Published in *Journal of Computer and System Sciences* 53 (1996), 575–582

Partition Triples: A Tool for Reduction of Data Sets

Jerzy W. **Grzymala-Busse** and Soe **Than**

Department of Electrical Engineering and Computer Science

University of Kansas

Lawrence, KS 66045

Abstract.

Data sets discussed in this paper are presented as tables with rows corresponding to examples (entities, objects) and columns to attributes. A partition triple is defined for such a table as a triple of partitions on the set of examples, the set of attributes, and the set of attribute values, respectively, preserving the structure of a table. The idea of a partition triple is an extension of the idea of a partition pair, introduced by J. Hartmanis and J. Stearns in automata theory. Results characterizing partition triples and algorithms for computing partition triples are presented. The theory is illustrated by an example of an application in machine learning from examples.

1. Introduction

Reduction of data is a useful way to economize on space in current computer technology. Problems with reducing the size of data sets occur in many areas. One of them is empirical machine learning, where large input data sets make learning difficult. This paper presents a methodology for reduction of data sets, where the reduced data set preserves the structure of the original data set. Two forms of data sets are discussed in this paper. First, it is assumed that a data set is given in the form of a table, called an information system (or instance space). Rows of the table are labeled with names of examples (entities, objects), columns with names of attributes. Every example is characterized by a tuple of values of all attributes. For example, such an information system may contain data about patients in a hospital. Attributes are tests, such as surface temperature, diastolic blood pressure, systolic blood pressure, etc. A row of the table represents a patient, characterized by a tuple of values of all tests.

The second form of a data set discussed in this paper is a decision table. A decision table is defined the same way as an information system, except that the table contains an additional column called a decision. Every example is additionally characterized by a value of the decision. A decision value is usually determined by an expert. For example, a decision table may contain hospital patient data, where a patient is characterized as being healthy or sick with some disease by an expert—a physician. This form of input data is common, e.g., for machine learning from examples.

The following problem is addressed in this paper: How to reduce the original data set to a smaller data set (containing fewer examples, attributes, and attribute values). At the same time, the smaller data set should preserve the structure of the original data set.

The proposed method of reduction of data sets is based on the idea of a partition triple, i.e., a triple of three partitions: on the set of examples, attributes, and attribute values. Every such partition clusters elements (examples, attributes, and attribute values) into blocks of elements. Additionally, examples, attributes, and attribute values are reduced into corresponding blocks in such a way that in the reduced table, where examples, attributes and attribute values are replaced by corresponding blocks, the block containing v is a value of the attribute block containing a for the example block containing x if and only if v is a value of attribute a for example x in the original table. This way the blocks of examples and attributes are transformed into blocks of attribute values in the same way that their members are transformed in the original table.

Also, a triple algebra theory, which is a basic algebraic structure for partition triples, is developed. An algorithm for computing partition triples is presented as well.

The idea of a partition triple is an extension of the idea of a partition pair, introduced by J. Hartmanis and R. Stearns in automata theory [5]. The special case of a partition triple was studied in [1]. Some preliminary results on partition triples were presented in [3, 4].

The theory of partition triples has many potential applications. One of the most obvious is relational data bases. Another application, machine learning from examples, is briefly illustrated in this paper. In this domain the input data sets are presented as decision tables. Large input data sets, representing examples, make learning difficult. This paper includes an example showing how partition triples may be used for inducing simpler rules from examples. The induced rule set represents the same knowledge as the rule set induced from the original data set.

2. Preliminary Definitions

Originally we will assume that the data are collected in the table, called an *information system*, and defined as follows [8, 9]. The information system S is a fourtuple (E, A, V, ρ), where

E is a finite nonempty set of *examples*,

A is a finite nonempty set of *attributes*,

V is a finite nonempty set of *attribute values*,

and ρ is a function, ρ: $E \times A \rightarrow V$.

An example of such an information system is presented in Table 1.

Let X be a nonempty finite set. A *partition* π on X is a family of disjoint subsets of X whose set union is X. Elements of partition π will be called *blocks* of π. If elements x and y are both members of the same block of π, it will be denoted by $x \equiv y$ (π). There are two trivial partitions 0_X and 1_X, where 0_X is the partition on X in which all blocks are one-element subsets of X and 1_X is the partition on X which contains only one block.

If π and τ are partitions on X, then the product of π and τ, denoted by $\pi \cdot \tau$, is a partition on X such that $x \equiv y$ $(\pi \cdot \tau)$ if and only if $x \equiv y$ (π) and $x \equiv y$ (τ). The sum of π and τ, denoted by $\pi + \tau$, is a partition on X such that $x \equiv y$ $(\pi + \tau)$ if and only if there exists a sequence $x = x_1$, $x_2, \ldots, x_n = y$ of elements of X such that $x_i \equiv x_{i+1}(\pi)$ or $x_i \equiv x_{i+1}(\tau)$ for $i = 1, 2, \ldots, n - 1$.

A partition π is said to be smaller than or equal to another partition τ, denoted by $\pi \le \tau$, if and only if for every block B of π there exists a block B' of τ such that $B \subseteq B'$. Obviously, the product of π and τ may be defined as the greatest lower bound (g. l. b.) of π and τ, and the sum of π and τ may be defined as the least upper bound (l. u. b.) of π and τ, see, e.g., [5].

3. Partition Triples and MMm triples

The idea presented here of reduction data, in its special case, was originally developed in automata theory, under the name of partition pairs, see e.g. [5]. Later on this idea, extended to input data of machine learning systems, was presented in [1]. In this paper the idea of a partition triple from [1] is generalized. In [1], in any triple of partitions, the partition on the set of all attributes was constant. Some preliminary results were discussed in [3, 4].

Definition. For a (E, A, V, ρ), let π be a partition on E, let τ be a partition on A, and let λ be a partition on V. A *partition triple* on an information system $S = (E, A, V, \rho)$ is an ordered triple of partitions (π, τ, λ) such that for all $x, y \in E$ and $a, b \in A$

$$x \equiv y\ (\pi) \text{ and } a \equiv b\ (\tau) \text{ implies that } \rho(x, a) \equiv \rho(y, b)\ (\lambda).$$

The set of all partition triples on S will be denoted by $L(S)$.

Table 1

	a_1	a_2	a_3	a_4
x_1	0	0	2	2
x_2	0	1	2	2
x_3	1	1	2	3
x_4	1	1	3	3
x_5	4	3	4	4

Definition. Let (π, τ, λ) be a partition triple on an information system $S = (E, A, V, \rho)$. The (π, τ, λ)-*image* of S is the information system $(\pi, \tau, \lambda, \rho')$ such that for all $B_\pi \in \pi$, $B_\tau \in \tau$, and $B_\lambda \in \lambda$

$$\rho'(B_\pi, B_\tau) = B_\lambda \text{ if } \rho(x, a) = v,$$

where x, a, and v are arbitrary members of B_π, B_τ, and B_λ, respectively.

An example of a partition triple of the information system from Table 1 is

$$(\{\{x_1, x_2\}, \{x_3, x_4\}, \{x_5\}\}, \{\{a_1\}, \{a_2\}, \{a_3, a_4\}\}, \{\{0, 1\}, \{2, 3\}, \{4\}\}).$$

The (π, τ, λ)-image of S from Table 1 is presented in Table 2.

Lemma 3.1. Let (π, τ, λ) and (π', τ', λ') be partition triples on $S = (E, A, V, \rho)$, then

(i) $(\pi \cdot \pi', \tau \cdot \tau', \lambda \cdot \lambda')$

and

(ii) $(\pi + \pi', \tau + \tau', \lambda + \lambda')$

are also partition triples on $S = (E, A, V, \rho)$.

Proof. Let $x, y \in E$ and $a, b \in A$.

(i) $x \equiv y\ (\pi \cdot \pi')$ and $a \equiv b\ (\tau \cdot \tau')$ implies $x \equiv y\ (\pi)$, $x \equiv y\ (\pi')$, $a \equiv b\ (\tau)$, and $a \equiv b\ (\tau')$. Since (π, τ, λ) and (π', τ', λ') are partition triples on S, $\rho(x, a) \equiv \rho(y, b)\ (\lambda)$, $\rho(x, a) \equiv \rho(y, b)\ (\lambda')$, i.e., $\rho(x, a) \equiv \rho(y, b)\ (\lambda \cdot \lambda')$. Therefore $(\pi \cdot \pi', \tau \cdot \tau', \lambda \cdot \lambda')$ is a partition triple on S.

(ii) Without loss of generality let us assume that $x \equiv y\ (\pi + \pi')$ implies that there exists a

Table 2

	$\{a_1\}$	$\{a_2\}$	$\{a_3, a_4\}$
$\{x_1, x_2\}$	$\{0, 1\}$	$\{0, 1\}$	$\{2, 3\}$
$\{x_3, x_4\}$	$\{0, 1\}$	$\{0, 1\}$	$\{2, 3\}$
$\{x_5\}$	$\{4\}$	$\{2, 3\}$	$\{4\}$

sequence $x = x_0, x_1, x_2, ..., x_n = y$ such that $x_i \equiv x_{i+1}$ (π) for even i and $x_i \equiv x_{i+1}$ (π') for odd i, where $i \in \{0, 1, ..., n-1\}$ and $a \equiv b$ ($\tau + \tau'$) implies that there exists a sequence $a = a_0, a_1, a_2, ..., a_m = b$ such that $a_j \equiv a_{j+1}$ (τ) for even j and $a_j \equiv a_{j+1}$ (τ') for odd j, where $j \in \{0, 1, ..., m-1\}$. As (π, τ, λ) and (π', τ', λ') are partition triples on S, $\rho(x_i, a_0) \equiv \rho(x_{i+1}, a_0)$ (λ), for even i, and $\rho(x_i, a_0) \equiv \rho(x_{i+1}, a_0)$ (λ'), for odd i. Therefore $\rho(x_0, a_0) \equiv \rho(x_n, a_0)$ ($\lambda + \lambda'$), i.e., $\rho(x, a) \equiv \rho(y, a)$ ($\lambda + \lambda'$). As (π, τ, λ) and (π', τ', λ') are partition triples on S, $\rho(y, a_j) \equiv \rho(y, a_{j+1})$ (λ), for even j, and $\rho(y, a_j) \equiv \rho(y, a_{j+1})$ (λ'), for odd j. Therefore $\rho(y, a_0) \equiv \rho(y, a_m)$ ($\lambda + \lambda'$), i.e., $\rho(y, a) \equiv \rho(y, b)$ ($\lambda + \lambda'$). So, $\rho(x, a) \equiv \rho(y, a) \equiv \rho(y, b)$ ($\lambda + \lambda'$), i.e., ($\pi + \pi'$, $\tau + \tau'$, $\lambda + \lambda'$) is a partition triple on S.

Lemma 3.2. For any partition π on E, for any partition τ on A, and for any partition λ on V, (π, τ, 1_V) and (0_E, 0_A, λ) are partition triples on S.

Proof. Let π be any partition on E, and τ be any partition on A. Let x and y be any elements of E such that $x \equiv y$ (π) and a and b be any elements of A such that $a \equiv b$ (τ). Then $\rho(x, a) \equiv \rho(y, b)$ (1_V) because the partition 1_V has only one block which is the whole set V. Therefore, (π, τ, 1_V) is a partition triple on S.

Let λ be any partition on V. Let x and y be any elements of E such that $x \equiv y$ (0_E) and a and b be any elements of A such that $a \equiv b$ (0_A). Since each block of partitions 0_E and 0_A contains only one element, $x = y$ and $a = b$. Therefore, $\rho(x, a) = \rho(y, b)$, i.e., $\rho(x, a)$ and $\rho(y, b)$ are identical element in V, and hence they are in the same block of any partition on V, i.e., $\rho(x, a) \equiv \rho(y, b)$ (λ). Therefore, (0_E, 0_A, λ) is a partition triple on S.

Lemma 3.3. The set $L(S)$ of all partition triples on $S = (E, A, V, \rho)$ is a lattice.

Proof. The proof follows directly from Lemma 3.1.

Definitions. For a given partition π on E, the minimal partition λ on V such that (π, 0_A, λ) is a partition triple on $S = (E, A, V, \rho)$ will be denoted $\mathbf{m_{ev}}(\pi)$. It is obvious that

$$\mathbf{m_{ev}}(\pi) = \prod \{\lambda \mid (\pi, 0_A, \lambda) \in L(S)\}.$$

Similarly, for a given partition τ on A, the minimal partition λ on V such that $(0_E, \tau, \lambda)$ is a partition triple on $S = (E, A, V, \rho)$ will be denoted $\mathbf{m_{av}}(\tau)$, and

$$\mathbf{m_{av}}(\tau) = \prod \{\lambda \mid (0_E, \tau, \lambda) \in L(S)\}.$$

For a given partition λ on V we may ask what are maximal partitions π on E and τ on A such that $(\pi, 0_A, \lambda)$ and $(0_E, \tau, \lambda)$ are partition triples on $S = (E, A, V, \rho)$. Such partitions will be denoted $\mathbf{M_{ev}}(\lambda)$ and $\mathbf{M_{av}}(\lambda)$, where

$$\mathbf{M_{ev}}(\lambda) = \sum \{\pi \mid (\pi, 0_A, \lambda) \in L(S)\}$$

and

$$\mathbf{M_{av}}(\lambda) = \sum \{\tau \mid (0_E, \tau, \lambda) \in L(S)\}.$$

In the preceding definitions, **m** stands for minimum and **M** for maximum. Partitions $\mathbf{m_{ev}}(\pi)$ and $\mathbf{m_{av}}(\tau)$ represent the largest amount of information about blocks of attribute values which can be drawn from the information about blocks of π and τ, respectively. Partitions $\mathbf{M_{ev}}(\lambda)$ and $\mathbf{M_{av}}(\lambda)$ represent the least amount of information about blocks of examples and attributes which must be supplied to identify blocks of λ.

A partition triple (π, τ, λ) on $S = (E, A, V, \rho)$ will be called a *MMm triple* if and only if

$$\pi = \mathbf{M_{ev}}(\lambda),\ \tau = \mathbf{M_{av}}(\lambda), \text{ and } \lambda = \mathbf{m_{ev}}(\pi) + \mathbf{m_{av}}(\tau).$$

The set of all MMm triples of $S = (E, A, V, \rho)$ will be denoted $K(S)$. An example of a MMm triple of the information system from Table 1 is

$$(\{\{x_1, x_2, x_3, x_4\}, \{x_5\}\}, \{\{a_1\}, \{a_2\}, \{a_3, a_4\}\}, \{\{0, 1\}, \{2, 3\}, \{4\}\}).$$

4. Triple Algebra

We can study properties of partition triples and MMm triples by analyzing the underlying algebraic structure, called a triple algebra. The results of the abstract structure can be applied not only to the partition triples but also to other, not yet discovered, interpretations. Let L_1, L_2, and L_3

be finite lattices. Then a subset Δ of $L_1 \times L_2 \times L_3$ is a *triple algebra* on $L_1 \times L_2 \times L_3$ if and only if the following postulates hold:

P1. (x_1, y_1, z_1) and (x_2, y_2, z_2) are in Δ implies that $(x_1 \cdot x_2, y_1 \cdot y_2, z_1 \cdot z_2)$ and $(x_1 + x_2, y_1 + y_2, z_1 + z_2)$ are in Δ,

P2. For any x in L_1, y in L_2, and z in L_3, $(x, y, 1_{L_3})$ and $(0_{L_1}, 0_{L_2}, z)$ are in Δ.

For (x, y, z) and (x', y', z') in $L_1 \times L_2 \times L_3$, we define $(x, y, z) \le (x', y', z')$ if and only if $x \le x'$, $y \le y'$, and $z \le z'$.

Lemma 4.1. If Δ is a triple algebra on $L_1 \times L_2 \times L_3$ and (x, y, z) is in Δ, then $x' \le x$, $y' \le y$, and $z' \ge z$ implies that (x', y', z') is in Δ.

Proof. Suppose that (x, y, z) is in Δ and $x' \le x$, $y' \le y$, and $z' \ge z$. By the property P2, $(x', y', 1_{L_3})$ is in Δ. Hence $(x \cdot x', y \cdot y', z \cdot 1_{L_3})$ is in Δ, by the property P1. Since $x' \le x$ and $y' \le y$, (x', y', z) is in Δ. By the property P2, $(0_{L_1}, 0_{L_2}, z')$ is in Δ and hence $(x' + 0_{L_1}, y' + 0_{L_2}, z + z')$ is in Δ. Since $z' \ge z$, (x', y', z') is in Δ.

Definitions. Let Δ be a triple algebra on $L_1 \times L_2 \times L_3$. For any x in L_1, y in L_2, and z in L_3, we define

$$\mathbf{m}_{13}(x) = \prod \{z \mid (x, 0_{L_2}, z) \in \Delta\},$$

$$\mathbf{m}_{23}(y) = \prod \{z \mid (0_{L_1}, y, z) \in \Delta\},$$

$$\mathbf{M}_{13}(z) = \sum \{x \mid (x, 0_{L_2}, z) \in \Delta\},$$

$$\mathbf{M}_{23}(z) = \sum \{y \mid (0_{L_1}, y, z) \in \Delta\},$$

$$\mathbf{m}_{123}(x, y) = \mathbf{m}_{13}(x) + \mathbf{m}_{23}(y),$$

$$\mathbf{M}_{123}(z) = (\mathbf{M}_{13}(z), \mathbf{M}_{23}(z)).$$

Lemma 4.2. For any x in L_1 and y in L_2, $\mathbf{m}_{123}(x, y) = \prod \{z \mid (x, y, z) \in \Delta\}$.

Proof. Let us define the following sets

$$R(x, y) = \{z \mid (x, y, z) \in \Delta\},$$

$$R_1(x) = \{z \mid (x, 0_{L_2}, z) \in \Delta\},$$

and

$$R_2(y) = \{z \mid (0_{L_1}, y, z) \in \Delta\}.$$

Then $z \in R(x, y) \Rightarrow (x, y, z) \in \Delta$

$\Rightarrow (x, 0_{L_2}, z) \in \Delta$ and $(0_{L_1}, y, z) \in \Delta$, because $0_{L_2} \leq y$ and $0_{L_1} \leq x$,

$\Rightarrow z \in R_1(x)$ and $z \in R_2(y)$.

Therefore $R(x, y) \subseteq R_1(x)$ and $R(x, y) \subseteq R_2(y)$.

Consequently,

$$\prod \{z \mid z \in R(x, y)\} \geq \prod \{z \mid z \in R_1(x)\}$$

and

$$\prod \{z \mid z \in R(x, y)\} \geq \prod \{z \mid z \in R_2(y)\},$$

i.e., $\prod \{z \mid z \in R(x, y)\} \geq \prod \{z \mid z \in R_1(x)\} + \prod \{z \mid z \in R_2(y)\}$, or, $\prod \{z \mid z \in R(x, y)\} \geq \mathbf{m_{123}}(x, y)$. Let $z_1 = \prod \{z \mid z \in R_1(x)\}$ and $z_2 = \prod \{z \mid z \in R_2(y)\}$. For any $z \in R_1(x)$, $(x, 0_{L_2}, z) \in \Delta$. By the property P1, $\prod \{(x, 0_{L_2}, z) \mid z \in R_1(x)\} \in \Delta$, i.e., $(x, 0_{L_2}, z_1) \in \Delta$. Similarly, $(0_{L_1}, y, z_2) \in \Delta$. By the property P1, $(x + 0_{L_1}, 0_{L_2} + y, z_1 + z_2) \in \Delta$, i.e., $(x, y, z_1 + z_2) \in \Delta$, or $z_1 + z_2 \in R(x, y)$. Therefore $z_1 + z_2 \geq \prod \{z \mid z \in R(x, y)\}$, i.e., $\mathbf{m_{123}}(x, y) \geq \prod \{z \mid z \in R(x, y)\}$. Therefore $\mathbf{m_{123}}(x, y) = \prod \{z \mid (x, y, z) \in \Delta\}$.

Lemma 4.3. For any z in L_3, $\mathbf{M_{123}}(z) = \sum \{(x, y) \mid (x, y, z) \in \Delta\}$.

Proof is similar to the proof of Lemma 4.2.

Definition. For any two elements (x, y, z) and (x', y', z') in $L_1 \times L_2 \times L_3$, we define $(x, y, z) \leq (x', y', z')$ if and only if $x \leq x'$ in L_1, $y \leq y'$ in L_2, and $z \leq z'$ in L_3.

Lemma 4.4. Any triple algebra Δ on $L_1 \times L_2 \times L_3$ is a lattice under the above ordering with zero element (0, 0, 0), unit element (1, 1, 1), and component-wise g.l.b. and l.u.b. operations.

Definition. An element (x, y, z) in a triple algebra Δ is called a *MMm triple* if and only if $x = \mathbf{M_{13}}(z)$, $y = \mathbf{M_{23}}(z)$, and $z = \mathbf{m_{13}}(x) + \mathbf{m_{23}}(y)$. The set of all MMm triples of Δ will be denoted Q_Δ.

Theorem 4.1. Let Δ be a triple algebra on $L_1 \times L_2 \times L_3$. For any x in L_1, y in L_2, and z in L_3,

(1) $(\mathbf{M_{13}}(z), 0_{L_2}, z)$, $(0_{L_1}, \mathbf{M_{23}}(z), z)$, $(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z), z)$, $(x, 0_{L_2}, \mathbf{m_{13}}(x))$, $(0_{L_1}, y, \mathbf{m_{23}}(y))$, and $(x, y, \mathbf{m_{123}}(x, y))$ are in Δ.

(2) $x_1 \leq x_2$ implies $\mathbf{m_{13}}(x_1) \leq \mathbf{m_{13}}(x_2)$, $y_1 \leq y_2$ implies $\mathbf{m_{23}}(y_1) \leq \mathbf{m_{23}}(y_2)$, and $x_1 \leq x_2$ and $y_1 \leq y_2$ imply $\mathbf{m_{123}}(x_1, y_1) \leq \mathbf{m_{123}}(x_2, y_2)$.

(3) $\mathbf{m_{13}}(x_1 + x_2) = \mathbf{m_{13}}(x_1) + \mathbf{m_{13}}(x_2)$, $\mathbf{m_{23}}(y_1 + y_2) = \mathbf{m_{23}}(y_1) + \mathbf{m_{23}}(y_2)$, and $\mathbf{m_{123}}(x_1 + x_2, y_1 + y_2) = \mathbf{m_{123}}(x_1, y_1) + \mathbf{m_{123}}(x_2, y_2)$.

(4) $\mathbf{m_{13}}(x_1 \cdot x_2) \leq \mathbf{m_{13}}(x_1) \cdot \mathbf{m_{13}}(x_2)$, $\mathbf{m_{23}}(y_1 \cdot y_2) \leq \mathbf{m_{23}}(y_1) \cdot \mathbf{m_{23}}(y_2)$, and $\mathbf{m_{123}}(x_1 \cdot x_2, y_1 \cdot y_2) \leq \mathbf{m_{123}}(x_1, y_1) \cdot \mathbf{m_{123}}(x_2, y_2)$.

(5) $z \geq \mathbf{m_{13}}(x)$ if and only if $(x, 0_{L_2}, z) \in \Delta$, $z \geq \mathbf{m_{23}}(y)$ if and only if $(0_{L_1}, y, z) \in \Delta$, and $z \geq \mathbf{m_{123}}(x, y)$ if and only if $(x, y, z) \in \Delta$.

(6) $z_1 \leq z_2$ implies that $\mathbf{M_{13}}(z_1) \leq \mathbf{M_{13}}(z_2)$ and $\mathbf{M_{23}}(z_1) \leq \mathbf{M_{23}}(z_2)$.

(7) $\mathbf{M_{13}}(z_1 + z_2) \geq \mathbf{M_{13}}(z_1) + \mathbf{M_{13}}(z_2)$ and $\mathbf{M_{23}}(z_1 + z_2) \geq \mathbf{M_{23}}(z_1) + \mathbf{M_{23}}(z_2)$.

(8) $\mathbf{M_{13}}(z_1 \cdot z_2) = \mathbf{M_{13}}(z_1) \cdot \mathbf{M_{13}}(z_2)$ and $\mathbf{M_{23}}(z_1 \cdot z_2) = \mathbf{M_{23}}(z_1) \cdot \mathbf{M_{23}}(z_2)$.

(9) $x \leq \mathbf{M_{13}}(z)$ if and only if $(x, 0_{L_2}, z) \in \Delta$, $y \leq \mathbf{M_{23}}(z)$ if and only if $(0_{L_1}, y, z) \in \Delta$, and $x \leq \mathbf{M_{13}}(z)$ and $y \leq \mathbf{M_{23}}(z)$ if and only if $(x, y, z) \in \Delta$.

(10) $\mathbf{M_{13}}(\mathbf{m_{13}}(x)) \geq x$, $\mathbf{M_{23}}(\mathbf{m_{23}}(y)) \geq y$, $\mathbf{M_{13}}(\mathbf{m_{123}}(x, y)) \geq x$, and $\mathbf{M_{23}}(\mathbf{m_{123}}(x, y)) \geq y$.

(11) $\mathbf{m_{123}}(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z)) \leq z$, $\mathbf{m_{13}}(\mathbf{M_{13}}(z)) \leq z$, and $\mathbf{m_{23}}(\mathbf{M_{23}}(z)) \leq z$.

(12) $\mathbf{M_{13}}(\mathbf{m_{13}}(\mathbf{M_{13}}(z))) = \mathbf{M_{13}}(z)$, $\mathbf{M_{23}}(\mathbf{m_{23}}(\mathbf{M_{23}}(z))) = \mathbf{M_{23}}(z)$, $\mathbf{M_{13}}(\mathbf{m_{123}}(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z))) = \mathbf{M_{13}}(z)$, and $\mathbf{M_{23}}(\mathbf{m_{123}}(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z))) = \mathbf{M_{23}}(z)$.

(13) $\mathbf{m_{13}}(\mathbf{M_{13}}(\mathbf{m_{13}}(x))) = \mathbf{m_{13}}(x)$, $\mathbf{m_{23}}(\mathbf{M_{23}}(\mathbf{m_{23}}(y))) = \mathbf{m_{23}}(y)$, and $\mathbf{m_{123}}(\mathbf{M_{13}}(\mathbf{m_{123}}(x, y)), \mathbf{M_{23}}(\mathbf{m_{123}}(x, y))) = \mathbf{m_{123}}(x, y)$.

(14) $(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z), \mathbf{m_{123}}(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z)))$ and $(\mathbf{M_{13}}(\mathbf{m_{123}}(x, y)), \mathbf{M_{23}}(\mathbf{m_{123}}(x, y)), \mathbf{m_{123}}(x, y))$ are in Q_Δ.

(15) If (x_1, y_1, z_1) and (x_2, y_2, z_2) are in Q_Δ, then $x_1 \le x_2$ and $y_1 \le y_2$ if and only if $z_1 \le z_2$.

(16) The set Q_Δ under the ordering on Δ is a lattice in which

g.l.b.$\{(x_1, y_1, z_1), (x_2, y_2, z_2)\} = (x_1 \cdot x_2, y_1 \cdot y_2, \mathbf{m_{123}}(x_1 \cdot x_2, y_1 \cdot y_2))$ and

l.u.b.$\{(x_1, y_1, z_1), (x_2, y_2, z_2)\} = (\mathbf{M_{13}}(z_1 + z_2), \mathbf{M_{23}}(z_1 + z_2), z_1 + z_2)$.

Proof

(1) By the definitions of $\mathbf{M_{13}}$ and $\mathbf{M_{23}}$, $(\mathbf{M_{13}}(z), 0_{L_2}, z)$ and $(0_{L_1}, \mathbf{M_{23}}(z), z)$ are in Δ. By the property P1, $(\mathbf{M_{13}}(z) + 0_{L_1}, 0_{L_2} + \mathbf{M_{23}}(z), z + z)$ is in Δ, i.e., $(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z), z)$ is in Δ. Similarly, $(x, 0_{L_2}, \mathbf{m_{13}}(x))$ and $(0_{L_1}, y, \mathbf{m_{23}}(y))$ are in Δ and so is their sum, i.e., $(x + 0_{L_1}, 0_{L_2} + y, \mathbf{m_{13}}(x) + \mathbf{m_{23}}(y))$ is in Δ, i.e., $(x, y, \mathbf{m_{123}}(x, y))$ is in Δ.

The proofs of (2) – (16) are either similar to the proof of (1) or straightforward.

The following result gives characterization of Δ in terms of Q_Δ.

Theorem 4.2. Let Δ be a triple algebra on $L_1 \times L_2 \times L_3$. Let x in L_1, y in L_2, and z in L_3. Then (x, y, z) is in Δ if and only if there exists (x', y', z') in Q_Δ such that $x \le x'$, $y \le y'$, and $z \ge z'$.

Proof. Suppose that (x, y, z) is in Δ. Let $x' = \mathbf{M_{13}}(z)$, $y' = \mathbf{M_{23}}(z)$, and $z' = \mathbf{m_{123}}(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z))$. By Theorem 4.1.9, $x \le \mathbf{M_{13}}(z)$ and $y \le \mathbf{M_{23}}(z)$, i.e., $x \le x'$ and $y \le y'$. By Theorem 4.1.11, $\mathbf{m_{123}}(\mathbf{M_{13}}(z), \mathbf{M_{23}}(z)) \le z$, i.e., $z' \le z$. By Theorem 4.1.14, (x', y', z') is in Q_Δ. Therefore, there exists (x', y', z') in Q_Δ such that $x \le x'$, $y \le y'$, and $z \ge z'$.

Now, suppose that there exists (x', y', z') in Q_Δ such that $x \le x'$, $y \le y'$, and $z \ge z'$. Since $Q_\Delta \subseteq \Delta$, (x', y', z') is in Δ and hence (x, y, z) is in Δ, by Lemma 4.1.

Lemma 4.5. If (x_1, y_1, z_1) and (x_2, y_2, z_2) are in Q_Δ, then the following three statements are equivalent:

(1) $(x_1, y_1, z_1) \ge (x_2, y_2, z_2)$,

(2) $x_1 \ge x_2$ and $y_1 \ge y_2$,

(3) $z_1 \geq z_2$.

Proof. By Theorem 4.1.15, statement (2) and statement (3) are equivalent. By the definition of ordering relation in $L_1 \times L_2 \times L_3$, statement (1) is equivalent to the combination of statements (2) and (3). Therefore, the three statements are equivalent.

Lemma 4.6. If (x, y, z) is in Δ, then $(\mathbf{M_{13}}(z), y, z)$ and $(x, \mathbf{M_{23}}(z), z)$ are also in Δ.

Proof. Assume that (x, y, z) is in Δ. Then $(0_{L_1}, y, z)$ is in Δ, by Lemma 4.1. By Theorem 4.1.1, $(\mathbf{M_{13}}(z), 0_{L_2}, z)$ is in Δ. Therefore, $(0_{L_1} + \mathbf{M_{13}}(z), y + 0_{L_2}, z + z)$ is in Δ, i.e., $(\mathbf{M_{13}}(z), y, z)$ is in Δ. Similarly, $(x, \mathbf{M_{23}}(z), z)$ is in Δ.

Lemma 4.7. $\mathbf{M_{13}}(\mathbf{m_{23}}(\mathbf{M_{23}}(z))) \leq \mathbf{M_{13}}(z)$ and $\mathbf{M_{23}}(\mathbf{m_{13}}(\mathbf{M_{13}}(z))) \leq \mathbf{M_{23}}(z)$.

Proof. By Theorem 4.1.11, $\mathbf{m_{23}}(\mathbf{M_{23}}(z)) \leq z$. By Theorem 4.1.6, $\mathbf{M_{13}}(\mathbf{m_{23}}(\mathbf{M_{23}}(z))) \leq \mathbf{M_{13}}(z)$. Similarly, $\mathbf{M_{23}}(\mathbf{m_{13}}(\mathbf{M_{13}}(z))) \leq \mathbf{M_{23}}(z)$.

Lemma 4.8. $\mathbf{m_{13}}(\mathbf{M_{13}}(\mathbf{m_{23}}(y))) \leq \mathbf{m_{23}}(y)$ and $\mathbf{m_{23}}(\mathbf{M_{23}}(\mathbf{m_{13}}(x))) \leq \mathbf{m_{13}}(x)$.

Proof. By Theorem 4.1.11, $\mathbf{m_{13}}(\mathbf{M_{13}}(z)) \leq z$, for any z in L_3. Since $\mathbf{m_{23}}(y)$ is in L_3, $\mathbf{m_{13}}(\mathbf{M_{13}}(\mathbf{m_{23}}(y))) \leq \mathbf{m_{23}}(y)$. Similarly, $\mathbf{m_{23}}(\mathbf{M_{23}}(\mathbf{m_{13}}(x))) \leq \mathbf{m_{13}}(x)$.

Theorem 4.3. The set $L(S)$ of all partition triples on $S = (E, A, V, \rho)$ is a triple algebra and, hence, satisfies the above propositions.

Proof. By Lemma 3.2 and by Lemma 3.1, the set $L(S)$ satisfies the properties P1 and P2, respectively, of the triple algebra.

Therefore, $L(S)$ is a triple algebra. The definitions $\mathbf{m_{ev}}$, $\mathbf{m_{av}}$, $\mathbf{M_{ev}}$, and $\mathbf{M_{av}}$ of the partitions are analogous to the definitions $\mathbf{m_{13}}$, $\mathbf{m_{23}}$, $\mathbf{M_{13}}$, and $\mathbf{M_{23}}$, respectively, of the triple algebra. So all the results on the triple algebra can be applied to partition triples by replacing $\mathbf{m_{13}}$, $\mathbf{m_{23}}$, $\mathbf{M_{13}}$, $\mathbf{M_{23}}$, Δ, and Q_Δ with $\mathbf{m_{ev}}$, $\mathbf{m_{av}}$, $\mathbf{M_{ev}}$, $\mathbf{M_{av}}$, $L(S)$, and $K(S)$, respectively.

Thus, by Theorem 4.2, any partition triple on S can be computed from an MMm triple on S by refining the first two partitions and coarsening the third partition of the MMm triple. Also, it is sufficient to compute set K(S) and then compute $L(S)$ from $K(S)$.

5. Algorithm for computing $K(S)$

An algorithm to determine all MMm triples is an extension of the algorithm to determine Mm pairs for automata [5]. Let $\pi_{x,y}$ denote the partition on E such that all blocks of $\pi_{x,y}$ except one are singletons, and the only block of $\pi_{x,y}$ that is not a singleton contains two elements: x and y. Similarly, let $\tau_{a,b}$ denote the partition on A such that all blocks of $\tau_{a,b}$ except one are singletons, and the only block of $\tau_{a,b}$ that is not a singleton contains two elements: a and b. Our algorithm is based on the following result:

Theorem 5.1. If (π, τ, λ) is a MMm triple then

$$\lambda = \Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\} + \Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\}.$$

Proof. First, $\pi \ge \pi_{x,y}$ and $\tau \ge \tau_{a,b}$, hence $(\pi_{x,y}, 0_A, \lambda)$ and $(0_E, \tau_{a,b}, \lambda)$ are partition triples, $\mathbf{m_{ev}}(\pi_{x,y}) \le \lambda$, and $\mathbf{m_{av}}(\tau_{a,b}) \le \lambda$. Then $\Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\} \le \lambda$ and $\Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\} \le \lambda$. Hence

$$\Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\} + \Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\} \le \lambda.$$

On the other hand, $(\Sigma\, \{\pi_{x,y} \mid \pi_{x,y} \le \pi\}, 0_A, \Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\})$ and $(0_E, \Sigma\, \{\tau_{a,b} \mid \tau_{a,b} \le \tau\}, \Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\})$ are partition triples and so is $(\Sigma\, \{\pi_{x,y} \mid \pi_{x,y} \le \pi\}, \Sigma\, \{\tau_{a,b} \mid \tau_{a,b} \le \tau\}, \Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\} + \Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\}) = (\pi, \tau, \Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\} + \Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\})$. Therefore,

$$\Sigma\, \{\mathbf{m_{ev}}\, (\pi_{x,y}) \mid \pi_{x,y} \le \pi\} + \Sigma\, \{\mathbf{m_{av}}\, (\tau_{a,b}) \mid \tau_{a,b} \le \tau\} \ge \mathbf{m_{av}}(\pi) + \mathbf{m_{av}}(\tau) = \lambda,$$

and the result is proved.

Since $\pi_{x,\,y}$ and $\tau_{a,\,b}$ are the smallest nontrivial partitions on E and A, respectively, partitions $\mathbf{m}_{\mathbf{ev}}(\pi_{x,\,y})$ and $\mathbf{m}_{\mathbf{av}}(\tau_{a,\,b})$ are the smallest m-type partitions on V.

In the first stage of the algorithm sets $R_1 = \{\mathbf{m}_{\mathbf{ev}}\,(\pi_{x,\,y}\,) \mid (x, y) \in E \times E\}$ and $R_2 = \{\mathbf{m}_{\mathbf{av}}\,(\tau_{a,\,b}) \mid (a, b) \in A \times A\}$ are computed. For computation of R_1 only $|E| \cdot (|E| - 1)/2$ steps are required, because $\mathbf{m}_{\mathbf{ev}}(\pi_{x,\,y}) = \mathbf{m}_{\mathbf{ev}}(\pi_{y,\,x})$ and $\mathbf{m}_{\mathbf{ev}}(\pi_{x,\,x}) = 0_V$, where $|X|$ denotes the cardinality of the set X. Similarly, computation of the set R_2 requires only $|A| \cdot (|A| - 1)/2$ steps.

In the second stage of the algorithm, the set

$$R^{(1)} = \{\lambda + \lambda' \mid \lambda \in R_1, \lambda' \in R_2\}$$

is computed. Obviously, $R_1 \subseteq R^{(1)}$ and $R_2 \subseteq R^{(1)}$, because $0_V \in R_1$ and $0_V \in R_2$. Then the set

$$R^{(2)} = \{\lambda + \lambda' \mid \lambda \in R^{(1)}, \lambda' \in R^{(1)}\}$$

should be computed. Also, $R^{(1)} \subseteq R^{(2)}$. Similarly, the set $R^{(k+1)}$ is computed from $R^{(k)}$ by

$$R^{(k+1)} = \{\lambda + \lambda' \mid \lambda \in R^{(k)}, \lambda' \in R^{(k)}\}.$$

The process stops when, for some k, $R^{(k)} = R^{(k+1)} = R$. Thus, the set $R^{(1)}$ is the set of generators for R. Every m-type partition $\lambda \in R$ determines two unique M-type partitions on E and A, respectively. These M-type partitions are $\mathbf{M}_{\mathbf{ev}}(\lambda)$ and $\mathbf{M}_{\mathbf{av}}(\lambda)$, respectively. Moreover, for any $\lambda \in R$, $\mathbf{M}_{\mathbf{ev}}(\lambda)$ and $\mathbf{M}_{\mathbf{av}}(\lambda)$ may be computed using the following formulas

$$\mathbf{M}_{\mathbf{ev}}(\lambda) = \Sigma\ \{\pi_{x,\,y} \mid \mathbf{m}_{\mathbf{ev}}(\pi_{x,\,y}) \le \lambda\},$$

and

$$\mathbf{M}_{\mathbf{av}}(\lambda) = \Sigma\ \{\tau_{a,\,b} \mid \mathbf{m}_{\mathbf{av}}(\tau_{a,\,b}) \le \lambda\}.$$

Finally, the set $K(S)$ of all MMm triples is

$$\{(\mathbf{M}_{\mathbf{ev}}(\lambda), \mathbf{M}_{\mathbf{av}}(\lambda), \lambda) \mid \lambda \in R\}.$$

6. Decision Tables

In this section we will assume that the data sets are presented in the form of a decision table. The following definition of a decision table is a slightly modified version of the definition introduced by Z. Pawlak [8, 9]. The decision table T is a sixtuple (E, A, V, d, W, ρ), where

E is a finite nonempty set of *examples*,

A is a finite nonempty set of *attributes*,

V is a finite nonempty set of *attribute values*,

d is a variable called a *decision*,

W is a finite nonempty set of decision values,

$\rho: E \times (A \cup \{d\}) \to V \cup W$, where if ρ is restricted to $E \times A$ it has values from V, and if ρ is restricted to $E \times \{d\}$ it has values from W.

For the sake of simplicity, restrictions of ρ to $E \times A$ and to $E \times \{d\}$ will also be denoted ρ.

An example of the decision table is presented in Table 3.

Definition. For a decision table (E, A, V, d, W, ρ), let π be a partition on E, let τ be a partition on A, and let λ be a partition on V. A *partition triple* on a decision table $T = (E, A, V, d, W, \rho)$ is an ordered triple of partitions (π, τ, λ) such that for all $x, y \in E$ and $a, b \in A$

$$x \equiv y\ (\pi) \text{ and } a \equiv b\ (\tau) \text{ implies that } \rho(x, a) \equiv \rho(y, b)\ (\lambda) \text{ and } \rho(x, d) = \rho(y, d).$$

The set of all partition triples on T will be denoted by $L(S)$.

Table 3

	Attributes				Decision
	a_1	a_2	a_3	a_4	d
x_1	0	0	2	2	0
x_2	0	1	2	2	0
x_3	1	1	2	3	1
x_4	1	1	3	3	1
x_5	4	3	4	4	1

Definition. Let (π, τ, λ) be a partition triple on a decision table $T = (E, A, V, d, W, \rho)$. The (π, τ, λ)-*image* of T is the decision table $(\pi, \tau, \lambda, d, W, \rho')$ such that for all $B_\pi \in \pi$, $B_\tau \in \tau$, and $B_\lambda \in \lambda$

$$\rho'(B_\pi, B_\tau) = B_\lambda \text{ and } \rho'(B_\pi, d) = \rho(x, d) \text{ if } \rho(x, a) = v,$$

where x, a, and v are arbitrary members of B_π, B_τ, and B_λ, respectively.

Lemma 6.1. Let (π, τ, λ) and (π', τ', λ') be partition triples on $T = (E, A, V, d, W, \rho)$, then

(i) $(\pi \cdot \pi', \tau \cdot \tau', \lambda \cdot \lambda')$

and

(ii) $(\pi + \pi', \tau + \tau', \lambda + \lambda')$

are also partition triples on $T = (E, A, V, d, W, \rho)$.

Proof is a straightforward extension of the proof of Lemma 3.1.

Lemma 6.2. For any partition π on E such that $\pi \le \{d\}^*$, for any partition τ on A, and for any partition λ on V, $(\pi, \tau, 1_V)$ and $(0_E, 0_A, \lambda)$ are partition triples on S.

Proof is a straightforward extension of the proof of Lemma 3.2.

Lemma 6.3. The set $L(T)$ of all partition triples on $T = (E, A, V, d, W, \rho)$ is a lattice.

Proof follows directly from Lemma 6.1 .

Operators $\mathbf{m_{ev}}$, $\mathbf{m_{av}}$, $\mathbf{M_{ev}}$, and $\mathbf{M_{av}}$ for decision tables may be defined in the same way as for information systems. Moreover, all previous results of the triple algebra are valid for decision tables as well. In particular, the algorithm for computing the set $K(S)$ of all MMm triples for information systems may be used for decision tables with little changes.

7. Applications—An Example

There are many possible applications of the theory presented. In any area where information systems or decision tables are used, the obvious benefits of simplification can be utilized. One of the evident areas of applications is relational data bases. Another area, less evident, is machine learning from examples.

Reduction of input data sets in machine learning from examples may be considered a kind of preprocessing. Other known approaches to preprocessing of input data in machine learning include selecting the most representative examples [6, 7] and a kind of refinement [10].

Let us illustrate the application of partition triple theory to machine learning from examples using the example of input data in the form of a decision table from Table 4.

Using machine learning system LERS [2], the following rules were induced:

(Quantitative, Excellent) & (Reading, Excellent) $\rightarrow$ (Admission, Accept),

Table 4

	Attributes					Decision
	Quantitative	Analytical	Advanced	Grammar	Reading	Admission
x_1	High	Excellent	High	Excellent	Excellent	Accept
x_2	Excellent	High	Excellent	High	High	Accept
x_3	Excellent	High	Excellent	High	Excellent	Accept
x_4	High	Excellent	Medium	Excellent	High	Accept
x_5	Excellent	High	Low	High	Excellent	Accept
x_6	Excellent	High	Medium	Medium	Low	Reject
x_7	High	Excellent	Low	Low	Medium	Reject
x_8	Low	Medium	Medium	Excellent	High	Reject
x_9	Medium	Low	Low	High	Excellent	Reject

(Grammar, Excellent) & (Quantitative, High) $\rightarrow$ (Admission, Accept),

(Advanced, Excellent) $\rightarrow$ (Admission, Accept),

(Grammar, Medium) $\rightarrow$ (Admission, Reject),

(Grammar, Low) $\rightarrow$ (Admission, Reject),

(Quantitative, Low) $\rightarrow$ (Admission, Reject),

(Quantitative, Medium) $\rightarrow$ (Admission, Reject).

One of the partition triples of the decision table from Table 4 is

$(\{\{x_1, x_2, x_3\}, \{x_4, x_5\}, \{x_6, x_7\}, \{x_8, x_9\}\},$

{{Quantitative, Analytical}, {Advanced}, {Grammar, Reading}},

{{Excellent, High}, {Medium, Low}}).

After assigning new names for the blocks of attributes and for the blocks of attribute values, the corresponding reduced decision table is presented in Table 5.

The rules induced by LERS from the reduced decision table are:

(Aptitude, Above_avg) & (Language, Above_avg) $\rightarrow$ (Admission, Accept),

(Aptitude, Below_avg) $\rightarrow$ (Admission, Reject),

(Language, Below_avg) $\rightarrow$ (Admission, Reject).

Table 5

	Attributes			Decision
	Aptitude	Advanced	Language	Admission
x_1	Above_avg	Above_avg	Above_avg	Accept
x_4	Above_avg	Below_avg	Above_avg	Accept
x_6	Above_avg	Below_avg	Below_avg	Reject
x_8	Below_avg	Below_avg	Above_avg	Reject

The above set of rules represents exactly the same knowledge as the set of rules induced from the original decision table, yet this set is much simpler and more evident.

8. Conclusions

The theory of partition triples of data sets is presented in the paper mostly for information systems, resembling relational databases. However, all results, with respective changes, are valid for decision tables as well. The theory may be used in an obvious way—for computing simpler data sets, while preserving the structure of the original data sets. The main idea is to compute the set K of all MMm triples of a data set. Any partition triple may be computed from a suitable member of K by refining the first two partitions and coarsening the third partition. The theory is illustrated by an example of application from the area of machine learning, showing that induced rules from the simplified data are more evident.

The disadvantage of the presented algorithms is their computational complexity. In general, the worst case time computational complexity for the algorithms to compute all MMm partitions is exponential. Therefore, new, less complex algorithms should be developed, producing only some partition triples, perhaps even only one good partition triple.

Acknowledgement

The authors thank the anonymous referee for many valuable comments on the presentation.

References

1. J. W. Grzymala-Busse, On the reduction of instance space in learning from examples, *in* "Proceedings, 5th International Symposium on Methodologies for Intelligent Systems, Knoxville, TN, October 1990", North Holland, New York, pp. 388–395.

2. J. W. GRZYMALA-BUSSE, LERS—A system for learning from examples based on rough sets, *in* "Intelligent Decision Support. Handbook of Applications and Advances of the Rough Set Theory", R. Slowinski (ed.), Kluwer Academic Publishers, Dordrecht, The Netherlands, 1992, pp. 3–18.

3. J. W. GRZYMALA-BUSSE AND S. THAN, Reduction of instance space in machine learning from examples, *in* "Proceedings, 5th International Symposium on Artificial Intelligence, Cancun, Mexico, December 1992", pp. 303–309.

4. J. W. GRZYMALA-BUSSE AND S. THAN, Data compression in machine learning applied to natural language, *Behavior Research Methods, Instruments, & Computers* **25** (1993), 318–321 .

5. J. HARTMANIS AND R. E. STEARNS, "Algebraic Structure Theory of Sequential Machines", Prentice-Hall, Inc., Englewood Cliffs, NJ, 1966.

6. D. KIBLER AND D. W. AHA, Learning representative exemplars of concepts: An initial case study, *in* "Proceedings, 4th International Workshop on Machine Learning, Irvine, CA, June 1987", pp. 24–30.

7. R. S. MICHALSKI AND R. L. CHILAUSKY, Knowledge acquisition by encoding expert rules versus computer induction from examples: A case study involving soybean pathology, *Int. J. Man-Machine Studies* **12** (1980), 63–87.

8. Z. PAWLAK, Rough sets, *Int. J. Computer and Information Sci.*, **11** (1982), 341–356.

9. Z. PAWLAK, Rough classification, *Int. J. Man-Machine Studies* **20** (1984), 469–483.

10. W. VAN DE VELDE, Learning through progressive refinement, *in* "Proceedings, EWSL 88, 3rd European Working Session on Learning, Glasgow, United Kingdom, October 1988", pp. 211–226.