^1$.

Theorem 4. Checking weather formulas $A_1, \ldots , A_K$ are true on the set $\omega = \{\omega_1, \ldots , \omega_t\}$ is equivalent to checking equivalences $p_j^1(x_j^1) \Leftrightarrow P_j^1(y_j^1)$ and weather formulas $A_1^1, \ldots , A_K^1$ are true on the same set.

For decreasing the number of steps of an algorithm solving the problem of analysis of compound object it is sufficient

$$n_1 t^r + t^{s1+n1} < t^m. \quad (6)$$

Conclusion

Hence level logical description of classes of objects is described. In the frameworks of such an approach the conditions of decreasing of upper bounds of number of steps of an algorithm solving various pattern recognition problems including recognition of compound objects (compound images and scenes, complex signals and so on) are done.

Bibliography

[1]. Timofeev A.V., Kosovskaya T.M. Logic-Axiomatical Method of Knowledge Representation and Recognition for Robots. In: Computers in Industry 15 (1990). Proc. Intelligent Manufactoring Systems – IMS'89. Elsevier Science Publishers B.V. 1990.

[2]. Kosovskaya T.M. Level descriptions of classes for providing a solution in pattern recognition problems. In: Proc. III International Conf. “Discrete models in the theory of control systems”. Moscow, Dialog-MSU, 1998. (in Russian)

[3]. Kosovskaya T.M. Proofs of the number of step bounds for solving of some pattern recognition problems with logical description. In: Herald of St.Petersburg University, Series 1 (Mathematics), 2007 (?). (be published, in Russian)

Authors' Information

Tatiana Kosovskaya – St.Petersburg Institute of Informatics and Automatics of Russian Academy of Sciences, e-mail: kosov@NK1022.spb.edu