NIS-Apriori-based rule generation with three-way decisions and its application system in SQL

doi: https://doi.org/10.1016/j.ins.2018.09.008

NIS-Apriori-Based Rule Generation with Three-way Decisions and Its Application System in SQL

Hiroshi Sakai[a,*], Michinori Nakata[b], Junzo Watada[c]

[a] *Department of Basic Sciences, Faculty of Engineering, Kyushu Institute of Technology, Tobata, Kitakyushu 804-8550, Japan*
[b] *Faculty of Management and Information Science, Josai International University, Gumyo, Togane, Chiba 283-0002, Japan*
[c] *Department of Computer & Information Sciences, Universiti Teknologi PETRONAS, 32610 Seri Iskandar, Perak Darul Ridzuan, Malaysia*

Abstract

In the study, non-deterministic information systems-Apriori-based (NIS-Apriori-based) rule generation from table data sets with incomplete information, SQL implementation, and the unique characteristics of the new framework are presented. Additionally, a few unsolved new research topics are proposed based on the framework. We follow the framework of NISs and propose certain rules and possible rules based on possible world semantics. Although each rule τ depends on a large number of possible tables, we prove that each rule τ is determined by examining only two τ-dependent possible tables. The NIS-Apriori algorithm is an adjusted Apriori algorithm that can handle such tables. Furthermore, it is logically sound and complete with regard to the rules. Subsequently, the implementation of the NIS-Apriori algorithm in SQL is described and a few new topics induced by effects of NIS-Apriori-based rule generation are confirmed. One of the topics that are considered is the possibility of estimating missing values via the obtained certain rules. The proposed methodology and the environment yielded by NIS-Apriori-based rule generation in SQL are useful for table data analysis with three-way decisions.

Keywords: Rule generation, Three-way decisions, Possible world

*Corresponding author.
Email addresses: sakai@mns.kyutech.ac.jp (Hiroshi Sakai), nakatam@ieee.org (Michinori Nakata), junzo.watada@gmail.com (Junzo Watada)

semantics, NIS-Apriori algorithm, Soundness and completeness,
Implementation in SQL, Imputation of missing values

1. Introduction

The study focuses on *Apriori algorithm* [3] based rule generation from tables with uncertainty and its actual application system. This field of research is closely related to rough sets [22, 24, 38], granular computing [25, 48], data mining [3, 5], and information incompleteness [11, 15, 16, 17, 20, 21, 23, 30, 34, 41, 43, 44, 45, 46]. Each field of research is related to others and we consider a review of the previous literature based on six classes, ranging from I to VI, as shown in Table 1. Our study belongs to Class VI. In the Introduction, we first review existing studies based on Table 1. We then describe the purpose of the study. In Table 1, the vertical heading is considered as information incompleteness and we consider exact data and inexact data. The horizontal heading is related to the purpose of the research. We consider information retrieval, model and approach for rules, and rule generation system.

Table 1: Classification of the previous studies related to our research.

	Information retrieval	Model and approach for rules	Rule generation system
Exact data	Class I	Class III	Class V
Inexact data	Class II	Class IV	Class VI

First, we review Class I (information retrieval with exact data) in Table 1. Marek and Pawlak [19] clarified the mathematical framework of information retrieval. The study appears to correspond to the origin of rough sets and table data analysis. The *definability* of a set (which is the basic concept of rough sets) was employed in the mathematical framework. Codd [7] also proposed *relational algebra* for table data management, and SQL systems were developed.

We subsequently focus on Class II (information retrieval with inexact data). Lipski [17] employed non-deterministic information to handle information incompleteness and investigated a question-answering system based on *possible world semantics* [14]. A query is transformed into a normal form

for the evaluation by possible worlds. Lipski proved that the set of axioms for the transformation corresponds to the system S4 in modal logic, and the set also corresponds to *sound* (a transformed query becomes the normal form) and *complete* (any query is transformed into the normal form) [17]. The property theoretically ensures the validity of the system. For example, the SLD-resolution algorithm in logic programs and deductive databases is sound and complete for logical consequences [18]. We agree with the property, and we introduce the property into rule generation, i.e., an algorithm generates a rule τ if and only if τ is defined as a rule. A rule generation system with soundness and completeness is rare. With respect to Prolog, Zadeh's fuzzy theory [48] was also applied to fuzzy databases and fuzzy Prolog, and Sakai [31] proposed a framework of logic programs with non-deterministic values.

Furthermore, Orłowska and Pawlak [21, 23] proposed *many valued information systems* and nondeterministic information systems to handle information incompleteness. We follow the systems and address their systems as *Non-deterministic Information Systems* (NISs). We correspondingly term a table without information incompleteness as *Deterministic Information Systems* (DISs).

We focus on Class III (model and approach for rules in exact data) as listed in Table 1. In the 1980s, the research trend appeared to shift from information retrieval to data mining and rule generation. Pawlak [22] proposed *rough set* theory that affords a mathematical framework of table data analysis. In rough set theory, *lower* and *upper approximations* of a target set X (of objects in a table) are generated via *equivalence classes*, and rules are obtained as a side effect. Several other related models and approaches to consider rules are investigated. Skowron et al. [38] proposed the *discernibility matrix* and *discernibility function* and proved that the problems of generating minimal (relative) *reducts* and of generating minimal dependencies are NP-hard. Greco and Słowiński [10] proposed a framework of *dominance-based rough sets* and handled rough sets for tables in which the attribute values are ordered. Ziarko [49] extended rough set models to *variable precision rough set models*. Komorowski et al. [12] surveyed the framework of rough sets. Tsumoto [42] applied rough set-based rule generation to medical data analysis. Yao [45, 46] extended rough sets to *three-way decisions* with probabilistic rough sets, and investigated rough set models in *multi-granulation spaces*. Ciucci [6] investigated lower and upper approximations for tables with rational numbers. Leung et al. [16] extended rough sets in tables to those in interval-valued information systems. Qian et al. [28] con-

sidered rules with a disjunctive decision part and termed the framework as *multi-granulation rough sets*. Zhu [47] proposed topological approaches to covering rough sets.

Information incompleteness is an extremely attractive issue, and thus it is natural that models and approaches for rules in exact data are extended to those in inexact data. We move to Class IV (model and approach for rules in inexact data) as shown in Table 1. Several other important models and approaches for rules in inexact data are investigated. Kryszkiewicz [15] characterised rules in incomplete information systems where incomplete attribute values '$*$' are introduced into DISs. Extended similarity relations obtained via the value '$*$' are applied to calculate the lower and upper approximations of a set X. Nakata et al. [20] followed Lipski's incomplete information and proposed rule generation based on possible world semantics. Stefanowski et al. [41] considered the relationship between incomplete information tables and rough classification. Wu et al. [43] investigated incomplete fuzzy information systems via a rough set approach. Yang et al. [44] examined the relationship between the dominance-based rough set approach and the incomplete interval-valued information system.

In the studies in Class III and Class IV, the main problem involves the characterization of rules via lower and upper approximations. Most studies focus on models and approaches. In order to discriminate the research on implementations from that on models and approaches, we consider Class V (rule generation system in exact data) and Class VI (rule generation system in inexact data). Predki et al. [27] developed a *Rough Set Data Explorer* for decision support, and Bazan et al. [4] created a *Rough Set Exploration System*, which is applicable to data exploration, classification support, and knowledge discovery. Grzymała-Busse [11] realised *Learning from Examples based on Rough Sets* (LERS). In LERS, a set X is covered by sets termed as *blocks*, and rules are generated as a side effect. Ślęzak et al. [39] considered the property of the distribution of attribute values and proposed the concept of packing in SQL. The technology is termed as *infobright*. In order to handle medical data sets, Tsumoto [42] generated the *PRIMEROSE system*. Recently, Riza et al. [29] developed a rough set-based package in R that employs rough set theory and fuzzy rough set theory.

Finally, we consider studies in Class VI. Although information incompleteness is extremely attractive, there is a paucity of studies on Class VI. Grzymała-Busse [11] introduced *missing values* into table data and extended the LERS system. It employs a few assumptions for the definition of blocks

that assume a role similar to that of equivalence classes in rough sets. After defining blocks, a covering algorithm is applied to a set X of objects.

We also dealt with studies in Class VI [32, 34]. We follow the framework of Lipski's incomplete information databases and Orłowska's NISs and propose a NIS-Apriori-based rule generation. The framework is related to rough sets in applying equivalence classes although the definition of lower and upper approximations is slightly different. Therefore, certain rules and possible rules in our framework are not investigated in Class IV. We develop the *NIS-Apriori algorithm* that corresponds to an adjusted Apriori algorithm [3] for the NIS case. The Apriori algorithm was proposed to obtain association rules from transaction data, and it is currently a representative algorithm for data mining [5].

We briefly reviewed related studies based on Class I through Class VI in Table 1. We now describe the purpose of the study. Our study belongs to Class VI, and the purpose involves realizing systems that handle tables with inexact data. More specifically, we describe the following.
(1) We clarify the difference between *Rough Set-based Rule Generation* (RSRG) and *Apriori-based Rule Generation* (APRG). Although they both handle rules from tables, the characteristics of the obtainable rules are different.
(2) We reconsider the theory of NIS-Apriori-based rule generation with respect to the aspect of aforementioned point (1). Thus, we solve a computational problem. Without the solution, it is difficult to address the rules from NISs.
(3) We present a prototype system in SQL that simulates the NIS-Apriori algorithm.
(4) We propose a few unsolved new research topics related to NIS-Apriori-based rule generation. As a topic, a plausible method to estimate the actual DIS from NIS is considered.
(5) By presenting (1) through (4), we demonstrate that NIS-Apriori-based rule generation is a significantly new framework and that it extends the research area of three-way decisions, rough sets, and granular computing.

The study is organised as follows. RSRG and APRG in DISs are reviewed in section 2. The rule generation in DISs are extended to NISs and the difference between RSRG and APRG is clarified in Section 3. This section also addresses the computational problem. In Section 4, the Apriori algorithm in DISs is adjusted to that in NISs. This modified algorithm is termed as the NIS-Apriori algorithm. In Section 5, a prototype system in SQL powered by the NIS-Apriori algorithm is presented. In Section 6, unsolved new topics

Table 2: Exemplary DIS ψ_{price}.

OB	$weight$	$size$	$price$
1	$light$	$small$	$high$
2	$light$	$medium$	$high$
3	$light$	$medium$	low
4	$heavy$	$large$	low
5	$heavy$	$large$	low

related to NIS-Apriori-based rule generation are introduced. In Section 7, a software tool to estimate the actual DIS from NIS is presented. Finally, in Section 8 the conclusions of the study are presented.

2. Rule Generation in DISs

This section clarifies Deterministic Information Systems (DISs), rules [12, 19, 22, 23, 24, 26, 38, 42, 49] in DISs, rough set-based rule generation (RSRG), and Apriori-based rule generation (APRG).

2.1. Rules in DISs

Specifically, the DIS ψ denotes a quadruplet ψ=$(OB, AT, \{VAL_A \mid A \in AT\}, f)$ where OB denotes a finite set in which the elements are termed as *objects*, AT denotes a finite set in which the elements are termed as *attributes*, VAL_A denotes a finite set in which the elements are termed as *attribute values*, and f is a mapping such that $f : OB \times AT \rightarrow \cup_{A \in AT} VAL_A$. We typically predefine a *decision attribute* $Dec \in AT$ and assume that CON is a subset of $AT \setminus \{Dec\}$. We assume that CON denotes a set of *condition attributes*. In ψ, a pair $[A, val]$ $(A \in AT, val \in VAL_A)$ is termed as a *descriptor*, and a formula $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$ $(val_A \in VAL_A, val \in VAL_{Dec})$ is termed as an *implication*. Table 2 is an exemplary DIS ψ_{price}.

Definition 1. *[12, 24, 38, 42] With respect to DIS ψ, two given threshold values $0 < \alpha, \beta \leq 1.0$, an implication τ that satisfy (1) and (2) are termed as (a candidate of) a rule in ψ.*
(1) support(τ) $(= |eq(\wedge_{A \in CON}[A, val_A] \wedge [Dec, val])|/|OB|) \geq \alpha$,
(2) accuracy(τ) $(= |eq(\wedge_{A \in CON}[A, val_A] \wedge [Dec, val])|/|eq(\wedge_{A \in CON}[A, val_A])|) \geq \beta$.

Here, $eq(*)$ denotes a set of the objects that satisfy formula $*$, and $|M|$ (M denotes a set of objects) denotes the cardinality of the set M. If $|eq(\wedge_{A \in CON}[A, val_A])|=0$, we define $support(\tau)=0$ and $accuracy(\tau)=0$.

Example 1. *We consider an exemplary DIS ψ_{price} in Table 2. With respect to two threshold values $\alpha=0.2$ and $\beta=0.7$, we assume that τ_1 is $[weight, light] \Rightarrow [price, high]$. Since $|eq([weight, light])|=|\{1,2,3\}|=3$ and $|eq([weight, light] \wedge [price, high])|= |\{1,2\}|$, $support(\tau_1)=2/5 > 0.2$ and $accuracy(\tau_1)=2/3 < 0.7$ hold. Thus, it is observed that τ_1 is not a rule. With respect to τ_2 : $[weight, light] \wedge [size, small] \Rightarrow [price, high]$, $|eq([weight, light] \wedge [size, small])|=1$ and $|eq([weight, light] \wedge [size, small] \wedge [price, high])|=1$ hold, and thus we obtain $support(\tau_2)=1/5 \geq 0.2$ and $accuracy(\tau_2)= 1/1 > 0.7$. Thus, the results indicate that τ_2 denotes a rule.*

2.2. Rough Set-Based Rule Generation (RSRG) in DISs

In rough set theory, the equivalence classes defined by descriptors are initially prepared, and a target set X is then covered by the equivalence classes. Generally, we consider that $eq(\wedge_{A \in CON}[A, val_A])$ denotes a *lower approximation* of a set X, if $eq(\wedge_{A \in CON}[A, val_A]) \subseteq X$. Conversely, consider that $eq(\wedge_{A \in CON}[A, val_A])$ is an *upper approximation* of a set X, if $X \subseteq eq(\wedge_{A \in CON}[A, val_A])$. The method is theoretically supported by the next proposition.

Proposition 1. *[24] In DISs and an implication $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$, the following (1) and (2) are equivalent:*
(1) $accuracy(\tau)=1$,
(2) $eq(\wedge_{A \in CON}[A, val_A]) \subseteq eq([Dec, val])$.

Ziarko [49] introduced variable precision rough sets to handle implications that satisfy $accuracy(\tau) \geq \beta$ ($0.5 < \beta \leq 1$). Based on the study by Skowron, it is NP-hard to determine all minimal reducts [38], and more than a conjunction $\wedge_{A \in CON}[A, val_A]$ that satisfies (2) in Proposition 1 exists. Thus, more than a rule exists for X specified by a descriptor. In RSRG without the backtracking functionality, completeness (any rule is obtained by the algorithm) is not ensured. However, in the study by Skowron, all conjunctions are handled in the discernibility function, and thus completeness is ensured.

2.3. Apriori-Based Rule Generation (APRG) in DISs

The Apriori algorithm for handling transaction data [3] was proposed by Agrawal. It is currently a representative data mining algorithm [5]. The algorithm is adjusted to table data sets by identifying each descriptor with an item. Here, we obtain the following two useful properties.

(Property 1) With respect to two threshold values α and β, each descriptor $[A, val_A]$ satisfying $|eq([A, val_A])|/|OB| < \alpha$ is ignored in APRG because any implication τ including $[A, val_A]$ does not satisfy $support(\tau) \geq \alpha$. Therefore, it is sufficient to consider descriptors in $\{[A, val_A] \mid |eq([A, val_A])|/|OB| \geq \alpha\}$. The property reduces the amount of meaningless implications. If we employ higher values for α, the amount of the obtainable rules is reduced.

(Property 2) With respect to $\eta_1 : [A, val_A] \Rightarrow [Dec, val]$ and $\eta_2 : [B, val_B] \Rightarrow [Dec, val]$ satisfying $accuracy(\eta_1) < \beta$ and $accuracy(\eta_2) < \beta$, $accuracy(\eta_3) \geq \beta$ occurs for $\eta_3 : [A, val_A] \wedge [B, val_B] \Rightarrow [Dec, val]$. Therefore, if $support(\eta_1) \geq$

Algorithm 1 Apriori algorithm adjusted for the DIS case

Input: DIS ψ, decision attribute Dec, threshold values α and β.
Output: $Rule(\psi)$.

1: $Rule(\psi) \leftarrow \{\}$, $i \leftarrow 1$
2: create IMP_i, where each $\tau_{i,j} \in IMP_i$ satisfies $support(\tau_{i,j}) \geq \alpha$
3: **while** $(|IMP_i| \geq 1)$ **do**
4: $Rest \leftarrow \{\}$
5: **for all** $\tau_{i,j} \in IMP_i$ **do**
6: **if** $accuracy(\tau_{i,j}) \geq \beta$ **then**
7: add $\tau_{i,j}$ to $Rule(\psi)$
8: **else**
9: add $\tau_{i,j}$ to $Rest$
10: **end if**
11: **end for**
12: $i \leftarrow i + 1$
13: generate IMP_i via $Rest$ and Property (2) in Section 2.3, where $\tau_{i,j} \in IMP_i$ satisfies
14: (i) $support(\tau_{i,j}) \geq \alpha$, and
15: (ii) $\tau_{i,j}$ is not a redundant implication for any implication in $Rule(\psi)$
16: **end while**
17: **return** $Rule(\psi)$

α and $support(\eta_2) \geq \alpha$, then it is necessary to consider η_3. Thus, the condition part of any rule $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$ consists of the descriptors in $\{[A, val_A] \mid support([A, val_A] \Rightarrow [Dec, val]) \geq \alpha\}$.

In Algorithm 1, IMP_{i+1} is generated from IMP_i by using the aforementioned properties. We say $\tau' : (\wedge_{A \in CON}[A, val_A]) \wedge [B, val_B]) \Rightarrow [Dec, val]$ is a *redundant* implication for $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$. If we consider that τ denotes a rule, we automatically consider that τ' also corresponds to a rule to reduce the number of rules. We only handle each rule in which the condition part is minimal.

Proposition 2. *Algorithm 1 is sound and complete for the defined rules in DIS ψ. Thus, $Rule(\psi)$=$\{\tau \mid support(\tau) \geq \alpha$ and $accuracy(\tau) \geq \beta$ in $\psi\}$ holds in Algorithm 1.*
(Proof)
(Soundness) Each implication $\tau \in IMP_i$ satisfies $support(\tau) \geq \alpha$, and τ is added to $Rule(\psi)$, if $accuracy(\tau) \geq \beta$ holds. This means each $\tau \in Rule(\psi)$ satisfies $support(\tau) \geq \alpha$ and $accuracy(\tau) \geq \beta$ in ψ.
(Completeness) Any implication τ (except the redundant case) is included in IMP_i, and the $accuracy(\tau)$ value is examined in Algorithm 1. So, there is no implication τ satisfying $support(\tau) \geq \alpha$, $accuracy(\tau) \geq \beta$ and $\tau \notin Rule(\psi)$.

2.4. Discussion of RSRG and APRG

Two types of rule generation handle rules from DISs although the obtainable rules are slightly different. The subsection clarifies the difference between RSRG and APRG from two aspects.
(Aspect 1: Characteristics of the obtainable rules)
RSRG: Let $eq([Dec, val])$ be a set X. The X is covered by other equivalence classes, and a few rules $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$ are generated. Therefore, the obtained rules are for the descriptor $[Dec, val]$ that defines the set X.
APRG: The set X is not specified, and it is observed that X=OB (the total object set). Therefore, a few representative rules with high *support* and *accuracy* values in table data sets are obtained.
(Aspect 2: Logical property of the rule generation algorithm)
RSRG: With respect to a covering of X, the set CON of the condition attributes is potentially not unique. This can lead to missing a few rules. Specifically, in ψ_{price} in Table 2, we consider a target set X=$\{3, 4, 5\}$ defined by the descriptor $[price, low]$. Given that $eq([weight, heavy])$=$\{4, 5\}$

holds, $\{4,5\} \subseteq X$ is derived. Thus, the set $\{4,5\}$ denotes a lower approximation of X, and we obtain a rule $[weight, heavy] \Rightarrow [price, low]$. However, $eq([size, large])=\{4,5\}$ also holds, and thus another rule is necessary $[size, large] \Rightarrow [price, low]$. Namely, a few rules exist for a lower approximation. In order to obtain all rules, backtracking functionality is required.
APRG: If τ satisfies $support(\tau) \geq \alpha$ and $accuracy(\tau) \geq \beta$, the implication τ is obtainable (with the exception of the redundant implications) by the Apriori algorithm. Therefore, completeness for the defined rules is ensured. However, the Apriori algorithm is time-consuming for the relatively lower threshold value α.

Both RSRG and APRG are useful rule generation methods although they are slightly different. In RSRG, most of the study focuses on determining the lower and upper approximations of X such as in Class III and IV in Table 1. However, the generation of two approximations does not directly obtain the set of rules. A gap exists between generating two approximations of X and rule generation. Therefore, we discriminate Classes III and IV with Classes V and VI in Table 1. In the study, we mainly investigate APRG.

3. Rule Generation in NISs

This section clarifies rule induction and Non-deterministic Information Systems (NISs) [21, 23] and then investigates the theoretical foundations of rules in NISs.

3.1. Rules in NISs

It is noted that NIS Φ is also a quadruplet $(OB, AT, \{VAL_A \mid A \in AT\}, g)$ where g denotes a mapping $g : OB \times AT \rightarrow P(\cup_{A \in AT} VAL_A)$ (a power set of $\cup_{A \in AT} VAL_A$) [21, 22]. Each set $g(x, A)$ is interpreted as the actual value in the set although the value is not known. We agree with that the introduction of the mapping g affords an approach to handle information incompleteness.

Remark 1. *In the Mammographic data set [9], there are several ? symbols. For example, the tuple of object 7 is (assessment=4,age=70,shape=?, margin=?,density=3,severity=0). The attribute value of shape is missing, and the domain of shape is* $\{round, oval, lobular, irregular\}$. *In this case, it is observed that* $g(7, shape)=\{round, oval, lobular, irregular\}$ *and a value in* $g(7, shape)$ *denotes the actual value due to the interpretation of the mapping* g *in NIS.*

Table 3: Exemplary NIS Φ_{salary}.

object	*age*	*depart(ment)*	*smoke*	*salary*
$x1$	$\{young\}$	$\{first\}$	$\{yes\}$	$\{low\}$
$x2$	$\{young, senior\}$	$\{first, second, third\}$	$\{yes\}$	$\{low\}$
$x3$	$\{senior\}$	$\{second\}$	$\{yes, no\}$	$\{high\}$
$x4$	$\{young, senior\}$	$\{second\}$	$\{no\}$	$\{high\}$
$x5$	$\{young\}$	$\{first, second, third\}$	$\{yes, no\}$	$\{high\}$
$x6$	$\{senior\}$	$\{third\}$	$\{no\}$	$\{high\}$

Table 3 is an exemplary NIS Φ_{salary}. With respect to Φ=$(OB, AT, \{VAL_A \mid A \in AT\}, g)$, we term each DIS ψ=$(OB, AT, \{VAL_A \mid A \in AT\}, h)$ satisfying $h(x, A) \in g(x, A)$ as a *derived DIS* from NIS Φ [32]. In Φ_{salary}, there are 144 (=$2^4 \times 3^2$) derived DISs. Let $DD(\Phi)$ denote a set of all derived DISs from NIS Φ.

Remark 2. *With respect to NIS Φ and $DD(\Phi)$, we observe (or assume that) that a DIS in $DD(\Phi)$ stores actual information, and we term it an unknown actual DIS ψ^{actual}. Without any additional information, a method to select $\psi^{actual} \in DD(\Phi)$ from Φ is absent due to information incompleteness.*

The concept of ψ^{actual} was introduced in incomplete information data bases [17], and we follow the same. Under the situation, we define the following.

Definition 2. *[32, 34] Two new rules are given in the following.*
(1) An implication τ denotes a certain rule, if τ denotes a rule in each $\psi \in DD(\Phi)$.
(2) An implication τ denotes a possible rule, if τ denotes a rule in at least a $\psi \in DD(\Phi)$.
(3) Let $CER(\Phi)$ denote the set of certain rules, and let $POS(\Phi)$ denote the set of possible rules.

3.2. Meaning of Certain Rules and Possible Rules

Evidently, a certain rule is also a possible rule, and the certainty and possibility follow the logical framework of possible world semantics [14, 17]. We observe NIS Φ, where $g(x, A)$ denotes a singleton set for each x and A, is DIS ψ. In this case, $DD(\Phi)$=$\{\psi\}$ holds, and thus the definitions of

certain rule and possible rule are identical. Therefore, Definition 2 includes the definition of rules in DISs, and Definition 2 denotes a natural extension from the rules in DISs. In a manner similar to that in Section 2.4, we clarify the meaning of certain rules and possible rules.

Proposition 3. *With respect to $\psi \in DD(\Phi)$, the threshold values α and β, and $Rule(\psi)$ in Algorithm 1, the following holds.*
(1) $CER(\Phi)$=$\cap_{\psi \in DD(\Phi)} Rule(\psi)$,
(2) $POS(\Phi)$=$\cup_{\psi \in DD(\Phi)} Rule(\psi)$,
(3) $CER(\Phi) \subseteq Rule(\psi^{actual}) \subseteq POS(\Phi)$.
(Proof)
(1) For any $\tau \in CER(\Phi)$, τ is an element in $Rule(\psi)$ for each ψ, so $\tau \in \cap_{\psi \in DD(\Phi)} Rule(\psi)$. The converse also holds.
(2) For any $\tau \in POS(\Phi)$, τ is an element in $Rule(\psi)$ at least one ψ, so $\tau \in \cup_{\psi \in DD(\Phi)} Rule(\psi)$. In this time, the converse also holds.
(3) Since $\cap_{\psi \in DD(\Phi)} Rule(\psi) \subseteq Rule(\psi^{actual}) \subseteq \cup_{\psi \in DD(\Phi)} Rule(\psi)$ clearly holds, the inclusion relation is derived.

Remark 3. *Based on (3) in Proposition 3, the set $CER(\Phi)$ is a lower approximation of $Rule(\psi^{actual})$ and the set $POS(\Phi)$ is an upper approximation of $Rule(\psi^{actual})$, namely we consider lower and upper approximations of $Rule(\psi^{actual})$. Although there are a few frameworks of rule generation with incomplete information, most studies attempt to cover a set X of objects via a few discernible classes and not equivalence ones. In the definition of such classes, a few assumptions are typically introduced [11, 15], and thus the validity of rules is characterised by the validity of the assumptions. Conversely, our purpose involves characterizing the rules that 'certainly hold' or 'possibly hold' in $\psi^{actual} \in DD(\Phi)$. The concept of ψ^{actual} is from Lipski's incomplete information databases [17], and to the best of the authors' knowledge, a framework of rule generation based on $Rule(\psi^{actual})$ is absent, and thus a framework is proposed in the current study. Therefore, rule generation in our framework significantly differs from the previous framework of rule generation with incomplete information.*

In Φ_{salary} in Table 3, there are 144 (=$3^2 \times 2^4$) derived DISs, and there exists $\psi^{actual} \in DD(\Phi_{salary})$. With respect to the threshold values α=0.3 and β=0.6, there are three certain rules and eight possible rules in Figure 1, and we conclude the following three-way decisions.
(1) The rule in $CER(\Phi)$ is also a rule in the unknown ψ^{actual},

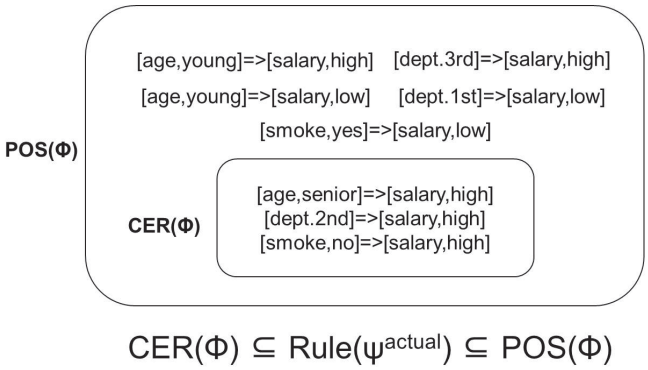


Figure 1: $CER(\Phi_{salary})$ is the lower approximation of $Rule(\psi^{actual})$ and $POS(\Phi_{salary})$ is the upper approximation of $Rule(\psi^{actual})$ for α=0.3 and β=0.6.

(2) The rule in $POS(\Phi) \setminus CER(\Phi)$ *is potentially a rule in the unknown* ψ^{actual},
(3) Any other implication is not a rule in the unknown ψ^{actual}.
Intuitively, each certain rule is the most reliable rule. Conversely, each possible rule is a rule in the unknown DIS ψ^{actual}. *All other implications with the exception of the possible rules are not a rule in the unknown DIS* ψ^{actual}. *Similarly, the framework of the certain rules and possible rules belongs to the three-way decisions [45].*

3.3. Computational Problem of Rule Generation in NISs and Its Solution

Although Definition 2 appears natural, we face the computational problem in rule generation in NISs. It applies APRG to each $\psi \in DD(\Phi)$, although the amount of the elements in $DD(\Phi)$ increases exponentially. For example, there are more than 10^{100} derived DISs for the Mammographic data set and the Hepatitis data set in the UCI machine learning repository [9]. Therefore, it is difficult to sequentially apply APRG to each $\psi \in DD(\Phi)$, and we propose another method to handle certain rules and possible rules.

Definition 3. *[32] In NIS* Φ *with a mapping* g, *two types of the granules* inf *and* sup *are defined as follows.*
(1) With respect to a descriptor $[A, val]$,

$inf([A, val])$=$\{x : object \mid g(x, A) = \{val\}\}$,
$sup([A, val])$=$\{x : object \mid val \in g(x, A)\}$.
(2) With respect to a conjunction $\wedge_{A \in CON}[A, val_A]$ *of descriptors,*
$inf(\wedge_{A \in CON}[A, val_A])$=$\cap_{A \in CON} inf([A, val_A])$,
$sup(\wedge_{A \in CON}[A, val_A])$=$\cap_{A \in CON} sup([A, val_A])$.

If $g(x, A)$ denotes a singleton set for each x and A, we see Φ denotes a DIS and two types of granules inf and sup define the same set. The set corresponds to an equivalence class in DIS. In NIS Φ, each equivalence class is extended to granules inf and sup. The unknown actual equivalence class $eq([A, val])$ satisfies $inf([A, val]) \subseteq eq([A, val]) \subseteq sup([A, val])$.

Definition 4. *[32] With respect to NIS* Φ *and an implication* τ*, we define the following.*
(1) $minsupp(\tau)$=$\min_{\psi \in DD(\Phi)}\{support(\tau)\ in\ \psi\}$,
(2) $minacc(\tau)$=$\min_{\psi \in DD(\Phi)}\{accuracy(\tau)\ in\ \psi\}$,
(3) $maxsupp(\tau)$=$\max_{\psi \in DD(\Phi)}\{support(\tau)\ in\ \psi\}$,
(4) $maxacc(\tau)$=$\max_{\psi \in DD(\Phi)}\{accuracy(\tau)\ in\ \psi\}$.

Lemma 1. *With respect to two natural numbers* P *and* Q $(P \geq Q)$, $\frac{Q}{P} \leq \frac{Q+1}{P+1}$ *holds.*

Proposition 4. *With respect to NIS* Φ *and an implication* $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$*, the following holds.*

$$\begin{array}{l} minsupp(\tau) = |inf(\wedge_{A \in CON}[A, val_A]) \cap inf([Dec, val])|/|OB|, \\ minacc(\tau) = \frac{|inf(\wedge_{A \in CON}[A, val_A]) \cap inf([Dec, val])|}{|inf(\wedge_{A \in CON}[A, val_A])| + |OUTACC|}, \\ OUTACC = \{sup(\wedge_{A \in CON}[A, val_A]) \setminus inf(\wedge_{A \in CON}[A, val_A])\} \\ \qquad \setminus inf([Dec, val]). \end{array} \tag{1}$$

Furthermore, a derived DIS $\psi_{min} \in DD(\Phi)$ *exists where* $support(\tau)$ *(in* ψ_{min}*)*=$minsupp(\tau)$ *and* $accuracy(\tau)$ *(in* ψ_{min}*)*=$minacc(\tau)$.
(Proof) The proof is given in Appendix A.

Proposition 5. *With respect to NIS* Φ *and an implication* $\tau : \wedge_{A \in CON}[A, val_A] \Rightarrow [Dec, val]$*, the following holds.*

$$\begin{array}{l} maxsupp(\tau) = |sup(\wedge_{A \in CON}[A, val_A]) \cap sup([Dec, val])|/|OB|, \\ maxacc(\tau) = \frac{|sup(\wedge_{A \in CON}[A, val_A]) \cap sup([Dec, val])|}{|inf(\wedge_{A \in CON}[A, val_A])| + |INACC|}, \\ INACC = \{sup(\wedge_{A \in CON}[A, val_A]) \setminus inf(\wedge_{A \in CON}[A, val_A])\} \\ \qquad \cap sup([Dec, val]). \end{array} \tag{2}$$

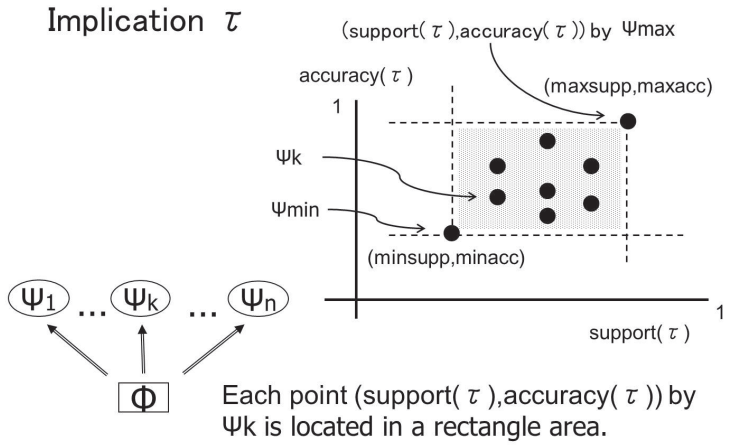


Figure 2: Each point $(support(\tau), accuracy(\tau))$ by $\psi \in DD(\Phi)$ is located in a rectangle area.

Furthermore, a derived DIS $\psi_{max} \in DD(\Phi)$ *exists where* $support(\tau)$ *(in* ψ_{max}*)=*$maxsupp(\tau)$ *and* $accuracy(\tau)$ *(in* ψ_{max}*)=*$maxacc(\tau)$*.*
(Proof) The proof is given in Appendix B.

Based on Propositions 4 and 5, each point $(support(\tau), accuracy(\tau))$ by $\psi \in DD(\Phi)$ is located in a rectangle area in Figure 2, and this leads to the following theorem.

Theorem 1. *With respect to an implication* τ*, the following holds.*
(1) τ *denotes a certain rule, if and only if* τ *denotes a rule in* ψ_{min}*, i.e.,* $minsupp(\tau) \geq \alpha$ *and* $minacc(\tau) \geq \beta$*.*
(2) τ *denotes a possible rule, if and only if* τ *denotes a rule in* ψ_{max}*, i.e.,* $maxsupp(\tau) \geq \alpha$ *and* $maxacc(\tau) \geq \beta$*.*
(3) Rule generation by (1) and (2) as mentioned above does not depend on the number of elements in $DD(\Phi)$*.*
(Proof) The proof is given in Appendix C.

Example 2. *We consider the same condition in Figure 1 and apply Proposition 4 to* $\tau : [age, senior] \Rightarrow [salary, high]$*. We assume that* $accuracy(\tau) = |NUME|/|DENO|$ *(*$NUME$ *denotes a set of objects that satisfy* τ*, and* $DENO$ *denotes a set of objects that satisfy* $[age, senior]$*. Both* $NUME$ *and*

$DENO$ are affected by information incompleteness in Φ_{salary}.). In order to reduce accuracy(τ), the minimum $NUME$ and maximum $DENO$ are considered. Given that $inf([age, senior])=\{x3, x6\}$ and $sup([age, senior])=\{x2, x3, x4, x6\}$, then $\{x3, x6\} \subseteq DENO \subseteq \{x2, x3, x4, x6\}$, and thus we adjust $x2, x4 \in (sup([age, senior]) \setminus inf([age, senior]))$.
(i) With respect to $x2$, since $inf([salary, high])=\{x3, x4, x5, x6\}$, $OUTACC=\{x2, x4\} \setminus \{x3, x4, x5, x6\}=\{x2\}$. Thus, it generates $[age, senior] \Rightarrow [salary, low]$ (the same condition value and different decision value) from $x2$. Hence, $x2 \in DENO$ and $x2 \notin NUME$, and the selection of the attribute value affects decreasing accuracy(τ).
(ii) With respect to $x4$, either τ or $[age, young] \Rightarrow [salary, high]$. If we consider the case in which τ holds, then $x4 \in DENO$ and $x4 \in NUME$. This indicates that accuracy(τ) is increased by Lemma 1. Thus, we consider another case, i.e., the case of $[age, young] \Rightarrow [salary, high]$ from $x4$, and $x4 \notin DENO$ and $x4 \notin NUME$.
Similarly, $DENO=\{x2, x3, x6\}$, and $NUME=inf([age, senior]) \cap inf([salary, high])=\{x3, x6\} \cap \{x3, x4, x5, x6\}=\{x3, x6\}$. With respect to the $DENO$ and $NUME$, accuracy(τ)=2/3, and the value is minacc(τ). Given that $NUME$ also corresponds to the minimum set for τ, the value support(τ)=2/6 is minsupp(τ). The above total calculation is performed via formula (1) in Proposition 4. Simultaneously, a few attribute values in Table 4 are specified.

Table 4: Specified attribute values (underlined part) related to τ : $[age, senior] \Rightarrow [salary, high]$ in Φ_{salary}. The table minimises both *support*(τ) and *accuracy*(τ).

object	*age*	*depart(ment)*	*smoke*	*salary*
$x1$	$\{young\}$	$\{first\}$	$\{yes\}$	$\{low\}$
$x2$	$\{young, \underline{senior}\}$	$\{first, second, third\}$	$\{yes\}$	$\{low\}$
$x3$	$\{senior\}$	$\{second\}$	$\{yes, no\}$	$\{high\}$
$x4$	$\{\underline{young}, senior\}$	$\{second\}$	$\{no\}$	$\{high\}$
$x5$	$\{young\}$	$\{first, second, third\}$	$\{yes, no\}$	$\{high\}$
$x6$	$\{senior\}$	$\{third\}$	$\{no\}$	$\{high\}$

4. NIS-Apriori-Based Rule Generation

In this section, *NIS-Apriori-based Rule Generation* in NISs are investigated. The NIS-Apriori algorithm (Algorithm 2) consists of two phases, i.e.,

Algorithm 2 NIS-Apriori algorithm

Input: NIS Φ, the decision attribute Dec, the threshold values α and β.
Output: Two sets $CRule(\Phi)$ (Certain rules) and $PRule(\Phi)$ (Possible rules).

$CRule(\Phi) \leftarrow \{\}$, $i \leftarrow 1$
create IMP_i, where each $\tau_{i,j} \in IMP_i$ satisfies $minsupp(\tau_{i,j}) \geq \alpha$
while $(|IMP_i| \geq 1)$ **do**
    $Rest \leftarrow \{\}$
    for all $\tau_{i,j} \in IMP_i$ **do**
        if $minacc(\tau_{i,j}) \geq \beta$ **then**
            add $\tau_{i,j}$ to $CRule(\Phi)$
        else
            add $\tau_{i,j}$ to $Rest$
        end if
    end for
    $i \leftarrow i + 1$
    generate IMP_i via $Rest$ and Property (2) in Section 2.3, where $\tau_{i,j} \in IMP_i$ satisfies
        (i) $minsupp(\tau_{i,j}) \geq \alpha$, and
        (ii) $\tau_{i,j}$ is not a redundant implication for any implication in $CRule(\Phi)$
end while
return $CRule(\Phi)$
$PRule(\Phi) \leftarrow \{\}$ ▷ $PRule(\Phi)$ generation follows $CRule(\Phi)$ generation. In this case, $CRule(\Phi)$, $minsupp(\tau_{i,j})$, and $minacc(\tau_{i,j})$ are replaced with $PRule(\Phi)$, $maxsupp(\tau_{i,j})$, and $maxacc(\tau_{i,j})$. The other program is the same, and thus we omit the same part.
return $PRule(\Phi)$

the $CRule(\Phi)$ generation phase and the $PRule(\Phi)$ generation phase. The first phase employs the criterion values $minsupp$ and $minacc$ in Proposition 4, and the second phase employs the criterion values $maxsupp$ and $maxacc$ in Proposition 5. We determine that the time complexity of the NIS-Apriori algorithm as exceeding twice the complexities of the Apriori algorithm. However, the NIS-Apriori algorithm does not depend on the amount of derived DISs as shown in Theorem 1. This indicates that the NIS-Apriori algorithm is applicable to NIS with a large amount of derived DISs.

Proposition 6. *The NIS-Apriori algorithm is sound and complete for the defined certain and possible rules, namely $CER(\Phi)$ and $POS(\Phi)$ in Definition 2 are equal to $CRule(\Phi)$ and $PRule(\Phi)$ in Algorithm 2, respectively.*
(Proof)
(Soundness: $CRule(\Phi) \subseteq CER(\Phi)$ and $PRule(\Phi) \subseteq POS(\Phi)$) Each implication $\tau \in CRule(\Phi)$ satisfies $minsupp(\tau) \geq \alpha$ and $minacc(\Phi) \geq \beta$, so $\tau \in CER(\Phi)$. As for the possible rules, the same is concluded.
(Completeness: $CER(\Phi) \subseteq CRule(\Phi)$ and $POS(\Phi) \subseteq PRule(\Phi)$) In Algorithm 2, any implication τ (except the redundant case like Algorithm 1) is included in IMP_i, and the condition of τ is examined. So, there is no implication τ satisfying the condition $minsupp(\tau) \geq \alpha$, $minacc(\tau) \geq \beta$ and $\tau \notin CRule(\Phi)$. The same is concluded for the possible rules.

We agreed with the framework of NISs and we consider that data analysis on NISs is a solution to process information incompleteness in DISs. Definition 2 extends the rules in DISs to the certain rules and possible rules in NISs and the extension is natural in the logical framework of possible world semantics. However, we are faced with the computational problem of handling the aforementioned rules. Although it applies the APRG to each $\psi \in DD(\Phi)$, there is a large number of derived DISs. There are more than 10^{100} derived DISs for the Mammographic data set and the Hepatitis data set. It is difficult to sequentially examine the condition of rules in each $\psi \in DD(\Phi)$.

With respect to the problem, the NIS-Apriori algorithm supported by Propositions 4 - 5 and Theorem 1 affords a solution. Propositions 4 and 5 show that it calculates four criterion values, namely $minsupp(\tau)$, $minacc(\tau)$, $maxsupp(\tau)$, and $maxacc(\tau)$, in polynomial time via the granules inf and sup for descriptors. Granules inf and sup follow the concept of the equivalence classes in rough sets and the concept of granular computing. We include the properties into the Apriori algorithm and proposed the NIS-Apriori algorithm.

5. NIS-Apriori Algorithm in SQL and the Prototype System

In this section, we focus on the implementation and the application of the NIS-Apriori Algorithm in SQL. Given that SQL exhibits high versatility, the environment yielded by NIS-Apriori-based rule generation in SQL is useful for table data analysis with three-way decisions. The section employs NIS Φ_{salary} in Table 3 to explain the execution.

The prototype system simulates NIS-Apriori algorithm, and consists of the following procedures in SQL. Currently, we handle rules with less than three condition attributes.
(1) The translation procedure $file\_nrdf$ from the csv format data set to the NRDF format data set,
(2) The $step1$ procedure generating the set of rules $\{\tau : [A, val_A] \Rightarrow [Dec, val]\}$,
(3) The $step2$ procedure generating the set of rules $\{\tau : [A, val_A] \wedge [B, val_B] \Rightarrow [Dec, val]\}$,
(4) The $step3$ procedure generating the set of rules $\{\tau : [A, val_A] \wedge [B, val_B] \wedge [C, val_C] \Rightarrow [Dec, val]\}$.

It is necessary to prepare a procedure in SQL for a csv format table data set although we uniformly handle the NRDF format data set after the translation. The prototype system is implemented on note PC and desktop PC via the phpMyAdmin tool. The program is implemented by the procedures in SQL, and thus we easily execute the prototype system on any PC with SQL [35, 36].

5.1. The NRDF Format

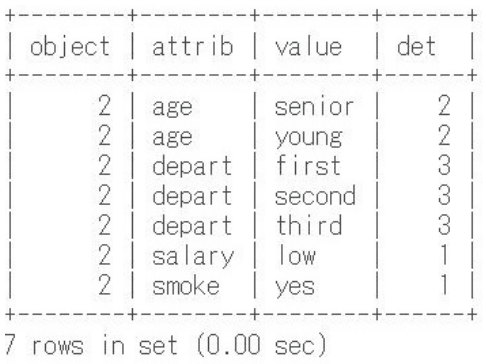
The csv format is often employed as an expression of table data sets. This is very familiar although it is necessary to prepare a program for a csv format data set because the amount of all attributes, the name of each attribute, and the attribute values are typically different. In order to handle various types of datasets, it is necessary to apply a unified format.

We apply the NRDF format [40] that corresponds to the extended RDF (resource description framework) format. The RDF format is observed as the EAV (entity-attribute-value) format [13]. Decision tree induction and the KDD-related tasks of attribute selection are implemented via the EAV format [13].

The NRDF format consists of four attributes, namely *object*, *attrib*, *value*, and *det*. In Figure 3, the NRDF format expression of the object 2 in Φ_{salary} is shown. Evidently, data in the NRDF format indicates a set of descriptors. With respect to specifying non-deterministic information, the 4th column *det* is added. The value of *det* denotes the amount of possible values. If *det*=1, the value is deterministic.

5.2. Procedures step1, step2, and step3

In Step 1, the procedure $step1$ generates the certain and possible rules in the form of $P_1 \Rightarrow Dec$. The procedure consists of the following subparts.



object	attrib	value	det
2	age	senior	2
2	age	young	2
2	depart	first	3
2	depart	second	3
2	depart	third	3
2	salary	low	1
2	smoke	yes	1

7 rows in set (0.00 sec)

Figure 3: Expression of object 2 in the NRDF format of Φ_{salary}.

1. Generation of a table *condi* (the condition of the rule generation),
2. Generation of a table *con_des* (the descriptors for the condition),
3. Generation of a table *dec_des* (the descriptors for the decision),
4. Generation of a few tables for implications, inf, *inacc*, and *outacc*),
5. Generation of a table $c1\_rule$
(certain rules satisfying $minsupp(\tau) \geq \alpha$ and $minacc(\tau) \geq \beta$),
6. Generation of a table $c1\_rest$
(implications satisfying $minsupp(\tau) \geq \alpha$ and $minacc(\tau) < \beta$),
7. Generation of a table $p1\_rule$
(possible rules satisfying $maxsupp(\tau) \geq \alpha$ and $maxacc(\tau) \geq \beta$),
8. Generation of a table $p1\_rest$
(implications satisfying $maxsupp(\tau) \geq \alpha$ and $maxacc(\tau) < \beta$).

In Step 2, the procedure *step*2 generates the certain and possible rules in the form of $P_1 \wedge P_2 \Rightarrow Dec$. Since $support(P_1 \wedge P_2 \Rightarrow Dec) \leq support(P_1 \Rightarrow Dec)$ holds, it is not necessary to handle $\tau \notin IMP_1$ in Algorithm 2. As shown in Step 1, a table $c1\_rest \subseteq IMP_1$ exists in Algorithm 2, and it is sufficient to consider the implications $P_1 \wedge P_2 \Rightarrow Dec$ that satisfy $(P_1 \Rightarrow Dec), (P_2 \Rightarrow Dec) \in c1\_rest$ in certain rule generation (the Property (2) in Section 2.3). The set IMP_2 in Algorithm 2 is generated by the table $c1\_rest$, and each implication in IMP_2 is examined to obtain the certain rules. In a manner similar to the tables $c1\_rest$ and $p1\_rest$, $c2\_rest$ and $p2\_rest$ are generated for Step 3.

Figures 4 and 5 show the execution log for the Mammographic data set Φ_{mammo} and the obtained certain rules. We also executed the NIS-Apriori

```
mysql> show tables;
+---------------------------+
| Tables_in_nrdf(mammo_test) |
+---------------------------+
| nrdf                      |
| table 1                   |
+---------------------------+
2 rows in set (0.00 sec)

mysql> call step1('severity',960,0.1,0.8);
Query OK, 0 rows affected (30.06 sec)

mysql> call step2('severity',960,0.1,0.8);
Query OK, 0 rows affected (1 min 28.81 sec)

mysql> call step3('severity',960,0.1,0.8);
Query OK, 0 rows affected (1 min 39.95 sec)
```

Figure 4: Execution for the Mammographic data set Φ_{mammo} in the command prompt.

```
mysql> select * from c1_rule;
+------------+------+----------+------------+---------+--------+
| att1       | val1 | deci     | deci_value | minsupp | minacc |
+------------+------+----------+------------+---------+--------+
| assess     | 5    | severity | 1          |   0.317 |  0.869 |
| margin     | 1    | severity | 0          |   0.329 |  0.859 |
| end_attrib | NULL | NULL     | NULL       |    NULL |   NULL |
+------------+------+----------+------------+---------+--------+
3 rows in set (0.00 sec)

mysql> select * from c2_rule;
+------------+------+--------+------+----------+------------+---------+--------+
| att1       | val1 | att2   | val2 | deci     | deci_value | minsupp | minacc |
+------------+------+--------+------+----------+------------+---------+--------+
| age        | 50   | assess | 4    | severity | 0          |   0.113 |  0.800 |
| age        | 70   | shape  | 4    | severity | 1          |   0.101 |  0.843 |
| assess     | 4    | shape  | 1    | severity | 0          |   0.174 |  0.884 |
| assess     | 4    | shape  | 2    | severity | 0          |   0.163 |  0.876 |
| end_attrib | NULL | NULL   | NULL | NULL     | NULL       |    NULL |   NULL |
+------------+------+--------+------+----------+------------+---------+--------+
5 rows in set (0.00 sec)

mysql> select * from c3_rule;
+------------+------+------+------+------+------+------+------------+---------+--------+
| att1       | val1 | att2 | val2 | att3 | val3 | deci | deci_value | minsupp | minacc |
+------------+------+------+------+------+------+------+------------+---------+--------+
| end_attrib | NULL | NULL | NULL | NULL | NULL | NULL | NULL       |    NULL |   NULL |
+------------+------+------+------+------+------+------+------------+---------+--------+
1 row in set (0.00 sec)
```

Figure 5: Obtained certain rules from the Mammographic data set Φ_{mammo}. The obtained certain rules satisfy $support(\tau) \geq 0.1$ and $accuracy(\tau) \geq 0.8$ in each of more than 10^{100} derived DISs.

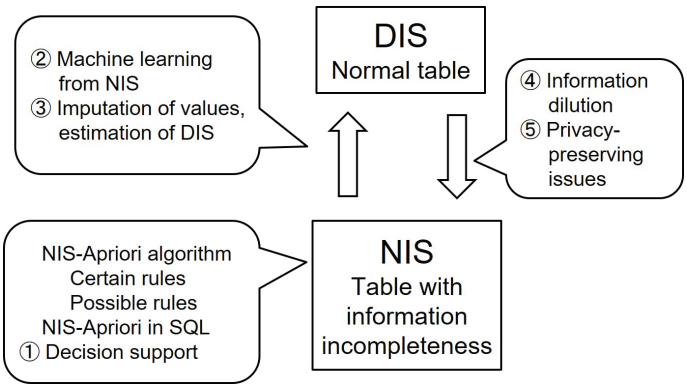


Figure 6: New topics under the background of NIS-Apriori-based rule generation.

previously implemented in Prolog and C, and compared rules by Prolog with rules by SQL corresponding to approximately 10 small size data sets. Thus, two systems generated the same rules with the exception of the redundant implications. This assures the validity of the implementation. The results on the Mammographic data set and the Flu data set by NIS-Apriori in Prolog and NIS-Apriori in SQL are uploaded to the web page [36].

6. New Topics Related to NIS-Apriori-Based Rule Generation

This section considers the possibility of applying NIS-Apriori-based rule generation. We discuss new topics in Figure 6 under the background of NIS-Apriori-based rule generation.
(1) In ① in Figure 6, it applies the certain rules and the possible rules to decision support based on tables with information incompleteness.
(2) In ②, it considers a possibility wherein we sequentially add additional information to obtain the actual unknown DIS ψ^{actual}.
(3) In ③, it considers the possibility wherein we impute missing values by using the obtained rules. In this case, additional information is not necessary.
(4) In ④, it considers the possibility wherein we intentionally replace original information with non-deterministic information to preserve a few constraints. This is a method to hide original information. We term this as *information dilution*. We handle the constraint wherein each rule must be obtained as a possible rule [33].
(5) In ⑤, it applies non-deterministic information to privacy-preserving data

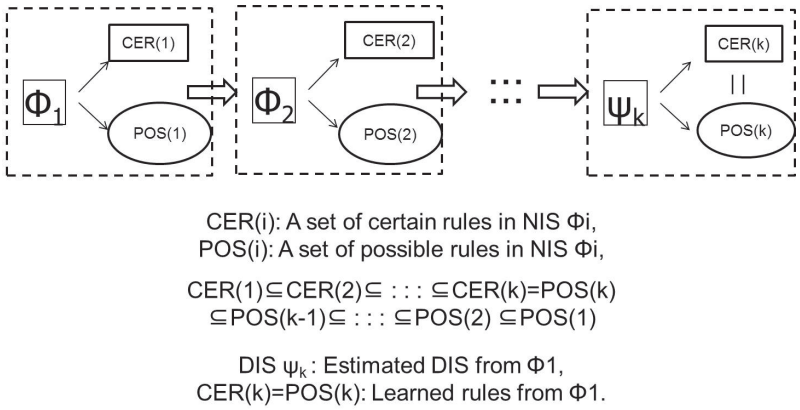


Figure 7: Process of machine learning from NIS Φ_1 (=Φ).

mining.

Here, we focus on ② in Figure 6 and clarify mathematical properties. We consider a sequence of NISs from NIS Φ (=Φ_1).

Definition 5. *With respect to i=1, 2, $\cdots$, we assume that* Φ_i=$(OB, AT, \{VAL_A \mid A \in AT\}, g_i)$, *and* Φ_{i+1}=$(OB, AT, \{VAL_A \mid A \in AT\}, g_{i+1})$ *satisfy* $g_{i+1}(x, A) \subseteq g_i(x, A)$ *for any* $x \in OB$ *and any* $A \in AT$. *This leads to NISs,* Φ_1, Φ_2, $\cdots$, Φ_n. *We term the NISs a sequence from* Φ_1. *Specifically, if* Φ_1=Φ, *then we assume that this denotes a sequence from* Φ. *If* $g_i(x, A)$ *denotes a singleton set for any* $x \in OB$ *and any* $A \in AT$, *then it is considered that* Φ_i *denotes a DIS* ψ_i.

Intuitively, it is observed that Φ_{i+1} is obtained by fixing non-deterministic information in Φ_i, and thus Φ_{i+1} is more informative than Φ_i. With respect to the aforementioned Φ_{i+1} and Φ_i, the following properties are obtained:
(1) In a sequence from Φ, $CER(\Phi_i) \subseteq POS(\Phi_i)$ holds in every Φ_i,
(2) In a sequence from Φ, $CER(\Phi_i) \subseteq CER(\Phi_{i+1})$ holds,
(3) In a sequence from Φ, $POS(\Phi_{i+1}) \subseteq POS(\Phi_i)$ holds.
Therefore, with respect to the fixed threshold values α and β, the following inclusion relation in a sequence from Φ_1 is obtained. The uncertainty is sequentially reduced, and we finally obtain a DIS ψ_k.

In Figure 7, the revision from Φ_i to Φ_{i+1} by the obtained rules exhibits a few variations. In the actual application, it is necessary to define how Φ_{i+1} is obtained from Φ_i and its obtained rules.

7. Trial: Estimation of Unknown Actual DIS ψ^{actual}

In this section, a plausible method to estimate the unknown actual DIS ψ^{actual} without any additional information is proposed. This topic is shown in Figure 6. We consider the imputation of missing values by the obtained certain rules. Although statistical methods are typically employed in the imputation of missing values [8, 30, 50], we employ the obtained certain rules. With respect to the side effect, we estimate the unknown actual DIS ψ^{actual} and its rules.

7.1. Imputation of Missing Values and Estimation of DIS ψ^{actual}

We employ the results in Section 6 and consider estimating DIS ψ^{actual}. The problem involves sequentially reducing the amount of non-deterministic information. With respect to this problem, we apply the obtained certain rules because the certain rules are the most reliable. We propose a strategy in which '*a value of non-deterministic information is fixed to create the maximum possible number of certain rules* '. The details are given below:
(Strategy 1) (Positive Unification) *In an object x, a value is assigned to the unfixed value to create as many higher ordered certain rules as possible.*
(Strategy 2) (Contradiction Prevention) *In an object x, a value is assigned to not contradict the higher ordered certain rule.*

The strategies also assume a role similar to that of *maximum likelihood estimation* in statistics [1]. Each parameter is estimated to maximise the likelihood function, and each missing value is imputed to create the higher ordered certain rules.

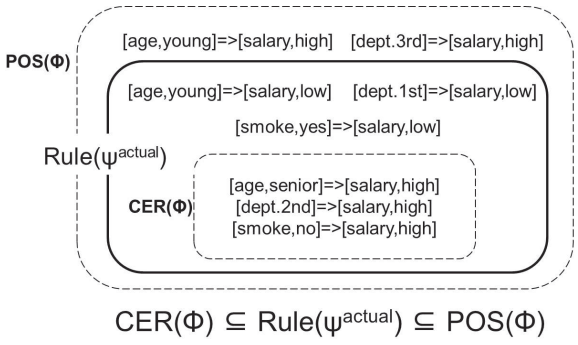


Figure 8: Relation between $CER(\Phi_{salary})$, $POS(\Phi_{salary})$ in Figure 1 and the estimated $Rule(\psi^{actual})$.

Algorithm 3 Algorithm to estimate ψ^{actual}

Input: NIS Φ, the decision attribute Dec, the threshold values $\alpha_1 \geq \alpha_2 \geq \cdots$ and $\beta_1 \geq \beta_2 \geq \cdots$.

Output: The estimated DIS ψ^{actual}.

1: generate a table $nrdf_1$ (the NRDF format table of Φ)
2: $i \leftarrow 1$
3: **while** (there is such a tuple that $det > 1$ in $nrdf_i$) **do**
4: generate a set $CER(\Phi_i)$ of certain rules $(\tau : [A, val_A] \Rightarrow [Dec, val])$ satisfying $minsupp(\tau) \geq \alpha_i$ and $minacc(\tau) \geq \beta_i$ ▷ We sequentially employ lower α_i and β_i values to obtain new certain rules
5: apply two strategies in Section 7.1 and generate $nrdf_{i+1}$ via $CER(\Phi_i)$ and $nrdf_i$
6: $i \leftarrow i + 1$
7: **end while**
8: generate ψ^{actual} via $nrdf_i$
9: **return** ψ^{actual}

7.2. Prototype System and Example of the Execution

This subsection describes the system to estimate DIS $\psi^{actual} \in DD(\Phi)$, and presents an example of the execution via Φ_{salary} in Table 3. As shown in Figure 1, $CER(\Phi_{salary})$ consists of three certain rules. Although we implemented procedures *step*1, *step*2, and *step*3 for NIS-Apriori-based rule generation, we currently employ the certain rules in the form of $[A, val_A] \Rightarrow [Dec, val]$ by the procedure *step*1 (Algorithm 3). In the actual execution, we obtain the following inclusion relation, and the set of estimated rules is given $Rule(\psi^{actual})$ in Figure 8.

$$\begin{aligned} &CER(\Phi_{salary,1}) \subseteq CER(\Phi_{salary,2}) \subseteq CER(\psi_{salary,3}) \\ &= POS(\psi_{salary,3}) \subseteq POS(\Phi_{salary,2}) \subseteq POS(\Phi_{salary,1}). \end{aligned} \tag{3}$$

7.3. Evaluation of the Proposed Imputation

We showed the possibility of applying the obtained certain rules to impute missing values. Our imputation exhibits the following properties.
(1) The proposed imputation is an application of the obtained rules. Based on [8, 30, 50], most imputation studies appear to depend on statistical functions and error functions. Conversely, the proposed imputation depends on data

dependency between attributes and not statistical functions.
(2) The proposed imputation does not require any additional information in a manner similar to the distribution functions and the error functions. It is sufficient to specify two threshold values α and β in each step.
(3) The total procedure in Figure 7 depends on NIS-Apriori in SQL, and thus the proposed imputation is not considered yet. Without the certain rules, it does not consider the imputation.
(4) It realises the software tool easily as an application of NIS-Apriori in SQL.

Remark 4. *We assume that* $[A, a] \Rightarrow [Dec, v]$ *is an obtained certain rule and that the degree of data dependency [24] from the attribute* A *to the attribute Dec is 1. Subsequently, we obtain the following.*
(1) If the mapping $h : VAL_A \to VAL_{Dec}$ *is one-to-one, then the missing values in* $[A, ?] \Rightarrow [Dec, v]$ *and* $[A, a] \Rightarrow [Dec, ?]$ *are correctly imputed. (Figure 9 shows that the mapping* $h : VAL_{age} \to VAL_{salary}$ *is one to one. In this table, any missing value that is randomly added is imputed correctly.)*

```
mysql> select * from rule1;
+------------+--------+--------+------------+---------+----------+
| att1       | val1   | deci   | deci_value | support | accuracy |
+------------+--------+--------+------------+---------+----------+
| age        | middle | salary | normal     |   0.250 |    1.000 |
| age        | senior | salary | high       |   0.350 |    1.000 |
| age        | young  | salary | low        |   0.400 |    1.000 |
| end_attrib | NULL   | NULL   | NULL       |    NULL |     NULL |
+------------+--------+--------+------------+---------+----------+
4 rows in set (0.00 sec)
```

Figure 9: The rules from ψ_{salary} indicate that the mapping is one-to-one.

(2) If the mapping $h : VAL_A \to VAL_{Dec}$ *is many-to-one, then the missing value in* $[A, a] \Rightarrow [Dec, ?]$ *is correctly imputed although the missing value in* $[A, ?] \Rightarrow [Dec, v]$ *is not correctly imputed.*

The highlight of the study involves clarifying the unique characteristics of NIS-Apriori-based rule generation, and the effective usage of the proposed imputation is currently in progress. However, we employ the Balloons data set [9] with less data dependency for an experiment. The data set is a DIS $\psi_{balloons}$, and it consists of 20 objects and 5 attributes $\{color, size, act, age, class\}$. Each attribute exhibits two attribute values. Figure 10 (we especially added

```
mysql> select * from rule1;
+--------+-----------+---------+-------+------------+---------+----------+
| ruleID | att1      | val1    | deci  | deci_value | support | accuracy |
+--------+-----------+---------+-------+------------+---------+----------+
|      1 | act       | DIP     | class | F          |   0.400 |    1.000 |
|      2 | act       | STRETCH | class | T          |   0.400 |    0.667 |
|      3 | age       | ADULT   | class | T          |   0.400 |    0.667 |
|      4 | age       | CHILD   | class | F          |   0.400 |    1.000 |
|      5 | color     | PURPLE  | class | F          |   0.300 |    0.600 |
|      6 | color     | YELLOW  | class | F          |   0.300 |    0.600 |
|      7 | size      | LARGE   | class | F          |   0.300 |    0.600 |
|      8 | size      | SMALL   | class | F          |   0.300 |    0.600 |
|      0 | end_attrib | NULL    | NULL  | NULL       |    NULL |     NULL |
+--------+-----------+---------+-------+------------+---------+----------+
9 rows in set (0.00 sec)
```

Figure 10: Obtained rules ($support(\tau) \geq 0.2$ and $accuracy(\tau) \geq 0.5$) in the Balloons data set $\psi_{balloon}$.

rule ID for the following explanation) shows the rules in $\psi_{balloons}$, and data dependency from any attribute to Dec=$class$ is absent. Based on Remark 4, each missing value in $[act, DIP] \Rightarrow [class, ?]$ and $[age, CHILD] \Rightarrow [class, ?]$ is correctly imputed via rule 1 and 4. However, the *accuracy* of rule 2: $[act, STRETCH] \Rightarrow [class, T]$ is 67%, and thus 33% of missing values by $[act, STRETCH] \Rightarrow [class, ?]$ are imputed incorrectly.

We followed the method in [8, 50] and randomly added missing values to $\psi_{balloons}$. Subsequently, we performed experiments in three cases.
(Case 1) In each condition attribute, two values (10%) were randomly changed to missing values. In Step 1, the certain rules 1, 4, 6, 5, 2, and 3 (ordered by *minacc* and *minsupp*, α=0.2, β=0.5) in Figure 10 were generated. In Step 2, the certain rules 1, 4, 5, 8, 2, 3, and 6 (α=0.2, β=0.5) were generated. In Step 3, the certain rules 1, 4, 5, 8, 2, 4, 6, 7, and four new implications (α=0.2, β=0.3) were generated. Six values were imputed correctly in which the recovery rate was 60% (=6/10). In other two experiments, the recovery rates were 60% and 70%, respectively. The total recovery rate was 63%.
(Case 2) In each attribute, five values (25%) were randomly changed to missing values. With respect to other two experiments, the recovery rates were 48% and 68%, respectively. In total, the recovery rate was 53%.
(Case 3) In each condition attribute, eight values (40%) were randomly changed to missing values. In the decision attribute, five values (25%) were randomly changed to missing values. The recovery rate was approximately 57% (=21/37). With respect to other two experiments, the recovery rates

were 38% and 62%, respectively. In total, the recovery rate was 52%.

We apply certain rules to impute the missing values. This is a topic in Figure 6, and our study focuses on the imputation as a preliminary trial.

8. Concluding Remarks and Discussion

In the study, we agreed with the framework of NISs, investigated NIS-Apriori-based rule generation, and then implemented prototype systems in SQL. The NIS-Apriori algorithm takes the core role of the proposed prototype, and the NIS-Apriori algorithm was the only practically applicable algorithm that could handle certain rules and possible rules based on possible world semantics. Thus, the NIS-Apriori-based rule generation afforded a new framework for table data analysis with three-way decisions.

In NIS-Apriori-based rule generation, we considered the sets $CER(\Phi)$ and $POS(\Phi)$ in Definition 2. The two sets correspond to the lower and upper approximations of $Rule(\psi^{actual})$ defined in the unknown actual DIS ψ^{actual}. The approximations are not for a set X of objects. An example is provided in Figure 1. Although our framework is different from that of the typical rough sets, and the lower approximation $CER(\Phi)$ and upper approximation $POS(\Phi)$ of $Rule(\psi^{actual})$ also follow the concept of approximations in rough sets.

The most important theoretic investigation is the solution of the computational problem for rule generation. The certain rules and possible rules depend on each $\psi \in DD(\Phi)$. There are more than 10^{100} derived DISs for the Mammographic data set and the Hepatitis data set, and thus it is difficult to sequentially examine the condition of rules in each $\psi \in DD(\Phi)$. We introduced two granules inf and sup to each descriptor, and proved formula (1) in Proposition 4 and formula (2) in Proposition 5. Based on the results, we calculate certain and possible rules in polynomial time. Theorem 1 summarises the solution for the computational problem.

In the implementation of the NIS-Apriori algorithm in SQL, we employed the NRDF format, and we handle any NISs after the translation from the csv format data set to the NRDF format data sets. We compared the result by NIS-Apriori in Prolog with the result by NIS-Apriori in SQL, and the results indicated that both results are the same with the exception of the redundant rules. In NIS-Apriori in Prolog, granules inf and sup were expressed by the lists. The lengths of the lists almost corresponded to the maximum in the practical Prolog system, and we observed that it was difficult to apply

NIS-Apriori in Prolog to the large data sets. The results suggested that the NIS-Apriori in SQL was applicable to large data sets as opposed to that in Prolog. The results and the log files were presented on the web page [36].

With respect to an application of NIS-Apriori in SQL, we considered a plausible method to estimate ψ^{actual} without any additional information. Each missing value was imputed to create the maximum possible obtained certain rules. This corresponded to a new attempt to impute missing values. In Figure 8, we show the estimated rules $Rule(\psi^{actual})$ from Figure 1. The validity of the estimated $Rule(\psi^{actual})$ depends on the reliable information by the certain rules. We are currently in the process of creating an actual application. It is necessary to add a few other constraints employed in [8, 50] to estimate missing values.

By presenting the contents, we revealed that NIS-Apriori-based rule generation is a significantly new framework and that it extends the research area of three-way decisions, rough sets, and granular computing. We also considered the application of NIS-Apriori-based rule generation to privacy-preserving data mining [2] by non-deterministic information. In NISs, we positively handled ambiguous information by non-deterministic information, and we intentionally dilutes the original value 'A' to 'either A or B' [33]. This was useful for data security, and NIS-Apriori-based rule generation handled even the diluted table data sets.

Acknowledgment: The authors would like to thank the editors and anonymous reviewers for their useful comments. The study was supported by JSPS (Japan Society for the Promotion of Science) KAKENHI Grant Number 26330277.

Appendix A. The proof of Proposition 4

We show the way to generate $\psi_{min} \in DD(\Phi)$. For p $(=\wedge_{A \in CON}[A, val_A])$ in the condition part of τ, we have three cases, i.e., $\{p\}$ (a singleton set with p), $\{p, \bar{p}\}$ (a set with p and at least one element $\bar{p} \neq p$), and $\{\bar{p}\}$ (a set without p). In two sets $\{p, \bar{p}\}$ and $\{\bar{p}\}$, there may be several elements for $\bar{p}$, however it is enough to consider a set with one $\bar{p}$. Then, any tuple in a table becomes one of $\{p\}$, $\{p, \bar{p}\}$, and $\{\bar{p}\}$. Similarly, we consider three cases for q $(=[Dec, val])$, i.e., $\{q\}$, $\{q, \bar{q}\}$ and $\{\bar{q}\}$, and we have nine sets in Table A.5.

In S_2, we may select either $p \Rightarrow q$ or $p \Rightarrow \bar{q}$, and then the related attribute values of the object is fixed for this selection. If we select $p \Rightarrow q$, the object is

Table A.5: Nine sets of the obtainable implications related to $\tau : p \Rightarrow q$.

	$\{q\}$	$\{q, \bar{q}\}$	$\{\bar{q}\}$
$\{p\}$	$S_1 : \{p \Rightarrow q\}$	$S_2 : \{p \Rightarrow q, p \Rightarrow \bar{q}\}$	$S_3 : \{p \Rightarrow \bar{q}\}$
$\{p, \bar{p}\}$	$S_4 : \{p \Rightarrow q,$ $\bar{p} \Rightarrow q\}$	$S_5 : \{p \Rightarrow q, p \Rightarrow \bar{q},$ $\bar{p} \Rightarrow q, \bar{p} \Rightarrow \bar{q}\}$	$S_6 : \{p \Rightarrow \bar{q},$ $\bar{p} \Rightarrow \bar{q}\}$
$\{\bar{p}\}$	$S_7 : \{\bar{p} \Rightarrow q\}$	$S_8 : \{\bar{p} \Rightarrow q, \bar{p} \Rightarrow \bar{q}\}$	$S_9 : \{\bar{p} \Rightarrow \bar{q}\}$

Table A.6: A selection of the implications from Table A.6.

	$\{q\}$	$\{q, \bar{q}\}$	$\{\bar{q}\}$
$\{p\}$	$S_1 : p \Rightarrow q$	$S_2 : p \Rightarrow \bar{q}$	$S_3 : p \Rightarrow \bar{q}$
$\{p, \bar{p}\}$	$S_4 : \bar{p} \Rightarrow q$	$S_5 : p \Rightarrow \bar{q}$	$S_6 : p \Rightarrow \bar{q}$
$\{\bar{p}\}$	$S_7 : \bar{p} \Rightarrow q$	$S_8 : \bar{p} \Rightarrow q$	$S_9 : \bar{p} \Rightarrow \bar{q}$

counted in both the denominator and the numerator of $accuracy(\tau)$. Based on Lemma 1, this selection causes $accuracy(\tau)$ to increase. So, in order to decrease $accuracy(\tau)$ we try to select the implication with 'the same condition values' and 'a different decision value' from the objects. We take the selection in Table A.6 from Table A.5, which makes the value of $accuracy(\tau)$ the minimum. Thus, $minacc(\tau)=|S_1|/(|S_1| + |S_2| + |S_3| + |S_5| + |S_6|)$ holds. Because $S_1=inf(p) \cap inf(q)$, $inf(p)=S_1 \cup S_2 \cup S_3$, and $S_5 \cup S_6=(sup(p) \setminus inf(p)) \setminus inf(q)$, we have the formula (1). At the same time, $support(\tau)$ is clearly the minimum, because $p \Rightarrow q$ does not occur except the definite case S_1. We have $minsupp(\tau)=|S_1|/|OB|=|inf(p) \cap inf(q)|/|OB|$.

Appendix B. The proof of Proposition 5

We similarly show the way to generate $\psi_{max} \in DD(\Phi)$. In order to increase $accuracy(\tau)$, we try to select the implication with 'the same condition values' and 'the same decision value' from the objects. We take the selection in Table B.7 from Table A.5, which makes the value of $accuracy(\tau)$ the maximum. At the same time, $support(\tau)$ is clearly the maximum, because $p \Rightarrow q$ occurs as many as possible.

Appendix C. The proof of Theorem 1

(1) If τ is a certain rule, τ is also a rule in ψ_{min}. Therefore, we have two

Table B.7: A selection of the implications from Table A.5.

	$\{q\}$	$\{q, \bar{q}\}$	$\{\bar{q}\}$
$\{p\}$	$S_1 : p \Rightarrow q$	$S_2 : p \Rightarrow q$	$S_3 : p \Rightarrow \bar{q}$
$\{p, \bar{p}\}$	$S_4 : p \Rightarrow q$	$S_5 : p \Rightarrow q$	$S_6 : \bar{p} \Rightarrow \bar{q}$
$\{\bar{p}\}$	$S_7 : \bar{p} \Rightarrow q$	$S_8 : \bar{p} \Rightarrow q$	$S_9 : \bar{p} \Rightarrow \bar{q}$

formulas, $\alpha \leq (support(\tau)$ in $\psi_{min})$ and $\beta \leq (accuracy(\tau)$ in $\psi_{min})$. Based on the definition of $minsupp(\tau)$ and $minacc(\tau)$, we have $\alpha \leq minsupp(\tau)$ and $\beta \leq minacc(\tau)$. On the other hand, $minsupp(\tau) \leq (support(\tau)$ in $\psi)$ and $minacc(\tau) \leq (accuracy(\tau)$ in $\psi)$. Therefore, if $\alpha \leq minsupp(\tau)$ and $\beta \leq minacc(\tau)$, we have $\alpha \leq (support(\tau)$ in $\psi)$ and $\beta \leq (accuracy(\tau)$ in $\psi)$. This means τ is a certain rule.
(2) We can similarly show that τ is a possible rule, if and only if $\alpha \leq (support(\tau)$ in $\psi_{max})$ and $\beta \leq (accuracy(\tau)$ in $\psi_{max})$.
(3) The calculations by Propositions 4 and 5 employ granules inf and sup. We do not enumerate any $\psi \in DD(\Phi)$. Namely, the detection of certain rules and possible rules does not depend upon the amount of elements in $DD(\Phi)$.

References

[1] J. Aldrich, R.A. Fisher and the making of maximum likelihood 1912–1922, Statistical Science 12(3) (1997) 162–176.

[2] C. Aggarwal, S. Yu, A general survey of privacy-preserving data mining models and algorithms, in: Privacy-Preserving Data Mining, Models and Algorithms (C. Aggarwal, S. Yu, Eds.), Springer, 2008, pp. 11–52.

[3] R. Agrawal, R. Srikant, Fast algorithms for mining association rules in large databases, in: Proc. VLDB'94 (J.B. Bocca, M. Jarke, C. Zaniolo, Eds.), Morgan Kaufmann, 1994, pp. 487–499.

[4] J. Bazan, M. Szczuka, The rough set exploration system, Transactions on Rough Sets 3 (2005) 37–56.

[5] A. Ceglar, J.F. Roddick, Association mining, ACM Computing Survey 38(2) (2006).

[6] D. Ciucci, T. Flaminio, Generalized rough approximations in PI 1/2, Int. Journal of Approximate Reasoning 48(2) (2008) 544–558.

[7] E.F. Codd, A relational model of data for large shared data banks, Communication of the ACM 13(6) (1970) 377–387.

[8] A. Farhangfar, L.A. Kurgan, W. Pedrycz, A novel framework for imputation of missing values in databases, IEEE Trans. Systems, Man, and Cybernetics, Part A 37(5) (2007) 692–709.

[9] A. Frank, A. Asuncion, UCI machine learning repository, Irvine, CA: University of California, School of Information and Computer Science, 2010. http://mlearn.ics.uci.edu/MLRepository.html

[10] S. Greco, B. Matarazzo, R. Słowiński, Granular computing and data mining for ordered data: The dominance-based rough set approach, in: Encyclopedia of Complexity and Systems Science (R.A. Meyers, Ed.), Springer, 2009, pp. 4283–4305.

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

[12] J. Komorowski, Z. Pawlak, L. Polkowski, A. Skowron, Rough sets: a tutorial, in: Rough Fuzzy Hybridization: A New Method for Decision Making (S. K. Pal, A. Skowron, Eds.), Springer, 1999, pp. 3–98.

[13] M. Kowalski, S. Stawicki, SQL-based heuristics for selected KDD tasks over large data sets, In: Proc. FedCSIS 2012 (M. Ganzha, L. Maciaszek, M. Paprzycki, Eds.), 2012, pp. 303–310.

[14] S.A. Kripke, Semantical considerations on modal logic, Acta Philosophica Fennica 16 (1963) 83–94.

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

[16] Y. Leung, M.M. Fischer, W.Z. Wu, J.S. Mi, A rough set approach for the discovery of classification rules in interval-valued information systems, Int. Journal of Approximate Reasoning 47(2) (2008) 233–246.

[17] W. Lipski, On semantic issues connected with incomplete information databases, ACM Transactions on Database Systems 4(3) (1979) 262–296.

[18] J.W. Lloyd, Foundations of Logic Programming, Springer, 1984, p. 124.

[19] M. Marek, Z. Pawlak, Information storage and retrieval systems: Mathematical foundations, Theoretical Computer Science 1(4) (1976) 331–354.

[20] M. Nakata, H. Sakai, Twofold rough approximations under incomplete information, Int. Journal General Systems 42(6) (2013) 546–571.

[21] E. Orłowska, Z. Pawlak, Representation of nondeterministic information, Theoretical Computer Science 29(1-2) (1984) 27–39.

[22] Z. Pawlak, Rough sets, Int. Journal of Computer and Information Sciences 11(5) (1982) 341–356.

[23] Z. Pawlak, Systemy informacyjne: Podstawy teoretyczne (in Polish), WNT Press, 1983, p. 186.

[24] Z. Pawlak, Rough sets: Theoretical aspects of reasoning about data, Kluwer Academic Publishers, 1991, p. 229.

[25] W. Pedrycz, A. Skowron, V. Kreinovich (Eds.), Handbook of Granular Computing, Wiley, 2008, p. 1116.

[26] L. Polkowski, A. Skowron (Eds.), Rough sets in knowledge discovery 1: Methodology and applications, Studies in Fuzziness and Soft Computing 18, Springer, 1998, p. 576.

[27] B. Predki, R. Słowiński, J. Stefanowski, R. Susmaga, S. Wilk, ROSE – software implementation of the rough set theory, in: Proc. Rough Sets and Current Trends in Computing (L. Polkowski, A. Skowron, Eds.) LNAI 1424, Springer, 1998, pp. 605–608.

[28] Y.H. Qian, J.Y. Liang, Y. Yao, C.Y. Dang, MGRS: A multi-granulation rough set, Information Sciences 180 (2010) 949–970.

[29] L.S. Riza, et al., Implementing algorithms of rough set theory and fuzzy rough set theory in the R package RoughSets, Information Sciences 287(10) (2014) 68–89.

[30] Z. Sahri, R. Yusof, J. Watada, FINNIM: Iterative imputation of missing values in dissolved gas analysis dataset, IEEE Transactions on Industrial Informatics 10(4) (2014) 2093–2102.

[31] H. Sakai, On a framework for logic programming with incomplete information, Fundamenta Informaticae 19(3/4) (1993) 223–234.

[32] H. Sakai, R. Ishibashi, K. Koba, M. Nakata, Rules and apriori algorithm in non-deterministic information systems, Transactions on Rough Sets 9 (2008) 328–350.

[33] H. Sakai, M. Wu, N. Yamaguchi, M. Nakata, Rough set-based information dilution by non-deterministic information, in: Proc. RSFDGrC2013, LNCS 8170, Springer, 2013, pp. 55–66.

[34] H. Sakai, M. Wu, M. Nakata, Apriori-based rule generation in incomplete information databases and non-deterministic information systems, Fundamenta Informaticae 130(3) (2014) 343–376.

[35] H. Sakai, C. Liu, X. Zhu, M. Nakata, On NIS-Apriori based data mining in SQL, in: Proc. Int. Conf. on Rough Sets, LNCS 9920, Springer, 2016, pp. 514–524.

[36] H. Sakai, Execution logs by RNIA software tools, 2016.
`http://www.mns.kyutech.ac.jp/~sakai/RNIA`

[37] H. Sakai, M. Nakata, J. Watada, A proposal of machine learning by rule generation from tables with non-deterministic information and its prototype system, in: Proc. Int. Conf. on Rough Sets, LNCS 10313, Springer, 2017, pp. 535–551.

[38] A. Skowron, C. Rauszer, The discernibility matrices and functions in information systems, in: Intelligent Decision Support - Handbook of Advances and Applications of the Rough Set Theory (R. Słowiński, Ed.), Kluwer Academic Publishers, 1992, pp. 331–362.

[39] D. Ślęzak, V. Eastwood, Data warehouse technology by infobright, in: Proc. ACM SIGMOD (U. Çetintemel, S.B. Zdonik, et. al. Eds.), 2009, pp. 841–846.

[40] D. Ślęzak, H. Sakai, Automatic extraction of decision rules from non-deterministic data systems: Theoretical foundations and SQL-based implementation, in: Database Theory and Application (D. Ślęzak, T.H. Kim, Y. Zhang, J. Ma, K.I. Chung, Eds.), CCIS 64, Springer, 2009, pp. 151–162.

[41] J. Stefanowski, A. Tsoukiàs, Incomplete information tables and rough classification, Computational Intelligence 17(3) (2001) 545–566.

[42] S. Tsumoto, Automated induction of medical expert system rules from clinical databases based on rough set theory, Information Sciences 112 (1998) 67–84.

[43] W.Z. Wu, W.X. Zhang, H.Z. Li, Knowledge acquisition in incomplete fuzzy information systems via the rough set approach, Expert Systems 20 (2003) 280–286.

[44] X. Yang, D. Yu, J. Yang, L. Wei, Dominance-based rough set approach to incomplete interval-valued information system, Data & Knowledge Engineering 68(11) (2009) 1331–1347.

[45] Y.Y. Yao, Three-way decisions with probabilistic rough sets, Information Sciences 180 (2010) 314–353.

[46] Y.Y. Yao, Y. She, Rough set models in multigranulation spaces, Information Sciences 327 (2016) 40–56.

[47] W. Zhu, Topological approaches to covering rough sets, Information Sciences 177(6) (2007) 1499–1508.

[48] L.A. Zadeh, Toward a theory of fuzzy information granulation and its centrality in human reasoning and fuzzy logic, Fuzzy Sets and Systems 90(2) (1997) 111–127.

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

[50] C. Zhong, W. Pedrycz, D. Wang, L. Li, Z. Li, Granular data imputation: A framework of granular computing, Applied Soft Computing, 46 (2016) 307–316.