Consistency of Incomplete Data

Patrick G. Clark[1] and Jerzy Grzymala-Busse[1,2]
[1]Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS 66045, U.S.A.
[2]Institute of Computer Science, Polish Academy of Sciences, 01-237 Warsaw, Poland

Keywords: Consistency, Incomplete Data, Missing Attribute Values, Rough Set Theory.

Abstract: In this paper we introduce an idea of consistency for incomplete data sets. Consistency is well-known for completely specified data sets, where a data set is consistent if for any two cases with equal all attribute values, both cases belong to the same concept. We generalize the definition of consistency for incomplete data sets using rough set theory. For incomplete data sets there exist three definitions of consistency. We discuss two types of missing attribute values: lost values and "do not care" conditions. We illustrate an idea of consistency for incomplete data sets using experiments on many data sets with missing attribute values derived from five benchmark data sets. Results of our paper may be applied for increasing the efficiency of mining incomplete data.

1 INTRODUCTION

A complete data set, i.e., a data set with specified all attribute values, is consistent if for any two cases with the same attribute values, both cases belong to the same concept (class). Yet another definition of consistency is based on rough set theory: a complete data set is consistent if for any concept its lower and upper approximations are equal (Pawlak, 1982; Pawlak, 1991). Consistency for incomplete data sets, i.e., data sets with some missing attribute values, was not defined in the accessible literature.

The main objective of this paper is to study consistency for incomplete data sets using rough set theory. For incomplete data sets there exist three definitions of approximations, called *singleton*, *subset* and *concept*. Additionally, for complete data sets an idea of the approximation was generalized by introducing *probabilistic* approximations, with an additional parameter, interpreted as a probability (Grzymala-Busse and Ziarko, 2003; Pawlak et al., 1988; Yao, 2007; Yao and Wong, 1992; Yao et al., 1990; Ziarko, 1993; Ziarko, 2008). Probabilistic approximations were extended to incomplete data sets by introducing singleton, subset and concept probabilistic approximations in (Grzymala-Busse, 2011). First results on experiments on probabilistic approximations were published in (Clark and Grzymala-Busse, 2011).

We discuss singleton, subset and concept consistency for incomplete data sets and study their basic properties. Additionally, we conducted experiments on five benchmark data sets converted to many incomplete data sets. In incomplete data sets, there are two kinds of missing attribute values: lost values and "do not care" conditions (Grzymala-Busse, 2003). Lost values indicate the original attribute values were not known and processing this kind of missing attribute values is conducted by taking into account only existing, specified attribute values. On the other hand, "do not care" conditions indicate an actual value was one of the attribute values and we assumed that the missing attribute values may be replaced by any attribute value.

2 COMPLETE DATA SETS

Our basic assumption is that the data sets are presented in the form of a *decision table*. An example of a decision table is shown in Table 1. Rows of the decision table represent *cases*, while columns are labeled by *variables*. The set of all cases is denoted by U. In Table 1, $U = \{1, 2, 3, 4, 5, 6, 7, 8\}$. Some variables are called *attributes* while one selected variable is called a *decision* and is denoted by d. The set of all attributes will be denoted by A. In Table 1, $A = \{Temperature, Headache, Cough\}$ and $d = Flu$. For an attribute a and case x, $a(x)$ denotes the value of the attribute a for case x. For example, *Temperature* (1) = *high*.

Clark P. and Grzymala-Busse J..
Consistency of Incomplete Data.
DOI: 10.5220/0004490300800087
In *Proceedings of the 2nd International Conference on Data Technologies and Applications* (DATA-2013), pages 80-87
ISBN: 978-989-8565-67-9

Table 1: A complete decision table.

	Attributes			Decision
Case	Temperature	Headache	Cough	Flu
1	high	yes	yes	yes
2	normal	yes	no	yes
3	high	no	no	yes
4	high	no	no	yes
5	high	no	no	no
6	normal	yes	no	no
7	normal	yes	no	no
8	normal	no	yes	no

A significant idea used for scrutiny of 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 following set

$$\{x \mid a(x) = v\}. \quad (1)$$

For Table 1, blocks of all attribute-value pairs are

[(Temperature, normal)] = {2, 6, 7, 8},
[(Temperature, high)] = {1, 3, 4, 5},
[(Headache, yes)] = {1, 2, 6, 7},
[(Headache, no)] = {3, 4, 5, 8},
[(Cough, yes)] = {1, 8},
[(Cough, no)] = {2, 3, 4, 5, 6, 7}.

A special block of a decision-value pair is called a *concept*. In Table 1, the concepts are [(Flu, yes)] = {1, 2, 3, 4} and [(Flu, no)] = {5, 6, 7, 8}.

Let B be a subset of the set A of all attributes. Complete data sets are characterized by the indiscernibility relation $IND(B)$ defined as follows: for any $x, y \in U$,

$$(x, y) \in IND(B) \text{ if and only if } a(x) = a(y) \text{ for any } a \in B \quad (2)$$

Obviously, $IND(B)$ is an equivalence relation. The equivalence class of $IND(B)$ containing $x \in U$ will be denoted by $[x]_B$ and called B-elementary set. A-elementary sets will be called *elementary*. We have

$$[x]_B = \cap\{[(a, a(x))] \mid a \in B\}. \quad (3)$$

The set of all equivalence classes $[x]_B$, where $x \in U$, is a partition on U denoted by B^*. For Table 1, A^* = {{1}, {2, 6, 7}, {3, 4, 5}, {8}}. All members of A^* are elementary sets.

The data set, presented in Table 1, contains *conflicting cases*, for example cases 2 and 6: for any attribute $a \in A$, a(2) = a(6), yet cases 2 and 6 belong to two different concepts. A data set containing conflicting cases will be called *inconsistent*. We may recognize that a data set is inconsistent comparing the partition A^* with a partition of all concepts: there exist an elementary set that is not a subset of any concept. For Table 1, the elementary set {2, 6, 7} is not a subset of any of the two concepts {1, 2, 3, 4} and {5, 6, 7, 8}. There exists yet another way to recognize inconsistency of data sets, based on ideas of B-lower and B-upper approximations. Let X be a subset of U. The B-*lower approximation* of X, denoted by $\underline{appr}_B(X)$, is defined as follows

$$\{x \mid x \in U, [x]_B \subseteq X\}. \quad (4)$$

The B-*upper approximation* of X, denoted by $\overline{appr}_B(X)$, is defined as follows

$$\{x \mid x \in U, [x]_B \cap X \neq \emptyset\}. \quad (5)$$

For the data set from Table 1 and the concept [(Flu, yes)] = {1, 2, 3, 4} = X,

$$\underline{appr}_A(X) = \{1\}$$

and

$$\overline{appr}_A(X) = \{1, 2, 3, 4, 5, 6, 7\}.$$

A data set is inconsistent if and only if there exists a concept X for which $\underline{appr}_A(X) \neq \overline{appr}_A(X)$.

3 INCOMPLETE DATA SETS

An example of the incomplete data set is presented in Table 2. Some attribute values are missing. Such values are denoted either by ?, denoting a lost value (the original attribute value was not known, we will try to use only existing, specified attribute values) or by $*$, denoting a "do not care" condition (we are assuming that the missing attribute value may be replaced by any attribute value).

For incomplete decision tables the definition of a block of an attribute-value pair is modified (Grzymala-Busse, 2003; Grzymala-Busse, 2004a; Grzymala-Busse, 2004b).

- 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 in-

Table 2: An incomplete decision table.

	Attributes			Decision
Case	Temperature	Headache	Cough	Flu
1	$*$	yes	yes	yes
2	normal	?	no	yes
3	?	no	$*$	yes
4	high	no	no	yes
5	high	$*$	?	no
6	normal	yes	$*$	no
7	normal	yes	no	no
8	normal	?	yes	no

cluded in blocks $[(a,v)]$ for all specified values v of attribute a.

Thus, for Table 2, the blocks of all attribute-value pairs are
[(Temperature, normal)] = {1, 2, 6, 7, 8},
[(Temperature, high)] = {1, 4, 5},
[(Headache, yes)] = {1, 5, 6, 7},
[(Headache, no)] = {3, 4, 5},
[(Cough, yes)] = {1, 3, 6, 8},
[(Cough, no)] = {2, 3, 4, 6, 7}.

Let B be a subset of the set A of all attributes. 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 $K(x,a) = U$.

Characteristic set $K_B(x)$ may be interpreted as the set of cases that are indistinguishable from x using all attributes from B, with a given interpretation of missing attribute values. Thus, $K_A(x)$ is the set of all cases that cannot be distinguished from x using all attributes.

For Table 2 and $B = A$,

$K_A(1) = U \cap \{1,5,6,7\} \cap \{1,3,6,8\} = \{1,6\}$,
$K_A(2) = \{1,2,6,7,8\} \cap U \cap \{2,3,4,6,7\} = \{2,6,7\}$,
$K_A(3) = U \cap \{3,4,5\} \cap U = \{3,4,5\}$,
$K_A(4) = \{1,4,5\} \cap \{3,4,5\} \cap \{2,3,4,6,7\} = \{4\}$,
$K_A(5) = \{1,4,5\} \cap U \cap U = \{1,4,5\}$,
$K_A(6) = \{1,2,6,7,8\} \cap \{1,5,6,7\} \cap U = \{1,6,7\}$,
$K_A(7) = \{1,2,6,7,8\} \cap \{1,5,6,7\} \cap \{2,3,4,6,7\} = \{6,7\}$, and
$K_A(8) = \{1,2,6,7,8\} \cap U \cap \{1,3,6,8\} = \{1,6,8\}$.

For incomplete data sets there exist three distinct definitions of approximations. Let X be a subset of U.

The B-*singleton lower approximation* of X, denoted by $\underline{appr}_B^{singleton}(X)$, is defined as follows

$$\{x \mid x \in U, K_B(x) \subseteq X\}. \tag{6}$$

The singleton lower approximations were studied in many papers, see, e.g., (Grzymala-Busse, 2003; Grzymala-Busse, 2004b; Kryszkiewicz, 1995; Kryszkiewicz, 1999; Lin, 1989; Lin, 1992; Slowinski and Vanderpooten, 2000; Stefanowski and Tsoukias, 1999; Stefanowski and Tsoukias, 2001; Yao, 1998).

The B-*singleton upper approximation* of X, denoted by $\overline{appr}_B^{singleton}(X)$, is defined as follows

$$\{x \mid x \in U, K_B(x) \cap X \neq \emptyset\}. \tag{7}$$

The singleton upper approximations, like singleton lower approximations, were also studied in many papers, e.g., (Grzymala-Busse, 2003; Grzymala-Busse, 2004b; Kryszkiewicz, 1995; Kryszkiewicz, 1999; Slowinski and Vanderpooten, 2000; Stefanowski and Tsoukias, 1999; Stefanowski and Tsoukias, 2001; Yao, 1998).

The B-*subset lower approximation* of X, denoted by $\underline{appr}_B^{subset}(X)$, is defined as follows

$$\cup \{K_B(x) \mid x \in U, K_B(x) \subseteq X\}. \tag{8}$$

The subset lower approximations were introduced in (Grzymala-Busse, 2003; Grzymala-Busse, 2004b).

The B-*subset upper approximation* of X, denoted by $\overline{appr}_B^{subset}(X)$, is defined as follows

$$\cup \{K_B(x) \mid x \in U, K_B(x) \cap X \neq \emptyset\}. \tag{9}$$

The subset upper approximations were introduced in (Grzymala-Busse, 2003; Grzymala-Busse, 2004b).

The B-*concept lower approximation* of X, denoted by $\underline{appr}_B^{concept}(X)$, is defined as follows

$$\cup \{K_B(x) \mid x \in X, K_B(x) \subseteq X\}. \tag{10}$$

The concept lower approximations were introduced in (Grzymala-Busse, 2003; Grzymala-Busse, 2004b).

The B-*concept upper approximation* of X, denoted by $\overline{appr}_B^{concept}(X)$, is defined as follows

$$\begin{aligned} &\cup \{K_B(x) \mid x \in X, K_B(x) \cap X \neq \emptyset\} \\ &= \cup \{K_B(x) \mid x \in X\}. \end{aligned} \tag{11}$$

The concept upper approximations were studied in (Grzymala-Busse, 2003; Grzymala-Busse, 2004b; Lin, 1992).

For Table 2 and X = {1, 2, 3, 4}, all A-singleton, A-subset and A-concept approximations are:

$\underline{appr}_A^{singleton}(X) = \{4\}$,

$$\overline{appr}_A^{singleton}(X) = \{1,2,3,4,5,6,8\},$$

$$\underline{appr}_A^{subset}(X) = \{4\},$$

$$\overline{appr}_A^{subset}(X) = U,$$

$$\underline{appr}_A^{concept}(X) = \{4\},$$

$$\overline{appr}_A^{concept}(X) = \{1,2,3,4,5,6,7\}.$$

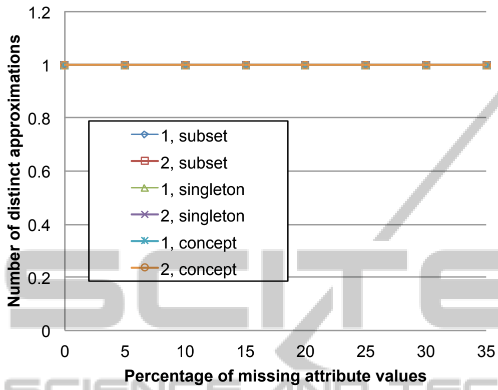

Figure 1: Bankruptcy data set with lost values.

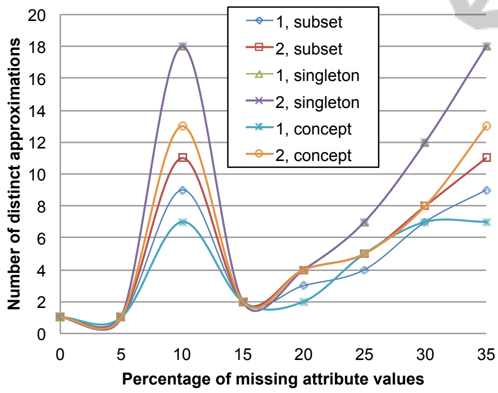

Figure 2: Bankruptcy data set with "do not care" conditions.

4 PROBABILISTIC APPROXIMATIONS

By analogy with lower and upper approximations defined using characteristic sets, we will introduce three kinds of probabilistic approximations: singleton, subset and concept. Again, let B be a subset of the attribute set A and X be a subset of U.

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

$$\{x \mid x \in U,\ Pr(X \mid K_B(x)) \geq \alpha\}, \qquad (12)$$

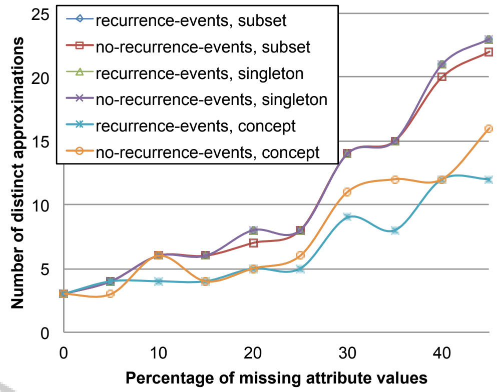

Figure 3: Breast cancer data set with lost values.

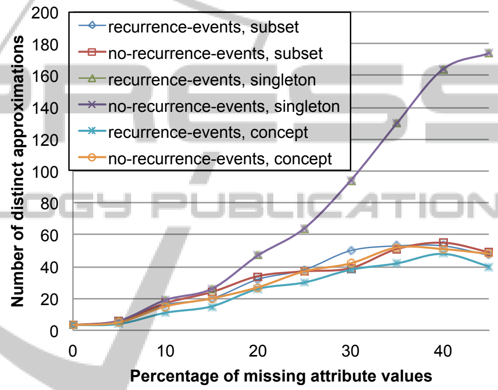

Figure 4: Breast cancer data set with "do not care" conditions.

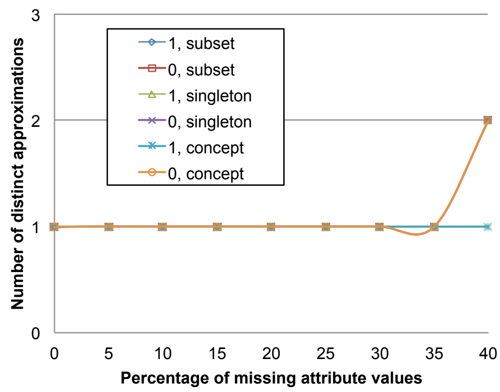

Figure 5: Echocardiogram data set with lost values.

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 as follows

$$\cup\{K_B(x) \mid x \in U,\ Pr(X \mid K_B(x)) \geq \alpha\}. \qquad (13)$$

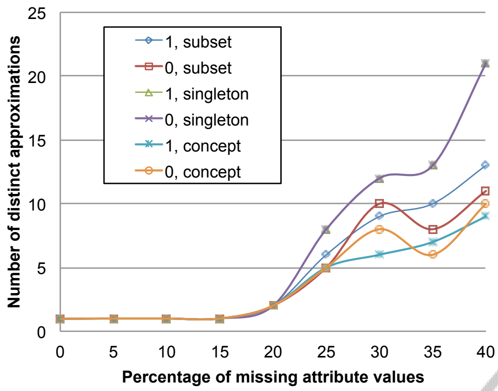

Figure 6: Echocardiogram data set with "do not care" conditions.

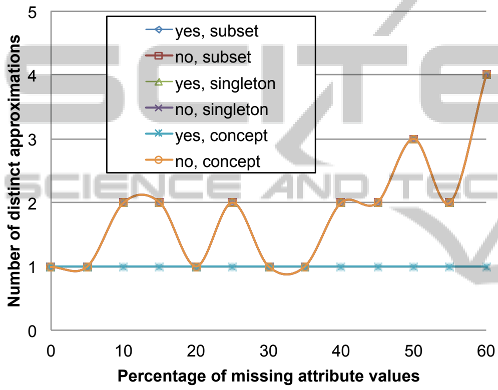

Figure 7: Hepatitis data set with lost values.

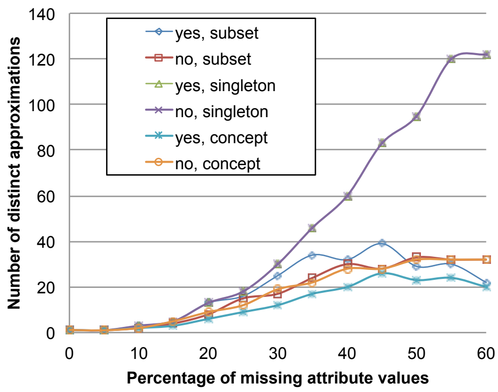

Figure 8: Hepatitis data set with "do not care" conditions.

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

$$\cup\{K_B(x) \mid x \in X,\ Pr(X \mid K_B(x)) \geq \alpha\}. \qquad (14)$$

Let $type \in \{singleton, subset, concept\}$. Note that

$$appr^{type}_{1,B}(X) = \underline{appr}^{type}_{B}(X) \qquad (15)$$

and for the smallest possible positive α (in our experiments such $\alpha = 0.001$)

$$appr^{type}_{\alpha,B}(X) = \overline{appr}^{type}_{B}(X). \qquad (16)$$

For Table 2, all distinct A-singleton, A-subset and A-concept approximations of the set $X = \{1, 2, 3, 4\}$ are

$$appr^{singleton}_{0.333,A}(X) = \{1,2,3,4,5,6,8\}$$

$$appr^{singleton}_{0.5,A}(X) = \{1,3,4,5\},$$

$$appr^{singleton}_{0.667,A}(X) = \{3,4,5\},$$

$$appr^{singleton}_{1,A}(X) = \{4\},$$

$$appr^{subset}_{0.333,A}(X) = U,$$

$$appr^{subset}_{0.5,A}(X) = \{1,3,4,5,6\},$$

$$appr^{subset}_{0.667,A}(X) = \{1,3,4,5\},$$

$$appr^{subset}_{1,A}(X) = \{4\},$$

$$appr^{concept}_{0.333,A}(X) = \{1,2,3,4,5,6,7\},$$

$$appr^{concept}_{0.5,A}(X) = \{1,3,4,5,6\},$$

$$appr^{concept}_{0.667,A}(X) = \{3,4,5\},$$

$$appr^{concept}_{1,A}(X) = \{4\}.$$

5 CONSISTENCY

Let X be a concept of the incomplete data set, i.e., a block of a decision-value pair. Let B be a subset of the set A of all attributes. Let $type \in \{singleton, subset, concept\}$. The concept X is *B-type consistent* in the data set if and only if for any $\alpha > 0$

$$appr^{type}_{\alpha,B}(X) = X. \qquad (17)$$

If the concept X is *A-type consistent* in the data set then it will be called *type consistent*.

For completely specified data sets, if $\underline{B}(X) = X$ or $\overline{B}(X) = X$ then $\underline{B}(X) = X = \overline{B}(X)$. An analogous result is not true for incomplete data sets.

A data set will be called *B-type consistent* when for any concept X the number of all distinct B-type probabilistic approximations of X is equal to one. If the data set is *A-type consistent* then it will be called *type consistent*.

6 EXPERIMENTS

We conducted experiments on eight benchmark data sets, taken from the University of California at Irvine *Machine learning Repository*, see Table 3. For any data set, a family of new data sets was derived, with randomly placed lost values, starting from 0%, with the number of lost values gradually increasing, with the increment of 5%, until in the process of adding lost values, for some case, all attribute values were missing. If so, we conducted an additional two random attempts and, if there was still a case with missing all attribute values, the process of deriving new data sets with lost values was terminated. For any derived data set with lost values we created a corresponding data set with "do not care" conditions by replacing all lost values by "do not care" conditions. Our objective was to test, for any data set from such set family, how many distinct singleton, subset and concept probabilistic approximations may be created when the α parameter changes from 0.001 to 1.

For any data set, the number of distinct singleton, subset and concept probabilistic approximations for all concepts were recorded, see Figures 1–14. On these figures the numbers of distinct singleton, subset and concept approximations are shown for all concepts. For example, on Figure 1, "1, subset" means the concept labeled in the data set by "1" combined with the subset probabilistic approximation. It is clear that for data sets with "do not care" conditions all three numbers were larger than for corresponding data sets with lost values.

All eight data sets created from the *bankruptcy* data, with percentage of lost values starting at 0% and ending with 35%, are singleton consistent (and hence subset and concept consistent). Additionally, the *bankruptcy* data set with 0% and 5% of "do not care" conditions are also singleton consistent. On the other hand, all 19 data sets created from the *breast cancer* data set, i.e., the original data set, 9 data sets with lost values and 9 data sets with "do not care" conditions are singleton inconsistent.

All data sets with up to 35% of lost values, derived from the *echocardiogram* data set, are singleton consistent. The data set derived from the *echocardiogram* data set, with 40% of lost values, is an example of the data set with one concept (labeled by "1") that is concept consistent but not singleton consistent. Four data sets derived from the *echocardiogram* data set, with 0, 5, 10 and 15% of "do not care" conditions, are singleton consistent.

For the data sets derived from *hepatitis* data set, both data sets with lost values and "do not care" conditions with 5% of missing attribute values are singleton consistent. Moreover, for the concept "yes", data sets with lost values derived from the *hepatitis* data set, up to 60%, are concept consistent. For data sets derived from the *wine recognition* data set, for the concept "3" with up to 5% of lost values and with up to 5% of "do not care" conditions, are singleton consistent.

Table 3: Data sets used for experiments.

Data set	Number of cases	Number of attributes	Number of concepts
Bankruptcy	66	5	2
Breast cancer	277	9	2
Echocardiogram	74	7	2
Hepatitis	155	19	2
Wine recognition	178	13	3

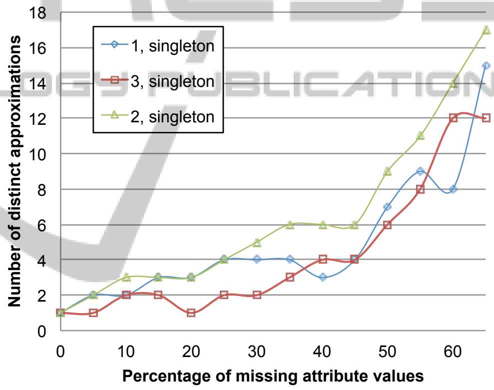

Figure 9: Singleton probabilistic approximations for wine recognition data set with lost values.

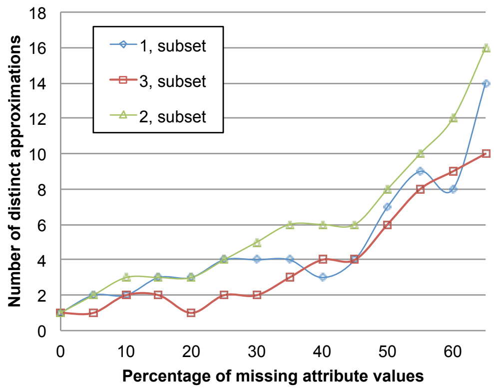

Figure 10: Subset probabilistic approximations for wine recognition data set with lost values.

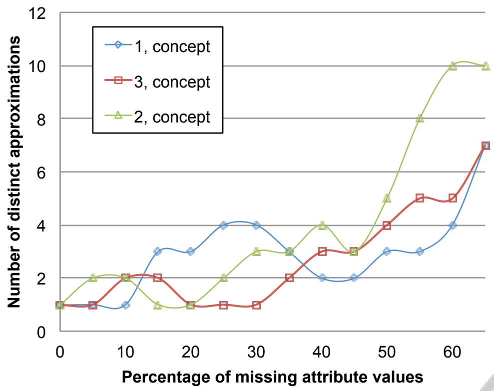

Figure 11: Concept probabilistic approximations for wine recognition data set with lost values.

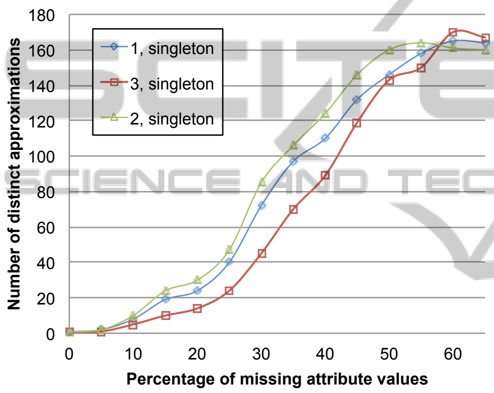

Figure 12: Singleton probabilistic approximations for wine recognition data set with "do not care" conditions.

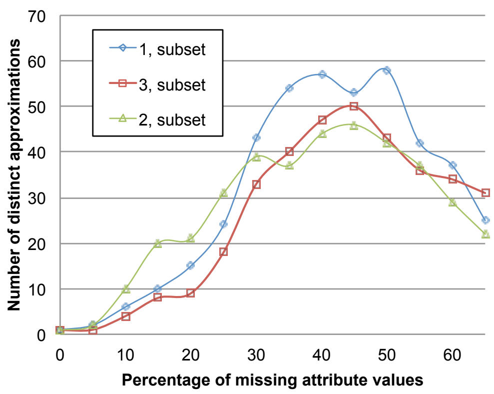

Figure 13: Subset probabilistic approximations for wine recognition data set with "do not care" conditions.

7 CONCLUSIONS

In this paper we introduced the idea of consistency for incomplete data sets. Consistency is defined for a single concept. For a given type of approximation,

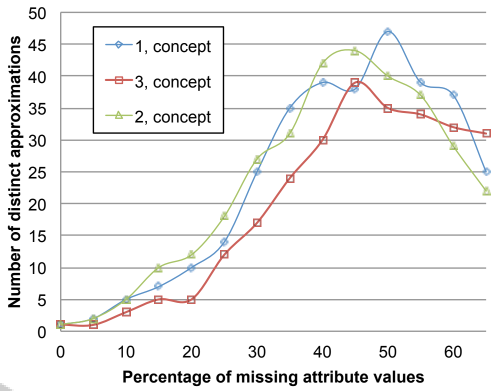

Figure 14: Concept probabilistic approximations for wine recognition data set with "do not care" conditions.

it is possible that some concepts are consistent while other concepts are not consistent. Consistency of the concept depends on type of approximation. A concept may be subset and concept consistent while it is singleton inconsistent.

Results of our experiments show that for a given data set (or a concept) there exist more consistent data sets, of all types, for data sets affected by lost values than for data sets affected by "do not care" conditions. For some data sets all derived incomplete data sets, for all concepts, are singleton consistent, for some data sets only some concepts are singleton consistent, while for some data sets all concepts are not singleton consistent.

REFERENCES

Clark, P. G. and Grzymala-Busse, J. W. (2011). Experiments on probabilistic approximations. In *Proceedings of the 2011 IEEE International Conference on Granular Computing*, pages 144–149.

Grzymala-Busse, J. W. (2003). Rough set strategies to data with missing attribute values. In *Workshop Notes, Foundations and New Directions of Data Mining, in conjunction with the 3-rd International Conference on Data Mining*, pages 56–63.

Grzymala-Busse, J. W. (2004a). Characteristic relations for incomplete data: A generalization of the indiscernibility relation. In *Proceedings of the Fourth International Conference on Rough Sets and Current Trends in Computing*, pages 244–253.

Grzymala-Busse, J. W. (2004b). Data with missing attribute values: Generalization of indiscernibility relation and rule induction. *Transactions on Rough Sets*, 1:78–95.

Grzymala-Busse, J. W. (2011). Generalized parameterized approximations. In *Proceedings of the RSKT 2011, the 6-th International Conference on Rough Sets and Knowledge Technology*, pages 136–145.

Grzymala-Busse, J. W. and Ziarko, W. (2003). Data mining based on rough sets. In Wang, J., editor, *Data Mining:*

Opportunities and Challenges, pages 142–173. Idea Group Publ., Hershey, PA.

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

Kryszkiewicz, M. (1999). Rules in incomplete information systems. *Information Sciences*, 113(3-4):271–292.

Lin, T. Y. (1989). Neighborhood systems and approximation in database and knowledge base systems. In *Proceedings of the ISMIS-89, the Fourth International Symposium on Methodologies of Intelligent Systems*, pages 75–86.

Lin, T. Y. (1992). Topological and fuzzy rough sets. In Slowinski, R., editor, *Intelligent Decision Support. Handbook of Applications and Advances of the Rough Sets Theory*, pages 287–304. Kluwer Academic Publishers, Dordrecht, Boston, London.

Pawlak, Z. (1982). Rough sets. *International Journal of Computer and Information Sciences*, 11:341–356.

Pawlak, Z. (1991). *Rough Sets. Theoretical Aspects of Reasoning about Data*. Kluwer Academic Publishers, Dordrecht, Boston, London.

Pawlak, Z., Wong, S. K. M., and Ziarko, W. (1988). Rough sets: probabilistic versus deterministic approach. *International Journal of Man-Machine Studies*, 29:81–95.

Slowinski, R. and Vanderpooten, D. (2000). A generalized definition of rough approximations based on similarity. *IEEE Transactions on Knowledge and Data Engineering*, 12:331–336.

Stefanowski, J. and Tsoukias, A. (1999). On the extension of rough sets under incomplete information. In *Proceedings of the RSFDGrC'1999, 7th International Workshop on New Directions in Rough Sets, Data Mining, and Granular-Soft Computing*, pages 73–81.

Stefanowski, J. and Tsoukias, A. (2001). Incomplete information tables and rough classification. *Computational Intelligence*, 17(3):545–566.

Yao, Y. Y. (1998). Relational interpretations of neighborhood operators and rough set approximation operators. *Information Sciences*, 111:239–259.

Yao, Y. Y. (2007). Decision-theoretic rough set models. In *Proceedings of the Second International Conference on Rough Sets and Knowledge Technology*, pages 1–12.

Yao, Y. Y. and Wong, S. K. M. (1992). A decision theoretic framework for approximate concepts. *International Journal of Man-Machine Studies*, 37:793–809.

Yao, Y. Y., Wong, S. K. M., and Lingras, P. (1990). A decision-theoretic rough set model. In *Proceedings of the 5th International Symposium on Methodologies for Intelligent Systems*, pages 388–395.

Ziarko, W. (1993). Variable precision rough set model. *Journal of Computer and System Sciences*, 46(1):39–59.

Ziarko, W. (2008). Probabilistic approach to rough sets. *International Journal of Approximate Reasoning*, 49:272–284.