Jiří Vaníček

Information aspects: A tool for redundancy control of data

Kybernetika, Vol. 24 (1988), No. 4, 293--306

Persistent URL: http://dml.cz/dmlcz/124418

KYBERNETIKA — VOLUME 24 (1988), NUMBER 4

INFORMATION ASPECTS
A Tool for Redundancy Control of Data

JIŘÍ VANÍČEK[1]

The contribution outlines the theory of information aspects, a tool for controlling the distinguishing capability of data in the initial stages of the data base design. Descriptors of an information space are defined as mappings from the set of objects into some set of attributes. The equivalence is defined on the set of descriptors by means of the same discerning capability. Classes of this equivalence are called aspects. Shared and compound aspects, independence, completeness, orthogonality, bases and support of an aspect with respect to the given basis are defined. The theory refers to the design of distributed data base on a conceptual level.

1. INTRODUCTION

One of the most important questions of the data base theory is the problem of independence and redundancy of data stored in the data base with respect to their discerning capability. Efficient tools for controlling the redundancy are prerequisite for any adequate design of distributed data base systems, both from the view of their efficiency and reliability.

Reducing the degree of redundancy decreases demands on memory space for storing data and facilitates their updating. After data independency is achieved, the problems of retaining consistency during maintenance are eliminated, but considerable price may be paid in the increased complexity of actual processing and in the cost of transferring the data between the nodes where the information is required and the nodes where it is stored. A controlled degree of redundancy may significantly enhance reliability of the system using information that can be derived from other nodes to reconstruct data stored in one processing node. Therefore the objective is not to eliminate the data redundancy completely, but to create tools for controlling it.

[1] The present work sums up research carried at the Research Institute of Mathematical Machines in Prague as a part of the project Research of Methods and Tools for Knowledge Representation.

This paper proposes principles of information aspects theory whose aim is to describe the independence and redundancy of data with respect to the discerning capatibility. The relationship between this capability and informational capability of attributes in the data base is described in [9] and [10]. All investigations in this paper have to do with the distinguishing capability in the sense of [9] regardless the term "information" is frequently used.

This paper employs some results and methods of the works [6], [7] and [8] which explore the so-called data spaces describing behaviour of abstract programs. The concept of data space is applied here as an ordered triple whose first number is the set of all possible states of a real computer, the second number is the set of all visible parameters (states of the virtual computer of a given level) and the third one is the set of mappings of the set of all states of a real computer into itself (abstract programs). In the above mentioned works the question of mutual independence of visible parameter has been studied from the view of need to describe behaviour of abstract programs. The present paper employs the method of mentioned works to describe the discerning capability of data while seeking a firmer mathematical basis for their ideas, especially with regard to the question of existence in the mathematical sense of the abstract objects used, which are defined by means of their properties.

2. INFORMATION SPACE, DESCRIPTORS, ASPECTS

2.1 The concept of information space (abbreviated as "space") is used to refer to an ordered couple (X, F), where X is a non-empty set of objects and F is a non-empty set of mappings f which are defined on some subset $D_f \subset X$, mapping it into an arbitrary set R_f. The elements of the set F are called information descriptors (abbreviated as "descriptors"). Ranges R_f of the values of individual descriptors f may differ.

2.2 By adding an artificial value denoted as ϑ ($\vartheta \notin R_f$ for each $f \in F$) to the range of values of all descriptors in the sense that $f(x) = \vartheta$ if and only if (abbreviated as "iff") $x \notin D_f$, i.e. iff the value of the descriptor f for the given object $x \in X$ has no sense, a state can be achieved whereby the descriptors are defined always over the whole set X, i.e. they are mappings of X onto the set $R_f^* = R_f \cup \{\vartheta\}$ for the individual $f \in F$. This addition of ϑ will be assumed throughout the following.

2.3 To examine information contents of the descriptors, it is necessary to realize that two descriptors f and g are equivalent from our point of view if they have the same discerning capability towards objects, i.e. if for any objects $x, y \in X$ there is $f(x) = f(y)$ iff $g(x) = g(y)$. This relation between descriptors occurs iff there exists a bijection φ of the set R_f onto the set R_g such that $g(x) = \varphi(f(x))$ for all $x \in X$. The complexity of the relevant realization of this bijection will be not significant for the following.

The relation $\approx \subset F \times F$, defined by $(f, g) \in \approx$ iff for all $x, y \in X$ it holds that

$$(f(x) = f(y)) \Leftrightarrow (g(x) = g(y)),$$

is apparently reflexive, transitive and symmetrical and therefore is an equivalence on F. This equivalence defines a decomposition of the set F into a family of mutually disjunct classes of equivalent descriptors. Such classes of mutually equivalent descriptors will be called information aspects (abbreviated as "aspects").

In the following, aspects and descriptors belonging to these classes will always be denoted by the same letters with the letter denoting an aspect in parentheses, i.e. $f \in (f)$. We shall focus on such properties of aspects which can be stated in terms of descriptors pertaining to them, but which do not depend on an actual choice of such descriptors. In such a case we shall write only e.g. "for $f \in (f)$ there is ..." without explicitly stating and verifying that the relevant property does not change if the descriptor f is replaced by another descriptor $f' \in (f), f' \neq f$.

The set of all aspects of the space (X, F) will be denoted by (F).

2.4 Every aspect $(f) \in (F)$ of the space (X, F) unambiguously specifies the relation $\sim_{(f)} \subset X \times X$ on the set X of all objects which is defined by $(x, y) \in \sim_{(f)}$ iff $f(x) = f(y)$ for $f \in (f)$.

For each $(f) \in (F)$ the relation $\sim_{(f)}$ is an equivalence on the set of all objects X. If we denote the set of all equivalences on X by $E(X) \subset \exp(X \times X)$, where $\exp(X \times X)$ is the set of all subsets of the Cartesian product $X \times X$, an injection defined by $\varphi((f)) = \sim_{(f)}$ of the set (F) into $E(X)$ results. The mapping assigning to each object $x \in X$ such a class (of the decomposition of X according to the equivalence $\varphi((f)) = \sim_{(f)}$) to which the element x belongs may be a new descriptor which, after being added to the set F, may be considered to be a natural representative of the aspect $(f) \in (F)$.

3. ORDERING OF ASPECTS

3.1 We shall use the term ordering in the sense of partial ordering, i.e. as a relation $\prec$ on a set A satisfying the following conditions:

(a) $a \prec a$ for any $a \in A$;

(2) if $a \prec b$ and $b \prec a$ for $a, b \in A$, then $a = b$;

(3) if $a \prec b$ and $b \prec c$ for $a, b, c \in A$, then also $a \prec c$.

The ordered couple $(A, \prec)$, where A is a non-empty set and $\prec$ is an ordering on A will be called an ordered set. Thus there may be incomparable elements in an ordered set.

Let us define on the set of all aspects of the space (X, F) the relation $\sqsubset \subset (F) \times (F)$ as follows:

$$((f), (g)) \in \sqsubset \text{ iff for every } x, y \in X, f \in (f), g \in (g) \text{ holds:}$$

$$(f(x) \neq f(y)) \Rightarrow (g(x) \neq g(y)).$$

Then $\sqsubset$ is an ordering on the set (F). The aspect (f) will be said to be coarser then the aspect (g) or (g) to be finer than (f), written as $(f) \sqsubset (g)$ or $(g) \sqsupset (f)$, iff for these aspects holds $((f), (g)) \in \sqsubset$.

3.2 If φ is the mapping assigning to $(f) \in (F)$ the equivalence $\sim_{(f)}$, described in the paragraph 2.4, i.e.

$$x \sim_{(f)} y \Leftrightarrow f(x) = f(y) \quad \text{for} \quad f \in (f)\,,$$

then for any $(f), (g) \in (F)$ holds $(f) \sqsubset (g)$ iff $\sim_{(f)} \supset \sim_{(g)}$. The injection φ is therefore a dual isomorphic mapping of $((F), \sqsubset)$ onto a subset of the ordered set $(E(X), \subset)$ of all equivalences on X, ordered by the set inclusion.

3.3 For any two equivalences $e_1, e_2 \in E(X)$ the relation $e_1 \cap e_2$ is again an equivalence on X. The set $e_1 \cup e_2 \subset X \times X$ is however the relation that need not comply with the law of transitivity (see (3) in 3.1) and in general is not an equivalence on X. However, the subset

$$E_{(e_1, e_2)} = \{e^* \in E(X) \colon e^* \supset e_1, e^* \supset e_2\}$$

of the set $E(X)$ has always the smallest element $(e_1 \cup e_2)^*$, for which the following holds:

(1) $(e_1 \cup e_2)^* \supset e_1, (e_1 \cup e_2)^* \supset e_2$;

(2) if $e^* \in E(X)$ and $e^* \supset e_1$, $e^* \supset e_2$ then $e^* \supset (e_1 \cup e_2)^*$.

This element $(e_1 \cup e_2)^*$ is called transitive cloasure of the union $e_1 \cup e_2$ and may also be characterized as satisfying the assertion:

$(x, y) \in (e_1 \cup e_2)^*$ iff there exists a finite sequence of objects $x_0, x_1, \ldots, x_n$ from X such that the following holds:

$$x_0 = x\,, \quad x_n = y\,, \quad \text{for every} \quad j = 0, 1, \ldots, n-1 \quad \text{either}$$

$$(x_j, x_{j+1}) \in e_1 \quad \text{or} \quad (x_j, x_{j+1}) \in e_2\,.$$

We may introduce the binary operations $\vee$ and $\wedge$ on the set of all equivalences on the set X of objects using following definitions:

$$e_1 \vee e_2 = (e_1 \cup e_2)^*\,, \quad e_1 \wedge e_2 = e_1 \cap e_2\,.$$

The ordered triple $(E(X), \vee, \wedge)$ is then a lattice, where for every set $E \subset E(X)$ there exists the smallest element of the set $\{e' \in E(X) \colon e' \supset e \text{ for every } e \in E\}$, denoted as

$$\bigvee_{e \in E} e = \sup E$$

and the greatest element of the set $\{e' \in E(X) \colon e' \subset e \text{ for every } e \in E\}$, denoted as

$$\bigwedge_{e \in E} e = \inf E\,.$$

Hence the triple $(E(X), \vee, \wedge)$ is a complete lattice. The smallest element in this lattice is the set $X \times X$, representing the equivalence in which any two objects are equivalent, the greatest element is the set $\{(x, x) \colon x \in X\}$, which represents the identity relation.

3.4 After an appropriate extension of the set of descriptors of the space (X, F) is made, the dual isomorphism between $((F), \sqsubset)$ and some subset of $(E(X), \subset)$ gives the possibility to find for any two aspects (f), $(g) \in (F)$ an aspect having exactly the compound discerning capability that units the discerning capability of (f) and (g), and the aspect having exactly the discerning capability shared by the discerning capability of (f) and (g).

The injection mapping φ from the paragraph 3.2 may be considered as dual isomorphic mapping of (F) into the lattice $(E(X), \vee, \wedge)$. Therefore there exists a isomorphic embading of (F) into a lattice $\mathscr{D}(E(X))$, which is dual to $(E(X), \vee, \wedge)$. As a required extension of (F) we can consider the whole lattice $\mathscr{D}(E(X))$ or some sublattice of $\mathscr{D}(E(X))$, which contains (F). In the first case the gratest element of the extension is the aspect discerning all different objects (a key), the smallest element is the so called trivial aspect (see 5.1), which is generated by constant descriptors.

The minimal possible extension of (F) may be obtained by the following construction. Let $\mathfrak{M}_{(F)}$ be a set of all sublattices of $\mathscr{D}(E(X))$, which contain $\varphi((F))$. Then

$$\bigcap_{L \in \mathfrak{M}_{(F)}} L$$

is the sublattice of $\mathscr{D}(E(X))$ with the required property. This lattice will be called the lattice generated by (F). The greatest and smallest elements of the lattice generated by (F) may differ from the aspect discerning all objects and the trivial aspect, respectively.

Hence the following definition may be given:

If (X, F) is a space, (F) is the set of all its aspects and (f), $(g) \in (F)$, then there exists a space (X, F^*) with the set of all its aspects $(F^*) \supset (F)$ such that:

(1) there is a unique aspect $(h) \in (F^*)$ such that:

(1.1) $(h) \sqsupset (f)$, $(h) \sqsupset (g)$;

(1.2) if $(h') \in (F^*)$, $(h') \sqsupset (f)$, $(h') \sqsupset (g)$ then $(h') \sqsupset (h)$;

(2) there is a unique aspect $(k) \in (F^*)$ such that:

(2.1) $(k) \sqsubset (f)$, $(k) \sqsubset (g)$;

(2.2) if $(k') \in (F^*)$, $(k') \sqsubset (f)$, $(k') \sqsubset (g)$ then $(k') \sqsubset (k)$.

The aspect (h) described in the point (1) of the above statement is called a compound aspect of the aspect (f) and (g) and denoted by $(h) = (f) \sqcup (g)$.

The aspect (k) described in the point (1) of the above statement is called a shared aspect of the aspects (f) and (g) and denoted by $(k) = (f) \sqcap (g)$.

3.5 The definition of compound and shared aspects can be extended from two aspects to an arbitrary set od aspects. Completeness of the lattice $(E(X), \vee, \wedge)$ and the dual isomorphism from the paragraphs 3.2 and 3.4 provide a possibility of embedding any information space into such a space in which the set of all its aspects is closed with respect to forming the compound and the shared aspects of an arbitrary set of aspects. For such a space (X, F^*) the triple $((F), \cup, \cap)$ is a complete lattice which is dual isomorphic to some complete sublattice of $(E(X), \vee, \wedge)$.

Since according to the above mentioned dual isomorphism the compound aspect corresponds to mere intersection of sets, its effective construction from descriptors of given aspects is easy. However, the construction of the shared aspect, corresponding to the transitive cloasure of the union of equivalences may be more complicated.

4. COMPLETENESS AND MINIMAL COMPLETENESS OF THE SET OF ASPECTS

4.1 An important question is whether objects are fully characterized by a given set of aspects.

We call a subset $(G) \subset (F)$ of the set of all aspects of the space (X, F) to be complete with respect to X iff for any different objects $x \neq y \in X$ there exists an aspect $(f) \in (G)$ such that $f(x) \neq f(y)$.

The set $\{(f)\}$, which contains only a single aspect (f) is obviously complete with respect to X iff the descriptior $f \in (f)$ is a key in X, that is iff it discerns arbitrary two diferent objects.

We call a subset $(G) \subset (F)$ of the set of all aspects of the space (X, F) to be minimal complete with respect to X iff it is complete with respect to X and leaving out any of its elements result in a set which is no longer complete with respect to X.

Hence if $(G) \subset (H) \subset (F)$ and if (G) is complete with respect to X, then (H) is also complete with respect to X. In case of finite set (F) of aspects of information space (X, F) a minimal complete set of aspects may always be selected from any complete set of aspects using the finite process. Nevertheless, the number of elements of such a minimal complete set with respect to X is not determined uniquely. The following example presents the space with complete (infinite) set of aspects from which no minimal complete subset may be selected.

Example 1. Let X be the set of all infinite sequences $x = \{x_j\}$, $j = 0, 1, \ldots$, such that for each $j = 0, 1, \ldots$, either $x_j = 0$ or $x_j = 1$.
Let

$$f_i(x_j) = \sum_{j=0}^{i} x_j 2^{-j}, \quad (f)_i = (f_i), \quad \text{for} \quad i = 0, 1, \ldots .$$

Then $\{(f)_0, (f)_1, \ldots\}$ is the set of aspects of the space $(X, \{(f)_0, (f)_1, \ldots\})$, which is complete with respect to X and contains no minimal complete subset.

4.2 For spaces closed with respect to the operation of forming the compound and shared aspects for arbitrary set of aspects, complete and minimal complete sets of aspects may be characterized as follows:

If $(G) \subset (F)$ and (G) is complete with respect to X then $\sup (G)$ is the finest aspect in the ordered set $((F), \sqsubset)$. The reverse statement to the latter is valid only when completeness with respect to X of the set (F) of all aspects of the space (X, F) is assumed. This condition, however, need not be satisfied in the common case.

A complete set of aspects is minimal complete with respect to X iff no aspect can be found in this set, which is coarser than the compound aspect of all other aspects of this set.

5. INDEPENDENCE AND ORTHOGONALITY OF ASPECTS

5.1 In this section we shall assume that the space (X, F) is closed with respect to operation of forming the compound and the shared aspects for arbitrary set of aspects.

We shall call trivial aspect an aspect which is generated by a constant descriptor, i.e. which does not convey any information. If (F) contains a trivial aspect, this aspect is of course the smallest element of the ordered set $(F), \sqsubset)$ and therefore for the corresponding equivalence the following holds $\sim_{(f)} = X \times X$.

It is natural to define independence of two aspects as follows: Aspects $(f), (g) \in (F)$ are called to be independent iff their shared aspect $(f) \sqcap (g)$ is trivial.

Example 2. Let $\mathbb{N}$ be set of all natural numbers,

$$X = \{(a, b) \subset \mathbb{N} \times \mathbb{N} : |a - b| \leqq 1\}, \quad f(a, b) = a, \quad g(a, b) = b.$$

Then the aspects (f) and (g) are independent although their values may differ at most by one.

5.2 Example 2 shows that the concept of independence is not in a full accordance with our intuitive idea about complete independence of conveyed information. Hence, we shall introduce a stronger concept of orthogonality, based on the requirement that no conclusion can be drawn about the value of descriptors of other aspects from the knowledge of values of one of the aspects.

Two aspects $(f), (g) \in (F)$ of the space (x, F) are called to be orthogonal iff any two different objects $x \neq y$ there exists an object $z \in X$ such that $f(z) = f(x)$ and $g(z) = g(y)$ for $f \in (f)$ and $g \in (g)$.

Example 3. Let $\mathbb{C}^+$ be a set of all complex numbers with positive real and imaginary parts.

$$\mathbb{C}^+ = \{x + iy : x, y \in \mathbb{R}, x > 0, y > 0\}.$$

Let us denote by

$R(x + iy) = x$ the real part of the number $x + iy$,

$I(x + iy) = y$ the imaginary part of the number $x + iy$ and

$A(x + iy) = x + iy = \sqrt{(x^2 + y^2)}$ the absolute value of $x + iy$.

Then aspects (R) and (I) are evidently orthogonal, but the aspects (R), (A) are independent but not orthogonal, because for $12 + 5i$ and $3 + 4i$ there is no complex number z such that $R(z) = R(12 + 5i) = 12$ and $A(z) = A(2 + 4i) = 5$ (the dependence $A(z) \geqq R(z)$ must take place).

5.3 The definition of orthogonality may be generalized in a natural way for arbitrary both finite and infinite sets of aspects.

The subset $(G) \subset (F)$ of the set of all aspect of the space (X, F) is called to be orthogonal iff for any injective mapping ψ of the set (G) into set X there exists an object $z_\psi \in X$ such that

$$f(z_\psi) = f(\psi((f))) \quad \text{for every} \quad (f) \in (G), \quad f \in (f).$$

A subset of an orthogonal set of aspect is always an orthogonal set. The folowing example shows the space in which from the orthogonality of all pairs $\{(f), (g)\}$, $\{(f), (h)\}$, $\{(g), (h)\}$ does not follow the orthogonality of the set $\{(f), (g), (h)\}$.

Example 4. Let (X, F) be an arbitrary space, (f) and (g) two orthogonal aspects in (F) and $x_0 \in X$ an arbitrary object. Let us define the descriptor h as follows:

If for $f \in (f)$ and $g \in (g)$ there is $((f(x) = f(x_0)$ **and** $g(x) = g(x_0))$
or $(f(x) \neq f(x_0)$ **and** $g(x) \neq g(x_0)))$
then $h(x) = 0$
else $h(x) = 1$.

The aspect (h) is orthogobal with respect to (f) and also with respect to (g) but the set $\{(f), (g), (h)\}$ is not orthogonal.

5.4 The following theorem postulates an important property of complete and orthogonal sets.

If the set $(G) \subset (F)$ of aspects of the space (X, F) is complete with respect to X, orthogonal and does not contain a trivial aspect, then (G) is minimal complete with respect to X.

Proof. Let (g) be an arbitrary aspect from (G). As (g) is not trivial, there exist $x, y \in X$ such that $g(x) \neq g(y)$ for $g \in (g)$. From the orthogonality of $(G) \dot{-} \{(g)\} = (G')$ follows the existence of $x_1, y_1 \in X$ such that $g(x_1) = g(x)$, $g(y_1) = g(y)$ and $f(x_1) = f(y_1)$ for all $(f) \in (G')$, $f \in (f)$. Therefore if (G') is complete with respect to X, then must be $x_1 = y_1$ which is the contradiction to $g(x_1) \neq g(y_1)$.

5.5 As the concepts alternative to the orthogonality we can have weak orthogonality and finite orthogonality. The former is based on the requirement of admissibility of local changes in values of descriptors, the latter on the orthogonality of any finite subset of the given set of aspects.

A set $(G) \subset (F)$ of aspects from the space (X, F) is called to be weak orthogonal iff for any two objects x. $y \in X$ and any aspect $(f) \in (G)$ there exists an object $z \in X$ such that:

(1) $\quad f(z) = f(x) \quad \text{for} \quad f \in (f);$

(2) $\quad g(z) = g(y) \quad \text{for} \quad (g) \in (G), \quad (g) \neq (f), \quad g \in (g).$

A set $(G) \subset (F)$ of aspects from the space (X, F) is called to be finite orthogonal if every finite subset of (G) is orthogonal.

5.6 The following relations hold between the various concepts of orthogonality introduced.

Any orthogonal set of aspects of the space (X, F) is weak orthogonal.

Any weak orthogonal set of aspects of the space (X, F) is finite orthogonal.

Proof. Let $(G) \subset (F)$ be a orthogonal set of aspects, $x, y \in X$ arbitrary objects and $(f) \in (G)$ an arbitrary fixed aspect in (G). Let us define in injection mapping φ of (G) into X by $\varphi((f)) = x$ and $\varphi((g)) = y$ for all $(g) \in (G) \dot{-} \{(f)\}$. The existence of $z \in X$ such that $f(z) = f(x)$ and $g(z) = g(y)$ for $(g) \in (G) - \{(f)\}$, $g \in (g)$ follows from the orthogonality of (G).

The second statement we shall prove by a contradiction. Let (G) be a weak orthogonal set of aspects which is not finite orthogonal. Let P be the set of all natural numbers k such that there exists a subset of (G) with k elements which is not orthogonal. Let n be a minimal element of P; $n > 0$. Let $\{(f_1), (f_2), \ldots, (f_n)\}$ be such a nonorthogonal set of aspects and $x_1, x_2, \ldots, x_n$ arbitrary objects in X. Since $n = \min P$ the set $\{(f_1), (f_2), \ldots, (f_{n-1})\}$ must be orthogonal and therefore there exists some $y \in X$ such that $f_j(y) = f_j(x_j)$ for $f_j \in (f_j)$, $j = 1, 2, \ldots, n - 1$. Since (G) is weak orthogonal, there exists an object $z \in X$ such that $f_n(x_n) = f_n(z)$ for $f_n \in (f_n)$ and $g(z) = g(y)$ for all $(g) \in (G) - \{(f_n)\}$, $g \in (g)$ for $j = 1, 2, \ldots, n - 1$. Therefore $f_j(z) = f_j(y) = f_j(x_j)$ for $f_j \in (f_j)$, $j = 1, 2, \ldots, n - 1$, which proves orthogonality of $\{(f_1), (f_2), \ldots, (f_n)\}$.

All three mentioned concepts of orthogonality coincide in the space with finite set of aspects. The following two examples demonstrate the existence of a (infinite) space with a weak orthogonal but not orthogonal set of aspects and the existence of a (infinite) space with a finite orthogonal but not weak orthogonal set of aspects.

Example 5. Let X be a set of all infinite sequences $\{x_j\}$, $j = 1, 2, \ldots$, or real numbers having only a finite number of non-zero elements. Let $f_i(\{x_j\}) = x_i$, for $i = 1, 2, \ldots$ and $F = \{f_1, f_2, \ldots\}$. Then the set $\{(f_1)(f_2), \ldots\}$ is weak orthogonal but not orthogonal.

Example 6. Let X be a set of all convergent infinite sequences $\{x_j\}$, $j = 1, 2, \ldots$, of real numbers. Let

$$f_i(x_j) = x_i \quad \text{for} \quad i = 1, 2, \ldots, g(\{x_j\}) = \lim_{j \to \infty} x_j ,$$

$F = \{g, f_1, f_2, \ldots\}$. Then the set F is finite orthogonal but not weak orthogonal.

5.7 The statement analogical to Theorem 5.4 (concerning minimal completness) is valid. For weak orthogonal sets of aspects:

If the set $(G) \subset (F)$ of aspects of the space (X, F) is complete with respect to X, weak orthogonal and does not contain a trivial aspect, then (G) is minimal complete with respect to X.

The proof is completely analogous to that in 5.4.

Unfortunately for finite orthogonal sets the analogous statement does not hold. The following example demonstrates the existence of the information space with finite orthogonal set of aspects complete with respect to the set of all objects, which does not contain a trivial aspect and does not contain even any minimal complete subset.

Example 7. Let A be an arbitrary infinite countable set and V the set formed by the empty set and all finite sequences of elements from A. Let us define the relation R on V by:

$$R_0 = \{(\emptyset, (a)) \colon a \in A\} ,$$

$$R_j = ((a_1, a_2, \ldots, a_j), (a_1, a_2, \ldots, a_j, a_{j+1})) \colon a_1, a_2, \ldots, a_{j+1} \in A\} ,$$

$$\text{for} \quad j = 1, 2, \ldots.$$

$$R = \bigcup_{j=1}^{\infty} R_j .$$

Therefore (V, R) is an directed tree in which each vertex $v \in V$ has countable many successors. Let X be the set of all mappings φ of the set V into the set $\{0, 1\}$ satisfying the following condition: $\varphi(u) = 1$ for $u \in V$ iff there exists an infinit number of $v \in V$ such that $(u, v) \in R$ and $\varphi(v) = 1$. Therefore as objects we consider such evaluations of the set of verteces of the tree (V, R) in which the vertex has value 1 iff there is infinite number of their successors having also the value 1. It is obvious that X is a non empty set.

For each vertex $v \in V$ let us define the mapping f_v by $f_v(\varphi) = \varphi(v)$ for all $\varphi \in X$ and let $F = \{f_v \colon v \in V\}$. Descriptors are therefore the values of our evaluations on individual verteces of the tree (V, R). If (F) is the set of all information aspects in the space (X, F), the set (F) is complete and finite orthogonal but need not contain any minimal complete subset, because the value $f(\varphi)$ for $\varphi \in X$ is the value of evaluation φ in some vertex $v \in V$, and therefore is fully determined by the values of this evaluation on the set of all successors of the vertex v.

6. THE BASIS AND THE SUPPORT

6.1 Using the concepts introduced, some kind of analogy between the coordinate system in linear spaces and respective phenomena in information spaces may be examined, which means generating all aspects in the information space by using some of the selected orthogonal aspects.

Let (X, F) be an information space, $(G) \subset (F)$ a set of its aspects and $(f) \in (F)$ an aspect. The set (G) is said to generate the aspect $(f) \in (F)$, written as $(f) \models (G)$ iff for any objects $x, y \in X$ the assertion ($g(x) = g(y)$ for every $(g) \in (G)$ and $g \in (g)$) implies $f(x) = f(y)$ for $f \in (f)$.

6.2 In the terminology of ordering of aspects and compound aspects the condition of generating the aspect (f) by the set $(G) \subset (F)$ may be expressed in the following form:

The aspect $f \in (F)$ is generated by the set $(G) \subset (F)$ of aspects (i.e. $(f) \models (G)$) iff (f) is coarser than the compound aspect of the set (G) (i.e. $(f) \sqsubset \bigcup_{(g)\in(G)} (g) = \sup (G)$).

Proof. If $(f) \sqsubset \bigcup_{(g)\in(G)} (g)$ and $x \in X$ then exists some $(g_0) \in (G)$ such that $(f) \sqsubset (g_0)$ and $g_0(x) = g_0(y)$ for $g_0 \in (g_0)$ implies $f(x) = f(y)$ for $f \in (f)$ and therefore $(f) \models (G)$.

If the relation $(f) \sqsubset \bigcup_{(g)\in(G)} (g)$ is not valid, then there exist two objects $x, y \in X$ such that $f(x) \neq f(y)$ for $f \in (f)$ and $h(x) = h(y)$ for $h \in \bigcup_{(g)\in(G)} (g)$. Therefore $g(x) = g(y)$ for all $(g) \in (G)$, $g \in (g)$ and $(f) \not\models (G)$.

6.3 The set $(\mathscr{B})$ of aspects of the space (X, F) is called an orthogonal basis of (X, F), or an orthogonal basis of (F), (abbreviated as basis), iff it is orthogonal, complete with respect to X and does not contain a trivial aspect.

It follows from foregoing that the basis is alwyas a minimal complete set with respect to X and generates an arbitrary aspect $(f) \in (F)$ of the space (X, F). Vice versa the orthogonal set of aspects, which does not contain a trivial aspect and generates all aspects of the space (X, F) is a basis iff the set (F) of all aspects of the space (X, F) is complete with respect to X. Example 1 of the paragraph 4.1 demonstrates that generally, there need not be any basis in the given (infinite) information space. For finite spaces a basis always exists and can be constructed. But the number of elements of a basis is not determined uniquely.

6.4 For generating aspects by using elements of a given basis it is important to examine which elements of the basis take real part in the generation of a given aspect. Therefore, we shall introduce the concept of support of an aspect with respect to a given basis of the information space.

Let (X, F) be a space with the basis $(\mathscr{B})$ and $(f) \in (F)$ be an aspect. A subset $\mathscr{S}((f), (\mathscr{B}))$ of the basis $(\mathscr{B})$ will be called the support of (f) with respect to basis $(\mathscr{B})$ iff:

(1) $\mathscr{S}((f), (\mathscr{B}))$ generates (f) and
(2) by excluding an arbitrary aspect of the set $\mathscr{S}(f), (\mathscr{B}))$ a set is obtained which longer generates the aspect (f).

The following example demonstrates that, generally, a support with respect to a basis need not exist for a given aspect, even if a basis of an information space exists. However, from the next paragraph it follows that if such a support exists, it is determined uniquely.

Example 8. Let X be an arbitrary infinite set of objects, $\{g_j\}$, $j = 1, 2, \ldots$, an infinite sequence of real functions on X, which generate an orthogonal and complete set of information aspects. Let us define the real function f on X as follows: if the sequence $g_j(x)$ $j = 1, 2, \ldots,$ is convergent then $f(x) = \lim_{j\to\infty} g_j(x)$ else $f(x) = 0$. If $F = \{f, g_1, g_2, \ldots\}$, $(\mathscr{B}) = \{(g_1), (g_2) \ldots\}$ then $(\mathscr{B})$ is the basis of (X, F) and for any finite subset $(\mathscr{B}_0) \subset (\mathscr{B})$ there is $(f) \models (\mathscr{B}) - (\mathscr{B}_0)$. Therefore (f) must not have a support with respect to $(\mathscr{B})$.

6.5 An important task is to be able to decide whether a given element of the basis is an element of the support of a given aspect. For thus purpose we shall prove the following assertion:

If an aspect (f) has with respect to the basis $(\mathscr{B})$ of the space (X, F) the support $\mathscr{S}((f), (\mathscr{B}))$, than all elements $(g) \in (\mathscr{B})$ such that the set $\{(f), (g)\}$ is not orthogonal are contained in the support $\mathscr{S}((f), (\mathscr{B}))$.

Proof. If $(g) \in \mathscr{S}((f), (\mathscr{B}))$ then from the orthogonality of the basis $(\mathscr{B})$ we obtain the existence of an object $z \in X$ such that $h(z) = h(x)$ and $g(z) = g(x)$ for all $(h) \in \mathscr{S}((f), (\mathscr{B}))$, $h \in (h)$ and $g \in (g)$. Since $\mathscr{S}((f), (\mathscr{B}))$ is the support of (f), it holds that $f(z) = f(x)$ and therefore $\{(g), (f)\}$ is orthogonal.

The reverse statement does not hold as follows from the following example.

Example 9. Let f and g be real functions on the set of objects X, such that the set of aspects $\{(f), (g)\}$ is a basis of the space (X, F). Let $h(x) = f(x) + g(x)$ for $x \in X$. Then $\mathscr{S}((h), (\mathscr{B})) = \{(f), (g)\} = (\mathscr{B})$, but the sets $\{(h), (f)\}$ and $\{(h), (g)\}$ are both orthogonal.

However, this statement can be used for the following complete characterisation of a support.

6.6 Let (X, F) be a space with basis $(\mathscr{B})$ and let the aspect $(f) \in (F)$ have a support $\mathscr{S}((f), (\mathscr{B}))$ with respect to the basis $(\mathscr{B})$. Then:

(1) $\mathscr{S}((f), (\mathscr{B})) = \{(g) \in (\mathscr{B}) : (f) \not\models (\mathscr{B}) \dot{-} \{(g)\}\}$;

(2) $\mathscr{S}((f), (\mathscr{B}))$ is the minimal subset of the basis $(\mathscr{B})$ generating the aspect (f), i.e. the smallest element of the set $\{(G) \subset (\mathscr{B}) : (f) \models (G)\}$ in the ordered set $(\exp(\mathscr{B}), \subset)$, where $\exp(\mathscr{B})$ is the set of all subsets of the basis $(\mathscr{B})$.

Proof. The equivalence of statements (1) and (2) is obvious; it is therefore sufficient to verify the equation (1).

Let (f) be an aspect such that $(f) \not\models (\mathscr{B}) \dot{-} \{(g)\}$ for some $(g) \in (\mathscr{B})$. Then $x, y \in X$ exist such that $f(x) \neq f(y)$ for $f \in (f)$, but $h(x) = h(y)$ for all $(h) \in (\mathscr{B}) \dot{-} \{(g)\}$, $h \in (h)$. If $(g) \in \mathscr{S}((f), (\mathscr{B}))$ then $h(x) = h(y)$ for all $(h) \in \mathscr{S}((f), (\mathscr{B}))$, $h \in (h)$ and therefore $f(y) = f(y)$ for $f \in (f)$, which is the contradiction.

If $(g) \in \mathscr{S}((f), (\mathscr{B}))$, then from the minimality of the support follows the existence of $x, y \in X$ such that $f(x) \neq f(y)$, $g(x) \neq g(y)$ and $h(x) = h(y)$ for $f \in (f)$, $g \in (g)$ and all $(h) \in \mathscr{S}((f), (\mathscr{B})) - \{(g)\}$, $h \in (h)$. Orthogonality of the basis $(\mathscr{B})$ implies the existence of $x', y' \in X$ such that $g(x') = g(x)$, $g(y') = g(y)$, $h(x') = h(x)$ and $h(y') = h(y)$ for all $(h) \in \mathscr{S}((f), (\mathscr{B})) - \{(g)\}$, $h \in (h)$ and $k(x') = k(y')$ for all $(k) \in (\mathscr{B}) \dot{-} \mathscr{S}((f), (\mathscr{B}))$, $k \in (k)$. The values of all descriptors from aspects from

$(\mathscr{B}) \dot{-} \{(g)\}$ coincides on x' and y', therefore $f(x') = f(y')$ for $f \in (f)$ and the values of all descriptors from aspects from the support $\mathscr{S}((f), (\mathscr{B}))$ coincides both on x and x' and y and y' therefore $f(x') = f(x)$ and $f(y') = f(y)$, which is the contradiction with $f(x) \neq f(y)$ for $f \in (f)$.

6.7 Orthogonality of two aspects can be verified by means of their supports as follows:

Let (X, F) be a space with basis $(\mathscr{B})$, (f), $(g) \in (F)$ two aspects which have with respect to $(\mathscr{B})$ the supports $\mathscr{S}((f), (\mathscr{B}))$ and $\mathscr{S}((g), (\mathscr{B}))$, respectively. Then if $\mathscr{S}((f), (\mathscr{B})) \cap \mathscr{S}((g), (\mathscr{B})) = \emptyset$ the set $\{(f), (g)\}$ is orthogonal.

Proof. Let $\mathscr{S}((f), (\mathscr{B}))$ and $\mathscr{S}((g), (\mathscr{B}))$ are disjunct. Then for any $x, y \in X$ there exists a $z \in X$ such that $h(z) = h(x)$ for all $(h) \in \mathscr{S}((f), (\mathscr{B}))$, $h \in (h)$ and $k(z) = k(x)$ for all $(k) \in \mathscr{S}((g), (\mathscr{B}))$, because the basis is an orthogonal set. Therefore $f(z) = f(x)$ and $g(z) = g(x)$ for $f \in (f)$ and $g \in (g)$ and $\{(f), (g)\}$ is an orthogonal set.

From the next example follows that the reverse statement is not valid.

Example 10. Let (X, F) be a space such that real functions h and k on X generate a basis $\{(h), (k)\} = (\mathscr{B})$ of the space (X, F). Let $f(x) = h(x) + k(x)$ and $g(x) = h(x) - k(x)$ for $x \in X$. Then $\{(f), (g)\}$ is orthogonal and $\mathscr{S}((f), (\mathscr{B})) = \mathscr{S}((g), (\mathscr{B})) = (\mathscr{B})$.

6.8 Using the notion of support even the compound and shared aspects of the given aspect can be characterized. The following assertion holds.

Let (X, F) be a space with a basis $(\mathscr{B})$ and $(f), (g) \in (F)$ two aspects having with respect to the basis $(\mathscr{B})$ the supports $\mathscr{S}((f), (\mathscr{B}))$ and $\mathscr{S}((g), (\mathscr{B}))$, respectively. Then if there exist a compound aspect $(f) \uplus (g)$ and the shared aspect $(f) \Cap (g)$ in (F) the following holds:

(1) $\mathscr{S}((f), (\mathscr{B})) \cup \mathscr{S}((g), (\mathscr{B})) = \mathscr{S}((f) \uplus (g), (\mathscr{B}))$;

(2) $\mathscr{S}((f), (\mathscr{B})) \cap \mathscr{S}((g), (\mathscr{B})) \subset \mathscr{S}((f) \Cap (g), (\mathscr{B}))$.

Proof. The proof of the assertion (1) is easy and may be omitted. To prove (2) let us select for each $(h) \in (\mathscr{B})$ some fixed descriptor $\varrho((h)) = h^0 \in (h)$ as a representation of the aspect (h) and let us define the mapping f^* of the set X of all objects into the set $\mathop{\times}\limits_{(f) \in \mathscr{S}((f),(\mathscr{B}))} R_{h^0}$, where $R_{h^0} = h^0(X)$ is the range of the descriptor $h^0 = \varrho((h))$, by the equation $f^*(x) = \{((h), h^0(x)) : (h) \in \mathscr{S}((f), (\mathscr{B}))\}$. For each $x \in X$ the value $f^*(x)$ is obviously a mapping of the support $\mathscr{S}((f), (\mathscr{B})$ into a set

$$\bigcup_{(h) \in \mathscr{S}((f),(\mathscr{B}))} R_{h^0} .$$

Similarly let $g(x) = \{((h), h^0(x)) \colon (h) \in \mathscr{S}((g), (\mathscr{B}))\}$ for $x \in X$ and $s^*(h) = \{((h), s^0(h)) \colon (h) \in \mathscr{S}((f), (\mathscr{B})) \cap \mathscr{S}((g), (\mathscr{B}))\}$ for $x \in X$.

It is easy to show that in the information space $(X, F \cup \{f^*\} \cup \{g^*\} \cup \{s^*\})$ the following conditions hold:

$$(f) \sqsubset (f^*), (g) \sqsubset (g^*), (s^*) = (f^*) \sqcap (g^*) ;$$
$$(s) \models \mathscr{S}((f), (\mathscr{B})) \cap \mathscr{S}((g), (\mathscr{B})) .$$

From this conditions follows: $(f) \sqcap (g) \sqsubset (f) \sqsubset (f^*)$ and $(f) \sqcap (g) \sqsubset (g) \sqsubset (g^*)$ and therefore $(f) \sqcap (g) \sqsubset (f^*) \sqcap (g) \models \mathscr{S}((f), (\mathscr{B})) \cap \mathscr{S}((g), (\mathscr{B}))$.

Example 10 shows that, generally, the set inclusion in the relation (2) cannot be replaced by equation, because not only independent, but also orthogonal aspects may have supports with nonempty intersection.

(Received April 27, 1987.)

REFERENCES

[1] ANSI/X3/SPARC Study group on Data Base Management Systems — Interim Report. FDT *7* (1975), 2.

[2] C. Beeri and M. Y. Vardi: A proof procedure for data dependencies. J. Assoc. Comput. Mach. *31* (1984), 4, 718—741.

[3] J. M. Cadiou: On semantic issues in the relational model of data. In: Mathematical Foundations of Computer Science (Lecture Notes in Computer Science 45). Springer-Verlag, Berlin—Heidelberg—New York 1976.

[4] E. F. Codd: Further normalization of the data base relational model. In: Proceeding of Courant Computer Science Symposium 6 on Data Base Systems (R. Rustin ed.). Prentice-Hall, Englewood Cliffs, N. J. 1971, 65—98.

[5] E. F. Codd: Extending the database relational model to capture more meaning. ACM Trans. Database Systems *4* (1979), 397—434.

[6] A. Crammers and T. N. Hibbard: On the relationship between a procedure and its data. In: Mathematical Foundations of Computer Science (Lecture Notes in Computer Science 45). Springer-Verlag, Berlin—Heidelberg—New York 1976.

[7] A. Crammers and T. N. Hibbard: Orthogonality of information structures. Acta Inform. *9* (1978), 243—261.

[8] A. Crammers and T. N. Hibbard: The semantic definition of programming languages in terms of their data spaces. In: Programmiersprachen Informatik-Fachberichte 1. Springer-Verlag, Berlin—Heidelberg—New York 1976.

[9] M. Duží and P. Materna: Attributes: Distinguishing capability versus informational capability. Computers and Artificial Intelligence (submitted).

[10] M. Duží and P. Materna: Four concepts of the distinguishing force of an attribute. Computers and Artificial Intelligence (submitted).

[11] J. Vaníček: On the relation between the information and data. In: Proc. of European Congress on Simulation, Vol. B. Academia, Prague 1987, 232—238.

RNDr. Jiří Vaníček, CSc., Výzkumný ústav matematických strojů, k.ú.o. (Research Institute of Mathematical Machines, C.P.O.), Loretánské nám. 3, 118 55 Praha 1. Czechoslovakia.