Pattern Recognition with Mathematical Logic
and Discrete Analysis

Levon Aslanyan
Institute for Informatics and Automation Problems
Information Society Technologies Center
Joint Laboratory of Computational Biology
Institute for Information Theories and Applications
1, P.Sevak, 0014 Yerevan, lasl@sci.am

Pattern recognition undergoes an important development for many years. As a research area this is not uni-modular like classic mathematical sciences, there are a number of sub disciplines inside such as – feature selection, object and feature ranking, analogy measuring, supervised and unsupervised classification, etc. The same time pattern recognition is indeed an integrated theory studying object descriptions and their diverse classification models. This is a collection of mathematical, statistical, heuristic and inductive techniques of fundamental role in executing the intellectual tasks, typical for a human being, – but on computers.

Many applied problems with multidimensional experimental data use the object description that is often non-classical, that is, - not exclusively in terms of only numerical or only categorical features, but simultaneously by both kinds of values. Sometimes, the missing or unreliable value is introduced so that finally we deal with mixed and incomplete descriptions of objects as elements of Cartesian product of feature values, without any algebraic, logical or topological properties assumed in applied area. How then, do we select in these cases the most informative features, classify a new object given a (training) sample or find the relationships between all objects based on a certain measure of similarity? Logic Combinatorial Pattern Recognition (LCPR) is a research area formed since 70’s that uses a mix of discrete descriptors with similarities, separation, frequencies, integration, corrections, and optimization, and solves the whole spectrum of pattern recognition tasks.

This approach is originated by the work [Dmitriev et al, 1966] that transfers the engineering domain technique of tests for electrical schemes [Chegis, Yablonskii, 1958] to the feature selection and object classification area. The applied task of [Dmitriev et al, 1966] address prognosis of mineral resources.

Testor theory [Dmitriev et al, 1966] is based on a restriction of feature sets - with preservation of the basic learning set property. A feature subset $T = \{x_{i_1}, x_{i_2}, \ldots, x_{i_k}\}$ is called a testor (or test) of $L$, if projection of $L$ on $T$ keeps the classes nonintersecting (learning set property). Irreducible is the testor where no one $x_{i_j}$ may be eliminated with conservation of the learning set property. Further a feature ranking is made taking into consideration the frequency of belonging $x_{i_j}$ to the testors (irreducible testors). Testor based supervised classification algorithms are constructed by use of frequency similarity measures. Which is the structural property used from the learning
set? This is the pairwise difference of learning set elements from different classes (learning set property). We may suppose that this is a consequence of the well accepted compactness hypothesis. On formal basis testor technology is well visible on binary tables. Now it is known that constructing all irreducible testors is an NP hard problem. Approximations are studied as well. There is a high similarity with association rule mining models, especially in learning the monotone set structures – be it with frequent itemsets in associative rule mining or irreducible tests and testors in this theory.

**KORA and Logic Separation.** [Vaintsvaig, 1973] introduced one of the basic concept in LCPR – the KORA algorithm. KORA is constructing elementary conjunctions $C_i$ that intersect with only one class $C_i$ of learning set $L$. It is easy to imagine the corresponding interval $I_C$ of $n$-dimensional binary cube $E^n$. $I_C$ does not contain contradictory knowledge of learning the set $L$. Contrary, [Aslanyan, 1975] considers all irreducible conjunction forms that intersect with only one class $C_i$ of the learning set $L$. This is not the generating idea of this work but is the consequence of the Logic Separation (LS) Principle. The work, factually for the first time, is considering learning set elements in a non-disjoint manner. The potential function concept is used that an element spreads its similarity, reducing by the distance measure, which interrupts facing the different class object. An extension may consider not only pairs of learning set elements – one spreading the similarity measure and the second interrupting that - but also arbitrary subsets/fragments of learning set. Several comments: logically, it is evident that the best learning algorithm is also best suited to the learning set (at least the learning set is reconstructable by the information used by the algorithm). This also may use the recognition hypothesis when available. The same learning set fragments play a crucial role in determining the best suited recognition algorithm to the given learning data set.

After the methodological discussion it is worth to mention further developments with the voting algorithm, algorithmic correction procedure and other approaches of advanced logic-combinatorial pattern recognition [Zhuravlev, 1998]

References


