

# USING RELATIVE LOGIC FOR PATTERN RECOGNITION

Juliusz L. Kulikowski

*Institute of Biocybernetics and Biomedical Engineering PAS, Warsaw, Poland, e-mail jlkulik@ibib.waw.pl*

**Abstract:** It is presented an approach to pattern recognition in the case of extremely poor primary information about the properties of patterns that are to be recognized. In such case algorithms based on reference sets of objects are widely used. However, in some situations experts cannot assign the names of patterns to the reference objects in a unique way. It is shown here that in such cases an approach based on topological logic can be used for relative logical evaluation of statements formulated by the experts. There is described a method of relative reference sets construction and of using them in an algorithm of pattern recognition admitting relative logical evaluation of statements concerning the recognized patterns.

**Key words:** pattern recognition, relative objects' similarity assessment, topological logic

## 1. INTRODUCTION

Pattern recognition is, generally speaking, an art of giving reasonable answers to the questions like: "*What does the given sign (image, object, string of data, etc.) represent?*" assuming that the answer should indicate the name of a class of objects the observed one belongs to and that there is at least one alternative class it might belong to as well. The *pattern* is thus a synonym of a *class of similar objects*. The principles that are used as a basis of solution of a pattern recognition problem depend on primary assumptions and information about the objects under observation, the patterns, and the criteria of pattern recognition quality. We call the set of such primary assumptions and information a *pattern recognition model*. In literature

pattern recognition models based on deterministic, probabilistic and/or statistical assumptions have been described [1]. Another class of models form those based on potential functions [2], artificial neural networks [3], “*k Nearest Neighbors*” (*k-NN*) approach and other learning concepts [4]. The aim of this paper is presentation of a pattern recognition approach useful in situation of extremely poor primary information concerning the recognized classes of objects. The approach is based on a concept of *topological logic*, originally formulated by C.G. Hempel [5].

## 2. PATTERN RECOGNITION METHODS UNDER PRIMARY INFORMATION DEFICIENCY

Let us start with consideration of the *k-NN* pattern recognition method as a one having opinion of being based on very wide primary assumptions. In fact, it can be used if the below-given conditions are satisfied:

a/ there is given an observation space  $X$  whose elements are some real  $n$ -component vectors  $\xi$ ;

b/  $X$  is a metric space i.e. there is defined a distance function:

$$d: X \times X \rightarrow R^+ \quad (1)$$

where  $\times$  stands for a Cartesian product,  $R^+$  denotes a non-negative real half-axis, and  $d$  satisfies the reciprocity, symmetry and triangle-inequality conditions;

c/ the number  $K$  of recognized patterns is fixed, finite and a priori known;

d/ each pattern is a random vector  $\xi_\kappa$ ,  $\kappa = 1, 2, \dots, K$ , whose conditional probability distribution is a priori unknown; however, 1° its mean value and variance are finite, 2° the mean values of the random vectors are significantly different.

In addition, it is assumed that there is given a family of *reference sets*  $S_1, S_2, \dots, S_K \subset X$  such that:

e/ they are finite and mutually disjoint;

f/ each reference set corresponds to a given pattern;

g/ the elements (vectors  $\xi$ ) of any reference set are observed and correctly classified values of a random vector corresponding to the given pattern;

h/ the cardinal numbers (numbers of elements)  $|S_\kappa|$ ,  $\kappa = 1, 2, \dots, K$ , of the reference sets can be increased so that the sets become representative for the corresponding statistical populations.

The classical *k-NN* method is thus based on strong assumptions which in practice are not obviously satisfied. In particular:

$A/X$  may be not a metric space;

$B/$  assignment of the reference set elements  $\xi$  to the patterns may be not univocally given;

$C/$  a similarity between the new-observed objects  $x$  and those of the reference set  $S$  may be evaluated with a limited accuracy.

Therefore, it arises the problem: how in the above-mentioned situation the pattern recognition problem could be solved. A solution based on a generalized similarity concept will be proposed below.

### 3. THE NOTION OF RELATIVE SIMILARITY

The notion of *similarity* in pattern recognition is a crucial one. For a given object  $x$  its pattern is recognized if a class  $\xi_x$  of objects similar to the given one is indicated. However, an interpretation of the notion of *similarity* in classical sets [6], *fuzzy sets* [7], *rough sets* [8] theory etc. is different. In the  $k$ -*NN* method the distance measure  $d$  is used as a decreasing function of objects' similarity. In [9] an extension of the similarity concept and a method of pattern recognition method, called "*k Most Similar Objects*" (*k-MSO*) has been proposed.

A new situation arises when stating: "Budapest reminds me of Vienna more than Prague". In this case no numerical similarity measure is used; two pairs of objects: [Budapest, Vienna] and [Budapest, Prague] are taken into account and their similarities are compared. This leads to a below described concept of *relative similarity*. For this purpose let us take into account a non-empty set  $A$  and a Cartesian product  $F = A \times A$ . A similarity of the elements of  $A$  can be defined as a sort of ordering of the elements of  $F$ . Any two elements of  $F$  in the sense of similarity can be:  $a/$  equivalent,  $b/$  one may be dominating with respect to the other one or  $c/$  mutually incomparable (this is not admitted if similarity is evaluated by a numerical similarity measure). Then the model is specified by the below-given definitions:

For any  $a, b, c, d, e, f \in A$  and for any their unordered pairs:

Definition 1.

A *relative similarity equivalence* between the pairs of elements of a set  $A$  is a relation  $\approx_{sim}$  (read: „are similar equivalently to”) described on a Cartesian product  $F \times F$  (i.e. on the pairs of pairs of the elements of  $A$ ) satisfying the following conditions:

- 1/  $\{a, a\} \approx_{sim} \{a, a\}$  (reciprocity),
- 2/ if  $\{a, b\} \approx_{sim} \{c, d\}$  then  $\{c, d\} \approx_{sim} \{a, b\}$  (strong symmetry),
- 3/ if  $\{a, b\} \approx_{sim} \{c, d\}$  and  $\{c, d\} \approx_{sim} \{e, f\}$  then  $\{a, b\} \approx_{sim} \{e, f\}$  (transitivity),

4/  $\{a,a\} \approx_{sim} \{b,b\}$  (equivalence of similarities of identical elements).

Any subset  $\Phi_k \subseteq F$  consisting of all pairs of elements satisfying the relation  $\approx_{sim}$  will be called a *class of relative similarity*.

Definition 2.

A *relative similarity between the pairs of elements of A* is a relation  $\leq_{sim}$  (read: „are similar at most like”) described on a Cartesian product  $F \times F$  and satisfying the following conditions:

- 1/  $\{a,a\} \leq_{sim} \{a,a\}$  (reciprocity),
- 2/  $\{a,b\} \leq_{sim} \{c,d\}$  and  $\{c,d\} \leq_{sim} \{a,b\}$  if and only if  $\{a,b\} \approx_{sim} \{c,d\}$  (weak symmetry),
- 3/ if  $\{a,b\} \leq_{sim} \{c,d\}$  and  $\{c,d\} \leq_{sim} \{e,f\}$  then  $\{a,b\} \leq_{sim} \{e,f\}$  (transitivity),
- 4/ there are no  $\{a,a\}$  and  $\{c,d\}$  such that  $\{a,a\} \leq_{sim} \{c,d\}$  and not  $\{a,a\} \approx_{sim} \{c,d\}$  (maximality of self-similarity).

Definition 3.

A *relative incomparability of the pairs of elements of A* is a relation  $\neq_{sim}$  (read: “are incomparable in similarities to”) described on a Cartesian product  $F \times F$  such that

$\{a,b\} \neq_{sim} \{c,d\}$  if and only if it is neither  $\{a,b\} \leq_{sim} \{c,d\}$  nor  $\{c,d\} \leq_{sim} \{a,b\}$ .

The above-defined relations can be illustrated by a directed graph  $G = [Q, L, \varphi]$ , where  $Q$  denotes a set of nodes,  $L$  is a set of arcs and  $\varphi$  is a function assigning arcs to ordered pairs of nodes. The graph, called a *relative similarity graph (RSG)*, should satisfy the conditions:

- a/ all its nodes in an unique way have been assigned to the subsets  $\Phi_k \subseteq F$  forming classes of relative similarity,
- b/ the function  $\varphi$  in an unique way assigns an arc  $\lambda_{kl}$  to an ordered pair  $[\Phi_k, \Phi_l]$  such that for each element  $\{a,b\} \in \Phi_k$  and each element  $\{c,d\} \in \Phi_l$  it is  $\{a,b\} \leq_{sim} \{c,d\}$  (but not the reverse),
- c/  $G$  is contourless.

It follows from the definition that to any pair  $\{\{a,b\}, \{e,f\}\}$  satisfying the condition  $\{a,b\} \leq_{sim} \{e,f\}$  it will be assigned in RSG a path  $\Lambda_{km}$  from  $\Phi_k$  to  $\Phi_m$  such that  $\{a,b\} \in \Phi_k$  and  $\{e,f\} \in \Phi_m$ . Any pair  $\{\{a,b\}, \{e,f\}\}$  such that in RSG it can not be represented in the above-described way is relatively incomparable. A RSG is thus a representation of relative similarities among the elements of A.

It also follows from the definition of *RSG* that  $Q$  contains a subset  $Q_{max}$  of *maximum nodes* such that no arc  $\lambda_{kl}$  starting from a node in  $Q_{max}$  connects it with another node outside  $Q_{max}$ .

### 4. RELATIVE REFERENCE SETS

The reference sets used in learning pattern recognition algorithms are usually collected under the recommendations of experts given in the form of statements:

$$\Theta_{\kappa}(\xi) : \text{“}\xi \text{ represents } E_{\kappa}\text{”}$$

which means that  $\xi$  being an observation (an object) from  $\Sigma$  should be included into  $S_{\kappa}$ , where  $S_{\kappa}$  is a reference set representing the  $\kappa$ -th pattern.

A problem arises when the experts are not convinced that the given objects can be assigned to reference sets in an unique way. The below-presented notion of *relative reference sets* based on *topological logic* [5] (see also [10], [11], [12]) is a solution of the problem.

The statements that should be logically assessed are:  $\Theta_1(\xi), \Theta_2(\xi), \dots, \Theta_{\kappa}(\xi), \dots, \Theta_K(\xi)$ . It is admitted that for a given  $\xi$  not only one of the statements is “true” and the logical values of the statements can not be evaluated but by relative comparisons. For this purpose let us assume that the expert has his intuitive concept of “typical” representatives  $\xi^*_1, \xi^*_2, \dots, \xi^*_{\kappa}, \dots, \xi^*_K$  of the corresponding patterns. Then, for any  $\xi^*_\kappa, \kappa = 1, 2, \dots, K$ , he takes into consideration a pair  $\{\xi^*_\kappa, \xi\}$ , and he establishes a semi-ordering relation among the pairs:

$\{\xi^*_\kappa, \xi_1\} \approx_{sim} \{\xi^*_\lambda, \xi_2\}$  meaning that the similarity between  $\xi^*_\kappa$  and  $\xi$  is equivalent to the one between  $\xi^*_\lambda$  and  $\xi$ , or

$\{\xi^*_\kappa, \xi\} \leq_{sim} \{\xi^*_\lambda, \xi\}$  meaning that the similarity between  $\xi^*_\kappa$  and  $\xi_1$  is at most like the one between  $\xi^*_\lambda$  and  $\xi$ , etc.

In practice, the relation can be established as follows.

1. For a given number  $K$  of patterns and for a given reference object  $\xi$  the expert is asked to assign to  $\xi$  a string of *rank position markers*:

$$n(\xi) = [v_1, v_2, \dots, v_{\kappa}, \dots, v_K] \tag{3}$$

where  $v_{\kappa} \in [0, 1, \dots, K-1]$ , denoting the length of the shortest path connecting  $\xi$  with  $\xi^*_\kappa$  in a *RSG*. For example, if  $K = 5$  and  $n(\xi) = [2, 1, 2, 3, 4]$ , then *RSG* will have the form shown in Fig. 1.

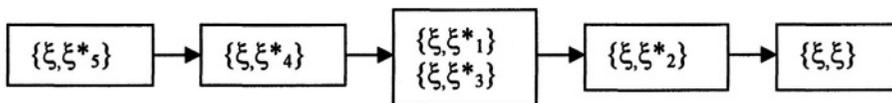


Fig.1. A graph  $G(\xi)$  representing relative similarities between objects.

$G(\xi)$  thus illustrates a ranking of patterns, as seen by the expert: the shorter is the path in  $G(\xi)$  between the nodes containing  $\{\xi, \xi^*_{\kappa}\}$  and  $\{\xi, \xi\}$ , the more representative is  $\xi$  as a member of  $S_{\kappa}$ .

Similar relationships should be established for all other members of  $\Sigma$ .

2. For the given  $\Sigma$  and a  $\kappa \in [1, 2, \dots, K]$  there will be denoted by  $\Phi^{\kappa}_0, \Phi^{\kappa}_1, \dots, \Phi^{\kappa}_{K-1}$  the subsets of  $\Sigma$  consisting of reference objects  $\xi$  such that the distance in the *RSG* between  $\{\xi, \xi^*_{\kappa}\}$  and  $\{\xi, \xi\}$  is equal, correspondingly, 0, 1, ...,  $K-1$ . So defined subsets form a partition of  $\Sigma$ :

$$\Sigma = \bigcup_{\lambda=0}^{K-1} \Phi^{\kappa}_{\lambda} \quad (4)$$

For each  $\kappa$  the partition is, in general, different. The subsets  $\Phi^{\kappa}_0, \Phi^{\kappa}_1, \dots, \Phi^{\kappa}_{K-1}$  form a semi-ordered family of sets that can be also represented by a graph.

3. The above-described operation of the reference set  $\Sigma$  partition into a semi-ordered family of subsets  $\Phi^{\kappa}_{\lambda}$  should be performed for all  $\kappa \in [1, 2, \dots, K]$ . This leads to the following definition:

Definition 4.

Let  $\Sigma$  be a finite subset of the elements of observation space  $X$  in which the relations  $\approx_{sim}$  and  $\leq_{sim}$  have been defined. Then for any fixed  $\kappa \in [1, 2, \dots, K]$  a partition of  $\Sigma$  into a semi-ordered family of subsets  $\{\Phi^{\kappa}_{\lambda}\}$ ,  $\lambda \in [0, 1, \dots, K-1]$ , induced by the relations  $\approx_{sim}$  and  $\leq_{sim}$  will be called a *relative reference set (RRS)* corresponding to the pattern  $\xi_{\kappa}$ .

A membership of a reference object  $\xi$  in the *RRS* is thus characterised by its relative similarity to the given patterns, as indicated by the expert.

## 5. PATTERN RECOGNITION

Let us assume that for the given observation space  $X$  it is given a family of *RRSs* corresponding to  $K$  patterns. Let it be also given a new object  $x \in X$ ,  $x \notin \Sigma$ . It arises the problem of assigning  $x$  to a pattern which means: indication of a pattern  $\xi_{\kappa}$  whose members are, in the sense of relative similarity, the most similar ones to  $x$ . The problem can be solved in two steps:

1. Finding out a subset of  $k$  reference objects, the most similar ones to  $x$ ;
2. Evaluation of relative similarity of  $x$  to the patterns through the relative similarity of reference objects the most similar ones to  $x$ .

The problem consists in logical evaluation of the pairs of statements:

$\Theta_{\kappa}(x)$ : " $x$  is similar to  $\xi_p$ "  $\wedge$  " $\xi_p$  is similar to  $\xi^*_{\kappa}$ ",

$\Theta_{\lambda}(x)$ : " $x$  is similar to  $\xi_q$ "  $\wedge$  " $\xi_q$  is similar to  $\xi^*_{\lambda}$ ",

where  $\xi_p, \xi_q \in \Sigma$  and  $\kappa, \lambda \in [1, 2, \dots, K]$ . This can be done using the general principles of topological logic ([5], [10], [11]), stating, in particular, that if  $A$  and  $B$  are some statements then

1°  $A \leq_{\log} B$  (read: " $A$  is logically valuable at most as  $B$ ") then  
 $(A \wedge B) \leq_{\log} A$  and  $(A \wedge B) \leq_{\log} B$ ;

2°  $A \approx_{\log} B$  (read: " $A$  is logically as valuable as  $B$ ") then  
 $(A \wedge B) \approx_{\log} A$  and  $(A \wedge B) \approx_{\log} B$ ;

3°  $A \leq_{\log} B$  then  $A \leq_{\log} (A \vee B)$  and  $B \leq_{\log} (A \vee B)$ ;

4°  $A \approx_{\log} B$  then  $A \approx_{\log} (A \vee B)$  and  $B \approx_{\log} (A \vee B)$ .

We are thus able to compare logically the pairs of statements like:

" $x$  is similar to  $\xi_p$ ", " $\xi_p$  is similar to  $\xi^*_{\kappa}$ ",

etc. for various  $x$ ,  $\xi_p, \xi^*_{\kappa}$ , as well as composite statements like:

" $x$  is similar to  $\xi_p$ "  $\wedge$  " $\xi_p$  is similar to  $\xi^*_{\kappa}$ ".

This leads, finally, to the following pattern recognition rule:

Algorithm:

For the given reference set  $\Sigma$ , set  $R$  of scaling objects and an object  $x \in X$ :

Using the scaling objects  $R$  find in  $\Sigma$  a subset  $\xi_1, \xi_2, \dots, \xi_k$  of  $k$  reference objects the most similar to  $x$ ;

For each  $\xi_p, p = 1, 2, \dots, k$ , if  $\xi_p \in \Phi^{\kappa}_{\lambda}$ , formulate the statement:

$\Theta_{\kappa}(x)$ : " $x$  is similar to  $\xi_p$ "  $\wedge$  " $\xi_p$  is similar to  $\xi^*_{\kappa}$ ",

Evaluate relatively the so obtained statements  $\Theta_{\kappa}(x)$  according to the logical rules 1°, 2°;

Select the statements of the highest logical value: if it is  $\Theta_{\kappa}(x)$  then  $\xi_{\kappa}$  is the recognized pattern, if there are more than one statement of the highest value, than all the corresponding patterns are logically mutually equivalent.

## 6. CONCLUSIONS

The methods of relative logical assessment of statements based on topological logic can be used in pattern recognition in the case of extremely poor primary information about the patterns being to be recognised. In such case relative similarities of observed objects can be used instead of exact numerical similarity scales.

## ACKNOWLEDGEMENTS

I would like to express my deep thanks to the Reviewers of this paper for their valuable comments and critical remarks.

## REFERENCES

1. G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition* (John Wiley and Sons, Inc., New York, 1992).
2. E.M. Braverman, *Experiments on Machine Learning of Recognition of Visual Pattern* (in Russian, *Avtomatika i telemekhanika*, No 3, 1962).
3. C.M. Bishop, *Neural Networks in Pattern Recognition* (Clarendon Press, Oxford, 1997).
4. E. Fix, J.L. Hodges, *Discriminatory Analysis: Nonparametric Discrimination Small Sample Performance* (Project 21-49-004, Rep. No 11, USAF School on Aviation Medicine, Randolph Field, Texas, 1952, pp. 280-322, ).
5. C.G. Hempel, *A purely topological form of non-Aristotelian logic* (*Journ. Symb. Logic*, 2(3), 1937).
6. H. Rasiowa, R. Sikorski, *The Mathematics of Meta-Mathematics* (PWN, Warsaw, 1968).
7. L.A. Zadeh, *Fuzzy Sets*. (*Inf. and Control*, vol. 8, 1965, pp. 338-353,).
8. Z. Pawlak, *Rough Sets. Theoretical Aspects of Reasoning About Data*” (Dordrecht, Kluwer Academic Publishers, 1991).
9. J.L. Kulikowski, *Pattern Recognition Based on Ambiguous Indications of Experts* (in: *Mat. Konferencji n.t. „Komputerowe Systemy Rozpoznawania KOSYR’2001”*, Wyd. Politechniki Wroclawskiej, Wroclaw, 2001, pp. 15-22).
10. Ch. A. Vessel, *“On the topological logic* (in Russian, in: *“Non-classical Logic”*, Nauka, Moscow, 1970, pp. 238-261).
11. J.L. Kulikowski, *Decision Making in a Modified Version of Topological Logic* (in: *Proc. of the Seminar on „Non-conventional Problems of Optimization”*, Part 1, Prace IBS PAN, No 134, Warszawa, 1984, pp. 24-44).
12. J.L. Kulikowski, *Topological logic and rough sets. A comparison of methods* (in Polish, in: *„Zbiory rozmyte i ich zastosowania*, Wyd. Politechniki Slaskiej, Gliwice, 2001, pp. 195-216).