DISTANCE BETWEEN OBJECTS DESCRIBED BY PREDICATE FORMULAS

Tatiana Kosovskaya

Abstract: *Functions defining a distance and a distinguish degree between objects described by predicate formulas are introduced. It is proved that the introduced function of distance satisfies all properties of a distance. The function of objects distinguish degree adequately reflects similarity of objects but does not define a distance because the triangle inequality is not fulfilled for it. The calculation of the introduced functions is based on the notion of partial deduction of a predicate formula.*

Keywords: *artificial intelligence, pattern recognition, distance between objects, predicate calculus.*

ACM Classification Keywords: *I.2.4 ARTIFICIAL INTELLIGENCE Knowledge Representation Formalisms and Methods – Predicate logic.*

Conference topic: *Computer-related Distances.*

Introduction

A solution of an Artificial Intelligence (and pattern recognition) problem is often based on the description of an investigated object by means of global features which characterize the whole object but not its parts. In such a case a space of features appears and the distance between objects may be introduced in some natural way, for example, by the formula $\sum_i |x_i - y_i|^k$ for some natural *k*. In the case *k=1* we deal with a well-known Hamming metric which is widely used in the information theory for comparison of the same length strings of symbols.

But how we can measure the distance between two objects described by some local features which characterize some parts of an object or relations between such parts? What is the distance between two identical images one of which is situated in the left upper corner of the screen and the second is in the right lower corner? Even if we use a monochrome screen with two degrees of lightness the distance calculated by the formula $\sum_i |x_i - y_i|$ will give the number of pixels in the image itself multiplied by two.

If we analyze a market situation with two participants *A* and *B* (the participants of the market are not ordered) with the feature values $(p^1_1, p^1_2, \ldots, p^1_n)$ and $(p^2_1, p^2_2, \ldots, p^2_n)$ then in the dependence of their order $((p^1_1, p^1_2, \ldots, p^1_n), (p^2_1, p^2_2, \ldots, p^2_n))$ or $((p^2_1, p^2_2, \ldots, p^2_n), (p^1_1, p^1_2, \ldots, p^1_n))$ we receive that these two situations are essentially different (the distance between them may be up to *2n)*.

To recognize objects from the done set Ω every element of which is a set $\omega = \{\omega_1, \ldots, \omega_t\}$, a logic-objective approach was described in [Kosovskaya, 2007]. Such an approach consists in the following. Let the set of predicates $p_1, \ldots, p_n$ (every of which is defined on the elements of ω) characterizes properties of these elements and relations between them. Let the set Ω is a union of (may be intersected) classes $\Omega = \cup_{k=1}^{K} \Omega_k$.

Logical description $S(\omega)$ of the object ω is a collection of all true formulas of the form $p_i(\tau)$ or $\neg p_i(\tau)$ (where τ is an ordered subset of ω) describing properties of ω elements and relations between them.

Logical description of the class Ω_k is such a formula $A_k(\mathbf{x})$ that if the formula $A_k(\omega)$ is true then $\omega \in \Omega_k$. The class description always may be represented as a disjunction of elementary conjunctions of atomic formulas.

Here and below the notation $\mathbf{x}$ is used for an ordered list of the set x. To denote that all values for variables from the list $\mathbf{x}$ are different the notation $\exists \mathbf{x}_{\neq} A_k(\mathbf{x})$ will be used.

The introduced descriptions allow solving many artificial intelligence problems [Kosovskaya, 2011]. These problems may be formulated as follows. **Identification problem:** to check out such a part of the object ω which belongs to the class Ω_k. **Classification problem:** to find all such class numbers k that $\omega \in \Omega_k$. **Analysis problem:** to find and classify all parts τ of the object ω. The solution of these problems is reduced to the proof of predicate calculus formulas $S(\omega) \Rightarrow \exists \mathbf{x}_{\neq} A_k(\mathbf{x})$, $S(\omega) \Rightarrow \vee_{k=1}^{K} A_k(\mathbf{x})$, $S(\omega) \Rightarrow \vee_{k=1}^{K} \exists \mathbf{x}_{\neq} A_k(\mathbf{x})$.

The proof of every of these formulas is based on the proof of the sequent

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

where $A(\mathbf{x})$ is an elementary conjunction.

The notion of partial deduction was introduced by the author in [Kosovskaya, 2009] to recognize objects with incomplete information. In the process of partial deduction instead of the proof of (1) we search such a maximal sub-formula $A'(\mathbf{x}')$ of the formula $A(\mathbf{x})$ that $S(\omega) \Rightarrow \exists \mathbf{x}'_{\neq} A'(\mathbf{x}')$ and there is no information that $A(\mathbf{x})$ is not satisfiable on ω.

Let a and a' be the numbers of atomic formulas $A(\mathbf{x})$ and $A'(\mathbf{x}')$ respectively, m and m' be the numbers of objective variables in $A(\mathbf{x})$ and $A'(\mathbf{x}')$ respectively. Then partial deduction means that the object ω is an r-th part ($r = m'/m$) of an object satisfying the description $A(\mathbf{x})$ with the certainty $q = a'/a$.

More precisely, the formula $S(\omega) \Rightarrow \exists \mathbf{x}_{\neq} A(\mathbf{x})$ is partially (q,r)-deductive if there exists a maximal sub-formula $A'(\mathbf{x}')$ of the formula $A(\mathbf{x})$ such that $S(\omega) \Rightarrow \exists \mathbf{x}'_{\neq} A'(\mathbf{x}')$ is deducible and τ is the string of values for the list of variables $\mathbf{x}'$, but the formula $S(\omega) \Rightarrow \exists \mathbf{x}_{\neq} [DA'(\mathbf{x})]^{\mathbf{x}'}_{\tau}$ is not deducible. Here $[DA'(\mathbf{x})]^{\mathbf{x}'}_{\tau}$ is obtained from $A(\mathbf{x})$ by deleting from it all conjunctive members of $A'(\mathbf{x}')$, substituting values of τ instead of the respective variables of $\mathbf{x}'$ and taking the negation of the received formula.

The defined below distance between objects takes into account the non-coincidence of their descriptions as the Hamming metric. It may be calculated not only for descriptions with the same number of atomic formulas which are ordered in some natural way, but for such ones which are sets (not ordered) of an arbitrary finite power.

Distance and Distinguish Degree Between Objects

Let ω_0 and ω_1 be two objects with logical descriptions $S(\omega_0)$ and $S(\omega_1)$ respectively and $A_0(\mathbf{x}_0)$ and $A_1(\mathbf{x}_1)$ be elementary conjunctions constructed according to these logical descriptions by changing different constants by different variables and putting the sign & between the atomic formulas. It is evident that $S(\omega_i) \Rightarrow \exists \mathbf{x}_{\neq} A_i(\mathbf{x}_i)$ i.e. ω_i satisfies the formula $A_i(\mathbf{x}_i)$ (for $i = 0, 1$). Let $A_i(\mathbf{x}_i)$ contains a_i atomic formulas and t_i variables.

Let us construct a partial deduction of a sequent $S(\omega_i) \Rightarrow \exists \mathbf{x}_{\neq} A_{1-i}(\mathbf{x})$ for every $i = 0,1$.

Let $A'_{i,1-i}(\mathbf{x}'_{i,1-i})$ be the maximal (under the number of variables) sub-formula of the formula $A_{1-i}(\mathbf{x}_{1-i})$ for which such a partial deduction exists.

Let the formula $A'_{i,1-i}(\mathbf{x}'_{i,1-i})$ contains $a'_{i,1-i}$ atomic formulas and $t'_{i,1-i}$ variables. Note that $A'_{01}(\mathbf{x}'_{01})$ and $A'_{01}(\mathbf{x}'_{10})$ coincide (up to the names of variables) and hence $a'_{01} = a'_{10}$. Let $\Delta a'_{i,1-i} = a_i - a'_{i,1-i}$. $\Delta a'_{01}$ is the number of non-coincidences of atomic formulas (up to the names of constants) in $S(\omega_0)$ with respect to $S(\omega_1)$.

Definition. The distance between objects ω_0 and ω_1 is the sum of the number of non-coincidences of atomic formulas (up to the names of constants)

$$\rho(\omega_0,\omega_1) = \Delta a'_{0,1} + \Delta a'_{1,0}.$$

Remember examples from the introduction. If the image on the display screen is described by a predicate $p(i,j,x)$ "pixel with the number (i,j) has the lightness x" then two identical images one of which is situated in the left upper corner of the screen and the second is in the right lower corner have the same (up to the names of constants) logical descriptions. Therefore the distance between them equals 0.

The logical description of two participants in the example with market participants is a set of atomic formulas $\{p_1(A), p_2(A),\ldots,p_n(A), p_1(B), p_2(B),\ldots,p_n(B)\}$ or $\{p_1(B), p_2(B),\ldots,p_n(B), p_1(A), p_2(A),\ldots,p_n(A)\}$ which are equal and the distance equals 0.

It is natural that in dependence of the chosen initial predicates the distance between objects may differ. Let's give an example of distance calculation between two contour images described by two predicate systems.

Example. Let we have two images A and B represented on the figures 1 and 2.

Consider two systems of initial predicates.

1. $V(x,y,z) \Leftrightarrow \angle yxz < \pi$,

   $I(x,y,z,u,v) \Leftrightarrow$ "the vertex x is a point of intersection of segments $[y,z]$ and $[u,v]$.

In such a case every of the points $a_1, a_2, a_5, a_9, a_{10}$ is represented in $A_a(x_1,\ldots,x_{10})$ by one formula. For example, the point a_1 is represented in $A_a(x_1,\ldots,x_{10})$ by $V(x_1,x_6,x_3)$.

Every of the points $b_1, b_2, b_5, b_9, b_{10}$ is represented in $A_b(x_1,\ldots,x_{10})$ by six formulas. For example, the point b_1 is represented in $A_b(x_1,\ldots,x_{10})$ by $V(x_1,x_2,x_3)$, $V(x_1,x_3,x_4)$, $V(x_1,x_4,x_5)$, $V(x_1,x_2,x_4)$, $V(x_1,x_3,x_5)$, $V(x_1,x_2,x_5)$.

Every of the points a_3, a_4, a_6, a_7, a_8 (as well as b_3, b_4, b_6, b_7, b_8) is represented in $A_a(x_1,\ldots,x_{10})$ (and in $A_b(x_1,\ldots,x_{10})$) by five formulas. For example, the point a_3 is represented in $A_a(x_1,\ldots,x_{10})$ by $V(x_3,x_4,x_1)$, $V(x_3,x_1,x_2)$, $V(x_3,x_2,x_6)$, $V(x_3,x_6,x_4)$, $I(x_3,x_1,x_6,x_2,x_4)$. So, the formula $A_a(x_1,\ldots,x_{10})$ has 30 atomic formulas and the formula $A_b(x_1,\ldots,x_{10})$ has 55 atomic formulas.

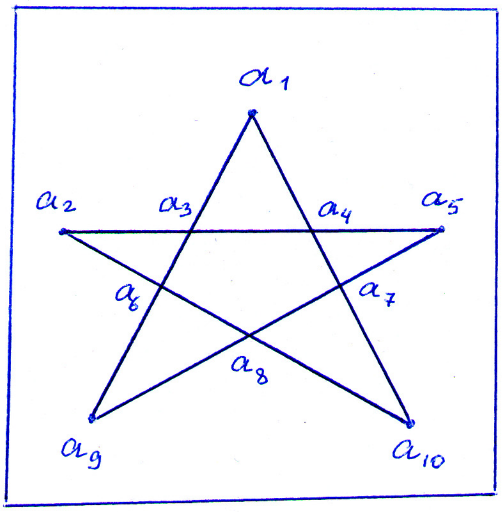

Fig. 1. Image *A*.

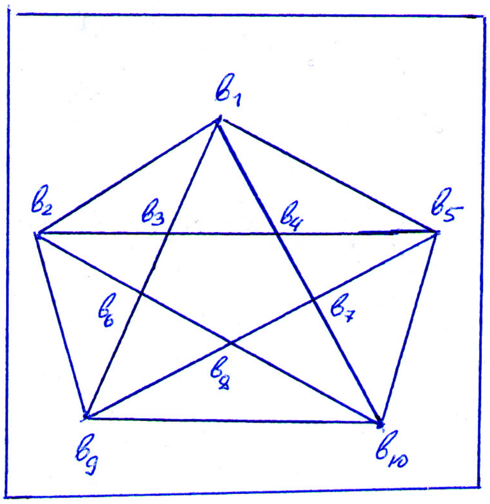

Fig. 2. Image *B*.

While construction partial deduction of $S(A) \Rightarrow \exists \mathbf{x}_{\neq} A_b(\mathbf{x}_b)$ all variables will receive some values but $A'_b(\mathbf{x}'_b)$ (coinciding with $A_a(\mathbf{x})$) contains only 30 atomic formulas. Hence $\Delta a'_{a,b} = 55 - 30 = 25$.

While construction partial deduction of $S(B) \Rightarrow \exists \mathbf{x}_{\neq} A_a(\mathbf{x}_a)$ all variables receive some values and all atomic formulas from $A_a(\mathbf{x}_a)$ are included into $A'_a(\mathbf{x}_a)$. Hence $\Delta a'_{a,b} = 0$.

In such a case $\rho(A,B) = 25 + 0 = 25$.

2. $E(x,y) \Leftrightarrow$ "x and y are adjacent".

This predicate describes not every point individually but a binary relation between them. As a fact we have a set of edges of a planar graph. The formula $A_a(x_1,\ldots,x_{10})$ has 15 atomic formulas and the formula $A_b(x_1,\ldots,x_{10})$ has 20 atomic formulas. While construction partial deductions of $S(A) \Rightarrow \exists \mathbf{x}_{\neq} A_b(\mathbf{x}_b)$ and $S(B) \Rightarrow \exists \mathbf{x}_{\neq} A_a(\mathbf{x}_a)$ it will be received that $\Delta a'_{a,b} = 20 - 15 = 5$, $\Delta a'_{b,a} = 0$. And $\rho(A,B) = 15$.

These examples demonstrate that besides the fact that different initial predicates provide different distances between objects; the value of the calculated distance does not illustrate the degree of their similarity. To overcome such a lack we may normalize the defined distance in order that it is not greater than 1. It may be done, for example, by dividing the distance by $a_0 + a_1$.

Definition. The degree of distinction between the objects ω_0 and ω_1 is the sum of the number of non-coincidences of atomic formulas (up to the names of constants) divided by the sum of numbers of atomic formulas in elementary conjunctions $A_0(\mathbf{x}_0)$ and $A_1(\mathbf{x}_1)$

$$d(\omega_0,\omega_1) = (\Delta a'_{0,1} + \Delta a'_{1,0}) / (a_0 + a_1).$$

The distinction degrees between the objects A and B in the previous example are $d(A,B) = 25/55 \approx 0.45$ for the first set of predicates and $d(A,B) = 15/20 = 0.75$ for the second set of predicates.

Properties of the introduced functions

The introduced functions ρ and d have the following properties.

Property 1. For every objects ω_0 and ω_1 $\rho(\omega_0, \omega_1) \geq 0$ and $d(\omega_0, \omega_1) \geq 0$.

This property is an immediate consequence of definitions.

Property 2. If $\omega_0 = \omega_1$ then $\rho(\omega_0, \omega_1) = 0$ and $d(\omega_0, \omega_1) = 0$.

The proof is based on the fact that in such a case the descriptions objects ω_0 and ω_1 contain the same (up to the names of constants) formulas.

Property 3. If ω_0 is a proper subset of ω_1 then $\rho(\omega_0, \omega_1) > 0$ and $0 < d(\omega_0, \omega_1) < 1$.

Proof. As ω_0 is a proper subset of ω_1 so the elementary conjunction $A_0(\mathbf{x}_0)$ is a corollary (but not equivalent) of the formula $A_1(\mathbf{x}_1)$. Therefore $a_1 > a_0$, $a_1 > \Delta a'_{0,1} > 0$, $\Delta a'_{1,0} = 0$ and $\rho(\omega_0, \omega_1) = \Delta a'_{0,1} + \Delta a'_{1,0} = \Delta a'_{0,1} > 0$, $d(\omega_0, \omega_1) = (\Delta a'_{0,1} + \Delta a'_{1,0})/(a_0 + a_1) = \Delta a'_{0,1}/a_1 < 1$.

Property 4. If ω_0 and ω_1 have no common (up to the names of constants) formulas in their descriptions then $\rho(\omega_0, \omega_1) = (a_0 + a_1)$ and $d(\omega_0, \omega_1) = 1$.

Proof. As ω_0 and ω_1 have no common (up to the names of constants) formulas in their descriptions so $A_0(\mathbf{x}_0)$ and $A_1(\mathbf{x}_1)$ also have no common atomic formulas and $a'_{0,1} = a'_{1,0} = 0$. Hence $\rho(\omega_0, \omega_1) = \Delta a'_{0,1} + \Delta a'_{1,0} = (a_0 - a'_{0,1}) + (a_1 - a'_{1,0}) = a_0 + a_1$ and. $\rho_n(\omega_0, \omega_1) = \rho(\omega_0, \omega_1)/(a_0 + a_1) = (a_0 + a_1)/(a_0 + a_1) = 1$.

Property 5. If ω_0 and ω_1 have common (up to the names of constants) formulas in their descriptions but neither of them is a part of the other then $\rho(\omega_0, \omega_1) > 0$, $0 < d(\omega_0, \omega_1) < 1$.

This property is evident.

Theorem 1.

Function ρ defines a distance between objects. I.e. it satisfies the properties of distance:

1. for every objects ω_0 and ω_1 $\rho(\omega_0, \omega_1) \geq 0$;
2. for every objects ω_0 and ω_1 $\rho(\omega_0, \omega_1) = \rho(\omega_1, \omega_0)$;
3. $\omega_0 = \omega_1$ if and only if $\rho(\omega_0, \omega_1) = 0$;
4. triangle inequality is fulfilled for the function ρ, i.e. for every objects ω_1, ω_2 and ω_3 $\rho(\omega_1, \omega_2) + \rho(\omega_2, \omega_3) \geq \rho(\omega_1, \omega_3)$.

Proof. Points 1 and 2 are direct corollaries of the definition of ρ. Point 3 follows from the properties 2 – 5. Let's prove the triangle inequality.

Let $\omega_1, \omega_2, \omega_3$ be objects which descriptions have a_1, a_2, a_3 atomic formulas respectively. $a'_{1,2}$, $a'_{2,3}$, $a'_{3,1}$ are the numbers of atomic formulas contained respectively in maximal sub-formulas $A'_1(\mathbf{x}'_1)$, $A'_2(\mathbf{x}'_2)$, $A'_3(\mathbf{x}'_3)$ obtained while partial deduction of the respective sequents.

Let δ be the number of atomic formulas coinciding (up to the names of variables) simultaneously in $A'_1(x'_1)$, $A'_2(x'_2)$, $A'_3(x'_3)$; a^-_i be the number of atomic formulas which do not take part in partial derivations of $S(\omega_i) \Rightarrow \exists \mathbf{x}_{\neq} A_j(\mathbf{x}_j)$ $(j = 1, 2, 3, i \neq j)$; $a'_{1,2} = a''_{1,2} + \delta$, $a'_{2,3} = a''_{2,3} + \delta$, $a'_{3,1} = a''_{3,1} + \delta$.

Then

$$a_1 = a^-_1 + a''_{1,2} + a''_{1,3} + \delta,$$

$$a_2 = a^-_2 + a''_{1,2} + a''_{2,3} + \delta,$$

$$a_1 = a^{-}_1 + a''_{1,2} + a''_{1,3} + \delta,$$

$$a_2 = a^{-}_2 + a''_{1,2} + a''_{2,3} + \delta,$$

$$a_3 = a^{-}_3 + a''_{1,3} + a''_{2,3} + \delta.$$

$$\rho(\omega_1, \omega_2) + \rho(\omega_2, \omega_3) = (a_1 - a'_{1,2}) + (a_2 - a'_{1,2}) + (a_2 - a'_{2,3}) + (a_3 - a'_{2,3}) =$$

$$= (a^{-}_1 + a''_{1,3}) + (a^{-}_2 + a''_{2,3}) + (a^{-}_2 + a''_{1,2}) + (a^{-}_3 + a''_{1,3}).$$

$$\rho(\omega_1, \omega_3) = (a_1 - a'_{1,3}) + (a_3 - a'_{1,3}) = (a^{-}_1 + a''_{1,2}) + (a^{-}_3 + a''_{2,3}).$$

The reminder after subtraction of these expressions is $\rho(\omega_1, \omega_2) + \rho(\omega_2, \omega_3) - \rho(\omega_1, \omega_3) = 2a^{-}_2 + 2a''_{1,3} \geq 0$.

The triangle inequality is proved. The theorem is proved.

Theorem 2.

Function d does not define a distance. It does not satisfy the triangle inequality but satisfies the properties

1. for every objects ω_0 and ω_1 $d(\omega_0, \omega_1) \geq 0$;
2. for every objects ω_0 and ω_1 $d(\omega_0, \omega_1) = d(\omega_1, \omega_0)$;
3. $\omega_0 = \omega_1$ if and only if $d(\omega_0, \omega_1) = 0$.

Proof. Fulfillment of points 1 – 3 is a corollary of such properties for the function ρ .

Let's give an example of such objects ω_1, ω_2 and ω_3 that $d(\omega_1, \omega_2) + d(\omega_2, \omega_3) < d(\omega_1, \omega_3)$.

Let ω_1 be a proper part of ω_2 (hence $a'_{1,2} = a_1$) and $(a_2 - a'_{1,2}) = 0.1\, a_1$ (i.e. $a_2 = 1.1\, a_1$). Then $d(\omega_1, \omega_2) = (a_2 - a'_{1,2})/(a_1 + a_2) = 0.1\, a_1/ 1.1 a_1 = 0.1 / 1.1$.

Let also ω_1 has no common elements with ω_3. Then $d(\omega_1, \omega_3) = 1$.

Let all elements of ω_2 which does not belong ω_1 are elements of ω_3 and $a_3 - a'_{2,3} = a_1$ (i.e. $a_3 = a_1 + a'_{2,3} = 1.1\, a_1$). Then $d(\omega_2, \omega_3) = (a_3 - a'_{2,3})/(a_2 + a_3) = a_1 / 2.1\, a_1 = 1/ 2.1$.

$d(\omega_1, \omega_2) + d(\omega_2, \omega_3) = 0.1/1.1 + 1/ 2.1 \approx 0.09 + 0.043 = 0.133 < 1 = d(\omega_1, \omega_3)$.

Conclusion

The presence of a metric between objects involved in an Artificial Intelligence problem allows to state an earlier investigated object which is the mostly similar to the given for investigation one. Algorithms based on the principle "the nearest neighbor" are well-known in pattern recognition, particularly in the training of a neural network.

But usual metrics used in Artificial Intelligence problems are metrics in the fixed-dimensional spaces. This dimension equals to the number of features which describe an object. Usually an object is considered as a single indivisible unit and such a feature is its global characteristic.

If an object is considered as a set of its parts and the features describe properties of its elements and relations between them, then we can't map an object into a fixed-dimensional space. Such descriptions may be simulated by discrete features but the number of such features exponentially depends of the number of elements in the largest object under consideration [Russel, 2003].

Hence, the introduction of a metric for comparison of objects considered as a set of their elements is an important direction in the development of Artificial Intelligence.

Bibliography

[Kosovskaya, 2007] T.M. Kosovskaya. Proofs of the number of steps bounds for solving of some pattern recognition problems with logical description. In: Vestnik of St.Petersburg University. Ser. 1. 2007. No. 4. P. 82 – 90. (In Russian)

[Kosovskaya, 2009] T.M. Kosovskaya. Partial deduction of a predicate formula as an instrument for recognition of an object with incomplete description. In: Vestnik of St.Petersburg University. Ser. 10. 2009. No. 1. P. 77 – 87. (In Russian)

[Kosovskaya, 2011] Kosovskaya T. Discrete Artificial Intelligence Problems and Number of Steps of their Solution // International Journal "Information Theories and Applications", Vol. 18, Number 1, 2011. P. 93 – 99.

[Russel, 2003] S.J.Russel and P.Norvig. Artificial Intelligence. A Modern Approach. Pearson Education, Inc. 2003. ISBN 0-13-790396-2.

Authors' Information

Tatiana Kosovskaya *– Dr., Senior researcher, St. Petersburg Institute of Informatics and Automation of Russian Academy of Science, 14 line, 39, St. Petersburg, 199178, Russia; Professor of St. Petersburg State Marine Technical University, Lotsmanskaya ul., 3, St. Petersburg, 190008, Russia; Professor of St. Petersburg State University, University av., 28, Stary Petergof, St. Petersburg, 198504, Russia, e-mail: kosov@NK1022.spb.edu*

Major Fields of Scientific Research: Logical approach to artificial intelligence problems, theory of complexity of algorithms.

The paper is published with financial support of the project ITHEA XXI of the Institute of Information Theories and Applications FOI ITHEA (www.ithea.org) and the Association of Developers and Users of Intelligent Systems ADUIS Ukraine (www.aduis.com.ua).