TECHNICAL SCIENCES
Abbrev.: Techn. Sc., 2011, 14(1)

STABILITY OF OPTIMAL PARAMETERS FOR CLASSIFIER BASED ON SIMPLE GRANULES OF KNOWLEDGE

Piotr Artiemjew

Chair of Mathematical Methods in Computer Sciences
University of Warmia and Mazury in Olsztyn

K e y w o r d s: rough sets, rough inclusions, granules of knowledge, classification of data, missing values, stability of optimal parameters.

A b s t r a c t

Searching for optimal parameters of a classifier based on simple granules of knowledge investigated recently by the author (ARTIEMJEW 2010) raises a question about stability of optimal parameters. In this article, we will check dependence of stability of the optimal radius of granulation on random damage of decision system. The results of experiments show the dependence of stability on size of damage and strategies of treating missing values. This kind of research aims at finding methods of protecting decision systems which are vulnerable to damage against decreasing their classification effectiveness, which means preserving classifying possibilities similar to undamaged decision systems.

BADANIE STABILNOŚCI OPTYMALNYCH PARAMETRÓW KLASYFIKATORA BAZUJĄCEGO NA PROSTYCH GRANULACH WIEDZY

Piotr Artiemjew

Katedra Metod Matematycznych Informatyki
Uniwersytet Warmińsko-Mazurski w Olsztynie

S ł o w a k l u c z o w e: zbiory przybliżone, inkluzje przybliżone, granule wiedzy, klasyfikacja danych, uszkodzenia systemu decyzyjnego, stabilność optymalnych parametrów.

A b s t r a k t

Przeprowadzone w ostatnim czasie badania (ARTIEMJEW 2010) zmierzające do wyszukiwania optymalnych parametrów klasyfikacji modułów decyzyjnych opartych na prostych granulach wiedzy zrodziły pytanie o stabilność optymalnych parametrów klasyfikacji. W pracy sprawdzono zależność stabilności optymalnych promieni granulacji od losowego uszkadzania systemu decyzyjnego. Wyniki badań wskazały jednoznacznie, że istnieje zależność między stabilnością a wielkością uszkodzenia i strategiami traktowania wartości uszkodzonych. Tego typu badania mają na celu szukanie metod zabezpieczania systemów decyzyjnych, które są podatne na uszkodzenia, przed zmniejszaniem ich efektywności klasyfikacyjnej. Celem było zachowanie możliwości klasyfikacyjnych zbliżonych do efektywności nieuszkodzonych systemów decyzyjnych.

Introduction

Knowledge is understood here as ability to classify and as far as real phenomena are considered, the frame is that of information systems each of which is a pair $I = (U, A)$ where U is a set of objects, entities, and A is a set of attributes; decision system is a triple $DS = (U, A, d)$, where $d \notin A$. The basic form of granulation in decision and information systems consists in partitioning U into classes of the indiscernibility relation IND(A) defined as IND(A) $= \{(u, v) : a(u) = a(v), \forall a \in A\}$. Each class $[u]_A = \{v \in U : \mathrm{IND}_A(u, v)\}$ is interpreted as a elementary granule and unions of elementary granules are granules of knowledge. Another approach to granulation, proposed by (POLKOWSKI 2008), consists in using rough inclusions, cf. (POLKOWSKI 2008).

A rough inclusion is a relation $\mu \subseteq U \times U \times [0,1]$ which can be regarded as graded similarity relation extending the indiscernibility relation by relaxing restrictions on attribute values. We let $\mathrm{IND}(u, v) = \{a \in A : a(u) = a(v)\}$.

In our approach we use rough inclusion proposed by (POLKOWSKI 2008), to classify test objects. Test object is classified by granules, which have been formed from training set as follows,

$$g_{r_{\mathrm{gran}}}(u) = \{v \in U : \frac{|\mathrm{IND}(u, v)|}{|A|} \geq r_{\mathrm{gran}}\}$$

where $r_{\mathrm{gran}} \in [0,1]$ is the granulation radius.

The most numerous decision class transfers decision to our testing object. If tie occurs, it is resolved by a random choice. This type of classification is the simplest among studied by the author cf. (ARTIEMJEW 2008, 2009, POLKOWSKI, ARTIEMJEW 2007). The results of research for optimal parameters for this method is available in (ARTIEMJEW 2010). In this work we continue this approach and our main purpose is to find the threshold of the optimal parameter stability for the random damage of the decision system.

In the work (ARTIEMJEW 2010) we have proposed a method for experimental detecting of the optimal radius value for a given data set, by means of multiple CV-5 and subsequent confirmation by means of Leave One Out method. Once the optimal value is found for the test data, it can be used for classifying incoming objects without any need for full granulation procedure. For the multiple damage of the decision system in fixed percentage, the stability of optimal parameter is checked by Leave One Out Method. Optimal radius of classification is stable if despite the damage of the decision system it is still the same and fulfills the criterion of optimality, which is a maximal value of accuracy within coverage in the range of [0.9, 1.0].

Classification by simple granules of knowledge theoretical background

Rough inclusion generally is a predicate of the form $\mu_\pi\ (x, y, r)$, where x, y are individual objects, $r \in [0,1]$, which satisfies the following requirements, relative to a given part relation π on a set U of individual objects, see (POLKOWSKI 2008),

$$1)\ \mu_\pi\ (x, y, 1) \Leftrightarrow \mathrm{ing}_\pi\ (x, y);$$

$$2)\ \mu_\pi\ (x, y, 1) \Rightarrow [\mu_\pi\ (z, x, r) \Rightarrow \mu_\pi\ (z, y, r)]$$

$$3)\ \mu_\pi\ (x, y, r) \wedge s < r \Rightarrow \mu_\pi\ (x, y, s)$$

Those requirements seem to be intuitively clear. 1) demands that the predicate μ_π is an extension to the relation ing_π of the underlying system of Mereology; 2) does express monotonicity of μ_π, and 3. assures the reading. “to degree at least r”. We use here only one rough inclusion, albeit a fundamental one, viz., (POLKOWSKI 2008) for its derivation,

$$\mu_L\ (u, v, r) \Leftrightarrow \frac{|\ \mathrm{IND}\ (u, v)\,|}{|A\,|} \geq r.$$

A granule $g_\mu\ (u, r)$ about $u \in U$ of the radius r, relative to μ, is defined by letting,

$$g_\mu\ (u, r) \text{ is ClsF } (u, r),$$

where the property $\mathrm{F}(u, r)$ is satisfied with an object v if and only if $\mu\ (v, u, r)$ holds, and Cls is the class operator, see, e.g., (POLKOWSKI 2008). Practically, in case of μ_L the granule $g(u, r)$ collects all $v \in U$ such that $|\,\mathrm{IND}(v, u)\,| \geq r \cdot |A\,|$.

Error evaluation

Classifiers are evaluated by error which is the ratio of the number of correctly classified objects to the number of recognized test objects (called also total accuracy) and total coverage, $\frac{\mathrm{rec}}{\mathrm{test}}$, where is the number of recognized test cases and test is the number of test cases.

The results by Leave One Out method should be modified. We build for Leave One Out method the confusion matrix, where testing objects from all folds are treated as one test decision system, and we compute accuracy as percentage of correctly classified testing objects. Coverage is a percent of objects which have been classified. The motivation to use Leave One Out method is to be found, among other places in (MOLINARO, SIMON, PFEIFFER 2005). This paper proves the effectiveness and almost unbiased character of this method.

Voting by granules on decision values

Given a granule $g = g_{r_{\mathrm{gran}}}(u)$ of test objects u in a training decision system (U_{trn}, A, d), for each test object u, the value of decision assigned to u by the granule g is defined as, $d(u) = d'$ such that, for D as the set of granule's $g_{r_{\mathrm{gran}}}(u)$ decision classes,

$$d' = \{d'' \in D : \max |\{v \in g_{r_{\mathrm{gran}}}(u) : d(v) = d''\}|\}$$

in case of tie, random choice is applied.

The procedure of classification by means of Standard Granules (CSG algorithm)

1. The training decision system (U_{trn}, A, d) and the test system (U_{tst}, A, d) has been input, where U_{tst}, U_{trn} is respectively universe of test and training objects, A is a set of attributes and d is a decision reflecting a partition of objects into classes.

2. The granular radius $r_{\mathrm{gran}} \in \{\frac{0}{\mathrm{card}\{A\}}, \frac{1}{\mathrm{card}\{A\}}, \ldots, \frac{\mathrm{card}\{A\}}{\mathrm{card}\{A\}}\}$ has been chosen.

3. For classified test object $u \in U_{\mathrm{tst}}$, the classification granule $g_{r_{\mathrm{gran}}}(u)$ has been found in the training set U_{trn} as follows, $g_{r_{\mathrm{gran}}}(u) = \{v \in U_{\mathrm{trn}} : \frac{|\mathrm{IND}(u,v)|}{|A|} \geq r_{\mathrm{gran}}\}$, where $r_{\mathrm{gran}} \in [0,1]$.

4. The most numerous decision class of granule $g_{r_{\mathrm{gran}}}(u)$ transfers decision to our test object. If tie occurs, it is resolved randomly.

Empirical method of checking the stability of the optimal parameters

By the x percent damage of the decision system (U, A, d), we mean that we insert a number of $x\% * |A| * |U|$ random stars inside the system. Credit Approval (UCI Repository) contains a little bit of original missing values marked as "?", but it doesn't matter for us, because we treat these missing values as additional descriptors. Only artificial damage created by stars is considered as missing values. As we can see in Figure 1, the stability of optimal parameter is checked by five times random damage of the original decision system and classified by the leave one out method, where finally we obtain an average result. Radius is stable when in spite of the damage is the same as original decision system optimal radius.

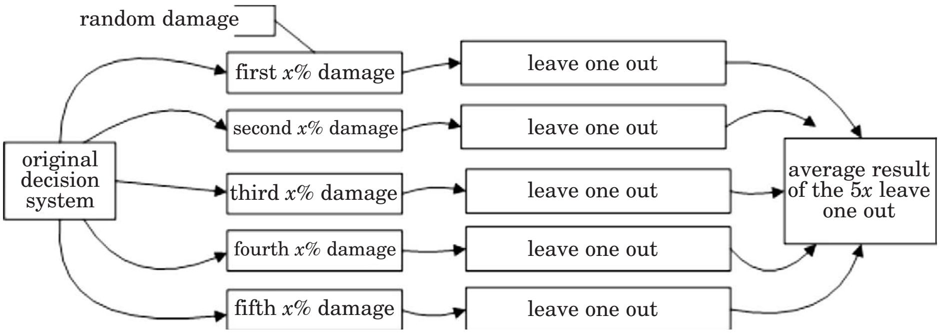

Fig. 1. Stability checking method of the optimal radius for the classifier based on simple granules of knowledge (CSG)

Methods for treating missing values

During the classification of the damaged decision system by simple granules of knowledge we are considering two ways of treating missing values. The first one is „*=*" and the second one is strategy „*=don't care". For both strategies we are using the same damaged decision systems. For instance for Credit Approval (UCI Repository) decision system we are damaging $x\%*15*690$ descriptors. Details of percentage damaged for the considered decision system are shown in Table 1.

Table 1

Detailed information regarding the percentage damage of Credit Approval (UCI Repository); Damper = Percentage damage of decision system, Starnum = Number of stars in damaged decision system

Damper [%]	Starnum	Damper [%]	Starnum	Damper [%]	Starnum
1	103	6	621	15	1552
2	207	7	724	20	2070
3	310	8	828	25	2587
4	414	9	931	30	3105
5	517	10	1035		

Results of experiments for the strategy „*=*”

Results for Credit Approval at percentage damage from interval [1% to 10%]

At first we would like to show results in details for the first one-percentage-damage, see Table 2. Short description for all 1 percentage damage and average result we can see in Table 3 and finally, in Figure 2 we can see visualization of error for all 1 percentage damage. The results for damages from interval 2% to 10% will be shown in the short form only, see Table 4 and 5.

Table 2

„*=*”; Leave One Out; First 1% damage; Credit Approval (UCI Repository); Classification by means of standard granules (CSG); r_{gran} = Granulation radius, $T._{acc}$ = Total accuracy, $T._{cov}$ = Total coverage, $M._{gran}$ = The mean percentage size of classification granule in training system; The best results for: $r_{gran} = 0.6$, $T._{acc} = 0.846497$, $T._{cov} = 0.972464$, $M._{gran} = 24.3942$

r_{gran}	$T._{acc}$	$T._{cov}$	$M._{gran}$
0	0.555072	1	689
0.0666667	0.555072	1	687.104
0.133333	0.555072	1	673.4
0.2	0.573913	1	633.51
0.266667	0.750725	1	559.814
0.333333	0.826087	1	442.325
0.4	0.824638	1	296.896
0.466667	0.833091	0.998551	165.516
0.533333	0.836972	0.995652	73.2928
0.6	0.846497	0.972464	24.3942
0.666667	0.837113	0.702899	5.53623
0.733333	0.838093	0.304348	0.802899
0.8	0.833333	0.0695652	0.110145
0.866667	0.750001	0.0173913	0.0376812
0.933333	0.799999	0.00724638	0.0115942
1	0	0	0

Table 3

„*=*"; Leave One Out; All results for 1% of the damage; Credit Approval (UCI Repository); Classification by means of standard granules (CSG); Optimal.r_{gran} = Optimal granulation radius, $T._{\text{acc}}$ = Total accuracy, $T._{\text{cov}}$ = Total coverage, $M._{\text{gran}}$ = The mean size of classification granule in training system, $M._{\text{trn}}$ = Mean training table size

	Optimal.r_{gran}	$T._{\text{acc}}$	$T._{\text{cov}}$	$M._{\text{gran}}$	$\frac{M._{\text{gran}}}{M._{\text{trn}}} \cdot 100\%$ [%]
nil	0.6	0.860327	0.975362	26.7913	3.89
First 1% damage	0.6	0.846497	0.972464	24.3942	3.54
Second 1%damage	0.6	0.856049	0.946377	23.8667	3.46
Third 1% damage	0.6	0.854572	0.966667	23.7594	3.45
Fourth 1% damage	0.6	0.861862	0.965217	23.7275	3.44
Fifth 1% damage	0.6	0.857576	0.956522	23.7768	3.45
Average result	0.6	0.855312	0.96145	23.905	3.47

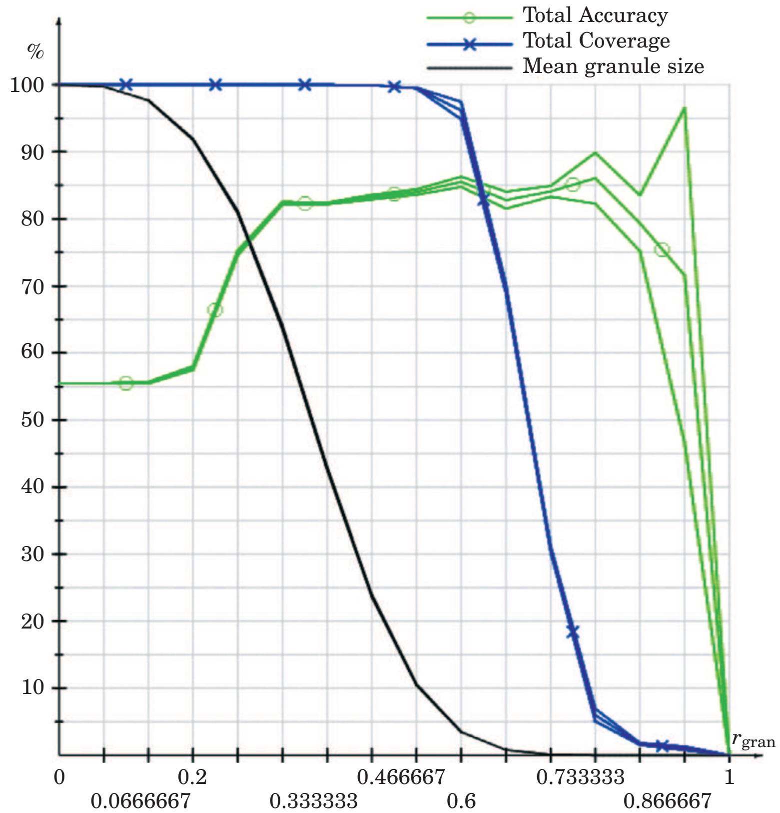

Fig. 2. „*=*"; Leave One Out; Average results of experiments and error visualization for all 1% of the damage; $\frac{(\max - \min)}{2}$; Credit Approval (UCI Repository); Classification by means of standard granules (CSG); r_{gran} = Granulation radius, $T._{\text{acc}}$ = Total accuracy, $T._{\text{cov}}$ = Total coverage, $M._{\text{gran}}$ = The mean percentage size of classification granule in training system; The best results for: r_{gran} = 0.6, $T._{\text{acc}}$ = 0.855312, $T._{\text{cov}}$ = 0.96145, $M._{\text{gran}}$ = 23.905

Table 4

„*=*"; Leave One Out; Average results for all damage; Credit Approval (UCI Repository); Classification by means of standard granules (CSG); Damper = Percentage damage of decision system, Optimal.r_{gran} = Optimal granulation radius, $T._{\text{acc}}$ = Total accuracy, $T._{\text{cov}}$ = Total coverage, $M._{\text{gran}}$ = The mean size of classification granule in training system, $M._{\text{trn}}$ = Mean training table size

Damper [%]	Optimal.r_{gran}	$T._{\text{acc}}$	$T._{\text{cov}}$	$M._{\text{gran}}$	$\frac{M._{\text{gran}}}{M._{\text{trn}}} \cdot 100\%$ [%]
nil	0.6	0.860327	0.975362	26.7913	3.89
1	0.6	0.855312	0.96145	23.905	3.47
2	0.6	0.850606	0.935074	20.5066	2.98
3	0.6	0.85161	0.92174	17.607	2.56
4	0.533333	0.84138	0.990436	52.3532	7.6
5	0.533333	0.844274	0.984638	46.987	6.82
6	0.533333	0.837318	0.98348	43.0766	6.25
7	0.533333	0.83784	0.975942	39.0534	5.67
8	0.533333	0.835582	0.973044	35.5524	5.16
9	0.533333	0.833924	0.97565	31.2922	4.54
10	0.466667	0.834204	0.99826	81.524	11.83

Table 5

„*=*"; Leave One Out; Table of errors' percentage for all damage; $\frac{(\max - \min.)}{2}$; Credit Approval (UCI Repository); Classification by means of standard granules (CSG); Damper = Percentage damage of decision system, Optimal.r_{gran} = Optimal granulation radius, $T._{\text{acc}}$ = Total accuracy, $T._{\text{cov}}$ = Total coverage, $M._{\text{gran}}$ = The mean size of classification granule in training system

Damper [%]	Optimal.r_{gran}	$T._{\text{acc}}$ [%]	$T._{\text{cov}}$ [%]	$M._{\text{gran}}$ [%]
1	0.6	0.76	(+-)1.3	(+-)0.0483
2	0.6	(+-)0.54	(+-)0.94	(+-)0.0538
3	0.6	(+-)1.48	(+-)0.79	(+-)0.0961
4	0.533333	(+-)0.69	(+-)0.43	(+-)0.2614
5	0.533333	(+-)1.08	(+-)0.43	(+-)0.1962
6	0.533333	(+-)0.67	(+-)0.57	(+-)0.2711
7	0.533333	(+-)1.61	(+-)0.65	(+-)0.2097
8	0.533333	(+-)1.15	(+-)0.5	(+-)0.2908
9	0.533333	(+-)1.15	(+-)0.28	(+-)0.1079
10	0.466667	(+-)0.73	(+-)0.07	(+-)0.3075

Additional result for Wisconsin Diagnostic Breast Cancer data set

In Table 6 additional results obtained from classification by means of simple granular structures, for the damage in range 1 to 7% are collected.

Table 6

„*=*"; Leave One Out; Average results for all damage; Wisconsin Diagnostic Breast Cancer (UCI Repository); Classification by means of standard granules (CSG); Damper = Percentage damage of decision system, Optimal.r_{gran} = Optimal granulation radius, $T._{acc}$ = Total accuracy, $T._{cov}$ = Total coverage, $M._{gran}$ = The mean size of classification granule in training system, $M._{trn}$ = Mean training table size

Damper [%]	Optimal.r_{gran}	$T._{acc}$	$T._{cov}$	$M._{gran}$	$\frac{M._{gran}}{M._{trn}} \cdot 100\%$ [%]
nil	0.258065	0.870722	0.924429	33.7047	5.93
1	0.258065	0.874908	0.921616	30.8408	5.43
2	0.258065	0.875358	0.907908	28.6496	5.04
3	0.258065	0.879718	0.908964	26.168	4.61
4	0.258065	0.877342	0.902638	23.6408	4.16
5	0.258065	0.873384	0.902286	22.9652	3.9
6	0.258065	0.87183	0.902284	21.084	3.58
7	0.225806	0.855392	0.96485	46.9272	7.97

A summary for the strategy „*=*"

At the fixed optimality condition: Optimal.r_{gran} = $\{r_{gran}$: for maximal accuracy at coverage $\in$ [0.9,1.0]$\}$, within „*=*" strategy for Credit Approval (UCI Repository) the optimality of the radius during classification is stable for damages of 1, 2, 3%. Above three percent, beginning from 4% coverage for radius r_{gran} = 0.6 is becoming smaller than 0.9, hence, in spite of maximal accuracy, our radius ceases being optimal. For damages of 4–7% accuracy is still maximal for radius 0.6, but coverage $\in$ [0.8; 0.9]. Starting from 8% to 10%, the best accuracy is for radius 0.533333. For 10% damage at r_{gran} = 0.6 coverage $\in$ [0.7, 0.8). Finally when we are applying the strategy „*=*" during the classification CSG, stability of the optimum is low and is kept for 1,2 and 3% damage. However, the level of accuracy is high for this damage. Classification error resulting from randomization of damage is following. for 1–3% damage $T._{acc.error} \in$ [0.54%, 1.48%], $T._{cov.error} \in$ [0.79%, 1.3%], for 4–7% damage $T._{acc.error} \in$ [0.88%, 1.29%], $T._{cov.error} \in$ [0.86%, 1.52%] and finally for 8–10% damage $T._{acc.error} \in$ [1.21%, 1.45%], $T._{cov.error} \in$ [0.79%, 2.17%]. Result shows that there is no bigger dependency between the classification error resulting from randomization and the percentage of damage.

Preserving the radius' optimality is dependent on the classification granule size. The bigger the damage is, the coverage and $M._{\text{gran}}$ for originally optimal radius becomes smaller. The optimum is moving to the direction of the radius 0.533333 for which classification granules have similar size to classification granules of the originally optimal radius. It is possible to predict that the optimum with the increase of the damage will fall down. The additional result for Wisconsin Diagnostic Breast Cancer (UCI Repository) see Table 6 confirm this hypothesis.

Results for the strategy „*=don't care"

Results for Credit Approval at percentage damage from interval [1 to 10%] and 15, 20, 25, 30%

The results for damages from interval 1 to 10% and 15, 20, 25, 30% will be shown in the short form only, see Table 7.

Table 7

„*=don't care"; Leave One Out; Average results for all damage; Credit Approval (UCI Repository); Classification by means of standard granules (CSG); Damper = Percentage damage of decision system, Optimal.r_{gran} = Optimal granulation radius, $T._{\text{acc}}$ = Total accuracy, $T._{\text{cov}}$ = Total coverage, $M._{\text{gran}}$ = The mean size of classification granule in training system, $M._{\text{trn}}$ = Mean training table size

Damper [%]	Optimal.r_{gran}	$T._{\text{acc}}$	$T._{\text{cov}}$	$M._{\text{gran}}$	$\frac{M._{\text{gran}}}{M._{\text{trn}}} \cdot 100\%$ [%]
nil	0.6	0.860327	0.975362	26.7913	3.89
1	0.6	0.854102	0.977392	29.2134	4.24
2	0.6	0.843688	0.975362	30.5474	4.43
3	0.6	0.84391	0.97855	32.446	4.71
4	0.6	0.843138	0.98319	35.3892	5.14
5	0.6	0.84488	0.9829	36.4368	5.29
6	0.6	0.837492	0.988116	40.021	5.81
7	0.533333	0.836474	0.99971	109.477	15.89
8	0.6	0.834058	0.993912	45.1226	6.55
9	0.6	0.833386	0.993332	45.9836	6.67
10	0.533333	0.827828	1	121.15	17.58
15	0.6	0.822166	0.99913	67.351	9.78
20	0.6	0.807246	1	91.8064	13.3
25	0.533333	0.797104	1	234.658	34.06
30	0.666667	0.794494	1	70.849	10.28

Additional result for Wisconsin Diagnostic Breast Cancer data set

We include in Table 8 additional results obtained from classification by means of simple granular structures, for the damage in range 1 to 5%.

A summary for the strategy „*=don't care"

In considered strategy for Credit Approval (UCI Repository) stability of optimal radius is higher than in strategy „*=*". Optimal radius is preserved for 1–6% damage. Granules are being extended in this strategy and at the moment when they have too large noise inside the optimum is being lost. Starting from 7% damage, optimal radius is instable and reaches values 0.6 or 0.533333. At this method at the increase of damage, the coverage is growing to 1 because of the classifying granules' extension. The optimum is moving to the direction of radius 0.666667.

It is possible to expect that the optimal radius with the increase of the damage will move up. The additional result for Wisconsin Diagnostic Breast Cancer (UCI Repository) see Table 8 confirm this hypothesis.

Table 8

„*=don't care"; Leave One Out; Average results for all damage; Wisconsin Diagnostic Breast Cancer (UCI REPOSITORY); Classification by means of standard granules (CSG); Damper = Percentage damage of decision system, Optimal.r_{gran} = Optimal granulation radius, $T._{acc}$ = Total accuracy, $T._{cov}$ = Total coverage, $M._{gran}$ = The mean size of classification granule in training system, $M._{trn}$ = Mean training table size

Damper [%]	Optimal.r_{gran}	$T._{acc}$	$T._{cov}$	$M._{gran}$	$\frac{M._{gran}}{M._{trn}} \cdot 100\%$ [%]
nil	0.258065	0.870722	0.924429	33.7047	5.93
1	0.258065	0.862936	0.956766	41.5476	7.31
2	0.290323	0.869142	0.919506	23.9922	4.22
3	0.290323	0.837724	0.9529	29.8426	5.25
4	0.322581	0.844062	0.914938	15.7237	2.77
5	0.322581	0.79595	0.939192	20.22	3.43%

Conclusions

It follows from experiments that for small damage of decision system in the range of few percent, optimal parameters for the classifier basing on simple granules of knowledge maintain optimality. Range of the optimality is dependent on the strategy of treating missing values during classification and the decision system structure. For the exemplary Credit Approval (UCI Repository) system at the damage interval from 1 to 6% and strategy of treating missing values „*=don't care", the decision system still contains the same optimal parameter. For the strategy „*=*" stability is saved only for 1 to 3% damage, but in this case we have obtained better accuracy of classification than in strategy „*=don't care". For Wisconsin Diagnostic Breast Cancer (UCI Repository) at strategy „*=don't care" optimal radius is saved only for 1% damage and for the strategy „*=*" optimum is preserved for damage from interval 1 to 6%. Generally we can say that at „*=*" strategy optimal radius with the increase of the damage will move down and for „*=don't care" strategy it will move up. Finally the optimum of CSG is more stable for the strategy „*=*".

From the engineering point of view, obtaining an answer to the question of how to preserve stable level of classification in the case of real software and hardware discrete decision systems is really important, particularly if these decision systems are susceptible to noise and to damage conditional information. Thanks to such knowledge it is possible to design decision modules which are damage-resistant and are automatically adaptable to new conditions. In the future work, we will check the relation between the amount of information and stability of optimal parameters of classification based on simple granules of knowledge.

Acknowledgements

The research has been supported by grant 1309-802 from Ministry of Science and Higher Education of the Republic of Poland.

Translated by Ewa Sienkiewicz

Accepted for print 22.03.2011

References

Artiemjew P. 2009. *On strategies of knowledge granulation and applications to decision systems*. PhD Dissertation, Polish Japanese institute of Information Technology, L. Polkowski, Supervisor, Warsaw.

Artiemjew P. 2008. *Natural versus granular computing. Classifiers from granular structures*. Proceedings of 6th International Conference on Rough Sets and Current Trends in Computing RSCTC'08, Akron, Ohio, USA, vol. 5306, Springer Berlin / Heidelberg, pp. 150–159.

Artiemjew P. 2010. *In search of optimal parameters for classifier based on simple granules of knowledge*. III International Interdisciplinary Technical Conference of Young Scientists (InterTech 2010), 3: 138–142, Poznan University Press.

POLKOWSKI L. 2008. *A Unified Approach to Granulation of Knowledge and Granular Computing Based on Rough Mereology. A Survey*. In.: *Handbook of Granular Computing*. Eds. W. Pedrycz, A. Skowron, V. Kreinovich. John Wiley & Sons, New York, pp. 375–401.

POLKOWSKI L., ARTIEMJEW P. 2008. *Rough mereology in classification of data. Voting by means of residual rough inclusions*. Proceedings of 6th International Conference on Rough Sets and Current Trends in Computing RSCTC'08, Akron, Ohio, USA.

POLKOWSKI L., ARTIEMJEW P. 2007. *On granular rough computing. Factoring classifiers through granular structures*. In: *Rough Sets and Intelligent Systems Paradigms, International Conference, RSEISP*. Eds. M. Kryszkiewicz, J.F. Peters, H. Rybiński, A. Skowron. LNAI, 4585: 280–290. Springer, Berlin.

POLKOWSKI L. 2007. *Granulation of knowledge in decision systems. The approach based on rough inclusions. The Method and its applications (plenary talk).* Lecture Notes in Artificial Intelligence (Proceedings RSEiSP'07), 4585: 69–79.

MOLINARO A.M., SIMON R., PFEIFFER R.M. 2005. *Prediction error estimation. A comparison of resampling methods*. Bioinformatics, 21(15): 3301–3307, Oxford University Press, Oxford, UK.

UCI Repository: http.//www.ics.uci.edu/mlearn/databases/