ON THE REPRESENTATIONS OF DATA

G. Ricci†

CS-74-20

Department of Computer Science
University of Waterloo
Waterloo, Ontario, Canada

October 1974

† Supported by NRC of Canada, Grant A-4096.
Present address: Istituto di Matematica, Università di Parma.

ON THE REPRESENTATIONS OF DATA

G. Ricci†

Department of Applied Analysis and
Computer Science
University of Waterloo

ABSTRACT

A representation of a set A is just an injection of A into a set of functions (considered as families of reference objects). Two distinct treatments of representations are provided.

The former treatment is of analytic nature. It breaks a representation into certain parts and studies a kind of homomorphism between them. This yields three properties of representations (Reducibility, Redundancy and Independence) which describe how a representation "transmits" information from the objects being represented to the reference ones. All these properties are characterized and their mutual relationships are shown.

The latter treatment is of synthetic nature. It characterizes representations in terms of lattices of equivalences. Thus, all representation properties are mapped into basic properties of lattices.

Besides these treatments, several commonly used representations are considered, in order to check how the formal definitions are able to model them together with their properties. Moreover, to check the practical value of the theory, the solutions of two problems are outlined.

The former problem is to characterize the structural features of computation which depend on the representation of the data. The latter is to extend an independence notion used in Probability Theory. Its solution gives a third characterization of representation independence.

---

† Supported by NRC of Canada, Grant A-4096.
Present address: Istituto di Matematica, Universitá di Parma.

Section 0 - INTRODUCTION

0.1 Motivation

Representations of numbers, state assignments for sequential machines, representations of functions by series or transforms and geometric reference systems are cases of representations. Each of them is treated in the literature by a distinct theory. The main purpose of this work is to begin a general theory of such representations regardless of the nature of the objects being represented.

The opportuneness of a general theory arises in treating the problem open in [11]. There, it is shown how certain local features of some algorithms depend both on the nature of the operation to perform and on the representation chosen for the data. This was an extension of the similar approach for sequential machines [7].

The open problem was to characterize the structural features of a larger variety of computations. This also required to study how a general representation transmits information from the set being represented [12].

Afterwards, B. Forte pointed out to the author a problem arising from Information Theory. Its solution (see 8.4 and 8.6) makes use of the very theory formerly devised for computational purposes.

0.2 Notation

Set notation is as in [9] except what listed below:

ϕ	the empty set
$\bar{A}$	the complement of A relative to a set understood from the context
i	the set $\{i\}$ (when safe from ambiguity)
A^2	$A \times A$
λ	$\{\phi\}$

N	the set of natural numbers
$\underline{n}$	the set $\{0,1,\ldots,n-1\} \subseteq N$
ω	the identity relation on some set
P(A)	the set of subsets of A
r(A)	the r-image of a relation or function r (if B is the domain and $B \cap P(B) \neq \phi$, this notation is ambiguous. However, the context will help to distinguish images from values)
B^A	the set of functions from A to B
fg	the composition of g and f
$(a_i)_I$	a family of index I, i.e. a function with domain I
$\prod_I A_i$	the product of the family $(A_i)_I$
f(a;b)	the value at $b \in B$ of function f(a) for $a \in A$ and $f:A \to C^B$
f(;b)	the function from A to C, which, for f as above, takes value f(a;b) at $a \in A$.
f(A';b)	the set of f(a;b) for $a \in A' \subseteq A$ and f as above
$\mathcal{P}(A)$	the boolean set lattice $(P(A), \subseteq)$
$\mathcal{E}(A)$	the lattice of equivalences on A, $(E(A), \subseteq)$

We always will use intersection in some $\mathcal{P}(A)$ (A being understood from the context). Thus, $\cap\phi = A$. Moreover, to denote surjections and injections, we will use the modified arrows $\twoheadrightarrow$ and $\rightarrowtail$ respectively.

0.3 Notions assumed

We assume the reader familiar with complete lattices [13], [2], [6]. However, we list some notions which are not frequent in the literature.

Sets of generators. In a complete lattice $(A,\leq)$ a set $G \subseteq A$ is called a set of (upper) generators if any $a \in A$ is infimum generable from G, $a = \wedge G'$ for some $G' \subseteq G$.

Basis. An (upper) basis of the above lattice is the partially ordered set $(B,\leq)$, where B is the smallest set of generators. If B exists, it is given by the set of infimum irreducibles.

Reducible lattice. If a complete lattice has a basis, it is called reducible. An example of an irreducible lattice is an oriented closed segment of the real line.

Branching. Let $\wedge : P(A) \to A$ be an infimum and let $\wedge' : P(P(A)) \to P(A)$ be the function which yields the images of $\wedge$. The identity $\wedge\wedge' = \wedge\cup$, where the union is on $P(P(A))$, will be referred as the branching property (see also the complete associativity in [2], p.118).

Moreover, the natural mapping of an equivalence will be referred as the block function of the corresponding partition. Thus, $f:A \twoheadrightarrow$ is a block function iff it satisfies the membership property: $b = f(a)$ iff $a \in b$ for all $a \in A$.

Section 1 - PRELIMINARIES

We introduce the preliminary notions about representations, namely embedding in a representation, isomorphism and equivalence among representations.

1.1 Assignments and Representations- Definition

We first introduce some auxiliary notation and terminology. Consider a function

(1) $r:A \to U^I$.

As an example, r might assign to each point P of the ordinary space A the triple of its coordinates in a cartesian reference system, $r(P) = (x_0,x_1,x_2)$. Here, $I = \underline{3}$, U is the set of real members and the triple is thought of as a family.

For any $J \subseteq I$ let $r_J:A \to U^J$ be the function such that $r_J(a)$ is the restriction to J of the function $r(a):I \to U$. We will call r_J the J-projection of r and $r_J(a)$ the J-projection of r(a). For a singleton $J = \{j\}$ we should consider two different functions, namely the j-projection $r_j:A \to U^{\{j\}}$ and the function $r(;j):A \to U$. However, we can safely identify these functions as well as their values, which will be called axes and coordinates respectively. Since the atoms in I determine the axes, also the elements of I/ω, or of I itself, will be called axes.

Function r will be called a (coordinate) assignment of A if no axis r_i is a constant. (If we delete this minor restriction, proposition 4.3 will become weaker, but all the remainder will not). An assignment will be called a representation if r is one to one. In this case, a value r(a) will be called the representation of a.

Let $C_i = r(A;i)$ for $i \in I$, then we have a family $(C_i)_I$ of sets of coordinates which will be called the components of r and r will be called an assignment on this family of components. Note that $r:A \to \prod_I C_i$ and that, since no axis is constant,

(2) $\phi \subset r_i^{-1}(c) \subset A$ for all $i \in I$ and $c \in C_i$.

Moreover, since r_ϕ is a constant with value ϕ,

(3) $r_\phi(B) = \lambda$ for all B with $\phi \subset B \subseteq A$,

whereas

(4) $r_J(\phi) = \phi$ for all $J \subseteq I$.

These limiting cases will be useful later on.

1.2 Embedding and Equivalence - Definition

Let r be an assignment as above. We will say that a set $L \subseteq I$ of axes is embedded into another set $K \subseteq I$ and we will write $K > L$, if

(5) $fr_K = r_L$ for some $f:r_K(A) \twoheadrightarrow r_L(A)$.

Function f will be called the embedding coefficient between K and L or between r_K and r_L. If f is a bijection we will say that K and L are equivalent, $K \simeq L$.

Embedding and equivalence are relations in P(I), but it is convenient to extend them. Given any $g:A \twoheadrightarrow X$, thought of as an additional axis, we will say that g is embedded into K and we write $r_K > g$, if $fr_K = g$ for some $f:r_K(A) \twoheadrightarrow X$. Let now $s:A \to V^J$ be another assignment. If (5) holds after substituting s_L, with $L \subseteq J$, for r_L, then we still say that $K > L$ and, when f is bijection, that $K \simeq L$. Thus embedding and equivalence become relations in $P(I) \cup P(J)$.

1.3 Embedding Properties - Proposition

Given assignments r and s as above, we have the following properties:

a) Embedding is a preorder on $P(I) \cup P(J)$.

b) Embedding contains the containment orders of $P(I)$ and $P(J)$.

c) The equivalence induced by the embedding preorder is equivalence $\simeq$.

Proof a) follows from (5) and from the composition associativity of functions applied to the embedding coefficients. b) follows from the fact that, if $K \supseteq L$, the values of L-projections are determined by restricting the K-projection values. c) composing all the embedding coefficients along a circuit from a set to itself, we must get the identity. Hence, these coefficients have to be one to one as required. Q.E.D.

1.4 Isomorphic and Equivalent Representations - Definition

Two assignments $r:A \to U^I$ and $s:A \to V^J$ are said to be isomorphic if there is a bijection of equivalent axes between the two indices, i.e. for some $p:I \twoheadrightarrow J$, $p(i) \simeq i$ for all $i \in I$.

Assume now that we have $p:P(I) \to P(J)$ and $q:P(J) \to P(I)$ such that $p(L) \simeq L$ and $q(K) \simeq K$ for all $L \subseteq I$ and $K \subseteq J$. Then r and s will be called equivalent. In other words, two assignments are equivalent iff every set of axes of the one is equivalent to some set of the other.

A trivial consequence of these definitions is the following.

1.5 Isomorphism and Equivalence - Proposition

If two assignments of a set are isomorphic, then they are equivalent. If they are equivalent, they need not be isomorphic.

Proof If two assignments are isomorphic, each axis of the one is equivalent to an axis of the other. Thus, given any set of axes of the one, the corresponding set of axes of the other is clearly equivalent to the former. Therefore, we have equivalence.

To prove that the converse is not true, take a representation r with a finite number of axes. We always can find an equivalent representation by adding to r an axis which is an exact copy of one of the axes of r. Yet the latter representation cannot be isomorphic to the former. Q.E.D.

Section 2 - REDUCIBILITY

Isomorphism and equivalence coincide in a class of representation which are called reduced. This class has a natural extension into a very large class of representations, called reducible, which still does not exhaust the class of all representations. This study of reducibility is included mainly for mathematical completeness.

2.1 Useless Axes and Reduced Representations

Definition An axis i of an assignment $r:A \to U^I$ will be called useless if it is equivalent to a set of axes J not containing it, i.e. $J \simeq i$ for some $J \subseteq \bar{i}$. If r has no useless axes, then r will be called reduced. Note that reduction is preserved under isomorphism.

2.2 Equivalence in the Reduced Case

Proposition Two reduced assignments are equivalent if and only if they are isomorphic.

Proof Let $r:A \to U^I$ and $s:A \to V^J$ be two assignments. By 1.5 we have only to show that, when they are reduced, equivalence implies isomorphism.

Let $p:P(I) \to P(J)$ relate equivalent sets of axes, then we may assume it to be "atomic", i.e. $p(L) = \cup p(L/\omega)$ for all $L \subseteq I$. The same is assumed for $q:P(J) \to P(I)$.

By transitivity, $i \simeq q(p(i))$ for all $i \in I/\omega$. Thus, since r is reduced,

(6) $i \subseteq q(p(i))$,

which implies by atomicity that, for some $j \in J/\omega$, $j \subseteq p(i)$ and $i \subseteq q(j)$. Therefore by 1.3, $i > j$ and $j > i$. This shows that for each $i \in I/\omega$ we have a $j \in J/\omega$ such that $i \simeq j$, i.e. we have a $p':I \to J$ relating equivalent axes.

We claim that ' is the required bijection. In fact, it must be one to one, because $p'(i) = j = p'(i')$ implies $i \simeq i'$ and, by reduction, $i = i'$. It also must be onto, because by repeating the reasonment for (6) we get $j \in p'(q(j)) \subseteq p'(I)$ for all $j \in J$. Note that the hypothesis of reduction on s is used only to prove this surjectivity.

In conclusion r and s are isomorphic. Q.E.D.

This theorem shows that the (possible) reduced representation in a class of equivalent representations is unique (up to an isomorphism). We want to characterize this reduced representation by means of the following definition and theorem.

2.3 Reducible Representations - Definition

An assignment $r:A \to U^I$ is said to be reducible if there exists a $K \subseteq I$ such that r_K is a reduced assignment of A which is equivalent to r. In other words, r is reducible iff we can get an equivalent reduced assignment only by deleting some axes. By 2.2 we are allowed to call r_K the reduced assignment of r.

2.4 Reducibility characterization - Proposition

An assignment s is equivalent to a reduced assignment r if and only if s is reducible. (Hence, the only way to get equivalent reduced assignments up to isomorphisms is through axis deletion.)

Proof ("If" part). Trivial.

("Only if" part). Recall the proof 2.2. The only missing hypothesis is the reduction of s. Thus, we still have the injection $p':I \rightarrowtail J$ Take $K = p'(I)$. Then s_K is an assignment isomorphic to r. Hence s_K is reduced and by transitivity, it is equivalent to s, which turns out to be reducible. Q.E.D.

The following is a direct corollary of definition 2.3. It characterizes equivalence among reducible representations in terms of isomorphism.

2.5 Equivalence for reducible assignments - Proposition

Two reducible assignments are equivalent if and only if their reduced assignments are isomorphic.

Proof Trivial. Q.E.D.

Since the reducible representations are just reduced representations with some useless axes added, we could restrict our attention to reduced representations (and by 2.2 avoid equivalence problems). However, there are representations which are not reducible.

2.6 Irreducible representation - Example

Let A be the set of continuous real functions of a real variable. Let R be the set of (finite) real numbers and for $i \in R$ let (i] be the open-closed interval of the reals up to i inclusive and let (i) be the open-open one.

Define a representation r of A on $(C_i)_I$ by

$$I = R,$$

$$C_i = \{f \mid f \in R^{(i]} \text{ and } f \text{ is continuous}\} \quad \text{for all } i \in I,$$

$$r_i(f) = f_{(i]} \quad \text{for all } i \in I \text{ and } f \in A.$$

Trivially r satisfies the conditions for being a representation. Here the coordinates of a function $f \in A$ are its "left-hand portions" obtained by considering f in a real interval from $-\infty$ to i inclusive. We could say that r gives us a "cumulative" representation of f, by an analogy with the terminology of the integral calculus; see also 3.5.

Observe that, since A is a set of continuous functions, an axis $i \in I$ can be equivalent to a set $X \subseteq I$ of axes if and only if i is the supremum of X according to the usual partial order of the real line, i.e. iff i is a "right accumulation point" for X.

We show now that r cannot be reducible. Suppose that for $J \subseteq I$, r_J is equivalent to r. Since each axis $i \in I$ must be equivalent to some subset X of J, then, by the previous observation, J must be everywhere dense on I. This implies that each $j \in J \subseteq I$ is equivalent to $(j) \cap J \subseteq J\text{-}j$. So all the axes in J are useless in r_J. Hence, no r_J can be reduced. Therefore, r is not reducible.

Section 3 - REDUNDANCY

Redundancy is the simplest notion arising from embedding. It appears in commonly used representations frequently and has a very natural characterization in terms of embedding.

3.1 Redundant Representations - Definition

An axis i of an assignment $r:A \to U^I$ will be called redundant if it is embedded into a set of axes not containing it, i.e.

(7) $J > i$ for some $J \subseteq \bar{i}$.

If r has redundant axes, it will be called a redundant assignment. An axis or an assignment which are not redundant will be called nonredundant.

In (7) any $J \in P(\bar{i})$ can be used. There are cases where a restricted choice of J is interesting. For instance, in 5.3 we will consider the finite sets in $P(\bar{i})$ and we will speak of finite-redundancy. However, these subcases of a definition will not be studied.

A useless axis is redundant. Hence a nonredundant assignment is reduced. However, the converse need not be true. The following is an example of a reduced and redundant representation.

3.2 Error-detecting Code - Example

The simplest error-detecting code which is a reduced and redundant representation is a three-bits encoding of a four-elements set as in Fig.1.

A	r_1	r_2	r_3
a	0	0	0
b	0	1	1
c	1	0	1
d	1	1	0

Fig.1

By inspection on this table we see that no axis is useless. Yet any axis is redundant. For instance, the third axis can be computed as the "parity-bit" of the first two, but we cannot derive the latters from the former.

As known, this representation has the power to detect errors which may affect a single bit. This can be done by comparing the values of r_3 with the values of r_1 and r_2. Essentially, this comparison makes use of the embedding coefficient between {1,2} and 3. Hence, here the embedding coefficients have the role of "control" function for error detection. A similar role might be seen for error-correcting codes like those using crossed parities.

The property of being nonredundant is characterized in terms of embedding by a converse of proposition 1.3b).

3.3 Redundancy and embedding - Proposition

An assignment $r:A \to U^I$ is nonredundant if and only if its embedding relation reduces itself to the contaminment order on P(I).

Proof ("only if" part). By 1.3 we know that $>$ contains $\supseteq$. Suppose that r is nonredundant and by contradiction that $\supseteq$ does not contain $>$, i.e. $J > K$ and $K \not\subseteq J$ for some $J,K \subseteq I$. Take an axis $k \in K-J$, then $K > k$ and by transitivity $J > k$ contrary to the nonredundancy of k.

("if" part). Suppose that $>$ and $\supseteq$ coincide. Thus $J > k$ for some $J \subseteq \bar{k}$ is a contradiction. Therefore, there are no redundant axes. Q.E.D.

3.4 Contractile assignments - Definition

An assignment $r:A \to U^I$ will be called contractile if there exists $J \subseteq I$ equivalent to I such that r_J is nonredundant. We find an example of this in 3.2, where we can delete any one of the axes and still get a two-bits representation of A which clearly is nonredundant.

We will not characterize the class of contractile assignments as we did for the reducible ones. We want only to point out the independency of this notion from the preceding ones. An irreducible assignment, indeed, can be contractile. Take, for instance, the irreducible representation of 2.6 and add an "identity" axis ∞ such that $r(f;\infty) = f$. The new r is still irreducible, but it is contractile, because r_J is still a representation, if we take $J = \{\infty\}$. Moreover, even a reduced representation need not be contractile as shown below.

3.5 Unavoidable redundancy - Example

Recall example 2.6 about continuous functions and modify it slightly. Take the coordinates to be reals, $C_i = R$ for all $i \in I = R$, and assume r to be the identity, $r(f;i) = f(i)$ for all $f \in A$ and $i \in I$. (Note that by a "left to right accumulation" of the axes we come back to the representation of 2.6).

r is still a representation, but it is reduced. In fact, since we cannot guess a value of a continuous function by another value at a different point, no axis can be equivalent to a different set of axes.

However, every axis is redundant. In fact, in order to know the value of a continuous function at a given point i it is sufficient to know its values in any set X of points which have i as an accumulation point. Thus $X > i$, though the converse fails. (Note also that every axis is finitely nonredundant as in 3.1).

If we throw away any single axis i, the set $\bar{i}$ remains equivalent to I. We can do the same for any set of i's provided that what remains is everywhere dense. Hence, again we will always have redundant axes in what must remain, since it will be dense. In conclusion, there is not a subset of axes equivalent to I which forms a nonredundant representation. Thus, r is not contractile.

Section 4 - INDEPENDENCE

Many commonly used representations satisfy an independency condition. This condition can be characterized in terms of embedding as the generalization of the idea of using the "right" number of coordinates for a geometric variety. The condition of independence, together with the conditions treated in the preceding sections, forms a hierarchy among representations.

4.1 The Independence Notion - Comment

In order to introduce the notion of independence between axes or set of axes, we consider first the various relationships introduced so far between sets of axes of a given representation. The first relationship was equivalence which means equality of transmitting power. This was an extreme case; we then considered a weaker case with the relation of embedding which means containment of transmitting power. If we go on in this direction, we have to consider the extreme case which is the opposite of equivalence, i.e. the case in which no transmitting power is shared between different sets of axes.

To say that no transmitting power can be shared is a stronger requirement than to say that no equivalence or embedding may exist. Indeed we are requiring that "nothing" can be done by a set of axes if it is done elsewhere.

Suppose that a set J of axes is able to "express" a property P of the set to be represented. (E.g. the last bit of the binary positional representation of a natural number expresses the fact that the number is odd or even). Suppose now that a set of axes K, disjoint from J, is able to express property P. Then J and K are able to express at least one property in common, and we will say that they are not independent. Hence, we will say that J and K are independent if no such a property P exists.

We can extend the previous notion from the case of disjoint J and K to the general case, by saying that no property P as above may exist unless it is expressed by the axes in $J \cap K$. So two arbitrary sets J and K are independent if the properties expressed both by J and by K are also expressible by the axes which are common to J and K. Finally, from a couple {J,K} we can pass to any set of sets of axes and this extension leads us to the following definition.

4.2 Independence - Definition

Let $r:A \rightarrow U^I$ be an assignment and $J \subseteq I$ be a set of axes. We will say that J expresses a property of certain elements of A or that it expresses the set $B \subseteq A$ of the elements which satisfy this property, if B can be identified through the J-coordinates, i.e. if $B = r_J^{-1}(r_J(B))$ or equivalently if $r_J(B) \cap r_J(\bar{B}) = \phi$. In this case we will write $J\eta B$.

We can extend the relation $\eta \subseteq P(J) \times P(A)$ to additional axes. Thus, if $g:A \rightarrow X$ can be substituted for r_J in the above equalities, we will write $g\eta B$. Note that $J > K$ and $K\eta B$ imply $J\eta B$ and that the same holds substituting g for J.

We will say that a set $Q \subseteq P(I)$ of sets of axes is independent or that the sets in Q are independent, if any property expressed by all sets in Q is expressed by their intersection, i.e. for all $B \subseteq A$.

(8) $J\eta B$ for all $J \in Q$ implies $\cap Q \eta B$.

Note that from (3) and (4) we get $\phi \eta B$ iff $B = \phi, A$. Thus condition (8) for $\cap Q = \phi$ tells us that no (nontrivial) property can be expressed by all sets in Q.

We will say that r is (completely) independent, if each $Q \subseteq P(I)$ is independent. If an assignment or a set are not independent, they will be called dependent.

Note that an independent assignment must be an (independent) representation. In fact, for $Q = \phi$ the hypothesis in (8) is always true and, since $\cap \phi = I$, we must have $I \eta B$ for all $B \subseteq A$. This requires that r is one to one.

As for the case of redundancy, weaker notions of independence are possible. This can be done by restricting the choice of sets Q (e.g. in order to define a "pairwise" independence) or the choice of the members of Q. In the latter case, because of (8) we would require that the class of these members is closed under intersection (e.g. finite sets, recursive sets, etc.). We could also restrict the choice of the sets B. Weaker independences may arise in some cases (see e.g. 5.5 and 5.3), but we will limit ourselves to the study of complete independence.

4.3 Independence and redundancy - Proposition

If a representation is redundant then it is not independent. However, there are nonredundant representations which are dependent.

Proof Let $i \in I$ be a redundant axis of representation r. Since $\bar{\bar{i}} > i$ and $\bar{i} > \bar{i}$, we easily get $\bar{i} > I$. (In general, we could show that embedding defines a "pair" lattice). Thus, since r is one to one, $\bar{\bar{i}} \eta B$ for all $B \subseteq A$ as observed in 4.2.

Take $B = r_i^{-1}(c)$ for some $c \in C_i$. From (2) we know that B cannot be trivial and by construction that $i \; \eta \; B$. Therefore, $Q = \{i, \bar{i}\}$ and B satisfy the hypothesis in (8). Since $\cap Q = \phi$, the conclusion in (8) cannot be fulfilled by a nontrivial B. Hence, r is dependent.

To prove that there are dependent assignments without redundant axes, consider the representation of the table in Fig.2 which is "reversed" by the matrix of Fig.3.

A	r_1	r_2
a	1	0
b	2	0
c	0	1
d	0	2

Fig.2

	0	1	2
0		c	d
1	a		
2	b		

Fig.3

From the table we see that no axis is redundant. Yet from the matrix we see that $1 \; \eta \; B$ and $2 \; \eta \; B$ for $B = \{a, b\}$. This contradicts the independence condition in the case of disjoint sets. Q.E.D.

The following example will introduce a characterization theorem for independence which will play an important role.

4.4 Independence and right dimensionality - Example

Let A be the surface of the three dimensional sphere $x^2+y^2+z^2 = 1$ and, as usual, denote the axes of the cartesian representation employed by x, y and z; see Fig.4

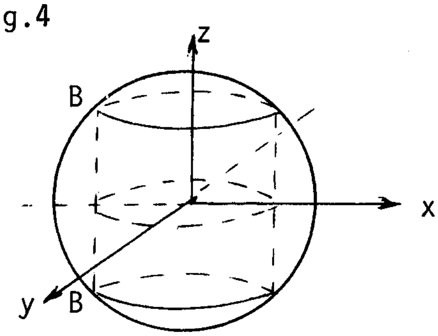

Fig.4

Consider the set $B \subseteq A$ of the points lying on two parallels of symmetric latitude. From the figure we easily see that $z \eta B$ and also that $\{x,y\} \eta B$. These two sets of axes are disjoint, but the set B is not trivial. Therefore, the representation is dependent.

(Note that this representation is nonredundant as well as the counterexample of 4.3. Moreover, that example is a simplification of the present one. In fact, the matrix of Fig.3 may be thought of as the circle $x^2+y^2 = 1$ in the field of remainders mod.3).

Let us modify our representation by taking the longitude of a given point rather than its x and y coordinates. Thus we have the axis z as above and a new axis ψ defined by $\psi(P) = \text{arctang } y(P)/x(P)$.

The new representation is independent. In fact, $z \eta B$ and $\psi \eta B$ are tantamount to saying that B is an union of parallel circles and that B is an union of meridion semicircles. Clearly, this implies that B must be either empty or the whole surface. Since the other cases of (8) are trivial, this shows that the independence condition is fulfilled.

In conclusion, the former representation with three axes was dependent whereas the latter (with as many axes as the dimension of the variety being represented) is independent. This suggests us that independence might be related to an idea of "right" number of axes. This, indeed, is the case as the following proposition will show by extending the above geometric concepts to our general set-theoretic case.

4.5 Independence and embedding - Proposition

Let $r:A \to U^I$ be a representation. A set $Q \subseteq P(I)$ is independent if and only if, for all additional axes $g:A \to X$.

(9) $r_J > g$ for all $J \in Q$ implies $r_{\cap Q} > g$.

(In other words, independence means minimal use of axes or, if we want, maximal expressive power for the axes used. Note that, for $\cap Q = \phi$, (9) means that there are no (nonconstant) additional axes which may be embedded into disjoint sets of axes).

Proof ("If" part). Suppose that (9) is satisfied for all g and that the hypothesis of (8) is true. To prove the conclusion of (8), define $g:A \to \{B,\bar{B}\}$ by $g(a) = B$ iff $a \in B$, i.e. g is the block function of the partition determined by B in A.

Clearly, $r_J > g$ for all $J \in Q$. Thus by (9), $r_{\cap Q} > g$, i.e. we can detect membership in B simply from the axes in $\cap Q$. Hence $\cap Q\ \eta\ B$.

("Only if" part). Assume the hypothesis of (9) is true for an independent Q as in (8). In order to show (9) we must provide the embedding coefficient for its conclusion.

Define a relation $f \subseteq r_{\cap Q}(A) \times X$ by

(10) $(c,x) \in f$ iff $c = r_{\cap Q}(a)$ and $x = g(a)$ for some $a \in A$.

Clearly, if f would be a function, then it will be the required embedding coefficient. Therefore, to prove the theorem we only have to show that $(c,x), (c,y) \in f$ imply $x = y$.

Let $a, b \in A$ satisfy

(11) $r_{\cap Q}(a) = r_{\cap Q}(b) = c$

(12) $g(a) = x, \quad g(b) = y.$

Define $B = g^{-1}(g(a))$. Since $g \,\eta\, B$ and since by hypothesis $J > g$ for all $J \in Q$, we get $J \,\eta\, B$ for all $J \in Q$. Thus from (8), $\cap Q \,\eta\, B$ which by (11) implies $b \in B$. Therefore, from (12) we get $x = y$ as required Q.E.D.

We conclude this section by collecting the results about the relationships among the properties studied.

4.6 The representation hierarchy - Proposition

Assignments or representations of sets can be classified into six classes with respect to embedding properties. These six classes form a hierarchy with respect to proper inclusion as indicated in the diagram below, where a class at the top is more general than one at the bottom.

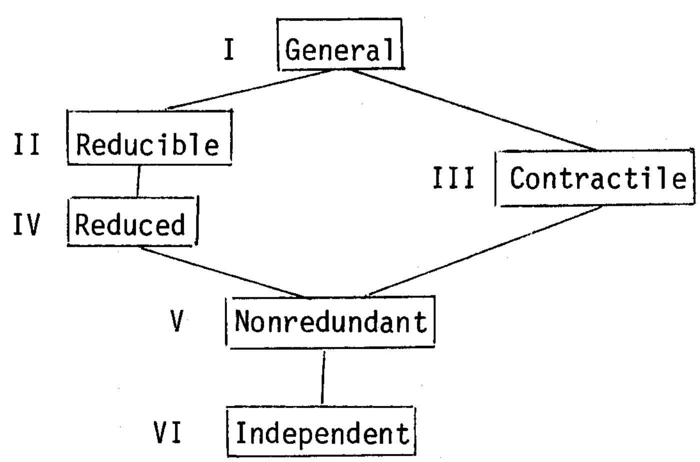

Fig.5

Proof Everything has been substantially proved before. We only give a table of reference where one can check inclusion and noninclusion.

Pair of classes	$\supseteq$	$\nsubseteq$	$\nsupseteq$
I, II	2.3	-	2.6
I, III	3.4	-	3.5
II, IV	2.3	-	Trivial
IV, V	3.1	-	3.2
III, V	3.4	-	3.4
V, VI	4.3	-	4.3
II, III	-	3.5	3.4
III, IV	-	3.2	3.5

Fig.6

Q.E.D.

Section 5 - MODELS OF REPRESENTATIONS

We describe how our formal definitions fit the "real" representations. This yields some further problems about independence. We also discuss how we can handle those "representations" which do not satisfy completely our definition.

5.1 Finiteness characters - Definition An assignment r of a set A on a family $(C_i)_I$ may satisfy some conditions of cardinality which we list below in order to provide a descriptive terminology for the following.

(13) I is a denumerable set.

(14) All the C_i are finite and we assume them to be finite ordinals.

(15) For each $a \in A$ there is a cofinite $J \subseteq I$ (i.e. $\bar{J}$ is finite) such that $r_j(a) = r_k(a)$ for all $j,k \in J$. In other words, there are only a finite number of significant coordinates for each a.

(16) There is a finite ordinal $\underline{u}$ such that $C_i \subseteq \underline{u}$ for all $i \in I$.

We will say that r is

discrete if (13) holds,

finite if (13) and (14) hold,

limited if (13), (14) and (15) hold,

bounded if (13), (14), (15) and (16) hold.

5.2 Positional representations - Example

Taking $I = N$, $C_i = \underline{10}$ for all $i \in I$ and $r(n;i) = [n/10^i]_{\text{mod } 10}$ for all $i \in I$ and $n \in N$ we get the decimal positional representation of the set $A = N$ of natural numbers. This is a case of bounded representations.

We easily see that r is nonredundant, since no set of digits of a natural number can determine another digit (not included in it). r also is independent, but this needs a more detailed proof as follows.

Observe at first that for each pair $b,b' \in A$ and for each set $L \subseteq I$ there exists a $b'' \in A$ such that

(17) $r_L(b) = r_L(b'')$ and $r_{\bar{L}}(b') = r_{\bar{L}}(b'')$.

By contradiction, let r be dependent, i.e. for some $B \subseteq A$ and $Q \subseteq P(I)$ we have

(18) $r_J(B) \cap r_J(\bar{B}) = \phi$ for all $J \in Q$

and $r_{\cap Q}(B) \cap r_{\cap Q}(\bar{B}) \neq \phi$. Therefore, there are $b \in B$ and $b' \in \bar{B}$ such that

(19) $r_{\cap Q}(b) = r_{\cap Q}(b')$.

Let $D(b,b') \subseteq I$ be the set of digits in which b and b' differ. Clearly, $D(b,b')$ is finite. Moreover, by (18) and by (19) we have that

(20) $D(b,b') \cap J \neq \phi$ for all $J \in Q$,

(21) $D(b,b') \cap \bigcap Q = \phi$

Let $K \in Q$; then we can take a $d \in D(b,b') \cap K \subseteq K$ such that by (21) there exists $J \in Q$ with $d \notin J$. Hence by (18) we can take a $b'' \in \bar{B}$ such that (17) holds for $L = d$:

(22) $r_{\bar{d}}(b'') = r_{\bar{d}}(b')$ and $r_d(b'') = r_d(b)$.

Consider now the pair (b,b''). The only change from (b,b') is that $D(b,b'')$ has one element less than $D(b,b')$. Indeed, (19) continues to hold for b'' also. Thus we can iterate our process until (20) fails, $D(b,c) \cap J = \phi$ for some $c \in \bar{B}$. This contradiction implies the independence of r.

The above demonstration holds for every representation which satisfies (22) and the finiteness of D(b,b'). Note also that $\prod_I C_i$ has the cardinality of the continuum whereas r(A) has the same cardinality of N, i.e. we have a poor utilization of the reference objects though the representation is independent.

We can get the decimal positional representation of non-negative real numbers simply by changing the index I to the set of integers. Thus, the representation is finite but not more limited since an irrational number has an infinity of significant digits. Adding to I a binary axis for the sign, we get a representation for all the reals and clearly other easy changes yield other representations related to the positional one, e.g. the floating-point representation. However, in the case of reals, independence opens a larger problem as in the following.

5.3 <u>The pencil and the onto conditions - Comment</u>

Condition (17) requires that whenever r(A) contains two elements, then it must also contain any element obtained by a "mixing" of the two. The following "pencil" condition will turn out to be equivalent to (17).

We will say that $r:A \to U^I$ satisfies the pencil condition if

(23) $$r_{\bar{J}}(r_J^{-1}(d)) = r_{\bar{J}}(A) \quad \text{for all } J \subseteq I \text{ and } d \in r_J(A).$$

This condition says that for each "handle" d we have in r(A) a pencil which has "bristles" in all $r_{\bar{J}}(A)$.

Clearly, if $r_J(b) = d$ then (23) yields (17) and conversely. Hence the pencil condition is equivalent to (17).

If we require that $r:A \to \prod_I C_i$ is onto, we clearly get a stronger condition than the pencil condition as shown by the cardinality considerations in 5.2. Neither this onto condition nor the pencil one are necessary to get independence. The representation of $A = \{a,b,c\}$ by two binary digits is not onto, but it is independent as one can check by the matrix below.

	0	1
0	a	b
1	c	

Fig.7

A problem which seems to be difficult (and we leave open) is proving (or disproving) that the pencil condition or the onto condition are sufficient to get independence. In favor of this conjecture we have the proof in the previous example. In fact, while we require the finiteness of $D(b,b')$, we do not employ the whole pencil condition (17) but a finite one as in (22).

Another point in favor of the conjecture is that the pencil condition is sufficient to prove some conditions of weak independence. E.g. it would not be difficult to prove pairwise independence.

The motivation of this problem lies in the role of independence in several applications (see section 8) and in the role of the onto condition in the literature (see 8.4).

5.4 Modular representation - Example

The modular representation of the set $A = N$ of natural numbers is defined by $I = N$, $C_i = \underline{p}_i$ for all $i \in I$ and

(24) $$r_i(n) = [n]_{\text{mod } p_i} \quad \text{for all } i \in I \text{ and } n \in A,$$

where p_i denotes the (i+1)-th prime number. This representation is limited, because for $i \geq n$, $r_i(n) = n$; but it is not bounded.

As far as the embedding properties are concerned, we easily see that r is dependent and moreover that it is redundant, because every axis is redundant. Indeed, the values of an axis i are determined by the infinite set of axes $j > i$ (which contains, for all $n \in A$, an axis j such that $p_j \geq n$). However, r is not finite-redundant as in 3.1. Indeed, if J is a finite set of axes and i is an axis to be embedded in J, then the set $J \cup i$ is also finite and by the Chinese remainder theorem it determines a finite set of equations (24) which always has a solution no matter what the $r_i(n)$ are. So J cannot contain enough information to determine i (if $i \notin J$).

5.5 Series - Example

An example of a representation which is discrete, but not finite, arises in considering the set A of functions analytic at 0. In this case, we can represent an element $a \in A$ by a series, e.g. a series of powers. I.e. we can use the representation r defined by $I = N$, $C_i = C$ (set of complex numbers) for all $i \in I$ and $r_i(a) = D^{(i)}(a;0)$ for all $a \in A$, $i \in I$, where $D^{(i)}$ denotes the operator of i-th derivative.

We cannot show that this representation is independent by using the proof we have used for the positional representation, because here also D(b,b') need not be finite. However, we clearly can use that proof in order to prove a kind of "finite" independence, namely the independence of all the sets $Q \subseteq P(I)$ which contain a finite set K of axes. Indeed (22) continues to hold since the change of a single coefficient in a power series does not change its convergence.

This reasoning could clearly be extended from analytic functions to functions with a singularity at 0 by the bilateral Laurent's series, or also (with some caution) to larger classes of functions by using "series with a continuum of coefficients" e.g. the Fourier transform.

5.6 Linear spaces and groups - Example

Consider the set A of the points of a linear space which we can assume, for simplicity, to be a pre-Hilbertian space. I.e. we assume there is a binary operation, the scalar product, which takes arguments in A and values in another set C. Then if G is a set of generators of A, we have a representation r of A by taking $I = G$, $C_i = C$ for all $i \in I$ and $r_i(a) = \langle a,i \rangle$ for all $a \in A$, $i \in I$, where $\langle a,i \rangle$ denotes the scalar product of a and i.

We can easily prove that r is independent if and only if G is a linearly independent set of generators, i.e. if it is a basis of A. In other words, our definition of independence is just the extension of the geometric one. (However, in this case independence and nonredundancy coincide.)

What we have said for linear spaces can be repeated (mutatis mutandis) for abelian groups. In fact, in this case, the carrier set of a group can be represented conveniently by representations which are independent and which correspond to the various bases of the group.

The case of abelian groups is perhaps more interesting than the case of a linear space because the reference sets C_i can be quite heterogeneous, e.g. in a normal basis there are both finite C_i, which correspond to the torsion coefficients, and denumerable C_i corresponding to the generators with infinite period. Moreover, for a given group there can be quite dissimilar bases [8]. In other words, groups are intermediate between geometry and sets as far as representations are concerned.

5.7 Assignments and representations - Comment

In practice assignments are more common than representations. To handle a real number we do not use its full positional representation as in 5.2; rather we use an assignment obtained by selecting, from its positional representation, a finite number of axes, which are considered more significant. Similarly, in order to treat functions, we use tabulations, i.e. assignments extracted from the "natural" representation of functions as in 3.5.

Since assignments are motivated by these "extractions", a theory of assignments should look at two problems. The first is to study the properties of such representations from which a possible assignment can be extracted. The second problem is to define, for a given representation, the criteria of selection of the axes forming an assignment.

Numerical analysis consider the latter problem, e.g. when it introduces the notion of truncation and of step of tabulation. It also considers the former problem, e.g. when it compares the representation of a function by series of powers or of polynomials.

Our main concern is the study of the general properties of representations and it is more related to the first of the two problems discussed. Indeed, the second problem should consider some kind of evaluation of the choice of the axes forming an assignment, i.e. it is related to the problem of measuring how information is transferred from a set to the reference sets. We prefer to limit ourselves to studying how information is transferred without measuring it.

5.8 Linked data structures - Example

Some practical representations do not satisfy the condition of having a unique representation of every object. An example of this arises when we consider the use of "linked data structures" for representing objects like strings, trees, graphs, etc.

Consider the case of a string, say $w = \sigma_0\sigma_1\sigma_2$, over an alphabet Σ. We may have a set $I = N$ of locations, each of them having a set $C_i = (\Sigma \cup \not{b}) \times (I \cup \not{b})$ of configurations, where $\not{b}$ is a symbol neither in Σ nor in I. In other words, any location may contain a pair (σ,i), where σ is a letter of Σ or is $\not{b}$ and i is a location of I or is $\not{b}$. Then we can "represent" w by starting from a fixed location, say $i = 0$, by any one of the allocations of I, depicted in Figure 8 (where $\not{b}$ is written as a space).

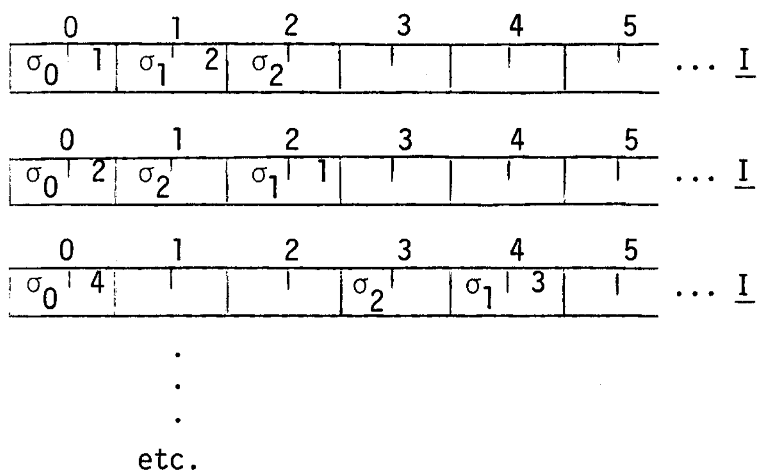

Figure 8

We see that we have a lot of different representations for our unique string w and we know from programming practice that this abundance means a very useful flexibility of representation and not a strange waste of our representation facilities.

Here it is still possible to get uniqueness of representation by enlarging the set under representation, i.e. by saying that the objects which we really want to represent are not simply the words over Σ, but something containing more information. In the following we will explain a way to do so, which also will suggest to us that trying to get uniqueness can sometimes be a good method for understanding the real nature of the representations considered.

If we want a representation of strings by linked data structures with their flexibility, then it means that we are interested in all the possible ways of assigning the index locations. In general, a location assignment for a string $w:\underline{n} \to \Sigma$will be any 1-1 mapping $\psi:\underline{n} \rightarrowtail I$, with $\psi(0) = 0$. In conclusion, when we want to represent a word w by a linked data structure, then we really want to represent a pair (w,ψ) containing a word and its location assignment.

Under the previous point of view, the representation we are considering is a usual representation r of the set A of pairs (w,ψ) as above. In fact, r is the (biunique) representation mapping defined by $r((w,\psi);\psi(i)) = (w(i),\psi(s(i)))$ where s is the successor function (which identifies our special kind of linking) and where the values are filled by $b\!\!\!/$ when they are undefined.

Note that in order to get the set A we have added to the set of strings the information about the location assignment. This is information about the representation itself! This trick is useful particularly for objects which are functions with a varying domain as strings or trees. (Moreover, Lagrange's multipliers technique might be interpreted in this way [12].)

Note also that we have got uniqueness of representation in a way reminding the "state splitting" technique of [7]. In fact, we have splitted a word w into a set of pairs (w,ψ). However, the study of "extended representations" $r \subseteq A \times U^I$, where r^{-1} is a function, is still an open problem.

Section 6 - REPRESENTATION LATTICE

This section introduces the notion of representation lattice. The representation lattice corresponds to an extension of some classical notions of synthetic geometry (sheaves of coordinate surfaces or coordinate lines, etc.). Well known facts of a geometric nature continue to hold in this extended case. The representation lattice is also an extension of the notion of information used in [7] and elsewhere, and it is a very natural aspect of the notion of representation. The precise relationship between representation lattices and representations is stated and proved.

6.1 Assignment lattice - Definition

Let $r:A \to U^I$ be an assignment. We will call induction map the function $e:P(I) \to E(A)$, such that $e(J)$ for $J \subseteq I$ is the equivalence induced by r_J, i.e. $(a,a') \in e(J)$ iff $r_J(a) = r_J(a')$. The values $e(J)$ will be called J-equivalences and their set $e(P(I))$ will be denoted by E_r.

In the limiting case, by (3) we have the total equivalence

(25) $$e(\phi) = A^2.$$

Moreover, when r is a representation, we have

(26) $$e(I) = \omega.$$

We can order the set E_r by set inclusion (in the same way as for the lattice of all equivalences $\mathcal{E}(A)$) and we get a partially ordered set $\mathcal{E}_r = (E_r,\subseteq)$ which is a lattice by the next proposition. Thus $\mathcal{E}_r$ will be called the lattice of the assignment r.

6.2 Infimum containment - Proposition

The set of J-equivalences E_r of an assignment $r:A \to U^I$ forms, under the inclusion order, a complete lattice which is infimum contained in the lattice $\mathcal{E}(A)$ of all the equivalences on A. Moreover, the induction map $e:P(I) \twoheadrightarrow E_r$ is a dual supremum-homomorphism between the corresponding lattices.

Proof To prove the first statement it is enough to prove that the infima of sets of J-equivalences in $\mathcal{E}(A)$ are J-equivalences. Now, since the infima in $\mathcal{E}(A)$ are intersections, we have for all $F \subseteq E_r$

$$\wedge F = \cap F = \{(a,b) | a,b \in A \text{ and } r_J(a) = r_J(b) \quad \text{for all } J \in G\}$$

where G is any set in P(I) such that e(G) = F. Thus

$$\wedge F = \{(a,b) | a,b \in A \text{ and } r_i(a) = r_i(b) \quad \text{for all } i \in \cup G\} = e(\cup G) \in E_r.$$

To prove that e is a dual supremum homomorphism, simply observe that for any $G \subseteq P(I)$ the set evaluation e(G) = F is contained in E_r. Thus by the previous argument

(27) $$e(\cup G) = \cap e(G) \quad \text{for all } G \subseteq P(I).$$

6.3 Representation lattice - Example

Consider a cartesian representation of the euclidean three-dimensional space as the reference system given by the axes x,y,z of Figure 9.

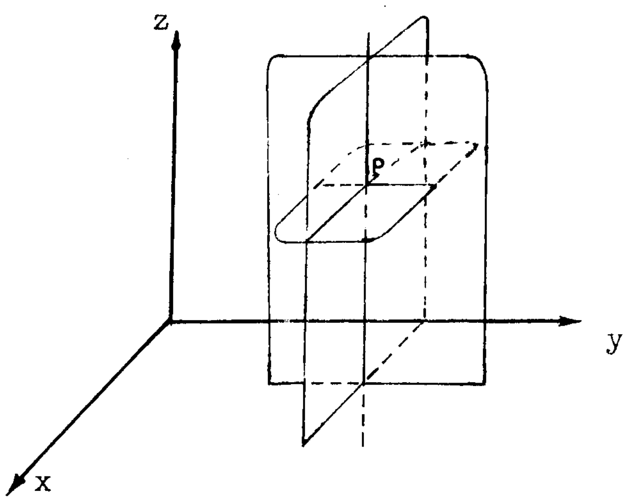

Fig.9

The J-equivalences are given e.g. by the sheaves of planes parallel to a given axis. (Each plane is a block of that equivalence). Other J-equivalences are given by the (bidimensional) sheaves of lines parallel to an axis or by the space or finally, by all the points of the space. In Fig.10 we have drawn the diagram of the representation lattice, where e.g.

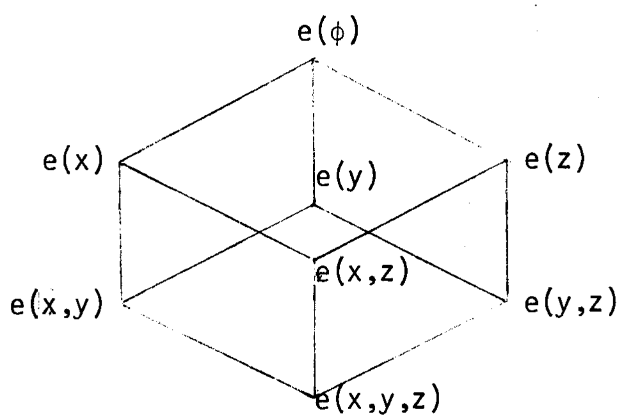

Fig.10

the sheaves of planes correspond to the equivalences e(x), e(y), e(z).

The interpretation of the previous theorem about infimum containment is now very simple. It states for instance that the intersections of coordinate planes are coordinate lines. This property is a very trivial one for our cartesian representation, but it holds for any representation of any object. As we will see in section 7, this property is characteristic for a representation lattice.

6.4 The synthetic approach - Comment

The previous example justifies the terms "synthetic" and "analytic" we are using for the present treatment and for the past (functional) treatment respectively.

A synthetic approach differs from an analytic one mainly because the former does not use "foreign" objects like coordinates to treat the space which is really considered. Axes and coordinates simply become equivalences and blocks defined on the space considered.

A synthetic approach allows a ready and clean exploitation of the intuitive knowledges we may have about the environment we are studying. Thus, setting a problem and interpreting its solution is easier when we already have a good knowledge of the objects being studied.

In [7] the notion of equivalence (or partition) induced by an axis is referred as the "information" conveyed by that axis. In the Theory of the Measure of Information [4] the "things" which are measured (i.e. the information) are again defined as partitions and their blocks. Therefore the synthetic approach which leads us to study representation lattices is the informational approach.

The analytic approach was concerned more with how information is transmitted than with the information itself. Its main advantage is its formalization ability.

We know that at least for existing Geometry, the analytic and the synthetic treatments are equivalent, i.e. they are two different ways of stating the same things. Hence we could ask whether studying representation lattices is as general as studying representations. The two following propositions will show that for many purposes we have such an equivalence.

6.5 The embedding lemma - Proposition

Let $e: P(I) \twoheadrightarrow E_r$ be an induction map as in 6.1. Then the embedding relation $>$ of r is mapped by e onto the inclusion order of E_r, i.e. $J > K$ iff $e(J) \subseteq e(K)$ for all $J, K \in P(I)$. (Consequently, two sets of axes are equivalent if and only if they have the same J-equivalence.)

Proof Suppose $J > K$, i.e. there exists a function

(28) $$f: r_J(A) \twoheadrightarrow r_K(A)$$

such that $fr_J = r_K$. Then clearly, if $r_J(a) = r_J(a')$, we have also $r_K(a) = r_K(a')$. In other words $(a,a') \in e(J)$ implies that $(a,a') \in e(K)$. Therefore, $e(J) \subseteq e(K)$.

Conversely, suppose that, for all $a, a' \in A$, $r_J(a) = r_J(a')$ implies $r_K(a) = r_K(a')$. Then the relation $f = r_K r_J^{-1}$ is a function as in (28). Hence, we have the required embedding. Q.E.D.

6.6 Representation lattice and equivalence - Proposition

Two assignments r and s of a set A determine the same representation lattice, $E_r = E_s$, if and only if they are equivalent.

Proof. If r and s are equivalent as in 1.4 then any set of axes of one of them is equivalent to a set of axes of the other. Hence, by the previous lemma the set of J-equivalences must be the same.

Conversely, if $E_r = E_s$ and if e and e' are the induction maps of r and s respectively, then mappings p and q as in definition 1.4 can be defined by the previous lemma through $e'^{-1}e$ and $e^{-1}e'$ respectively. So r and s are equivalent. Q.E.D.

Note that, since equivalence is a weaker notion than isomorphism, as in 1.5, the synthetic approach is weaker than the analytic one. However, the difference between the two approaches is interesting mainly in the case of irreducible representations as in 2.6. Moreover, we could avoid this difference by using families of J-equivalences rather than sets of J-equivalences.

Section 7 - SYNTHETIC CHARACTERIZATION

This section shows how the analytic properties of representations can be translated in terms of a representation lattice. Hence, we give synthetic characterizations of reducibility, redundancy and independence. We also show how to define a representation, given some essential information about the representation lattice.

7.1 Existence and uniqueness - Proposition

If a complete lattice $\mathcal{E}'$ of equivalences on a set A is infimum contained in the lattice $\mathcal{E}(A)$ of all the equivalences on A, then there exists an assignment r of A which has $\mathcal{E}'$ as representation lattice, $\mathcal{E}' = \mathcal{E}_r$. (This assignment r is unique up to equivalences.)

Proof. Let $\mathcal{E}' = (E', \subseteq)$ be a complete lattice, and let $G \subseteq E'$ be an upper generating set. Since the upper bound given by $A^2 = \wedge\phi$, is trivially reducible, we can always take an upper generating set I which does not contain A^2, $I = G - A^2$.

We can define an assignment $r: A \to \prod_I C_i$ by $C_i = A/i$ for all $i \in I$ and $r(a;i) = [a]_i$, for all $a \in A$, $i \in I$. In other words, the coordinates are the blocks of the partitions and the axes are given by the block functions.

Since $A^2 \notin I$ condition (2) is satisfied and r is an assignment. The infimum containment of $\mathcal{E}'$ and (27) make $E_r \subseteq E'$. In fact, for all $J \subseteq I$, $e(J) = e(\cup J/\omega) = \cap e(J/\omega) \in E'$. Moreover, since I is an upper generating set for $\mathcal{E}'$, we can reread the previous passages in reverse and conclude that $E' \subseteq E_r$. Therefore, $E_r = E'$ as required.

By 6.6 the assignment found is unique up to equivalence. However, this "uniqueness" is rather large. As an example, we could take $I = E'-\{A^2\}$ and get a very non reduced assignment.

7.2 <u>The synthetic characterization corollary - Proposition</u>

A set of equivalences determines an assignment lattice if and only if it contains all its infima. It determines a representation lattice if and only if it contains the null equivalence ω and its infima.

<u>Proof</u>. Trivial by 7.1 and (26). Q.E.D.

7.3 <u>Reducibility characterization - Proposition</u>

An assignment is reducible if and only if its representation lattice is reducible (as in 0.3).

<u>Proof.</u> ("<u>Only if</u>"). If an assignment s of a set A is reducible, then by 2.4 it is equivalent to a reduced assignment $r:A \to U^I$. Therefore, it is enough to prove that the assignment lattice E_r has a basis.

Consider the set $B \subseteq E_r$, $B = e(I/\omega)$ determined by the single axes. Since any $J \subseteq I$ is an union of atoms, by 6.2 we have that B is an upper generating set of E_r. The equivalences in B must be infimum irreducible (otherwise in I/ω we should have useless axes by the embedding lemma 6.5 contrary to the reduction of r). Hence, B forms the required basis of E_r.

("<u>If" part</u>). If an assignment lattice is reducible, we can take its basis B as the upper generating set as in the proof of 6.2. We get in this way an assignment r (equivalent to the original one) which is reduced by the embedding lemma, as in the only if part. Q.E.D.

The following is partly a corollary of the previous characterization of reducibility.

7.4 The basis corollary - Proposition

The basis of a reducible lattice containing its infima, determines a unique assignment lattice which determines a unique (up to isomorphism) reduced assignment.

Proof. An upper generating set (or a basis) of a lattice of equivalences containing its infima determines it. In the case of a basis, i.e. in the case of a reducible lattice, by 7.3 we know that there exists a corresponding reduced-assignment lattice. By 2.2 the equivalence between assignments and the isomorphism coincide in the case of reduced assignments. Hence, we get the uniqueness. Q.E.D.

The previous corollary tells us that for the reduced representations, the bases of the corresponding representation lattices are a very "representative" concept. Thus, we might ask how to characterize such bases.

7.5 Bases characterizations - Proposition

A set $B \subseteq E(A)$ of equivalences over A is a basis of an assignment lattice, if and only if

$$\cap B' \in B \text{ implies } \cap B' \in B' \text{ for all } B' \subseteq B,$$

i.e. if and only if B does not contain its infimum reducibles.

Proof ("Only if" part). If for some $B' \subseteq B$ we have $\cap B' \in B-B'$, then no assignment lattice with basis B may exist. Indeed, this lattice should be infimum contained in $E(A)$, but this means that $\cap B'$ is a reducible element, contrary to the fact that it belongs to the basis.

("If" part). Suppose that B does not contain its infimum reducibles, then define C as the set of all the equivalences obtained by intersection from some subsets of B. We claim that C is closed under intersection (i.e. it defines an assignment lattice), and that its basis is B.

Let $C' \subseteq C$ and let $P' \subseteq P(B)$ be the set of subsets which generate the equivalences of C' by intersection. Consider now $B' = \cup P'$ which is a subset of B. By the branching property as in 0.3 we have $\cap C' = \cap B' \in C$. Thus C is closed under intersections.

It remains to show that the lattice determined by C has B as basis. Clearly, B is an upper generating set. Moreover, it cannot have reducible elements because $b = \cap C'$ for $C' \subseteq C$ means $b = \cap B'$ for some $B' \subseteq B$ as we have just shown. Therefore, B is the required basis and we conclude that C is the required set of equivalences. Q.E.D.

7.6 Redundancy characterization - Proposition

A reduced assignment is nonredundant if and only if its assignment lattice is a boolean set lattice. (This lattice is dually isomorphic to the boolean lattice of the sets of axes,)

Proof. ("Only if" part). Let $e:P(I) \twoheadrightarrow E_r$ be the induction map of an assignment r. It is enough to prove that, if r is nonredundant, e is a dual isomorphism of $P(I)$ onto E_r. Indeed, the embedding relation > on P(I) reduces to the inclusion order $\supseteq$ by 3.3. By the embedding lemma 6.5, e must be 1-1. Thus, if r is nonredundant, e is a bijection which preserves the order (dually), i.e. it is a dual isomorphism. Hence, E is a boolean set lattice.

Note that we can now add to (27) the similar formula for the case of intersection

(29) $\quad e(\cap G) = \vee e(G) \quad$ for all $G \subseteq P(I)$.

("If" part). If E_r is a boolean set lattice, its basis is defined by the set B of its dual atoms. According to 7.3 and 7.4 the (unique) reduced assignment r is given by taking I = B in the construction of the proof 7.1.

By contradiction, assume redundancy, $J > i$ for some $J \subseteq I - i$. By the embedding lemma we should have $e(J) \subseteq e(i)$, but e(i) is a dual atom, $e(i) = i$. This is a contradiction, because in a boolean set lattice any element is exactly the intersection of the dual atoms which contain it, but $e(J) = \cap e(J/\omega)$ and $i = e(i) \neq e(j) = j$ for all $j \in J/\omega$. Therefore, this booleanity implies nonredundancy. Q.E.D.

7.7 Nonredundant representations - Comment

Formula (29), as well as (27), expresses in the geometric case a well known fact. It says, for instance, that the coordinate planes in a cartesian representation of a space, can be obtained as "spans" of couples of coordinate lines.

Almost all the representations used in classical mathematics are nonredundant, though they may be or may not be independent as in example 4.4

concerning the sphere. Redundancy does not have a great deal of importance for many theorical purposes. So nonredundant representations can be used, and we get the nice property that the representation lattice is isomorphic to the lattice of the sets of axes. In this situation we can handle a representation by considering the latter boolean set lattice only. This means that we could introduce representations or reference systems (as we do in classical mathematics), only by speaking of single axes which are the atoms of that lattice.

In our generalized approach we have not used single axes, but we have defined the main notions and properties of representations in terms of sets of axes. This indeed enabled us to handle representations in general.

7.8 Modularity and distributivity - Comment

In our classification of representation lattices, we have gone directly from the reducible lattices to the boolean ones. On our way we have missed some intermediate classes of lattices, e.g. the modular and distributive ones. The properties of being modular or distributive are algebraic properties of lattices, which can be characterized in terms of the absence of some sublattices [2].

For instance, a lattice is not modular iff it contains a sublattice as in Fig.11, it is not distributive iff it contains a sublattice as in Fig.11, or in Fig.12.

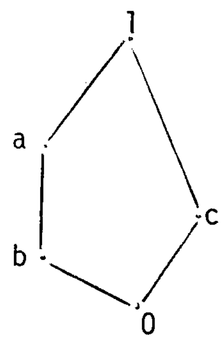

Fig.11

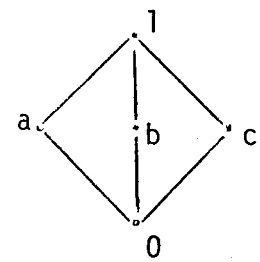

Fig.12

We have not given a formal definition of representations which are characterizable in terms of modular or distributive lattices. However, it should be clear that such "analytic" notions can be introduced and that they correspond to meaningful facts. For instance, the nondistributive case of Fig.12 corresponds to the representation lattice of an error detecting code as in 3.2. Indeed, the elements a, b and c in Fig.12 correspond to the equivalences e(1), e(2) and e(3).

Error detecting codes such as those using parity bits, have a redundancy well shared between the axes as Fig.12 shows. On the contrary, the case of Fig.11 represents a case of "concentrated" redundancy.

However, we will not consider such a detailed classification of representations. Instead, we go on to consider the last case of the previous chapter, namely, independent representations.

7.9 Independence characterization - Proposition

A reduced representation r of a set A is independent if and only if its representation lattice is a boolean set lattice contained in the equivalence lattice of A.

Proof. ("Only if") If $r:A \mapsto U^I$ is an independent representation, then by 4.3 it is nonredundant and by 7.7 E_r is a boolean set lattice dually isomorphic to $P(I)$ by e. Since the infimum containment has already been shown, we need to show only that E_r is supremum contained in $E(A)$:

(30) $e(\cap Q) = V'e(Q)$ for all $Q \subseteq P(I)$,

where V' denotes the supremum in $E(A)$.

Observe that $e(\cap Q)$ is an upper bound, $e' \subseteq e(\cap Q)$ for all $e' \in e(Q)$, because e is a dual isomorphism. Consider now any equivalence $e'' \in E(A)$ such that $e' \subseteq e''$ for all $e' \in e(Q)$. Let g be the block function of e" and consider the hypothesis in (9). Since we can apply the embedding lemma 6.5 to also an additional axis g, this hypothesis is fulfilled and by the independence of r we get $r_{\cap Q} > g$. Therefore, using the embedding lemma again, we get $e(\cap Q) \subseteq e''$. Hence, $e(\cap Q)$ is the required supremum as in (30).

("If" part). In order to prove that containment implies independence it is enough to prove that it implies embedding condition (9). This is immediate by the embedding lemma. Q.E.D.

7.10 Independence and the synthetic approach - Comment

We have shown how the main properties of representations can be recovered by the synthetic approach. This means that essentially we can study a representation simply by some structures defined on the set to be represented. This is particularly easy to do in the case of independent representations, because the representation lattice is a sublattice of the "natural" lattice of equivalences.

On the contrary, when the representation is dependent, one has to use some caution. For instance, in the dependent representation of the sphere 4.4, one finds that the blocks of e({y,z}) v e({x,z}) in E(A) are not parallel circles ℓ, but 4-tuples of points; see Fig.13.

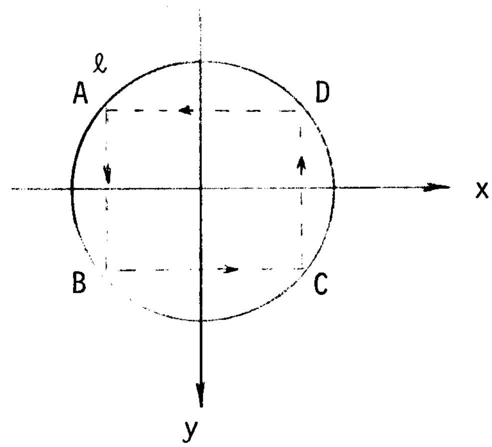

Fig.13

Section 8 - APPLICATIONS

We describe two applications of the present theory. The former which motivated this work [12], studies certain properties of computation which depend on the representation chosen for the data. It is treated here by few examples. The latter is related to probability theory and provides a further characterization of independence.

8.1 Sequentiality structures - Comment

Assume we have a representation r of a set A and an operation $\alpha : A^s \to A$, where s is another set (usually it is a finite ordinal representing the number of arguments). To compute α we determine the coordinates of the result by those of the arguments. This process may compell us to consider certain coordinates before certain others.

For instance, to add natural numbers in the decimal positional representation, we must proceed from right to left. In fact, we can compute a digit of the result only after we have got the carry from its right.

In general, it is possible to define a relation which expresses such a sequentiality among the axes (or sets of axes) of r [12] and we say that α induces a sequentiality structure on r. By an analytic treatment this structure can be characterized in terms of a very natural structure over the "invariant" sets of axes [12]. Here, invariance is just the extension of the geometric invariance of axes under a motion of the referenced space.

Under this extension, the (geometric) eigenvalues become the local operations to perform on the coordinates of the invariant axes [12]. Given

α,a change of r yields a change of the invariant sets and, hence, a change of the sequentiality structure as well as of the local operations.

A corollary of our synthetic treatment yields a further characterization. It states that a sequentiality structure is dually isomorphic to the basis of a certain lattice of congruences [12]. We will not precise these two characterizations nor the definitions involved, but we exhibit them by a simple example.

8.2 Isomorphic basis - Example

Consider the (monadic) operation α on a set A and a binary representation r of A as in Fig.14. Under this encoding we can realize α by a sequential machine with three complementation flip-flops and an "and" gate as in Fig.15.

A	r	α
a	0 0 0	g
b	0 0 1	h
c	0 1 0	e
d	0 1 1	f
e	1 0 0	c
f	1 0 1	d
g	1 1 0	b
h	1 1 1	a

Fig.14

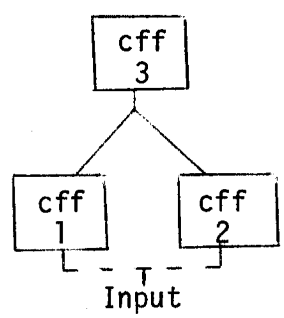

Fig.15

Clearly, we must logically operate first in cff 1 and cff 2 to get their present states and then in cff 3. In this case, the logical constraints expressing the sequentiality structure are materialized in the hardware structure of our device. The sets ϕ, {1,2}, {1}, {2}, {1,2,3} are the "closed" sets of [7] and correspond to our invariant sets. However, the first two ϕ, {1,2} can be deduced from the last three by union. The last three (which are union-irreducible) are somehow sufficient to identify the sequentiality structure of Fig.15 [12]. This is the essential meaning of the former characterization.

As far as the latter characterization is concerned, we consider the J-equivalences of the invariant sets(which are congruences). These form a lattice as in Fig.16, where partitions denote the corresponding congruences. The basis of this lattice (shown in Fig.16 by thick lines) is dually isomorphic to the structure of Fig.15. In other words, the latter characterization states that we can detect sequentiality from synthetic information.

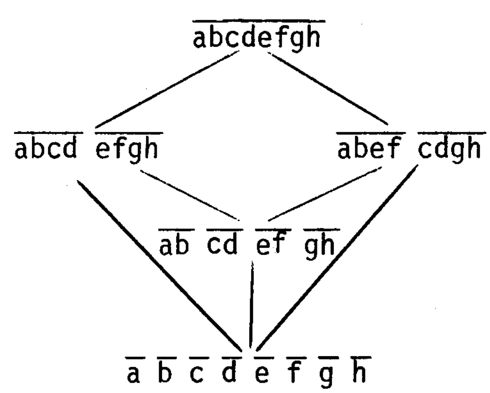

Fig.16

The theory of sequential machines [7] has a somehow similar result (it relates the lattice of congruences to that of closed sets only). However, it holds under certain conditions of nonredundancy. These are more restrictive than the nonredundancy of our theory (which is sufficient to get the outlined synthetic characterization [12]).

Both characterizations of 8.1 hold for any operation $\alpha : A^s \to A$, whatever the cardinalities of A and s are. Thus, they hold for tree-automata as well as for infinitary operations. Hence, this application of representation theory continues the structural study of these objects begun in [11]. In this more general setting, of course, we cannot speak of hardware structure, but of sequentiality structure only, as in the addition case of 8,1.

Though the representations we are considering are not extended representations as in 5.8, they can give rise to unusual structural properties. One of these can be interpreted in terms of structural reliability as shown in the following example.

8.3 <u>Reliability - Example</u>

Consider the (monadic) operation α on A and the representation r of A as in Fig.17.

A	r	α
a	0 1	b
b	0 2	a
c	1 0	d
d	2 0	c

Fig.17

	f	f'
0	0	0
1	1	2
2	2	1

Fig.18

We can use two identical "components" C' and C" with two operations as in Fig.18. To operate each of them, we use the identity (delay) operation f if the argument (state) of the other is not 0, otherwise, we use f'. Thus α can be realized by a structure as in Fig.19 (solid lines).

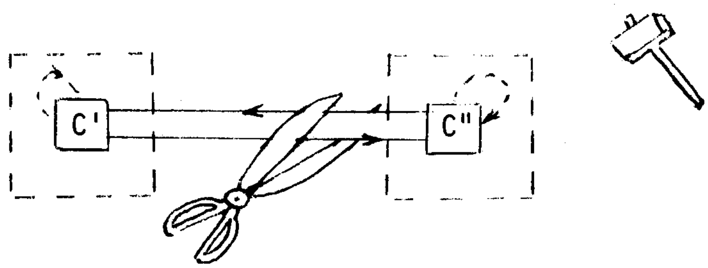

Fig.19

If at each component we become unaware of the state of the other (scissor), we can still operate the device correctly. In fact, since the state of a component is 0 if and only if the other state is not, we can make each single component able to decide about f or f' by itself through some additional device (dotted lines). In this way, we get a structure with two macrocomponents. If one breaks the additional devices (hammer), one can use the (uncut) connections.

In conclusion, the machine described has the property that a single failure either in the (macro)-components or in their connections can be recovered. Note that we have no repetition nor redundancy in the components.

Representation r is the dependent representation of the proof 4.3. One can check that an independent representation (e.g. by two binary axes) would not give rise to such a reliability against failures. It is possible to prove that, in general, being dependent is a characterizing condition for representations allowing such a reliability [12].

8.4 The Onto Condition in Switching and Probability - Comment

The onto condition, $r(A) = \prod_I C_i$, discussed in 5.3, seems to be much more known than independence. Switching Theory [3], [10] relates it to the realizability of functions (on a represented domain) by simpler boolean polynomials. However, as far as structural properties are concerned, the previous example shows that independence rather than the onto condition plays the main role.

In Probability Theory, the onto condition is sometimes called algebraic independence [4]. Essentially, it requires that the block functions b' and b" of two partitions A' and A" (called algebraically independent) on a set A form an onto assignment of A. This is a necessary and sufficient condition for the existence of "stochastic independence".

Stochastic independence requires that any probability function $p:A \to [0,1]$ can be computed by any pair of (marginal) probability functions $p':A' \to [0,1]$ and $p'':A'' \to [0,1]$ by the product formula $p(a) = p'(a')\cdot p''(a'')$ where a' and a" are the blocks determined by a. This also is the basis for the additivity of information measures.

Though foundamental and important, algebraic independence has severe drawbacks. If the cardinality of A is a prime, algebraic independence can never hold. The same is true, if the cardinalities of A and A' are relatively prime. Even the very important set of integers cannot have (non-trivial) onto representations (its cardinality is a "prime").

Since, we must give up algebraic independence in an infinity of cases, some of them being very important, we might ask to state a condition weaker than stochastic independence without these drawbacks. This implies that we cannot require that the product formula (or information additivity) holds (everywhere). However, we still would be able to freely assign marginal probabilities.

For sake of simplicity we state such a weaker condition for general real valued functions on finite partitions.

8.5 Cumulative independence - Definition

Two finite nontrivial partitions A' and A" on the same set are said to be cumulatively independent if, given any two real valued functions $f':A' \to R$, $f'':A'' \to R$ such that

$$(31) \qquad \sum_{x \in A'} f'(x) = \sum_{y \in A''} f''(y),$$

there always exists a real valued function $f:A \to R$, where $A = A' \wedge A''$, such that

$$(32) \qquad \sum_{a \subseteq x} f(a) = f'(x)\ , \quad \text{for all } x \in A',$$

$$(33) \qquad \sum_{a \subseteq y} f(a) = f''(y)\ , \quad \text{for all } y \in A''.$$

In other words, cumulatively independent partitions are those which allow a free choice of marginal functions. These turn out to be well-known partitions.

8.6 Cumulative independence and independence - Proposition

Two partitions are cumulatively independent if and only if their block functions are independent.

Proof. ("Only if") Let A' and A" be two cumulatively independent partitions with block functions b' and b" respectively. By contradiction, assume that b' η B and b" η B for a nontrivial B. Thus we can choose $f':A' \to R$ and $f'':A'' \to R$ such that $f'(x) = 0$ for $x \subseteq B$, $f'(x) > 0$ elsewhere and $f''(y) = 0$ for $x \subseteq \bar{B}$. Clearly, we can make f" satisfying (31).

Consider now a function f as in 8.5. Since

$$\sum_{a\in A} f(a) = \sum_{a\subseteq B} f(a) + \sum_{a\subseteq\bar{B}} f(a) = \sum_{x\subseteq B} f'(x) + \sum_{y\subseteq\bar{B}} f''(y),$$

this sum is zero. But, using (32) only, we get a positive value. This contradiction implies that b' and b" are independent.

("If" part). To prove cumulative independence we will show the solvability of the system of all linear equations (32) and (33) with unknowns f(a) whenever the given terms satisfy (31). Therefore, it is enough to show that there are no linear dependences among the rows of coefficients in this system, except when the sum of left-hand sides of (32) equals the corresponding sum on (33).

No linear dependence may exist among rows either in (32) only or in (33) only. In fact, in each one of these two subsystems, there is no column with two or more nonzero elements. Hence, any possible linear dependence must have the form:

$$(34) \qquad \sum_{x\in A'} u'(x)\cdot \sum_{a\subseteq x} f(a) = \sum_{y\in A''} u''(y)\cdot \sum_{a\subseteq y} f(a),$$

where u' and u" should be non-constant column vectors, $u':A' \to R$ and $u'':A'' \to R$.

Consider now the block functions b' and b" of A' and A" which are independent. We can set b' = d'b and b" = d"b where b is the block function of A. Thus, d' and d" form a representation of A which is independent as one can easily check.

From characterization 4.5 we get that d' > g and d" > g, for all $g: A \to R$, implies that g is a constant. This minimality condition will finally yield the required solvability.

By associativity, we can rewrite (34) as

$$\sum_{a \in A} u'(d'(a))f(a) = \sum_{a \in A} u''(d''(a))f(a).$$

Thus the function g = u'd' = u"d" satisfies the hypothesis of the minimality condition and it must be constant. Hence, u' and u" also must be constant. This makes (34) never fulfilled. Therefore, our system is always solvable

Q.E.D.

ACKNOWLEDGEMENTS

I wish to thank K. Culik II who helped me in many ways while thesis supervisor of [12] and afterwards, I. Fleischer for his comments on the algebraic significance of characterization 7.9, D.S. Scott for a reference in 0.3, and B. Forte who pointed out the opportuneness of overcoming the drawbacks of classical stochastic independence in certain problems of information theory and incited me to the application in 8.4-8.6.

REFERENCES

[1] M. Arbib, Theories of Abstract Automata, Prentice-Hall, Englewood Cliffs, New Jersey, 1969.

[2] G. Birkhoff, Lattice Theory, Amer. Math. Soc. Colloq. Publ., Vol.25, New York, 1967.

[3] T.L. Booth, Digital Networks and Computer Systems, J. Wiley & Sons, New York, 1971.

[4] B. Forte, Measures of information: the general axiomatic theory, R.I.R.O., 3, 1969.

[5] G. Grätzer, Universal Algebra, Van Nostrand, Princeton, New Jersey, 1968.

[6] G. Grätzer, Lattice Theory - first concepts and distributive lattices, Freeman, San Francisco, California, 1971.

[7] J. Hartmanis and R.E. Stearns, Algebraic Structure Theory of Sequential Machines, Prentice-Hall, Englewood Cliffs, New Jersey, 1966.

[8] A.G. Kurosh, The Theory of Groups, Chelsea, New York, 1955.

[9] J.D. Monk, Introduction to Set Theory, McGraw-Hill, New York, 1969.

[10] M. Phister Jr., Logical design of digital computers, John Wiley & Sons, New York, 1958.

[11] G. Ricci, Cascades of tree-automata and computations in universal algebras, Mathematical Systems Theory, Vol.7, N.3, 1973.

[12] G. Ricci, The Abaci, Ph.D. dissertation, University of Waterloo, Ontario, 1973.

[13] G. Szasz, Introduction to lattice theory, Academic Press, New York, 1963.