Acta Cybernetica 14 (2000) 619-630.

Difference Functions of Dependence Spaces

Jouni Järvinen *

Abstract

Here the reduction problem is studied in an algebraic structure called dependence space. We characterize the reducts by the means of dense families of dependence spaces. Dependence spaces defined by indiscernibility relations are also considered. We show how we can determine dense families of dependence spaces induced by indiscernibility relations by applying indiscernibility matrices. We also study difference functions which connect the reduction problem to the general problem of identifying the set of all minimal Boolean vectors satisfying an isotone Boolean function.

1 Introduction

Z. Pawlak introduced his notion of information systems in the early 1980's [11]. Information concerning properties of objects is the basic knowledge included in information systems, and it is given in terms of attributes and values of attributes. For example, we may express statements concerning the color of objects if the information system includes the attribute "color" and a set of values of this attribute consisting of "green", "yellow", etc. It should be noted that relational databases can be viewed as information systems in the sense of Pawlak.

In an information system each subset of attributes defines an indiscernibility relation, which is an equivalence on the object set such that two objects are equivalent when their values of all attributes in the set are the same. It may turn out that a proper subset of a set of attributes classifies the objects with the same accuracy as the original set, which means that some attributes may be omitted. An attribute set C is a reduct of an attribute set B, if C is a minimal subset of B which defines the same indiscernibility relation as B. The reduction problem means that we want to enumerate all reducts of a given subset of attributes.

This work is devoted to the reduction problem in a dependence space. It is based on some papers of the same author, in particular on [5]. The fundamental notion appearing in the present paper is the concept of a dense family of a dependence space. We prove that our definition of dense families agrees with the definition presented earlier in the literature [10]; this result appeared also in [7]. Proposition 4.1 characterizes reducts in dependence spaces by the means of dense families. Also

*Turku Centre for Computer Science TUCS, Lemminkäisenkatu 14, FIN-20520 Turku, Finland. Email: jjarvine@cs.utu.fi

difference functions are defined by using dense families (cf. [5]). Proposition 5.2 characterizes reducts by the means of difference functions and Proposition 6.1 contains a construction of a dense family in the dependence space of an information system starting with its indiscernibility matrix; this result appeared also in [6].

As stated above, this paper gives a survey of some results concerning reducts and their construction, but the presented formulations are simpler than the formulations published earlier. It also completes proofs of some theorems published without proofs in the quoted papers.

2 Preliminaries

All general lattice theoretical and algebraic notions used in this paper can be found in [2, 4], for example. An oredered set (many authors use the shorthand poset) $(P, \leq)$ is a *join-semilattice* if the join $a \vee b$ exists for all $a, b \in P$. An equivalence relation Θ on P is a *congruence* relation on the semilattice $(P, \vee)$ if $a_1 \Theta b_1$ and $a_2 \Theta b_2$ imply $(a_1 \vee a_2)\Theta(b_1 \vee b_2)$ for all $a_1, a_2, b_1, b_2 \in P$. We denote by a/Θ the *congruence class* of a, that is, $a/\Theta = \{b \in P \mid a\Theta b\}$.

An ordered set $(P, \leq)$ is a *lattice* if $a \vee b$ and $a \wedge b$ exist for all $a, b \in P$. Let us consider a lattice $(P, \leq)$. An element $a \in P$ is *meet-irreducible* if $a = b \wedge c$ implies $a = b$ or $a = c$. We denote the set of all meet-irreducible elements $a \neq 1$ (in case P has a unit) of $(P, \leq)$ by $\mathcal{M}(P)$. The following lemma can be found in [2], for example.

Lemma 2.1. *If $(P, \leq)$ is a finite lattice, then*

$$a = \bigwedge\{x \in \mathcal{M}(P) \mid a \leq x\}$$

for all $a \in P$. □

Let $(P, \leq)$ be an ordered set. A subset S of P is *meet-dense* (see e.g. [2]), if for all $x \in P$ there exists a subset X of S such that $x = \bigwedge_P X$. Now the following lemma holds.

Lemma 2.2. *If $(P, \leq)$ is a finite lattice, then $S (\subseteq P)$ is meet-dense if and only if $\mathcal{M}(P) \subseteq S$.*

Proof. Let $S \subseteq P$ be meet-dense and $a \in \mathcal{M}(P)$. Since S is meet-dense and $a \neq 1$, there exists a finite nonempty subset $X = \{a_1, \ldots, a_n\}$ of S such that $a = a_1 \wedge \cdots \wedge a_n$. Because a is meet-irreducible, we obtain that $a \in X$ and so $a \in S$. Hence, $\mathcal{M}(P) \subseteq S$ holds.

Conversely, suppose that $\mathcal{M}(P) \subseteq S \subseteq P$. Then for all $a \in P$,

$$\{x \in \mathcal{M}(P) \mid a \leq x\} \subseteq \{x \in S \mid a \leq x\} \subseteq \{x \in P \mid a \leq x\},$$

which implies

$$\begin{aligned} a &= \bigwedge\{x \in \mathcal{M}(P) \mid a \leq x\} \geq \bigwedge\{x \in S \mid a \leq x\} \\ &\geq \bigwedge\{x \in P \mid a \leq x\} = a. \end{aligned}$$

Hence $a = \bigwedge\{x \in S \mid a \leq x\}$. This means that S is meet-dense. □

Let $(P, \leq)$ be an ordered set and $a, b \in P$. We say that a is *covered by* b, and write $a \prec b$, if $a < b$ and $a \leq c < b$ implies $a = c$. It is known (see e.g. [2]) that in a finite lattice $(P, \leq)$ the set of the elements of $(P, \leq)$ covered by exactly one element of P is $\mathcal{M}(P)$. Thus, by Lemma 2.2, a subset of a finite lattice $(P, \leq)$ is meet-dense if and only if it contains all elements of P which are covered by exactly one element of P.

A family $\mathcal{L}$ of subsets of a set A is said to be a *closure system* on A if $\mathcal{L}$ is closed under intersections, which means that for all $\mathcal{H} \subseteq \mathcal{L}$, we have $\bigcap \mathcal{H} \in \mathcal{L}$. We denote by $\wp(A)$ the *power set* of A, i.e., the set of all subsets of A. A *closure operator* on a set A is an extensive, idempotent and isotone map $\mathcal{C}\colon \wp(A) \to \wp(A)$; that is to say, $B \subseteq \mathcal{C}(B)$, $\mathcal{C}(\mathcal{C}(B)) = \mathcal{C}(B)$, and $B \subseteq C$ implies $\mathcal{C}(B) \subseteq \mathcal{C}(C)$ for all $B, C \subseteq A$. A subset B of A is *closed* (with respect to $\mathcal{C}$) if $\mathcal{C}(B) = B$. A closure system $\mathcal{L}$ on A defines a closure operator $\mathcal{C}_\mathcal{L}$ on A by the rule

$$\mathcal{C}_\mathcal{L}(B) = \bigcap\{X \in \mathcal{L} \mid B \subseteq X\}.$$

Conversely, if $\mathcal{C}$ is a closure operator on A, then the family

$$\mathcal{L}_\mathcal{C} = \{B \subseteq A \mid \mathcal{C}(B) = B\}$$

of closed subsets of A is a closure system. The relationship between closure systems and closure operators is bijective; the closure operator induced by the closure system $\mathcal{L}_\mathcal{C}$ is $\mathcal{C}$ itself, and the closure system induced by the closure operator $\mathcal{C}_\mathcal{L}$ is $\mathcal{L}$. It is well-known that if $\mathcal{L}$ is a closure system on A, then the ordered set $(\mathcal{L}, \subseteq)$ is a lattice in which

$$X \wedge Y = X \cap Y \quad \text{and} \quad X \vee Y = \mathcal{C}_\mathcal{L}(X \cup Y)$$

for all $X, Y \in \mathcal{L}$.

Next we consider meet-dense subsets of the lattice $(\mathcal{L}, \subseteq)$, where $\mathcal{L}$ is a closure system on a finite set.

Proposition 2.3. *Let $\mathcal{T}$ be a meet-dense subset of a lattice $(\mathcal{L}, \subseteq)$, where $\mathcal{L}$ is a closure system on a finite set A.*

(a) *For all $B \subseteq A$, $\mathcal{C}_\mathcal{L}(B) = \bigcap\{X \in \mathcal{T} \mid B \subseteq X\}$.*

(b) *For all $B, C \subseteq A$ the following three conditions are equivalent:*

- (i) *$\mathcal{C}_\mathcal{L}(B) \subseteq \mathcal{C}_\mathcal{L}(C)$;*
- (ii) *for all $X \in \mathcal{T}$, $C \subseteq X$ implies $B \subseteq X$;*
- (iii) *for all $X \in \mathcal{T}$, $B - X \neq \emptyset$ implies $C - X \neq \emptyset$.*

Proof. (a) Because $\mathcal{C}_\mathcal{L}(B) \in \mathcal{L}$, and $B \subseteq X$ if and only if $\mathcal{C}_\mathcal{L}(B) \subseteq X$ for all $X \in \mathcal{L}$, we obtain by Lemmas 2.1 and 2.2 that

$$\begin{aligned} \mathcal{C}_\mathcal{L}(B) &= \bigcap\{X \in \mathcal{M}(\mathcal{L}) \mid \mathcal{C}_\mathcal{L}(B) \subseteq X\} = \bigcap\{X \in \mathcal{M}(\mathcal{L}) \mid B \subseteq X\} \\ &\supseteq \bigcap\{X \in \mathcal{T} \mid B \subseteq X\} \supseteq \bigcap\{X \in \mathcal{L} \mid B \subseteq X\} = \mathcal{C}_\mathcal{L}(B). \end{aligned}$$

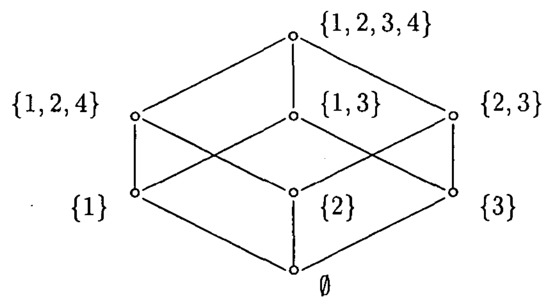

Figure 1: The closure lattice $(\mathcal{L}_\mathcal{D}, \subseteq)$

Hence, $\mathcal{C}_\mathcal{L}(B) = \{X \in \mathcal{T} \mid B \subseteq X\}$.

(b) Let $\mathcal{C}_\mathcal{L}(B) \subseteq \mathcal{C}_\mathcal{L}(C)$. If $X \in \mathcal{T}$ and $C \subseteq X$, then $B \subseteq \mathcal{C}_\mathcal{L}(B) \subseteq \mathcal{C}_\mathcal{L}(C) \subseteq \mathcal{C}_\mathcal{L}(X) = X$. Conversely, if for all $X \in \mathcal{T}$, $C \subseteq X$ implies $B \subseteq X$, then $\{X \in \mathcal{T} \mid C \subseteq X\} \subseteq \{X \in \mathcal{T} \mid B \subseteq X\}$. Hence by (a), $\mathcal{C}_\mathcal{L}(B) = \bigcap\{X \in \mathcal{T} \mid B \subseteq X\} \subseteq \bigcap\{X \in \mathcal{T} \mid C \subseteq X\} = \mathcal{C}_\mathcal{L}(C)$. Thus, (i) and (ii) are equivalent. Also (ii) and (iii) are equivalent since for all $X, Y \subseteq A$, $Y \subseteq X$ if and only if $Y - X = \emptyset$. □

3 Dense families of dependence spaces

We recall Novotný's and Pawlak's [9] definition of dependence spaces. We note that in [7] Järvinen studied infinite dependence spaces.

Definition. If A is a finite nonempty set and Θ is a congruence on the semilattice $(\wp(A), \cup)$, then the ordered pair $\mathcal{D} = (A, \Theta)$ is said to be a *dependence space.*

Let $\mathcal{D} = (A, \Theta)$ be a dependence space. Recalling the finiteness of A, it is clear that for every $B(\subseteq A)$, the congruence class B/Θ has a greatest element $\mathcal{C}_\mathcal{D}(B) = \bigcup B/\Theta$. It was noted in [8] that for all $B, C \subseteq A$,

$$B \Theta C \quad \text{if and only if} \quad \mathcal{C}_\mathcal{D}(B) = \mathcal{C}_\mathcal{D}(C).$$

In [8] it was also observed that $\mathcal{C}_\mathcal{D}: \wp(A) \to \wp(A), B \mapsto \bigcup B/\Theta$ is a closure operator on A. We denote by $\mathcal{L}_\mathcal{D}$ the closure system corresponding to the closure operator $\mathcal{C}_\mathcal{D}$. Hence, the family $\mathcal{L}_\mathcal{D}$ consists of the greatest elements of the Θ-classes.

Example 3.1. Let $A = \{1,2,3,4\}$ and Θ be the congruence relation on $(\wp(A), \cup)$ whose congruence classes are $\{\emptyset\}$, $\{\{1\}\}$, $\{\{2\}\}$, $\{\{3\}\}$, $\{\{4\}, \{1,2\}, \{1,4\}, \{2,4\}, \{1,2,4\}\}$, $\{\{1,3\}\}$, $\{\{2,3\}\}$ and $\{\{3,4\}, \{1,2,3\}, \{1,3,4\}, \{2,3,4\}, \{1,2,3,4\}\}$. The closure lattice $(\mathcal{L}_\mathcal{D}, \subseteq)$ corresponding to the dependence space $\mathcal{D} = (A, \Theta)$ is presented in Figure 1. Moreover, $\mathcal{M}(\mathcal{L}_\mathcal{D}) = \{\{1,2,4\}, \{1,3\}, \{2,3\}\}$.

Dense families of dependence spaces were introduced in [10]. Here we define them differently as meet-dense subsets of the lattice $(\mathcal{L}_\mathcal{D}, \subseteq)$; recall that in $(\mathcal{L}_\mathcal{D}, \subseteq)$, $X \wedge Y = X \cap Y$ for all $X, Y \in \mathcal{L}_\mathcal{D}$. We will also show that our definition agrees with Novotný's definition of dense families.

Definition. Let $\mathcal{D} = (A, \Theta)$ be a dependence space. A family $\mathcal{T} \subseteq \wp(A)$ is *dense* in $\mathcal{D}$ if it is a meet-dense subset of the lattice $(\mathcal{L}_\mathcal{D}, \subseteq)$.

By Lemma 2.2, a family $\mathcal{T}$ is dense in $\mathcal{D}$ if and only if it is a subfamily of $\mathcal{L}_\mathcal{D}$ which contains all elements of the lattice $(\mathcal{L}_\mathcal{D}, \subseteq)$ which are covered by exactly one element of $\mathcal{L}_\mathcal{D}$.

Example 3.2. Let us consider the dependence space $\mathcal{D} = (A, \Theta)$ of Example 3.1. The Hasse diagram of $(\mathcal{L}_\mathcal{D}, \subseteq)$ is given in Figure 1. The dense families of $\mathcal{D}$ are the 32 families $\mathcal{T}$ such that $\mathcal{M}(\mathcal{L}_\mathcal{D}) \subseteq \mathcal{T} \subseteq \mathcal{L}_\mathcal{D}$.

Let A be a set. Each family $\mathcal{T} \subseteq \wp(A)$ defines a binary relation $\Gamma(\mathcal{T})$ on $\wp(A)$:

$$(B, C) \in \Gamma(\mathcal{T}) \text{ if and only if } (\forall X \in \mathcal{T})\ B \subseteq X \iff C \subseteq X.$$

We note that in [10] dense families were defined by the condition presented in the next proposition.

Proposition 3.3. *Let $\mathcal{D} = (A, \Theta)$ be a dependence space. A family $\mathcal{T} \subseteq \wp(\mathcal{A})$ is dense in $\mathcal{D}$ if and only if $\Gamma(\mathcal{T}) = \Theta$.*

Proof. Let $\mathcal{T}$ be dense and $B \Theta C$. Then $\mathcal{C}_\mathcal{D}(B) = \mathcal{C}_\mathcal{D}(C)$, which implies by Proposition 2.3(b) that for all $X \in \mathcal{T}$, $B \subseteq X$ iff $C \subseteq X$. Thus, $\Theta \subseteq \Gamma(\mathcal{T})$. Conversely, if $(B, C) \in \Gamma(\mathcal{T})$, then

$$\mathcal{C}_\mathcal{D}(B) = \bigcap\{X \in \mathcal{T} \mid B \subseteq X\} = \bigcap\{X \in \mathcal{T} \mid C \subseteq X\} = \mathcal{C}_\mathcal{D}(C),$$

which is equivalent to $B \Theta C$. Hence, also $\Gamma(\mathcal{T}) \subseteq \Theta$.

On the other hand, let $\Gamma(\mathcal{T}) = \Theta$. We will show that $\mathcal{M}(\mathcal{L}_\mathcal{D}) \subseteq \mathcal{T} \subseteq \mathcal{L}_\mathcal{D}$, which implies by Lemma 2.2 that $\mathcal{T}$ is a meet-dense subset of $(\mathcal{L}_\mathcal{D}, \subseteq)$. Suppose that $X \in \mathcal{T}$. Because $X \Theta \mathcal{C}_\mathcal{D}(X)$ and $X \subseteq X$, we obtain $\mathcal{C}_\mathcal{D}(X) \subseteq X$, which implies $X \in \mathcal{L}_\mathcal{D}$. Hence, $\mathcal{T} \subseteq \mathcal{L}_\mathcal{D}$.

Assume that $\mathcal{M}(\mathcal{L}_\mathcal{D}) \not\subseteq \mathcal{T}$. This means that there exists a $Y \in \mathcal{M}(\mathcal{L}_\mathcal{D})$ such that $Y \notin \mathcal{T}$. Since $Y \in \mathcal{M}(\mathcal{L}_\mathcal{D})$, there exists exactly one $Z \in \mathcal{L}_\mathcal{D}$ such that $Y \prec Z$ holds in $\mathcal{L}_\mathcal{D}$. For all $X \in \mathcal{T}$, $Z \subseteq X$ implies obviously that $Y \subseteq X$. Suppose that there is an $X \in \mathcal{T}$ such that $Y \subseteq X$ but $Z \not\subseteq X$. Since $X, Z \in \mathcal{L}_\mathcal{D}$, we get $X \cap Z \in \mathcal{L}_\mathcal{D}$ and $Y \subseteq X \cap Z \subset Z$. The fact that $Y \prec Z$ holds in $\mathcal{L}_\mathcal{D}$ implies $Y = X \cap Z$. Because Y is meet-irreducible, we obtain $Y = X$ or $Y = Z$. Obviously both of these equalities lead to a contradiction! Hence, for all $X \in \mathcal{T}$, also $Y \subseteq X$ implies $Z \subseteq X$. Thus, $(Y, Z) \in \Gamma(\mathcal{T}) = \Theta$, which means that $Y = \mathcal{C}_\mathcal{D}(Y) = \mathcal{C}_\mathcal{D}(Z) = Z$, a contradiction! Therefore, also $\mathcal{M}(\mathcal{L}_\mathcal{D}) \subseteq \mathcal{T}$ holds. □

4 Independent sets and reducts

In this section we review independent sets and reducts defined in dependence spaces. Further references can be found in [8, 9, 10], for example. Our main result of this section gives a characterization of the reducts of a given subset of a dependence space in terms of dense families.

Let $\mathcal{D} = (A, \Theta)$ be a dependence space. A subset $B(\subseteq A)$ is called *independent*, if B is minimal with respect to inclusion in its Θ-class. We denote the set of all independent subsets of $\mathcal{D}$ by $IND_{\mathcal{D}}$.

The notion of reducts is important in the theory of Pawlak's information systems. Here we study reducts in the more algebraic setting of dependence spaces. For any $B(\subseteq A)$ a set $C(\subseteq A)$ is called a *reduct* of B, if $C \subseteq B$, $B\Theta C$ and $C \in IND_{\mathcal{D}}$. The set of all reducts of B will be denoted by $RED_{\mathcal{D}}(B)$. In the other words, a subset C $(\subseteq B)$ is a reduct of B, if C is minimal in B/Θ with respect to inclusion. Because A is finite, it is obvious that every set has at least one reduct.

Finding all reducts of a given set is called the *reduction problem.* Our next proposition, which appears without a proof also in [6], characterizes the reducts of a given set by the means of dense families.

Proposition 4.1. *Let $\mathcal{T}$ be a dense family in a dependence space $\mathcal{D} = (A, \Theta)$. If $B \subseteq A$, then $C \in RED_{\mathcal{D}}(B)$ if and only if C is minimal set with respect to the property of containing an element from each nonempty difference $B - X$, where $X \in \mathcal{T}$.*

Proof. Let C be a minimal set which contains an element from each nonempty difference $B - X$, $X \in \mathcal{T}$. First we show that $C \subseteq B$. If $C \not\subseteq B$, then $B \cap C \subset C$ and $(B \cap C) \cap (B - X) = C \cap (B - X) \neq \emptyset$ whenever $B - X \neq \emptyset$, a contradiction! Thus, $C \subseteq B$. Now $C - X = (B \cap C) - X = C \cap (B - X) \neq \emptyset$ for all $X \in \mathcal{T}$ such that $B - X \neq \emptyset$. This implies by Proposition 2.3(b) that $\mathcal{C}_{\mathcal{D}}(B) \subseteq \mathcal{C}_{\mathcal{D}}(C)$. The inclusion $\mathcal{C}_{\mathcal{D}}(C) \subseteq \mathcal{C}_{\mathcal{D}}(B)$ is obvious. Hence, $B\Theta C$. Assume that $C \notin IND_{\mathcal{D}}$. Then there exists a $D \subset C$ such that $C\Theta D$. Since Θ is transitive, we obtain $B\Theta D$ and in particular $\mathcal{C}_{\mathcal{D}}(B) \subseteq \mathcal{C}_{\mathcal{D}}(D)$. This implies by Proposition 2.3(b) that $D \cap (B - X) = D - X \neq \emptyset$ whenever $B - X \neq \emptyset$, a contradiction! Hence, C is independent and thus C is a reduct of B.

On the other hand, suppose $C \in RED_{\mathcal{D}}(B)$. Then $C \subseteq B$, $B\Theta C$, $C \in IND_{\mathcal{D}}$, and especially $\mathcal{C}_{\mathcal{D}}(B) \subseteq \mathcal{C}_{\mathcal{D}}(C)$. This implies that $C \cap (B - X) = (B \cap C) - X = C - X \neq \emptyset$ for all $X \in \mathcal{T}$ which satisfy $B - X \neq \emptyset$. Assume that there exists a $D \subset C$ which contains an element from each nonempty difference $B - X$, where $X \in \mathcal{T}$. Then $D - X = (B \cap D) - X = D \cap (B - X) \neq \emptyset$ for all $X \in \mathcal{T}$ such that $B - X \neq \emptyset$. Hence, $\mathcal{C}_{\mathcal{D}}(B) \subseteq \mathcal{C}_{\mathcal{D}}(D)$. Since $D \subset B$ also $\mathcal{C}_{\mathcal{D}}(D) \subseteq \mathcal{C}_{\mathcal{D}}(B)$ holds. This implies $B\Theta D$, and because $C\Theta B$ we obtain $C\Theta D$, a contradiction! □

Example 4.2. Let us consider the dependence space $\mathcal{D} = (A, \Theta)$ defined in Example 3.1. We have already noted that $\mathcal{M}(\mathcal{L}_{\mathcal{D}}) = \{\{1, 2, 4\}, \{1, 3\}, \{2, 3\}\}$ is the smallest dense family.

Next we find the reducts of A. The differences $A - X$, where $X \in \mathcal{M}(\mathcal{L}_{\mathcal{D}})$, are

$$A - \{1,2,4\} = \{3\}, \;\; A - \{1,3\} = \{2,4\}, \text{ and } \; A - \{2,3\} = \{1,4\}.$$

They are all nonempty. Because the reducts of A must contain an element from all of these differences, each reduct must include 3. It can be easily seen that $\{1,2,3\}$ and $\{3,4\}$ are the reducts of A.

5 The difference function

In this section we study the notion of difference function. Difference functions were introduced in [5]. Here we give an equivalent, but a clearer definition. First we recall some notions concerning Boolean functions (see e.g. [1], where further references can be found). A *Boolean function*, or a *function* for short, is a mapping $f : \{0,1\}^n \to \{0,1\}$. An element $v \in \{0,1\}^n$ is called a *Boolean vector* (a *vector* for short). If $f(v) = 1$ (resp. 0), then v is called a *true* (resp. *false*) vector of f. The set of all true vectors (resp. false vectors) of f is denoted by $T(f)$ (resp. $F(f)$).

Let $u = (u_1, \ldots, u_n)$ and $v = (v_1, \ldots, v_n)$ be vectors. We set $u \leq v$ if and only if $u_i \leq v_i$, for $1 \leq i \leq n$. A function f is *isotone* if $u \leq v$ always implies $f(u) \leq f(v)$.

In the sequel we assume that f is an isotone function. A true vector v of f is *minimal* if there is no true vector w such that $w < v$, and let $\min T(f)$ denote the set of all minimal true vectors of f. A *maximal* false vector is symmetrically defined and $\max F(f)$ denotes the set of all maximal false vectors of f.

Let $\mathcal{D} = (A, \Theta)$ be a dependence space such that $A = \{a_1, \ldots, a_n\}$ and let $\mathcal{T}$ be dense in $\mathcal{D}$. For any $B \subseteq A$, let $\delta(B)$ denote the disjunction of all variables y_i, where $a_i \in B$. We define the *difference function* $f_B^{\mathcal{T}}(y_1, \ldots, y_n)$ as the conjunction

$$\bigwedge_{\substack{X \in \mathcal{T} \\ B - X \neq \emptyset}} \delta(B - X).$$

Clearly, the function $f_B^{\mathcal{T}}$ is isotone. A function $\chi : \wp(A) \to \{0,1\}^n$ is defined by

$$B \mapsto (\chi_1(B), \ldots, \chi_n(B)),$$

where

$$\chi_i(B) = \begin{cases} 0 & \text{if } a_i \notin B \\ 1 & \text{if } a_i \in B \end{cases}$$

for all i, $1 \leq i \leq n$. The value $\chi(B)$ is called the *characteristic vector* of B.

Now the following lemma holds.

Lemma 5.1. *Let $\mathcal{T}$ be a dense family in a dependence space $\mathcal{D} = (A, \Theta)$. For all $B, C \subseteq A$, the following conditions are equivalent:*
(a) $\chi(C) \in T(f_B^{\mathcal{T}})$;
(b) *C contains an element from each nonempty difference $B - X$, $X \in \mathcal{T}$.*

Proof. Let $B, C \subseteq A$ and $\{X \in \mathcal{T} \mid B - X \neq \emptyset\} = \{X_1 \ldots X_k\}$.

(a) $\Rightarrow$ (b) Assume that $f_B^{\mathcal{T}}(\chi(C)) = 1$. If $C \cap (B - X_i) = \emptyset$ for some i, $1 \leq i \leq k$, then obviously the disjunction $\delta(B - X_i)$ has the value 0 for $\chi(C)$. This implies that also the conjunction $\bigwedge_{1\leq i\leq k}\delta(B-X_i)$ has the value 0 for $\chi(C)$, a contradiction! Hence, $C\cap(B-X_i)\neq\emptyset$ for all i, $1\leq i\leq k$.

(b) $\Rightarrow$ (a) Suppose that $C\cap(B-X_i)\neq\emptyset$ for all $1\leq i\leq n$. Then for all $1\leq i\leq n$, the disjunction $\delta(B-X_i)$ has the value 1 for $\chi(C)$. This implies that the conjunction $\bigwedge_{1\leq i\leq k}\delta(B-X_i)$ has the value 1 for $\chi(C)$, i.e., $f_B^{\mathcal{T}}(\chi(C))=1$. □

Now we can write the following proposition. Note that for any $X\subseteq A$, $X^{\mathsf{C}}=A-X$ is the *complement* of X.

Proposition 5.2. *Let $\mathcal{T}$ be a dense family in a dependence space $\mathcal{D}=(A,\Theta)$. If $B\subseteq A$, then*

(a) $\min T(f_B^{\mathcal{T}})=\{\chi(C)\mid C\in RED_{\mathcal{D}}(B)\}$ *and*

(b) $\max F(f_B^{\mathcal{T}})=\max\{\chi((B-X)^{\mathsf{C}})\mid X\in\mathcal{T},B-X\neq\emptyset\}$.

Proof. Let us denote $f_B^{\mathcal{T}}$ simply by f.

(a) Let $v\in\min T(f)$ and let C be the subset of A which satisfies $\chi(C)=v$. By Lemma 5.1 C contains an element from each nonempty difference $B-X$, where $X\in\mathcal{T}$. Assume that C is not minimal set with respect to that property, that is, there exist a $D\subset C$ which also contains an element from each nonempty difference $B-X$, where $X\in\mathcal{T}$. By Lemma 5.1 this implies that $\chi(D)\in T(f)$. But $D\subset C$ implies $\chi(D)<\chi(C)$ and hence $\chi(C)\notin\min T(f)$, a contradiction! Therefore, C is minimal set with respect to the property of containing an element from each nonempty difference $B-X$, where $X\in\mathcal{T}$. This implies that C is a reduct of B by Proposition 4.1.

On the other hand, suppose that C is a reduct of B. Then C contains an element from each nonempty difference $B-X$, where $X\in\mathcal{T}$, and thus $\chi(C)\in T(f)$. Suppose that $\chi(C)\notin\min T(f)$. This means that there exists a vector $v\in T(f)$ such that $v<\chi(C)$. Let D be the subset of A which satisfies $\chi(D)=v$. Then obviously D is a set which contains an element from each nonempty difference $B-X$, where $X\in\mathcal{T}$. Since $D\subset C$, C is not a reduct of B, a contradiction! Hence, $\chi(C)\in\min T(f)$.

(b) By Lemma 5.1 it is obvious that $f(\chi(C))=0$ if and only if there exist an $X\in\mathcal{T}$ such that $B-X\neq\emptyset$ and $C\cap(B-X)=\emptyset$. This is equivalent to the condition that $f(\chi(C))=0$ if and only if there exist an $X\in\mathcal{T}$ such that $B-X\neq\emptyset$ and $C\subseteq(B-X)^{\mathsf{C}}$.

Suppose that $\chi(C)\in\max F(f)$. This implies that $C\subseteq(B-X)^{\mathsf{C}}$ for some $X\in\mathcal{T},B-X\neq\emptyset$, and hence $\chi(C)\leq\chi((B-X)^{\mathsf{C}})$. Assume that $\chi(C)<\chi((B-X)^{\mathsf{C}})$. Since $\chi((B-X)^{\mathsf{C}})\in F(f)$, this implies $\chi(C)\notin\max F(f)$, a contradiction! Hence, $\chi(C)\in\{\chi((B-X)^{\mathsf{C}})\mid X\in\mathcal{T},B-X\neq\emptyset\}$. Assume that there exists a $\chi(D)\in\{\chi((B-X)^{\mathsf{C}})\mid X\in\mathcal{T},B-X\neq\emptyset\}$ such that $\chi(C)<\chi(D)$. Clearly, this implies that $\chi(D)\in F(f)$ and hence $\chi(C)\notin\max F(f)$, a contradiction! Thus, $\chi(C)\in\max\{\chi((B-X)^{\mathsf{C}})\mid X\in\mathcal{T},B-X\neq\emptyset\}$.

Conversely, suppose that $\chi(C)\in\max\{\chi((B-X)^{\mathsf{C}})\mid X\in\mathcal{T},B-X\neq\emptyset\}$. Then obviously $\chi(C)\in F(f)$. Assume that there exists a $\chi(D)\in F(f)$ such that $\chi(C)<\chi(D)$. This implies that there exists an $X\in\mathcal{T}$ such that $D\subseteq(B-X)^{\mathsf{C}}$

and $B - X \neq \emptyset$. We obtain that $\chi(C) < \chi(D) \leq \chi((B-X)^{\complement})$ for some $X \in \mathcal{T}$ such that $B - X \neq \emptyset$, a contradiction! Hence, $\chi(C) \in \max F(f)$. □

Hence, the minimal true vectors of the difference function of $B(\subseteq A)$ are the characteristic vectors of the reducts of B.

Example 5.3. Let us consider the dependence space $\mathcal{D} = (A, \Theta)$ defined in Example 3.1. The family $\mathcal{T} = \{\{1,2,4\},\{1,3\},\{2,3\}\}$ is known to be dense in $\mathcal{D}$. The differences $A - X$ are all nonempty for all $X \in \mathcal{T}$. Hence,

$$\begin{aligned} f_A^{\mathcal{T}} &= \delta(A - \{1,2,4\}) \wedge \delta(A - \{1,3\}) \wedge \delta(A - \{2,3\}) \\ &= 3 \wedge (2 \vee 4) \wedge (1 \vee 4), \end{aligned}$$

where i stands for y_i. The function $f_A^{\mathcal{T}}$ has the minimal true vectors $(0,0,1,1)$ and $(1,1,1,0)$, which implies by Proposition 5.2 that $RED_{\mathcal{D}}(A) = \{\{3,4\},\{1,2,3\}\}$.

The *dual* of a Boolean function f, denoted by f^d, is defined by $f^d(x) = \overline{f}(\overline{x})$, where $\overline{f}$ and $\overline{x}$ denote the complements of f and x, respectively. It is well-known that $(f^d)^d = f$ and that the DNF expression of f^d is obtained from that of f by exchanging $\vee$ and $\wedge$ as well as constants 0 and 1, and then expanding the resulting formula. For example, the dual of $g = 3\vee(1\wedge4)\vee(2\wedge4)$ is $g^d = 3\wedge(1\vee4)\wedge(2\vee4) = (3\wedge4)\vee(1\wedge2\wedge3)$.

It is known (see e.g. [1]) that for any isotone Boolean function f, $\min T(f^d) = \{\overline{v} \mid v \in \max F(f)\}$. Let us denote $f_B^{\mathcal{T}}$ simply by f. By Proposition 5.2:

$$\begin{aligned} v \in \min T(f^d) &\iff \overline{v} \in \max F(f) \\ &\iff \overline{v} \in \max\{\chi((B-X)^{\complement}) \mid X \in \mathcal{T}, B - X \neq \emptyset\} \\ &\iff v \in \min\{\chi(B-X) \mid X \in \mathcal{T}, B - X \neq \emptyset\}. \end{aligned}$$

The family $\mathcal{T} = \{\{1,2,4\},\{1,3\},\{2,3\}\}$ is known to be dense in the dependence space $\mathcal{D}$ of Example 3.1. Let us denote by f the difference function of the set A. Then

$$\begin{aligned} \min(f^d) &= \min\{\chi(A-X) \mid X \in \mathcal{T}, A - X \neq \emptyset\} \\ &= \{(0,0,1,0),(0,1,0,1),(1,0,0,1)\}. \end{aligned}$$

This means that $f^d = 3 \vee (1\wedge 4) \vee (2 \wedge 4)$ and $f = (f^d)^d = (3\wedge4)\vee(1\wedge2\wedge3)$. Hence, $\min T(f) = \{(0,0,1,1),(1,1,1,0)\}$, as stated in Example 5.3.

Remark. Let $f = f(x_1,\ldots,x_n)$ and $g = g(x_1,\ldots,x_n)$ be a pair of isotone Boolean functions given by their minimal true vectors $\min T(f)$ and $\min T(g)$, respectively. Let us consider the following problem; test whether f and g are mutually dual. In [3] Fredman and Khachiyan showed that this problem can be solved in time $k^{o(\log k)}$, where $k = |\min T(f)| + |\min T(g)|$.

This implies that for an isotone Boolean function f given by its minimal true vectors and for a subset $G \subset \min T(f^d)$, a new vector $v \in \min T(f^d) - G$ can be computed in time $nk^{o(\log k)}$, where $k = |\min T(f)| + |G|$ (see [1], for example). This means also that for any isotone Boolean function f given by its minimal true vectors, f^d can be computed in time $nk^{o(\log k)}$, where $k = |\min T(f)| + |\min T(f^d)|$.

6 An application to information systems

An *information system* is a triple $\mathcal{S} = (U, A, \{V_a\}_{a \in A})$, where U is a set of *objects*, A is a set of *attributes*, and $\{V_a\}_{a \in A}$ is an indexed set of *value sets of attributes.* All these sets are assumed to be finite and nonempty. Each attribute is a function $a: U \to V_a$ which assigns a value of the attribute a to objects (see e.g. [9, 10, 11]).

For any $B \subseteq A$, the *indiscernibility relation* of B is defined by

$$I(B) = \{(x, y) \in U^2 \mid a(x) = a(y) \text{ for all } a \in B\}.$$

It is known that $I(B)$ is an equivalence relation on U such that its equivalence classes consist of objects which are indiscernible with respect to all attributes in B. Let us define the following binary relation $\Theta_{\mathcal{S}}$ on the set $\wp(A)$:

$$(B, C) \in \Theta_{\mathcal{S}} \iff I(B) = I(C).$$

So, two subsets of attributes are in the relation $\Theta_{\mathcal{S}}$ if and only if they define the same indiscernibility relation. It is known (see e.g. [8, 9]) that $\Theta_{\mathcal{S}}$ is a congruence on the semilattice $(\wp(A), \cup)$. Hence, the pair $\mathcal{D}_{\mathcal{S}} = (A, \Theta_{\mathcal{S}})$ is a dependence space. It can be easily seen that C ($\subseteq A$) is a reduct of B ($\subseteq A$) in the dependence space $\mathcal{D}_{\mathcal{S}}$ if and only if C is a minimal subset of B which defines the same indiscernibility relation as B.

Assume that $U = \{x_1, \ldots, x_m\}$. Then the *indiscernibility matrix* of $\mathcal{S}$ is an $m \times m$-matrix $\mathbf{M}_{\mathcal{S}} = (c_{ij})_{m \times m}$ such that

$$c_{ij} = \{a \in A \mid a(x) = a(y)\}$$

for all $1 \leq i, j \leq m$. Thus, the entry c_{ij} consists of the attributes which do not discern objects x_i and x_j (cf. discernibility matrices defined in [12]). It is now trivial that

$$(x_i, x_j) \in I(B) \iff B \subseteq c_{ij}.$$

Next we show how matrices of preimage relations induce dense families.

Proposition 6.1. *If* $\mathcal{S} = (U, A, \{V_a\}_{a \in A})$ *is an information system and* $\mathbf{M}_{\mathcal{S}} = (c_{ij})_{m \times m}$ *is the indiscernibility matrix of* $\mathcal{S}$*, then the family*

$$\mathcal{T}_{\mathcal{S}} = \{c_{ij} \mid 1 \leq i, j \leq m\}$$

is dense in the dependence space $\mathcal{D}_{\mathcal{S}} = (A, \Theta_{\mathcal{S}})$.

Proof. By Proposition 3.3 it suffices to show that $\Gamma(\mathcal{T}_S) = \Theta_S$. If $(B, C) \in \Theta_S$, then for all $1 \le i, j \le m$, $B \subseteq c_{ij}$ iff $(x_i, x_j) \in I(B)$ iff $(x_i, x_j) \in I(C)$ iff $C \subseteq c_{ij}$, which implies $(B, C) \in \Gamma(\mathcal{T}_S)$. Hence, $\Theta_S \subseteq \Gamma(\mathcal{T}_S)$.

If $(B, C) \in \Gamma(\mathcal{T}_S)$, then for all $1 \le i, j \le m$, $(x_i, x_j) \in I(B)$ iff $B \subseteq c_{ij}$ iff $C \subseteq c_{ij}$ iff $(x_i, x_j) \in I(C)$, which implies $I(B) = I(C)$. Thus, also $\Gamma(\mathcal{T}_S) \subseteq \Theta_S$ and hence $\Gamma(\mathcal{T}_S) = \Theta_S$. □

We conclude this paper by an example.

Example 6.2. Let us consider an information system $\mathcal{S} = (U, A, \{V_A\}_{a \in A})$, where the object set $U = \{1, 2, 3, 4, 5\}$ consists of five persons, the attribute set A consists of the attributes Age, Eyes, and Height, and the corresponding value sets are $V_{\text{Age}} = \{\text{Young, Middle, Old}\}$, $V_{\text{Eyes}} = \{\text{Blue, Brown, Green}\}$, and $V_{\text{Height}} = \{\text{Short, Normal, Tall}\}$.

Let the values of the attributes be defined as in the following table.

	Age	Eyes	Height
1	Young	Blue	Short
2	Young	Brown	Normal
3	Middle	Brown	Tall
4	Old	Green	Normal
5	Young	Brown	Normal

For example, the indiscernibility relation $I(A)$ of the attribute set A is an equivalence on U which has the equivalence classes $\{1\}$, $\{2, 5\}$, $\{3\}$, and $\{4\}$.

If we denote a = Age, b = Eyes, and c = Height, then the indiscernibility matrix of $\mathcal{S}$ is the following:

$$\mathbf{M}_S = \begin{pmatrix} A & \{a\} & \emptyset & \emptyset & \{a\} \\ \{a\} & A & \{b\} & \{c\} & A \\ \emptyset & \{b\} & A & \emptyset & \{b\} \\ \emptyset & \{c\} & \emptyset & A & \{c\} \\ \{a\} & A & \{b\} & \{c\} & A \end{pmatrix}$$

By Proposition 6.1, the family $\mathcal{T}_S = \{\emptyset, \{a\}, \{b\}, \{c\}, A\}$ consisting of the entries of $\mathbf{M}_S$ is dense in the dependence space $\mathcal{D}_S = (A, \Theta_S)$. Let us denote by f the difference function of the set A. Then $\min(f^d) = \min\{\chi(A - X) \mid X \in \mathcal{T}, A - X \neq \emptyset\} = \{(0,1,1), (1,0,1), (1,1,0)\}$. This means that $f^d = (b \wedge c) \vee (a \wedge c) \vee (a \wedge b)$ and $f = (b \vee c) \wedge (a \vee c) \wedge (a \vee b) = (a \wedge b) \vee (a \wedge c) \vee (b \wedge c)$. Obviously, $(1,1,0)$, $(1,0,1)$ and $(0,1,1)$ are the minimal true vectors of f. Thus, $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$ are the reducts of A in $\mathcal{D}_S$.

Acknowledgement

The author thanks the referee for his valuable comments and suggestions.

References

[1] J. C. Bioch, T. Ibaraki, *Complexity of identification and dualization of positive Boolean functions*, Information and Computation 123 (1995), 50–63.

[2] B. A. Davey, H.A. Priestley, *Introduction to lattices and order*, Cambridge University Press, Cambridge, 1990.

[3] M. Fredman, L. Khachiyan, *On the complexity of dualization of monotone disjunctive normal forms*, Journal of Algorithms 21 (1996), 618–628.

[4] G. Grätzer, *Lattice theory: first concepts and distributive lattices*, W. H. Freeman and company, San Francisco, 1971.

[5] J. Järvinen, *A representation of dependence spaces and some basic algorithms*, Fundamenta Informaticae 29 (1997), 369–382.

[6] J. Järvinen, *Preimage relations and their matrices*. In L. Polkowski, A. Skowron (eds.), *Rough sets and current trends in computing*, Lecture Notes in Artificial Intelligence 1424, Springer-Verlag, Berlin/Heidelberg, 139–146, 1998.

[7] J. Järvinen, *Knowledge representation and rough sets*, PhD Dissertation, University of Turku, Department of Mathematics, March 1999 (available at `http://www.cs.utu.fi/jjarvine`).

[8] J. Novotný, M. Novotný, *Notes on the algebraic approach to dependence in information systems*, Fundamenta Informaticae 16 (1992), 263–273.

[9] M. Novotný, Z. Pawlak, *Algebraic theory of independence in information systems*, Fundamenta Informaticae 14 (1991), 454–476.

[10] M. Novotný, *Applications of dependence spaces*. In E. Orłowska (ed.), *Incomplete information: rough set analysis*, Physica-Verlag, Heidelberg/New York, 247–289, 1998.

[11] Z. Pawlak, *Information systems. Theoretical foundations*, Informations Systems 6 (1981), 205–218.

[12] A. Skowron, C. Rauszer, *The discernibility matrices and functions in information systems*. In R. Slowinski (ed.), *Intelligent decision support. Handbook of applications and advances of the rough set theory*, Kluwer Academic Publisher, Dordrecht, 331–362, 1991.

Received November, 1999