

# Q I QUÆSTIONES INFORMATICÆ

Volume 5 • Number 3

December 1987

---

M.E. Orłowska	<b>Common Approach to Some Informational Systems</b>	<b>1</b>
S.P. Byron-Moore	<b>A Program Development Environment for Microcomputers</b>	<b>13</b>
N.C.K. Phillips S.W. Postma	<b>Pointers as a Data Type</b>	<b>21</b>
P.J.S. Bruwer J.J. Groenewald	<b>A Model to Evaluate the Success of Information Centres in Organizations</b>	<b>24</b>
J. Mende	<b>Three Packaging Rules for Information System Design</b>	<b>32</b>
T. D. Crossman	<b>A Comparison of Academic and Practitioner Perceptions of the Changing Role of the Systems Analyst: an Empirical Study</b>	<b>36</b>
P.J.S. Bruwer	<b>Strategic Planning Models for Information Systems</b>	<b>44</b>
S.H. von Solms	<b>Generating Relations Using Formal Grammars</b>	<b>51</b>
A.L. du Plessis C.H. Bornman	<b>The ELSIM Language: an FSM-Based Language for ELSIM SEE</b>	<b>67</b>
	<i>BOOK REVIEW</i>	<b>56</b>
	<i>CONFERENCE ABSTRACTS</i>	<b>57</b>

---

An official publication of the Computer Society of South Africa and of the South African Institute of Computer Scientists

'n Amptelike tydskrif van die Rekenaarvereniging van Suid-Afrika en van die Suid-Afrikaanse Instituut van Rekenaarwetenskaplikes

# QUÆSTIONES INFORMATIÆ

An official publication of the Computer Society of South Africa  
and of the South African Institute of Computer Scientists

'n Amptelike tydskrif van die Rekenaarvereniging van Suid-Afrika  
en van die Suid-Afrikaanse Instituut van Rekenaarwetenskaplikes

## Editor

Professor G. Wiechers  
INFOPLAN  
Private Bag 3002  
Monument Park 0105

## Editorial Advisory Board

Professor D.W. Barron  
Department of Mathematics  
The University  
Southampton SO9 5NH, UK

Professor J.M. Bishop  
Department of Computer Science  
University of the Witwatersrand  
1 Jans Smuts Avenue  
2050 WITS

Professor K. MacGregor  
Department of Computer Science  
University of Cape Town  
Private Bag  
Rondebosch, 7700

Prof H. Messerschmidt  
University of the Orange Free State  
Bloemfontein, 9301

Dr P.C. Pirow

Graduate School of Business Admin.  
University of the Witwatersrand  
P.O. Box 31170, Braamfontein, 2017

Professor S.H. von Solms  
Department of Computer Science  
Rand Afrikaans University  
Auckland Park  
Johannesburg, 2001

Professor M.H. Williams  
Department of Computer Science  
Herriot-Watt University, Edinburgh  
Scotland

## Production

Mr C.S.M. Mueller  
Department of Computer Science  
University of the Witwatersrand  
2050 WITS

## Subscriptions

Annual subscription are as follows:

	SA	US	UK
Individuals	R10	\$7	£5
Institutions	R15	\$14	£10

*Computer Society of South Africa*  
*Box 1714 Halfway House*

Quæstiones Informatiæ is prepared by the Computer Science Department of the University of the Witwatersrand and printed by Printed Matter, for the Computer Society of South Africa and the South African Institute of Computer Scientists.

# COMMON APPROACH TO SOME INFORMATIONAL SYSTEMS

M.E. Orlowska

*Department of Computer Science and Information Systems  
University of South Africa*

## ABSTRACT

In this paper we consider relationships between a complete system [6], an L-system [3], a stochastic system [13] and an N-system [8]. We introduce a notion of the induction of one system by another. We prove the important fact that an N-system is a common generalization of all the systems mentioned above. This fact is clearly demonstrated using geometric characterization of the set of subcontinua of the interval  $\langle 0,1 \rangle$ .

## 1. INTRODUCTION

In [6] W. Marek and Z. Pawlak have described a mathematical model (introduced previously by Pawlak [12], and independently by E. Wong and T.C. Chiang [15]) of an information storage and retrieval system, abbreviated to ISR system (see also W. Lipski and W. Marek [5]). An ISR system is a mathematical model of a relational database with complete information. The theory of ISR systems is based on the formalized language which is intermediate between propositional and predicate calculi. However, the theory corresponds to Codd's theory of relational databases [1].

In [3] and [4] W. Lipski has extended the model of an ISR system to cover the case of some incomplete information (called the L-system for short). In the theory of L-systems only the following cases of information incompleteness are considered:

- We know that a fact  $F$  is true,
- We know that a fact  $F$  is false,
- We don't know whether or not  $F$  is true.

W. Lipski has not considered any measure of the lack of information. All probabilistic interpretations, and also cases — sometimes handled with the use of fuzzy sets — fall beyond the scope of his study.

Another important study of the problems related to incomplete information has been undertaken by Jaegerman [2]. His work differs both in the methods used (algebraic methods rather than the logical ones exploited by Lipski) and in the results obtained. Jaegerman's model reveals the relation of pseudo-Boolean algebras and intuitionistic logic to the problems of incomplete information, while in Lipski's study, the relevant notions turn out to be those of topological Boolean algebras and modal logic  $S_4$ .

In [14] T. Traczyk and W. Marek and in [13] T. Traczyk have shown that Pawlak's model of an ISR system is a particular instance of a much more general case — a stochastic information system. They have shown that numerical Boolean algebras might be used as a convenient tool in dealing with informational systems. In particular, they have presented a natural way of introducing fuzzy informational systems, which only produce a (probability) measure that an object  $x$  happens to have a property  $a$ .

In [8] the present author has presented an attempt at a common approach to all the above-mentioned models, and has shown that it is possible to express the ISR system, the L-system and the stochastic informational system as quadruple  $S = \langle X, A, R_I, V \rangle$  where  $X$  is a set of objects,  $A$  is a nonempty set of descriptors,  $R_I$  is an equivalence relation on  $A$  and  $V$  is a map defined on the product  $A \times X$  with the values in the set of all subcontinua of the interval  $\langle 0,1 \rangle$ .

In this paper a very natural consequence of these considerations is presented. A generalization of all the above-mentioned models is described and analyzed. The generalized

system is called an N-system. An N-system is a relational model of a database with incomplete information.

The process of changing (increasing) the information plays a vital role in our theory. The “dynamic” approach allows our study to be considered not only from an information retrieval point of view, but also in the broader context of modelling the process of collecting and representing knowledge. Some important algebraic properties of N-systems representing “dynamic” changes of information have been studied in [8] and also in [9].

## 2. A COMPLETE SYSTEM

Historically this subject was originally conceived independently by Wong, Chiang [15] and Pawlak [12]. The ideas of [12] were elaborated upon by Marek and Pawlak in [6] and finally in [16]. These papers were concerned mainly with a formalization of the subject and investigation of logical properties of the adjoined language. Then in [16], combinational problems were attached, connected with organization of the memory of a computer while implementing an information storage and retrieval system. Further developments along this line were presented in [3], [4], [5] and [8].

In this section we will present basic definitions and some notions that enable us to introduce a common approach to known models of databases and their generalization.

### Definition 2.1 (ISR System)

An information storage and retrieval system (ISR system) is a quadruple  $S = \langle X, A, R_I, U \rangle$  where

- i)  $X$  is a set called the set of objects,
- ii)  $A$  is a nonempty set of descriptors, and  $R_I$  is an equivalence on  $A$  with equivalence classes  $A_i$  ( $i \in I$ ),
- iii)  $U : A \rightarrow \rho(X)$  is a function satisfying the following two conditions:
  - 1) If  $aR_I b$  and  $a \neq b$  then  $U(a) \cap U(b) = \emptyset$
  - 2)  $\cup \{U(b) : bR_I a\} = X$  for each  $a \in A$  where  $\rho(X)$  is the set of subsets of the set  $X$ .

Note that conditions (1) and (2) can be expressed equivalently by one condition, namely  $U(a) = X \setminus \cup \{U(b) : aR_I b, a \neq b\}$  for each  $a \in A$

It is obvious that any set can be identified with its characteristic function. Then let us put

$$(1) \quad V(a,x) = \begin{cases} 1 & \text{if } x \in U(a) \\ 0 & \text{if } x \notin U(a) \end{cases}$$

The function  $V : A \times X \rightarrow \{0,1\}$  is the characteristic function of the set  $U(a)$ , for each  $a \in A$ .

### Lemma 2.2

The function  $V : A \times X \rightarrow \{0,1\}$  which satisfies the condition:

$$(2) \quad \sum_{a \in A_i} V(a,x) = 1 \text{ for each } x \in X \text{ and } i \in I$$

uniquely determines the function  $U : A \rightarrow \rho(X)$  which satisfied the conditions (1) and (2) in Definition 2.1.

*Proof:* Let the function  $U: A \rightarrow \rho(X)$  be defined as follows:  $U(a) = \{x \in X; V(a,x) = 1\}$ . From condition (2) it follows that for each  $x \in X$  there is exactly one element  $a_0 \in A_i, i \in I$ , such that  $V(a_0, x) = 1$  but for all  $b \neq a, b \in A_i, V(b, x) = 0$  then, for each  $x \in X$  there is exactly one descriptor  $a \in A_i$  such that  $x \in U(a)$ . Moreover, there are no two different elements  $a, b \in A_i$  such that  $V(a, x) = V(b, x) = 1$  then  $U(a) \cap U(b) = \emptyset$  for each pair of different elements  $a, b \in A_i$ .

### Corollary 2.3

The function  $U$  in the definition of the ISR system can be replaced by the function  $V: A \times X \rightarrow \{0,1\}$  with the condition  $\sum_{a \in A_i} V(a, x) = 1$  for each  $x \in X$  and  $i \in I$ .

## 3. AN L-INCOMPLETE SYSTEM

Generally speaking, the problem of information incompleteness seems to be very complex. First of all, let us notice that there are many kinds of information incompleteness or, at least, situations which superficially look like information incompleteness. Let us mention some of them:

- 1) "We don't know but we shall know in a specified time."
- 2) "We don't know and we shall never know."
- 3) "We don't know and it is not known whether or not we shall ever know."
- 4) "We do not even know whether the question makes any sense."
- 5) "It is very probable (almost sure) that ..."
- 6) "With probability  $p$ ..."
- 7) "With probability  $p \in \langle \alpha, \beta \rangle \subseteq \langle 0, 1 \rangle$ "
- 8) "To some extent (to some degree)..." et cetera.

In [3] and [4] W. Lipski has extended the ISR model to cover only some cases of information incompleteness. Roughly speaking, the theory of L-incomplete systems concerns cases (1), (2) and (3), between which we do not involve any distinction. For a fact  $F$  we shall consider only three cases:

- i) We know  $F$  is true,
- ii) We know  $F$  is false,
- iii) We don't know whether or not  $F$  is true.

Here we make the following basic assumption: *If we know at a certain time that an object  $x$  has property  $a$  (does not have property  $a$ ) then it is not possible for it to turn out not to have property  $a$  (to have property  $a$ ) later on.* To satisfy this condition we can assume that: *Our information may be incomplete but is correct, and the objects are time-invariant with regard to the properties.*

To deal with the process of increasing the knowledge of information in his system, Lipski [3] has introduced many algebraic notions and has investigated their properties but this falls beyond the scope of our study.

Now we shall precisely specify what we mean by an information L-incomplete system. Below we give a brief description of a mathematical model of such an L-incomplete system, which later will call the L-system.

Let  $A$  be a "disjoint union" of all attribute domains of  $A = \{\langle i, a \rangle; i \in I \wedge a \in A_i\}$

The elements of  $A$  can be called elementary properties: We say that  $x$  has property  $\langle i, a \rangle$  if for object  $x$  attribute  $i$  takes value  $a$ . Our knowledge of properties of objects is represented by two functions:

$u : A \rightarrow \rho(X)$  (the lower valuation)       $U : A \rightarrow \rho(X)$  (the upper valuation)  
where  $\rho(X)$  denotes the set of all subsets of  $X$ .

Intentionally,  $u(i,a)$  is the set of objects known to have property  $\langle i,a \rangle$ , whereas  $U(i,a)$  is the set of objects which possibly has this property, i.e.  $X \setminus U(i,a)$  is exactly the set of objects known not to have property  $\langle i,a \rangle$ . If our information is to be consistent then

(1)  $u(i,a) \subseteq U(i,a)$  for all  $\langle i,a \rangle \in A$ .

Moreover, we may assume that

(2)  $u(i,a) = X \setminus \{U(i,b) : b \in A_i \wedge b \neq a\}$  for all  $(i,a) \in A$ .

Indeed, since  $A_i$  contains all possible values attribute  $i$  can take, we know that  $i$  takes the value  $a$  exactly when we know that  $i$  does not take any other value  $b \in A_i \setminus \{a\}$ . From the above two conditions on  $u$  and  $U$  it follows easily that

(3)  $u(i,a) \cap U(i,b) = \emptyset$  for all  $a \neq b, a, b \in A_i$

and

(4)  $\cup\{U(i,a) : a \in A_i\} = X$

Note that, by (2) the function  $u$  is uniquely determined by  $U$  (but not conversely!).

Now we are ready to summarize a formal definition of an informational L-system. We will go back to the same notation as we have been using in the ISR system definition.

### Definition 3.1 (L-system)

An information system with incomplete information (L-system for short) is a quadruple

$$S = \langle X, \{A_i\}_{i \in I}, U, u \rangle$$

where

$X$  is the set of objects

$I$  is the finite set of the numbers of attributes

$A_i$  is a nonempty set called a domain of the  $i$ -th attribute.

$u, U : A \rightarrow \rho(X)$  are the functions called lower and upper value (respectively), which satisfy the two following conditions:

$$(1) u(a) \subseteq U(a) \text{ for each } a \in A, A = \bigcup_{i \in I} A_i$$

$$(2) u(a) = X \setminus \bigcup_{b \in A_i \setminus \{a\}} U(b) \text{ for each } a \in A_i, i \in I.$$

To introduce a common definition of all the kinds of systems considered, let us define the following function  $V : A \times X \rightarrow 2^{\{0,1\}} \setminus \{\emptyset\}$  such that

$$V(a,x) = \begin{cases} \{1\} & \text{if } x \in u(a) \\ \{0,1\} & \text{if } x \in U(a) \setminus u(a) \\ \{0\} & \text{if } x \in X \setminus U(a) \end{cases}$$

The function  $V$  satisfies the following conditions:

$$(*) \forall (x \in X) [V(a,x) = \{1\} \Leftrightarrow \forall (b \neq a) a, b \in A_i \Rightarrow V(b,x) = \{0\}]$$

The above condition follows directly from the condition (2) of Definition 3.1.

### Lemma 3.2

Each function  $V : A \times X \rightarrow 2^{\{0,1\}} \setminus \{\emptyset\}$  which satisfies condition (\*) uniquely determines the functions  $u, U : A \rightarrow \rho(X)$  which satisfy conditions (1) and (2) of Definition 3.1.

*Proof:* Let us put  $u(a) = \{x \in X; V(a,x) = \{1\}\}$ ,  $U(a) = \{x \in X; V(a,x) \in \{\{0,1\}, \{1\}\}\}$

It is obvious that  $u(a) \subset U(a)$  for each  $a \in A$ . From the condition (\*) it follows that if  $a=b$  and  $a, b \in A_i$  then  $u(a) \cap U(b) = \emptyset$  then  $u(a) \subset X \setminus \bigcup_{b \in A_i \setminus \{a\}} U(b)$ .

We will now show the contrary inclusion “ $\supset$ ”. Let  $x \in u(a)$ . Assuming that from the condition (\*) we have  $\forall (b \neq a) a, b \in A_i \wedge V(b, x) \neq \{0\}$ . Let  $b_0$  be such an element. It means that  $x \in U(b_0)$

and in the conclusion  $u(a) = X \setminus \bigcup_{b \in A_i \setminus \{a\}} U(b)$ . It completed the proof of the Lemma.

$$\begin{matrix} b \neq a \\ b \in A_i \end{matrix}$$

### Corollary 3.3

The functions  $u$  and  $U$  in Definition 3.1 of the L-system may be replaced by the function  $V : A \times X \rightarrow 2^{\{0,1\}} \setminus \{\emptyset\}$  which satisfied the condition (\*).

## 4. STOCHASTIC INFORMATIONAL SYSTEM

In [7] M. Maczynski has introduced a notion of a numerical Boolean algebra and he has investigated some algebraic properties of this notion on a purely theoretical basis.

In [14] T. Traczyk and M. Marek and in [13] T. Traczyk have shown that the numerical Boolean algebras might be used as a convenient tool in dealing with the representation of knowledge in the informational systems.

In particular, Traczyk has presented a natural way of introducing a fuzzy informational system which only produces a (probability) measure that an object  $x$  has a property  $a$ . Firstly, let us recall the notion of a numerical Boolean algebra.

### Definition 4.1 (Numerical Boolean Algebra)

A Boolean algebra  $B$  is said to be numerical (see [7]) over  $X$  if

- i)  $B \subseteq \langle 0, 1 \rangle^X$  (the set of functions from  $X$  into the closed interval  $\langle 0, 1 \rangle$ ),
- ii) the natural ordering of functions
 
$$f \leq g \text{ iff } (\forall x \in X) f(x) \leq g(x)$$
 coincides with the Boolean ordering
 
$$f \leq g \text{ iff } f \vee g = g$$
- iii) if  $f \wedge g = 0$  (the zero-constant function) then  $f \vee g = f + g$  (the arithmetical addition),
- iv)  $(1 - f) \wedge f = 0$  for all  $f \in B$  where  $1 - f$  denotes the arithmetical subtraction.

It is proved in [7] that every Boolean algebra has a numerical representation and the construction of such a representation is very simple. Namely, if we have a full set of measures (A set of measures,  $M$ , on a Boolean algebra  $B$  is said to be full provided  $(\forall m \in M) m(a) \leq m(b)$  implies  $a \leq b$  for all  $a, b \in B$ ), say  $M$ , on a given Boolean algebra  $B$  and  $\tilde{B}$  is the set of all functions  $\tilde{a} : M \rightarrow \langle 0, 1 \rangle$  which are defined by

$$v) \quad \tilde{a}(m) = m(a), m \in M, a \in B,$$

then  $\tilde{B}$  can easily be endowed in the structure of a numerical Boolean algebra over  $M$  in such a way that the correspondence  $a \rightarrow \tilde{a}$  is an isomorphism between  $B$  and  $\tilde{B}$ . On the other hand, if

B in a numerical Boolean algebra over any set X then  $m_x(a) = a(x)$  for all  $x \in X$  and  $a \in B$  defines a measure on B and  $\tilde{X} = \{m_x : x \in X\}$  is a full set of measures on B. For details see [7].

We shall now present the formal definition of a stochastic informational system.

### Definition 4.2 (Stochastic System)

By a stochastic informational system we mean a triple  $\langle X, \{B_i\}_{i \in I}, B \rangle$  where

- i) X is a nonempty set (the set of objects);
- ii)  $\{B_i\}$  is a family of numerical Boolean algebras over X (algebras of attributes);
- iii) B is a numerical Boolean algebra over X as well as a coproduct of the family  $\{B_i\}_{i \in I}$  (the algebra of descriptors).

For each  $x \in X$  and  $a \in B_i$  the number  $m_x^{(i)}(a) = a(x)$  may be thought of as the logical value of the information that x has the description a. It may also be considered as the (probability) measure that the observation of the attribute  $B_i$  on x results in a.

It is convenient to consider  $B_i, i \in I$  as subalgebras of B. Then, for each  $a \in B$  of the form

$$a = \bigwedge_{i \in I} a_i \text{ where } a_i \in B_i, \text{ we have the equality } a(x) = m_x(a) = \prod_{i \in I} m_x(a_i) = \prod_{i \in I} m_x^{(i)}(a_i) \text{ where } m_x$$

denotes the Cartesian product of measures  $m_x^{(i)}, i \in I$ .

If B is a finite Boolean algebra and  $\{a_i, \dots, a_{i_n}\}$  is the set of atoms in  $B_i, i \in I$ , and

$I = \{1, \dots, m\}$ , then the following double sequence of numbers

$$A(x) = \begin{bmatrix} a_{11}(x) & \dots & a_{1n_1}(x) \\ \dots & \dots & \dots \\ a_{m1}(x) & \dots & a_{mn_m}(x) \end{bmatrix}$$

it said to be a full description of  $x \in X$ . It has the following two properties:

$$(a_{i1} \vee \dots \vee a_{in_i})(x) = a_{i1}(x) + \dots + a_{in_i}(x) = 1 \text{ for all } x \in X, i \in I,$$

and

$$(a_{1j_1} \wedge \dots \wedge a_{mj_m})(x) = a_{1j_1}(x) \cdot \dots \cdot a_{mj_m}(x) \text{ for each permutation } (j_1, \dots, j_m) \text{ of } I.$$

We say that an ISR system  $\langle X, A, R_I, U \rangle$  induces a stochastic informational system

$\langle \tilde{X}, \{B_i\}_{i \in I}, B \rangle$  provide  $X \subset \tilde{X}$  and there exists a set of isomorphisms  $h_i$  of  $\rho(A_i)$  onto  $B_i, i \in I$ , such that for every  $i \in I$  and for every  $a \in A_i$

$$h_i(a)(x) = \begin{cases} 1 & \text{if } x \in U(a) \\ 0 & \text{otherwise} \end{cases}$$

### Theorem 4.3

- (1) Each ISR system  $\langle X, A, R_I, U \rangle$  induces a stochastic informational system  $\langle X, \{B_i\}_{i \in I}, B \rangle$ .
- (2) If B is a finite numerical Boolean algebra over X and  $a(x) = 1$  or 0 for every  $a \in B$  and  $x \in X$ , then the stochastic informational system  $\langle X, \{B_i\}_{i \in I}, B \rangle$  is induced by an ISR system.

*Proof:* follows directly from the above considerations.

In the case of stochastic systems we may also introduce some function V, as an element of common definition of all the kinds of information systems considered. Let

$A_i = \{a_{i_1}, a_{i_2}, \dots, a_{i_{n_i}}\}$  be a set of the atoms of the algebra  $B_i$ , and  $A = \bigcup_i A_i$ .

We will define the function  $V : A \times X \rightarrow \langle 0,1 \rangle$ ;  $V(a,x) = a(x)$  for each  $a \in A$ ,  $x \in X$ , the function  $V$  satisfies condition

$$(*) \sum_{a \in A_i} V(a,x) = 1 \text{ for each } x \in X, i \in I.$$

This is a direct consequence of the facts about numerical Boolean algebras [7]. Note that the function  $V$  is uniquely defined by the numerical Boolean algebra  $B$ , but not every function  $V$  that satisfied the above condition (\*) will identify the set of atoms of some numerical Boolean algebra over  $X$ .

## 5. AN N-INCOMPLETE INFORMATIONAL SYSTEM

A very natural consequence of the above considerations is a generalization of all the systems mentioned. What it entails is a construction of the mathematical model of databases with more general incompleteness of stored information. We will call the new system an N-incomplete system or N-system for short.

An N-system is a relational model of a database with incomplete information. The following cases of noncompleteness of information are admissible:

- It is not known whether or not an object  $x \in X$  has a property  $a \in A$ .
- We know that the probability that the object  $x \in X$  has a property  $a \in A$  is a number from the interval  $\langle \alpha, \beta \rangle \subseteq \langle 0,1 \rangle$ .

Let  $C(\langle 0,1 \rangle)$  denote the set of all subcontinua of interval  $\langle 0,1 \rangle$ . Let  $V : Z \rightarrow C(\langle 0,1 \rangle)$  be a map. For several  $z \in Z$ ,  $V(z)$  is a segment  $\langle \alpha, \beta \rangle \subseteq \langle 0,1 \rangle$ . The number  $\alpha$  and  $\beta$  will be called the lower and upper value of  $V(z)$  and will be denoted by symbols  $\alpha = V_*(z)$ ,  $\beta = V^*(z)$ .

### Definition 5.1 (N-system)

By N-system we mean a quadruple  $S = \langle X, A, R_I, V \rangle$  where

$X$  is a finite, nonempty set of objects

$A$  is a finite set of descriptors

$R_I$  is an equivalence relation on  $A$ ; the equivalence classes of  $R_I$  we mean as attributes and are denoted by  $\{A_i\}_{i \in I}$

$V : D_V \rightarrow C(\langle 0,1 \rangle)$  where  $D_V$  is a subset of product  $A \times X$  and the following conditions are satisfied:

$$(1) \sum_{a \in A_i} V_*(a,x) \leq 1 \text{ for } i \in I$$

$$(2) \text{ if for some } x \in X \text{ and } i \in I; \{a \in A_i; (a,x) \in D_V\} = A_i \text{ then } \sum_{a \in A_i} V^*(a,x) \geq 1$$

## 6. THE RELATIONSHIPS BETWEEN INFORMATIONAL SYSTEMS

In this section we will show that the N-system previously defined is a common generalization of the ISR system, the L-system and the stochastic system.

Firstly we will define an induction of an N-system by the L-system and the stochastic system and then we will prove that each L-system and each stochastic system induces some N-system.

### Definition 6.1

We will say that an L-system  $\langle X, \{D_i\}_{i \in I}, U, u \rangle$  induces an N-system  $\langle X, A, R_I, V \rangle$  if there exists a function  $h_i$  such that  $h_i : D_i \xrightarrow{1-1} A_i$  and

$$V(h_i(a), x) = \begin{cases} \{1\} & \text{if } x \in u(a) \\ \langle 0, 1 \rangle & \text{if } x \in U(a) \setminus u(a) \\ \{0\} & \text{if } x \notin U(a) \end{cases}$$

### Definition 6.2

We will say that a stochastic system  $\langle X, \{B_i\}_{i \in I}, B \rangle$  induces an N-system  $\langle X, A, R_I, V \rangle$  if there exists an isomorphism  $h_i : \tilde{A}_i \rightarrow B_i$  such that  $[h_i(b)](x) = \tilde{V}(b, x)$  for  $b \in \tilde{A}_i, i \in I$  where  $\tilde{V}$  is an extension of the function  $V$  on the set of algebraic expressions  $\tilde{A}$ ;

$\tilde{V} : \tilde{A} \times X \rightarrow C(\langle 0, 1 \rangle)$  defined as:

If  $b = a_{i_1} \vee \dots \vee a_{i_k}$  then  $\tilde{V}(b, x) = \langle \tilde{V}^*(b, x), \tilde{V}^*(b, x) \rangle$  where

$$\tilde{V}^*(b, x) = \max\left(\sum_{j=1}^k V^*(a_{i_j}, x), 1 - \sum_{i \in \{i_1, \dots, i_k\}} V^*(a_i, x)\right)$$

$$\tilde{V}^*(b, x) = \min\left(\sum_{j=1}^k V^*(a_{i_j}, x), 1 - \sum_{i \in \{i_1, \dots, i_k\}} V^*(a_i, x)\right)$$

### Theorem 6.3

Every L-system induces some N-system. Moreover, each N-system  $S = \langle X, A, R_I, V \rangle$  such that  $V(a, x) \in \{\{1\}, \langle 0, 1 \rangle, \{0\}\}$  which satisfies two conditions

$$V^*(a, x) + \sum_{b \in A_i \setminus \{a\}} V^*(b, x) \geq 1 \qquad V^*(a, x) + \sum_{b \in A_i \setminus \{a\}} V^*(b, x) \leq 1$$

for each  $i \in I, a \in A_i, x \in X$  is induced by some L-system.

*Proof:* Let us recall two important facts about L-systems.

(1)\* if  $a$  and  $b \in D_i$ , and  $a \neq b$  then  $u(a) \cap u(b) = \emptyset$ .

(2)\*  $\bigcup_{a \in D_i} U(a) = X$ .

Let a L-system be:  $\langle X, \{D_i\}_{i \in I}, U, u \rangle$  and  $A_i = D_i, A = \bigcup_{i \in I} A_i, R_I = \bigcup_{i \in I} (A_i)^2$  and let the function

$V : A \times X \rightarrow C(\langle 0, 1 \rangle)$  be the function defined as follows:

$$V(a,x) = \begin{cases} \{1\} & \text{for } x \in u(a) \\ \langle 0,1 \rangle & \text{for } x \in U(a) \setminus u(a) \\ \{0\} & \text{for } x \notin U(a) \end{cases}$$

It is clear that  $h_i = \text{id}_{A_i}$  is the function which satisfied the conditions of Definition 6.1, and also that the function  $V(a,x)$  satisfies the conditions of Definition 6.1. We must just check whether or not the system  $S = \langle X, A, R_I, V \rangle$  is an N-system.

The equivalence classes of the relation  $R_I$  are the sets  $A_i$ . We will show that

$\sum_{a \in A_i} V^*(a,x) \leq 1$  and  $\sum_{a \in A_i} V^*(a,x) \geq 1$  for each  $i \in I$  and  $x \in X$ . If  $\sum_{a \in A_i} V^*(a,x) > 1$  then there would exist elements  $a, b \in A_i$ ,  $a \neq b$ , such that  $V(a,x) = V(b,x) = \{1\}$ , so  $x \in u(a) \cap u(b)$  but this is in conflict with (1)\*. If  $\sum_{a \in A_i} V^*(a,x) < 1$  then  $V(a,x) = \{0\}$  for each  $a \in A_i$ . Hence, it

follows that  $x \notin U(a)$  for each  $a \in A_i$ , and then  $x \notin \bigcup_{a \in A_i} U(a)$  but this is a contradiction of (2)\*.

We will now prove the second part of the theorem. Let  $S = \langle X, A, R_I, V \rangle$  be the N-system satisfying the assumptions. Let us put  $u(a) = \{x \in X; V(a,x) = \{1\}\}$  and  $U(a) = \{x \in X; 1 \in V(a,x)\}$ . We will check that  $u$  and  $U$  satisfy the conditions from the definition of an N-system (Definition 5.1).

It is obvious that  $u(a) \subset U(a)$  for each  $a \in A$ . Let  $a \in A_i$  be fixed. If  $x \in u(a)$  or  $V^*(a,x) = V^*(a,x) = 1$  then from the assumptions about the function  $V$ , it follows that for each

$b \neq a$ ,  $b \in A_i$ ,  $1 \notin V(b,x)$  or if  $x \in u(a)$  then  $x \notin \bigcup_{b \in A \setminus \{a\}} U(b)$  or  $x \in X \setminus \bigcup_{b \in A \setminus \{a\}} U(b)$

On the other hand, if  $x \in X \setminus \bigcup_{b \in A \setminus \{a\}} U(b)$  then  $x \notin U(b)$  or  $1 \notin V(b,x)$ , hence it follows that

$V^*(a,x) = V^*(a,x) = 1$  so  $x \in u(a)$ . In this way we have proved that  $u(a) = X \setminus \bigcup_{b \in A_i \setminus \{a\}} U(b)$  for

each  $i \in I$ ,  $a \in A_i$

### Corollary 6.4

The N-system induced by an L-system is the greatest element in its equivalence class of the relation  $\rho$  (see [9]).

This follows from Corollary 2.21 in [9] and the second part of Theorem 6.5.

### Theorem 6.5

Every stochastic system induces some N-system.

*Proof:* Let  $A_i = \{a_1^i, a_2^i, \dots, a_{k_i}^i\}$  be the set of atoms of the algebra  $B_i$  and let us put

$$A = \bigcup_{i \in I} A_i, \quad R_I = \bigcup_{i \in I} (A_i)^2, \quad V: A \times X \rightarrow C(<0,1>), \quad V^*(a_j^i, x) = V^*(a_j^i, x) = a_j^i(x).$$

Since  $A_i \cap A_j = \emptyset$  then  $A_i$  are equivalence classes of the relation  $R_I$  in  $A$ .

We will now prove that  $S = \langle X, A, R_I, V \rangle$  is an N-system.

$$\sum_{j=1}^{k_i} V^*(a_j^i, x) = \sum_{j=1}^{k_i} V^*(a_j^i, x) = \sum_{j=1}^{k_i} a_j^i(x) = 1$$

The last equality follows from the following property of a full description of  $x \in X$ :

$$(a_{i1} \vee \dots \vee a_{in_i})(x) = a_{i1}(x) + \dots + a_{in_i}(x) = 1 \quad \text{for all } x \in X \text{ and } i \in I.$$

It is easy to see that the function  $V$  satisfies the conditions (1) and (2) in Definition 5.1.

We will show that  $S$  is induced by  $\langle X, \{B_i\}_{i \in I}, B \rangle$ . The assignment  $\tilde{a}^{i_j} \rightarrow a_j^i$  generates a natural homeomorphism  $h_i: \tilde{A}_i \rightarrow B_i$ .

Let  $b = a_{j1}^i \vee \dots \vee a_{jm}^i \in \tilde{A}_i$ .

$$[h_i(b)](x) = (h_i(a_{j1}^i) + \dots + h_i(a_{jm}^i))(x) = h_i(a_{j1}^i)(x) + \dots + h_i(a_{jm}^i)(x) = a_{j1}^i(x) + \dots + a_{jm}^i(x) =$$

$\tilde{V}^*(b, x) = \tilde{V}^*(b, x)$ . The consecutive equalities above follow from the fact that  $B_i$  is a numerical Boolean algebra, from the definition  $h_i$  and from Corollary 2.3 [9] ( $S$  is the N-system with single-element representations of objects).

## Corollary 6.6

Every ISR system (as a particular case of the L-system) induces some N-system.

Moreover, every N-system  $S = \langle X, A, R_I, V \rangle$  such that  $V(a, x) = \{\{1\}, \{0\}\}$  and  $\sum_{a \in A_i} V^*(a, x) = 1$

for each  $i \in I$  and  $x \in X$ , is induced by some ISR-system.

The proof of this Corollary is merely a slight modification of the proof of Theorem 6.3.

Finally we may state that: For each storage and information retrieval system considered, there is always some corresponding N-system with the function  $V$  which, apart from the conditions (1) and (2) in Definition 5.1, must satisfy some additional relationships related to its counter-domains. These are:

1) An N-system induced by an ISR system

$$V: A \times X \rightarrow \{(0,0), (1,1)\} \text{ such that } \sum_{a \in A_i} V^*(a, x) = \sum_{a \in A_i} V^*(a, x) = 1 \text{ for each } i \in I, x \in X.$$

(2) An N-system induced by a stochastic system

$$V: A \times X \rightarrow \{(t, t); t \in <0,1>\} \text{ such that } \sum_{a \in A_i} V^*(a, 1) = \sum_{a \in A_i} V^*(a, x) = 1 \text{ for each } i \in I, x \in X.$$

(3) An N-system induced by an L-system

$$V: A \times X \in \{(0,0), (1,1), (0,1)\} \text{ such that } V^*(a, x) + \sum_{b \in A_i \setminus \{a\}} V^*(b, x) \geq 1, V^*(a, x) + \sum_{b \in A_i \setminus \{a\}} V^*(b, x) \leq 1 \quad \text{for each } i \in I, x \in X.$$

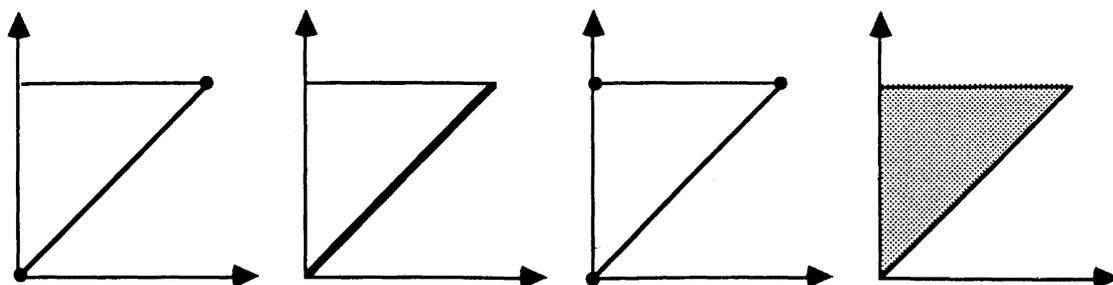
We will now illustrate the counter-domains of the function  $V$  in the N-system induced by different systems. We use the geometric characterization of the set of subcontinua of the interval  $<0,1>$ . See [10].

1) ISR system

2) Stochastic system

3) L-system

4) N-system



We have therefore shown that the notion of an N-system is the generalization of all the informational systems considered.

## 7. CONCLUSION

The notion of information incompleteness seems to be inherent in the domain of data bases. The emergence of expert systems as one of the major areas of activity within artificial intelligence has resulted in a rapid growth of interest within the artificial intelligence community in issues relating to the management of uncertainty and incomplete information. However, very little has been done towards clarifying the problems connected with incomplete information and creating a theoretical background for studying them.

We have presented a mathematical model of incomplete information data bases, which we call an N-system. The notion of an N-system is the generalization of a complete system, an L-system and a stochastic system.

The theory of N-systems can be rephrased within Codd's relational model of data.

The model presented here is intended to provide a basic notion and logical foundation for studying problems connected with sufficiently general information incompleteness. We have shown that for each informational system considered there is always some corresponding N-system.

## REFERENCES

1. Codd E.F., [1970], A relational Model of data for large shared data banks, *CACM*, **13**, 6, 377-387.
2. Jaegermann M., [1975], Information storage and retrieval systems — mathematical foundations. Systems with incomplete information, *CC PAS Report*, **214**.
3. Lipski W., [1981], On databases with incomplete information, *JACM*, **28**, 1.
4. Lipski W., [1979], On semantic Issues connected with incomplete information databases, *ACM TODS*, **4**, 3.
5. Lipski W. and Marek W., [1979], Information systems: On queries involving cardinalities, *Information Systems*, **4**, 3, Elmsford, NY: Pergamon Press, Inc.
6. Marek W. and Pawlak Z., [1979], Information storage and retrieval systems: Mathematical foundations, *Theoretical Computer Science*, **1**, 4.
7. Maczynski, M., [1973], On some numerical characteristics of Boolean algebras, *Coll. Math*, **27**, 207-210.
8. Orłowska M.E., [1980], Algebraic and topological properties of data bases within complete information. D.Sc. thesis. Technical University of Warsaw, Poland.
9. Orłowska M.E., [1987], Dynamic properties of N-incomplete databases, *Demonstratio Mathematica* (to appear). (Preliminary version available as Unisa Research Report RR/87-03.)
10. Orłowska M.E., [1987], Geometric characterization of the set of subintervals of the interval  $\langle 0,1 \rangle$ , *Unisa Research Report RR/87-01*.

11. Orłowska M.E., [1986], Geometric representation of objects in relational databases. Paper presented at *SACLA conference*.
12. Pawlak Z., [1973], Mathematical foundations of information retrieval, *CC.PAS Reports* , 101.
13. Traczyk T., [1978], Common extension of Boolean information systems, *Fundamenta Informaticae*, II.
14. Traczyk, T. and Marek W., [1976], Stochastic information systems, *CC.PAS*, 247.
15. Wong E. and Chiang T.C., [1971], Canonical structure in attribute based file organisation, *CACM*, 14.
16. Lipski W. and Marek W., [1975], On Information storage and retrieval systems, *CA.PAS Reports*, 200.

## NOTES FOR CONTRIBUTORS

The purpose of the journal will be to publish original papers in any field of computing. Papers submitted may be research articles, review articles and exploratory articles of general interest to readers of the journal. The preferred languages of the journal will be the congress languages of IFIP although papers in other languages will not be precluded.

Manuscripts should be submitted in triplicate to:

Prof. G. Wiechers  
INFOPLAN  
Private Bag 3002  
Monument Park 0105  
South Africa

### Form of manuscript

Manuscripts should be in double-space typing on one side only of sheets of A4 size with wide margins. Manuscripts produced using the Apple Macintosh will be welcomed. Authors should write concisely.

The first page should include the article title (which should be brief), the author's name and affiliation and address. Each paper must be accompanied by an abstract less than 200 words which will be printed at the beginning of the paper, together with an appropriate key word list and a list of relevant Computing Review categories.

### Tables and figures

Tables and figures should not be included in the text, although tables and figures should be referred to in the printed text. Tables should be typed on separate sheets and should be numbered consecutively and titled.

Figures should also be supplied on separate sheets, and each should be clearly identified on the back in pencil and the authors name and figure number. Original line drawings (not photocopies) should be submitted and should include all the relevant details. Drawings etc., should be submitted and should include all relevant details. Photographs as illustrations should be avoided if possible. If this cannot

be avoided, glossy bromide prints are required.

### Symbols

Mathematical and other symbols may be either handwritten or typewritten. Greek letters and unusual symbols should be identified in the margin. Distinction should be made between capital and lower case letters; between the letter O and zero; between the letter I, the number one and prime; between K and kappa.

### References

References should be listed at the end of the manuscript in alphabetic order of the author's name, and cited in the text in square brackets. Journal references should be arranged thus:

1. Ashcroft E. and Manna Z., The Translation of 'GOTO' Programs to 'WHILE' programs., *Proceedings of IFIP Congress 71*, North-Holland, Amsterdam, 250-255, 1972.
2. Bohm C. and Jacopini G., Flow Diagrams, Turing Machines and Languages with only Two Formation Rules., *Comm. ACM*, 9, 366-371, 1966.
3. Ginsburg S., *Mathematical Theory of Context-free Languages*, McGraw Hill, New York, 1966.

### Proofs

Proofs will be sent to the author to ensure that the papers have been correctly typeset and *not* for the addition of new material or major amendment to the texts. Excessive alterations may be disallowed. Corrected proofs must be returned to the production manager within three days to minimize the risk of the author's contribution having to be held over to a later issue.

Only original papers will be accepted, and copyright in published papers will be vested in the publisher.

### Letters

A section of "Letters to the Editor" (each limited to about 500 words) will provide a forum for discussion of recent problems.

