

Algorithms for Algebraic and Logical Correction and Their Applications

Yu. I. Zhuravlev^a, S. V. Ablameiko^b, A. S. Biryukov^a, A. A. Dokukin^a,
V. V. Krasnoproshin^b, V. V. Obratsov^b, M. Yu. Romanov^a, and V. V. Ryazanov^a

^a *Dorodnitsyn Computing Center, Russian Academy of Sciences, ul. Vavilova 40, Moscow, 119333 Russia*

^b *Belarussian State University, pr. Nezavisimosti 4, Minsk, 220030 Belarus*

e-mail: zhuravlev@ccas.ru, ablameyko@bsu.by, asbirykov@yandex.ru, dalex@ccas.ru, krasnoproshin@bsu.by, Obratsov@cosmostv.by, mromanov@gmail.com, rvv@ccas.ru

Abstract—Practical algorithms for precedent-based recognition are considered that are based on the logical or algebraic correction of various heuristic algorithms or models of recognition. The recognition problem is solved in two stages. First, algorithms from a certain group are independently applied to recognition of an arbitrary object, and then an appropriate corrector is used to calculate the final collective solution. General concepts of the algebraic approach, descriptions of practical algorithms for logical and algebraic correction, and the results of comparison of these algorithms are presented.

Keywords: precedent-based recognition, classification, algebraic and logical corrections, correct algorithm.

DOI: 10.1134/S105466181002001X

INTRODUCTION

There are three main stages of development in the history of the theory of precedent-based recognition:

– the development of heuristic methods and algorithms—both universal ones that are designed for solving a wide spectrum of problems, and special methods and algorithms aimed at processing of a given type of information (for example, the algorithms “nearest neighbor,” “test algorithm,” “Kora algorithm,” and the Fischer discriminant);

– the development of parametric models and the search for the best algorithms within fixed models (for example, estimate evaluation models and neural network models).

– the development of approaches to solving recognition problems by sets of algorithms and by extension of the existing models.

The development of the third stage is primarily related to the algebraic theory of recognition, the fundamentals of which were developed by Yu. I. Zhuravlev in the late 1970s [1, 2]. He introduced the definition of a recognition algorithm as an algorithm that converts original information on the classes and descriptions of objects to be recognized into a matrix of binary answers to the questions on the membership of an object in each class (for simplicity, we assume that algorithms do not fail). Then one can set up a problem of recognition by a group of given algorithms in two stages by means of logical correction: first, each algorithm independently solves recognition problems for

objects from a given sample (i.e., it calculates a certain Boolean matrix), and then a certain Boolean function (a logical corrector) is applied to these Boolean matrices, which calculates the final classification. There are various approaches to the choice of the type of logical correctors and various methods for searching these correctors. Various methods of logical correction of recognition algorithms are described in [3–5].

Zhuravlev also showed that an arbitrary algorithm could be represented as a product (successive execution) of two algorithms: a recognition operator and a decision rule. The recognition algorithm converts original information and descriptions of objects to be recognized into a number matrix. The decision rule converts the number matrix into a binary matrix of final answers. He showed that the crucial role is played by the recognition operator. Over the sets of recognition operators, he defined algebraic operations that allow one to construct algebraic extensions of the given sets of operators in the form of operator polynomials. The products of these operator polynomials and a standard fixed threshold decision rule allow one to construct new recognition algorithms, including correct algorithms (that recognize elements of a given test sample without error). To date, necessary and sufficient conditions for the existence of correct algorithms and estimates for the degrees of operator polynomials have been obtained in algebraic recognition theory [1, 2, 6, 7]. In the present paper, we describe some practical algorithms for constructing algebraic and logical correctors that allow one to construct correct and quasi-correct algorithms and their approximations. We present examples of solving practical problems,

Received January 11, 2010

including the results of comparison with known recognition algorithms.

1. ALGEBRAIC AND LOGICAL CORRECTION OF THE SETS OF RECOGNITION ALGORITHMS

Following [1], we will use the following system of notations and concepts. Consider a standard recognition problem with l classes K_1, K_2, \dots, K_l . Denote by $P_i(S)$ a predicate " $S \in K_i$ " defined on the set $\{S\}$ of admissible objects. Suppose given a finite collection of admissible objects $\tilde{S} = \{S'_1, \dots, S'_q\}$. We will use the notation $\tilde{K}_i = \tilde{S} \cap K_i, i = 1, 2, \dots, l$, and assume that $\tilde{K}_i \neq \emptyset, i = 1, 2, \dots, l$.

A matrix $\|\alpha_{ij}\|_{q \times l}$, where $\alpha_{ij} = P_j(S'_i)$, is called an information matrix of the collection \tilde{S} with respect to the system of predicates P_1, \dots, P_l . The row $(\alpha_{i1}, \dots, \alpha_{il})$ is called the information vector of the object S'_i .

The recognition problem consists in finding the information matrix $\|\alpha_{ij}\|_{q \times l}$ by the initial information I on the classes and the sample \tilde{S} produced for recognition. A real practical algorithm $A(I, S'_1, \dots, S'_q) = \|\beta_{ij}\|_{q \times l}$ solves this problem with error: $\|\alpha_{ij}\|_{q \times l} \neq \|\beta_{ij}\|_{q \times l}$ (here $\beta_{ij}, \beta_{ij} \in \{0, 1\}$, is the value of the predicate P_j on the object S'_i calculated by algorithm A). For simplicity, we restrict ourselves to the case of absence of recognition failures.

Suppose there is a set $\{A\}$ of algorithms that convert $(I, S'_1, S'_q), I \in \{I\}, \tilde{S} \subseteq \{S'\}$, into matrices $A(I, S'_1, S'_q) = \|\beta_{ij}\|_{q \times l}$. The algorithm A is said to be correct for the recognition problem if the equality $A(I, S'_1, \dots, S'_q) = \|\alpha_{ij}\|_{q \times l}$ holds.

In [1], it was proved that each algorithm $A \in \{A\}$ can be represented as a successive execution (product) of algorithms B and C ($A = BC$), where $B(I, S'_1, \dots, S'_q) = \|\alpha_{ij}\|_{q \times l}$, α_{ij} are real numbers, $C(\|\alpha_{ij}\|_{q \times l}) = \|\beta_{ij}\|_{q \times l}$, $\beta_{ij} \in \{0, 1\}$, $B = B(A)$, and $C = C(A)$.

The set $\{A\}$ generates the sets $\{B\}$ and $\{C\}$. The elements of $\{B\}$ are called recognition operators, and the elements of $\{C\}$, decision rules. The number matrices $B(I, S'_1, \dots, S'_q) = \|a_{ij}\|_{q \times l}$ are called matrices of estimates of objects for classes, or estimate matrices for short.

A decision rule C is said to be correct on $\{\tilde{S}\}$ if, for any finite collection $\tilde{S} \subseteq \{S'\}$, there exists at least one

number matrix $\|a_{ij}\|_{q \times l}$ such that $C(\|a_{ij}\|_{q \times l}) = \|\alpha_{ij}\|_{q \times l}$.

In [1], Zhuravlev introduced the operations of addition, multiplication, and multiplication by a scalar on the set $\{B\}$ of recognition operators, which allow one to construct operators of the form $B' = \sum b_i B_{i1} \cdot \dots \cdot B_{ik}$, where $B'(I, S'_1, \dots, S'_q) = \|a'_{uv}\|$, $a'_{uv} = \sum b_i a_{uv}^{i1} \cdot \dots \cdot a_{uv}^{ik}$, if $B_{ik}(I, S'_1, \dots, S'_q) = \|a_{uv}^{ik}\|_{q \times l}$. Then we can consider the algorithms

$$A = B'C^*, \quad (1)$$

where C^* is the correct decision rule and the operators B' are constructed on the basis of the operators of parametric recognition models. The construction of algorithms of type (1) is called an algebraic correction of the algorithms $\{A\}$.

Suppose there is a group of n recognition algorithms $A^1, A^2, \dots, A^n, A^k(I, S'_1, \dots, S'_q) = \|\beta_{ij}^k\|_{q \times l}, k = 1, 2, \dots, n$. The sets of matrices $\{\|\beta_{ij}^k\|_{q \times l}, k = 1, 2, \dots, n, \|\alpha_{ij}\|_{q \times l}\}$ define l Boolean functions $f_j(y_1, y_2, \dots, y_n), j = 1, 2, \dots, l$, that are not everywhere defined. Here, $f_j(\beta_{ij}^1, \beta_{ij}^2, \dots, \beta_{ij}^n) = \alpha_{ij}, i = 1, 2, \dots, q$, and the function $f_j(y_1, y_2, \dots, y_n)$ is not defined on the other $2^n - q$ sets. The main problem consists in defining a function over the entire discrete unit cube E^n so that additional meaningfully justified conditions that guarantee a certain unique and optimal choice of such a function are maximally satisfied. These functions are called logical correctors.

In the present paper, we present the results of investigations related to the practical construction of algebraic and logical correctors for recognition algorithms. Some of these correctors are integrated into the software system RECOGNITION for intelligent data analysis and recognition [8]. We illustrate the results by comparing them with the example of a number of applied problems.

2. POLYNOMIAL CORRECTION ON THE BASIS OF MAXIMUM-HEIGHT TERMS

As mentioned above, the proof of the existence of a correct polynomial over a family of estimate evaluation algorithms (EEAs) is one of the main results of algebraic theory of recognition [2]. However, the practical application of this theorem to the construction of polynomial correctors gives rise to a few complicated problems. First, one should construct a collection of basic algorithms, which should be as stable as possible. Next, one should minimize the complexity of the polynomial, i.e., its degree and the number of terms.

For solving these problems, a new characteristic of the algorithm, the so-called height, was introduced [9]:

$$\min_{(i,j) \in M_1} \Gamma_j(S^i) - \max_{(u,v) \in M_0} \Gamma_v(S^u),$$

where M_1 is the set of correct (object, class) pairs, i.e., pairs in which an object belongs to a class, and M_0 is the set of incorrect pairs.

In [10, 11], a collection of methods was developed for the optimization of the height of algorithms in the EEA family with variable thresholds of the proximity functions. In particular, the authors described an exact method and a collection of fast approximate algorithms whose comparative test was carried out on the collection of specially generated samples [11].

Given a collection of basic algorithms B_1, \dots, B_q (without loss of generality, we assume that the basic collection contains one algorithm for each test object S_1, \dots, S_q), we can easily construct a correct polynomial

$$\text{mial} \left(\sum_{i=1}^q c_i B_i^{k_i} \right) C.$$

To minimize the degree of this polynomial, we suggest using a gradient scheme. Let us write out the optimization problem formally. For simplicity, consider the case when classes do not intersect (although this is not essential). The optimization is carried out with respect to the variables $k_1, \dots, k_q \in \mathbb{Z}, k_i > 0$. We should minimize the quantity $\sum_i k_i$ (or $\max_i k_i$); here, the fol-

lowing inequalities should be satisfied, which guarantee the correctness of the polynomial, namely,

$$\left\| \sum_{i=1}^q c_i B_i^{k_i} C(S_j) \right\|_{K(S_j)} > \left\| \sum_{i=1}^q c_i B_i^{k_i} C(S_j) \right\|_d, \\ j = 1, \dots, q, \quad d \neq K(S_j),$$

where $K(S_j)$ is the class to which the j th test object belongs, $\|B(S)\|_d$ is the estimate of object S for the class K_q by the operator B . We will construct the optimization method according to the gradient principle.

(1) Take a first-degree polynomial as the initial solution.

(2) If the correctness inequalities are satisfied, then stop; otherwise go to step 3.

(3) In the case of the functional $\sum_i k_i$, increase by

one the degree of the term whose increase leads to the correction of the maximal number of incorrect inequalities, and then go to step 2.

In the case of the functional $\max_i k_i$, we increase by one the degree of the term whose increase leads to the correction of the maximal number of incorrect in-

qualities among all the terms whose degree is less than $\max_i k_i + 1$, and then go to step 2.

It is easily seen that the method converges.

Originally, this scheme was developed for a specially constructed collection of EEAs; however, it can easily be generalized for an arbitrary number of algorithms that form a basis for a given test sample.

Nevertheless, the application of this scheme is complicated by the high complexity of constructing the basic collection of algorithms by known methods (it is assumed that a collection of EEAs of maximum height is used as a basis). Actually, we decided to give up the correctness and use a collection of "good" terms, whose height can easily (i.e., by fast approximate methods) be made positive. When solving a large number of applied problems, we noted that algorithms of large height on a given subsample tend to possess good recognition quality in the neighborhood of the objects of this subsample. As a result, we obtained an EEA-polynomial method, which constructs a second-degree polynomial over an EEA of the form

$$\left(\sum_{i=1}^q B_i D_{i,q} \right) C,$$

where $D_{i,q}$ is an EEA with a special proximity function that is inversely proportional to the Euclidean distance to the central object of the basic EEA.

3. CONSTRUCTION OF A POWER-LAW ALGEBRAIC CORRECTOR

Correctors of this type were introduced in [1, 2]. Let a certain model of algorithms $\{A\}$ be given. Denote the corresponding model of recognition operators by $\{B\}$, and fix a decision rule c in the class of threshold decision rules $C(c_0, c_1)$, which, as is known, are correct for $c_0 \leq c_1$ (provided that recognition failures are eliminated). Then, power-law algebraic correctors in the general form can be represented as

$$c \circ \left(\sum_{(i,j)} \gamma_{ij} B_{ij}^t \right), \tag{2}$$

where γ_{ij} is a real number (R), t is a natural number (N), and the index (i, j) runs over a set of values that significantly depends on the problem to be solved. A pair (i, j) is chosen from the set $I = \{1, 2, \dots, q\} \times \{1, 2, \dots, l\}$. The degree t of the corresponding recognition operator B_{ij} is determined componentwise, i.e., in a linear algebra that is commutative with respect to multiplication. Denote the matrices $B_{ij} = \|b_{uv}^{ij}\|_{q \times l}$, where the indices (u, v) belong to $\{1, 2, \dots, q\} \times \{1, 2, \dots, l\}$ and run over the entire set of admissible values for every fixed $(i, j) \in I$.

Model (2) is correct under not-too-stringent conditions imposed on the results of recognition in $\{B\}$. The number of recognition operators in (2) is redundant in the general case. However, it is known precisely that it is senseless to use more than $q \cdot l$ operators when constructing (2), because the basis of the space R^{ql} consists of $q \cdot l$ matrices. We will consider this form of the model to be canonical and present the above-mentioned correctness conditions for it.

We will need one more notation. The set of indices I for a given information matrix $\|\alpha_{ij}\|_{q \times l}$ is obviously divided into two subsets

$$M_0 = \{(i, j) | (i, j) \in I, \alpha_{ij} = 0\}$$

$$\text{and } M_1 = \{(i, j) | (i, j) \in I, \alpha_{ij} = 1\}.$$

Then, model (2) is correct provided that

$$\forall (i_0, j_0) \in M_1 \exists B_{ij} \in \{B\} (b_{i_0 j_0}^{ij} > \max_{(u, v) \in M_0} b_{uv}^{ij}); \quad (3)$$

moreover, in this case

$$t \geq [(\ln c_1 - \ln c_0) \times (\min_{(i, j) \in M_1} (\ln b_{i_0 j_0}^{ij} - \ln \max_{(u, v) \in M_0} b_{uv}^{ij}))^{-1}] + 1, \quad (4)$$

where $[x]$ is the integral part of the number $x \in R$ and $(i_0, j_0) \in M_1$ is chosen from condition (3).

Of course, conditions (3) and (4) are sufficient conditions. A specific feature of such a canonical model is that one should construct operators satisfying condition (3). The model constructed has drawbacks. It may become unstable for sufficiently large t . Of course, condition (3) is more constructive than the condition for constructing the basis of the space R^{ql} . A certain analog of this condition was considered in [2, 12]. In these papers, the authors introduced the condition of quasi-completeness and showed that models $\{A\}$ satisfying this condition exist. Obviously, these models $\{A\}$ satisfy condition (3). Nevertheless, in spite of their feasibility, the condition of quasi-completeness and condition (3) are redundant.

Below, we will show that a small modification of the power-law corrector can make condition (3) less stringent, while the degree t can be set equal to 1.

To this end, we introduce a matrix-inversion operation

$$\forall R = \|r_{ij}\| \in R^{ql} (r_{ij} \neq 0) (R^{-1} = \|r_{ij}^{-1}\|) \\ \text{or } E = RR^{-1}$$

in the space R^{ql} ; here E is the unit matrix with respect to the commutative inversion of the space R^{ql} .

We modify model (2) as follows:

$$c \circ \left(\sum_{(i, j)} \gamma_{ij} \tilde{B}_{ij}^t(i_0, j_0) \right), \quad (5)$$

and define the operator $\tilde{B}_{ij}(i_0, j_0)$ using the operation of inversion (the index (i_0, j_0) belongs to I and is taken from (3))

$$\tilde{B}_{ij}(i_0, j_0) = (E + \beta |b_{i_0 j_0}^{ij} E - B_{ij}|)^{-1}. \quad (6)$$

It is easily seen that, under condition (3) for the operator B_{ij} , the matrix $\tilde{B}_{ij}(i_0, j_0) = \|\tilde{b}_{i_0 j_0}^{ij}\|$ possesses the following property:

$$\begin{cases} \tilde{b}_{i_0 j_0}^{ij} = 1, \\ \tilde{b}_{uv}^{ij} < 1 \quad \forall (u, v) \neq (i_0, j_0). \end{cases} \quad (7)$$

Hence, condition (4) for this operator is rewritten as

$$t \geq [(\ln c_1 - \ln c_0) \times (\min_{(i, j) \in M_1} (\ln(1 + \min_{(u, v) \neq (i_0, j_0)} \beta |b_{i_0 j_0}^{ij} - b_{uv}^{ij}|))^{-1})] + 1.$$

From the last inequality, we can easily find that, for

$$\beta > \left(\frac{c_1}{c_0} - 1 \right) \left(\min_{\substack{(i, j) \in M_1 \\ (u, v) \neq (i_0, j_0)}} |b_{i_0 j_0}^{ij} - b_{uv}^{ij}| \right)^{-1},$$

model (5) is correct for $t = 1$.

It is easily seen that condition (7) for the recognition operator $\tilde{B}_{ij}(i_0, j_0)$ (6) is satisfied not only when the operator B_{ij} satisfies (3). In fact, this condition is easily generalized to

$$\forall (i_0, j_0) \in M_1 \exists B_{ij} \in \{B\} (b_{i_0 j_0}^{ij} \neq \max_{(u, v) \in M_0} b_{uv}^{ij}), \quad (8)$$

which also implies (7) for the operator (6).

Thus, in model (5) one can get rid of the main drawbacks of model (2). Condition (8) imposes much less stringent conditions on the model of algorithms $\{A\}$. Moreover, while in the first model it was relevant to speak of the instability of the results, in model (5) the correctness is achieved by the choice of an appropriate parameter β .

4. CONSTRUCTION OF A SECOND-DEGREE ALGEBRAIC CORRECTOR FOR A GIVEN COLLECTION OF ALGORITHMS

Suppose there is a certain set of recognition algorithms A^1, A^2, \dots, A^n and a test sample $\tilde{S}^n = \{S_1^t, \dots, S_q^t\}$ with the information matrix $\|\alpha_{ij}\|_{q \times l}$. We assume that the results of recognition of the objects A^1, A^2, \dots, A^n are defined in the form of estimates for classes: $A^t(S^t) =$

$(\Gamma_{i1}^t, \Gamma_{i2}^t, \dots, \Gamma_{il}^t)$, $A^t(\tilde{S}) = \|\Gamma_{ij}^t\|_{q \times l}$, $\Gamma_{ij}^t \in [0, 1]$. We will seek a second-degree algebraic corrector

$$A = x_0 + \sum_{t=1}^n x_t A^t + \sum_{t=n+1}^N x_t A^{\nu(t)} A^{\mu(t)}, \quad (9)$$

where $N = n + 1 + n(n + 1)/2$,

$$\nu(t), \mu(t) = 1, 2, \dots, n.$$

As a decision rule, we take the operator C^* : $(\Gamma_1, \Gamma_2, \dots, \Gamma_l) \rightarrow (\beta_1, \beta_2, \dots, \beta_l)$, where $\beta_j = \begin{cases} 1, & \Gamma_j \geq 0.5, \\ 0, & \Gamma_j < 0.5. \end{cases}$

For a given test sample, $B^t(I, S_1^t, \dots, S_q^t) = \|\Gamma_{ij}\|_{q \times l} = \left\| x_0 + \sum_{t=1}^n x_t \Gamma_{ij}^t + \sum_{t=n+1}^N x_t \Gamma_{ij}^{\nu(t)} \Gamma_{ij}^{\mu(t)} \right\|_{q \times l}$. The problem consists in finding unknown values of the parameters x_t for which the matrix $\|\Gamma_{ij}\|_{q \times l}$ is the best approximation to the information matrix $\|\alpha_{ij}\|_{q \times l}$. After an obvious change of notation

$$\Gamma_{ij} = x_0 + \sum_{t=1}^T x_t a_{ij}^t, \quad (10)$$

we solve the problem of finding an algebraic corrector of the form (9) by the least squares method, as a problem of minimization of the quadratic residual between the information matrix $\|\alpha_{ij}\|_{q \times l}$ and $\|\Gamma_{ij}\|_{q \times l}$.

To guarantee the stability of the algebraic corrector, it seems expedient to seek second-degree polynomials that contain a small number of terms. To this end, we should eliminate the least informative terms from the sums $\Gamma_{ij} = x_0 + \sum_{t=1}^T x_t a_{ij}^t$ and take only one term from

among a few strongly correlated ones.

To construct correctors of the form of (10) with a small number of terms, there exists the following multistep heuristic procedure.

Normalize the matrices $\|a_{ij}^t\|$, $t = 1, 2, \dots, T$, and estimate the informativeness of the matrices $I_0 = \{1, 2, \dots, T\}$, $\varphi_0(x_0^*, x_1^*, \dots, x_T^*) =$

$$\arg \min_{x_0, x_1, \dots, x_T} \sum_{i=1}^q \sum_{j=1}^l \left(x_0 + \sum_{t=1}^T x_t a_{ij}^t - \alpha_{ij} \right)^2.$$

At step $\nu = 0, 1, 2, \dots$, there is a set $I_\nu \subseteq \{1, 2, \dots, T\}$, a residual

$$\varphi_\nu(x_0^*, x_{\tau_1}^*, \dots, x_{\tau_k}^*) \arg \min_{x_0, x_{\tau_1}, \dots, x_{\tau_k}} \sum_{i=1}^q \sum_{j=1}^l \left(x_0 + \sum_{t \in I_\nu} x_t a_{ij}^t - \alpha_{ij} \right)^2 \quad (11)$$

and a solution $(x_0^*, x_{\tau_1}^*, \dots, x_{\tau_k}^*)$, corresponding to the set I_ν , such that $|\varphi_0 - \varphi_\nu| \leq \varepsilon$.

Step $\nu + 1$.

(1) Choice of a set of weakly correlated matrices $\|a_{ij}^t\|$, $t \in I_{\nu+1} \subseteq I_0$, and solution of problem (11) with respect to the parameters corresponding to $I_{\nu+1}$.

(2) If $|\varphi_0 - \varphi_{\nu+1}| > \varepsilon$, it is assumed that a solution $(x_0^*, x_{\tau_1}^*, \dots, x_{\tau_k}^*)$ corresponding to the set of indices I_ν is found. Otherwise go to the next step.

Remark. The sequence I_ν , $\nu = 1, 2, \dots$, is constructed so that the algorithm is finite.

The normalization, estimation of informativeness, correlation, and the choice of matrices are performed as follows. Instead of the matrices $\|a_{ij}^t\|$, $t = 1, 2, \dots, T$, we take matrices $\|\lambda_t a_{ij}^t\|$, $t = 1, 2, \dots, T$, that minimize

the criterion $\phi \|a_{ij}^t\| = \sum_{i=1}^q \sum_{j=1}^l (\lambda_t a_{ij}^t - \alpha_{ij})^2$, $t = 1, 2, \dots,$

T , i.e., $\lambda_t = \left(\sum_{i=1}^q \sum_{j=1}^l (a_{ij}^t \alpha_{ij}) \right) / \left(\sum_{i=1}^q \sum_{j=1}^l (a_{ij}^t)^2 \right)$. We

assume that the closer the matrix $\|a_{ij}^t\|$ is to the information matrix of the sample, the more informative it is.

Next, we assume that the matrices $\|a_{ij}^t\|$, $t = 1, 2, \dots, T$, are already ordered in increasing quality criterion

$\phi \|a_{ij}^t\| = \sum_{i=1}^q \sum_{j=1}^l (a_{ij}^t - \alpha_{ij})^2$, $t = 1, 2, \dots, T$ (or $\phi \|a_{ij}^t\| =$

$\sum_{i=1}^q \sum_{j=1}^l |a_{ij}^t - \alpha_{ij}|$, $t = 1, 2, \dots, T$). Suppose there is a certain metric ρ for number matrices of dimension $q \times l$

and that the parameter δ is fixed.

Then, we find an informative subset $\|b_{ij}^t\|$, $t = 1, 2, \dots, \tau$, of different matrices of the set $\tilde{A}^1 = \{\|a_{ij}^t\|, t = 1, 2, \dots, T\}$ by the following rule.

Set $\|b_{ij}^1\| = \|a_{ij}^1\|$. Eliminate from the set $\|a_{ij}^t\|$, $t = 1, 2, \dots, T$, all the matrices such that $\rho(\|b_{ij}^1\|, \|a_{ij}^t\|) \leq \delta$, $t = 1, 2, \dots, T$. Enumerate the remaining matrices

again, and denote them as $\tilde{A}^2 = \{\|a'_{ij}\|, t = 1, 2, \dots, T\}$. Set $\|b_{ij}^2\| = \|a_{ij}^1\|$. Eliminate from \tilde{A}^2 all the elements such that $\rho(\|b_{ij}^2\|, \|a'_{ij}\|) \leq \delta, \|a'_{ij}\| \in \tilde{A}^2$. The process is continued until it reaches the situation $\tilde{A}^k = \emptyset$. It is convenient to choose the parameter δ as $\delta = d \left(\sum_{i,j=1}^{q,l} \alpha_{ij} \right)$, where $0 \leq d \leq 1$ is a control parameter.

$$\bar{U}(B^*) = L(\{B_1^{x_1} B_2^{x_2} \dots B_s^{x_s} : B_i \in B^*, x_i \in R, i = \overline{1, s}, 1 \leq s \leq k\}).$$

The elements of this set are called generalized polynomials over B^* , and the terms of this polynomial are called generalized monomials.

Similarly, the set $\bar{U}(A^*)$ of algorithms $A = BC$ such that $B \in \bar{U}(B^*)$ is called a generalized algebraic closure of A^* .

Let a set of recognition operators (ROs) $B^* = \{B_1, B_2, \dots, B_n\}$ be given. The degree of the generalized monomial $B_1^{x_1} B_2^{x_2} \dots B_s^{x_s}$ is $x_1 + x_2 + \dots + x_s$. The degree of the generalized polynomial $B \in \bar{U}(B^*)$ is the maximum of the degrees of its constituent monomials. The degree of the algorithm $A = BC$ from the generalized algebraic closure is the degree of the generalized polynomial B .

Any polynomial over ROs is also a generalized polynomial. The degree of the polynomial coincides with the degree obtained when considering this polynomial as a generalized one. All the main results remain valid for the concepts introduced.

Suppose that, for the recognition problem, a collection of ROs B_k is given that construct estimate matrices $B_k(Z) = \|\Gamma_k^{uv}\|_{q \times l}$, and the operator C is defined by a standard threshold decision rule, $C(\Gamma_k^{uv}) = \{0, \Gamma_k^{uv} \leq c_1; 1, \Gamma_k^{uv} \geq c_2; \Delta, c_1 < \Gamma_k^{uv} < c_2\}$, which is defined by the parameters c_1 and c_2 .

Remark. Without loss of generality, we will assume that the parameters c_1 and c_2 satisfy the condition $c_1 + c_2 = 1$. Otherwise, these parameters should be divided by the sum $c_1 + c_2$.

Let us divide the elements of the answer matrix into two sets, $\|\beta_{ij}\|_{q \times l} = M_0 \cup M_1$, where $M_i = \{(u, v) : \beta_{uv} = i\}$, q is the number of objects, and l is the number of classes. A pair $(u, v) \in M_1$ such that $|\Gamma_{uv}(B) - \Gamma_{ij}(B)|$ for any $(i, j) \in M_0$ is said to be a marked pair in B . Denote the set of marked pairs in B by $M(B)$. In fact, this means that marked pairs are distinguished from all the pairs (i, j) such that the object S^i does not belong to the class K_j .

5. CONSTRUCTION OF A GENERALIZED POLYNOMIAL ALGEBRAIC CORRECTOR OF MINIMAL DEGREE

In [13, 14], a generalization of polynomial algebraic correctors is developed in which the degree of the recognition operator is assumed to be a real number rather than an integer. Below we give the main definitions and the description of the algorithm for the practical construction of these correctors.

The generalized algebraic closure of a set of recognition operators $B^* = \{B_1, B_2, \dots, B_n\}$ is the set

A collection of ROs $\{B_k\}$ is said to be regular if it satisfies the following conditions.

- Each operator is admissible; i.e., there exists at least one pair $(u, v) \in M_1$ that is marked in B .
- Each operator is normalized; i.e., if (u, v) is marked in B , then $\Gamma_k^{uv} \geq 1$; otherwise $\Gamma_k^{uv} < 1$.
- The system $\{B_k\}$ is basic for Z ; i.e., $M_1 = \bigcup_{k=1}^n M(B_k)$.

Otherwise, the collection of operators is called irregular.

Thus, operators that are not admissible are eliminated from the original collection, and the remaining operators are normalized. The normalization coefficients of the operators are defined by

$$\Gamma(B) = [\Gamma_{\min}^1(B)]^{-1},$$

$$\text{where } \Gamma_{\max}^0(B) = \max_{(i,j) \in M_0} |\Gamma_{ij}(B)|,$$

$$\Gamma_{\min}^1(B) = \min_{(i,j) \in M(B)} |\Gamma_{ij}(B)|.$$

If the system obtained is basic, then the resulting polynomial operator is correct.

To determine the degrees of operators in a polynomial of the form $A = \left\{ \sum_{k=1}^n B_k^{x_k} \right\} \circ C(c_1, c_2)$, we intro-

duce the change of variables $y_k = e^{x_k}, \gamma_k^{uv} = \ln \Gamma_k^{uv}$,

$\phi_{uv}(\tilde{y}) = \sum_{k=1}^n y_k^{\gamma_k^{uv}}$ and solve the following optimization problem:

$$U = \left\{ \begin{array}{l} \phi_{uv}(\tilde{y}) \leq c_1, \text{ for any } (u, v) \in M_0; \\ \tilde{y} |\phi_{uv}(\tilde{y}) \geq c_2, \text{ for any } (u, v) \in M_1; \\ 0 \leq y_k \leq M = ql, \text{ for any } k = \overline{1, n} \end{array} \right\};$$

$$R = \{\tilde{y} \in U, \max_{k=\overline{1, n}} y_k \rightarrow \min\}.$$

For any nonempty set of indices $I = \{i_1, \dots, i_m\} \subseteq \{1, \dots, n\}$, consider the auxiliary optimization problem $R_I = \{\tilde{y} \in U_I, y_{i_1} \rightarrow \min\}$ with the set of constraints $U_I = \{\tilde{y} \in U; y_{i_1} = y_{i_2} = \dots = y_{i_m}; y_{i_1} \geq y_k, k \notin I\}$, which is effectively solved by the method of linearization.

For problem R_I , we introduce the following notation:

$$F_I = \max_{k=\overline{1, n}} \tilde{y}(R_I)|_k; \quad G_I = \tilde{y}(R_I)|_{i_1}.$$

Let us describe an algorithm for searching solutions that employs auxiliary problems.

- Consider an ordered set of collections $H = \{I: I \subseteq \{1, 2, \dots, n\}; I \neq \emptyset\}$, which is ordered so that collections with lower cardinality precede the collections with higher cardinality. Assign the initial value $F_{\min} = +\infty$.
- Take the first unconsidered collection I in H and solve problem R_I .
- Calculate the values of G_I and F_I . If $F_I < F_{\min}$, then assign $F_{\min} = F_I$.
- If $G_I \geq F_{\min}$, then remove from H all the collections that contain a subcollection I .
- If not all the collections are considered in H , go to step 2.
- The sought solution to problem R is a solution to the auxiliary problem defined by the collection I^* for which $F_{I^*} = F_{\min}$.

If the collection of ROs is not basic, then, by solving the auxiliary problem, one should find a solution that gives the least error; i.e., one should solve the problem under the set of constraints

$$U_\varepsilon = \left\{ \begin{array}{l} \phi_{uv}(\tilde{y}) \leq c_1 + \varepsilon, \text{ for any } (u, v) \in M_0; \\ \tilde{y}|\phi_{uv}(\tilde{y}) \geq c_2 - \varepsilon, \text{ for any } (u, v) \in M_1; \\ 0 \leq y_k \leq M = ql, \text{ for any } k = \overline{1, n} \end{array} \right\},$$

and find a minimal number ε for which the system is consistent.

6. CONSTRUCTION OF A LOGICAL CORRECTOR ON THE BASIS OF THE POTENTIALS OF TEST OBJECTS

The algorithm is a modification of the general approach of [3]. For an object S to be recognized, one calculates its estimate for each class according to the

$$\text{formula } \Gamma_i(S) = \frac{1}{|\overline{K}_i|} \sum_{t \in \overline{K}_i} \Phi_t(S'_t, S), i = 1, 2, \dots, l.$$

The quantity $\Phi_t(S'_t, S)$ is called the value of the potential function for the pair S'_t and S . We will calculate this quantity by one of the following methods:

$$(a) \quad \Phi_t(S'_t, S) = \begin{cases} 1, & |\{\beta'_j(S) \geq \beta'_j(S'_t), j = 1, 2, \dots, n\}|/n \geq \sigma, \\ 0, & \text{otherwise.} \end{cases}$$

or

$$(b) \quad \Phi_t(S'_t, S) = \{\text{the number of satisfied inequalities } \beta'_j(S) \geq \beta'_j(S'_t), j = 1, 2, \dots, n\}.$$

$$\text{Here } \beta'_j(S'_t) = \beta_{it}^j.$$

The first method is called the monotonic potential correction (it depends on the control parameter $\delta, 0 < \delta \leq 1$), and the second method is called the weakly monotonic potential correction.

A logical corrector, as a set of functions $f_j(y_1, y_2, \dots, y_n), j = 1, 2, \dots, l$, is defined as follows: $f_j(\beta_i^1(S),$

$$\beta_i^2(S), \dots, \beta_i^n(S)) = \begin{cases} 1, & \Gamma_i(S) \geq \Gamma_j(S), j = 1, \dots, l, \\ 0, & \text{otherwise} \end{cases}.$$

7. EXPERIMENTAL INVESTIGATION OF TWO-LEVEL RECOGNITION SCHEMES WITH A LOGICAL CORRECTOR

We investigated an approach to the construction of recognition algorithms that is based on two-level recognition schemes with a corrector and employs the synthesis of reliable schemes consisting of unreliably operating units. Two-level algorithms are, in a sense, invariant with respect to the problem because, on the first level, one can combine algorithms from different models, thus increasing the reliability of the recognition procedure.

According to the general principles of constructing two-level schemes, we can distinguish two logically completed stages of the experiment [1]. At the first stage, one solves test problems by algorithms from the corrected collection, determines the accuracy of the algorithms, and reveals the degree of statistical dependence between them. As a measure of similarity between algorithms, one uses a sample correlation coefficient between the events that consist of correct recognition of a random object.

Using the results of a fixed collection of algorithms, one compiles the table

$$\|A|\beta\| = \left\| \begin{array}{c|c} \alpha_{11} \dots \alpha_{1n} & \beta_1 \\ \dots & \dots \\ \alpha_{m1} \dots \alpha_{mn} & \beta_m \end{array} \right\|,$$

where $\alpha_{ij}, 1 \leq i \leq q, 1 \leq j \leq n$, are elements of a classification vector that correspond to the results of recognition of objects from a test sample by the algorithms and β_i are elements of the information vector that correspond to the true classification of objects.

Table 1

| 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|-----------|------|------|------|------|------|------|------|------|------|------|
| EA1 | NN1 | EA2 | NN2 | EA3 | EA4 | T | SS1 | PF1 | PF2 | SS2 |
| Problem 1 | | | | | | | | | | |
| 0.74 | 0.74 | 0.77 | 0.73 | 0.64 | 0.81 | 0.80 | 0.62 | 0.72 | 0.69 | 0.79 |

At the second stage, one analyzes the properties of the corrector being tested. The main focus is placed on the study of the effect of the statistical dependence between algorithms on the quality of recognition.

The problems for the experiments were chosen with regard to the difference in the subject fields and the physical bases, as well as to the complexity of their solution (the accuracy of algorithms for problems do not satisfy the given requirements).

When choosing models of recognition algorithms, we took into account their universality, familiarity, difference in construction principles, and simplicity of implementation. By these principles, we chose algorithm models of the following types: (1) estimate evaluation (EA), (2) nearest neighbors (NN), (3) potential functions (PF), (4) separating surfaces (SS), and (5) taxonomy (T).

In each model, we constructed one or several algorithms for a specific problem; these algorithms differed in the set of optimizing parameters and in the method of their determination. The training of the recognition algorithms and the checking of the accuracy of their operation (for problems with a large volume of information) were carried out by the training and test samples, respectively. The latter samples were formed by a random partition of the entire set of objects of the problem into two disjoint subsets. For problems with a small number of objects, we used the sliding exam method. As a result of the first stage, we constructed tables $\|A|\beta\|$ by the results of recognition of test objects.

To carry out the experiments, the logical corrector (the corresponding Boolean function) minimized the empirical risk in the class of all Boolean functions. Then we obtained a simple definition of a partial Boolean function $\|A|\beta\|$.

The *first problem* consisted in the classification of the underlying surfaces of the Earth by aerospace spectroscopy data [15, 16]. The set of objects of the problem consisted of 300 spectrograms of the surface

water of two oceans (Atlantic and Pacific). Among the features of the problem were the physical homogeneity of the characteristics that describe the object of recognition and the fact that the classes are not sharply different (underlying surfaces are of the same type).

The set of objects is randomly divided into two samples (disjoint sets): a training sample, consisting of 100 objects, and a test sample, consisting of 200 objects. The following heuristics were used to solve the problem: EA1–EA4, PF1, PF2, NN1, NN2, SS1, SS2, and T. After the stage of training, the accuracy of algorithms was estimated as the rate of correct recognition of objects from the test sample. The accuracy of the algorithms as applied to Problem 1 is illustrated in Table 1.

The *second problem* consisted in the recognition of oil-bearingness of some oilfields by the results of the analysis of nearby water-bearing strata. There were 64 objects in total. Specific features of the problem are the qualitative difference in the character of description of the objects and small volume of information.

The training and the accuracy estimate of the algorithms were performed by the sliding exam method. This approach is the most reliable under conditions of small volumes of information. Seven algorithms were used in the experiments on this problem: EA1–EA3, NN, PF, SS, and T (see Table 2).

By the results of the algorithms, we compiled tables of pair correlation coefficients on each test sample. For problem 1, the results obtained are shown in Table 3. As expected, the strongest correlation is observed between algorithms of the same type. For example, in Table 3, strongly related blocks form variants of the estimate evaluation, nearest neighbors, and potential function algorithms. At the same time, the correlation is insignificant for the heuristics PF2 and T from different blocks.

Thus, the execution of the first stage of the experiment makes it possible to form different sets of correctable heuristics and the corresponding tables $\|A|\beta\|$ for studying the properties of a logical corrector.

Using the data of Tables 3 and 4 for all the problems, we formed about 100 sets of algorithms and the corresponding tables $\|A|\beta\|$ with different quantitative and qualitative characteristics. The training and the accuracy estimate of the correctors were performed with the use of the tables $\|A|\beta\|$ by the sliding exam method. We investigated the possibility of correcting

Table 2

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|-----------|------|------|------|------|------|------|
| NN | SS | T | EA1 | EA2 | EA3 | SS |
| Problem 2 | | | | | | |
| 0.67 | 0.74 | 0.70 | 0.67 | 0.60 | 0.66 | 0.70 |

Table 3

| | | | | | | | | | | | |
|----|---|------|------|----------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|
| | 1 | 3 | 5 | 8 | 2 | 4 | 6 | 9 | 10 | 7 | 11 |
| 1 | 1 | 0.63 | 0.81 | 0.68 | 0.40 | 0.44 | 0.60 | 0.20 | 0.40 | 0.23 | 0.47 |
| 3 | | 1 | 0.50 | 0.52 | 0.38 | 0.38 | 0.50 | 0.32 | 0.20 | 0.26 | 0.54 |
| 5 | | | 1 | 0.65 | 0.50 | 0.54 | 0.46 | 0.07 | 0.09 | 0.12 | 0.33 |
| 8 | | | | 1 | 0.36 | 0.39 | 0.48 | 0.18 | 0.06 | 0.09 | 0.42 |
| 2 | | | | | 1 | 0.97 | 0.66 | -0.1 | -0.2 | 0.03 | 0.07 |
| 4 | | | | | | 1 | 0.63 | 0.06 | 0.15 | 0.08 | 0.49 |
| 6 | | | | | | | 1 | 0.12 | 0.01 | 0.23 | 0.32 |
| 9 | | | | | | | | 1 | 0.83 | 0.13 | 0.54 |
| 10 | | | | | | | | | 1 | 0.08 | 0.37 |
| 7 | | | | | | | | | | 1 | 0.05 |
| 11 | | | | | | | | | | | 1 |

Table 4

| | | | | | | | |
|---|-------------|-------------|-------------|-------------|----------|-------------|-------------|
| | 1 | 4 | 6 | 3 | 5 | 2 | 7 |
| 1 | 1 | 0.50 | 0.47 | 0.42 | 0.40 | 0.26 | 0.35 |
| 4 | 0.50 | 1 | 0.19 | 0.35 | 0.53 | 0.18 | 0.13 |
| 6 | 0.47 | 0.19 | 1 | 0.47 | 0.50 | 0.08 | 0.46 |
| 3 | 0.42 | 0.35 | 0.47 | 1 | 0.32 | 0.15 | 0.18 |
| 5 | 0.40 | 0.53 | 0.50 | 0.32 | 1 | 0.17 | 0.04 |
| 2 | 0.26 | 0.18 | 0.08 | 0.15 | 0.17 | 1 | 0.22 |
| 7 | 0.35 | 0.13 | 0.46 | 0.18 | 0.04 | 0.22 | 1 |

Table 5

| | | | | | |
|--|--------------|---------|---------|---------------|---------------|
| | <i>n</i> = 3 | | | <i>n</i> = 5 | |
| Collection of correctable heuristics | 2, 5, 7 | 2, 4, 7 | 2, 4, 6 | 2, 3, 4, 5, 7 | 2, 4, 5, 6, 7 |
| Accuracy of the best heuristic | 0.70 | 0.74 | 0.74 | 0.74 | 0.74 |
| Average accuracy of the collection of heuristics | 0.68 | 0.70 | 0.69 | 0.68 | 0.67 |
| Logical corrector | 0.77 | 0.81 | 0.77 | 0.80 | 0.81 |

collections with the minimum (maximum) possible pair correlation coefficients between algorithms and randomly formed collections with a certain average value of statistical proximity. The accuracy of the results obtained was compared with that of the best algorithm in the collection and with the average accuracy of heuristics of the collection.

Under the experimental conditions, a further increase in the volume of correctable collections gives rise to algorithms with higher values of pair correlation coefficients in these collections.

The results of the experiments allow us to conclude that the formation of a collection of algorithms with strong correlation rarely leads to an increase in the accuracy of recognition compared with the accuracy

of the best algorithm. However, the application of trainable correctors can be justified even in this case. The accuracy of these correctors is higher than the average accuracy of the algorithms of a collection, which corresponds to the accuracy of an algorithm randomly chosen from the collection.

In practice, one often cannot guarantee the independent operation of some algorithms or precisely estimate the degree of their independence. As a result, correctable collections may contain algorithms with different degrees of statistical proximity. Table 7 shows the results of some experiments with such collections by the example of Problem 1.

The analysis of the data shows that the correction accuracy increases with *n*.

Table 6

| | $N = 3$ | | | $n = 5$ | | $n = 7$ |
|--|---------|---------|----------|---------------|-----------------|----------------------|
| | | | | | | |
| Collection of correctable heuristics | 2, 4, 6 | 1, 5, 8 | 3, 9, 10 | 1, 2, 4, 5, 6 | 1, 3, 9, 10, 11 | 2, 3, 4, 5, 6, 9, 10 |
| Accuracy of the best heuristic | 0.74 | 0.67 | 0.70 | 0.74 | 0.74 | 0.74 |
| Average accuracy of the collection of heuristics | 0.70 | 0.61 | 0.66 | 0.67 | 0.68 | 0.67 |
| Logical corrector | 0.74 | 0.67 | 0.71 | 0.74 | 0.75 | 0.74 |

Table 7

| | $n = 3$ | | | $n = 5$ | | $n = 7$ |
|--|---------|---------|---------|---------------|---------------|---------------------|
| | | | | | | |
| Collection of correctable heuristics | 1, 3, 4 | 4, 5, 6 | 1, 4, 5 | 1, 4, 5, 6, 7 | 1, 2, 3, 6, 7 | 1, 2, 3, 4, 5, 6, 7 |
| Accuracy of the best heuristic | 0.70 | 0.67 | 0.67 | 0.70 | 0.74 | 0.74 |
| Average accuracy of the collection of heuristics | 0.68 | 0.64 | 0.65 | 0.66 | 0.69 | 0.68 |
| Logical corrector | 0.72 | 0.69 | 0.70 | 0.73 | 0.81 | 0.84 |

In this paper, we obtained experimental data on the correlation of the results of recognition between algorithms applied to real problems. Results of this kind are obtained for the first time. Although one cannot make valid conclusions based on a limited number of experiments, these results are of great interest because here correlation plays the role of a natural measure of statistical proximity between algorithms.

When solving applied problems, as a rule, one should guarantee that the following operations are executed in the automatic mode:

- formation of training information for individual problems;
- solution of an applied problem by individual algorithms;
- enhancement of the reliability of recognition at the expense of the combination of several algorithms (a two-level recognition scheme with a corrector);
- comparative analysis of the efficiency of algorithms, and some other operations.

These problems were solved within a software complex, which can be considered as a testing area for recognition algorithms with a control system [17]. The architecture of this complex consists of the following basic parts: control subsystems, user programs, and a data manager (DM).

The control subsystem organizes the operation of the complex and controls the operation of its components. The user programs implement specific recognition algorithms (an algorithm library) and perform pre- and postprocessing of data. The database consists of collections of data of the object file, an indexer, and a subproblem file.

The software complex of this architecture was used for organizing and carrying out an experiment on the

study of the properties of recognition algorithms constructed on the basis of a two-level scheme.

These types of experiments require considerable expenses related to the preparation, storing, and manipulation of various data. The functional capabilities and the composition of the software technology allow one to automate many routine processes and increase the efficiency of experiments.

8. RESULTS OF PRACTICAL COMPARISON OF CORRECTORS AND ALGORITHMS

Comparison of various recognition algorithms and the choice of new objects that are most promising for recognition have always been central questions of this field; however, these questions do not have a unique answer. Each approach has its own preferences and limitations, which depend on the length of the training sample, the number of classes and features, the representativeness of the description of each class, the type of features, the compactness of the classes, correlation of features, configuration of classes, etc. A method that is the best for a certain applied problem may turn out to be mediocre when applied to other problems. A natural comparison criterion here is an estimate for the number of applied problems for which a certain method occupies leading positions. Table 8 below illustrates such comparisons. For calculations, we used the software system RECOGNITION into which the methods listed below were integrated. A row of the table corresponds to a recognition (correction) method, and a column (except for the column N), to an applied problem:

DT is an algorithm for voting by deadlock tests [1];

LR is an algorithm for voting by systems of logical regularities [18];

BT stands for binary decision trees [19];

Table 8

| Method/problem | Abalone | Eco | Breast | Credit | Home | Image | Ion | Pnevmo | Year | N |
|----------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|-------------|----------|
| DV | 60.6 | 63.1 | 92.4 | 85.3 | 63.3 | 86.5 | 87.9 | 61.4 | 44.3 | 2 |
| LR | 56.8 | 61.5 | 96.1 | 82.8 | 71.6 | 90.8 | 90.1 | 59.6 | 13.9 | 1 |
| BT | 56.7 | 73.2 | 92.1 | 85.1 | 67.4 | 90.2 | 87.4 | 57.9 | 45.2 | 1 |
| Fisher | 62.4 | 77.1 | 94.4 | 84.8 | 76.5 | 89.4 | 83.5 | 54.4 | 53.8 | 2 |
| LM | 65.5 | 79.3 | 94.9 | 85.9 | 61.7 | 92.7 | 86.8 | 54.4 | 32.0 | 5 |
| MSV | 66.5 | 81.6 | 94.6 | 83.3 | 77.7 | 91.8 | 94.0 | 63.2 | 65.5 | 6 |
| NeurC | 64.7 | 80.4 | 95.5 | 82.8 | 73.5 | 92.0 | 87.9 | 57.9 | 59.7 | 3 |
| Collective | | | | | | | | | | |
| Vote | 63.5 | 79.3 | 94.9 | 82.2 | 76.5 | 93.9 | 93.4 | 57.9 | 48.3 | 4 |
| LogC | 65.4 | 76.5 | 95.2 | 84.8 | 78.8 | 94.0 | 90.7 | 68.4 | 58.8 | 7 |
| AlgC | 65.0 | 81.6 | 95.2 | 81.0 | 76.9 | 92.3 | 91.2 | 61.4 | 60.6 | 7 |
| GenC | 63.5 | 80.4 | 94.9 | 82.8 | 78.0 | 94.3 | 94.0 | 66.7 | 61.2 | 7 |

FISHER is Fisher's linear discriminant [20];

LM is a linear machine [20];

MSV is a method of support vectors [21];

NeurC is a multilayer perceptron [22];

Vote is a collective solution as majority voting [8];

LogC is a logical corrector (Section 6);

AlgC is a second-degree algebraic corrector (Section 4);

GenC is a generalized algebraic corrector (Section 5);

"Abalone" is the problem of determining the age of abalone, $l = 3$, $q = 2079$, $n = 8$ (data are supplied by Sam Waugh (Sam.Waugh@cs.utas.edu.au), Department of Computer Science, University of Tasmania, GPO Box 252C, Hobart, Tasmania 7001, Australia);

"Eco" is the protein localization problem, $l = 8$, $q = 179$, $n = 7$ [23];

"Breast" is the mammary gland cancer diagnosis problem, $l = 2$, $q = 355$, $n = 9$ [24];

"Credit" is the credit card endorsement problem $l = 2$, $q = 348$, $n = 15$ [25];

"Home" is the problem of assessing housing prices, $l = 5$, $q = 264$, $n = 13$ [26];

"Image" is the problem of recognition of standard images, $l = 7$, $q = 2100$, $n = 16$ [25];

"Ion" is the radio signal recognition problem, $l = 2$, $q = 182$, $n = 34$ [27];

"Pnevmo" is the problem of assessing the degree of severity of pneumonia, $l = 4$, $q = 57$, $n = 41$;

"Year" is the protein localization sites problem, $l = 10$, $q = 747$, $n = 8$ [23].

Columns 2–10 show estimates for the recognition accuracy of various methods (percentage of correctly recognized objects in a certain sample) in different problems. For each problem, the table presents the number of classes, the number of objects to be recognized, and the number of features. Note that the parameters of the algorithms were used by default (i.e., they were specified automatically), rather than chosen from the training data (for example, by the minimization of the number of errors on the training data in the sliding test mode). Boldface numbers in each table show the four best results, the prize-winning places. The column N indicates how many times an appropriate method took a prize-winning place.

The results of testing the EEA-polynomial method are presented in Table 9.

The interpretation of the problems is given in [25] and, in part, in [8].

Table 9

| Task | AVO-polynom |
|----------------|-------------|
| Abalone | 62.3 |
| Breast cancer | 96.1 |
| Ionosphere | 98.7 |
| Echocardiogram | 77.4 |
| Hepatitis | 88.0 |
| Image | 89.4 |
| Credit | 86.2 |

9. CONCLUSIONS

In this paper, we have presented some practical methods of algebraic and logical correction of heuristic recognition algorithms, as well as the results of their comparison. Although the number of practical problems was small, the results obtained provide a good illustration of the expediency of using correcting algorithms for increasing the reliability of the results of recognition. This is especially important when recognition software is used by unqualified personnel in the automatic mode.

ACKNOWLEDGMENTS

This work was supported by the National Academy of Sciences of the Republic of Belarus and the Russian Foundation for Basic Research within joint project no. 08-01-90016 “Methods for Constructing Efficient Feature Spaces for Monitoring and Recognition of Complex Objects in Images and Scenes,” by the Russian Foundation for Basic Research (project nos. 08-01-00636-a and 09-01-00409-a), by a grant of the President of the Russian Federation (project no. NSh-5294.2008.1), and by the Presidium of the Russian Academy of Sciences and the Division of Mathematical Sciences, Russian Academy of Sciences, within the programs for supporting fundamental research.

REFERENCES

1. Yu. I. Zhuravlev, “On Algebraic Approach to Solve the Problems of Recognition and Classification,” in *Problems of Cybernetics* (Nauka, Moscow, 1978), Vol. 33, pp. 5–68.
2. Yu. I. Zhuravlev, “Correct Algebras for Incorrect (Heuristic) Sets of Algorithms. I,” *Kibernetika*, No. 4, 5–17 (1977); II, No. 6 (1977); III, No. 2, 35–43 (1978).
3. Yu. A. Zuev, “Methods to Improve the Classification Reliability, in the Presence of Several Classifiers, Based on Monotony Principle,” *Zh. Vychisl. Mat. Mat. Fiz.* **21** (1), 157–167 (1981).
4. A. M. Veshtort, Yu. A. Zuev, and V. V. Krasnoprosin, “Two-Level Recognition System With Logical Corrector,” in *Recognition, Classification, Prediction. Mathematical Methods and Their Application* (Nauka, Moscow, 1989), issue 2, pp. 73–98.
5. L. A. Aslanyan, L. F. Mingo, J. B. Castellanos, F. B. Chelnokov, A. A. Dokukin, and V. V. Ryazanov, “On Logical Correction of Neural Network Algorithms for Pattern Recognition,” in *Proceedings of the 4th International Conference “Information Research & Application”* (Foi-Commerce, Sofia, 2006).
6. A. G. D'yakov, “Algebra over Estimation Algorithms: The Minimal Degree of Correct Algorithms,” *Zh. Vychisl. Mat. Mat. Fiz.* **45** (6), 1134–1145 (2005) [*Comput. Math. Math. Phys.* **45**, (6), 1095 (2005)].
7. A. G. D'yakov, “Algebra over Estimation Algorithms: Normalization with Respect to the Interval,” *Zh. Vychisl. Mat. Mat. Fiz.* **49**, No. 1, 200–208 (2009) [*Comput. Math. Math. Phys.* **49** (1), 194 (2009)].
8. Yu. I. Zhuravlev, V. V. Ryazanov, and O. V. Sen'ko, *Recognition. Mathematical Methods. Program System. Practical Applications* (Fazis, Moscow, 2006) [in Russian].
9. Yu. I. Zhuravlev and I. V. Isaev, “The Way to Generate the Recognition Algorithms, Which are Correct for the Given Control Sample,” *Zh. Vychisl. Mat. Mat. Fiz.* **19** (3), 726 (1979).
10. A. A. Dokukin, “A Method for Constructing an Optimal Estimate Calculation Algorithm,” *Zh. Vychisl. Mat. Mat. Fiz.* **46**, No. 4, 754–760 (2006) [*Comput. Math. Math. Phys.* **46** (4), 719 (2006)].
11. A. A. Dokukin, “On the Construction of Samples for Testing Approximate Optimization Methods for Estimate Calculation Algorithms,” *Zh. Vychisl. Mat. Mat. Fiz.* **46** (5), 978–983 (2006) [*Comput. Math. Math. Phys.* **46**, (5), 914 (2006)].
12. V. V. Krasnoprosin and V. A. Obratsov, “Problems of Decision-Making according to the Precedent: Solvability and Algorithm Selection,” in *Selected Scientific Works of Belarus State University* (2001), Vol. 6, pp. 285–312.
13. M. Yu. Romanov, “A Method for Constructing a Recognition Algorithm in Algebra over an Estimate Calculation Set,” *Zh. Vychisl. Mat. Mat. Fiz.* **47** (8), 1426–1430 (2007) [*Comput. Math. Math. Phys.* **47** (8), 1368 (2007)].
14. M. Yu. Romanov, “Implementation of a Method for Constructing a Recognition Algorithm in Algebra over an Estimate Calculation Set,” *Zh. Vychisl. Mat. Mat. Fiz.* **48** (9) (2008) [*Comput. Math. Math. Phys.* **48** (9), 1680 (2001)].
15. A. M. Veshtort, S. I. Kashkevich, S. B. Kostyukevich, V. V. Krasnoprosin, and S. G. Sinyakovich, “Principles for Generating the Information Automated Data System for Spectrometric Aerial Survey for Physical-Geographical Zoning,” *Izv. Akad. Nauk USSR. Ser. Geogr.*, No. 2, 89–94 (1988).
16. V. S. Glushenkov, A. A. Kovalev, O. L. Konovalov, S. B. Kostyukevich, V. V. Krasnoprosin, and B. A. Yukhimenko, “Display System and the System for Editing the Digital Subject Radiation and Ecological Maps by Means of Personal Computer,” in *News of The National Academy of Sciences of Belarus. Series Physical and Mathematical Sciences* (Minsk, 1992), vols. 5–6, pp. 102–107.
17. S. V. Gafurov and V. V. Krasnoprosin, “Programming Technique for Generating the Systems to Solve the Recognition Problems with Complex Structure,” *Iskusstv. Intell.*, No. 1, 30–37 (2008).
18. V. V. Ryazanov, “Logical Regularities in Pattern Recognition (Parametric Approach),” *Zh. Vychisl. Mat. Mat. Fiz.* **47** (10), 1793 (2007) [*Comput. Math. Math. Phys.* **47** (10), 1720 (2007)].
19. V. I. Donskoi and A. I. Bashta, *Discrete Model of Decision-Making under Incomplete Information* (Tavriya, Simferopol, 1992) [in Russian].
20. R. O. Duda and P. E. Hart, *Pattern Recognition and Scene Analysis* (John Wiley, New York, 1973; Mir, Moscow, 1976).
21. C. J. C. Burges, “A Tutorial on Support Vector Machines for Pattern Recognition,” in *Data Mining and Knowledge Discovery* (1998), Vol. 2, pp. 121–167.
22. P. Wasserman, *Neurocomputing: Theory and Practice* (Van Nostrand Reinhold, New York, 1990; Mir, Moscow, 1992).
23. P. Horton and Nakai Kenta, “A Probabilistic Classification System for Predicting the Cellular Localizations Sites of Proteins,” in *Intelligent Sys. Mol. Biol.* (St. Lois, 1996), pp. 109–115.
24. O. L. Mangasarian and W. H. Wolberg, “Cancer Diagnosis via Linear Programming,” *SIAM News* **23** (5), 1 (1990).
25. <http://www.isc.uci.edu/~mlearn/MLRepository.html>
26. D. Harrison and D. L. Rubinfeld, “Hedonic Prices and the Demand for Clean Air,” *J. Environ. Econ. & Management* **5**, 81 (1978).
27. V. G. Sigillito, S. P. Wing, L. V. Hutton, and K. B. Baker, *Classification of Radar Returns from the Ionosphere Using Neural Networks* (Johns Hopkins APL Technical Digest, 1989), vol. 10, pp. 262–266.



Yurii Ivanovich Zhuravlev. Born 1935. Graduated from the Moscow State University in 1957. Received doctoral degree in 1965, Professor since 1967, and Academician of the Russian Academy of Sciences since 1992. Currently Deputy Director of the Dorodnitsyn Computing Center, Russian Academy of Sciences; Chair in the Mathematics Department of the Russian Academy of Sciences; and Head of the Department at Moscow State University. Editor-in-Chief of *Pattern Recognition and Image Analysis*. Foreign member of the Spanish Royal Academy of Sciences, the National Academy of Sciences of Ukraine, and the European Academy of Sciences. Winner of the Lenin and Lomonosov prizes. Scientific interests: mathematical logic; control systems theory; mathematical theory of pattern recognition, image analysis, and forecasting; operations research; and artificial intelligence.

computing technologies. Author of more than 180 papers, including 5 books.



Viktor Vladimirovich Krasnoproshin. Born 1947. Graduated from Belarussian State University in 1974. Received candidate's degree in 1979 and doctoral degree in 2007. Currently Head of the Department of Information Control Systems, Faculty of Applied Mathematics and Informatics, Belarussian State University. Scientific interests: decision making theory, pattern recognition, artificial intelligence, and information and

computing technologies. Author of more than 180 papers, including 5 books.



Vladimir Vasil'evich Ryazanov. Born 1950. Graduated from the Moscow Institute of Physics and Technology in 1973. Received candidates degree in 1979 and doctoral degree in 1994. Academician of the Russian Academy of Natural Sciences, Professor. Since 1976 has been with the Dorodnitsyn Computing Center, Russian Academy of Sciences. Currently is Head of the Department of Mathematical Problems of Recognition and Methods of Combinatorial Analysis. Scientific interests: recognition theory, cluster analysis, data analysis, optimization of recognition models, and applied systems of analysis and prediction.

recognition and prediction.

Mikhail Yur'evich Romanov. Born 1980. Received candidate's degree in 2008. Currently is a manager in Svyaznoi Company. Scientific interests: intelligent data analysis, discrete optimization problems, applied systems of recognition and prediction. Author of six papers.



Sergei Vladimirovich Ablameiko. Born 1956. Graduated from Belarussian State University in 1978. Received candidate's degree in 1984 and doctoral degree in 1990. Currently Rector of Belarussian State University, Academician of the National Academy of Sciences of the Republic of Belarus. Scientific interests: image processing and analysis, pattern recognition, numerical geometry, and geoinformation systems. Author of more than 350 papers, including 11 books.



Andrei Sergeevich Biryukov. Born 1979. Graduated from the Moscow Institute of Physics and Technology in 2002. Received candidates degree in 2005. Currently is a leading programmer at a branch of the Aelita Software Corporation. Scientific interests: cluster analysis; optimization of multiparameter recognition models; software systems for intelligent data analysis; and practical problems of recognition, classification, and prediction.



Aleksandr Aleksandrovich Dokukin. Born 1980. Graduated with honors from the Department of Computational Mathematics and Cybernetics, Moscow State University, in 2002. Received candidates degree in 2005. Currently is a researcher at the Dorodnitsyn Computing Center, Russian Academy of Sciences.



Vladimir Alekseevich Obratsov. Born 1953. Graduated from Belarussian State University in 1979. Received candidate's degree in 1986. Currently Associate Professor in the Department of Information Control Systems, Faculty of Applied Mathematic and Informatics, Belarussian State University. Scientific interests: pattern recognition, artificial intelligence, and inductive logic. Author of more than 60 papers.