PRACE CO PAN • CC PAS REPORTS

Wiktor Marek, Zdzisław Pawlak

Mathematical foundations of information storage and retrieval

Part 3

137

1973

WARSZAWA

CENTRUM OBLICZENIOWE POLSKIEJ AKADEMII NAUK
COMPUTATION CENTRE POLISH ACADEMY OF SCIENCES
WARSAW PKiN, P. O. Box 22, POLAND

Wiktor Marek, Zdzisław Pawlak

MATHEMATICAL FOUNDATIONS OF INFORMATION STORAGE AND RETRIEVAL

Part 3

137

Warszawa 1973

Mailing address: Dr. Wiktor Marek
Institute of Mathematics PAS
ul. Śniadeckich 8
00-950 Warszawa
P.O. Box 187

Prof. Dr. Zdzisław Pawlak
Computation Centre PAS
00-901 Warszawa PKiN
P.O. Box 22

P r i n t e d   a s   a   m a n u s c r i p t

Nakład 450 egz. Ark. wyd. 0,3; ark. druk. 0,5,
Papier offset. kl. III, 70 g, 70 x 100. Oddano do
druku w październiku 1973 r. W. D. N. Zam. nr 757/0
R-30

Abstract • Содержание • Streszczenie

We show how to incorporate within our formalism hierarchical aspects of information retrieval.

Математическое описание процесса поиска
и хранения информации. Третья часть

В работе показываем способ, как можно представить иерархические аспекты поиска и переработки информации.

Matematyczne podstawy wyszukiwania i gromadzenia
informacji. Część 3

Pokazujemy, w jaki sposób w naszym formalizmie można ująć hierarchiczne aspekty wyszukiwania i przechowywania informacji.

§ 1. Hierarchization

Definition 1.1. Let R, S be equivalences on X. We say that $R < S$ iff $R \subseteq S$ i.e.

$$(\forall x)(\forall y)(xRy \rightarrow xSy)$$

It is clear that $<$ is partial ordering

Definition 1.2. Let R be an equivalence on A and S an equivalence on A/R. We define a relation $S*R$ on A as follows:

$$xS*Ry \longleftrightarrow x/R\ S\ y/R$$

Lemma 1.1. $R < S*R$

Proof: Assume xRy. Then $x/R = y/R$ and since S is reflexive we get $x/R\ S\ y/R$ i.e. $xS*Ry$.

Lemma 1.2. Assume $S*(T*R)$ is defined. Then $(S*T)*R$ is defined and $S*(T*R) = (S*T)*R$.

Proof: Assume T is defined on A/R and S defined on A/R/T. Then $S*T$ is defined on A/R and so $(S*T)*R$ is defined.

Let $xS*(T*R)y$. Then $x/T*R\ S\ y/T*R$. Having in mind that $x/T*R$ consists of all those y/R which are (with x/R) in relation T we find that

$$x/R\ S*T\ y/R$$

which is desired result.

Lemma 1.3. If $S<R$ then there is T such that

$$R = T * S$$

Proof: We define $x/S \ T \ y/s$ iff xRy.
It is enough to prove that T is equivalence. Clearly T is reflexive and symmetric. If $x/S \ T \ y/S \ T \ z/S$ then xRy and yRz and so xRz i.e. $x/S \ T \ z/S$.
As we remember $A = \bigcup_{i=I} A_i$ was a decomposition. Since decomposition is nothing else but family of equivalence classes of some equivalence, we have relation R_I which determines the decomposition $\{A_i\}_{i \in I}$. We may identify I with A/R_I.

Definition 1.3. Let $S<R_I$, $\mathfrak{I}$ be i.s.r. system, $\mathfrak{I} = \langle X,A,I,U \rangle$
We define $\mathfrak{I}/S$ as follows:

$$\mathfrak{I}/S = \langle X, A/S, I, U/S \rangle$$

where $U/S(a/S) = \bigcup \{U(b) : bSa\}$
Clearly $\mathfrak{I}/S$ determines language $\mathcal{L}(A/S)$. The unique relation T such that $R_I = T * S$ comes into the axioms of corresponding theory as follows:

$$\sum_{b/S \ T \ a/S} c_{b/S} = V$$

Let us now enrich the language $\mathcal{L}(A)$ by constants $c_{a/S}$.
Our theory is also enriched as follows

$$(*) \qquad c_{a/S} = \sum_{bSa} c_b$$

In this way we obtain hierarchical i.s.r. system of rank 2 generated by S as follows: $\mathfrak{I}_S = \langle X, A \cup A/S, I, U \cup U/S \rangle$

(Notice that this is nothing else but $\mathfrak{I} \oplus \mathfrak{I}/S$! ! !)

Lemma 1.4. $\mathfrak{I}_2$ satisfies (*)

Full power of the operation $\oplus$ is seen when we have a sequence of relations:

$$S_1 < S_2 , \dots , < S_n < R_I$$

Definition 1.4. Under above assumptions we define

$$\mathfrak{I}_{S_1,\dots,S_n} = \mathfrak{I} \oplus \left(\bigoplus_{i=1}^{n} \mathfrak{I}/S_i \right)$$

Let $T_1,\dots,T_{n-1}$ be relations such that $S_{i+1} = T_i * S_i$

Lemma 1.5. $\mathfrak{I}_{S_1,\dots,S_n}$ satisfies the following formula

$$c_{a/S_{i+1}} = \sum_{b/S_i \ T_i \ a/S_i} c_{b/S_i}$$

Proof: It is clear that it is enough to give the proof for the case $S_1 < S_2 < R_I$, $S_2 = T * S_1$.

Indeed, for $a \in A$ $c_{a/S_2} = \sum_{b S_2 a} c_b$ and $c_{a/S_1} = \sum_{b S_1 a} c_b$

Since however $S_1 < S_2$ we have, for $a,b \in A$.

$$c_{a/S_1} \leq c_{a/S_2}$$

Using idempotence law we get

$$c_{a/S_2} = \sum_{b S_2 a} c_{b/S_1}$$

But this exactly reduces to the desired equation.

REFERENCES

[1] Marek W., Pawlak Z.: Mathematical Foundation of Information Storage and Retrieval I, CC PAS, Reports, Warsaw 1973

[2] Marek W., Pawlak Z.: Mathematical Foundation of Information Storage and Retrieval II, CC PAS, Reports, Warsaw 1973.

Mathematical Insritute of Polish Academy of Sciences
Computing Center of Polish Academy of Sciences

Received October 25, 1973