

ON THE APPLICATION OF CYBERNETIC PREDICTION SYSTEMS IN THE SEARCH FOR NEW MAGNETIC MATERIALS

E. M. SAVITSKII, V. B. GRIBULYA and N. N. KISELYOVA

A. A. Baikov Institute of Metallurgy, Academy of Sciences of the U.S.S.R., Moscow (U.S.S.R.)

(Received November 20, 1979)

Summary

Cybernetic computer-learning methods are proposed for predicting the existence of intermetallic compounds with a Laves-phase or a CaCu_5 -type crystal lattice and of ferromagnetics with a spinel structure. In this context computer learning is a process of analysing experimental evidence on the presence or absence of a property of interest in various physicochemical systems. As a result of machine learning, a model is produced of characteristic exhibition of a property (*e.g.* the formation of a particular type of chemical compound); the model depends on the properties of the system components. The advantage of this approach is that knowledge of the properties of the components is quite sufficient for cybernetic prediction. The accuracy of prediction is not less than 90% for compounds with a Laves-phase or the CaCu_5 structure and not less than 75% for compounds with a spinel structure.

Of promising magnetic materials, much attention has been given recently to intermetallic compounds with a Laves-phase or a CaCu_5 -type crystal lattice, oxide ferromagnetics and binary chalcogenides of the CdCr_2S_4 type with a spinel structure. We propose to apply cybernetic machine-learning methods to predict the existence of these and other compounds [1, 2].

1. Statement of the problem and methods of solution

We were presented with the task of predicting a particular property Y_i of physicochemical systems from the properties of their components. We characterize each system by a set of values of its component properties. For example, in terms of the electronic composition of the separate atoms, this description will be

Description	System
$1s^2 2s^2 2p^6 3s^2 3p^0 \dots 7s^0 1s^2 2s^2 2p^6 3s^2 3p^0 \dots 7s^0$	Mg-Al
$1s^2 2s^2 2p^6 3s^1 3p^0 \dots 7s^0 1s^2 2s^2 2p^6 3s^2 3p^5 \dots 7s^0$	Na-Cl
$1s^2 2s^1 2p^0 3s^0 3p^0 \dots 7s^0 1s^2 2s^2 2p^5 3s^0 3p^0 \dots 7s^0$	Li-F
$1s^2 2s^2 2p^1 3s^0 3p^0 \dots 7s^0 1s^2 2s^2 2p^3 3s^0 3p^0 \dots 7s^0$	B-N
.	.
.	.
.	.
$1s^2 2s^2 2p^0 3s^0 3p^0 \dots 7s^0 1s^2 2s^2 2p^6 3s^2 3p^4 \dots 7s^0$	Be-S

(1) (2)

Thus we obtain an information matrix X_i (1) describing a complete set of physicochemical systems (2). For certain systems the value of a property Y_i is known from experiment. From the matrix X_i we sample the lines (of the physicochemical system) for which Y_i is known and incorporate them into a matrix X_{i0} . The remaining lines of matrix (1) are incorporated into a matrix X_{iP} , which can be sampled for prediction. The X_{i0} matrix and its corresponding vector, the Y_i column, constitute the learning sample. The dependence

$$Y_i = F(X_{i0}) \quad (3)$$

connecting Y_i with the description of the system X_{i0} must be constructed by analysing the learning sample.

The problem of retrieving dependences from experimental data is very important in applied mathematics. It is necessary to introduce a number of restrictions on the type of the F function, as well as on the input and output data, in order to find the dependence (3). The various methods of searching for dependences can be classified according to the character of the restrictions introduced. For instance the set of restrictions for regression-retrieving methods involves the assumption that the observed values of the output properties $Y_{i1}, Y_{i2}, \dots, Y_{iN}$ are independent random magnitudes which assume real values [3]. In certain cases (regression analysis) it is necessary to assume a normal distribution of the observed values, equality of the $\sigma^2(Y_{ij})$ dispersions (where $j = 1, 2, \dots, N$) and determinacy of the independent variables. In addition, restrictions on the type of the F function, which must be represented as a polynomial, are introduced.

The set of restrictions on the pattern recognition methods includes the assumption of compactness of the object classes which restricts the observed results for $Y_{i1}, Y_{i2}, \dots, Y_{iN}$ to, in general, only two values. Depending on the particular algorithm, restrictions are introduced on the type of multi-dimensional statistical distributions of observed results in each class of objects, the covariance matrix values, the discreteness of the properties of the components etc.

The mathematical methods of retrieving the functions are to search the existing regularities on breaking a set of facts applied in quantum mechanics or thermodynamics, which can apparently also be classified according to the set of restrictions used.

If we consider a property of physicochemical systems of the same type, we can apply the principle of retention of information complexity. In this case the principle can be stated as follows: in the consideration of a definite property of systems of the same level of complexity, the removal of one restriction from a set of non-trivial and non-equivalent restrictions assumed in the retrieving of the dependence (3) results either in the introduction of new restrictions or in complication of the calculation involved in the search for dependences (3). This is in line with the principle of complexity retention discussed in ref. 4. It should be noted that the information complexity principle is only approximate, mainly because no clear definition of complexity is available, *i.e.* we cannot derive a common measure of it for different systems.

The natural consequence of the principle is that any attempt to construct the dependence (3) of a certain property Y_i of a particular class of physicochemical systems on the properties of chemical elements (*e.g.* the problem of calculating the properties of a solid from "first principles") results either in the assumption of a set of restrictions on the type of the F function, or in involved calculation, or in the introduction of restrictions on the output data for Y_i (*e.g.* in pattern recognition Y_i may assume only two values). Between these two extreme cases numerous methods exist in which moderate restrictions both on the type of the F function and on the character of the input and output data are assumed. The principal aim of the research chemist is to select the restrictions which will best suit a given problem and which will not result in involved calculation.

The methods of pattern recognition and concept formation which we use to retrieve dependence (3) in the learning sample introduce rather "rigid" restrictions on the character of the output data and the type of the F function. The Y_i property may assume only two values and the F function must be represented either as certain linear algebraic expressions or as a logical expression. It only describes the boundaries of classes, almost without reference to their internal structure. However, such a set of restrictions reduces to a minimum the requirements on the input data (properties of chemical elements or simple compounds are used) and results in a relatively simple calculation. In addition, the application of computer learning reduces the requirements on the compactness of classes in the space of properties of chemical elements; this is important when using physicochemical properties of elements which are rather uninformative with regard to separation into classes.

To solve the problems of constructing the dependence (3) and of prediction in chemistry and metallurgy, we use various machine-learning algorithms for pattern recognition and concept formation. Algorithms are very widely used [5, 6]. In these algorithms the application of piecewise linear

separation functions and the construction of generalized informative features are common; in addition no assumptions with regard to the type of distribution and the independence of observation results for the output property are needed. Our experience indicates that algorithms of this type are effective in the solution of prediction problems in chemistry and metallurgy.

2. The prediction of Laves phases and CaCu_5 -type phases

Substances with definite properties and a definite crystal structure have long been the object of attention of physical chemists as a basis for creating new materials. Extensive research has therefore been carried out in the field of synthesizing such substances, in particular metallic compounds exhibiting magnetic properties or compounds with Laves-phase and CaCu_5 -type structures. A detailed discussion of these substances together with rare earth metals with regard to their magnetic properties is given in ref. 7 and in subsequent reviews [8].

Cybernetic prediction makes the field of search for such chemical compounds appreciably narrower, giving a general idea of their frequency of occurrence.

In this case of machine learning the chemical elements were described by data on the electronic construction of the separate atoms A and B either forming or not forming the predicted phase of composition close to the stoichiometry A_2B or A_5B . The total number of learning examples in different training alternatives was 60 - 200. Less than half of these (30 - 60) were positive (*i.e.* examples of the possibility of phases found experimentally).

In a prediction in 1972 [1, 2] the total number of possible binary Laves phases (both equilibrium and non-equilibrium) was estimated to be about 2000; according to a 1974 prediction [1], the total number of CaCu_5 -type phases is about 1100. The reserve of binary compounds may be appreciably increased by the application of a greater number of components. Figure 1 shows some of the predicted Laves phases with rare earth elements. It can be seen that most of these substances have now been synthesized. Certain Laves phases of rare earths unknown to us have been synthesized under pressure (Fe_2Pr , Fe_2Nd , Fe_2Yb [9]). The prediction shows that Laves phases of rare earths with gold, silver, palladium and mercury, as well as with alkali metals, are possible although most of them have not been synthesized. McMasters *et al.* [10] believe that in systems with gold the electron concentration is unfavourable for the formation of the predicted structure. Nevertheless, there are no grounds for ruling out the possibility that polymorphism of the structures which they have found may result in the formation of Laves phases.

An example of the possible realization of Laves phases which have not been found in the corresponding binary systems, but which have been predicted, may be silicon-stabilized phases [11], *i.e.* Ni_2Zr , Nb_2Zn , Ta_2Zr , Cr_2W etc.

Figure 2 shows an example of the prediction of phases with the CaCu_5 structure. Data on the electronic structure of the separate A and B atoms

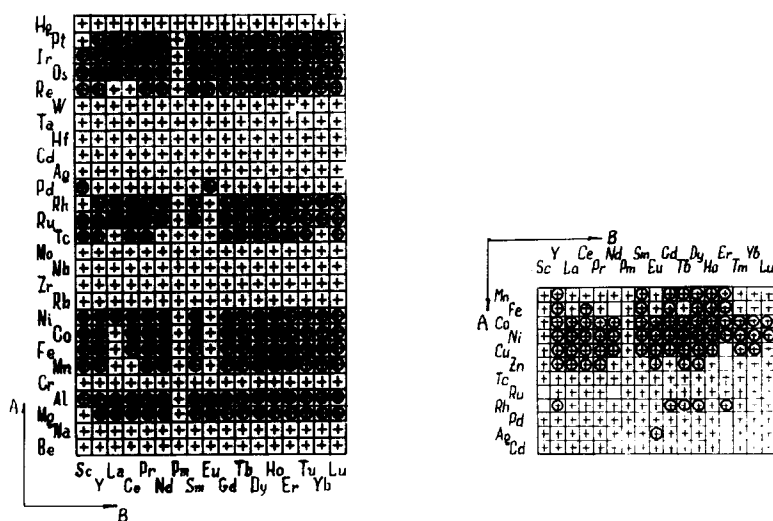


Fig. 1. The prediction of binary Laves phases. The symbol + indicates the prediction of a Laves phase. The symbol \oplus corresponds to known chemical compounds which could have been used for training the computer.

Fig. 2. The prediction of binary CaCu_5 -type phases. The symbol + indicates the prediction of a CaCu_5 -type phase. The symbol \oplus corresponds to known chemical compounds which could have been used for training the computer.

were used to describe the compounds. The prediction indicates that in principle the existence of such compounds in binary metallic systems is possible. The actual phase diagram obtained under particular experimental conditions need not reflect the data obtained by prediction. Therefore synthesis technology and stabilization of the required structure are of principal importance in actual conditions. As is evident from Fig. 2, most of the synthesized compounds of rare earths are with metals of groups VII and VIII of the 4th period. The prediction shows that nearly all d elements should form compounds of the CaCu_5 type with rare earth metals. The accuracy of the prediction is 99%.

3. The prediction of spinels of the composition AB_2X_4

The spinel structure is favourable for exhibiting magnetism. We predicted the possible formation of compounds with an AB_2X_4 composition in physicochemical systems the components of which include any of the elements A, B and X ($\text{X} \equiv \text{O}, \text{S}, \text{Se}, \text{Te}$). In addition, a spinel crystal structure was predicted for these compounds.

The properties of elements A and B and of their simple oxides (selenides, sulphides, tellurides), as well as reference data on the existence of AB_2X_4 compounds, were used as initial data for the prediction. The computer-

programming method of concept learning [6] was used to retrieve dependences for the possible formation of the compounds (or the spinel structure). The essence of machine learning consists in analysis of numerous experimental data in order to retrieve empirical dependences. A compound was considered to be unavailable if there were experimental indications that a heterogeneous mixture or a solid solution corresponded to the AB_2X_4 composition. If polymorphism was present, the compound was assumed to have the crystal structure which is realized under normal conditions.

As a result of machine learning, dependences were obtained in the form of logical expressions, or models whose variables were quantized values of the properties of the components. The characteristics found were used further to predict unknown compounds and the spinel crystal structure. All the calculations were carried out with a BESM-6 computer.

In the prediction system used [6] we succeeded to a large extent in solving the problems of initial data redundancy, of the large number of scans and of the ignored values of the matrix (1). This was achieved by automatically separating from the learning data with which the computer was loaded only those physicochemical systems (patterns) which carried new information on the class of systems being analysed, by using a hierarchic data structure (*i.e.* growing pyramidal networks) in the computer memory and by reasonably combining principles of mathematical logic and heuristic programming.

The accuracy of prediction was estimated by examining the results of the prediction, as in Section 2. In the examination the programmed computer was loaded with data on studied physicochemical systems describing the values of properties of their components which had not been used for training. The ratio of the number of correctly recognized patterns to the total number of sampling patterns in the examination is a measure of the accuracy of the prediction. The next stage involved further training of the computer using the learning sample material, the characteristics obtained being used for subsequent prediction of unknown compounds.

Starting from physicochemical models describing ternary systems we selected the following systems of component properties:

(1) the distribution of electrons in the energy levels of separate atoms, the formal valencies of elements in compounds (from electroneutrality considerations) and, for prediction of the type of crystal structure, the covalent or ionic (for oxide systems) radii of elements (cations in oxides);

(2) the first four ionization potentials, the standard isobaric thermal capacities, the types of incomplete electronic shells, the covalent or ionic (for oxides) radii of the elements (cations) and the formal valency of the element in the compound;

(3) the covalent (or ionic) radii, the enthalpies of the formation of simple oxides (sulphides, selenides, tellurides), the entropies of simple sulphides (selenides, tellurides) under standard conditions or the thermal capacities of simple oxides and the formal valencies of elements (in oxidic compounds).

A fourth system of properties was used in the prediction of the availability and properties of complex chalcogenides:

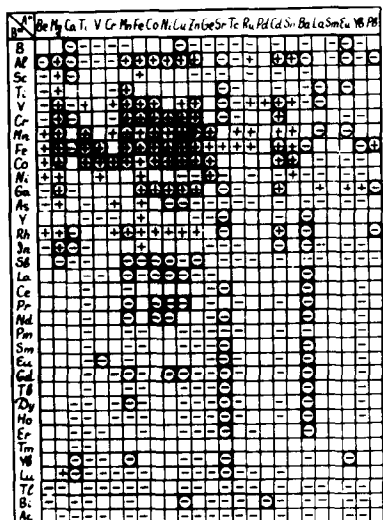


Fig. 3. The prediction of the spinel structure in compounds of the AB_2O_4 composition. The symbols \oplus and \ominus correspond to known chemical compounds which could have been used for training the computer.

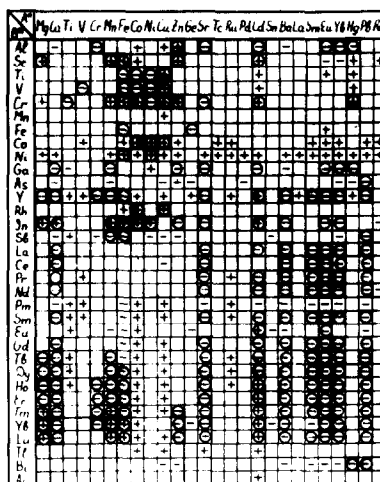


Fig. 4. The prediction of the spinel structure in compounds of the AB_2S_4 composition. The symbols \oplus and \ominus correspond to known chemical compounds which could have been used for training the computer.

(4) the types of incomplete electronic shells, the electronegativities, the covalent radii, the formal valencies exhibited in an AB_2X_4 compound and the enthalpies of formation of corresponding simple sulphides (selenides, tellurides) under standard conditions.

The characteristics obtained, involving the values of each of the four systems of properties, were used for the prediction. The predictions obtained for three or four systems of properties were compared and a decision as to whether the system belonged to a particular class was made unless the results were contradictory. Figures 3 - 5 show the results of such total predictions for spinels of the AB_2X_4 composition. We considered compounds in which the A and B elements formally have valencies of two and three respectively. Compounds for which the spinel structure is predicted are indicated by + and those for which any other structure is predicted are indicated by -.

An examination prediction showed that the accuracy of predicting the possibility of formation of the compounds is not less than 80% and the accuracy of predicting the occurrence of the spinel structure is not less than 75%. This is our second attempt to predict the existence and structure of binary oxides. We obtained similar predictions in 1974 [12]. In 5 years over 30 systems in the domain of the AB_2O_4 composition have been investigated [13, 14]. Only three of these systems have failed to confirm our predictions. Nearly 20 compounds have been studied crystallographically. Only for two compounds (including one used in computer learning) have our data failed to coincide with experiment. Results of predicting spinels after further

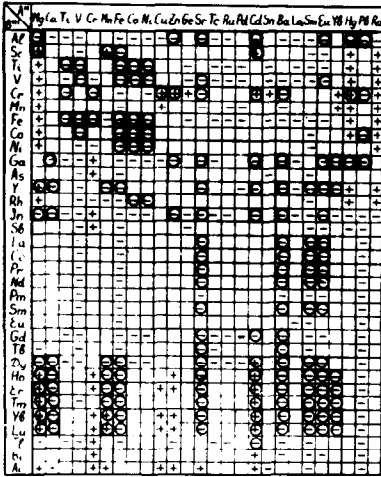


Fig. 5. The prediction of the spinel structure in compounds of the AB_2Se_4 composition. The symbols \oplus and \ominus correspond to known chemical compounds which could have been used for training the computer.

computer training with new experimental data are listed in Fig. 3. In 1978 we predicted the formation of binary tellurides of the AB_2Te_4 composition [15]. Of seven systems in the domain of the AB_2Te_4 composition studied since then [16, 17], for only one has the prediction failed to coincide with the experimental results.

Considering Figs. 3 - 5, the spinel crystal structure is most characteristic of oxide compounds of the AB_2O_4 composition (approximately 45% of the known compounds of this composition are spinels). The frequency of occurrence of the spinel structure in AB_2X_4 compounds drops sharply as the covalent character of the bonds increases in the transition to complex sulphides, selenides and tellurides. Only about 20% of AB_2S_4 compounds, 13% of AB_2Se_4 compounds and fewer than 1% of AB_2Te_4 compounds have a spinel crystal structure. Unlike complex oxides, structures such as those of $CaFe_2O_4$, Th_3P_4 , Cr_3Se_4 and Ag_2HgI_4 [18, 19] become competitive for these compounds. Empirical attempts to determine factors which would allow the domains of existence of these structures to be separated with sufficient accuracy have not been successful. For example, to determine the stability domains of crystal structures Iglesias and Steinfink [19] have proposed a plot of the dependence of the power constant $K = x_A x_B / r_e^2$ (where x_A and x_B are the electronegativities of the cations and r_e is the equilibrium distance between them) on the ratio r_A / r_B . In the diagram obtained the domains corresponding to the structures of olivine, K_2SO_4 , MnY_2S_4 , $CaFe_2O_4$ and Th_3P_4 and the monoclinic structure are separated relatively well; however, the stability domains of the spinel, Cr_3Se_4 and Ag_2HgI_4 structures overlap considerably. Thus the mere applica-

tion of geometrical factors and electronegativities for separating stability domains is clearly insufficient. We believe that the introduction of multi-dimensional separation functions that take into account all the properties of chemical elements and simple compounds is a more effective way of solving similar problems. As to the sampling of the most informative subspaces, it must be performed on a computer using a special program. This is the approach we apply when making use of learning prediction systems.

According to our predictions, the spinel structure is not realized in compounds of the $A^{VI}B^I_2X_4$ and $A^{IV}B^{II}_2X_4$ compositions. However, this is far from being the case with AB_2O_4 compounds. The spinels Ag_2MoO_4 , Na_2WO_4 , Mg_2SnO_4 , Zn_2PtO_4 etc. are known. The small number of spinels obtained and predicted for $A^{II}B^{III}_2X_4$ compounds is further proof that this type of crystal structure is not characteristic of ternary compounds of sulphur, selenium and tellurium.

The predictions obtained make it possible to develop a strategy for the search for new inorganic phases, including those possessing interesting magnetic properties. They also appreciably reduce the time needed to develop new magnetic materials.

References

- 1 Ye. M. Savitskii and V. B. Gribulya, *Computer-aided Prediction of Inorganic Compounds*, Nauka, Moscow, 1977 (in Russian).
- 2 Ye. M. Savitskii, V. B. Gribulya and N. N. Kiselyova, Applications of diagrams in metallurgy and ceramics, Vol. 2, *Natl Bur. Stand. (U.S.), Spec. Publ. 496*, 1978, p. 113.
- 3 V. N. Vapnik, *Retrieving of Dependences by Empirical Data*, Nauka, Moscow, 1979 (in Russian).
- 4 V. S. Tyukhtin, *The Theory of Automatic Identification and Gnoseology*, Nauka, Moscow, 1976 (in Russian).
- 5 Yu. V. Devingtal, *Izv. Akad. Nauk SSSR, Tekh. Kibern.*, 3 (1971) 139 (in Russian).
- 6 V. P. Gladun, *Heuristic Search in Complex Media*, Naukova Dumka, Kiev, 1977 (in Russian).
- 7 K. N. Taylor, *Adv. Phys.*, 20 (1971) 551.
- 8 K. J. Strnat (ed.), *Proc. 2nd Int. Workshop on Rare Earth-Cobalt Permanent Magnets and their Applications*, Dayton, Ohio, U.S.A., 1976.
- 9 J. F. Cannon, D. L. Robertson and H. T. Hall, *Mater. Res. Bull.*, 7 (1972) 5 - 11.
- 10 O. D. McMasters and K. A. Gschneidner, G. Brussone and A. Palenzona, *J. Less-Common Met.*, 25 (2) (1971) 135 - 160.
- 11 R. C. Mittal, S. K. Si and K. P. Gupta, *J. Less-Common Met.*, 60 (1978) 75 - 82.
- 12 N. N. Kiselyova, B. I. Pokrovskii, L. N. Komissarova and N. D. Vashchenko, *Zh. Neorg. Khim.*, 22 (1977) 883.
- 13 V. B. Lazarev and I. S. Shaplygin, *Izv. Akad. Nauk SSSR, Neorg. Mater.*, 14 (1978) 1942.
- 14 B. Latourrette, M. Devalette, F. Quillen and C. Fouassier, *Mater. Res. Bull.*, 13 (1978) 567.
- 15 Ye. M. Savitskii, N. N. Kiselyova and N. D. Vashchenko, *Dokl. Akad. Nauk SSSR*, 239 (1978) 1154.
- 16 O. M. Aliyev and P. G. Rustamov, *Zh. Neorg. Khim.*, 23 (1978) 2800.
- 17 P. G. Rustamov, J. I. Zulfugarly, M. G. Safarov and T. N. Guliyev, *Azerb. Khim. Zh.*, 4 (1978) 100.
- 18 J. Flahaut, *Prog. Sci. Technol. Rare Earths.*, 3 (1968) 209.
- 19 J. E. Iglesias and M. Steinfink, *J. Solid State Chem.* (1973) 119.