BULLETIN OF THE POLISH
ACADEMY OF SCIENCES
TECHNICAL SCIENCES
Vol. 33, No. 9–10, 1985

VARIA

COMPUTER SCIENCE

# Rough Sets and Multiexpert Systems

by

Zdzisław PAWLAK

*Presented by Z. PAWLAK on April 25, 1985*

**Summary.** In this paper the expert systems are considered as a special case of information systems and it is shown that the rough sets approach to multiexpert system analysis yields new and fruitful results.

**1. Introduction.** In this paper multiexpert systems are considered. We assume that the knowledge of every expert is identified with the expert ability to classify objects of a certain universe. We show that this class of expert systems can be regarded as a special case of information systems (see [1]) and can easly analysed in terms of rough set concepts (see [2, 3]).

**2. Multiexpert systems.** A multiexpert system is a 4-tuple

$$S = (U, A, V, f)$$

where

$U$ – is a set called the **universe**

$A = D \cup E\ (D \cap E = \emptyset)$ is the set of **attributes**; $D$ – is the set of **description** attributes and $E$ – **experts** attributes

$V = \bigcup_{a \in A} V_a$ – is the **domain (set of values)** of attribute $a$

$f: U \times A \to V$ is a total function, such that $f(x, a) \in V_a$, for every $x \in U$ and $a \in A$

In what follows, we shall identify experts attributes with experts, i.e. if $a \in E$ then $a$ will be called an expert.

The function $f_a: U \to V_a$ such that $f_a(x) = f(x, a)$ for every $a \in E$ is called the **opinion** of an expert $a$.

If $f(x, a) = f(y, a)$ for every $a \in B$, $B \subseteq A$ we say that $x$ and $y$ are **indiscernible** with respect to $B$ in $S$. Thus, each $B \subseteq A$ generates an **indiscernibility** relation $\tilde{B}$, such that $(x, y) \in \tilde{B}$ if $x$ and $y$ are indiscernible with respect to $B$ in $S$. Obviously, $\tilde{B}$ is an equivalence relation, for every $B$. The family of equivalence classes of $\tilde{B}$ is denoted by $B^*$ and equivalence classes of $\tilde{B}$ are called $B$-**elementary** sets in $S$.

We say that set $B \subseteq A$ is **independent** in $S$, if for every $C \subset B$, $\tilde{C} \supset \tilde{B}$; otherwise $B$ is **dependent** in $S$.

The maximal independent subset of $B$ is called **reduct** of $B$ in $S$. If reduct of $B$ is equal to $B$ we say that $B$ is **reduced** and the corresponding system is said to be $B$-**reduced**. Thus, we can reduce both set $D$ and set $E$. Reducing the set $D$ we simplify the system with regard to the descriptions of objects, and the reduction of set of experts $E$ provides the minimal set of experts with the same decisive power as the whole set of experts $E$.

The classification $a^*$, $a \in E$ is called the **expert** $a$ **decion** and the classification $B^*$, $B \subseteq E$ is called the **decision** of **group** of **experts** $B$.

**3. Approximations.** Let $S = (U, A, V, f)$ be an multiexpert system, $X \subseteq U$ and $B \subseteq A$.

We define two sets

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

called the $B$-**lower** and $B$-**upper approximation** of $X$ in $S$.

The set $Bn_B(X) = \bar{B}X - \underline{B}X$ is called the $B$-**boundary** of $X$ in $S$.

The number

$$\mu_B(X) = \frac{\text{card } \underline{B}X}{\text{card } \bar{B}X}$$

is called the **accuracy** of **approximation** of $X$ by $B$ in $S$.

Let $F = \{X_1, X_2, \ldots, X_m\}$ be a family of subsets of $U$, i.e. $X_i \subseteq U$ for every $i$, $1 \leqslant i \leqslant m$. By $B$-**lower** and $B$-**upper approximation** of $F$ we mean the sets

$$\underline{B}F = \{\underline{B}X_1, \underline{B}X_2, \ldots, \underline{B}X_m\}$$
$$\bar{B}F = \{\bar{B}X_1, \bar{B}X_2, \ldots, \bar{B}X_m\}.$$

The number

$$\beta_B(F) = \frac{\operatorname{card} \bigcup_{i=1}^{n} \underline{B}X_i}{\sum_{i=1}^{n} \operatorname{card} \bar{B}X_i}$$

is called the accuracy of the **approximation** of $F$ by $B$ in $S$ and the number

$$\gamma_B(F) = \frac{\operatorname{card} \bigcup_{i=1}^{n} \underline{B}X_i}{\operatorname{card} U}$$

is called the **quality** of the approximation of $F$ by $B$ in $S$.

**4. Dependency of attributes.** Let $S = (U, A, V, f)$ be a multiexpert system, $k$ — a number such that $0 \leqslant k \leqslant 1$, and $B, C \subseteq A$.

We say that $C$ **depends in a degree** $k$ on $B$ in $S$, in symbols $B \xrightarrow{k} C$, if $k = \gamma_B(C^*)$. If $k = 1$ we write $B \to C$ instead of $B \xrightarrow{1} C$, and we say that $C$ **totally depends** on $B$ in $S$; if $0 < k < 1$ we say that $C$ **roughly depends** on $B$ in $S$ and if $k = 0$ we say that $C$ is **totally independent** on $B$ in $S$.

If $D \to E$ we say that the multiexpert system is **deterministic**; otherwise the multiexpert system is **nondeterministic**.

It means that if the multiexpert system is deterministic we are able to describe experts decisions exactly by the description attributes available in the multiexpert system; if the multiexpert system is nondeterministic we are able to describe experts decisions only with a certain approximation (roughly), using the description attributes available in the system.

The numbers $\beta_D(E^*)$ and $\gamma_D(E^*)$ are called the **accuracy** and the **quality** of the multiexpert system, respectively. In other words, these numbers say how precise the experts decisions can be described in terms of attributes available in the system.

If the set of experts $E$ is dependent in a multiexpert system $S$, and $C$ is a reduct of $E$ in $S$ then $C \to E\text{-}C$, which is to mean that the set of experts $E\text{-}C$ is **superfluos** in $S$.

*Property 1.* If $B \subseteq D$ and $C \subseteq E$ are reducts of $D$ and $E$, respectively, then

$$\gamma_D(E^*) = \gamma_B(C^*)$$
$$\beta_D(E^*) = \beta_B(E^*).$$

Let $B \subseteq D$ and $C \subseteq E$. The set

$$\operatorname{Pos}_B(C^*) = \bigcup_{i=1}^{n} \underline{B}X_i$$

where $C^* = \{X_1, X_2, ..., X_n\}$ will be called the **positive region** of the classification $C^*$ with respect to $B$ ($B$-positive region) in $S$.

The set

$$Bn_B(C^*) = \bigcup_{i=1}^{n} \bar{B}X_i$$

is called the **doubtful region** of the classification $C^*$ with respect to $B$ ($B$-doubtful region) in $S$.

Positive region of the classification is the subset of the universe on which the experts decisions agree, and can be described in terms of attributes considered; the doubtful region of the classification is the subset of the universe on which it is impossible to recognize as to the experts opinion, whether they agree or disagree – eploying the attributes of the system.

Of course,

$$\mathrm{Pos}_B(C^*) \cup Bn_B(C^*) = U$$

for every $B \subseteq D$ and $C \subseteq E$.

**5. Rough equality of experts.** In this section we shall compare sets of experts using the concepts introduced previously.

Let $S = (U, A, V, f)$ be a multiexpert system, $A = D \cup E$, $B \subseteq D$ and $E_1, E_2 \subseteq E$.

We shall employ the following definitions

A1) $E_1$ and $E_2$ are **bottom equal** with respect to $B$ in $S$, in symbols $E_1 \eqsim_B E_2$ if $\underline{B}(E_1^*) = \underline{B}(E_2^*)$

A2) $E_1$ and $E_2$ are **top equal** with respect to $B$ in $S$, in symbols $E_1 \simeq_B E_2$ if $\bar{B}(E_1^*) = \bar{B}(E_2^*)$

A3) $E_1$ and $E_2$ are **roughly equal** with respect to $B$ in $S$, in symbols $E_1 \approx_B E_2$ if $E_1 \eqsim_B E_2$ and $E_1 \simeq_B E_2$.

The bottom equality preserves the positive region of the classifications $E_1^*$ and $E_2^*$; the top equality preserves the upper approximation of each class in both classifications $E_1^*$ and $E_2^*$, and the rough equality preserves the doubtful region of the classification $E_1^*$ and $E_2^*$.

*Property 2.*

B1) If $E_1 \eqsim_B E_2$ then $\gamma_B(E_1^*) = \gamma_B(E_2^*)$

B2) If $E_1 \simeq_B E_2$ then $\beta_B(E_1^*) = \beta_B(E_2^*)$

B3) If $E_1 \approx_B E_2$ then $\gamma_B(E_1^*) = \gamma_B(E_2^*)$ and $\beta_B(E_1^*) = \beta_B(E_2^*)$.

**6. Rough inclusion of experts.** In this section we shall compare sets of experts with respect to their decisive "power". In order to do so, we shall employ the following definitions.

Let $S=(U, A, V, f)$ be a multiexpert system, $A = D \cup E$, $B \subseteq D$ and let $F = \{X_1, X_2, ..., X_m\}$, $G = \{Y_1, Y_2, ..., Y_m\}$, $X_i, Y_i \subseteq U$ be two families of subset of $U$.

We shall use the following notations:

C1) $\underset{\sim}{\sqsubseteq}_B G$ if $\underline{B}X_i \subseteq \underline{B}Y_i$

C2) $F \tilde{\sqsubseteq}_B G$ if $\bar{B}X_i \subseteq \bar{B}Y_i$

C3) $F \sqsubseteq_B G$ if $F \underset{\sim}{\sqsubseteq}_B G$ and $F \tilde{\sqsubseteq}_B G$

for every $i$, $1 \leqslant i \leqslant m$.

It is easy to see that all three relations defined above are the partial ordering relations.

Let $E_1, E_2 \subseteq E$. We say that

D1) The set of experts $E_1$ is **internally better** then the set of experts $E_2$ with respect to $B$ in $S$ if $E_1^* \underset{\sim}{\sqsubseteq}_B E_2^*$

D2) The set of experts $E_1$ is **externally better** then the set of experts $E_2$ with respect to $B$ in $S$ if $E_1^* \tilde{\sqsubseteq}_B E_2^*$

D3) Set of experts $E_1$ is **roughly better** than the set of experts $E_2$ with respect to $B$ in $S$ if $E_1^* \sqsubseteq_B E_2^*$.

There is also possible another comparison of experts

E1) Set of experts $E_1$ is **internally finer** than the set of experts $E_2$ with respect of $B$ in $S$ in symbols $E_1 \gtrsim_B E_2$, if $\gamma_B(E_1^*) > \gamma_B(E_2^*)$

E2) Set of experts $E_1$ is **externally finer** than the set of experts $E_2$ with respect to $B$ in $S$, in symbols $E_1 \gtrsim_B E_2$ if $\beta_B(E_1^*) < \beta_B(E_2^*)$

E3) Set of experts $E_1$ is **roughly finer** than the set of experts $E_2$ with respect to $B$ in $S$, in symbols $E_1 >_B E_2$, if $E_1 \gtrsim_B E_2$ and $E_1 \gtrsim_B E_2$.

*Property 3.*

a) If $E_1 \underset{\sim}{\sqsupseteq}_B E_2$ then $E_1 \gtrsim_B E_2$

b) If $E_1 \tilde{\sqsupseteq}_B E_2$ then $E_1 \gtrsim_B E_2$

c) If $E_1 \sqsupseteq_B E_2$ then $E_1 >_B E_2$.

The converse implication is not true.

INSTITUTE OF COMPUTER SCIENCE, POLISH ACADEMY OF SCIENCES, PKIN, PO BOX 22, 00-901 WARSAW
(INSTYTUT PODSTAW INFORMATYKI PAN)
UNIVERSITY OF NORTH CAROLINA, DEPARTMENT OF COMPUTER SCIENCE, CHARLOTTE, NC 28223 USA

REFERENCES

[1] Z. Pawlak, *Information systems: theoretical foundations*, Inf. Syst., **6** (1981) 205–218.