Granulation of Knowledge in Decision Systems: The Approach Based on Rough Inclusions. The Method and Its Applications

Lech Polkowski

Polish –Japanese Institute of Information Technology
Koszykowa 86 02008 Warszawa, Poland
polkow@pjwstk.edu.pl

Abstract. Rough set approach to knowledge entails its granulation: knowledge represented as a collection of classifications by means of indiscernibility of objects consists of indiscernibility classes that form elementary granules of knowledge. Granules of knowledge that emerge as unions of elementary granules are also characterized as exact concepts that are described with certainty. Relaxing of indiscernibility relations has led to various forms of similarity relations. In this lecture, we discuss the approach to similarity rooted in mereological theory of concepts, whose primitive notion is that of a rough inclusion. Rough inclusions are predicates/relations of a part to a degree. Partial containment is the basic underlying phenomenon related to uncertainty, therefore rough inclusions allow for a formalization of a wide spectrum of contexts in which reasoning under uncertainty is effected.

Granules are formed by means of rough inclusions as classes of objects close to a specified center of the granule to a given degree; formally, they resemble neighborhoods formed with respect to a certain metric. Classes of objects in turn are defined by the class operator borrowed from mereology. The usage of mereological techniques based on the notion of a part is justified by its greater elegance and transparency in comparison to the naive theory of concepts based on the notion of an element.

At IEEE GrC 2005, 2006 the Author put forth the idea of a granular information/decision system whose objects are granules formed from the original information/decision system; the idea was issued along with the hypothesis that granular systems at sufficiently large radii of granulation, should preserve information about objects coded in the attribute–value language to a sufficiently high degree. This idea is here discussed along with results of some tests that bear it out.

The second application that is reflected in the lecture is about missing values; the approach discussed here is also based on granulation and the idea is to absorb objects with missing values into granules of knowledge in order to replace in a sense the missing value with a defined one decided by the granule.

Keywords: granulation of knowledge, rough sets, rough inclusions, granular decision systems.

M. Kryszkiewicz et al. (Eds.): RSEISP 2007, LNAI 4585, pp. 69–79, 2007.

1 The Idea of Computing Via Rough Sets

Classical ideas about representations of uncertainty, expressed respectively by Gottlob Frege and Max Black, found realization respectively in rough and fuzzy concept theories. Despite their formal differences and distinct starting points, both compute with granules of objects: rough sets with indiscernibility classes of objects, fuzzy sets with inverse images of fuzzy membership functions.

Rough sets represent knowledge by means of *information systems*, i.e., pairs of the form (U, A) where U is a set of *objects* and A is a set of *attributes* with each $a \in A$ a mapping $a : U \rightarrow V_a$ on U into the value set V_a. Objects are coded by their *information sets* of the form $inf(u) = \{(a = a(u)) : a \in A\}$. Objects u, v with $inf(u) = inf(v)$ are called *indiscernible* and they are regarded as identical with respect to the given set A. The *B–indiscernibility relation relative to a set* $B \subseteq A$ is $ind(B) = \{(u, v) : \forall a \in B.a(u) = a(v)\}$. Classes $[u]_B = \{v : (u, v) \in ind(B)\}$ are *B−elementary granules* of knowledge. Their unions are *B−granules* of knowledge.

A formula $(a = v)$ is an *elementary descriptor*; *descriptors* are formed as the smallest set containing all elementary descriptors and closed under sentential connectives $\vee, \wedge, \neg, \Rightarrow$. The meaning $[a = v]$ of an elementary descriptor is defined as the set $\{u : a(u) = v\}$ and it is recursively extended to meaning of descriptors [8].

Decision systems are information systems of the form $(U, A \cup \{d\})$ with a singled out attribute d called the *decision* that does represent a description of objects by an external informed source (say, an expert). Description of d in terms of *conditional* attributes in the set A is effected by means of *decision rules* [8] of the form

$$\bigwedge_{a \in B} (a = v_a) \Rightarrow (d = v). \tag{1}$$

The rule (1) is *true* whenever $[\bigwedge_{a \in B}(a = v_a)] \subseteq [d = v]$; otherwise it is partially true; see, e.g., [10] for a review of this topic.

2 Granulation of Knowledge

The issue of granulation of knowledge as a problem on its own, has been posed by L.A.Zadeh [23]. The issue of granulation has been a subject of intensive studies within rough set community, as witnessed by a number of papers, see, e.g., [17], [18].

Granules defined by indiscernibility and their direct generalizations to various similarity classes of tolerance, asymmetric similarity relations and general binary relations were subject to an intensive research, see, e.g. [7], [22]. Granulation of knowledge by means of rough inclusions was studied in [16].

Granulation of knowledge and applications to knowledge discovery in the realm of approximation spaces were studied, among others, in [20].

A study of granule systems was also carried out in [11], [12], [13], [14], in order to find general properties of granules. In proofs of those properties, techniques of mereology were applied as more simple and elegant than those of naive set theory.

2.1 The Technique of Mereology

Fundamental in mereology [6] is the relation of a *part*, π, that given a universe U, does satisfy the following conditions ,

$$1.\neg(x\pi x).2.x\pi y \wedge y\pi z \Rightarrow x\pi z, \tag{2}$$

i.e., it is transitive (cond. 2) and irreflexive (cond. 1).

The notion of an *element*, associated with the part relation π, is expressed with the help of the notion of an ingredient ing_π, informally an "improper part",

$$x\ ing_\pi\ y \Leftrightarrow\ x\ \pi\ y\ \vee\ x = y. \tag{3}$$

Mereology is a theory of individual objects, that decompose into parts, and passing to it from Ontology - theory of distributive concepts, is realized by means of the *set/class operator* [6]; given a non–empty collection F of objects, i.e., an ontological concept F, the individual representing F is given as the class of F, $Cls_\pi F$, subject to the following conditions,

$$\begin{array}{c} 1.\ u \in\ F \Rightarrow u\ ing_\pi\ Cls_\pi F. \\ 2.\ u\ ing_\pi\ Cls_\pi F \Rightarrow \forall v.[v\ ing_\pi\ u \Rightarrow \exists w,t.\ w\ ing_\pi\ v, w\ ing_\pi\ t, t \in\ F]. \end{array} \tag{4}$$

In the sequel, the subscript π will be mostly omitted.

In plain words, $ClsF$ consists of those objects whose each part has a part in common with an object in F; the reader will easily recognize that the union $\bigcup F$ of a family F of sets is the class of F with respect to the part relation $\subset$.

2.2 Rough Inclusions

A *rough inclusion* is a generic term introduced in [16] for a class of relations on the universe U; any rough inclusion μ is a ternary relation, a subset of the product $U \times U \times [0,1]$; see [11], [12], [13], [14], for details and discussion along with the extensive reference list.

A *rough inclusion* $\mu_\pi(x,y,r)$, where x, y are individual objects, $r \in [0,1]$, does satisfy the following requirements, relative to a given part relation π on a set U of individual objects,

$$\begin{array}{c} 1.\ \mu_\pi(x,y,1) \Leftrightarrow x\ ing_\pi\ y; \\ 2.\ \mu_\pi(x,y,1) \Rightarrow [\mu_\pi(z,x,r) \Rightarrow \mu_\pi(z,y,r)]; \\ 3.\ \mu_\pi(x,y,r) \wedge s < r \Rightarrow \mu_\pi(x,y,s). \end{array} \tag{5}$$

2.3 Examples of Rough Inclusions

Apart from a general theory, we give here some examples of rough inclusions, cf. [11], [13], [14].

1. **Rough inclusions from Archimedean t–norms.** They are induced from Archimedean t–norms, see, e.g, [3], [10]. We describe the one we are going to use in the sequel. The Łukasiewicz t–norm

$$L(x, y) = max\{0, x + y - 1\}, \tag{6}$$

admits a characterization,

$$L(x, y) = g(f(x) + f(y)), \tag{7}$$

with $f = 1 - x = g$. We define the set, $DIS(u, v) = \{a \in A : a(u) \neq a(v)\}$, and its complement $IND(u, v) = U \times U \setminus DIS(u, v)$.

We define the rough inclusion μ_L,

$$\mu_L(u, v, r) \Leftrightarrow g(\frac{|DIS(u, v)|}{|A|}) \geq, \tag{8}$$

i.e.,

$$\mu_L(u, v, r) \Leftrightarrow \frac{|IND(u, v)|}{|A|} \geq r. \tag{9}$$

The formula (9) witnesses that the reasoning based on the rough inclusion μ_L is the probabilistic one. At the same time, we have given a logical proof for formulas like (9) that are very frequently applied in Data Mining and Knowledge Discovery.

μ_L is *transitive* [11]: $\mu_L(u, v, r)$ *and* $\mu_L(v, w, s)$ *imply that* $\mu_L(u, w, L(r, s))$.

2. **Rough inclusions and metrics.** For a metric $d(u, v)$ on the set of objects U, i.e., 1. $d(u, u) = 0$; 2. $d(u, v) = d(v, u)$; 3. $d(u, v) \leq d(u, w) + d(w, v)$, we let $\mu_d(u, v, r) \Leftrightarrow d(u, v) \leq 1 - r$. Then, μ_d *is a rough inclusion*, transitive with respect to the t–norm L.

Conversely, consider a transitive symmetric rough inclusion μ_T; let $d_\mu(u, v) \leq r \Leftrightarrow \mu(u, v, 1-r)$. Then, clearly, $d_\mu(u, u) = 0$, $d_\mu(u, v) = d_\mu(v, u)$; concerning triangle inequality 3., if $d_\mu(u, v) \leq r$ and $d_\mu(v, w) \leq s$, then by transitivity of μ, $d_\mu(u, w) \leq 1 - T(1 - r, 1 - s) = S_T(r, s)$, where S_T is the t–*conorm*, induced by T, see,e.g., [10]; thus, d_μ is a generalized metric. Particular cases encompass: in case of $T = min$, $S_T = max$, hence $d_{min}(u, w) \leq max\{d_{min}(u, v), d_{min}(v, w)\}$, i.e., d_{min} is an Archimedean metric; in case of L, $S_L(r, s) = min\{1, r+s\} \leq r+s$, i.e., d_L is a metric satisfying 3., restricted by 1.

2.4 Granules Induced from Rough Inclusions

The general scheme of our own for inducing granules is as follows. We fix an information system (U, A), and a rough inclusion μ on U.

For an object u and a real number $r \in [0, 1]$, we define the granule $g_\mu(u, r)$ about u of the radius r, relative to μ, by letting,

$$g_\mu(u, r) \text{ is } ClsF(u, r), \tag{10}$$

where the property $F(u,r)$ is satisfied with an object v if and only if $\mu(v,u,r)$ holds.

It was shown, see [11], Theorem 4, that in case of a transitive μ,

$$v\ ing\ g_\mu(u,r) \Leftrightarrow \mu(v,u,r). \quad (11)$$

By (11), the granule $g_{\mu_L}(u,r)$ consists of objects v such that $\mu_L(v,u,r)$, i.e, $|IND(u,v)| \geq r \cdot |A|$;

For a given granulation radius r, and the rough inclusion μ_L, we form the collection $U^G_{r,\mu_L} = \{g_{\mu_L}(u,r)\}$.

3 Granular Decision Systems

The idea of a granular decision system was posed in [13]; for a given information system (U,A), a rough inclusion μ, and $r \in [0,1]$, the new universe $U^G_{r,\mu}$ is given. We apply a strategy $\mathcal{G}$ to choose a covering $Cov^G_{r,\mu}$ of the universe U by granules from $U^G_{r,\mu}$.

We apply a strategy $\mathcal{S}$ in order to assign the value $a^*(g)$ of each attribute $a \in A$ to each granule $g \in Cov^G_{r,\mu}$: $a^*(g) = \mathcal{S}(\{a(u) : u \in g\})$. The granular counterpart to the information system (U,A) is a tuple $(U^G_{r,\mu}, \mathcal{G}, \mathcal{S}, \{a* : a \in A\})$; analogously, we define granular counterparts to decision systems by adding the factored decision $d*$.

4 Factoring Classifiers Through Granular Systems

As objects in a granule are related one to another by similarity, the granule as a whole should determine a new object; and a judiciously chosen set of the new objects should guarantee the satisfactory quality of classification [13]. To test the validity of this hypothesis, experiments have been carried out with real data sets. We select here the Primary tumor data set [21] and we test it with exhaustive algorithm of RSES package [19] and with LEM2 algorithm with the parameter p=0.5 [2], [19]. We adopt random choice as the strategy $\mathcal{G}$, majority voting with random resolution of ties as $\mathcal{S}$, and train–and–test at ratio 1:1 as the method of test performing. Quality of classification is measured by *total accuracy* being the ratio of the number of correctly classified cases to the number of recognized cases, and *total coverage*, i.e, the ratio of the number of recognized test cases to the number of test cases. Results are given in Table 1. *nil* denotes results without granulation to which granular results are compared.

The procedure has been as follows.

1. the data table (U,A) has been input;
2. classification rules have been found on the training subtable of 50 percent of objects by means of each of the three algorithms;
3. classification of dataset objects in the test subtable of remaining 50 percent of objects has been found for each of the three classifications found at point 2;

4. given the granule radius, granules of that radius have been found on the training subtable;
5. a granular covering of the training subtable has been chosen;
6. the corresponding granular decision system has been determined;
7. granular classifiers have been induced from the granular system in point 6 by means of each of algorithms in point 2;
8. classifications of objects in the test subtable have been found by means of each of classifiers in point 7;
9. classifications from points 3,8 have been compared with respect to adopted global measures of quality: total accuracy and total covering.

Table 1. Primary tumor dataset:r=granule radius,tst=test sample size,trn=training sample size,rulex=number of rules with exhaustive algorithm, rullem=number of rules with LEM2, aex=total accuracy with exhaustive algorithm,cex=total coverage with exhaustive algorithm,alem=total accuracy with LEM2, clem=total coverage with LEM2

r	*tst*	*trn*	*rulex*	*rullem*	*aex*	*cex*	*alem*	*clem*
nil	170	169	4186	43	0.253	0.976	0.5	0.259
0.0	170	1	0	0	0.0	0.0	0.0	0.0
0.0588235	170	1	0	0	0.0	0.0	0.0	0.0
0.117647	170	1	0	0	0.0	0.0	0.0	0.0
0.176471	170	1	0	0	0.0	0.0	0.0	0.0
0.235294	170	1	0	0	0.0	0.0	0.0	0.0
0.294118	170	1	0	0	0.0	0.0	0.0	0.08
0.352941	170	1	0	0	0.0	0.0	0.0	0.0
0.411765	170	2	0	1	0.0	0.0	1.0	0.188
0.470588	170	3	0	1	0.0	0.0	1.0	0.188
0.529412	170	5	0	1	0.0	0.0	1.0	0.188
0.588235	170	8	0	1	0.0	0.0	1.0	0.188
0.647059	170	12	11	1	0.547	0.376	0.0	0.0
0.705882	170	17	40	3	0.457	0.476	0.667	0.035
0.764706	170	33	108	4	0.468	0.553	0.769	0.076
0.823529	170	54	1026	11	0.434	0.759	0.586	0.171
0.882353	170	75	3640	17	0.308	0.859	0.579	0.224
0.941176	170	107	4428	24	0.295	0.976	0.466	0.341

Conclusions for Primary tumor. For exhaustive algorithm,accuracy is better with granular than original training set from the radius of 0.647059 on where reduction in size of training set is 92.9 percent and reduction in size of rule set is almost 100 percent (11 versus 4186). Coverage falls within error bound of 22.3 percent from the radius of 0.823529 on, where reduction in training st size is 68.2 percent and reduction in size of rule set is 75.5 percent; it becomes the same as in non–granular case at $r = .941$ with reduction in object size of 36.7 percent.

LEM2 exceeds accuracy of classifier trained on original training table with accuracy of granular classifier from the radius of 0.705882 on where reduction in training set size is 89.95 percent and reduction in rule set size is 93 percent. Coverage for granular classifier is better or within error of 13.5 percent from the radius of 0.882353 where reduction in size of the training set is 55.6 percent and reduction in size of rule set is 60.5 percent.

Thus, granular approach provides results on par with those obtained in non–granular case.

5 A Granular Approach to Missing Values

An information/decision system is *incomplete* in case some values of conditional attributes from A are not known. Analysis of systems with missing values requires a decision on how to treat missing values; Grzymala–Busse in his work [2], analyzes nine such methods, among them, *4. assigning all possible values to the missing location*, *9. treating the unknown value as a new valid value*, etc. etc. Results in [2] indicate that methods *4,9* perform very well among all nine methods. In this work we consider and adopt two methods, i.e.*4, 9*. Analysis of this problem has been given also in Kryszkiewicz [4] and Kryszkiewicz–Rybinski [5].

We will use the symbol $*$ commonly used for denoting the missing value; we will use two methods *4, 9* for treating $*$, i.e, either $*$ is a *don't care* symbol meaning that any value of the respective attribute can be substituted for $*$, thus $* = v$ for each value v of the attribute, or $*$ is a new value on its own, i.e., if $* = v$ then v can be only $*$.

Table 2. Strategy A. CV–5; Hepatitis; exhaustive algorithm. r=radius,macc=mean accuracy, mcov=mean coverage, mrul=mean number of rules, mtrn=mean training granular sample size

r	*macc*	*mcov*	*mrul*	*mtrn*
0.0	0.0	0.0	0.0	1.0
0.0526316	0.0	0.0	0.0	1.0
0.105263	0.0	0.0	0.0	1.0
0.157895	0.0	0.0	0.0	1.0
0.210526	0.0	0.0	0.0	1.0
0.263158	0.0	0.0	0.0	1.4
0.315789	0.0	0.0	0.0	2.0
0.368421	0.0	0.0	0.0	2.4
0.421053	0.0	0.0	0.0	3.8
0.473684	0.2012	0.3548	6.4	3.4
0.526316	0.5934	1.0	29.6	7.4
0.578947	0.4992	0.7872	33.8	7.6
0.631579	0.5694	0.9872	176.2	20.0
0.684211	0.5852	0.9936	167.6	17.8
0.736842	0.6102	0.9936	263.0	22.8
0.789474	0.6130	1.0	911.0	49.4
0.842105	0.6258	1.0	989.6	46.8
0.894737	0.6386	1.0	1899.0	77.0
0.947368	0.6774	1.0	2836.2	105.8
1.0	0.6710	1.0	3286.4	123.4

Table 3. Strategy B. CV–5; Hepatitis; exhaustive algorithm. r=radius,macc=mean accuracy, mcov=mean coverage, mrul=mean number of rules, mtrn=mean training granular sample size

r	$macc$	$mcov$	$mrul$	$mtrn$
0.0	0.0	0.0	0.0	1.0
0.0526316	0.0	0.0	0.0	1.0
0.105263	0.0	0.0	0.0	1.0
0.157895	0.0	0.0	0.0	1.0
0.210526	0.0	0.0	0.0	1.2
0.263158	0.0	0.0	0.0	1.2
0.315789	0.0	0.0	0.0	1.6
0.368421	0.1104	0.1870	1.0	2.6
0.421053	0.0904	0.2000	1.6	3.4
0.473684	0.3938	0.5806	7.2	4.4
0.526316	0.4234	0.7936	26.2	7.6
0.578947	0.6302	0.9936	59.4	10.8
0.631579	0.6708	1.0	126.4	15.4
0.684211	0.6038	0.9742	253.4	24.4
0.736842	0.6292	0.9936	367.6	35.2
0.789474	0.6166	0.9936	947.0	52.2
0.842105	0.6324	1.0	1417.2	71.8
0.894737	0.6386	1.0	1797.0	79.6
0.947368	0.6450	1.0	3081.8	113.4
1.0	0.6646	1.0	3354.2	123.4

Our procedure for treating missing values is based on the granular structure $(U^G_{r,\mu}, \mathcal{G}, \mathcal{S}, \{a* : a \in A\})$; the strategy $\mathcal{S}$ is the majority voting, i.e., for each attribute a, the value $a^*(g)$ is the most frequent of values in $\{a(u) : u \in g\}$. The strategy $\mathcal{G}$ consists in random selection of granules for a covering.

For an object u with the value of $*$ at an attribute a, and a granule $g = g(v, r) \in U^G_{r,\mu}$, the question whether u is included in g is resolved according to the adopted strategy of treating $*$: in case $* = don't\ care$, the value of $*$ is regarded as identical with any value of a hence $|IND(u, v)|$ is automatically increased by 1, which increases the granule; in case $* = *$, the granule size is decreased. Assuming that $*$ is sparse in data, majority voting on g would produce values of a^* distinct from $*$ in most cases; nevertheless the value of $*$ may appear in new objects g^*, and then in the process of classification, such value is repaired by means of the granule closest to g^* with respect to the rough inclusion μ_L, in accordance with the chosen method for treating $*$.

In plain words, objects with missing values are in a sense absorbed by close to them granules and missing values are replaced with most frequent values in objects collected in the granule; in this way the method *3* or *4* in [2] is combined with the idea of a frequent value, in a novel way.

We have thus four possible strategies:

– Strategy A: in building granules $*$=*don't care*, in repairing values of $*$, $*$=*don't care*;

– Strategy B: in building granules $*$=*don't care*, in repairing values of $*$, $* = *$;
– Strategy C: in building granules $* = *$, in repairing values of $*$, $*$=*don't care*;
– Strategy D: in building granules $* = *$, in repairing values of $*$, $* = *$.

5.1 Results of Test with Real Data Set Hepatitis with Missing Values

We record here results of tests with Hepatitis data set [21] with 155 objects, 20 attributes and 167 missing values. We apply the exhaustive algorithm of RSES system [19] and 5–fold cross–validation (CV–5). Below we give averaged results for strategies A, B, C, and D. As before, radius *nil* indicates non–granulated case.

Now, we record in Tables 2–5 the results of classification for Hepatitis with exhaustive algorithm and CV–5 cross–validation for strategies A, B, C, D, respectively.

For comparison, we include results of tests with Hepatitis recorded in [1]; the method was modified LERS algorithm with additional parameters like strength and specificity of a rule and the approach 9. gave error rate of 0.1935 i.e. accuracy 0.8065. Best result given by strategy C based on the same treatment of $*$ is accuracy 0.6838. Naive LERS algorithm [1] gave for this data set and method 9 error of 0.3484 i.e. accuracy of 0.6516. Interestingly, granular method gives better than [1] results for Breast cancer data set, as reported in [15], these Proceedings.

Table 4. Strategy C. CV–5; Hepatitis; exhaustive algorithm. r=radius,macc=mean accuracy, mcov=mean coverage, mrul=mean number of rules, mtrn=mean training granular sample size

r	*macc*	*mcov*	*mrul*	*mtrn*
0.0	0.0	0.0	0.0	1.0
0.0526316	0.0	0.0	0.0	1.0
0.105263	0.0	0.0	0.0	1.2
0.157895	0.0	0.0	0.0	1.2
0.210526	0.0	0.0	0.0	1.8
0.263158	0.0	0.0	0.0	2.0
0.315789	0.2560	0.3936	2.4	4.0
0.368421	0.4486	0.6838	7.4	5.6
0.421053	0.4766	0.7870	19.2	7.8
0.473684	0.5806	1.0	58.4	10.6
0.526316	0.6580	1.0	136.6	17.4
0.578947	0.64902	0.9936	332.4	32.0
0.631579	0.6568	0.9936	991.6	47.4
0.684211	0.6646	1.0	1751.6	70.2
0.736842	0.6902	1.0	2648.8	93.2
0.789474	0.6322	1.0	3208.8	112.6
0.842105	0.6776	1.0	3297.8	120.2
0.894737	0.6710	1.0	3297.4	123.4
0.947368	0.6838	1.0	3305.4	124.0
1.0	0.6774	1.0	3327.2	124.0

Table 5. Strategy D. CV–5; Hepatitis; exhaustive algorithm. r=radius,macc=mean accuracy, mcov=mean coverage, mrul=mean number of rules, mtrn=mean training granular sample size

r	$macc$	$mcov$	$mrul$	$mtrn$
0.0	0.0	0.0	0.0	1.0
0.0526316	0.0	0.0	0.0	1.0
0.105263	0.0	0.0	0.0	1.0
0.157895	0.0	0.0	0.0	1.4
0.210526	0.0	0.0	0.0	1.6
0.263158	0.0	0.0	0.0	2.6
0.315789	0.3886	0.5162	6.0	3.8
0.368421	0.5730	0.9032	16.6	4.8
0.421053	0.6328	0.9418	23.8	6.8
0.473684	0.5740	0.9740	60.6	10.6
0.526316	0.6170	0.9936	120.6	16.8
0.578947	0.6888	0.9936	354.0	30.6
0.631579	0.6388	1.0	922.0	47.4
0.684211	0.6646	1.0	1828.6	70.8
0.736842	0.6450	1.0	2648.2	93.4
0.789474	0.6516	1.0	3182.0	112.4
0.842105	0.6710	1.0	3299.2	120.4
0.894737	0.6710	1.0	3333.8	123.4
0.947368	0.6646	1.0	3327.2	124.0
1.0	0.6710	1.0	3338.6	124.0

Conclusions for Hepatitis data set. Results for particular strategies compared radius by radius show that the ranking of strategies is $C > D > B > A$ with the average number of ranks respectively 1.3, 1.8., 3.1, 3.8; thus, the strategy C is most effective with D giving slightly worse results. Results by our granular approach are midway between results for naive and new LERS in [1] showing the potential of the method as well as the need for further development.

6 Conclusion

The results of tests reported in this work bear out the hypothesis that granulated data sets preserve information allowing for satisfactory classification. Also the novel approach to the problem of data with missing values has proved to be very effective. Further studies will lead to novel algorithms for rule induction based on granules of knowledge.

Acknowledgement

The author wishes to express his thanks to Professors A. Skowron and J. W. Grzymala–Busse who provided in the system RSES rule inducing algorithms used in this work.

References

1. Grzymala–Busse, J.W., Hu, M.: A comparison of several approaches to missing attribute values in Data Mining. Lecture Notes in Artificial intelligence, vol. 2005, pp. 378–385. Springer, Berlin (2000)
2. Grzymala–Busse, J.W.: Data with missing attribute values: Generalization of rule indiscernibility relation and rule induction. In: Transactions on Rough Sets I, pp. 78–95. Springer, Berlin (2004)
3. Hájek, P.: Metamathematics of Fuzzy Logic. Kluwer, Dordrecht (1998)
4. Kryszkiewicz, M.: Rules in incomplete information systems. Information Sciences 113, 271–292 (1999)
5. Kryszkiewicz, M., Rybiński, H.: Data mining in incomplete information systems from rough set perspective. In: Polkowski, L., Tsumoto, S., Lin, T.Y. (eds.) Rough Set Methods and Applications, pp. 568–580. Physica Verlag, Heidelberg (2000)
6. Leśniewski, S.: On the foundations of set theory. Topoi 2, 7–52 (1982)
7. Lin, T.Y.: Granular computing: Examples, intuitions, and modeling, in: [17], pp. 40–44
8. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer, Dordrecht (1991)
9. Pawlak, Z.: Rough sets. Int. J. Computer and Information Sci. 11, 341–356 (1982)
10. Polkowski, L.: Rough Sets. Mathematical Foundations. Physica Verlag, Heidelberg (2002)
11. Polkowski, L.: Toward rough set foundations. Mereological approach (a plenary lecture). In: Tsumoto, S., Słowiński, R., Komorowski, J., Grzymała-Busse, J.W. (eds.) RSCTC 2004. LNCS (LNAI), vol. 3066, pp. 8–25. Springer, Berlin (2004)
12. Polkowski, L.: Rough–fuzzy–neurocomputing based on rough mereological calculus of granules. Intern. J. Hybrid Intell. Systems 2, 91–108 (2005)
13. Polkowski, L.: Formal granular calculi based on rough inclusions (a feature talk). In: [17], pp. 57–62
14. Polkowski, L.: A model of granular computing with applications (a feature talk). In: [18], 9–16.
15. Polkowski, L., Artiemjew, P.: On granular rough computing with missing values. In: these Proceedings
16. Polkowski, L., Skowron, A.: Rough mereology: a new paradigm for approximate reasoning. International Journal of Approximate Reasoning 15(4), 333–365 (1997)
17. In: Proceedings of IEEE 2005 Conference on Granular Computing,GrC05, Beijing, China, July 2005, IEEE Press, New York (2005)
18. In: Proceedings of IEEE 2006 Conference on Granular Computing, GrC06, Atlanta, USA, May 2006, IEEE Press, New York (2006)
19. A. Skowron et al., RSES: A system for data analysis; available at http: logic.mimuw.edu.pl/ rses/
20. Skowron, A., Stepaniuk, J.: Information granules and rough–neural computing. In: Pal, S.K., Polkowski, L., Skowron, A. (eds.) Rough Neural Computing, pp. 43–84. Springer, Berlin (2004)
21. `http://www.ics.uci.edu./~mlearn/databases/`
22. Yao, Y.Y.: Perspectives of granular computing. In: [17], pp. 85–90
23. Zadeh, L.A.: Fuzzy sets and information granularity. In: Gupta, M., Ragade, R., Yager, R. (eds.): Advances in Fuzzy Set Theory and Applications. North–Holland, Amsterdam, pp. 3–18 (1979)