APPLICATION OF BINARY COMPOSITIONS IN MULTIDIMENSIONAL RECOGNITION TASKS
BINĀRĀS KOMPOZĪCIJAS IZMANTOŠANA DAUDZDIMENSIJU ATPAZĪŠANAS UZDEVUMOS

Romans Grekovs, Decision Support Systems Group, Institute of Information Technology,
Riga Technical University, 1 Kalku Street, Riga LV−1658, Latvia

Abstract. *This paper examines the possibility of using pattern recognition method, which is based on compositions of fuzzy relations, to work with multiple feature selections such as where the number of features is greater than 3. Investigation of problems, which are connected to multidimensional pattern recognition, is also applied. Some practical part concerning proposals on experiments is provided.*

Keywords. *Pattern recognition, compositions of fuzzy relations, n−dimensional description space.*

1. Introduction

A pretty big number of papers present a concept of binary compositions and their application for fuzzy pattern recognition. Most of them are dedicated to the problems, solutions and improvement of that method. But less practical attention has been made to create a conception of using of binary compositions for recognition in the multidimensional description space. This paper actually is a try of research in this branch and gives some ideas of usage of the following method for a wide range of tasks.

2. Using of Binary Compositions in Multidimensional Feature Searching Tasks

All the major descriptions of using binary compositions in fuzzy pattern recognition are presented in [1], [2] and [3]. The main idea is creation of shadows of objects into each subspace of the description space. Then the compositions of shadows of fuzzy sets get made. The results of each composition are the degrees of reduction δ, which characterise the extent of increase in power of starting sets while obtaining reduction sets. So for every feature in a task space the following degrees of reductions have been created (this example represents the reductions made between X_1X_3 and X_3X_2 2−dimensional hyperspaces):

$$S^A_{X_1X_3} \circ S^B_{X_3X_2} \to \delta^A_{X_1X_3}, \delta^B_{X_3X_2}$$

$$S^B_{X_1X_3} \circ S^A_{X_3X_2} \to \delta^B_{X_1X_3}, \delta^A_{X_3X_2}$$

where S^A_{XZ} is a representation of shadow of fuzzy set;
$\circ$ – operation of composition;
$\delta^A_{X_1X_3}$ – degree of reduction.

Then, by summarising this values using the following expression one can get evaluation of parameter X_2:

$$\delta_{X_3} = (\delta^A_{X_1X_3} \wedge \delta^B_{X_3X_2}) \vee (\delta^B_{X_1X_3} \wedge \delta^A_{X_3X_2}).$$

This is how it works for evaluation of just one feature – X_2. For 3−dimensional tasks it's pretty simple to understand the physical meaning of the shadows and all the steps of recognition look logical and don't require any serious proof. But in case if we deal with n−dimensional description task where n is greater than 0 the situation will become more and more complicated depending on how big value of n is present.

ISBN 9984 - 585 - 36 - 0

First of all let us imagine the conceptions of n–dimensional space reduction to sub–dimensions. So the question is: How many sub–dimensional spaces one can get from n–dimensional space? In case of 3–dimensional space three 2–dimensional subspaces (shadows) will be created on each plane. If it is 2–dimensional space we will get two 1–dimensional subspaces as shadows. So it seems to be very logical to assume that in case of 4–dimensional space one can get four 3–dimensional shadows and so on. Thus, for n–dimensional space n $(n-1)$–dimensional subspaces will be created. But we get another problem here: If for n–dimensional space we have n $(n-1)$–dimensional subspaces (i.e. shadows), but binary compositions work only on 2–dimensional subspaces, then how to get the finite number of 2–dimensional subspaces from any n–dimensional space? There is only one visible solution for situation like this: Using projections every time move from high dimension spaces to low dimension spaces while finite number of 2–dimensional shadows gets found. This task seems to be pretty complicated but no other choice exists.

To describe the method let us use 4–dimensional description space X^4, in which fuzzy sets A and B are set through membership functions f_A and f_B, respectively. The space of descriptions is represented by Cartesian product $X_1\times X_2\times X_3\times X_4$, where X_1, X_2 X_3 and X_4 are parameters of objects' descriptions.

1. First, four 3–dimensional shadows get constructed:

$$S^A_{X_1X_2X_3},\ S^B_{X_1X_2X_3};\ S^A_{X_1X_2X_4},\ S^B_{X_1X_2X_4};\ S^A_{X_1X_3X_4},\ S^B_{X_1X_3X_4};\ S^A_{X_2X_3X_4},\ S^B_{X_2X_3X_4}.$$

2. Second, for each 3–dimensional shadow three 2–dimensional shadows get created:

$$\begin{array}{lll}
S^A_{X_1X_2X_3} & \to & S^A_{X_1X_2X_3\to X_1X_2},\ S^A_{X_1X_2X_3\to X_1X_3},\ S^A_{X_1X_2X_3\to X_2X_3};\\
S^B_{X_1X_2X_3} & \to & S^B_{X_1X_2X_3\to X_1X_2},\ S^B_{X_1X_2X_3\to X_1X_3},\ S^B_{X_1X_2X_3\to X_2X_3};\\
S^A_{X_1X_2X_4} & \to & S^A_{X_1X_2X_4\to X_1X_2},\ S^A_{X_1X_2X_4\to X_1X_4},\ S^A_{X_1X_2X_4\to X_2X_4};\\
S^B_{X_1X_2X_4} & \to & S^B_{X_1X_2X_4\to X_1X_2},\ S^B_{X_1X_2X_4\to X_1X_4},\ S^B_{X_1X_2X_4\to X_2X_4};\\
S^A_{X_1X_3X_4} & \to & S^A_{X_1X_3X_4\to X_1X_3},\ S^A_{X_1X_3X_4\to X_1X_4},\ S^A_{X_1X_3X_4\to X_3X_4};\\
S^B_{X_1X_3X_4} & \to & S^B_{X_1X_3X_4\to X_1X_3},\ S^B_{X_1X_3X_4\to X_1X_4},\ S^B_{X_1X_3X_4\to X_3X_4};\\
S^A_{X_2X_3X_4} & \to & S^A_{X_2X_3X_4\to X_2X_3},\ S^A_{X_2X_3X_4\to X_2X_4},\ S^A_{X_2X_3X_4\to X_3X_4};\\
S^B_{X_2X_3X_4} & \to & S^B_{X_2X_3X_4\to X_2X_3},\ S^B_{X_2X_3X_4\to X_2X_4},\ S^B_{X_2X_3X_4\to X_3X_4};
\end{array}$$

So, from 4–dimensional description space 24 2–dimensional shadows get created (12 for class A, 12 for class B). In case of 5–dimensional space one can get ten 4–dimensional subspaces (5 for class A, 5 for class B) and each of them represents four 3–dimensional subspaces, and finally each of 3–dimensinal subspaces represents three 2–dimensional shadows needed. So for 5–feature task 10*4*3*2 = 240 binary shadows get constructed (120 for class A, 120 for class B). This gives us the common formula of number of shadows constructed from n–dimensional space:

$$N=2*\prod_{i=0}^{n-2}(n-i),$$

where n is the dimensionality of description set.

Using such approach gives pretty complicated task to solve and the main question appears: How to combine such a big number of 2–dimensional subspaces for feature selection or recognition? Of

ISBN 9984 - 585 - 36 - 0

course the next step is a choosing of 2–dimensional shadows to perform binary compositions. Suggestion is to combine shadows in a sequence, like they appear. It means if (say, in case of 4–dimensional description space) we have four 3–dimensional subspaces, then for each of the 3–dimensional subspaces it looks logical to process the corresponding 2–dimensional shadows and get separate reduction degrees and then combine them to get summarised evaluation of all corresponding degrees of reduction and then get the final evaluation of the feature. For example, for subspaces $S^A_{X_1X_2X_3}$ and $S^B_{X_1X_2X_3}$ shadows can be combined in the following way:

$$S^A_{X_1X_2X_3\to X_1X_2} \circ S^B_{X_1X_2X_3\to X_1X_3} ;\ S^B_{X_1X_2X_3\to X_1X_2} \circ S^A_{X_1X_2X_3\to X_1X_3} ; \quad \to \quad \delta^{X_1X_2X_3}_{X_1}$$

$$S^A_{X_1X_2X_3\to X_1X_2} \circ S^B_{X_1X_2X_3\to X_2X_3} ;\ S^B_{X_1X_2X_3\to X_1X_2} \circ S^A_{X_1X_2X_3\to X_2X_3} ; \quad \to \quad \delta^{X_1X_2X_3}_{X_2}$$

$$S^A_{X_1X_2X_3\to X_1X_3} \circ S^B_{X_1X_2X_3\to X_2X_3} ;\ S^B_{X_1X_2X_3\to X_1X_3} \circ S^A_{X_1X_2X_3\to X_2X_3} ; \quad \to \quad \delta^{X_1X_2X_3}_{X_3}.$$

So, after all compositions have been made the following degrees of reduction are created:

$\delta^{X_1X_2X_3}_{X_1}$, $\delta^{X_1X_2X_3}_{X_2}$, $\delta^{X_1X_2X_3}_{X_3}$, $\delta^{X_1X_2X_4}_{X_1}$, $\delta^{X_1X_2X_4}_{X_2}$, $\delta^{X_1X_2X_4}_{X_4}$, $\delta^{X_1X_3X_4}_{X_1}$, $\delta^{X_1X_3X_4}_{X_3}$, $\delta^{X_1X_3X_4}_{X_4}$, $\delta^{X_2X_3X_4}_{X_2}$, $\delta^{X_2X_3X_4}_{X_3}$, $\delta^{X_2X_3X_4}_{X_4}$

Now, combining the values of these degrees of reductions it is possible to get evaluation of each feature. There can be a lot of possibilities how to combine the values. Let us take that a simple sum by feature is meaningful. In this case the following evaluations are derived:

$$\delta^{X_1X_2X_3}_{X_1} + \delta^{X_1X_2X_4}_{X_1} + \delta^{X_1X_3X_4}_{X_1} = \delta_{X_1};$$

$$\delta^{X_1X_2X_3}_{X_2} + \delta^{X_1X_2X_4}_{X_2} + \delta^{X_2X_3X_4}_{X_2} = \delta_{X_2};$$

$$\delta^{X_1X_2X_3}_{X_3} + \delta^{X_1X_3X_4}_{X_3} + \delta^{X_2X_3X_4}_{X_3} = \delta_{X_3};$$

$$\delta^{X_1X_2X_4}_{X_4} + \delta^{X_1X_3X_4}_{X_4} + \delta^{X_2X_3X_4}_{X_4} = \delta_{X_4}.$$

3. Conclusions

Of course, the method of application of binary relations to fuzzy recognition tasks has some negative sides like complexity and requires a lot of machine resources. But it is pretty useful and gives high results of recognition. Application of this method to n–dimensional description spaces gives a great opportunity in using this method in tasks like various diagnostic tasks. Work in this direction requires a lot of serious research and is mostly experimenting. This work is reflecting just a small part of common method of fuzzy recognition based on binary relations and actually is included to a software system, which gets developing for experimental part of connected researches.

ISBN 9984 - 585 - 36 - 0

References

1. Ozols, Y. and Borisov, A. (1997), "Pattern Classification and Feature Extraction on the Basis of Composition of Fuzzy Relations." *Proceedings of the Seventh International Fuzzy Systems Association World Congress*, IFSA'97, Vol. III, Prague, June 25–29, pp. 129–134.
2. Zadeh, L. A. (1965), "Fuzzy sets and systems." *Proceedings of the Symposium on Systems Theory*, Polytechnic Institute of Brooklyn, pp. 29–37.
3. Borisov, A., Ozols, Y., and Grekovs, R. (2001), "Construction of Features and Decision Rules in Fuzzy Pattern Recognition Tasks." *Int. J. General Systems*, Vol. 30(1), pp. 23–43.

DATU NOLIKTAVAS REALIZĒŠANAS UN IZSTRĀDES RĪKI, TO NOVĒRTĒŠANAS KRITĒRIJI

DATA WAREHOUSE IMPLEMENTATION AND DEVELOPMENT TOOLS AND CRITERIA TO EVALUATE THEM

Aleksejs Jevsejenko, LU Datorzinātņu maģistrantūras maģistrants,
A/S "Dati", LR VID datu noliktavas projekta grupas vadītājs, aleksejs.jevsejenko@dati.lv
Pēteris Guļāns, RTU ASTF Datorzinātņu maģistrantūras maģistrants,
A/S "Dati", LR VID datu noliktavas projekta sistēmanalītiķis, e–pasts: peteris.gulans@dati.lv

Abstract. *This paper provides an overview of data warehouse systems development and production tools. It also contains tools evaluation methodology. There are four types of tools used in data warehouse environment. The group of unique criteria is used for every tool evaluation. Each criterion has its own severity (critical, important, additional). Project specific features, chosen criteria set and its severity is a base for tools evaluation matrix. In addition, the unique criteria set for each type of tools is developed and described in this paper.*

The results of research are used in Latvia State Revenue Service data warehouse system development.

1. Ievads

Datu noliktavas (Data Warehouse) koncepcija radās astoņdesmito gadu sākumā. Pirms tam lielākais vairums informācijas apstrādes sistēmu bija transakciju sistēmas, kuras saturēja tikai pašreizējus datus, kuri nepārtraukti mainījās, kas ļāva risināt operatīvus, pašreizējus jautājumus un bija šauri orientētas viena biznesa uzdevuma risināšanai. Transakcijas sistēmu īpašības rada vairākas problēmas.

- ✓ Vēsturisko datu trūkums sistēmā.
- ✓ Liels informācijas apstrādes sistēmu skaits vienas organizācijas ietvaros.
- ✓ Datu integritātes problēma. Dažādās sistēmās dati tika glabāti un apstrādāti, izmantojot dažādus algoritmus, tāpēc gadījumos, kad tika veikti mēģinājumi apvienot sistēmas, tas bija neiespējams atšķirīgo datu struktūru dēļ.
- ✓ Problēmas ar lielu datu apjoma apstrādi, jo galvenais bija nodrošināt ātru un efektīvu ievades, glabāšanas un apstrādes iespēju, nevis nodrošināt nestandarta analīzes iespējas.
- ✓ Izmaiņu veikšana sistēmās bija dārgs un darbietilpīgs process.

ISBN 9984 - 585 - 36 - 0