HYPOTHESIS OF SIMPLICITY IN PATTERN RECOGNITION

Zagoruyko N.G. & Samokhvalov K.F.

Laboratory of Pattern Recognition
Institute of Mathematics
Siberian Division
USSR Academy of Sciences
U.S.S.R.

ABSTRACT

This paner deals with the problem of constructing a general method for pattern recognition. It is proposed to realize this method based on a hypothe sis of simplicity which is formulated in an appropriate manner. The proposed method is illustrated with an example.

Introduction

In this paper the problem of pattern recognition is considered as that of determining a method for predicting future empirical results based on investigation of only preliminary information of an object and phenomena which are in previous experiments (in training samle). Solution of the problem stated, in principle, depends on the acceptance of a natural-scientific hypothesis that already has been mentioned in the literature (1). We assume that such a natural-scientific hypothesis, in appropriate way, must explain the historical fact that mathematically simpler natural-scientific theories, as a rule, are more preferable as methods of predicting future experimental facts. The concern of this paper is how one can use the conception of simplicity formulated only on intuitive level for the solution of pattern recognition problems.

§1. Training sample in given feature space

Let us define more exactly the notions"feature"and "feature space". Any feature is determined wholly by a choice of empirical procedure for its measurement. A measurement procedure is a comparison of the measuring object with some set X of standards. The comparison is determined by a finite set of empirical operations $F= \{f_1, f\}$ and empirical relations $P= \{p_1 \dots , p_a\}$ on a set of standards. Relative to these operations and relations the measuring object must behave in the same way as one of the standards. Let us introduce the following definition: the feature $\{x\}$ is a class of all algebraic systems isomorphic to that $X=\langle X; F; P\rangle$ in which domain X is a set of standards, F and P are collection of empirical operations and relations on this set, correpondingly.

For example, the feature "weight" is a class of all algebraic systems isomorphic to that $\langle W; (+) ; \quad \rangle$, where W is an accounting set of standards (weights), operation © is that of obtaining any standard by putting together on a pan two different standards, relation is "not heavier than" (3).

Since an abstract class of algebraic system is quite characterized by any of its representatives Including a system in which the domain is a set of numbers, operations and relations - suitable numerical operations and relations, then one can describe, for instance, the feature "weight" as a class of all algebraic systems isomorphic to a numerical one $\langle N; + ; \leq\rangle$, where N is a set of numbers, operation $+$ is an arithmetical operation of natural numbers, addition, and the relation $\leqq$ is "not greater than". For all this, we must not for/ret, however, that it is possible to consider the system $\langle N; +; \leqq \rangle$ as representation of the feature "weight" if and only if this system is isomorphic to the empirical one $\langle W; \oplus; \prec \rangle$

Tf it is known of some object £ only that one can perform measurements of its feature , then it is evident that this object does not differ from some standard $X_£$ which is an element of domain X of the corresponding algebraic system.

Let us choose this standard $x_£$. For a system isomorphic to that with the chosen element, $\mathcal{X}_z = \langle X; x_z, F, P \rangle$ is equivalent to writing: "the object £ takes value $x_£$ for feature $\{\mathcal{X}\}$". Given, for instance, that the object £ weighs 2 kg, one can write as an algebraic system isomorphic to that $W_z = \langle W, 2kg, \odot; \prec \rangle$, where the object 2 kg belongs to the set W of weight standards.

Now, let us assume we are given a finite set of features a , . . \ x) \ x \ corresponding to empirical nlgebraic systems

$$\mathcal{X}_1 = \langle X_1, F_1, P_1 \rangle$$
$$\mathcal{X}_i = \langle X_i, F_i, P_i \rangle$$
$$\mathcal{X}_n = \langle X_n, F_n, P_n \rangle$$

'e shall name a class of all algebraic systems isomorphic to an n-basic one (2)

$$\mathcal{X}^{(n)} = \langle X_1, F_1, P_1, \quad, X_n, P_n \rangle$$

for an n-dimensional feature space $\{\mathcal{X}^{(n)}\}$.

Here, as in the one-limensional case, instead of an object with known values of n features, one can consider a class of all systems isomorphic to

$$\mathcal{X}_z^{(n)} = \langle X_1, x_{1z}, F_1, P_1; \quad, X_n, x_{nz}, F_n, P_n \rangle$$

The realizations set $\{z_q\}$ of a training sample of power M generates the same power set of abstract classes of algebraic systems $\{\{\mathcal{X}_{zq}^{(n)}\}\}$, $q = 1, \ldots, M$. It is called a training sample for K patterns if the set $\{\{\mathcal{X}_{zq}^{(n)}\}\}$ is broken up into K classes of equivalence for the given feature space $\{\mathcal{X}^{(n)}\}$.

§2, Scheme of general theory of recognition for the given feature space

^he problem of recognition in general can be formulated thus: the training sample represents the following R classes of equivalence: *)

The 1-st class of equivalence:

$$\begin{cases} \mathcal{X}_{z_1}^{(n)} = \langle X_1, x_{1z_1}, F_1, P_1, \quad, X_n, x_{nz_1}, F_n; P_n \rangle \\ \vdots \\ \mathcal{X}_{z_k}^{(n)} = \langle X_1, x_{1z_k}; F_1, P_1, X_n; x_{nz_k}, F_n, P_n \rangle \end{cases}$$

$\vdots$

The k-th class of equivalence:

$$\begin{cases} \mathcal{X}_{z_q}^{(n)} = \langle X_1, x_{1z_q}, F_1, P_1, \quad, X_n, x_{nz_q}, F_n; P_n \rangle \\ \ldots \\ \mathcal{X}_{z_M}^{(n)} = \langle X_1; x_{1z_M}, F_1; P_1, \ldots, X_n, x_{nz_M}, F_n, P_n \rangle \end{cases}$$

Recognizing object y is given by the system

$$\mathcal{X}_y = \langle X_1, x_{1y}, F_1, P_1; \ldots, X_n; x_{ny}, F_n, P_n \rangle.$$

Tt is required, proceeding from the properties of algebraic systems of training sample and control realization *if* , to predict which of K classes of equivalence (patterns) system Xy belongs bo. As has been stated earlier, exclusively logical methods for this problem solution are not sufficient ones,

Tt is necessary to introduce an additional assumption which we postulate as the following many-stage process,

*)Here and further, instead of an abstract class of algebraic systems we mean any concrete representative of the class

considering it as an implicit formulation of the hypothesis of simplicity.

At the first stnge it is possible to consider the class of systems of equal complexity corresponding to minimum permissible level of complexity which is determined as stated below. If it is not successful in obtaining a solution, then class of systems having the next higher level of complexity is considered, and so on until the first solution is obtained. The choice of solution on each stage consists in the following: at first, systems which are homomorphous images of a system X_y are chosen from systems class of relevant complexity. Let us denote the collection of such systems by $HomX_Y$.

Realization y belongs to the i-th class of equivalence (pattern) if two conditions are fulfilled:

1).There is a system Q_i in the collection $HomX_y$ which is a homomorphous image even if for only one system from the i-th class of equivalence and there is none for any system of another class of equivalence.

2).Among the members remaining of $HomN_y$ there is no system $0;$ which with respect to the J-class of equivalence behaves the same way as Q_1 - to the i-th one (i≠j).

Now, we only need to break up the set Q of all non-isomorphic finite systems of a given signature $\langle \mathfrak{X}_1, F_1, P_1, \ldots; \mathfrak{X}_n, F_n; P_n \rangle$ into levels by complexity in order to provide a description of this process. Let us compare to each finite n-basic system Q a sequence of numbers

$$\langle \bar{\bar{X}}_1^{(Q)}, 0, \ell_{11}, \ \ , \ell_{1s_1}, m_{11}, \ \ , m_{1z_1}; \ \ \bar{\bar{X}}_n^{(Q)}, 0, \ell_{n_1}, \ldots, \ell_{ns_n}; m_{n_1}, \ \ , m_{nz_n} \rangle$$

in which every member of kind $X_i^{(Q)}$ is the power of domain $X_i^{(Q)}$, every member of kind m_{ij} is the number of places of relation $\rho_j \in P_i$ and every member of kind ℓ_{ij} is the number of places of operation $f_j \in F_i$.

Breaking up a set $\{Q\}$ into levels by its complexity is accomplished thus: for any system Q_1, Q_2 from $\{Q\}$, the system Q_1 is less than (or equal to) the system Q_2 by complexity if and only if the quantity

$$U_1(Q_1) = \bar{\bar{X}}_1^{(Q)} + (\bar{\bar{X}}_1^{(Q_1)})^{\ell_{11}+1} + \ldots + (\bar{\bar{X}}_1^{(Q_1)})^{\ell_{1s_1}+1} + (\bar{\bar{X}}_1^{(Q_1)})^{m_{11}} + \ldots + (\bar{\bar{X}}^{(Q_1)})^{\bar{m}_{1z_1}} + \ldots + \bar{\bar{X}}_n^{(Q_1)} + (\bar{\bar{X}}_n^{(Q_1)})^{\ell_{n_1}+1} + \ldots + (\bar{\bar{X}}_n^{(Q_1)})^{\ell_{ns_n}+1} + (\bar{\bar{X}}_1^{(Q_1)})^{m_{n_1}+1} + (\bar{\bar{X}}_n^{(Q_1)})^{m_{nz_n}}$$

is less than (or equal to) the quantity

$$U_1(Q_2) = \bar{\bar{X}}_1^{(Q)} + (\bar{\bar{X}}_1^{(Q_2)})^{\ell_{11}+1} + \ldots + (\bar{\bar{X}}_1^{(Q_2)})^{\ell_{1s_1}+1} + (\bar{\bar{X}}_1^{(Q_2)})^{m_{11}} + \ldots + (\bar{\bar{X}}^{(Q_2)})^{\bar{m}_{1z_1}} + \ldots + \bar{\bar{X}}_n^{(Q_2)} + (\bar{\bar{X}}_n^{(Q_2)})^{\ell_{n_1}+1} + \ldots + (\bar{\bar{X}}_n^{(Q_2)})^{\ell_{ns_n}+1} + (\bar{\bar{X}}_n^{(Q_2)})^{m_{n_1}} + \ldots + (\bar{\bar{X}}_n^{(Q_2)})^{m_{nz_n}}$$

It is evident that the set $\{Q\}$ is well ordered with relation "less (or equal) by complexity"; therefore, it makes sense to use a many-stage process corresponding to sequential consideration of complexity levels of systems from $\{Q\}$, beginning with the minimum one.

Example.

Let $\{\mathfrak{X}\}$ be a one-dimensional feature space whose representative is a one basic numerical algebraic system, that is algebra $\mathfrak{X} = \langle \{1, 2, \ \} ; + \rangle$. Let us assume that a training sample for two patterns is represented by two one-element sets { 3} and { 2} . Now, we determine to which of these two patterns the number 5 belongs.

In the feature space the training sample is represented by two systems:

$$\mathfrak{X}_3 = \langle \{1, 2, 3, \ldots\} ; 3, + \rangle \quad \text{for } 3$$

$$\mathfrak{X}_2 = \langle \{1, 2, 3, \ldots\} ; 2, + \rangle \quad \text{for } 2$$

and recognizing element y = 5 by system

$$\mathfrak{X}_5 = \langle \{1, 2, 3, \ldots\} ; 5, + \rangle$$

These systems have only two onerations: the binary one of addition + and the zero-ary one of element choice. The minimum power of domain is equal to 1. Thus, the minimum level of complexity containing one system $Q_1 = \langle \{a\}; a, \varphi_1^{(2)} \rangle$ where $\varphi_1^{(2)}$ is the binary oneration on set $\{a\}: \varphi_1^{(2)}(a, a) = a$ is determined in the class.

The class *Horn* X_s also includes only one algebra Q_1 at this stage. It is evident that Q_1 is a homomorphous image of both $X.$ and X_2 systems, therefore, the first one of the two conditions mentioned above is not fulfilled and there is no solution at this stage.

The second stage corresponds to the next larger complexity level which is given if the domain consists of two elements a and b .

The complexity level at this stage consists of eight non-isomorphic systems

$$Q_1 = \langle \{a, b\}, \ a, f_1^{(2)} \rangle ,$$

$$Q_2 = \langle \{a, b\}; \ b, f_1^{(2)} \rangle; \dots, Q_7 = \langle \{a, b\}; a, f_4^{(2)} \rangle ,$$

$$Q_8 = \langle \{a, b\}, \ b, f_4^{(2)} \rangle .$$

Here, the binary operations of "addition" $f^{(2)}$ on set $\{a, b\}$ are given by the following table.

TABLE 1.

	$f_1^{(2)}$	$f_2^{(2)}$	$f_3^{(2)}$	$f_4^{(2)}$
$a+a$	a	a	a	a
$a+b$	a	a	b	b
$b+a$				
$b+b$	a	b	a	b

Which of these eight systems generate the class *Horn X$* ?

Using known theorems for generating elements, it is easily to prove the class $HomX_5 = \{Q_1, Q_6\}$. In the same way one can be convinced that Q_1 is a homomorphous image of both X_2 and X_s which contradicts condition stated above in the rule of decision acceptance. The system Q_6 is a homomorphous image only of system $X.$, which permits us to accept a one-to-one decision: 5 belongs to pattern {3}.

We have given this example to emphasize that this process can be used, in some cases, in a feature space corresponding to infinite algebraic systems. However, we can not guarantee its use for any feature space. Nevertheless, since an adequate theory of measurement must not refer to infinite algebraic systems from an empirical viewpoint, the stated above restriction on use of this process has no importance in principle.

Conclusions.

In our paper we have intended to show the possibility of constructing an universal method of pattern recognition based on the hypothesis of simplicity. Questions concerning the practical realization of this method have not been considered here because there, apparently, is not at present an elaborate enough theory of measurement which excludes references to infinity.

REFERENCES

1. 3aropyttKO H.T., CaMOXBajiOB K.¢.npwpofla npoojieMLi pacno3HQBaHMH oopaaoB. "BhRMCJIMTCJIBIlbie CWCTeMbl", HOBOCl/I-dwpcK, nun.36, 1969.
2. njiOTKMH E.H. Tpynnbi aBTOMOp<; MHMOB ajiroGpaimecKMX CKCTeM. "Hayim", M., 1966.
3. Cyimec IL, 3nHec J{%. OCIJOBM Teopww wnMepeHHii. CO. "flcwxojioriwecKwe M3MepeHMH",MOCKBQ,W3fl."Mnp", 1967.