

REPORT DOCUMENTATION PAGE

Form Approved OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing the burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.
PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS.

| | | |
|--|---------------------------------------|--|
| 1. REPORT DATE (DD-MM-YYYY) 13-03-2001 | 2. REPORT TYPE Final Report | 3. DATES COVERED (From - To) 14 March 2000 - 01-Jun-01 |
|--|---------------------------------------|--|

| | |
|--|---|
| 4. TITLE AND SUBTITLE Search for Efficient Methods of Prediction of Multicomponent Inorganic Compounds | 5a. CONTRACT NUMBER F61775-00-WE014 |
| | 5b. GRANT NUMBER |
| | 5c. PROGRAM ELEMENT NUMBER |

| | |
|--|-----------------------------|
| 6. AUTHOR(S) Dr. Nadezhda Nikolaevna Kiselyova | 5d. PROJECT NUMBER |
| | 5d. TASK NUMBER |
| | 5e. WORK UNIT NUMBER |

| | |
|--|--|
| 7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Russian Academy of Sciences 49 Leninskii Prospekt Moscow 117334 Russia | 8. PERFORMING ORGANIZATION REPORT NUMBER N/A |
|--|--|

| | |
|---|--|
| 9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) EOARD PSC 802 BOX 14 FPO 09499-0014 | 10. SPONSOR/MONITOR'S ACRONYM(S) |
| | 11. SPONSOR/MONITOR'S REPORT NUMBER(S) SPC 00-4014 |

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

14. ABSTRACT

This report results from a contract tasking Russian Academy of Sciences as follows: The contractor will investigate methods for predicting the properties of new binary, ternary and quaternary inorganic chemical systems using computer-based semantic and neural networks. Details are provided in the proposal statement of work.

20020108 141

15. SUBJECT TERMS
EOARD, Materials

| | | | | | |
|--|------------------------------|-------------------------------|---|--------------------------------------|---|
| 16. SECURITY CLASSIFICATION OF: | | | 17. LIMITATION OF ABSTRACT UL | 18. NUMBER OF PAGES 32 | 19a. NAME OF RESPONSIBLE PERSON Carl A. Kutsche |
| a. REPORT UNCLAS | b. ABSTRACT UNCLAS | c. THIS PAGE UNCLAS | | | 19b. TELEPHONE NUMBER (Include area code) +44 (0)20 7514 4505 |

**SPECIAL PROJECT SPC-00-4014
CONTRACT F61775-00-WE014**

**SEARCH FOR EFFICIENT METHODS OF PREDICTION OF
MULTICOMPONENT INORGANIC COMPOUNDS**

(final report)

Principal Investigator: Dr. Nadezhda N.Kiselyova
Co-Principal Investigator: Prof. Victor P.Gladun

Members of the research team:
Dr.Neonila D.Vashchenko
Vitalii Yu.Velichko

February 2001

AQ F02-04-0465

ABSTRACT

The project's goal: investigation of the methods for predicting the properties of new binary, ternary and quaternary inorganic chemical systems using computer-based network structures. The final goal of investigations in this field is selection of the most efficient methods for development of the inorganic compound computer design systems for 21st century.

Subject terms (Key Words) Concept learning; Knowledge discovery; Computer learning; Prediction; Inorganic compound; Physical-chemical system; Pattern recognition; Artificial intelligence; Neural network; Semantic network; Growing pyramidal network.

FOREWORD

This final report was prepared by the above identified research team under EOARD Special Project SPC-00-4014 (Contract F61775-00-WE014). This work was carried out in close contact with researchers of the Materials Directorate, Wright Laboratory (USA). We appreciate the leadership of Dr. Steven R. LeClair. We also acknowledge the support provided by EOARD, for enabling the interaction between ourselves and our colleagues from Wright Laboratory.

TABLE OF CONTENTS

| <i>SECTION</i> | <i>PAGE</i> |
|---|-------------|
| 1. Introduction | 6 |
| 2. Learning and Testing Sets | 6 |
| 3. Attribute (Elementary Properties) Sets | 6 |
| 4. Program Systems Description | 9 |
| 4.1. Some Important Procedures of Attribute Set Formation | 9 |
| 4.1.1. Discretization of Numerical Attributes | 9 |
| 4.1.2. Formation of Derivative Attributes | 10 |
| 4.2. Investigations on the Basis of the <i>ANALOGY</i> System | 11 |
| 5. Description of Experiments | 12 |
| 5.1. Binary Systems Prediction | 12 |
| 5.1.1. Formation of Derivative Attributes | 14 |
| 5.1.1.1. Concatenation of All Attributes Arranged in Pairs | 14 |
| 5.1.1.2. Differences of Attribute Values | 14 |
| 5.2. Ternary Systems Prediction | 16 |
| 5.3. Quaternary Systems Prediction | 18 |
| CONCLUSIONS | 29 |
| KEY TERM LIST | 32 |
| REFERENCES | 33 |
| APPENDICIES 1-23 | |

1. Introduction

The algorithms of computer learning are used for computer design of inorganic substances the most frequently. The complexity of these problems is so great, that there is no universal criterion of choice of algorithm, best suited to a solution of certain problem. The present work is a part of investigations devoted to the search for efficient learning algorithms for computer design of new inorganic compounds. The distinctive feature of these investigations was the common learning sets and elementary properties for description of materials systems. The problems of prediction of compound formation and non-formation binary, ternary and quaternary systems were solved. Taking into account the great sizes of the proposed learning sets the investigators chose efficient algorithms of large information bulk analysis. Our colleagues had tested the by them developed program systems based on the method K-nearest neighbors and neuro-computing already (Villars et al. (2000)). In individual cases their accuracy of compound formation predicting exceeded 99%. This clearly demonstrates an efficiency of the approach to computer design of new inorganic materials based on learning algorithms, which was put forward by specialists from Baikov Institute more than 30 years ago (Savitskii et al. (1968)). The present investigation is aimed at testing of algorithms much used in inorganic materials computer design, which are based on the growing pyramidal networks (Gladun (2000), Gladun et al. (2000), Velichko & Moskal'kova (1999)).

2. Learning and Testing Sets

Customers proposed the data sets of three kinds:

1. Binary Systems:

B3: set consisting of 1333 binary systems, a mixture of formers and nonformers (Appendix 1, from this point former is designated by "1" and nonformer is designated as "0").

B4: set consisting of 692 binary systems, a mixture of formers and nonformers (Appendix 2).

2. Ternary Systems:

T3: 455 (mixed) ternary systems judged to be 80% reliable (Appendix 3).

T4: a set (mixed) of 4278 ternary systems (Appendix 4).

T5: a set (mixed) of 2156 ternary systems (Appendix 5).

3. Quaternary Systems:

Q1: a body of 2747 quaternary formers (Appendix 6).

Q2: a body of 4751 quaternary nonformers (Appendix 7).

Q3: set consisting of 4963 quaternary systems, a mixture of formers and nonformers (Appendix 8).

Q4: set consisting of 2536 quaternary systems, a mixture of formers and nonformers (Appendix 9).

3. Attribute (Elementary Properties) Sets

Each chemical element was described by the set of 87 attribute values. They were grouped as follows (Villars et al. (2000)):

- Size factor (10 properties) – Appendix 10.
 - o S1 Radii pseudo-potential (Zunger) (a.u.).
 - o S2 Radii ionic (Yagoda) (A).
 - o S3 Radii covalent (pm).
 - o S4 Radii metal (Waber) (A).
 - o S5 Distance valence electron (Schubert) (A).
 - o S6 Distance core electron (Schubert) (A).
 - o S7 Volume atom (Villars, Daams) (10^{-2} nm^3).
 - o S8 $V^{2/3}$ Miedema (cm^2).
 - o S9 Atomic environment number (Villars, Daams).

- S10 Atomic concentration (10^{23} m^{-3}).
- **Cohesion-energy factor** (12 properties) - Appendix 11.
 - C1 Temperature melting (K).
 - C2 Temperature boiling (K).
 - C3 Enthalpy vaporization (kJ mol^{-1}).
 - C4 Enthalpy melting (kJ mol^{-1}).
 - C5 Enthalpy atomization (kJ mol^{-1}).
 - C6 Enthalpy surface Miedema (kJ mole^{-1}).
 - C7 Enthalpy vacancies Miedema (kJ mole^{-1}).
 - C8 Energy cohesive Brewer (J mol^{-1}).
 - C9 Modulus compression (GPa).
 - C10 Modulus bulk (GPa).
 - C11 Modulus rigidity (GPa).
 - C12 Modulus Young (GPa).
- **Electro-chemical factor** (12 properties) - Appendix 12.
 - E1 Electronegativity (Martynov&Batsanov).
 - E2 Electronegativity (Pauling).
 - E3 Electronegativity (Alfred-Rochow).
 - E4 Electronegativity absolute.
 - E5 Energy ionization first (kJ mol^{-1}).
 - E6 Energy ionization second (kJ mol^{-1}).
 - E7 Energy ionization third (kJ mol^{-1}).
 - E8 Chemical potential Miedema.
 - E9 Work function (eV).
 - E10 $nWS^{1/3}$ Miedema (a.u.^{-1/3}).
 - E11 Nuclear charge effective Slater.
 - E12 Charge nuclear effective (Clementi).
- **Group number factor** (2 properties) - Appendix 13.
 - G1 Group number.
 - G2 Valence electron number.
- **Atomic number factor** (11 properties) - Appendix 14.
 - A1 Atomic number start counting left top, left-right sequence.
 - A2 Periodic number start counting left bottom, left-right sequence.
 - A3 Periodic number start counting top right, right-left sequence.
 - A4 Periodic number start counting bottom right, right-left sequence.
 - A5 Quantum number.
 - A6 Atomic weight (10^{-3} kg).
 - A7 Mass attenuation coefficient for $\text{MoK}\alpha$ ($\text{cm}^2 \text{ g}^{-1}$).
 - A8 Mass attenuation coefficient $\text{CrK}\alpha$ ($\text{cm}^2 \text{ g}^{-1}$).
 - A9 Mass attenuation coefficient for $\text{CuK}\alpha$ ($\text{cm}^2 \text{ g}^{-1}$).
 - A10 Mass attenuation coefficient $\text{FeK}\alpha$ ($\text{cm}^2 \text{ g}^{-1}$).
 - A11 atomic electron scattering factor at 0.5.
- **Mendeleev number factor** (11 properties) - Appendix 15.
 - M1 Mendeleev H t-d start left.
 - M2 Mendeleev H t-d start right.
 - M3 Mendeleev H d-t start left.
 - M4 Mendeleev H d-t start right.
 - M5 Mendeleev Pettifor.
 - M6 Mendeleev Pettifor regular.

- M7 Mendeleev chemists sequence.
- M8 Mendeleev t-d start left.
- M9 Mendeleev t-d start right.
- M10 Mendeleev d-t start left.
- M11 Mendeleev d-t start right.
- **Another properties** (29 properties) – Appendix 16.
 - I1 Magnetic resonance (MHz).
 - I2 Magnetic frequency of nuclei.
 - I3 Magnetic susceptibility ($\text{m}^3 \text{kg}^{-1}$).
 - I4 Spin nuclei (h).
 - I5 Density (kg m^{-3}).
 - I6 Resistivity electrical (10^{-8} Ohm m).
 - I7 Conductivity electrical ($10^4 \text{ Ohm}^{-1} \text{ cm}^{-1}$).
 - I8 Conductivity thermal ($\text{J mole}^{-1} \text{ K}^{-1}$).
 - I9 Sound velocity (m s^{-1}).
 - I10 Molar heat capacity ($\text{J mol}^{-1} \text{ K}^{-1}$).
 - I11 Entropy of solid ($\text{J mol}^{-1} \text{ K}^{-1}$).
 - I12 Electrochemical weight equivalent (mg K^{-1}).
 - I13 γ -free electron to specific heat ($\text{mJ mol}^{-1} \text{ K}^{-2}$).
 - I14 Temperature Debye (K).
 - I15 Thermal neutron capture cross section (barns).
 - I16 Hardness Brinell.
 - I17 Linear thermal expansion coefficient (10^{-6} K^{-1}).
 - I18 Poisson's ratio.
 - I19 Spectral lines n_o .
 - I20 Vapour pressure at T_m (N m^{-2}).
 - I21 Surface tension at T_m (mJ m^{-2}).
 - I22 Electron affinity (kJ mole^{-1}).
 - I23 Moment nuclear magnetic (μ).
 - I24 ΔH interface S in M (kJ mole^{-1}).
 - I25 ΔH interface O in M Miedema (kJ mole^{-1}).
 - I26 ΔH interface M in O Miedema (kJ mole^{-1}).
 - I27 Δ interface M in S Miedema (kJ mole^{-1}).
 - I28 Normal atomic configuration Thaler.
 - I29 Oxidation state first.

Consequently the description of each binary physical-chemical system contains $87 \cdot 2 = 174$ attribute values, ternary system – $87 \cdot 3 = 261$, quaternary system – $87 \cdot 3 = 348$ attribute values.

4. Program Systems Description

Prediction of chemical compounds on the basis of training sets is a typical problem for attributive analyses. In attributive analysis the objects are represented as sets of attribute values and the basic analytical operation is comparing attributive object descriptions.

Two approaches are used to solve the prediction problem: classification and sequential reasoning. Here we used the classification one. The investigation was performed on the basis of the research analytical complex intended for solution of the main problems of attributive analyses, that is knowledge discovery, classification, diagnostics and prediction. The complex was created in V.M.Glushkov Institute of Cybernetics of National Academy of Science of Ukraine.

There are two factors influencing the final result of analytical problem solution:

- 1) choice of the space of attributes used for object descriptions;
- 2) choice of such tools for representation of data processed that could provide acceleration of operations of attribute comparison.

In this connection, when solving problems of prediction of chemical compound forming systems, we concentrated the most attention on methods of attribute space formation and data representation.

4.1. Some Important Procedures of Attribute Set Formation

4.1.1. Discretization of Numerical Attributes

To avoid errors, which are appeared as a result of using different measure units, numerical attributes are transformed into nominal ones. This problem is referred to as a problem of discretization.

The discretization of the scale of a numerical attribute D_i means formation of mapping into finite ordered set of intervals of numerical values.

$B_i = \{b_m | (d_i^k \in b_m) : b_m^l \leq d_i^k < b_m^r\}$, $m=1, 2, \dots, M_i$, where b_m^l and b_m^r are left and right limits of interval b_m .

The discretization is performed on scales of numerical attributes by analysis of distributions of learning set objects belonging to different classes. Every object of the learning set is marked on the scale with its attribute value.

In the process of discretization intervals of the following types can be formed:

- 1) empty intervals not containing attribute values of objects belonging to the learning set;
- 2) homogeneous intervals containing marked attribute values of objects of the same class;
- 3) even intervals in which there are marked attribute values of different classes and the difference between numbers of marked values of the classes does not exceed the given threshold;
- 4) uneven intervals in which there are marked attribute values of different classes and the difference between the number of marked values of some class and the number of marked values of any other one exceeds the given threshold.

The discretization algorithm for attribute D_i ($i=1, 2, \dots, N$) consists in consecutive performance of the following operations:

1. Formation of the scale of attribute values describing objects of the learning set.
 - 1.1. Definition of the scale limits. The operation involves finding the highest (d_i^{max}) and the lowest (d_i^{min}) values of the attribute among all objects of the learning set.
 - 1.2. Definition of the size δ of the initial interval $\delta(d_i^{max} - d_i^{min})/L$, where L is the number of initial intervals. The algorithm makes it possible choosing L by an user.
 - 1.3. Numbering initial intervals and definition of their limits. Left and right limits of initial intervals are calculated in such a way: $b_1^l = d_i^{min}$; $b_1^r = d_i^{min} + \delta$; $b_m^l = b_{m-1}^r$; $b_m^r = b_m^l + \delta$; $m=2, \dots, L$.
2. Formation of distributions of learning set objects. In each of initial intervals the attribute values belonging to objects of the learning set are marked. Attribute value d_i^k ($i=1, 2, \dots, N$, $k=1, 2, \dots, K_i$) belongs to interval $b_m \in B_i$ ($m=1, 2, \dots, (M_i - 1)$) if $b_m^l \leq d_i^k < b_m^r$. For the last interval $b_{M_i}^l \leq d_i^k \leq b_{M_i}^r$.
3. Unification of neighboring empty intervals. The scale of the attribute is examined from the left limit to the right one. Neighboring empty intervals are united.

4. Generalization of the distribution. Nonempty parts of the scale are extended at the expense of joining several empty initial intervals from both sides. The number of initial intervals, that should be joined from each side, is defined by an user. The operation results in disappearance of short empty intervals.
5. Unification of neighboring homogeneous intervals of the same class. As a result the largest homogeneous intervals are formed.
6. Unification of neighboring uneven intervals.
7. Unification of neighboring even intervals.

The algorithm of discretization is aimed at providing separability and compactness of classes in the attribute space. It is achieved by revealing intervals of numerical values that are the most characteristic for classes, namely, intervals containing attribute values of one class only (homogeneous intervals) and intervals with predominance of attribute values of one class (uneven intervals).

The algorithm is aimed at selection of intervals of the largest size, i.e. at reducing of the total number of nominal attributes being formed.

In contrast to known algorithms of discretization the algorithm provides completely automatic discretization reflecting peculiarities of mutual distributions of different classes.

The algorithm of discretization is realized in the program *DISCRET*, which is a part of analytical complex.

4.1.2. Formation of Derivative Attributes

Results of prediction improve if the most characteristic essential attributes are used for object descriptions. It is the very important condition for division of different object classes in the attribute space. Sometimes division of classes is provided by using derivative attributes that are attributes whose values are defined by means of some function of values of initial attributes. Sometimes for this purpose arithmetic functions are used. For well-grounded choice of such function it is necessary to understand how specific attributes influence the predicted parameter. In our investigations we used attribute differences. This operation allowed obtaining the good predicting results.

4.2. Procedures of Attributive Analyses

There are two approaches to solving the classification problem:

- 1) just on the basis of the training set by investigation of the neighborhood of a point representing recognized object in the attribute space;
- 2) on the basis of generalized regularity characterizing the investigated object class.

The first approach has success for representative sets when distribution of object classes in the attribute space is compact. The trouble with this approach is that results of classification (prediction) of new objects depend on the parameter, which defines sizes of investigated neighborhood of the point, that represent the classified object in the attribute set (for example, on number k of the investigated nearest points in the method of k -nearest neighbors). The choice of k should provide the best classification of the examination set. Such k can ideally divide examination set and, nevertheless, give bad results when classifying (predicting) new objects. Prediction results for new object are of less dependency on k if the training set is large and representative. But representative sets are not always available.

The process of regularity revealing (the second approach) is complicated when the attribute space is redundant ("littered" with unessential attributes). Such is indeed our case in the worst variant. Naturally, no normal researcher search for regularity for prediction in the space, having some hundreds attributes. The preliminary reasonable selection of the most essential attributes is needed.

Investigations in the frames of this project were done with the help of program systems *DISCRET* and *ANALOGY* that are components of the analytical complex.

4.2. Investigations on the Basis of the *ANALOGY* System

The *ANALOGY* system is aimed:

- at design of complex objects with predefined properties;
- at design of complex objects that are analogous to specified ones;
- classification of objects;
- determination of unknown properties of complex objects;
- search for objects that are analogous to specified ones in databases.

The system can be adapted to different application domains by means of filling the appropriate database. The database of the system includes descriptions of:

- known properties of complex objects;
- composition of complex objects;
- properties of elements that are constituents of complex objects;
- rules for constructing of complex objects (for problems of design of composition).

The *ANALOGY* system provides:

- filling the data base in a dialogue mode in the restricted natural language;
- filling the data base from files;
- adaptation of the man-machine dialogue to the application domain;

The system includes a number of different procedures for comparison of objects, which are represented with their descriptions (as sets of values of nominal attributes), as well as procedures for evaluation of similarity level of such descriptions.

The investigations were done in two stages:

- discretization of attributes with the help of the *DISCRET* system;
- classification of objects with the help of the *ANALOGY* system by means of analysis of k-nearest neighbors.

5. Description of Experiments

In all tasks the program systems *ANALOGY* and *DISCRET* were used for testing tasks solution. The attribute sets include all elementary properties listed in Appendices 10-16.

5.1. Binary Systems Prediction

Table 1.

Results of Examination for Learning Set B3 and Testing Set B4

| Learning Set | | Testing Set | | Violations | | Accura-cy, % |
|--------------|----------------|--------------|-----------------|--------------|----------------------|-----------------|
| For- mers | Nonform ers | For- mers | Non- formers | Formers | Nonformers | |
| B3 | | B4 | | As-B; Au-Hg; | Ag-Cr; Ag-Mn; As-Au; | 95.0867 |

| | | | | | | |
|-----|-----|-----|-----|---|--|--|
| 923 | 410 | 473 | 219 | B-Re; Bi-In; Bi-Rh; Cs-P; Ge- Ir; Ge-Os; Hg-Rh; Mn-Nd; Mn-Pr; Mn-Y; Mo-U; Pb- Pd; Pd-Tl; Re-V; Sb-Tl; | Au-Ge; B-Cu; Cd-Fe; Cr-W; Cs-Pr; Cu-Fe; Cu-Ni; Fe-In; Ge-Sb; In-V; Li-Pu; Mg-Mo; Nb-Tl; Rh-Tc; | |
|-----|-----|-----|-----|---|--|--|

Table 2.

Results of Examination for Learning Set T4 and Testing Set B3+B4

| Learning Set | | Testing Set | | Violations | | Accuracy, % |
|--------------|------------|-------------|------------|--|---|----------------|
| Formers | Nonformers | Formers | Nonformers | Formers | Nonformers | |
| T4 | | B3+B4 | | Ac-O; Ag-Be; Ag-Hg; Al-W; Am-O; As-B; As-Cd; As-Ta; As-Tc; As-W; Au-B; Au-Be; Au-Cr; Au-Hg; Au-In; Au-Mn; Au-Nb; Au-Ta; Au-V; Au-Zn; Be- Co; Be-Cu; Be-Fe; Be-Hf; Be-Ir; Be- Mn; Be-Os; Be-P; Be-Pd; Be-Pt; Be- Re; Be-Rh; Be-Ru; Be-Tc; Bi-Rh; C- Na; Ce-Mg; Cr-Ni; Cr-Ru; Cr-Ti; Cr- Zn; Ga-Ta; Ga-W; Ge-Ta; Ge-W; Hg- Mn; Hg-Ni; Hg- Rh; In-Mn; In-Ta; Ir-Sb; Ir-W; La-Mg; Mg-Nd; Mg-Pr; Mg-Sm; Mo-Os; Mo-Pt; Mo-Re; Mo-Rh; Mo-Sb; Mo-Sn; Mo-Tc; Mo-Zn; Nb-Pb; Nb-Pt; Nb- Sb; Nb-Zn; Ni-V; O-Pm; O-Tc; Pb-V; Pt-V; Rh-W; Sb- Ta; Sb-V; Sn-Ta; Ta-Zn; W-Zr; | Ag-Ge; Ag-Si; Al-Cd; Al-Ge; Al-K; Al-Pb; As-Pb; As- Tl; Ba-Mo; Ba-Ti; Ba-Y; Bi-Cd; Bi-Co; Bi-Zn; C-Ge; C-Pt; C-Rh; Ce-Cr; Ce-Mo; Ce-Nb; Cr-Cs; Cr- Er; Cr-Gd; Cr-K; Cr-La; Cr-Mo; Cr- Nd; Cr-Np; Cr-Pr; Cr-Rb; Cr-Sm; Cr-Th; Cr-U; Cr- Y; Cr-Yb; Dy-Er; Dy-Gd; Dy-Mo; Dy-Sm; Dy-Ti; Dy-W; Dy-Yb; Er-Ho; Er-Ti; Er-V; Eu-Hf; Eu-V; Eu-Zr; Fe-In; Fe-K; Fe-Pb; Gd- Ho; Gd-Mo; Gd-Sc; Gd-Tm; Gd-W; Gd-Y; Ge-Pb; Ho-Mo; Ho-Sc; Ho-Sm; Ho-Yb; K-Ti; K-V; La-Nb; La-Re; La-Sc; La-Yb; Li-Ti; Lu-Mo; Lu-Nd; Lu-Sm; Lu-Ti; Lu-Yb; Mg-Ti; Mg-V; Mo-Np; Mo-Pa; Mo-Ta; Mo-Th; Mo-Tm; Mo-Y; Na-Rb; Nb-Rb; Nb-Ta; Nb-Ti; Ni-Tl; P-Sb; PbSi; PbZn; PuW; Sc-Ta; Sc-Tb; Sc-Ti; Sc-U; Sc-Y; Sm- Tm; Th-Zr; Ti-Tm; U-V; U-Yb; Y-Yb | 91.0617 |
| 2835 | 1443 | 1396 | 629 | | | |

5.1.1. Formation of Derivative Attributes

To raise predicting accuracy we used transformation of initial elementary properties. These expectations were realized.

5.1.1.1. Concatenation of All Attributes Arranged in Pairs

Operation of concatenation forms derivative attributes: $A_1A_2 = A_1, A_2$, where A_1A_2 – new attribute value, A_1 – attribute value of the 1st element, A_2 – attribute value of the 2nd element of system. The predicting results were worse than in previous case.

Table 3.
Results of Examination for Learning Set B3 and Testing Set B4 with Use of Concatenation of All Attributes

| Learning Set | | Testing Set | | Violations | | Accuracy, % |
|--------------|------------|-------------|------------|---|--|----------------|
| Formers | Nonformers | Formers | Nonformers | Formers | Nonformers | |
| B3 | | B4 | | Al-Au; Al-Cu; Al-Nb; As-Tc; Au-Hg; B-Re; Ba-C; Ba-Mg; Bi-Cs; Bi-In; Bi-Rh; Ge-Mo; Ge-Os; Hg-Rh; In-Nb; Mo-U; Mo-Zr; Pb-Rh; Pd-Tl; Re-Si; V-Zr; | Ag-Cr; Ag-Mn; Ag-Si; Au-Ge; B-Cu; Bi-Ta; Cd-Mn; Cr-W; Cs-Pr; Cu-Ni; Fe-In; Ge-Sb; Hf-Mg; In-V; K-Mg; Li-Mg; Li-Pu; Mg-Mo; Nb-Tl; Sc-U; | 94.0751 |
| 923 | 410 | 473 | 219 | | | |

The two-dimensional projection for binary systems (Figure 1) was constructed. White and blue circles mark testing systems. The greatest value of similarity degree for each testing system to learning systems with compound formation is plotted on the Y-axis. The greatest value of similarity degree for each testing system to learning systems without compound formation is plotted on the X-axis. Figure 1 illustrates that systems marked by white circles, which are above the line drawn from the origin of the coordinates at an angle of 45°, and systems marked by blue circles, which are under this line, were recognized with error.

5.1.1.2. Differences of Attribute Values

From the physical-chemical standpoint the differences of elementary properties are the best parameters, which determines the compound formation. Indeed this operation allowed obtaining the good predicting results.

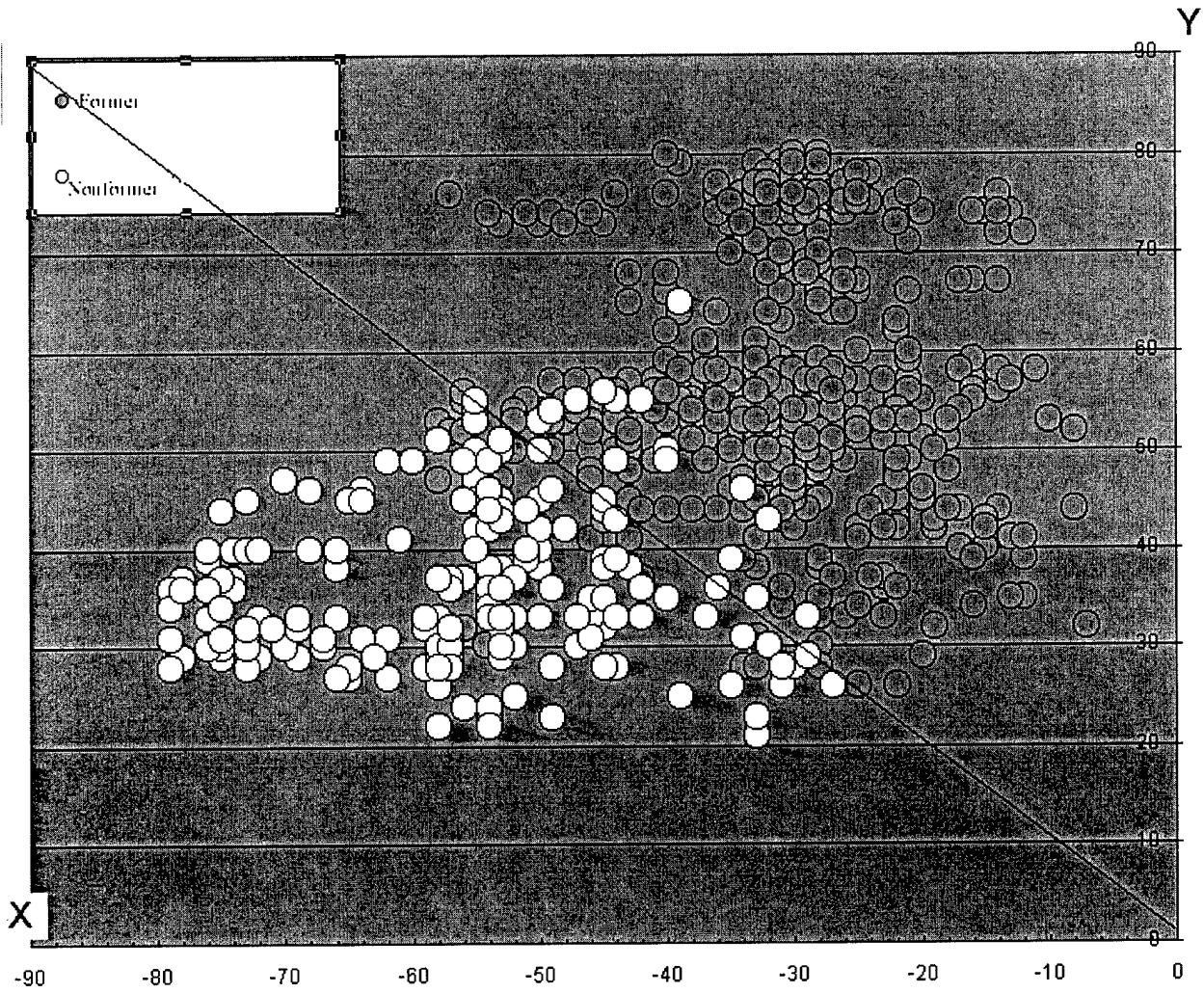


Figure 1. Illustration of Results 5.1.1.1.

The initial elementary properties E_i ($i = 1-12$) and S_i ($i = 1-4$) were substituted for differences of corresponding properties of elements A and B:

$$\begin{aligned} E_{iA} - E_{iB} \\ S_{iA} - S_{iB} \end{aligned}$$

Table 4.

Results of Examination for Learning Set B3 and Testing Set B4 with Use of Differences of Attribute Values

| Learning Set | | Testing Set | | Violations | | Accuracy, % |
|--------------|------------|-------------|------------|--|----------------------|-------------|
| Formers | Nonformers | Formers | Nonformers | Formers | Nonformers | |
| B3 | | B4 | | Al-Pt; Be-Cu; Be-V; Bi-Pd; Cd-Ni; Pb-Pd; Pb-Rh; Pd-Tl; Pt-Si; Re-V; Sb-Tl; | Ag-Cu; Ag-Si; Ca-Pu; | 97.8852 |
| 923 | 410 | 473 | 219 | | | |

Appendices 17-23 contain some interesting two-dimensional maps for binary and ternary systems. The transformed attributes are used as axes of these maps. The red stars correspond to formers but green circles are nonformers.

It should be noted that the difference of electronegativities of elements A and B determines chemical bonding strength. The difference of radii of these elements is determining factor in formation of their homogeneous solution also. By this means in deciding on the best transformation one should take account of physical-chemical ideas on nature of chemical bonding.

5.2. Ternary Systems Prediction

We predicted two testing sets: T3 and T5 using T4 learning set.

Table 5.

Results of Examination for Learning Set T4 and Testing Set T3

| Learning Set | | Testing Set | | Violations | | Accuracy, % |
|--------------|------------|-------------|------------|--|--|-------------|
| Formers | Nonformers | Formers | Nonformers | Formers | Nonformers | |
| 1 | | 2 | | 3 | 4 | 5 |
| T4 | | T3 | | | | |
| 2835 | 1443 | 64 | 391 | Ag-Al-Mn; Ag-Al-Si; Al-Mo-Ti; Al-Ta-Ti; B-Cr-V; Bi-In-Sb; C-Cr-V; C-Th-Y; Cu-Li-Mg; Ge-Nb-Zr; Ge-Ni-Pd; Nb-Ni-V; Pd-Pt-Sn; | Ag-Al-Co; Ag-Al-Fe; Ag-Al-Ti; Ag-C-Zr; Ag-Cd-Sn; Ag-Sb-Zn; Al-B-V; Al-C-W; Al-Cd-Mg; Al-Fe-Nb; Al-Mg-Zr; Al-Ni-Pt; Al-Ni-V; Al-Ni-W; As-Cd-Te; As-Cu-Sn; As-Ga-Si; As-In-Sn; As-Te-Zn; Au-Bi-Se; Au-C-Zr; Au-In-Sb; Au-Sb-Si; B-C-Hf; B-C-Zr; B-Ce-V; B-Cr-N; B-Cr-Nb; B-Fe-Mn; B-Ga-Mn; B-Gd-Zr; B-Ge-Mn; B-N-W; B-Re-W; Bi-Cd-Pb; Bi-In-Pb; Bi-Pb-Sn; C-Cr-Hf; C-Cr-Re; C-Cr-Ti; C-Cr-W; C-Cr-Zr; C-Fe-Ti; C-Ge-Si; C-Hf-Mo; C-Ir-Mo; C-Mn-Ni; C-Mo-Os; C-Mo-Ru; C-Nb-Re; C-Pt-Si; C-Re-Si; C-Re-V; C-W-Zr; Ca-Cu-Mg; Ca-Li-Mg; Cd-In-Sn; | 66.5934 |

(continuation)

| 1 | 2 | 3 | 4 | 5 |
|---|---|---|--|---|
| | | | Cd-Sb-Se; Cd-Sb-Sn; Co-Cr-Fe; Co-Cr-Ni; Co-Cr-Re; Cd-Se-Sn; Ce-Fe-Hf; Ce-Fe-Nb; Ce-Fe-Re; Ce-Fe-Ta; Ce-Mn-Y; Co-Cr-Zr; Co-Cu-Mn; Co-Cu-Mo; Co-Fe-Mn; Co-Fe-Mo; Co-Fe-Nb; Co-Fe-V; Co-Fe-W; Co-Fe-Zn; Co-Fe-Zr; Co-Mn-Pd; Co-Mo-Ta; Co-Nb-Pt; Co-Nb-Ta; Co-Ni-Sb; Co-Ni-Ta; Co-Ni-V; Co-Ni-W; Cr-Fe-Mn; Cr-Fe-Mo; Cr-Fe-Nb; Cr-Fe-Ni; Cr-Fe-Re; Cr-Fe-Ti; Cr-Fe-W; Cr-Mn-Ti; Cr-Nb-Zr; Cr-Ni-Pt; Cr-Ni-Ru; Cr-Ni-Ti; Cr-Ni-W; Cu-Fe-Mn; Cu-Fe-Mo; Cu-Fe-Zn; Cu-La-Mn; Cu-Mg-Ni; Cu-Mn-Zn; Cu-Mo-Ni; Cu-Nb-Ni; Cu-Ni-V; Cu-Ni-W; Cu-Ni-Zn; Fe-Mn-Ni; Fe-N-Si; Fe-Ni-Ti; Fe-Ni-W; Fe-Pd-Y; Fe-Sm-Y; Fe-Ta-Y; Ga-Pb-Te; Hf-Mo-Re; Hf-Mo-Ru; Hf-Re-W; Hf-Ru-Ti; Hf-Ru-V; Hf-Ru-Zr; In-Sb-Sn; Ir-Ru-Ti; Ir-Ru-Zr; Mo-Ni-V; Mo-Re-Zr; Mo-U-Zr; Ni-Rh-Ta; Ni-Ti-Zr; Ni-W-Zr; Re-Ru-Ti; Re-V-Zr; Re-W-Zr; Si-Ti-V; Si-Ti-W; Te-Tl-Zn; | |

Table 6.

Results of Examination for Learning Set T4 and Testing Set T5

| Learning Set | | Testing Set | | Violations | | Accuracy, % |
|--------------|------------|-------------|------------|--|--|-------------|
| Formers | Nonformers | Formers | Nonformers | Formers | Nonformers | |
| | T4 | | T5 | Al-B-Be; Al-Fe-Si; Al-Mn-Si; Al-Mn-Zn; As-Cd-Ge; As-Co-Re; As-Ni-Re; As-Sn-Zn; B-Ir-Zn; Bi-Ni-Rh; Bi-Rh-Si; Ca-O-Ti; Cd-Pt-Zn; Co-Rh-Sn; Ga-Ge-Mn; Ga-Nb-Ti; H-Th-Ti; Li-Mg-Zn; Mo-Nb-S; Sb-Sn-Zn; Te-Ti-Zr; | Ag-Au-Si; Cr-Cu-Mo; Fe-K-Mg; Fe-Li-Na; Pb-Si-Tl; | 98.7941 |
| 2835 | 1443 | 1429 | 727 | | | |

5.3. Quaternary Systems Prediction

Problem presented to computer program system was stated as: "Evaluate" Q1 using results learned from B3 and/or T4. In addition the evaluation of Q2 using B3 and T4 learning sets was carried out. The recognition on the base B3 and T4 learning sets furnished the desired 100% accuracy result. The test set Q2 was derived from 634 binary nonformers contained in the BINARY ALLOY PHASE DIAGRAMS CD-ROM (Villars et al. (2000)). Thus the program system *ANALOGY* is able to deduce Miedema rule: a ternary or quaternary system is a nonformer when all its binary (or ternary) boundary systems are also nonformers (Villars et al. (2000)).

Table 7.

Results of Examination for Learning Set B3 and Testing Set Q1

| Learning Set | | Testing Set | | Violations | Accuracy, % |
|--------------|------------|-------------|------------|---|-------------|
| Formers | Nonformers | Formers | Nonformers | Formers | |
| 1 | | 2 | | 3 | 4 |
| 923 | 410 | 2747 | Q1 | Ag-Al-Cs-F; Ag-As-Pb-S; Ag-Au-Ba-O; Ag-Au-Cl-Cs; Ag-Au-I-Rb; Ag-Ba-Ge-S; Ag-Ba-Se-Y; Ag-Bi-Cr-O; Ag-Br-O-Pb; Ag-C-Co-N; Ag-C-Co-O; Ag-Cd-O-V; Ag-Cl-O-Pb; Ag-Co-O-P; Ag-Cs-F-Fe; Ag-Er-O-W; Ag-F-Na-Zr; Ag-Ge-O-S; Ag-La-Nb-O; Ag-La-O-Ti; Ag-Mo-O-P; Ag-O-P-V; Ag-O-Tl-V; Ag-P-S-V; Ag-P-Se-V; Ag-Pb-S-Sb; Al-As-K-Na; Al-As-Na-O; Al-B-K-O; Al-Ba-Ge-O; Al-Ba-In-O; Al-Be-Mg-O; Al-Be-O-Si; Al-Cd-O-Si; Al-Cs-F-Na; Al-Cs-H-N; Al-Cs-N-O; Al-Cs-O-Si; Al-Cu-F-K; Al-F-K-Li; Al-F-K-Na; Al-F-Na-Ni; Al-F-Na-Rb; Al-F-Na-Zn; Al-Ge-K-O; Al-Ge-Li-O; Al-Ge-Na-O; Al-Ge-O-Y; Al-H-K-N; Al-H-K-O; Al-H-N-Na; Al-Ho-O-Pb; Al-K-Li-O; Al-K-Na-P; Al-K-O-P; Al-K-O-S; Al-K-O-Si; Al-Lu-O-Pb; Al-Na-Nb-O; Al-Na-O-P; Al-Na-O-Si; Al-Nb-O-Rb; As-Bi-Br-Hg; As-C-F-N; As-Cl-O-Pb; As-Cl-P-Sb; As-Cs-Li-O; As-Cs-O-Ti; As-Cu-O-Pb; As-Cu-Pb-S; As-Eu-O-Ta; As-F-S-Sb; As-F-Sb-Se; As-H-O-Tl; As-Hg-S-Tl; As-K-Nb-O; As-Li-Mn-O; As-Li-Na-O; As-Mn-O-Sr; As-Mn-S-Tl; As-Na-O-Sb; As-Na-O-V; As-O-Tl-Zn; As-Pb-S-Tl; As-S-Sb-Tl; As-S-Sn-Tl; Au-Ba-Na-O; Au-C-Cl-O; Au-C-Cl-S; Au-C-N-Tl; Au-Co-Na-O; B-Ba-Ca-O; B-Ba-O-Yb; | 71.8864 |

(continuation)

| 1 | 2 | 3 | 4 |
|---|---|--|---|
| | | B-Ba-O-Zn; B-Ca-Mn-O; B-Ca-O-Sn; B-Cd-Li-O; B-Cl-O-Pb; B-Cs-I-Zr; B-Cs-Li-O; B-Fe-Mg-O; B-H-N-Zn; B-H-O-Pb; B-H-O-Zn; B-K-Li-O; B-K-Nb-O; B-K-O- Ta; B-Li-Na-O; B-Mg-Mn-O; B-Na-O-Rb; B-Nd-O-Sr; B-Ni-O-Sn; B-O-Pb-S; B-O-Se-Zn; Ba-Be-La-O; Ba-Be-Nd-O; Ba-Be-O-Si; Ba-Bi-O-Ru; Ba-Bi-O-V; Ba-Br-Fe-O; Ba-C-Ca-O; Ba-Ca-Cr-F; Ba-Ca-Fe-O; Ba-Ca-Ir-O; Ba-Ca-O-Pd; Ba-Ca-O-Ru; Ba-Ca-O-Si; Ba-Ca-O-W; Ba-Cd-Ge-S; Ba-Cl-Fe-O; Ba-Co-Eu-O; Ba-Co-F-Fe; Ba-Co-O-Ru; Ba-Cr-Cs-F; Ba-Cr-F-Li; Ba-Cr-F-Na; Ba-Cu-F-Fe; Ba-Cu-Li-O; Ba-Cu-Lu-O; Ba-Cu-O-Tl; Ba-Cu-O-Tm; Ba-Cu-O-V; Ba-Cu-O-W; Ba-Cu-O-Yb; Ba-Eu-O-Re; Ba-F-Fe-Mn; Ba-F-Fe-Na; Ba-F-Fe-Zn; Ba-F-Li-Zr; Ba-F-Na-Zr; Ba-F-O-Ti; Ba-F-O-V; Ba-Fe-H-O; Ba-Fe-Li-N; Ba-Fe-O-Si; Ba-Fe-O-Sr; Ba-H-O-V; Ba-In-O-Zn; Ba-Ir-Ni-O; Ba-Ir-O-Y; Ba-La-O-Re; Ba-Li-N-Ta; Ba-Li-Nb-O; Ba-Li-O-Ta; Ba-Lu-O-W; Ba-Mg-O-V; Ba-Mn-Nd-O; Ba-Mn-O-V; Ba-Mo-Nd-O; Ba-Na-Nb-O; Ba-Na-Ni-O; Ba-Nb-O-Si; Ba-Nb-O-Ti; Ba-Nd-O-Si; Ba-Nd-O-Ti; Ba-Ni-O-Os; Ba-Ni-O-Ru; Ba-O-P-Ti; Ba-O-P-V; Ba-O-Pd-Y; Ba-O-Re-Sm; Ba-O-Re-Tm; Ba-O-Re-Yb; Ba-O-Si-Sn; Ba-O-Si-V; Ba-O-Si-Zn; Ba-O-Sr-U; Ba-O-Tl-V; Ba-O-Y-Zn; Be-Ca-O-Si; Be-Cd-F-Rb; Be-Cs-F-Li; Be-Dy-Ge-O; Be-Er-Ge-O; Be-Eu-Gd-O; Be-F-Ho-K; Be-F-K-Li; Be-F-K-Mg; Be-F-Li-Na; Be-F-Li-Rb; Be-F-Li-Zr; Be-F-Na-Rb; Be-F-Na-Th; Be-Gd-Ge-O; Be-Ge-La-O; Be-Ge-Na-O; Be-Ge-O-Pr; Be-Ge-O-Sm; Be-Ge-O-Y; Be-K-Na-O; Be-Li-O-Si; Be-Na-O-P; Be-Na-O-Rb; Be-Na-O-Sb; Be-Na-O-Si; Be-Nd-O-Si; Be-O-Rb-Si; Bi-C-Ir-O; Bi-C-Mn-O; Bi-C-O-Ru; Bi-Ca-Nb-O; Bi-Ca-O-V; Bi-Cl-Cs-Sb; | |

(continuation)

| 1 | 2 | 3 | 4 |
|---|---|--|---|
| | | Bi-Cl-In-S; Bi-Cl-Nb-O; Bi-Co-O-S; Bi-Cr-H-O; Bi-Cr-O-Rb; Bi-Cr-O-Tl; Bi-Cs-F-Tl; Bi-Cs-Mo-O; Bi-Cu-O-V; Bi-Cu-O-W; Bi-Cu-Pb-S; Bi-Fe-Mn-O; Bi-Fe-Mo-O; Bi-H-Mo-O; Bi-In-Pb-S; Bi-K-O-V; Bi-Mg-O-V; Bi-Mn-O-Sr; Bi-Mo-O-Rb; Bi-O-P-Pb; Bi-O-Sr-V; Bk-Cl-Cs-Na; Br-C-O-Os; Br-Cl-Cs-Mo; Br-Cs-Er-Nb; Br-Cs-Eu-Nb; Br-O-Pb-Tl; C-Ca-O-Sr; C-Cl-Cu-O; C-Cl-N-Sb; C-Cl-Ni-O; C-Cl-O-Os; C-Cl-O-Rh; C-Cl-Pb-Pb; C-Co-Cu-O; C-Co-O-Pb; C-Co-O-Ru; C-Co-O-Tl; C-Cs-I-Zr; C-Cs-O-Rh; C-Cu-Fe-N; C-Cu-H-N; C-Cu-H-O; C-Cu-K-N; C-Cu-K-O; C-Cu-N-S; C-Cu- Na-O; C-Cu-O-Tl; C-F-Ge-S; C-F-Ge-Se; C-Fe-N-Tl; C-Fe-O-Pb; C-H-N-Ni; C-H-O- Os; C-Hg-O-Os; C-In-O-Re; C-K-N-Nb; C-K-N-Ni; C-Li-Na-O; C-Mg-Na-O; C-Mo-N-Tl; C-N-Ni-S; C-N-Tl-Zn; C-O-Os-Pt; C-O-Os-S; C-O-Pt-Re; Ca-Ce- Li-O; Ca-Cl-Fe-O; Ca-Cl-O-V; Ca-Cr-F-Li; Ca-Cs-H-N; Ca-Cs-Nb-O; Ca-Cu-Mn-O; Ca-Cu-O-V; Ca-Fe-H-Mg; Ca-Fe-O-Si; Ca-H-O-V; Ca-K-O-V; Ca-Mg-O-V; Ca-Mn-O-Si; Ca-Mo-O-Sr; Ca-Mo-O-U; Ca-N-Sr-W; Ca-Na-O-V; Ca-O-P-V; Ca-O-Sc-Si; Ca-O-Si-Sn; Ca-O-Si-Sr; Ca-O-Si-Zn; Ca-O-Sr-U; Ca-O-Sr-W; Cd-Cl-K-Rb; Cd-Cu-O-V; Cd-Ga-In-S; Cd-Ge-H-O; Cd-Ge-Li-O; Cd-H-O-Pb; Cd-H-O-Si; Cd-H-O-V; Cd-K-O-V; Cd-Li-O-Si; Cd-Mo-O-Th; Cd-N-O-Tl; Cd-Na-O-V; Cd-O-P-Si; Cd-O-P-V; Cd-O-S-Tl; Cd-O-Sr-V; Ce-Cl-O-Ta; Ce-Cr-H-O; Ce-Cr-O-S; Ce-H-O-Ti; Cl-Cr-Cs-O; Cl-Cr-F-K; Cl-Cs-K-Sc; Cl-Cs-Li-Lu; Cl-Cs-Li-Sc; Cl-Cs-Li-Y; Cl-Cs-Lu-Nb; Cl-Cs-Mn-Rb; Cl-Cs-Na-U; Cl-Cs-Na-Y; Cl-Cs-Sb-Tl; Cl-Dy-O-Ti; Cl-F-P-Sb; Cl-Fe-K-Na; Cl-Fe-O-Pb; Cl-Gd-O-W; Cl-Hg-O-Pb; Cl-In-Li-Nb; Cl-La-Nb-O; Cl-La-O-Ta; Cl-La-O-Ti; Cl-La-O-U; Cl-La-O-W; | |

(continuation)

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| | | Cl-Na-Rb-Tm; Cl-Nd-O-Ta; Cl-Nd-O-Ti; Cl-Nd-O-U; Cl-O-P-Sb; Cl-O-Pb-Sb; Cl-O-Pb-Si; Cl-O-Pr-Ti; Cl-O-Pr-U; Cl-O-Pr-W; Cl-O-Sm-Ti; Cl-O-Sr-V; Co-Cs-F-K; Co-F-Fe-Na; Co-F-Li-Na; Co-Ge-Na-O; Co-K-Na-O; Co-Li-Na-O; Co-N-Na-O; Co-Na-O-Rb; Co-Ni-O-P; Cr-Cs-F-K; Cr-Cs-F-Na; Cr-Cs-Li-O; Cr-Cs-Nd-O; Cr-Cu-F-K; Cr-Cu-F-Na; Cr-Cu-H-O; Cr-Cu-P-S; Cr-Er-H-O; Cr-Eu-H-O; Cr-F-K-Na; Cr-F-K-Pb; Cr-F-K-Rb; Cr-F-Li-Na; Cr-F-Mn-Na; Cr-F-Na-Rb; Cr-F-Na-Sr; Cr-Gd-H-O; Cr-H-La-O; Cr-H-Na-O; Cr-H-Nd-O; Cr-H-O-Pr; Cr-H-O-Tb; Cr-H-O-Th; Cr-H-O-U; Cr-H-O-Yb; Cr-K-La-O; Cr-K-Li-O; Cr-K-Na-O; Cr-K-O-P; Cr-K-O-Tb; Cr-La-O-S; Cr-Li-O-Rb; Cr-Na-O-P; Cr-Na-O-Rb; Cr-Na-O-Si; Cr-Nd-O-S; Cr-O-P-Pb; Cr-O-P-Tl; Cr-S-Tl-V; Cs-Er-F-Na; Cs-F-Fe-Na; Cs-F-Fe-Tl; Cs-F-K-Mo; Cs-F-K-Pd; Cs-F-K-Rh; Cs-F-K-Ti; Cs-F-K-V; Cs-F-K-Y; Cs-F-Mn-Na; Cs-F-Mo-O; Cs-F-Mo-Tl; Cs-F-Na-Sc; Cs-F-Na-Ti; Cs-F-Na-Y; Cs-F-NiP-d; Cs-F-O-V; Cs-F-Pd-Rb; Cs-F-Rb-Si; Cs-F-Rb-Y; Cs-F-Tl-V; Cs-H-Li-N; Cs-H-Li-O; Cs-H-Mg-N; Cs-H-N-P; Cs-H-O-V; Cs-I-Mn-Zr; Cs-Li-Mn-O; Cs-Li-Mo-O; Cs-Li-O-P; Cs-Li-O-S; Cs-Li-O-Tb; Cs-Li-O-Ti; Cs-Lu-O-W; Cs-Mn-O-V; Cs-Mo-Nd-O; Cs-Mo-O-P; Cs-Mo-O-Pr; Cs-Mo-O-S; Cs-Mo-O-V; Cs-Mo-O-Zr; Cs-Na-O-V; Cs-Nb-O-P; Cs-Nb-O-S; Cs-Nb-O-Si; Cs-Nb-O-U; Cs-O-P-Sb; Cs-O-P-Ti; Cs-O-P-V; Cs-O-S-V; Cs-O-Sb-Si; Cs-O-U-V; Cu-Dy- O-W; Cu-Er-O-W; Cu-Eu-Mo-O; Cu-F-Fe-Na; Cu-F-Nb-O; Cu-F-O-W; Cu-Fe-O-V; Cu-Gd-Mo-O; Cu-Gd-O-W; Cu-H-Mo-O; Cu-H-Na-O; Cu-Ho-Mo-O; Cu-Ho-O-W; Cu-I-Na-O; Cu-In-Mo-O; Cu-In-O-W; Cu-K-Mo-O; Cu-K-Nb-S; Cu-K-Nb-Se; Cu-K-O-Ta; Cu-K-O-V; | |

(continuation)

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| | | Cu-K-Se-Ta; Cu-La-Mo-O; Cu-La-O-W; Cu-Li-O-P; Cu-Li-O-Se; Cu-Li-O-Si; Cu-Li-O-V; Cu-Li-P-Tb; Cu-Lu-O-W; Cu-Mg-Mn-O; Cu-Mo-Na-O; Cu-Mo-Nd-O; Cu-Mo-O-Pr; Cu-Mo-O-Sm; Cu-Mo-O-Tb; Cu-Mo-O-Y; Cu-Mo-O-Yb; Cu-O-Pb-Se; Cu-O-S-Tl; Cu-O-Sm-W; Cu-O-Sr-V; Cu-O-Sr-W; Cu-O-Tb-W; Cu-O-TmW; Cu-O-W-Y; Cu-O-W-Yb; Cu-Pb-S-Sb; Dy-O-Rb-W; Er-F-Na-Rb; Er-H-O-V; Er-K-O-P; Er-Na-O-P; Er-Na-O-V; Eu-H-Mo-O; Eu-K-Mo-O; F-Fe-K-Na; F-Fe-Li-Mn; F-Fe-Li-Na; F-Fe-Li-Rb; F-Fe-Mn-Pb; F-Fe-Na-Ni; F-Fe-Na-Rb; F-Fe-Na-Sr; F-H-Na-Ti; F-Hg-O-Zn; F-Hg-S-Si; F-K-Li-Nd; F-K-Li-Y; F-K-Mo-Na; F-K-Mo-Rb; F-K-Mo-Tl; F-K-Na-Rh; F-K-Na-Sc; F-K-Na-Th; F-K-Na-Y; F-K-Nb-O; F-K-O-Ta; F-K-O-V; F-K-Pd-Rb; F-K-Rb-Rh; F-K-Rb-Ti; F-K-Rb-V; F-K-Rb-Y; F-Li-Mn-Ti; F-Li-Na-Sc; F-Li-Na-Th; F-Li-Na-Ti; F-Li-Na-V; F-Li-Na-Y; F-Li-Na-Yb; F-Mn-Rb-Zr; F-Mo-Na-Rb; F-Mo-Na-Tl; F-Na-Nb-O; F-Na-Rb-Rh; F-Na-Rb-V; F-Na-Rb-Y; F-O-P-Sb; F-O-Pb-Ta; F-O-Re-Sb; Fe-Ge-La-O; Fe-Ge-Na-O; Fe-H-Mg-Yb; Fe-K-Li-O; Fe-K-Na-O; Fe-La-Mn-S; Fe-La-O-Se; Fe-Li-Na-O; Fe-Mg-O-Si; Fe-Na-O-P; Fe-Na-O-Se; Fe-Na-O-Si; Fe-Na-O-Ti; Fe-Pb-S-Sb; Ga-Ge-O-Pb; Ga-La-Mn-S; Ga-O-S-Tl; Gd-K-O-W; Ge-K-Nb-O; Ge-K-O-Ta; Ge-Li-Na-O; Ge-Li-O-Ta; Ge-Na-Nb-O; Ge-Na-O-Sr; Ge-Na-O-Ti; Ge-Na-O-Zr; Ge-O-P-Si; Ge-O-Rb-Sn; Ge-O-Sn-Tl; Ge-O-Ta-Tl; Ge-Pb-S-Tl; H-Hg-O-Sb; H-K-Li-N; H-K-Mg-N; H-La-O-Re; H-La-O-V; H-Li-Mn-O; H-Li-N-Na; H-Mg-N-Rb; H-Mn-O-Pb; H-Mn-O-Sr; H-Mo-Na-O; H-N-Na-Ni; H-Na-Nb-O; H-Na-O-V; H-Nd-O-V; H-O-P-Sb; H-O-Rb-V; H-O-Si-Zn; H-O-Sn-Zn; H-O-Sr-V; H-O-Ti-V; H-O-V-Y; H-O-V-Yb; Hg-I-Pb-S; Ho-K-O-P; I-K-Li-O; | |

(continuation)

| 1 | 2 | 3 | 4 |
|---|---|--|---|
| | | I-Na-Ni-O; In-K-Mo-O; In-Li-Mo-O; In-Mo-Na-O; In-Mo-O-Rb; Ir-N-O-Tl; Ir-Na-O-Rb; K-La-O-P; K-La-O-Ti; K-Li-Mn-O; K-Li-Nb-O; K-Li-O-P; K-Li-O-S; K-Li-O-Si; K-Li-O-Ta; K-Li-O-Te; K-Li-O-U; K-Li-O-V; K-Li-O-Zr; K-Mg-Mo-O; K-Mg-O-V; K-Mn-Mo-O; K-Mo-OTh; K-Mo-O-Y; K-Mo-O-Zr; K-Na-O-P; K-Na-O-Ti; K-Na-O-V; K-Nb-O-P; K-Nb-O-S; K-Nb-O-Si; K-O-P-Sb; K-O-P-Ta; K-O-P-Y; K-O-P-Zr; K-O-S-Ta; K-O-Sb-Si; K-O-Si-Sn; K-O-Si-Ta; K-O-Th-V; K-O-U-V; K-O-W-Y; K-Rb-Re-S; La-Li-Mo-O; La-Li-Nb-O; La-Li-O-Ta; La-Mo-Na-O; La-Na-O-V; La-Na-O-W; La-Ni-O-Ru; La-O-S-Ta; La-O-S-V; Li-Mg-O-S; Li-Mg-O-V; Li-Mn-O-P; Li-Mn-O-V; Li-Mn-O-Zn; Li-Mo-O-Rb; Li-Mo-O-Y; Li-Na-O-P; Li-Na-O-S; Li-Na-O-V; Li-Na-O-W; Li-Nb-O-W; Li-O-P-V; Li-O-P-Zr; Li-O-Sb-W; Li-O-Si-Zn; Li-O-Te-V; Li-O-Te-Zr; Li-O-W-Yb; Lu-Mo-Na-O; Lu-Na-O-W; Mg-Mn-O-P; Mg-Mn-O-Si; Mg-Na-O-P; Mg-Na-O-S; Mg-Na-O-Si; Mg-Na-O-V; Mg-O-Si-V; Mn-Na-Ni-O; Mn-O-Pb-Si; Mo-N-Na-O; Mo-Na-O-P; Mo-Na-O-Rb; Mo-Na-O-Tb; Mo-Na-O-Zr; Mo-O-P-Rb; Mo-O-P-Sr; Mo-O-P-Tl; Mo-O-Pr-Rb; Mo-O-Rb-S; N-Ni-O-Pb; Na-Nb-O-P; Na-Nd-O-Ti; Na-Nd-O-V; Na-Ni-O-Rb; Na-O-P-Sb; Na-O-P-Sr; Na-O-P-Th; Na-O-P-Ti; Na-O-P-V; Na-O-P-Zr; Na-O-Pb-Rb; Na-O-Ru-Sr; Na-O-S-Ti; Na-O-S-V; Na-O-Sb-Si; Na-O-Sb-Ti; Na-O-Sc-V; Na-O-Si-Sr; Na-O-Si-Ti; Na-O-Si-Zr; Na-O-Tb-W; Na-O-Te-V; Na-O-W-Y; Na-O-W-Zr; Nb-Nd-O-Sr; Nb-O-Rb-S; Nb-O-Rb-Si; Nb-O-Sm-Sr; Nb-O-Sr-Tm; Nb-O-Ti-U; Ni-O-P-Pb; Ni-O-Re-Sr; O-P-Pb-Zn; O-P-Rb-V; O-P-Sr-Ti; O-P-Sr-V; O-P-Tl-Zn; O-Pb-Si-Zn; O-Pr-Se-Ta; O-Rb-Si-Sn; O-Re-Sr-Y; O-Si-Sr-V; O-Sm-Tl-W; O-Sr-Tl-V; P-Rb-S-V; P-S-Sn-Tl; | |

Table 8.

Results of Examination for Learning Set T4 and Testing Set Q1

| Learning Set | | Testing Set | | Violations | Accuracy, % |
|--------------|------------|-------------|------------|--|-------------|
| Formers | Nonformers | Formers | Nonformers | Formers | |
| 1 | | 2 | | 3 | 4 |
| T4 | | Q1 | | | |
| 2835 | 1443 | 2747 | | Ag-As-Cu-O; Ag-Au-Cl-Cs; Ag-Ba-La-S; Ag-Bi-Cr-O; Ag-C-Co-O; Ag-Cd-O-V; Ag-In-O-W; Ag-La-O-Ti; Ag-O-Tl-V; Al-Ba-La-O; Al-Ca-F-Na; Al-Cd-O-Si; Al-Co-O-Si; Al-Fe-Mg-Si; Al-Fe-O-Si; As-Bi-Br-Hg; As-Bi-H-O; As-Cl-P-Sb; As-Cu-O-Pb; As-Eu-O-Ta; As-F-Sb-Sn; As-Pb-S-Tl; As-S-Sb-Tl; Au-C-Cl-S; B-Ba-Ca-O; B-C-H-Ni; B-Ca-Na-O; B-Cl-O-Pb; B-Co-Cu-O; B-Cs-Li-O; B-H-O-Pb; B-H-O-Tl; B-H-O-Zn; B-K-Li-O; B-Nd-O-Sr; B-Ni-O-Sn; Ba-Be-La-O; Ba-Be-Nd-O; Ba-C-Ca-O; Ba-Ca-Cr-F; Ba-Ca-F-Ga; Ba-Ca-Fe-O; Ba-Ca-Ir-O; Ba-Ca-O-Pd; Ba-Ca-O-Ru; Ba-Cr-Cs-F; Ba-Cr-F-Li; Ba-Cr-F-Na; Ba-Cu-La-S; Ba-Eu-O-Re; Ba-F-Fe-Na; Ba-F-O-V; Ba-Fe-O-Sr; Ba-Ga-La-O; Ba-Gd-Mn-O; Ba-Ho-Mn-O; Ba-Ir-O-Sr; Ba-La-O-Pt; Ba-La-O-Re; Ba-La-O-Ru; Ba-La-O-Zn; Ba-Mg-O-V; Ba-Mn-Nd-O; Ba-Mn-O-Sm; Ba-Mo-Nd-O; Ba-Na-Ni-O; Ba-Nd-O-Ti; Ba-Nd-O-Zn; Ba-Ni-O-Os; Ba-O-Ru-Sr; Be-Cs-F-Li; Be-F-K-Zn; Be-F-Li-Na; Be-F-Na-Rb; Be-F-Na-Th; Be-F-Rb-Zn; Be-Ge-O-Sm; Be-K-Na-O; Be-Na-O-P; Be-Na-O-Rb; Be-Na-O-Sb; Be-Na-O-Si; Bi-C-Co-O; Bi-C-O-Ru; Bi-Cr-H-O; Bi-Cr-O-Tl; Bi-Cu-O-V; Bi-Cu-O-W; Bi-F-K-Rb; Bi-Fe-Mo-O; Bi-H-Mo-O; Bi-Nb-O-Pb; Bi-O-Pb-V; Bi-O-Sr-V; Br-C-O-Os; Br-C-O-Ru; Br-O-Pb-Tl; C-Ca-Na-O; C-Cl-Cu-O; C-Cl-O-Os; C-Cl-O-Ru; C-Co-Cu-O; C-Co-O-Ru; C-Cu-Fe-N; C-Cu-H-O; C-Cu-O-Tl; C-H-O-Os; C-Hg-O-Os; C-Hg-O-Ru; C-I-O-Os; C-I-O-Pt; C-I-O-Ru; C-Li-Na-O; C-N-Tl-Zn; C-O-Os-Pt; C-O-Os-S; C-O-Pt-Re; C-O-Ru-S; C-O-Ru-Te; Ca-Cl-O-V; Ca-Cr-F-Li; Ca-Cs-F-Ni; Ca-Cu-O-Ti; Ca-Mg-O-V; Ca-Mo-O-Sr; Ca-Mo-O-U; Ca-N-Sr-W; Ca-Na-O-V; Ca-O-Si-Ti; Ca-O-Sr-W; Cd-Ga-In-S; Cd-H-O-V; Cd-O-P-V; Cd-O-Sr-V; Ce-Cl-O-Ta; Ce-H-O-Ti; Cl-Cr-F-K; Cl-Cs-K-O; Cl-Cs-K-Sc; Cl-Cs-Li-Lu; Cl-Cs-Li-Sc; Cl-Cs-Li-Y; Cl-Fe-K-Na; Cl-Fe-O-Pb; Cl-Gd-O-W; Cl-La-Nb-O; Cl-La-O-Ta; Cl-La-O-Ti; Cl-Nd-O-Ta; Cl-Nd-O-Ti; Cl-Nd-O-U; | 90.7138 |

(continuation)

| 1 | 2 | 3 | 4 |
|---|---|---|---|
| | | Cl-O-Pr-Ti; Cl-O-Pr-U; Co-F-Fe-Na; Cr-Cs-F-K; Cr-Cs-F-Na; Cr-Cs-Li-O; Cr-Cs-Nd-O; Cr-F-K-Na; Cr-F-K-Pb; Cr-F-K-Rb; Cr-F-Li-Na; Cr-F-Na-Rb; Cr-F-Na-Sr; Cr-H-O-Tb; Cr-K-Li-O; Cr-K-Na-O; Cr-Li-O-Rb; Cr-Na-O-Rb; Cr-O-P-Pb; Cs-F-Fe-Na; Cs-F-K-Y; Cs-F-Na-Ti; Cs-F-O-V; Cs-F-Rb-Y; Cs-H-O-V; Cs-Li-O-S; Cs-Li-O-Tb; Cs-Mn-O-V; Cs-Mo-Nd-O; Cs-Mo-O-Pr; Cs-Mo-O-V; Cs-Nb-O-U; Cu-F-Nb-O; Cu-F-O-W; Cu-Fe-O-V; Cu-H-Mo-O; Cu-In-Mo-O; Cu-Mo-O-Sb; Cu-O-Sr-V; Dy-O-Rb-W; Er-H-O-V; Eu-K-Mo-O; F-Fe-K-Na; F-Fe-K-Rb; F-Fe-Mn-Pb; F-Fe-Na-Ni; F-Fe-Na-Rb; F-Fe-Na-Sr; F-H-Na-Ti; F-K-Mo-Na; F-K-Rb-Rh; F-K-Rb-V; F-K-Rb-Y; F-Li-Na-Sc; F-Li-Na-Th; F-Li-Na-Ti; F-Li-Na-V; F-Mo-Na-Rb; F-Mo-Na-Tl; F-Mo-Nb-O; F-Mo-O-Sb; F-O-Re-Sb; Fe-K-Li-O; Fe-K-Na-O; Ga-Ge-O-Pb; Gd-K-O-W; Ge-O-Sn-Tl; H-La-O-V; H-Mn-O-Pb; H-Mn-O-Sr; H-Nd-O-V; H-O-Ti-V; H-O-V-Yb; I-K-Li-O; I-Li-O-Rb; Ir-K-Li-O; K-La-O-Ti; K-Li-Nb-O; K-Li-O-S; K-Li-O-U; K-Li-O-V; K-Mg- Mo-O; K-Mo-O-Sc; K-Mo-O-Th; K-Na-O-Ti; K-O-Th-V; La-Li-Mo-O; La-Li-Nb-O; La-Li-O-Ta; La-Na-O-V; La-O-S-Ta; Li-Mo-O-Sc; Li-Na-O-S; Li-Na-O-V; Lu-Mo-Na-O; Mo-Na-O-Rb; Mo-O-Pr-Rb; Mo-O-Sr-Zn; Na-Nd-O-Ti; Na-Nd-O-V; Na-O-Sc-V; O-P-Tl-Zn; O-Pb-Si-Zn; O-Pb-V-Zn | |

We solved to use no mixed learning set B3+T4 but to compare the results of predictions for set Q1 obtained with the use of learning sets B3 and T4 separately. The basic idea was: the predictions, which have been obtained using different learning sets, must be consistent. Lack of consistency is caused by poor discretization of property values and, consequently, by the "fuzzy" boundaries of the classes of formers and nonformers included into learning sets. Further inaccuracy of measurement of the elementary properties contributes to this fuzziness. If the results of prediction, using the different learning sets, contradict one another in the process of the comparison, then the predictions will be indeterminate. This way is an efficient method of improvement of predicting. It raises an indeterminacy of predicting but decreases the predicting error. This approach simulates the tactics of prudent investigator who prefers do not carry out an experiment that will not give the desired result. Table 9 contains the results of comparison of Q1 set predictions for learning sets B3 and T4.

Table 9.

Results of Comparison of Predictions of Testing Set Q1 for Learning Sets B3 and T4

| Violations | Indetermination |
|---|---|
| 1 | 2 |
| Ag-Au-Cl-Cs; Ag-C-Co-O; Ag-Cd-O-V; Ag-La-O-Ti; Ag-O-Tl-V; Al-Cd-O-Si; As-Bi-Br-Hg; As-Cl-P-Sb; As-Cu-O-Pb; As-Eu-O-Ta; As-Pb-S-Tl; As-S-Sb-Tl; Au-C-Cl-S; B-Ba-Ca-O; B-Cl-O-Pb; B-Cs-Li-O; B-H-O-Pb; B-H-O-Zn; B-K-Li-O; B-Nd-O-Sr; B-Ni-O-Sn; Ba-Be-La-O; Ba-Be-Nd-O; Ba-C-Ca-O; Ba-Ca-Cr-F; Ba-Ca-Fe-O; Ba-Ca-Ir-O; Ba-Ca-O-Pd; Ba-Ca-O-Ru; Ba-Cr-Cs-F; Ba-Cr-F-Li; Ba-Cr-F-Na; Ba-Eu-O-Re; Ba-F-Fe-Na; Ba-F-O-V; Ba-Fe-O-Sr; Ba-La-O-Re; Ba-Mg-O-V; Ba-Mn-Nd-O; Ba-Mo-Nd-O; Ba-Na-Ni-O; Ba-Nd-O-Ti; Ba-Ni-O-Os; Be-Cs-F-Li; Be-F-Li-Na; Be-F-Na-Rb; Be-F-Na-Th; Be-Ge-O-Sm; Be-K-Na-O; Be-Na-O-P; Be-Na-O-Rb; Be-Na-O-Sb; Be-Na-O-Si; Bi-C-O-Ru; Bi-Cr-H-O; Bi-Cr-O-Tl; | Ag-Al-Cs-F; Ag-As-Pb-S; Ag-Au-Ba-O; Ag-Au-I-Rb; Ag-Ba-Ge-S; Ag-Ba-Se-Y; Ag-Bi-Cr-O; Ag-Br-O-Pb; Ag-C-Co-N; Ag-Cl-O-Pb; Ag-Co-O-P; Ag-Cs-F-Fe; Ag-Er-O-W; Ag-F-Na-Zr; Ag-Ge-O-S; Ag-La-Nb-O; Ag-Mo-O-P; Ag-O-P-V; Ag-P-S-V; Ag-P-Se-V; Ag-Pb-S-Sb; Al-As-K-Na; Al-As-Na-O; Al-B-K-O; Al-Ba-Ge-O; Al-Ba-In-O; Al-Be-Mg-O; Al-Be-O-Si; Al-Cs-F-Na; Al-Cs-H-N; Al-Cs-N-O; Al-Cs-O-Si; Al-Cu-F-K; Al-F-K-Li; Al-F-K-Na; Al-F-Na-Ni; Al-F-Na-Rb; Al-F-Na-Zn; Al-Ge-K-O; Al-Ge-Li-O; Al-Ge-Na-O; Al-Ge-O-Y; Al-H-K-N; Al-H-K-O; Al-H-N-Na; Al-Ho-O-Pb; Al-K-Li-O; Al-K-Na-P; Al-K-O-P; Al-K-O-S; Al-K-O-Si; Al-Lu-O-Pb; Al-Na-Nb-O; Al-Na-O-P; Al-Na-O-Si; Al-Nb-O-Rb; As-C-F-N; As-Cl-O-Pb; As-Cs-Li-O; As-Cs-O-Ti; As-Cu-Pb-S; As-F-S-Sb; As-F-Sb-Se; As-H-O-Tl; As-Hg-S-Tl; As-K-Nb-O; As-Li-Mn-O; As-Li-Na-O; As-Mn-O-Sr; As-Mn-S-Tl; As-Na-O-Sb; As-Na-O-V; As-O-Tl-Zn; As-S-Sn-Tl; Au-Ba-Na-O; Au-C-Cl-O; Au-C-N-Tl; Au-Co-Na-O; B-Ba-O-Yb; B-Ba-O-Zn; B-Ca-Mn-O; B-Ca-O-Sn; B-Cd-Li-O; B-Cs-I-Zr; B-Fe-Mg-O; B-H-N-Zn; B-K-Nb-O; B-K-O-Ta; B-Li-Na-O; B-Mg-Mn-O; B-Na-O-Rb; B-O-Pb-S; B-O-Se-Zn; Ba-Be-O-Si; Ba-Bi-O-Ru; Ba-Bi-O-V; Ba-Br-Fe-O; Ba-Ca-O-Si; Ba-Ca-O-W; Ba-Cd-Ge-S; Ba-Cl-Fe-O; Ba-Co-Eu-O; Ba-Co-F-Fe; Ba-Co-O-Ru; Ba-Cu-F-Fe; Ba-Cu-Li-O; Ba-Cu-Lu-O; Ba-Cu-O-Tl; Ba-Cu-O-Tm; Ba-Cu-O-V; Ba-Cu-O-W; Ba-Cu-O-Yb; Ba-F-Fe-Mn; Ba-F-Fe-Zn; Ba-F-Li-Zr; Ba-F-Na-Zr; Ba-F-O-Ti; Ba-Fe-H-O; Ba-Fe-Li-N; Ba-Fe-O-Si; Ba-H-O-V; Ba-In-O-Zn; Ba-Ir-Ni-O; Ba-Ir-O-Y; Ba-Li-N-Ta; Ba-Li-Nb-O; Ba-Li-O-Ta; Ba-Lu-O-W; Ba-Mn-O-V; Ba-Na-Nb-O; Ba-Nb-O-Si; Ba-Nb-O-Ti; Ba-Nd-O-Si; Ba-Ni-O-Ru; Ba-O-P-Ti; Ba-O-P-V; |

(continuation)

| 1 | 2 |
|--|---|
| Bi-Cu-O-V; Bi-Cu-O-W; Bi-Fe-Mo-O; Bi-H-Mo-O; Bi-O-Sr-V; Br-C-O-Os; Br-O-Pb-Tl; C-Cl-Cu-O; C-Cl-O-Os; C-Co-Cu-O; C-Co-O-Ru; C-Cu-Fe-N; C-Cu-H-O; C-Cu-O-Tl; C-H-O-Os; C-Hg-O-Os; C-Li-Na-O; C-N-Tl-Zn; C-O-Os-Pt; C-O-Os-S; C-O-Pt-Re; Ca-Cl-O-V; Ca-Cr-F-Li; Ca-Mg-O-V; Ca-Mo-O-Sr; Ca-Mo-O-U; Ca-N-Sr-W; Ca-Na-O-V; Ca-O-Sr-W; Cd-Ga-In-S; Cd-H-O-V; Cd-O-P-V; Cd-O-Sr-V; Ce-Cl-O-Ta; Ce-H-O-Ti; Cl-Cr-F-K; Cl-Cs-K-Sc; Cl-Cs-Li-Lu; Cl-Cs-Li-Sc; Cl-Cs-Li-Y; Cl-Fe-K-Na; Cl-Fe-O-Pb; Cl-Gd-O-W; Cl-La-Nb-O; Cl-La-O-Ta; Cl-La-O-Ti; Cl-Nd-O-Ta; Cl-Nd-O-Ti; Cl-Nd-O-U; Cl-O-Pr-Ti; Cl-O-Pr-U; Co-F-Fe-Na; Cr-Cs-F-K; Cr-Cs-F-Na; Cr-Cs-Li-O; Cr-Cs-Nd-O; Cr-F-K-Na; Cr-F-K-Pb; Cr-F-K-Rb; Cr-F-Li-Na; Cr-F-Na-Rb; Cr-F-Na-Sr; Cr-H-O-Tb; Cr-K-Li-O; Cr-K-Na-O; Cr-Li-O-Rb; Cr-Na-O-Rb; Cs-F-K-Y; Cs-F-Na-Ti; Cs-F-O-V; Cs-F-Rb-Y; Cs-H-O-V; Cs-Li-O-S; Cs-Li-O-Tb; Cs-Mn-O-V; Cs-Mo-Nd-O; Cs-Mo-O-Pr; Cs-Mo-O-V; Cs-Nb-O-U; Cu-F-Nb-O; Cu-F-O-W; Cu-Fe-O-V; Cu-H-Mo-O; Cu-In-Mo-O; | Ba-O-Pd-Y; Ba-O-Re-Sm; Ba-O-Re-Tm; Ba-O-Re-Yb; Ba-O-Si-Sn; Ba-O-Si-V; Ba-O-Si-Zn; Ba-O-Sr-U; Ba-O-Tl-V; Ba-O-Y-Zn; Be-Ca-O-Si; Be-Cd-F-Rb; Be-Dy-Ge-O; Be-Er-Ge-O; Be-Eu-Gd-O; Be-F-Ho-K; Be-F-K-Li; Be-F-K-Mg; Be-F-Li-Rb; Be-F-Li-Zr; Be-Gd-Ge-O; Be-Ge-La-O; Be-Ge-Na-O; Be-Ge-O-Pr; Be-Ge-O-Y; Be-Li-O-Si; Be-Nd-O-Si; Be-O-Rb-Si; Bi-C-Ir-O; Bi-C-Mn-O; Bi-Ca-Nb-O; Bi-Ca-O-V; Bi-Cl-Cs-Sb; Bi-Cl-In-S; Bi-Cl-Nb-O; Bi-Co-O-S; Bi-Cr-O-Rb; Bi-Cs-F-Tl; Bi-Cs-Mo-O; Bi-Cu-Pb-S; Bi-Fe-Mn-O; Bi-In-Pb-S; Bi-K-O-V; Bi-Mg-O-V; Bi-Mn-O-Sr; Bi-Mo-O-Rb; Bi-O-P-Pb; Bk-Cl-Cs-Na; Br-Cl-Cs-Mo; Br-Cs-Er-Nb; Br-Cs-Eu-Nb; C-Ca-O-Sr; C-Cl-N-Sb; C-Cl-Ni-O; C-Cl-O-Rh; C-Cl-Pb-Pb; C-Co-O-Pb; C-Co-O-Tl; C-Cs-I-Zr; C-Cs-O-Rh; C-Cu-H-N; C-Cu-K-N; C-Cu-K-O; C-Cu-N-S; C-Cu-Na-O; C-F-Ge-S; C-F-Ge-Se; C-Fe-N-Tl; C-Fe-O-Pb; C-H-N-Ni; C-In-O-Re; C-K-N-Nb; C-K-N-Ni; C-Mg-Na-O; C-Mo-N-Tl; C-N-Ni-S; Ca-Ce-Li-O; Ca-Cl-Fe-O; Ca-Cs-H-N; Ca-Cs-Nb-O; Ca-Cu-Mn-O; Ca-Cu-O-V; Ca-Fe-H-Mg; Ca-Fe-O-Si; Ca-H-O-V; Ca-K-O-V; Ca-Mn-O-Si; Ca-O-P-V; Ca-O-Sc-Si; Ca-O-Si-Sn; Ca-O-Si-Sr; Ca-O-Si-Zn; Ca-O-Sr-U; Cd-Cl-K-Rb; Cd-Cu-O-V; Cd-Ge-H-O; Cd-Ge-Li-O; Cd-H-O-Pb; Cd-H-O-Si; Cd-K-O-V; Cd-Li-O-Si; Cd-Mo-O-Th; Cd-N-O-Tl; Cd-Na-O-V; Cd-O-P-Si; Cd-O-S-Tl; Ce-Cr-H-O; Ce-Cr-O-S; Cl-Cr-Cs-O; Cl-Cs-Lu-Nb; Cl-Cs-Mn-Rb; Cl-Cs-Na-U; Cl-Cs-Na-Y; Cl-Cs-Sb-Tl; Cl-Dy-O-Ti; Cl-F-P-Sb; Cl-Hg-O-Pb; Cl-In-Li-Nb; Cl-La-O-U; Cl-La-O-W; Cl-Na-Rb-Tm; Cl-O-P-Sb; Cl-O-Pb-Sb; Cl-O-Pb-Si; Cl-O-Pr-W; Cl-O-Sm-Ti; Cl-O-Sr-V; Co-Cs-F-K; Co-F-Li-Na; Co-Ge-Na-O; Co-K-Na-O; Co-Li-Na-O; Co-N-Na-O; Co-Na-O-Rb; Co-Ni-O-P; Cr-Cu-F-K; Cr-Cu-F-Na; Cr-Cu-H-O; Cr-Cu-P-S; Cr-Er-H-O; Cr-Eu-H-O; Cr-F-Mn-Na; Cr-Gd-H-O; Cr-H-La-O; Cr-H-Na-O; Cr-H-Nd-O; Cr-H-O-Pr; Cr-H-O-Th; Cr-H-O-U; Cr-H-O-Yb; Cr-K-La-O; Cr-K-O-P; Cr-K-O-Tb; Cr-La-O-S; Cr-Na-O-P; Cr-Na-O-Si; Cr-Nd-O-S; Cr-O-P-Pb; Cr-O-P-Tl; Cr-S-Tl-V; Cs-Er-F-Na; Cs-F-Fe-Na; Cs-F-Fe-Tl; Cs-F-K-Mo; Cs-F-K-Pd; Cs-F-K-Rh; Cs-F-K-Ti; Cs-F-K-V; Cs-F-Mn-Na; Cs-F-Mo-O; Cs-F-Mo-Tl; Cs-F-Na-Sc; Cs-F-Na-Y; Cs-F-Ni-Pd; Cs-F-Pd-Rb; Cs-F-Rb-Si; Cs-F-Tl-V; Cs-H-Li-N; Cs-H-Li-O; Cs-H-Mg-N; Cs-H-N-P; Cs-I-Mn-Zr; Cs-Li-Mn-O; Cs-Li-Mo-O; Cs-Li-O-P; Cs-Li-O-Ti; Cs-Lu-O-W; Cs-Mo-O-P; Cs-Mo-O-S; Cs-Mo-O-Zr; Cs-Na-O-V; Cs-Nb-O-P; Cs-Nb-O-S; Cs-Nb-O-Si; Cs-O-P-Sb; Cs-O-P-Ti; Cs-O-P-V; Cs-O-S-V; Cs-O-Sb-Si; Cs-O-U-V; Cu-Dy-O-W; Cu-Er-O-W; Cu-Eu-Mo-O; Cu-F-Fe-Na; Cu-Gd-Mo-O; Cu-Gd-O-W; Cu-H-Na-O; Cu-Ho-Mo-O; |

(continuation)

| 1 | 2 |
|--|--|
| <p>Cu-O-Sr-V; Dy-O-Rb-W; Er-H-O-V; Eu-K-Mo-O; F-Fe-K-Na; F-Fe-Mn-Pb; F-Fe-Na-Ni; F-Fe-Na-Rb; F-Fe-Na-Sr; F-H-Na-Ti; F-K-Mo-Na; F-K-Rb-Rh; F-K-Rb-V; F-K-Rb-Y; F-Li-Na-Sc; F-Li-Na-Th; F-Li-Na-Ti; F-Li-Na-V; F-Mo-Na-Rb; F-Mo-Na-Tl; F-O-Re-Sb; Fe-K-Li-O; Fe-K-Na-O; Ga-Ge-O-Pb; Gd-K-O-W; Ge-O-Sn-Tl; H-La-O-V; H-Mn-O-Pb; H-Mn-O-Sr; H-Nd-O-V; H-O-Ti-V; H-O-V-Yb; I-K-Li-O; K-La-O-Ti; K-Li-Nb-O; K-Li-O-S; K-Li-O-U; K-Li-O-V; K-Mg-Mo-O; K-Mo-O-Th; K-Na-O-Ti; K-O-Th-V; La-Li-Mo-O; La-Li-Nb-O; La-Li-O-Ta; La-Na-O-V; La-O-S-Ta; Li-Na-O-S; Li-Na-O-V; Lu-Mo-Na-O; Mo-Na-O-Rb; Mo-O-Pr-Rb; Na-Nd-O-Ti; Na-Nd-O-V; Na-O-Sc-V; O-P-Tl-Zn; O-Pb-Si-Zn;</p> | <p>Cu-Ho-O-W; Cu-I-Na-O; Cu-In-O-W; Cu-K-Mo-O; Cu-K-Nb-S; Cu-K-Nb-Se; Cu-K-O-Ta; Cu-K-O-V; Cu-K-Se-Ta; Cu-La-Mo-O; Cu-La-O-W; Cu-Li-O-P; Cu-Li-O-Se; Cu-Li-O-Si; Cu-Li-O-V; Cu-Li-P-Tb; Cu-Lu-O-W; Cu-Mg-Mn-O; Cu-Mo-Na-O; Cu-Mo-Nd-O; Cu-Mo-O-Pr; Cu-Mo-O-Sm; Cu-Mo-O-Tb; Cu-Mo-O-Y; Cu-Mo-O-Yb; Cu-O-Pb-Se; Cu-O-S-Tl; Cu-O-Sm-W; Cu-O-Sr-W; Cu-O-Tb-W; Cu-O-Tm-W; Cu-O-W-Y; Cu-O-W-Yb; Cu-Pb-S-Sb; Er-F-Na-Rb; Er-K-O-P; Er-Na-O-P; Er-Na-O-V; Eu-H-Mo-O; F-Fe-Li-Mn; F-Fe-Li-Na; F-Fe-Li-Rb; F-Hg-O-Zn; F-Hg-S-Si; F-K-Li-Nd; F-K-Li-Y; F-K-Mo-Rb; F-K-Mo-Tl; F-K-Na-Rh; F-K-Na-Sc; F-K-Na-Th; F-K-Na-Y; F-K-Nb-O; F-K-O-Ta; F-K-O-V; F-K-Pd-Rb; F-K-Rb-Ti; F-Li-Mn-Ti; F-Li-Na-Y; F-Li-Na-Yb; F-Mn-Rb-Zr; F-Na-Nb-O; F-Na-Rb-Rh; F-Na-Rb-V; F-Na-Rb-Y; F-O-P-Sb; F-O-Pb-Ta; Fe-Ge-La-O; Fe-Ge-Na-O; Fe-H-Mg-Yb; Fe-La-Mn-S; Fe-La-O-Se; Fe-Li-Na-O; Fe-Mg-O-Si; Fe-Na-O-P; Fe-Na-O-Se; Fe-Na-O-Si; Fe-Na-O-Ti; Fe-Pb-S-Sb; Ga-La-Mn-S; Ga-O-S-Tl; Ge-K-Nb-O; Ge-K-O-Ta; Ge-Li-Na-O; Ge-Li-O-Ta; Ge-Na-Nb-O; Ge-Na-O-Sr; Ge-Na-O-Ti; Ge-Na-O-Zr; Ge-O-P-Si; Ge-O-Rb-Sn; Ge-O-Ta-Tl; Ge-Pb-S-Tl; H-Hg-O-Sb; H-K-Li-N; H-K-Mg-N; H-La-O-Re; H-Li-Mn-O; H-Li-N-Na; H-Mg-N-Rb; H-Mo-Na-O; H-N-Na-Ni; H-Na-Nb-O; H-Na-O-V; H-O-P-Sb; H-O-Rb-V; H-O-Si-Zn; H-O-Sn-Zn; H-O-Sr-V; H-O-V-Y; Hg-I-Pb-S; Ho-K-O-P; I-Na-Ni-O; In-K-Mo-O; In-Li-Mo-O; In-Mo-Na-O; In-Mo-O-Rb; Ir-N-O-Tl; Ir-Na-O-Rb; K-La-O-P; K-Li-Mn-O; K-Li-O-P; K-Li-O-Si; K-Li-O-Ta; K-Li-O-Te; K-Li-O-Zr; K-Mg-O-V; K-Mn-Mo-O; K-Mo-O-Y; K-Mo-O-Zr; K-Na-O-P; K-Na-O-V; K-Nb-O-P; K-Nb-O-S; K-Nb-O-Si; K-O-P-Sb; K-O-P-Ta; K-O-P-Y; K-O-P-Zr; K-O-S-Ta; K-O-Sb-Si; K-O-Si-Sn; K-O-Si-Ta; K-O-U-V; K-O-W-Y; K-Rb-Re-S; La-Mo-Na-O; La-Na-O-W; La-Ni-O-Ru; La-O-S-V; Li-Mg-O-S; Li-Mg-O-V; Li-Mn-O-P; Li-Mn-O-V; Li-Mn-O-Zn; Li-Mo-O-Rb; Li-Mo-O-Y; Li-Na-O-P; Li-Na-O-W; Li-Nb-O-W; Li-O-P-V; Li-O-P-Zr; Li-O-Sb-W; Li-O-Si-Zn; Li-O-Te-V; Li-O-Te-Zr; Li-O-W-Yb; Lu-Na-O-W; Mg-Mn-O-Si; Mg-Na-O-P; Mg-Na-O-S; Mg-Na-O-Si; Mg-Na-O-V; Mg-O-Si-V; Mn-Na-Ni-O; Mn-O-Pb-Si; Mo-N-Na-O; Mo-Na-O-P; Mo-Na-O-Tb; Mo-Na-O-Zr; Mo-O-P-Rb; Mo-O-P-Sr; Mo-O-P-Tl; Mo-O-Rb-S; N-Ni-O-Pb; Na-Nb-O-P; Na-Ni-O-Rb; Na-O-P-Sb; Na-O-P-Sr; Na-O-P-Th; Na-O-P-Ti; Na-O-P-V; Na-O-P-Zr; Na-O-Pb-Rb; Na-O-Ru-Sr; Na-O-S-Ti; Na-O-S-V; Na-O-Sb-Si; Na-O-Sb-Ti; Na-O-Si-Sr; Na-O-Si-Ti; Na-O-Si-Zr; Na-O-Tb-W; Na-O-Te-V; Na-O-W-Y; Na-O-W-Zr; Nb-Nd-O-Sr; Nb-O-Rb-S; Nb-O-Rb-Si; Nb-O-Sm-Sr; Nb-O-Sr-Tm; Nb-O-Ti-U; Ni-O-P-Pb; Ni-O-Re-Sr; O-P-Pb-Zn; O-P-Rb-V; O-P-Sr-Ti; O-P-Sr-V;</p> |

(continuation)

| 1 | 2 |
|---|---|
| | O-Pr-Se-Ta; O-Rb-Si-Sn; Ag-As-Cu-O; Ag-Ba-La-S; Ag-Bi-Cr-O; Ag-In-O-W; Al-Ba-La-O; Al-Ca-F-Na; Al-Co-O-Si; Al-Fe-Mg-Si; Al-Fe-O-Si; As-Bi-H-O; As-F-Sb-Sn; B-C-H-Ni; B-Ca-Na-O; B-H-O-Tl; Ba-Ca-F- Ga; Ba-Cu-La-S; Ba-Eu-O-Re; Ba-F-Fe-Na; Ba-F-O-V; Ba-Fe-O-Sr; Ba-Ga-La-O; Ba-Gd-Mn-O; Ba-Ho-Mn-O; Ba-Ir-O-Sr; Ba-La-O-Pt; Ba-La-O-Ru; Ba-La-O-Zn; Ba-Mn-O-Sm; Ba-Nd-O-Zn; Ba-O-Ru-Sr; Be-F-K-Zn; Be-F-Rb-Zn; Bi-C-Co-O; Bi-F-K-Rb; Bi-Nb-O-Pb; Bi-O-Pb-V; Br-C-O-Ru; C-Ca-Na-O; C-Cl-O-Ru; C-Hg-O-Ru; C-I-O-Os; C-I-O-Pt; C-I-O-Ru; C-O-Ru-S; C-O-Ru-Te; Ca-Cs-F-Ni; Ca-Cu-O-Ti; Ca-O- Si-Ti; Cl-Cs-K-O; Cr-O-P-Pb; Cs-F-Fe-Na; Cu-Mo-O-Sb; F-Fe-K-Rb; F-Mo-Nb-O; F-Mo-O-Sb; I-Li-O-Rb; Ir-K-Li-O; K-Mo-O-Sc; Li-Mo-O-Sc; Mo-O-Sr-Zn; O-Pb-V-Zn |

In addition we solved the task stated as: "Evaluate" Q4 using results learned from Q3. The recognition furnished the 99.6844 % accuracy result (Table 10). This accuracy differs by the value of 0.32 % from accuracy published in (Villars et al. (2000)) for similar problem. However Villars and co-authors used twice as large as our learning set.

Table 10.

Results of Examination for Learning Set Q3 and Testing Set Q4

| Learning Set | | Testing Set | | Violations | | Accura- cy, % |
|--------------|----------------|--------------|----------------|---|------------|------------------|
| For mers | Nonfor mers | For- mers | Nonfor mers | Formers | Nonformers | |
| Q3 | | Q4 | | AlCdOSi; BaNbOTi; BaOPtTi; CaMoOSr; CuNiSSi; GePbSTl; NbOTiU; OPbSiZn | | 99.6844 |
| 1814 | 3149 | 934 | 1602 | | | |

Conclusions

The results, obtained in the present investigation, testify to high efficiency of our computer learning programs. The worst result (the predicting accuracy is equal 66.5934 %) was obtained in the case of the learning set T4 and the testing set T3 (Table 11). However it should be noted that the set T3 was derived from T1 and T2 sets, but judged to be 80 % reliable. It is worth noting also that the use of regularity, obtained in the analysis of the learning set T4, has furnished the pinpoint accuracy of the testing recognition of set T5 - 98.7941 %.

The accuracies of the paper (Villars et al. (2000)) were calculated relatively to definite answers. To make comparison of approaches more correct we recalculated some numbers twice: relatively to definite answers and relatively to all answers.

Task 1. Training set: B3, T1+T2, Q1+Q2, test set B4. Accuracy 95.6% relatively to definite answers and 89.11% relatively to all answers.

Task 2. Training set: B1+B2, T3, Q1+Q2, test set T5. Accuracy 99.57% relatively to definite answers and 97.62% relatively to all answers.

Task 3. Training set: B1+B2, T1+T2, Q3, test set Q4. Accuracy 99.96% relatively to definite answers and 99.32% relatively to all answers.

Table 11.

The Best Results of Examination Obtained with the Help of Our Programs

| Training set | Examination set | Transformation | Accuracy relatively to all answers, % |
|--------------|-----------------|---|---------------------------------------|
| B3 | B4 | $E_{iA}-E_{iB}$ (i=1-12); $S_{iA}-S_{iB}$ (i=1-4) | 97.8852 |
| T4 | T3 | No | 66.5934 |
| T4 | T5 | No | 98.7941 |
| B3 | Q1 | No | 71.8864 |
| T4 | Q1 | No | 90.7138 |
| T4 | B3+B4 | No | 91.0617 |
| B3 | Q2 | No | 100 |
| T4 | Q2 | No | 100 |
| Q3 | Q4 | No | 99.6844 |

We calculated our results relatively to all answers. These results are in a good agreement with the accuracy of Villars' results but our colleagues have used the learning set which was one order more. It should be pointed out that one of the promising line of predicting accuracy rise is **an use of some algebraic and logic functions of different element properties as attributes**. It was confirmed by the rise of set B4 testing accuracy (97.8852 %) with the use of property value differences: $E_{iA} - E_{iB}$ (i = 1 - 12) and $S_{iA} - S_{iB}$ (i = 1 - 3) instead of the appropriate properties of elements A and B. In addition the classifying possibilities of attribute description rise sharply with **the use of more simple compound properties** (for example, simple oxides, chalcogenides, halogenides, etc.) (Kiselyova (1993 (a, b), 1995, 1997, 2000 (a, b)), Kiselyova and Savitskii (1977)). This is because the properties "of a "pure" element determine to a lesser degree it as partner for chemical interaction than the properties, for example, of its binary compound whose physical-chemical parameters are determined by bonding strength of this element with another element. Unfortunately the customers have offered to use only alone property of more simple systems - possibility of formation of compound with any composition.

More worse results than those of our colleagues, were obtained with learning sets B3 (71.8864 %) and T4 (90.7138 %) and testing set Q1 containing the data on quaternary systems only with formation of compounds (Table 11) (it should be borne in mind that Villars and his co-authors have used the learning sets, in some times exceeding the our learning files). However the use of regularity, obtained in the analysis of sets B3 and T4, for testing recognition of set Q2, containing the information on quaternary systems without formation of compound, has yielded 100 % accuracy of a prediction.

As already noted, the purpose of algorithm testing was choosing the best method for computer design of inorganic materials. In (Kiselyova (2000 (a))) we offered the criteria for validity of computer learning algorithms for the solution of chemical problems:

- a possibility of large data bulks analysis;
- a possibility of detection of qualitative classifying regularities for small learning sets;
- an automatic rejection of attributes (or attribute intervals) which are unimportant for a classification;
- a possibility of operation in conditions of weak fulfilling of hypothesis of compactness;
- a fast learning;
- a fast predicting;
- a possibility of operation with missing values of attribute;
- a possibility of operation with qualitative attributes;
- a high accuracy of predicting in solution of chemical problem.

It must be emphasized that our analytical complex **satisfied the all these criteria**.

All learning sets, offered for testing, contained more than thousand and even some thousand examples. Such a situation is extremely rare in computer design of inorganic compounds. Our experience shows that even some hundreds of training examples are rather exception than a rule in the solution of the majority of inorganic materials science problems. As a rule in the case of large learning sets even the most primitive computer learning algorithms can yield good results. Therefore it would be expedient to carry out repeated testing of offered algorithms for practical sets containing small number (1-2 hundred) of learning examples. It will provide a choosing of multi-purpose algorithms capable to

yield the good predictions independently of the learning set sizes.

It is worthy of note that the use of thousands training examples in the present testing has allowed to neglect by possible errors in sets to be analyzed and in properties of elements. For example, it was completely not clearly the charge of ions corresponding to given ionic radius by Yagoda. It is impossible to understand why it was necessary to double the parameter G1 (Group number) by parameter G2 (Valence electron number) taking into account that the information on a maximal valence contained a set of errors (He^{+8} for atom with two electrons is one of such errors). Doubtless that computer learning program is capable to process very large data bulks however **the task of the expert on certain subject field is statement of problem, selection of objects for computer learning and property for description of these objects.**

The use of systems, which are different in a chemical bond, both for learning and for recognition makes incompetent a search for individual properties of elements exhibiting the most promise for classification of systems on formers and nonformers. It is apparent that the formation or non-formation of compounds at standard conditions in systems of different types is determined by different sets of element properties. For example, it is generally believed that the proximity of the atom sizes promotes formation of solid solutions. This criterion perfectly works, e.g., for the binary system of Zr and Hf. However in order to predict a compound formation in a binary mixture, formed by elements with atomic radii having close values but the first element (metalloid) has boiling point below than the melting point of the second element (metal), it is necessary to take into account, as a minimum, some properties of elements: melting and boiling points, size factors, potentials of ionization, electronegativities. Only **system approach** to forming of attribute set for description of systems, in which an investigator uses the physical-chemical conceptions of a nature of objects to be analyzed, can yield a reliable prediction.

The large bulks of learning sets in the present testing have smoothed over a problem of quality of data to be analyzed. In most cases computer learning algorithms pick out a basic tendency in an arrangement of classes of objects in multi-dimensional attribute space, rejecting some objects which do not fit in the classification schema (unfortunately sometimes unique systems, which could become ancestors of new material families, can be rejected also). The error of prediction is proportional to the relation of number of "erroneous" objects to total of objects of a learning set. Thus, data estimation by the qualified experts is of great importance at small sizes of files to be analyzed. In this case the most fruitful way of an improvement in quality of prediction is **the use of databases containing the information evaluated by the experts** (Kiselyova (2000 (b))).

One of a possibility of predicting quality improvement can be **the comparison of results of the predictions with the use of different computer learning methods**. The various methods provide results, which supplement one another. It is similar to the comparison of results of Q1 set examination on the basis of learning sets B3 and T4 in the present investigation. Similar approach was advanced in (Kiselyova and Savitskii (1977)) however in this and subsequent our investigations we compared results of the prediction obtained with use of different attribute descriptions. The decision on object belonging to one or another class can be made on the basis of voting of the predictions. Such approach allows reducing an influence of features of different algorithms but in specific cases the procedure of voting is associated with the increase of number of the indeterminate predictions.

Another promising approach to prediction of new materials is to replace procedures of classification with procedures of compound objects design. Such sort of procedures is available in the *ANALOGY* program that was used in this investigation.

Thus, it is expedient to carry out new testing of algorithms with real learning sets containing few hundreds of objects belonging to several classes. In this case it is should be remembered that statement of problem and preparation of an initial information (learning sets and properties for description of objects) are of major importance. The last problem can be solved on the basis of

enlisting of expert-chemists of high qualification. This testing will allow choosing the algorithms or classes of algorithms most promising for computer design of inorganic materials.

KEY TERM LIST:

Artificial intelligence - is an artificial system, usually constructed on base of the use of computer technology, which simulates a human solution of the complicated tasks. It is intended for perception, processing and storage of information, and also for forming solutions of problems in an expedient behavior.

Attribute - is a property of the constituent component of the physical-chemical system.

Class - is a set of objects chosen according to any property (properties).

Computer learning (method of artificial intelligence) - is a process of a modification of parameters of a classifying system on base of the use of experimental data with the purpose of improving quality of a classification.

Former - is a physical-chemical system with the formation of a compound(s).

Learning set - is a multidimensional array of feature values and a column vector of the desired property. Each row corresponds to some physical-chemical system already known, whose class is indicated by row position of the column vector.

Conformer - is a physical-chemical system without formation of a compound(s).

Object - is a physical-chemical system, which is described as a set of property (feature) values of the constituent elements.

Physical-chemical system - is a system (e.g., compound or solid solution), which is formed from chemical elements.

Prediction - is an identification (classification) of belonging of new object to the certain class in compliance with the fixed classification scheme.

REFERENCES

Gladun, V.P., (2000) "Partnership with Computer". Port-Royal. Kiev. 118 p. (russ.).

Gladun, V.P., Velichko, V.Yu., Kiselyova, N.N., Moskal'kova, N.M., (2000), "Formation of hypotheses about composition and properties of objects by the inference based on analogy" *Artificial Intelligence*, 1, 44 (russ.).

(a) Kiselyova, N.N., (1993) "Information-predicting system for the design of new materials" *J. Alloys and Compounds*, 197, 159.

(b) Kiselyova, N.N., (1993) "Prediction of inorganic compounds: experiences and perspectives" *MRS Bull.*, 28, 40.

Kiselyova, N.N. (principal investigator), (1995) "The Design of Inorganic Compounds: Searching for New Electro-optical, Ferro-electric, Superconducting, and Semiconducting Materials (final report)". EOARD SPC-94-4097. 191 p.

Kiselyova, N.N., (1997) "Application of artificial intelligence methods to inorganic compounds design" *Perspektivnyye Materialy*, 4, 5 (russ.).

(a) Kiselyova, N.N. (principal investigator), (2000) "Search for Efficient Methods of Prediction of Multicomponent Inorganic Compounds ((preliminary report)". EOARD SPC-00-4014. 29 p.

(b) Kiselyova, N.N. (2000) "Databases and semantic networks for the inorganic materials computer design" *Engineering Application of Artificial Intelligence*, 3, 533.

Kiselyova, N.N., Savitskii, E.M., (1977) "Prediction of inorganic compounds with composition ABO_2 using computer learning" *Dokladi Akad.Nauk SSSR*, 235, 1367 (russ.).

Savitskii, E.M., Devingtal, Yu.V., and Gribulya, V.B., (1968) "About recognition of binary phase diagrams of intermetallic systems using computer" *Dokladi Akad. Nauk SSSR*, 178, 79 (russ.).

Velichko, V.Yu., Moskal'kova, N.M., (1999) "The solution of property prediction problem for the structural attributive models using the inference based on analogy" *Artificial Intelligence*, 2, 378 (ukr.).

Villars, P., Brandenburg, K., Berndt, M., et al., (2000) "Recent success of the discovery of the Mendeleev number and its powerful application for binary + ternary + quaternary compound former/nonformer prediction", the 13th International Conference on Solid Compounds of Transition Elements - SCTE2000, Stresa, Italy 4-7 April 2000.