APPEARANCE OF IMPLICATIVE REGULARITIES IN BOOLEAN CRITERION SPACE AND PATTERN RECOGNITION

A. D. Zakrevskii

Introduction

Traditional pattern recognition methods are ordinarily associated (without any explicit mention) with the hypothesis of the functional dependence of the criterion of a class on other criteria in the description of the objects being described, and with the hypothesis of the existence of a certain decision function for these criteria. Moreover, a specific type of decision function is also usually considered known, it is taken as an axiom that it assigns a dividing plane in the multidimensional space of criteria [1], is potential [2], or is something else. The selection of the kind of decision function often relies on some heuristic reasoning, but it generally remain logically without foundation.

Without a priori assumption of any hypothesis on the existence of a decision function, we shall use a more general terminology, a "decision rule." Another most important problem of recognition theory, the problem of estimating the confidence of predictions realized on the basis of the decision rule is closely associated with the problem of its selection. It is reasonable to differentiate predictions, and in certain situations it is more useful to reject them than to draw on them without sufficient foundations.

A new approach to the combined solution of both problems is noted within the framework of the more general problem of the appearance of regularities in the data flow, and of the problem of converting the latter into "knowledge," is some adequate model of the class of objects under investigation that operationally permit finding the solution of diverse problems of pattern recognition, classification, empirical prediction, filling in the blanks in experimental tables, etc. [3-6].

In this paper this approach is developed in application to the case of binary criteria, which permits efficient utilization of the apparatus of Boolean function theory [7-9], but allows extension also of the case of multivalued criteria. A method is proposed for the determination, by a training sample, of the general properties of a single class of "real" objects representable by appropriate points of the Boolean space M of all criteria (including even the targets). The method relies on a unique a priori hypothesis about the preference of regularities connecting minimal groups of criteria. The expediency of the appearance of sufficiently strong regularities of the type "elementary exclusions" that yield implicative relations between criteria and the construction of differentiated prediction procedures on their basis, which permit extrapolation of partially assigned properties of objects not in the training sample with a foundation will follow logically from this hypothesis.

Problem of Reconstructing Sets of Real Objects

Let us assume that these objects are described by sets of values of binary criteria \( x_1, x_2, \ldots, x_n \) so that each object is identified with a certain specific value of the Boolean vector \( x = (x_1, x_2, \ldots, x_n) \). All \( 2^n \) distinct values of the vector \( x \) form a Boolean space \( M \), however, only some of them, comprising the set \( M_r \), correspond to real objects. Let us allow a certain subset \( M_f \) from \( M_r \), called the training sample of the experimental data, to be known. It is convenient to represent this information in the form of a Boolean matrix by setting up its columns in conformity to criteria, and its rows to objects. By formulating the problem of reconstructing (restoring, finding) the set \( M_r \) by means of the given set \( M_f \) and solving it correctly, we can successfully solve many other derivative problems as well. The reconstruction problem is especially interesting in the case when the powers of the sets introduced in the consideration are quite distinctive: \( |M_f| \ll |M_r| \ll |M| \).

In principle, the solution of this problem is evidently possible only in that case when the structure of the set \( M_r \) is subject to regularities of sufficient strength so that they could be reflected in the set \( M_f \). We shall hence consider that the set \( M_f \) is a random sample from \( M_r \) with a uniform distribution law.

In the general case only an approximate solution of this problem could be computed by assuming that the
degree of approximation will depend on the parameter $m$, the volume of the experimental data ($m = \lvert M_f \rvert$). The
larger the elements the set $M_f$ contains, the "finer" the regularities to be implied from it.

Regularities of the "Elementary Exclusion" Type

The solution of the problem of sampling a mode or type of regularities controlling the structure of the
set $M_f$ relies on a certain hypothesis about the preference for some modes over others. Such a hypothesis, that
plays the part of an axiom in the subsequent logical foundation of the proposed method to reconstruct the set
$M_f$, should necessarily start from nonformal reasoning of empirical nature which is sufficiently reasonable
although not provable on a purely logical level. The volume of a priori information included in the hypothesis
is here reduced usefully to a minimum by expanding the domain of logical deduction in such a manner.

Any regularity can be considered as a certain connection or relation between criteria, and the fewer the
criteria to be enclosed in some specific connection, the more preferable will we consider it. This assertion is
indeed taken as an axiom.

The existence of a connection between criteria forming a certain subset $X_1$ from $X = \{x_1, x_2, \ldots, x_n\}$
assumes that there is an element possessing them in the set $M_f$ for not every set of their values. In other
words, there exists at least one exclusion in the value of the vector $x_1$ formed from appropriate components
of the vector $x$. Let us call such an exclusion elementary, and let us represent it by the elementary conjunction
$k_1$ that becomes one in the excluded value of the vector $x_1$. The rank of the conjunction, the number of letters
therein, will be considered to be simultaneously the rank of the elementary exclusion it represents. An ele-
mentary exclusion of rank $l$ generates a domain of exclusion of intensity $2^{n-l}$, the characteristic set of the
elementary conjunction $k_1$ or the corresponding interval of the space $M$. For example, the elementary conjunction
$k_1 = x_2x_5x_8$ gives an elementary exclusion of rank 3 to the value 101 of the vector $x_1 = (x_2, x_5, x_8)$, i.e., to
the set of values $x_2 = 1, x_5 = 0, x_8 = 1$. In the space of Boolean variables $x_1, x_2, \ldots, x_n$ the domain of this ex-
clusion is an interval of eight elements which it is convenient to express by the triple vector $-1-01$.

The mode of the regularity chosen turns out to be universal since any regularity is equivalent to the as-

Assignment of a certain domain of exclusion in the space $M$, and this latter can always be represented as the
union of certain intervals and any regularity is therefore in the form of an appropriate set of elementary ex-
cclusions.

The concepts of simplicity and force of a regularity, ordinarily introduced at an intuitive level, are
easily made specific for elementary exclusions. It is natural to estimate the simplicity of a regularity by the
dimension of the expression it describes, and the force by the volume of the exclusion domain. In this case,
this results in the fact that the simplicity of an elementary exclusion turns out to equal its rank $l$, while the
force is the quantity $2^{n-l}$, i.e., the intensity of the appropriate exclusion interval.

Elementary exclusions can be interpreted as implicative relations between criteria, where each ele-
mentary exclusion of rank $l$ dissociates into $l$ implications. For instance, the implication $x_2x_5 \Rightarrow x_6$ follows from
the elementary exclusion $x_2x_5x_8$ (if $x_2 = 1, x_5 = 0, x_8 = 0$), as do also the implications $x_2x_6 \Rightarrow x_5$ and $x_2x_6 \Rightarrow x_5$. In
particular, for $l = 2$ the elementary exclusions are equivalent to simple propositions considered in an Aristo-
telian syllogistic.

The implicative relations are a more general type of regularity as compared with functional relations,
which are just a particular case of the implicative. Taking account of implicative relations can underlie meth-
ods permitting the realization of the effective prediction even in the absence of functional relations between cri-
teria. Hence, the considerations become clear that underlie the proposed a priori hypothesis on the preferabil-
ity of regularities of a specific kind, low-rank elementary exclusions.

Confirmation of the Hypothesis of Elementary Exclusions

When the information contained in the set $M_f$ is available, hypotheses can be advanced on the existence of
different elementary exclusions in the set $M_f$. The foundation for such hypotheses will be the fact that the
exclusions under consideration are not spoiled in the set $M_f$, however this is insufficient. Let us pose the ques-
tion: What is the probability that at least one hypothesis on an elementary exclusion of rank $l$ can be advanced
on this foundation where there are no regularities in the set $M_f$? We shall consider that both $M_f$ and $M_f$
are random samples from $M$, subject to a uniform distribution law, in this case. If the probability mentioned turns
cut to be high, then the corresponding hypothesis should evidently not be taken.