A Rough Set Approach to Incomplete Data

Jerzy W. Grzymala-Busse[1,2(✉)]

[1] Department of Electrical Engineering and Computer Science,
University of Kansas, Lawrence, KS 66045, USA
[2] Department of Expert Systems and Artificial Intelligence,
University of Information Technology and Management,
35-225 Rzeszow, Poland
`jerzy@ku.edu`

Abstract. This paper presents main directions of research on a rough set approach to incomplete data. First, three different types of lower and upper approximations, based on the characteristic relation, are defined. Then an idea of the probabilistic approximation, an extension of lower and upper approximations, is presented. Local probabilistic approximations are also discussed. Finally, some special topics such as consistency of incomplete data and a problem of increasing data set incompleteness to improve rule set quality, in terms of an error rate, are discussed.

Keywords: Incomplete data · Characteristic relation · Singleton concept and subset approximations · Probabilistic approximations · Local probabilistic approximations

1 Introduction

It is well-known that many real-life data sets are incomplete, i.e., are affected by missing attribute values. Recently many papers presenting a rough set approach in research on incomplete data were published, see, e.g., [4,7,9–17,21–27,31–34,37,38,40–42,44,47–62,68–72,74–78,80,81].

Most of the rough set activity in research on incomplete data is conducted in data mining. Using a rough set approach to incomplete data, we may distinguish between different interpretations of missing attribute values.

If an attribute value was accidentally erased or is unreadable, we may use the most cautious approach to missing attribute values and mine data using only specified attribute values. The corresponding type of missing attribute values is called *lost* and is denoted by "?". Mining incomplete data affected by lost values was studied for the first time in [44], where two algorithms for rule induction from such data were presented. The same data sets were studied later, see, e.g., [76,77].

Another type of missing attribute values happens when a respondent refuses to answer a question that seems to be irrelevant. For example, a patient is tested for a disease and one of the questions is a color of hair. The respondent may consider the color of hair to be irrelevant. This type of missing attribute values is called a *"do not care" condition* and is denoted by "*". The first study of "do

not care" conditions, again using rough set theory, was presented in [17], where a method for rule induction in which missing attribute values were replaced by all values from the domain of the attribute was introduced. "Do not care" conditions were also studied later, see, e.g. [50,51].

In yet another interpretation of missing attribute values, called an *attribute-concept* value, and denoted by "−", we assume that we know that the corresponding case belongs to a specific concept X and, as a result, we may replace the missing attribute value by attribute values for all cases from the same concept X. A *concept* (class) is a set of all cases classified (or diagnosed) the same way. For example, if for a patient the value of an attribute *Temperature* is missing, this patient is sick with *Flu*, and all remaining patients sick with *Flu* have *Temperature* values *high* then using the interpretation of the missing attribute value as the attribute-concept value, we will replace the missing attribute value by the value *high*. This approach was introduced in [24].

An approach to mining incomplete data presented in this paper is based on the idea of an attribute-value block. A characteristic set, defined as an intersection of such blocks, is a generalization of the elementary set, well-known in rough set theory [63–65]. A characteristic relation, defined by characteristic sets, is, in turn, a generalization of the indiscernibility relation. As it was shown in [21], incomplete data are described by three different types of approximations: singleton, subset and concept.

For rule induction from incomplete data it is the most natural to use the MLEM2 (Modified Learning form Examples Module, version 2) [2,18–20] since this algorithm is also based on attribute-value pair blocks. A number of extensions of this algorithm were developed in order to process incomplete data sets using different definitions of approximations, see, e.g., [5,31,43,45].

One of the fundamental concepts of rough set theory is lower and upper approximations. A generalization of such approximations, a probabilistic approximation, introduced in [79], was applied in variable precision rough set models, Bayesian rough sets and decision-theoretic rough set models [46,66,67,73,82–86]. These probabilistic approximations are defined using the indiscernibility relation. For incomplete data, probabilistic approximations were extended to characteristic relation in [30]. The probabilistic approximation is associated with some parameter α (interpreted as a probability). If α is very small, say $1/|U|$, where U is the set of all cases, the probabilistic approximation is reduced to the upper approximation; if α is equal to 1.0, the probabilistic approximation is equal to the lower approximation. Local probabilistic approximations, based on attribute-value blocks instead of characteristic sets, were defined in [7], see also [31].

2   Fundamental Concepts

A basic tool to analyze incomplete data sets is a *block of an attribute-value pair*. Let (a, v) be an attribute-value pair. For *complete* data sets, i.e., data sets in which every attribute value is specified, a block of (a, v), denoted by $[(a, v)]$, is the set of all cases x for which $a(x) = v$, where $a(x)$ denotes the value of the

attribute a for the case x. For incomplete data sets the definition of a block of an attribute-value pair is modified.

- If for an attribute a there exists a case x such that $a(x) = ?$, i.e., the corresponding value is lost, then the case x should not be included in any blocks $[(a, v)]$ for all values v of attribute a,
- If for an attribute a there exists a case x such that the corresponding value is a "do not care" condition, i.e., $a(x) = *$, then the case x should be included in blocks $[(a, v)]$ for all specified values v of attribute a.
- If for an attribute a there exists a case x such that the corresponding value is an attribute-concept value, i.e., $a(x) = -$, then the corresponding case x should be included in blocks $[(a, v)]$ for all specified values $v \in V(x, a)$ of attribute a, where

$$V(x, a) = \{a(y) \mid a(y) \text{ is specified}, y \in U, \ d(y) = d(x)\}.$$

For a case $x \in U$ the *characteristic set* $K_B(x)$ is defined as the intersection of the sets $K(x, a)$, for all $a \in B$, where the set $K(x, a)$ is defined in the following way:

- If $a(x)$ is specified, then $K(x, a)$ is the block $[(a, a(x)]$ of attribute a and its value $a(x)$,
- If $a(x)) = ?$ or $a(x) = *$, then the set $K(x, a) = U$.
- If $a(x) = -$, then the corresponding case x should be included in blocks $[(a, v)]$ for all known values $v \in V(x, a)$ of attribute a. If $V(x, a)$ is empty, $K(x, a) = U$.

The *characteristic relation* $R(B)$ is a relation on U defined for $x, y \in U$ as follows

$$(x, y) \in R(B) \text{ if and only if } y \in K_B(x).$$

The characteristic relation $R(B)$ is reflexive but—in general—does not need to be symmetric or transitive.

3   Lower and Upper Approximations

For incomplete data sets there is a few possible ways to define approximations [24, 43]. Let X be a concept, let B be a subset of the set A of all attributes, and let $R(B)$ be the characteristic relation of the incomplete decision table with characteristic sets $K_B(x)$, where $x \in U$. A *singleton B-lower approximation* of X is defined as follows:

$$\underline{B}X = \{x \in U \mid K_B(x) \subseteq X\}.$$

A *singleton B-upper approximation* of X is

$$\overline{B}X = \{x \in U \mid K_B(x) \cap X \neq \emptyset\}.$$

We may define lower and upper approximations for incomplete data sets as unions of characteristic sets. There are two possibilities. Using the first way, a *subset B*-lower approximation of X is defined as follows:

$$\underline{B}X = \cup\{K_B(x) \mid x \in U, K_B(x) \subseteq X\}.$$

A *subset B*-upper approximation of X is

$$\overline{B}X = \cup\{K_B(x) \mid x \in U, K_B(x) \cap X \neq \emptyset\}.$$

The second possibility is to modify the subset definition of lower and upper approximation by replacing the universe U from the subset definition by a concept X. A *concept B*-lower approximation of the concept X is defined as follows:

$$\underline{B}X = \cup\{K_B(x) \mid x \in X, K_B(x) \subseteq X\}.$$

The subset B-lower approximation of X is the same set as the concept B-lower approximation of X. A *concept B*-upper approximation of the concept X is defined as follows:

$$\overline{B}X = \cup\{K_B(x) \mid x \in X, K_B(x) \cap X \neq \emptyset\} = \cup\{K_B(x) \mid x \in X\}.$$

Two traditional methods of handling missing attribute values: Most Common Value for symbolic attributes and Average Values for numeric attributes (MCV-AV) and Concept Most Common Values for symbolic attributes and Concept Average Values for numeric attributes (CMCV-CAV), for details see [36], were compared with three rough-set interpretations of missing attribute values: lost values, attribute-concept values and "do not care" conditions combined with concept lower and upper approximations in [27]. In turned out that there is no significant difference in performance in terms of an error rate measured by ten-fold cross validation between the traditional and rough-set approaches to missing attribute values.

In [26] the same two traditional methods, MCV-AV and CMCV-CAV, and other two traditional methods (Closest Fit and Concept Closest Fit), for details see [36], were compared with the same three rough set interpretations of missing attribute values combined with concept approximations. The best methodology was based on the Concept Closest Fit combined with rough set interpretation of missing attribute values as lost values and concept lower and upper approximations.

Additionally, in [28], a CART approach to missing attribute values [1] was compared with missing attribute values interpreted as lost values combined with concept lower and upper approximations. In two cases CART was better, in two cases rough set approach was better, and in one case the difference was insignificant. Hence both approaches are comparable in terms of an error rate.

In [29,39], the method CMCV-CAV was compared with rough set approaches to missing attribute values, the conclusion was that CMCV-CAV was either worse or not better, depending on a data set, than rough-set approaches.

4 Probabilistic Approximations

In this section we will extend definitions of singleton, subset and concept approximations to corresponding probabilistic approximations. The problem is how useful are *proper* probabilistic approximations (with α larger than $1/|U|$ but smaller than 1.0). We studied usefulness of proper probabilistic approximations for incomplete data sets [3], where we concluded that proper probabilistic approximations are not frequently better than ordinary lower and upper approximations.

A B-singleton probabilistic approximation of X with the threshold α, $0 < \alpha \leq 1$, denoted by $appr_{\alpha,B}^{singleton}(X)$, is defined by

$$\{x \mid x \in U, \ Pr(X \mid K_B(x)) \geq \alpha\},$$

where $Pr(X \mid K_B(x)) = \frac{|X \ \cap \ K_B(x)|}{|K_B(x)|}$ is the conditional probability of X given $K_B(x)$ and $|Y|$ denotes the cardinality of set Y.

A B-subset probabilistic approximation of the set X with the threshold α, $0 < \alpha \leq 1$, denoted by $appr_{\alpha,B}^{subset}(X)$, is defined by

$$\cup\{K_B(x) \mid x \in U, \ Pr(X \mid K_B(x)) \geq \alpha\}$$

A B-concept probabilistic approximation of the set X with the threshold α, $0 < \alpha \leq 1$, denoted by $appr_{\alpha,B}^{concept}(X)$, is defined by

$$\cup\{K_B(x) \mid x \in X, \ Pr(X \mid K_B(x)) \geq \alpha\}.$$

In [6] *ordinary* lower and upper approximations (singleton, subset and concept), special cases of singleton, subset and concept probabilistic approximations, were compared with *proper* probabilistic approximations (singleton, subset and concept) on six data sets with missing attribute values interpreted as lost values and "do not care" conditions, in terms of an error rate. Since we used six data sets, two interpretations of missing attribute values and three types of probabilistic approximations, there were 36 combinations. Among these 36 combinations, for five combinations the error rate was smaller for proper probabilistic approximations than for ordinary (lower and upper) approximations, for other four combinations, the error rate for proper probabilistic approximations was larger than for ordinary approximations, for the remaining 27 combinations, the difference between these two types of approximations was not statistically significant.

Results of experiments presented in [9] show that among all probabilistic approximations (singleton, subset and concept) and two interpretations of missing attribute values (lost values and "do not care" conditions) there is not much difference in terms of an error rate measured by ten-fold cross validation. On the other hand, complexity of induced rule sets differs significantly. The simplest rule sets (in terms of the number of rules and the total number of conditions in the rule set) we accomplished by using subset probabilistic approximations combined with "do not care" conditions.

In [8] results of experiments using all three probabilistic approximations (singleton, subset and concept) and two interpretations of missing attribute values: lost values and "do not care" conditions were compared with the MCV-AV and CMCV-CAV methods in terms of an error rate. For every data set, the best among six rough-set methods (combining three kinds of probabilistic approximations with two types of interpretations of missing attribute vales) were compared with the better results of MCV-AV and CMCV-CAV. Rough-set methods were better for five (out of six) data sets.

5   Local Approximations

An idea of the local approximation was introduced in [41]. A local probabilistic approximation was defined in [7]. A set T of attribute-value pairs, where all attributes belong to the set B and are distinct, is called a *B-complex*. In the most general definition of a local probabilistic definition we assume only an existence of a family $\mathcal{T}$ of B-complexes T with the conditional probability $Pr(X|[T])$ of $X \geq \alpha$, where $Pr(X|[T]) = \frac{|X \cap [T]|}{|[T]|}$.

A *B-local probabilistic approximation* of the set X with the parameter α, $0 < \alpha \leq 1$, denoted by $appr_\alpha^{local}(X)$, is defined as follows

$$\cup\{[T] \mid \exists \ a \ family \ \mathcal{T} \ of \ B - complexes \ T \ of \ X \ with \ \forall \ T \in \mathcal{T}, \ Pr(X|[T]) \geq \alpha\}.$$

In general, for given set X and α, there exists more than one A-local probabilistic approximation. However, for given set X and parameter α, a B-local probabilistic approximation given by the next definition is unique.

A *complete B-local probabilistic approximation* of the set X with the parameter α, $0 < \alpha \leq 1$, denoted by $appr_\alpha^{complete}(X)$, is defined as follows

$$\cup\{[T] \mid T \ is \ a \ B - complex \ of \ X, \ Pr(X|[T]) \geq \alpha\}.$$

Due to computational complexity of determining complete local probabilistic approximations, yet another local probabilistic approximation, called a *MLEM2 local probabilistic approximation* and denoted by $appr_\alpha^{mlem2}(X)$, is defined by using A-complexes Y that are the most relevant to X, i.e., with $|X \cap Y|$ as large as possible, etc., following the MLEM2 algorithm.

In [31] concept probabilistic approximations were compared with complete local probabilistic approximations and with MLEM2 local probabilistic approximations for eight data sets, using two interpretations of missing attribute values: lost vales and "do not care" conditions, in terms of an error rate. Since two interpretations of missing attribute values and eight data sets were used, there were 16 combinations. There was not clear winner among three kinds of probabilistic approximations. In four combinations the best was the concept probabilistic approximation, in three cases the best was the complete local probabilistic approximation, and in four cases the best was the MLEM2 local probabilistic approximation. For remaining five combinations the difference in performance between all three approximations was insignificant.

6   Special Topics

When replacing existing, specified attribute values by symbols of missing attribute values, e.g., by "?"s, an error rate computed by ten-fold cross validation may be smaller than for the original, complete data set. Thus, increasing incompleteness of the data set may improve accuracy. Results of experiments showing this phenomenon were published, e.g., in [34, 35].

Yet another problem is associated with consistency. For complete data sets, a data set is consistent if any two cases, indistinguishable by all attributes, belong to the same concept. The idea of consistency is more complicated for incomplete data. This problem was discussed in [4] and also in [54, 60, 72].

7   Conclusions

Research on incomplete data is very active and promising, with many open problems and potential for additional progress.

References

1. Breiman, L., Friedman, J.H., Olshen, R.A., Stone, C.J.: Classification and Regression Trees. Wadsworth & Brooks, Monterey (1984)
2. Chan, C.C., Grzymala-Busse, J.W.: On the attribute redundancy and the learning programs ID3, PRISM, and LEM2. Technical report, Department of Computer Science, University of Kansas (1991)
3. Clark, P.G., Grzymala-Busse, J.W.: Experiments on probabilistic approximations. In: Proceedings of the 2011 IEEE International Conference on Granular Computing, pp. 144–149 (2011)
4. Clark, P.G., Grzymala-Busse, J.W.: Consistency of incomplete data. In: Proceedings of the Second International Conference on Data Technologies and Applications, pp. 80–87 (2013)
5. Clark, P.G., Grzymala-Busse, J.W.: A comparison of two versions of the MLEM2 rule induction algorithm extended to probabilistic approximations. In: Cornelis, C., Kryszkiewicz, M., Ślęzak, D., Ruiz, E.M., Bello, R., Shang, L. (eds.) RSCTC 2014. LNCS, vol. 8536, pp. 109–119. Springer, Heidelberg (2014)
6. Clark, P.G., Grzymala-Busse, J.W., Hippe, Z.S.: An analysis of probabilistic approximations for rule induction from incomplete data sets. Fundam. Informaticae **55**, 365–379 (2014)
7. Clark, P.G., Grzymala-Busse, J.W., Kuehnhausen, M.: Local probabilistic approximations for incomplete data. In: Chen, L., Felfernig, A., Liu, J., Raś, Z.W. (eds.) ISMIS 2012. LNCS, vol. 7661, pp. 93–98. Springer, Heidelberg (2012)
8. Clark, P.G., Grzymala-Busse, J.W., Kuehnhausen, M.: Mining incomplete data with many missing attribute values. A comparison of probabilistic and rough set approaches. In: Proceedings of the Second International Conference on Intelligent Systems and Applications, pp. 12–17 (2013)
9. Clark, P.G., Grzymala-Busse, J.W., Rzasa, W.: Mining incomplete data with singleton, subset and concept approximations. Inf. Sci. **280**, 368–384 (2014)

10. Cyran, K.A.: Modified indiscernibility relation in the theory of rough sets with real-valued attributes: application to recognition of fraunhofer diffraction patterns. Trans. Rough Sets **9**, 14–34 (2008)
11. Dai, J.: Rough set approach to incomplete numerical data. Inf. Sci. **241**, 43–57 (2013)
12. Dai, J., Xu, Q.: Approximations and uncertainty measures in incomplete information systems. Inf. Sci. **198**, 62–80 (2012)
13. Dai, J., Xu, Q., Wang, W.: A comparative study on strategies of rule induction for incomplete data based on rough set approach. Int. J. Adv. Comput. Technol. **3**, 176–183 (2011)
14. Dardzinska, A., Ras, Z.W.: Chasing unknown values in incomplete information systems. In: Workshop Notes, Foundations and New Directions of Data Mining, in Conjunction with the 3-rd International Conference on Data Mining, pp. 24–30 (2003)
15. Dardzinska, A., Ras, Z.W.: On rule discovery from incomplete information systems. In: Workshop Notes, Foundations and New Directions of Data Mining, in Conjunction with the 3-rd International Conference on Data Mining, pp. 24–30 (2003)
16. Greco, S., Matarazzo, B., Slowinski, R.: Dealing with missing data in rough set analysis of multi-attribute and multi-criteria decision problems. In: Zanakis, H., Doukidis, G., Zopounidis, Z. (eds.) Decision Making: Recent developments and Worldwide Applications, pp. 295–316. Kluwer Academic Publishers, Dordrecht (2000)
17. Grzymala-Busse, J.W.: On the unknown attribute values in learning from examples. In: Raś, Z.W., Zemankova, M. (eds.) ISMIS 1991. LNCS, vol. 542, pp. 368–377. Springer, Heidelberg (1991)
18. Grzymala-Busse, J.W.: LERS–a system for learning from examples based on rough sets. In: Slowinski, R. (ed.) Intelligent Decision Support, pp. 3–18. Handbook of Applications and Advances of the Rough Set Theory. Kluwer Academic Publishers, Dordrecht (1992)
19. Grzymala-Busse, J.W.: A new version of the rule induction system LERS. Fundamenta Informaticae **31**, 27–39 (1997)
20. Grzymala-Busse, J.W.: MLEM2: A new algorithm for rule induction from imperfect data. In: Proceedings of the 9th International Conference on Information Processing and Management of Uncertainty in Knowledge-Based Systems, pp. 243–250 (2002)
21. Grzymala-Busse, J.W.: Rough set strategies to data with missing attribute values. In: Notes of the Workshop on Foundations and New Directions of Data Mining, in Conjunction with the Third International Conference on Data Mining, pp. 56–63 (2003)
22. Grzymała-Busse, J.W.: Characteristic relations for incomplete data: a generalization of the indiscernibility relation. In: Tsumoto, S., Słowiński, R., Komorowski, J., Grzymała-Busse, J.W. (eds.) RSCTC 2004. LNCS (LNAI), vol. 3066, pp. 244–253. Springer, Heidelberg (2004)
23. Grzymala-Busse, J.W.: Data with missing attribute values: generalization of indiscernibility relation and rule induction. Trans. Rough Sets **1**, 78–95 (2004)
24. Grzymala-Busse, J.W.: Three approaches to missing attribute values–a rough set perspective. In: Proceedings of the Workshop on Foundation of Data Mining, in Conjunction with the Fourth IEEE International Conference on Data Mining, pp. 55–62 (2004)

25. Grzymała-Busse, J.W.: Incomplete data and generalization of indiscernibility relation, definability, and approximations. In: Slezak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFDGrC 2005. LNCS (LNAI), vol. 3641, pp. 244–253. Springer, Heidelberg (2005)
26. Grzymala-Busse, J.W.: A comparison of traditional and rough set approaches to missing attribute values in data mining. In: Proceedings of the 10-th International Conference on Data Mining, Detection, Protection and Security, Royal Mare Village, Crete, pp. 155–163 (2009)
27. Grzymala-Busse, J.W.: Mining data with missing attribute values: A comparison of probabilistic and rough set approaches. In: Proceedings of the 4-th International Conference on Intelligent Systems and Knowledge Engineering, pp. 153–158 (2009)
28. Grzymala-Busse, J.W.: Rough set and CART approaches to mining incomplete data. In: Proceedings of the International Conference on Soft Computing and Pattern Recognition, IEEE Computer Society, pp. 214–219 (2010)
29. Grzymala-Busse, J.W.: A comparison of some rough set approaches to mining symbolic data with missing attribute values. In: Kryszkiewicz, M., Rybinski, H., Skowron, A., Raś, Z.W. (eds.) ISMIS 2011. LNCS, vol. 6804, pp. 52–61. Springer, Heidelberg (2011)
30. Grzymała-Busse, J.W.: Generalized parameterized approximations. In: Yao, J.T., Ramanna, S., Wang, G., Suraj, Z. (eds.) RSKT 2011. LNCS, vol. 6954, pp. 136–145. Springer, Heidelberg (2011)
31. Grzymala-Busse, J.W., Clark, P.G., Kuehnhausen, M.: Generalized probabilistic approximations of incomplete data. Int. J. Approximate Reasoning **132**, 180–196 (2014)
32. Grzymala-Busse, J.W., Grzymala-Busse, W.J.: Handling missing attribute values. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook, pp. 37–57. Springer, Heidelberg (2005)
33. Grzymala-Busse, J.W., Grzymala-Busse, W.J.: An experimental comparison of three rough set approaches to missing attribute values. Trans. Rough Sets **6**, 31–50 (2007)
34. Grzymala-Busse, J.W., Grzymala-Busse, W.J.: Improving quality of rule sets by increasing incompleteness of data sets. In: Proceedings of the Third International Conference on Software and Data Technologies, pp. 241–248 (2008)
35. Grzymala-Busse, J.W., Grzymala-Busse, W.J.: Inducing better rule sets by adding missing attribute values. In: Chan, C.-C., Grzymala-Busse, J.W., Ziarko, W.P. (eds.) RSCTC 2008. LNCS (LNAI), vol. 5306, pp. 160–169. Springer, Heidelberg (2008)
36. Grzymala-Busse, J.W., Grzymala-Busse, W.J.: Handling missing attribute values. In: Maimon, O., Rokach, L. (eds.) Data Mining and Knowledge Discovery Handbook, 2nd edn, pp. 33–51. Springer, Heidelberg (2010)
37. Grzymala-Busse, J.W., Grzymala-Busse, W.J., Goodwin, L.K.: A comparison of three closest fit approaches to missing attribute values in preterm birth data. Int. J. Intell. Syst. **17**(2), 125–134 (2002)
38. Grzymala-Busse, J.W., Grzymala-Busse, W.J., Hippe, Z.S., Rzasa, W.: An improved comparison of three rough set approaches to missing attribute values. In: Proceedings of the 16-th International Conference on Intelligent Information Systems, pp. 141–150 (2008)
39. Grzymala-Busse, J.W., Hippe, Z.S.: Mining data with numerical attributes and missing attribute values–a rough set approach. In: Proceedings of the IEEE International Conference on Granular Computing, pp. 144–149 (2011)

40. Grzymała-Busse, J.W., Hu, M.: A comparison of several approaches to missing attribute values in data mining. In: Ziarko, W.P., Yao, Y. (eds.) RSCTC 2000. LNCS (LNAI), vol. 2005, p. 378. Springer, Heidelberg (2001)

41. Grzymala-Busse, J.W., Rzasa, W.: Local and global approximations for incomplete data. In: Greco, S., Hata, Y., Hirano, S., Inuiguchi, M., Miyamoto, S., Nguyen, H.S., Słowiński, R. (eds.) RSCTC 2006. LNCS (LNAI), vol. 4259, pp. 244–253. Springer, Heidelberg (2006)

42. Grzymala-Busse, J.W., Rzasa, W.: Local and global approximations for incomplete data. Trans. Rough Sets **8**, 21–34 (2008)

43. Grzymala-Busse, J.W., Rzasa, W.: A local version of the MLEM2 algorithm for rule induction. Fundamenta Informaticae **100**, 99–116 (2010)

44. Grzymala-Busse, J.W., Wang, A.Y.: Modified algorithms LEM1 and LEM2 for rule induction from data with missing attribute values. In: Proceedings of the 5-th International Workshop on Rough Sets and Soft Computing in conjunction with the Third Joint Conference on Information Sciences, pp. 69–72 (1997)

45. Grzymala-Busse, J.W., Yao, Y.: Probabilistic rule induction with the LERS data mining system. Int. J. Intell. Syst. **26**, 518–539 (2011)

46. Grzymala-Busse, J.W., Ziarko, W.: Data mining based on rough sets. In: Wang, J. (ed.) Data Mining: Opportunities and Challenges, pp. 142–173. Idea Group Publ, Hershey (2003)

47. Guan, L., Wang, G.: Generalized approximations defined by non-equivalence relations. Inf. Sci. **193**, 163–179 (2012)

48. Hong, T.P., Tseng, L.H., Chien, B.C.: Learning coverage rules from incomplete data based on rough sets. In: Proceedings of the IEEE International Conference on Systems, Man and Cybernetics, pp. 3226–3231 (2004)

49. Hong, T.P., Tseng, L.H., Wang, S.L.: Learning rules from incomplete training examples by rough sets. Expert Syst. Appl. **22**, 285–293 (2002)

50. Kryszkiewicz, M.: Rough set approach to incomplete information systems. In: Proceedings of the Second Annual Joint Conference on Information Sciences, pp. 194–197 (1995)

51. Kryszkiewicz, M.: Rules in incomplete information systems. Inf. Sci. **113**(3–4), 271–292 (1999)

52. Latkowski, R.: On decomposition for incomplete data. Fundamenta Informaticae **54**, 1–16 (2003)

53. Latkowski, R., Mikołajczyk, M.: Data decomposition and decision rule joining for classification of data with missing values. In: Tsumoto, S., Słowiński, R., Komorowski, J., Grzymała-Busse, J.W. (eds.) RSCTC 2004. LNCS (LNAI), vol. 3066, pp. 254–263. Springer, Heidelberg (2004)

54. Leung, Y., Wu, W., Zhang, W.: Knowledge acquisition in incomplete information systems: a rough set approach. Eur. J. Ope. Res. **168**, 164–180 (2006)

55. Li, D., Deogun, I., Spaulding, W., Shuart, B.: Dealing with missing data: algorithms based on fuzzy set and rough set theories. Trans. Rough Sets **4**, 37–57 (2005)

56. Li, H., Yao, Y., Zhou, X., Huang, B.: Two-phase rule induction from incomplete data. In: Wang, G., Li, T., Grzymala-Busse, J.W., Miao, D., Skowron, A., Yao, Y. (eds.) RSKT 2008. LNCS (LNAI), vol. 5009, pp. 47–54. Springer, Heidelberg (2008)

57. Li, T., Ruan, D., Geert, W., Song, J., Xu, Y.: A rough sets based characteristic relation approach for dynamic attribute generalization in data mining. Knowl. Based Syst. **20**(5), 485–494 (2007)

58. Li, T., Ruan, D., Song, J.: Dynamic maintenance of decision rules with rough set under characteristic relation. In: Proceedings of the International Conference on Wireless Communications, Networking and Mobile Computing, pp. 3713–3716 (2007)
59. Meng, Z., Shi, Z.: A fast approach to attribute reduction in incomplete decision systems with tolerance relation-based rough sets. Inf. Sci. **179**, 2774–2793 (2009)
60. Meng, Z., Shi, Z.: Extended rough set-based attribute reduction in inconsistent incomplete decision systems. Inf. Sci. **204**, 44–69 (2012)
61. Nakata, M., Sakai, H.: Rough sets handling missing values probabilistically interpreted. In: Slezak, D., Wang, G., Szczuka, M.S., Düntsch, I., Yao, Y. (eds.) RSFD-GrC 2005. LNCS (LNAI), vol. 3641, pp. 325–334. Springer, Heidelberg (2005)
62. Nakata, M., Sakai, H.: Applying rough sets to information tables containing missing values. In: Proceedings of the 39-th International Symposium on Multiple-Valued Logic, pp. 286–291 (2009)
63. Pawlak, Z.: Rough sets. Int. J. Comput. Inf. Sci. **11**, 341–356 (1982)
64. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Kluwer Academic Publishers, Dordrecht (1991)
65. Pawlak, Z., Grzymala-Busse, J.W., Slowinski, R., Ziarko, W.: Rough sets. Commun. ACM **38**, 89–95 (1995)
66. Pawlak, Z., Skowron, A.: Rough sets: some extensions. Inf. Sci. **177**, 28–40 (2007)
67. Pawlak, Z., Wong, S.K.M., Ziarko, W.: Rough sets: probabilistic versus deterministic approach. Int. J. Man-Mach. Stud. **29**, 81–95 (1988)
68. Peng, H., Zhu, S.: Handling of incomplete data sets using ICA and SOM in data mining. Neural Comput. Appl. **16**, 167–172 (2007)
69. Qi, Y.S., Wei, L., Sun, H.J., Song, Y.Q., Sun, Q.S.: Characteristic relations in generalized incomplete information systems. In: International Workshop on Knowledge Discovery and Data Mining, pp. 519–523 (2008)
70. Qi, Y.S., Sun, H., Yang, X.B., Song, Y., Sun, Q.: Approach to approximate distribution reduct in incomplete ordered decision system. J. Inf. Comput. Sci. **3**, 189–198 (2008)
71. Qian, Y., Dang, C., Liang, J., Zhang, H., Ma, J.: On the evaluation of the decision performance of an incomplete decision table. Data Knowl. Eng. **65**, 373–400 (2008)
72. Qian, Y., Li, D., Wang, F., Ma, N.: Approximation reduction in inconsistent incomplete decision tables. Knowl. Based Syst. **23**, 427–433 (2010)
73. Ślęzak, D., Ziarko, W.: The investigation of the bayesian rough set model. Int. J. Approx. Reason. **40**, 81–91 (2005)
74. Song, J., Li, T., Ruan, D.: A new decision tree construction using the cloud transform and rough sets. In: Wang, G., Li, T., Grzymala-Busse, J.W., Miao, D., Skowron, A., Yao, Y. (eds.) RSKT 2008. LNCS (LNAI), vol. 5009, pp. 524–531. Springer, Heidelberg (2008)
75. Song, J., Li, T., Wang, Y., Qi, J.: Decision tree construction based on rough set theory under characteristic relation. In: Proceedings of the ISKE 2007, the 2-nd International Conference on Intelligent Systems and Knowledge Engineering Conference, pp. 788–792 (2007)
76. Stefanowski, J., Tsoukiàs, A.: On the extension of rough sets under incomplete information. In: Zhong, N., Skowron, A., Ohsuga, S. (eds.) RSFDGrC 1999. LNCS (LNAI), vol. 1711, pp. 73–82. Springer, Heidelberg (1999)
77. Stefanowski, J., Tsoukias, A.: Incomplete information tables and rough classification. Computat. Intell. **17**(3), 545–566 (2001)

78. Wang, G.: Extension of rough set under incomplete information systems. In: Proceedings of the IEEE International Conference on Fuzzy Systems, pp. 1098–1103 (2002)
79. Wong, S.K.M., Ziarko, W.: INFER–an adaptive decision support system based on the probabilistic approximate classification. In: Proceedings of the 6-th International Workshop on Expert Systems and their Applications, pp. 713–726 (1986)
80. Yang, X., Yang, J.: Incomplete Information System and Rough Set Theory: Model and Attribute Reduction. Springer, Heidelberg (2012)
81. Yang, X., Zhang, M., Dou, H., Yang, J.: Neighborhood systems-based rough sets in incomplete information systems. Knowl. Based Syst. **24**, 858–867 (2011)
82. Yao, Y.Y.: Probabilistic rough set approximations. Int. J. Approx. Reason. **49**, 255–271 (2008)
83. Yao, Y.Y., Wong, S.K.M.: A decision theoretic framework for approximate concepts. Int. J. Man Mach. Stud. **37**, 793–809 (1992)
84. Yao, Y.Y., Wong, S.K.M., Lingras, P.: A decision-theoretic rough set model. In: Ras, Z.W., Zemankova, M., Emrich, M.L. (eds.) Methodologies for Intelligent Systems, North-Holland, pp. 388–395 (1990)
85. Ziarko, W.: Variable precision rough set model. J. Comput. Syst. Sci. **46**(1), 39–59 (1993)
86. Ziarko, W.: Probabilistic approach to rough sets. Int. J. Approx. Reason. **49**, 272–284 (2008)