BULLETIN OF THE POLISH
10. ACADEMY OF SCIENCES
TECHNICAL SCIENCES
Vol. 32, No. 3-4, 1984

VARIA
COMPUTER SCIENCES

On Superfluous Attributes in Knowledge Representation System

by

Zdzisław PAWLAK

Presented by Z. PAWLAK on February 6, 1984

Summary. In expert systems the objects are described by means of attributes. We consider in this paper how to reduce the set of attributes without loss of information about objects.

1. Introduction. In expert systems we describe objects by means of their properties expressed by attributes. The question arises whether all attributes available in the system are necessary to describe any subset of objects. The problem is considered here on the basis of rough set approach (see [1]).

2. Knowledge representation system. By knowledge representation system (see [2]) we mean a system

$$S = (U, A, V, \varrho)$$

where: U—is a set of **objects**, A—is a set of **attributes**, $V = \bigcup_{a \in A} V_a$—is a set of **values** of attributes, $\varrho: UxA \to V$—is an **information function**.

Set $V_a, a \in A$ will be referred to as **domain** of the attribute a.

Function $\varrho_x: A\ V$ such that $\varrho_x(a) = \varrho(x, a)$ for every $a \in A$, $x \in U$ will be called **information** about x in S.

3. Indiscernibility relation. Let B be a nonempty subset of attributes A. We say that objects $x, y \in U$ are B – **indiscernible** in S, $x \underset{B}{\sim} y$, if

$$\varrho_x(a) = \varrho_y(a) \text{ for every } a \in B.$$

Obviously B is an equivalence relation for any $B \subseteq A$.

Equivalence classes of relation B are called B – **elementary** sets in S. A – elementary sets in S are called simply **elementary** sets in S.

B – elementary set containing object $x \in U$ will be denoted by $[x]_B^S$, or $[x]_{\tilde{B}}$ when S is understood.

Subset $X \subseteq U$ will be called a B – **definable** set in S if X is union of some B – elementary sets in S; an empty set is B – definable for every $B \subseteq A$.

4. Approximation of sets in knowledge representation system. Let $S = (U, A, V, \varrho)$ be a knowledge representation system, let $X \subseteq U$ and let $B \subseteq A$ $(B \neq \emptyset)$.

A **lower** B–**approximation** of X in S ($\underline{B}_S(X)$ or $\underline{B}(X)$ when S is understood) we define as follows:

$$\underline{B}(X) = \{x \in U : [x]_{\tilde{B}} \subset X\}$$

An **upper** B – **approximation** of X in S ($\bar{B}_S(X)$ or $\bar{B}(X)$ when S is understood) we mean set

$$\bar{B}(X) = \{x \in U : [x]_{\tilde{B}} \cap X \neq \emptyset\}$$

5. Attributes superfluous in S. Let $S = (U, A, V, \varrho)$ be a knowledge representation system and let $B \subset A$ $(B \neq \emptyset)$ be a subset of attributes. We say that B is **superfluous** in S if $\widetilde{A-B} = \tilde{A}$.

If $\widetilde{A-B} \supset \tilde{A}$ we say that set B is **indispensable** in S. The following is valid:

—if $B, C \subset A$ are superfluous in S then $B \cup C$ may be not superfluous in S,

—if $B, C \subset A$ are indispensable in S then $B \cup C$ is also indispensable in S.

From the above properties it follows that if $a_1, \ldots, a_n \in A$ are superfluous attributes in S, then set $\{a_1, \ldots, a_n\}$ may be not superfluous in S; if $a_1, \ldots, a_n \in A$ are indispensable in S, then set $\{a_1, \ldots, a_n\}$ is also indispensable in S.

Let $B, C \subset A$ be two subsets of attributes in A. We say that set C is **dependent** of set B in S $(B \underset{S}{\rightarrow} C)$ if $\tilde{B} \subset \tilde{C}$.

One can show that if $B \rightarrow C$ then C is superfluous in S.

The set of all indispensable attributes in S will be called **core** of A in S and will be denoted A^*. Let us notice that there exists at most exactly one nonempty core for each A.

The least subset $B \subset A$ such that $\tilde{B} = \tilde{A}$ will be called **reduct** of A in S. Of course any set of attributes A may have more then one reduct.

Obviously if B reduct of A in S, then $A-B$ is superfluous in S.

The following property is true for every A

$$A^* = \bigcap B,$$

where $\bigcap B$ is intersection of all reducts of A in S.

Moreover we have the following property: for every $X \subseteq U$ and $B \subset A$, $\underline{A}(X) = \underline{A-B}(X)$ and $\bar{A}(X) = \overline{A-B}(X)$ if B is superfluous in S.

6. Attributes superfluous with respect to X in S. Sometimes we might be interested in checking whether some attributes are superfluous not for a whole system S, but for a certain subset $X \subset U$ of objects.

We say that subset $B \subset A$ of attributes is **superfluous with respect to** X in S if

$$\underline{A-B}(X) = \underline{A}(X) \text{ and } \overline{A-B}(X) = \bar{A}(X).$$

Evidently if B is superfluous with respect to X in S, then $\widetilde{A-B} = \tilde{A}$. Thus subset $B \subset A$ may be indispensable in S but superfluous with respect to X in S.

Let $X \subset U$ be a set of objects and let $B \subset A$ be a set of attributes. We define set **decidable** by set of attributes B with respect to X in S, in the following way (see [4]).

$$X_B = Fr_{A-B}(X) - Fr_A(X),$$

where $Fr_B(X) = \bar{B}(X) - \underline{B}(X)$.

The following conclusion is obvious from the definition: subset $B \subset A$ of attributes is superfluous with respect to X in S if $X_B = \emptyset$.

INSTITUTE OF COMPUTER SCIENCE, POLISH ACADEMY OF SCIENCES, P.O. BOX 22, 00-901 WARSAW, PKiN (INSTYTUT PODSTAW INFORMATYKI PAN, WARSZAWA)
DEPARTMENT OF COMPUTER SCIENCE, NORTH CAROLINA UNIVERSITY, CHARLOTTE, N.C. 28223, USA

REFERENCES

[1] Z. Pawlak, *Information Systems—Theoretical Foundations*, Information Systems, **6** (1981) 205–218.

[2] Z. Pawlak, *Rough Sets*, International Journal of Information and Computer Sciences, **11**, (1982) 341–356.

[3] Z. Pawlak, *Rough Classification*, International Journal of Man – Machine Studies (1984) (to appear).

[4] Z. Pawlak, *On discernibility of Objects in Knowledge Representation Systems*, Bull. Polish Academy of Sciences (1984) (to appear).

[5] E. Orłowska, Z. Pawlak, *Foundations of Knowledge Representation*, Springer – Verlag (1984) (submitted).

3. Павляк, **Об избыточных свойствах в системе представлений знаний**

В системах экспертов объекты описываются с помощью отличительных черт. В настоящей статье доказывается, как сократить множество свойств без потери информации об объектах.