Intelligent Information Systems 2008
ISBN 978-83-60434-44-4, pages 397–403

Relations in decision tables

Barbara Marszał-Paszek and Piotr Paszek

Institute of Computer Science, University of Silesia
Będzińska 39, 41–200 Sosnowiec, Poland
{bpaszek,paszek}@us.edu.pl

Abstract

In the present article, we show the possibility of using rules based on templates to get additional information from the decision table. This knowledge helps the classification process. The proposed method was tested on the group of decision tables from the UCI Machine Learning Repository, and the results are presented in this article.

Keywords: Rough set theory, Dempster-Shafer theory, templates

1 Introduction

The Dempster-Shafer theory (Dempster, 1967; Shafer, 1976) is based on belief functions, and plausible reasoning is used to combine separate pieces of information (evidence) to calculate the probability of an event. This theory is called a mathematical theory of evidence.

The rough set theory was proposed by Pawlak (1982) as a mathematical tool for describing the uncertain knowledge.

The basic functions of the evidence theory were defined based on the notation from rough set theory in Grzymała-Busse (1987) and Skowron and Grzymała-Busse (1991). These definitions allow us to find useful dependencies in the decision tables. These dependencies were used by Marszał-Paszek and Paszek (2006) to find a minimal template in a given decision table.

2 Main Notions

Let $\mathbb{A}$ be an information system and let $A = C \cup D$ where C and D are non-empty, be disjoint subsets of A. The set C is called the set of *condition* attributes, and the set D is called the set of *decision* attributes. The triple $\mathbb{A} = (U, A, C, D)$ is referred to as a *decision table.* A simplified version of a decision table has the form $\mathbb{A} = (U, A \cup \{d\})$, where the set of decision attributes D is limited to one decision attribute only. The decision d creates a partition of the universe U into decision classes $X_1, \ldots, X_{r(d)}$, where $r(d) = |\{k : \exists_{x \in U} : d(x) = k\}|$ is the number of different values of the decision attribute and is called the *rank* of the decision d.

Let $\Theta_A = \{1, 2, ..., r(d)\}$ be the frame of discernment defined by the decision d in the decision table $\mathbb{A}$.

For any $\theta \in \Theta_A$ the following equalities hold (Skowron and Grzymała-Busse, 1991):

$$Bel_A(\theta) = \frac{\left|\underline{A} \bigcup_{i \in \theta} X_i\right|}{|U|},$$

$$Pl_A(\theta) = \frac{\left|\overline{A} \bigcup_{i \in \theta} X_i\right|}{|U|}.$$

A *template* T (Nguyen *et al.*, 1997) in a decision table is any sequence $v_1, \ldots, v_n$, where

$$v_i \ \in \ V_{a_i} \ \cup \ \{*\}.$$

The symbol $'*'$ appearing in a given template means that the value of the marked attribute is not restricted by the template.
Alternatively, a template can be defined as the conjunction of a certain number of descriptors, e.g.,

$$T = (c = 0) \wedge (e = 1) \wedge (f = 1).$$

A given object matches a given template if $a_i(x) = v_i$, for each i such that $v_i \neq' *'$.

The following minimal templates problem (MTP) was considered (Marszał-Paszek and Paszek, 2006): Let A be a decision table $\mathbb{A}$; thresholds $\varepsilon_1, \varepsilon_2 \in (0, 1)$ and a natural number $1 \leq k < r(d)$. We would like to get minimal templates T (with respect to the length) for which a set $\theta \subseteq \Theta_{\mathbb{A}_T}$ exists with, at most, k elements ($|\theta| \leq k$) that satisfy the following conditions:

$$|Pl_{\mathbb{A}_T}(\theta) - Bel_{\mathbb{A}_T}(\theta)| < \varepsilon_1 \quad for \quad \varepsilon_1 \in (0,1); \quad \theta \subseteq \Theta_{A_T}; \tag{1}$$

$$|Pl_{\mathbb{A}_T}(\theta)| > 1 - \varepsilon_1 \quad for \quad \varepsilon_1 \in (0,1); \quad \theta \subseteq \Theta_{A_T}; \tag{2}$$

$$\frac{|U_T|}{|U|} > \varepsilon_2 \quad for \quad \varepsilon_2 \in (0,1). \tag{3}$$

Based on conditions 1-3 that direct the searching process of the minimal templates, the following rule is obtained: $T \Rightarrow \theta$, where the conditional part of the rule is a template, and the decision part is a set θ.

For example, when $\theta \subset \Theta_{\mathbb{A}} = \{1, 2, 3, 4, 5\}$ the rule in the shape of $(a = 1) \wedge (b = 0) \wedge (c = 1) \Rightarrow \neg(d = 3) \wedge \neg(d = 4) \wedge \neg(d = 5)$ seems to be more useful in the case where there are no strong rules (with the right hand side described by a single decision value) in a given decision table that have satisfactory support.

3 Classifications

In order to use the generated rules based on the templates, the following classification methods were proposed: 1) voting on templates, 2) the excluding method, and 3) the expanding method. `Voting on templates` is defined in the following way: having the set of rules in the shape of $\{T \Rightarrow \theta\}$ and the object x, that we want to classify, we decide to choose:

- these rules, in that T fits to the conditional part of the object x,
- then, we choose T with the largest support,
- after the T is selected, the set of the decision θ will be determined,
- during the classification process, all decisions that enter the set θ are rewarded.

In the classification process, in the excluding method and expanding method, we use minimal rules that are generated in RSES (Bazan *et al.*, 2004) and rules based on templates. On rules from RSES, we start standard voting, and on rules based on templates, we start `voting on templates`. The information obtained from the rules and templates is used in the classification process. Let the set of possible decisions for the given decision table be of the shape $\{1, 2, 3, 4, 5\}$. The following process is started for the proposed methods:

`Excluding classification method:`

1. When we receive from the rules (standard voting) {1} and from the templates (voting on templates) {1,2,3}, then we assign the decision {1,2,3}. That is, in the process of the classification, if the new object has the decision 4 or 5, it is a classification error.
2. When we receive from the rules (standard voting) {4} and from the templates (voting on templates) {1,2,3} `(conflict)`, then we assign the decision {1,2,3,4}. Justification: $(\neg 4 \wedge \neg 5) \vee 4 \equiv \neg 5 \vee 4 \equiv 1 \vee 2 \vee 3 \vee 4$.
3. When we receive from the rules {1} and from the templates (voting on templates) empty set (doesn't match to any templates), then we assign the decision {1}. We also proceed in a such way in the opposite situation.
4. In the remaining cases the object is not classified.

`Extending classification method:`

1. When we receive from the rules (standard voting) {1} and from the templates (voting on templates) {1,2,3} `(no conflict)`, then we assign the single decision {1}.
2. When we receive from the rules (standard voting) {4} and from the templates (voting on templates) {1,2,3} `(conflict)`, then we assign the decision {4}, if rule support is larger than template support.
   In the opposite case, we assign the single decision from the set { 1,2, 3 }, that wins the standard voting on the rules.
3. When we receive from the rules {1} and from the templates (voting on templates) an empty set (doesn't match to any templates), then we assign the decision {1}. We also proceed in the same way in the opposite situation.
4. In the remaining cases the object is not classified.

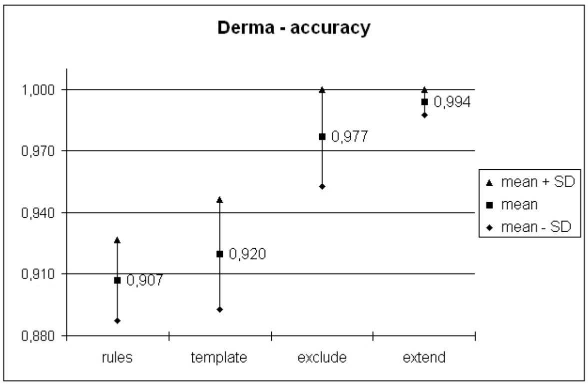

FIGURE 1: Accuracy of the classification for the dataset Derma

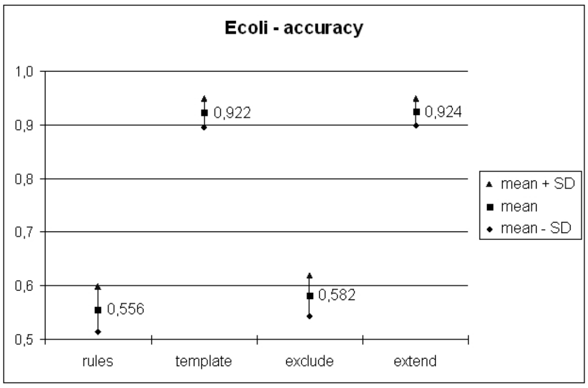

FIGURE 2: Accuracy of the classification for the dataset Ecoli

4 Results

The classification methods proposed in the previous section were tested on the group of decision tables from the UCI Machine Learning Repository (Blake and Merz, 1998). During the tests, four classification schemes were used:

- standard voting on the rules from RSES (only rules),
- voting on templates (only templates),
- excluding method (rules and templates),
- expanding method (rules and templates).

Table 1 contains basic information about tested datasets and information about the success rates for all classification schemes. For the datasets presented in Table 1, the "train & test" method was used several times. Then, the average error (success) rates with standard deviation were calculated.

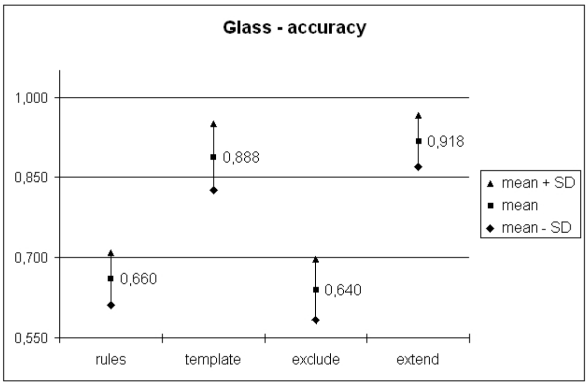

FIGURE 3: Accuracy of the classification for the dataset Glass

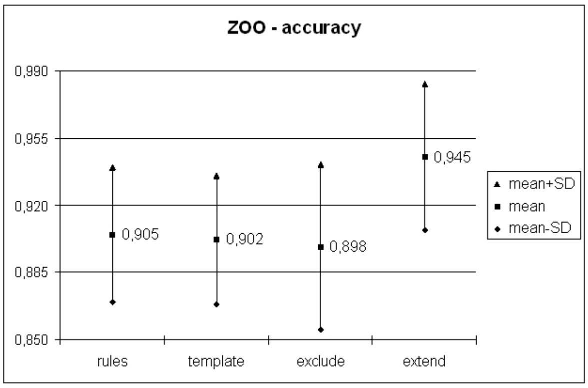

FIGURE 4: Accuracy of the classification for the dataset ZOO

Our work demonstrated that the use of the classification on templates themselves improved the quality of the classifications in comparison to the rules themselves for datasets Derma, Ecoli, Glass and ZOO. The number of rules based on templates for the given decision table is small. Despite this, the classification error is comparable from minimal rules error rate.

The classification on rules and templates (excluding the classification method) gave better results for the Derma and Ecoli datasets. The extending classification method always improved the quality of the classification (reduced the error rate).

Figure 1 contains information on the accuracy of the classification with standard deviation for the dataset Derma. Figures 2, 3 and 4 contain information on the accuracy of the classification for the datasets Ecoli, Glass, and ZOO respectively.

Table 1: General information on the dataset and the success rate with standard deviation for four classification schemes

Data	General Info			"Train & Test" – succes rate			
	attribute	example	class	rules	template	exlude	extend
Derma	34	366	6	90.7±1.2	97.7±2.3	92.0±2.7	99.4±0.6
Ecoli	7	336	8	55.6±4.2	92.2±2.7	58.2±3.8	92.4±2.6
Glass	9	214	3	66.0±4.9	88.8±6.1	64.0±5.6	91.8±4.8
ZOO	17	101	7	90.5± 3.5	89.8± 4.3	90.2± 3.3	94.5± 3.8

5 Conclusions

In this paper, we have demonstrated that the relationships between rough-set theory and evidence theory can be used to find the relations for a given decision table. These relations are mean rules based on templates. Such rules are additional information that can be obtained from the decision table. This information is used to support classification process. In the article, it was shown that templates used in the process of classification increased classification accuracy (reduced error rates) of new examples.

6 Acknowledgments

The research has been partially supported by the project "Decision support — new generation systems" of Innovative Economy Operational Programme 2007-2013 (Priority Axis 1. Research and development of new technologies) managed by Ministry of Regional Development of the Republic of Poland.

References

J.G. Bazan, M.S. Szczuka, A. Wojna, and M. Wojnarski (2004), On the Evolution of Rough Set Exploration System, in S. Tsumoto, R.W. Słowiński, J. Komorowski, and J.W. Grzmala-Busse, editors, *Rough Sets and Current Trends in Computing*, *Lecture Notes in Artificial Intelligence*, vol. 3066, pp. 592–601, Springer-Verlag, Berlin, Germany.

C.L. Blake and C.J. Merz (1998), UCI repository of machine learning databases, Technical report.

A.P. Dempster (1967), Upper and lower probabilities induced from a multivalued mapping, *Annals of Mathematical Statistics*, (38):325–339.

J.W. Grzymała-Busse (1987), Rough-set and Dempster-Shafer Approaches to Knowledge Acqusition uder Uncertainty-a Comparison, Technical report, Department of Computer Science University of Kansas.

B. Marszał-Paszek and P. Paszek (2006), Minimal Templates Problem, in M.A. Kłopotek, S.T. Wierzchoń, and K. Trojanowski, editors, *Intelligent Information Processing and Web Mining*, *Advances in Soft Computing*, pp. 397–402, Springer-Verlag, Heidelberg, Germany.

S.H. Nguyen, A. Skowron, P. Synak, and J. Wróblewski (1997), Knowledge Discovery in Databases: Rough Set Approach, vol. 2, pp. 204–209, Seventh International Fuzzy Systems Association World Congress (IFSA'1997), Academia, Prague, Czech Republic.

Z. Pawlak (1982), Rough Sets, *In T. J. of Information and Computer Sci.*, (11): 344–356.

G. Shafer (1976), *A mathematical theory of evidence*, Princeton University Press, Princeton.

A. Skowron and J.W. Grzymała-Busse (1991), From the Rough Set Theory to the Evidence Theory, Technical report 8/91, Institute of Computer Science Warsaw University of Technology, Poland.