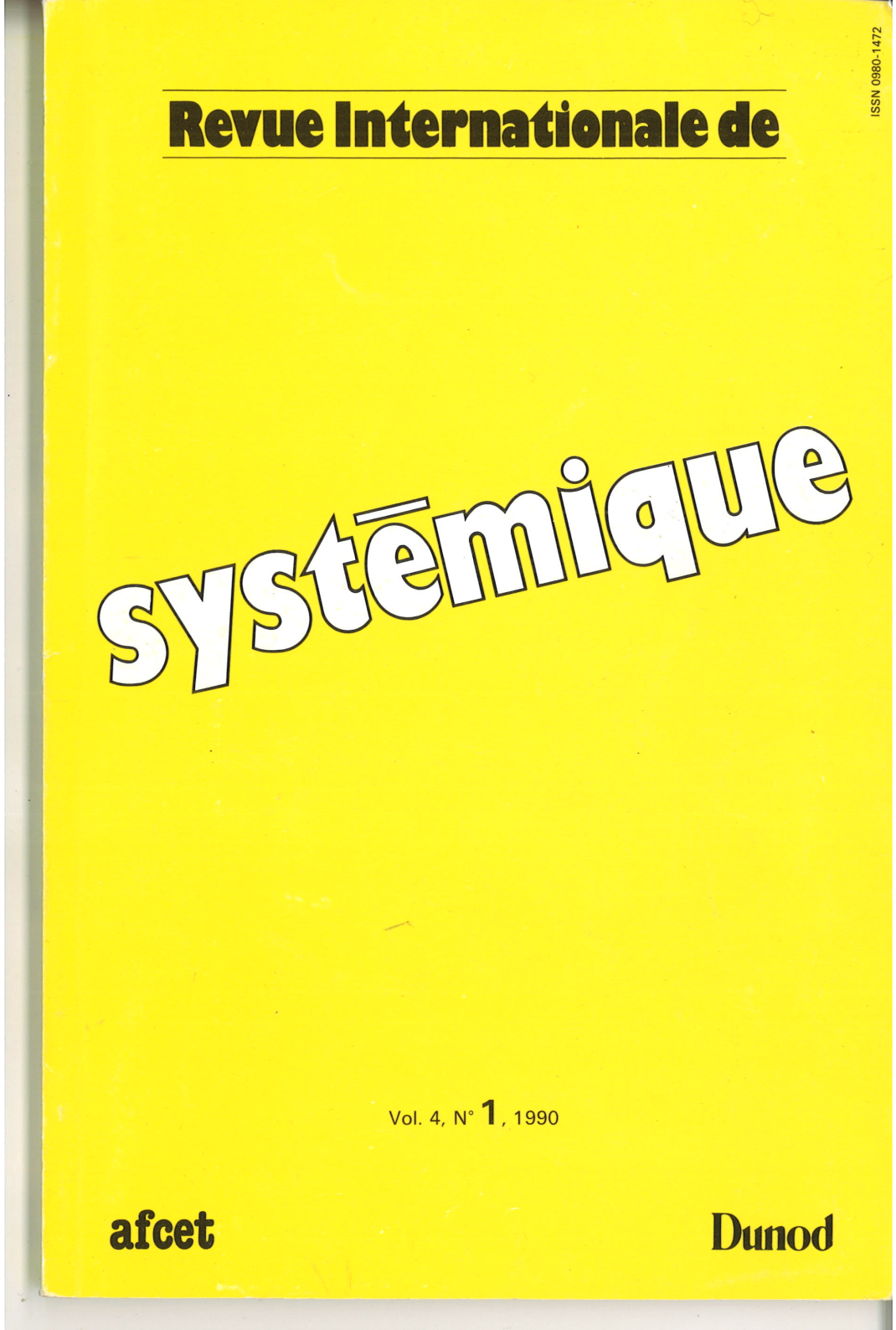

AFSCET

Revue Internationale de systēmique

Revue
Internationale
de Sytémique

volume 04, numéro 1, page 5 - 32, 1990

Reconstructability Analysis: an Overview

Michael Pittarelli

Numérisation Afscet, janvier 2016.

Creative Commons

REVUE INTERNATIONALE DE SYSTEMIQUE
Vol. 4, N° 1, 1990, pp. 5 à 32

RECONSTRUCTABILITY ANALYSIS: AN OVERVIEW

Michael PITTARELLI

State University of New York Institute of Technology [1]

Abstract
An overview of the main results and methods of reconstructability analysis is presented. Reconstructability analysis consists of techniques developed for the purpose of analyzing the degree to which a system is decomposable into a collection of subsystems and, conversely, the degree to which a system is identifiable from a collection of its subsystems. Reconstructability analysis has been successfully applied in a number of fields. The intent of this article is to stimulate both wider application of currently available techniques and research on open problems in the methodology of reconstructability analysis itself.

Résumé
Nous présentons les principaux résultats et méthodes de l'analyse de la reconstructibilité. Ce genre d'analyse fait appel à des techniques ayant comme but de trouver le degré de décomposabilité d'un système en un ensemble de sous-systèmes et, réciproquement, de trouver le degré d'identification d'un système à partir d'un ensemble de ses sous-systèmes. L'analyse de la reconstructibilité s'applique avec succès à de nombreux domaines. Nous souhaitons encourager l'application plus étendue des techniques existantes et aussi des recherches en vue de la résolution de problèmes méthodologiques non réglés.

1. Utica, New York 13504-3050, U.S.A., Department of Computer and Information Science.

Rev. intern. systémique. 0980-1472 Vol. 4/90/01/ 5 /28/$ 4.80/© Afcet Gauthier-Villars

1. Introduction

At the highest level of generality, reconstructability analysis may be said to be concerned with two complementary problems:

1. *The reconstruction problem.* – To what extent is a given system description inductively inferable from descriptions of subsystems deductively inferable from it?

2. *The identification problem.* – To what extent do the descriptions of a given collection of systems determine (deductively) the description of a single system of which they are each subsystems?

While recognizing what by now is a commonplace in science, that "the whole is greater than the sum of its parts", *i.e.*, that it is (usually) not possible to infer from detailed knowledge of subsystems similarly detailed knowledge of a system of which they are parts, practitioners of reconstructability analysis seek decompositions of systems into subsystems such that as great a simplification as possible is achieved while keeping the loss of information regarding the original system within acceptable limits. Hence the characterization by Conant (1988) of reconstructability analysis as "a form of enlightened reductionism".

The origins of reconstructability analysis can be traced to a paper by W. R. Ashby on "constraint analysis" in which a procedure is developed for determining the degree to which an *n*-ary relation is decomposable into subrelations (1964). Ashby (with Madden, 1972) later considered what is essentially the identification problem for relational systems. (Many results in *relational database theory* as well were anticipated by Ashby in these two papers.) Probabilistic constraints were first considered by Lewis (1959). Procedures utilizing a lattice of possible decompositions of a system described in terms of a set of variables were first formulated by Klir (1976) and later elaborated by Cavallo and Klir (1979, 1981*b*). Cavallo (1980) was the first to investigate the probabilistic identification problem, and Cavallo and Klir (1982*b*) extended the framework to accommodate possibilistic constraints. The identification problem for possibilistic systems was treated by Higashi *et al.* (1984).

Reconstructability analysis has been employed in a number of fields: ecological modeling, industrial engineering, computer engineering, and medical and agricultural research (Bard, 1980*a*, 1980*b*; Higashi, 1984; Klir, 1986; Shaffer and Cahoon, 1987; Uyttenhove, 1982). In what follows, a survey of some of the main results and techniques of reconstructability analysis is presented with the hope that researchers in a wide variety of disciplines will find them applicable to their work.

2. Definitions

For purposes of reconstructability analysis, a system is characterized in terms of a state space that is the Cartesian product of a set of finite variable domains. The characterization takes the form of a mapping from this Cartesian product to a set of real numbers such that the value of the mapping applied to a given state tuple indicates the degree to which that tuple, or state, is typical of the states of the system (e.g., the probability that the system is in that state).

Definition 2.1. – Formally, a system is defined as a four-tuple $F=(V, \Delta, \mathrm{dom}, f)$, where

- V is a finite non-empty set of *variables* (also referred to as *observables* or *attributes*), and V is referred to as the scheme for F;
- Δ is a non-empty set of finite sets of values called *domains*;
- $\mathrm{dom} : V \to \Delta$ is an onto function that associates a domain with each variable;
- $\mathrm{dom}(A) = \underset{v \in A}{\times} \mathrm{dom}(v)$, where $A \subseteq V$, is the set of system substates (subtuples) over variables A [dom(V) is the set of system states, or system tuples];
- for $w \in \mathrm{dom}(W)$, $b \in \mathrm{dom}(B)$, and $B \subseteq W$, $w[B]=b$ iff b and w agree on all attributes in B;
- $f : \mathrm{dom}(V) \to Q$, where $Q \subseteq \mathbf{R}$, is a mapping characterizing the degree to which a tuple $t \in \mathrm{dom}(V)$ is representative of the behavior of the system. Such functions are sometimes referred to as *behavior functions* and systems F are sometimes called *behavior systems*. [For convenience, a total ordering on dom(V) is assumed, and f is used to denote the correspondingly ordered tuple of images of the function f. Also, since f determines V, Δ, and dom, "f" is sometimes used to refer to the system F.] □

In mathematical terms, reconstructability analysis is concerned with the relation that holds between a behavior function $f : \mathrm{dom}(V) \to Q$ and a collection of *projections*

$$\pi_A(f) : \mathrm{dom}(A) \to Q,$$

where $A \subseteq V$. The value of $\pi_A(f)(a)$, for $a \in \mathrm{dom}(A)$, is a function of $\{f(t) \mid t \in \mathrm{dom}(V), a = t[A]\}$. The operation defining $\pi_A(f)(a)$ depends on the nature of f.

When f is a *probability distribution function*,

$$f : \mathrm{dom}(\mathrm{V}) \to [0, 1], \qquad \sum_{t \in \mathrm{dom}(\mathrm{V})} f(t) = 1,$$

the subtuples $a \in \mathrm{dom}(\mathrm{A})$ may be viewed as events equal to the union of the set of elementary events $\{ f(t) \mid t \in \mathrm{dom}(\mathrm{V}), a = t[\mathrm{A}] \}$. By the axioms of probability,

$$\pi_{\mathrm{A}}(f)(a) = \sum_{t \in \mathrm{dom}(\mathrm{V}),\, t[\mathrm{A}] = a} f(t).$$

When f is the characteristic function of a relation,

$$f : \mathrm{dom}(\mathrm{V}) \to \{ 0, 1 \},$$

where $f(t) = 1$ is interpretable as the statement "the system has been observed to be in state t" or "t is a possible state for the system" [note that the possibility of each substate $x \in \mathrm{dom}(v)$ for all $v \in \mathrm{V}$ does not imply the possibility of each system state $t \in \mathrm{dom}(\mathrm{V})$], and $f(t) = 0$ as its negation,

$$\pi_{\mathrm{A}}(f)(a) = \max_{t \in \mathrm{dom}(\mathrm{V}),\, t[\mathrm{A}] = a} f(t).$$

Such systems are referred to as "relational" or "set-theoretic". (Observe that a relational database instance is a collection of relational behavior systems.)

When f is a *possibility distribution*, $f : \mathrm{dom}(\mathrm{V}) \to [0, 1]$, the projection formula coincides with that for relational systems (Cavallo and Klir, 1982*b*). Since $\{ 0, 1 \} \subseteq [0, 1]$, a relational system is a special type of possibilistic system. Relational systems are sometimes referred to as "crisp possibilistic" systems and their characteristic functions as "crisp possibility distributions" (Klir and Folger, 1988).

Example 2.1. – The table below illustrates three behavior systems derived from the same frequency distribution $n : \mathrm{dom}(\mathrm{V}) \to \mathbf{N}$ – a probabilistic system (f_{pr}), a possibilistic system (f_{po}), and a relational system (f_{r}), calculated from n via the formulas

$$f_{\mathrm{pr}}(t) = n(t) / \sum_{t \in \mathrm{dom}(\mathrm{V})} n(t),$$

$$f_{\mathrm{po}}(t) = n(t) / \max_{t \in \mathrm{dom}(\mathrm{V})} n(t),$$

$$f_{\mathrm{r}}(t) = \max(1, n(t)).$$

(*See* Klir, 1989, for discussion of information-preserving transformations between different types of behavior functions.)

A	B	C	$n(t)$	$f_{\mathrm{pr}}(t)$	$f_{\mathrm{po}}(t)$	$f_{\mathrm{r}}(t)$
a_1	b_1	c_1	125	0.125	$0.3\overline{3}$	1
a_1	b_1	c_2	75	0.075	0.2	1
a_1	b_2	c_1	50	0.05	$0.1\overline{3}$	1
a_1	b_2	c_2	50	0.05	$0.1\overline{3}$	1
a_2	b_1	c_1	375	0.375	1.0	1
a_2	b_1	c_2	225	0.225	0.6	1
a_2	b_2	c_1	50	0.05	$0.1\overline{3}$	1
a_2	b_2	c_2	50	0.05	$0.1\overline{3}$	1

(As discussed by Jones [1985], arbitrary functions $f : \mathrm{dom}(\mathrm{V}) \to \mathrm{Q}$, including frequency distribution functions, may be analyzed using existing techniques of reconstructability analysis, for example, by transforming them to probability distributions. However, for arbitrary functions, the results of such analyses cannot be interpreted in the same manner as those for the standard types of behavior functions, e.g., in terms of conditional probabilistic independence relations among variables.) For the systems F above, $\mathrm{V} = \{ \mathrm{A}, \mathrm{B}, \mathrm{C} \}$, $\Delta = \{ \{ a_1, a_2 \}, \{ b_1, b_2 \}, \{ c_1, c_2 \} \}$, $\mathrm{dom}(\mathrm{A}) = \{ a_1, a_2 \}$, etc. The projections of the functions f onto the set $\{ \mathrm{B}, \mathrm{C} \} \subseteq \{ \mathrm{A}, \mathrm{B}, \mathrm{C} \} = \mathrm{V}$ are the functions $\pi_{\{\mathrm{B},\mathrm{C}\}}(f)$ below:

B	C	$\pi_{\{\mathrm{B},\mathrm{C}\}}(f_{\mathrm{pr}})(a)$	$\pi_{\{\mathrm{B},\mathrm{C}\}}(f_{\mathrm{po}})(a)$	$\pi_{\{\mathrm{B},\mathrm{C}\}}(f_{\mathrm{r}})(a)$
b_1	c_1	0.5	1.0	1
b_1	c_2	0.3	0.6	1
b_2	c_1	0.1	$0.1\overline{3}$	1
b_2	c_2	0.1	$0.1\overline{3}$	1

□

Each of the systems $\pi_{\{\mathrm{B},\mathrm{C}\}}(\mathrm{F})$ with behavior function $\pi_{\{\mathrm{B},\mathrm{C}\}}(f)$ is a *subsystem* of the associated system F.

Definition 2.2. – The system $\pi_{\mathrm{A}}(\mathrm{F}) = (\mathrm{A}, \Delta', \mathrm{dom}', f')$ is a *subsystem* of $\mathrm{F} = (\mathrm{V}, \Delta, \mathrm{dom}, f)$ if

- $\mathrm{A} \subseteq \mathrm{V}$;
- $\mathrm{dom}' : \mathrm{A} \to \Delta'$ is onto and $\mathrm{dom}'(v) = \mathrm{dom}(v)$, for each $v \in \mathrm{A}$;
- $f' = \pi_{\mathrm{A}}(f)$. □

A collection of behavior systems is a *structure system*.

Definition 2.3. – $\mathrm{S} = \{ \mathrm{F}_1, \ldots, \mathrm{F}_m \}$ is a *structure system* iff

(i) each of $F_1, \ldots, F_m$ is characterized by a function of the same type (e.g., a probability distribution function);

(ii) $v \in V_i \cap V_j$ implies $\mathrm{dom}_i(v) = \mathrm{dom}_j(v)$.

The set $X = \{V_1, \ldots, V_m\}$, where V_i is the scheme for F_i, is the *structural scheme* (or *structure*) for S. □

Example 2.2. – The table below represents a structure system S derived from the system f_{pr} of Example 2.1 as $S = \{\pi_{\{A,B\}}(f_{pr}), \pi_{\{B,C\}}(f_{pr})\}$:

A	B	$\pi_{\{A,B\}}(f_{pr})(.)$	B	C	$\pi_{\{B,C\}}(f_{pr})(.)$
a_1	b_1	0.2	b_1	c_1	0.5
a_1	b_2	0.1	b_1	c_2	0.3
a_2	b_1	0.6	b_2	c_1	0.1
a_2	b_2	0.1	b_2	c_2	0.1

Definition 2.4. – Suppose that V is a set of variables. The set $X = \{V_1, \ldots, V_m\}$ is a *model* of V if

(i) $V_1 \cup \ldots \cup V_m \subseteq V$;

(ii) $i \neq j$ implies $V_i \not\subseteq V_j$.

If the condition

(iii) $V_1 \cup \ldots \cup V_m = V$ is also met, *i.e.*, X is a cover of V, then X is a *reduced hypergraph* (Berge, 1973) over V. □

The structure $\{\{A, B\}, \{B, C\}\}$ of Example 2.2 is a model of (actually, a reduced hypergraph over) the set $\{A, B, C\}$.

A behavior function may be projected onto a model of its scheme. If $X = \{V_1, \ldots, V_m\}$ is a model of the scheme V for f, let $\pi_X(f)$ denote the set of projections of f onto the elements of X:

$$\pi_X(f) = \{\pi_{V_1}(f), \ldots, \pi_{V_m}(f)\}.$$

Notice that $\pi_X(f)$ is a structure system.

The identification and reconstruction problems may now be more precisely characterized in terms of these operations. For the identification problem, a structure system $S = \{F_1, \ldots, F_m\}$ is given, and the task is to characterize in various ways the set of behavior systems f such that $\pi_X(f) = S$: e.g., calculate a measure of its size, devise an efficient test for membership in the set, etc. (The scheme of F is usually taken to be the union of the elements of X, but need not be.) The reconstruction problem is that of identifying a structure X, or set of structures X, such that X is as simple as possible (e.g., the average cardinality of its elements is minimized) but the information loss when f is replaced by its projection onto X does not exceed some maximum

allowable value. These two complementary problems will now be separately considered in detail.

3. Identification

It is sometimes the case that data regarding some phenomenon are available not in the form of a single behavior function over the set V of variables of interest, but in the form of a structure system S. This may be the case for various reasons: S may represent the results of separate partial studies conducted independently which it is desired later to integrate; for some system it is technically not feasible to observe simultaneously behavior over all of the variables in terms of which it is defined, etc.

The identification problem encompasses all aspects of inference from a structure system S with structure X to behavior functions f such that $S = \pi_X(f)$. In extreme cases, there is a unique f such that $S = \pi_X(f)$. The function f is then said to be *identifiable* from X (or S). This situation is extremely rare (Madden and Ashby, 1972; Pittarelli, 1989*a*). There is, in general, a *set* of functions f such that $\pi_X(f) = S$.

Definition 3.1. – The *extension* of a structure system $S = \{f_1, \ldots, f_m\}$ over V, where the structure $X = \{V_1, \ldots, V_m\}$ of S is a model of V, is the set

$$E_V(S) = \{f \mid \pi_X(f) = S \text{ and } f : \mathrm{dom}(V) \to Q\}.$$

When X is a cover of V, $E_V(S)$ may be abbreviated E(S), and is referred to as the *reconstruction family* of S.

When a structure system S is derived from projection of a given behavior function onto a model of its scheme, it is guaranteed that the reconstruction family of S is non-empty: clearly,

$$f \in E_V(\pi_X(f)).$$

This is not guaranteed when a structure system represents independently derived behavior functions which it is desired to combine. There are many reasons why this may be so: the behavior functions may represent the opinions of different experts, or the same expert in multiple contexts (where the different elements of a structure system represent different contexts); measuring instruments (e.g., counters) vary in sensitivity, etc. Thus, empirically derived structure systems are likely to be inconsistent.

Two types of inconsistency are distinguished.

Definition 3.2. – A system S with structure $X=\{V_1, \ldots, V_m\}$ is *locally inconsistent* iff

$$\pi_{V_i \cap V_j}(f_i) \neq \pi_{V_i \cap V_j}(f_j),$$

for some pair (i, j) for which $V_i \cap V_j \neq \varnothing$.

(A structure system is locally consistent iff it is not locally inconsistent.)

A structure system S is *globally inconsistent* iff $E(S)=\varnothing$. □

Global consistency implies local consistency, but not conversely.

Approaches to inconsistency resolution are explored by Mariano (1987). They are roughly characterized as follows. Given an inconsistent structure system $S=\{f_1, \ldots, f_m\}$, find behavior functions $f'_1, \ldots, f'_m$ such that

$$E(\{f'_1, \ldots, f'_m\}) \neq \varnothing$$

and

$$\sum_{i=1}^{m} d(f'_i, f_i)$$

or

$$\max_{i \in \{1, \ldots, m\}} d(f'_i, f_i)$$

is minimized, for some suitable measure d of dissimilarity between behavior functions. For large systems, such procedures are expensive. An alternative approach would be to replace the single-number values of (possibilistic and probabilistic) behavior functions with intervals. Details of this method are beyond the scope of this paper, and are discussed in (Pittarelli, 1989 *b*).

The degree to which an arbitrary $f \in E(S)$ is identified by knowledge of S may be characterized in various ways. At the simplest extreme, it is possible to determine whether or not f is strictly identifiable from S. It would be well at this point to examine the mathematical structure of reconstruction families E(S) for each of the three types of behavior functions under discussion: probabilistic, possibilistic, and relational.

When S consists of probability distribution functions, E(S) is the convex polyhedron of solutions to a system of linear equations and inequalities. This follows from the nature of the projection operation. For the structure system $S=\pi_{\{\{A, B\}, \{B, C\}\}}(f_{pr})$ of Example 2.2, E(S) is the set of solutions p to the

system

$$\begin{aligned}
p(a_1 b_1 c_1)+p(a_1 b_1 c_2)&=0.2\\
p(a_1 b_2 c_1)+p(a_1 b_2 c_2)&=0.1\\
p(a_2 b_1 c_1)+p(a_2 b_1 c_2)&=0.6\\
p(a_2 b_2 c_1)+p(a_2 b_2 c_2)&=0.1\\
p(a_1 b_1 c_1)+p(a_2 b_1 c_1)&=0.5\\
p(a_1 b_1 c_2)+p(a_2 b_1 c_2)&=0.3\\
p(a_1 b_2 c_1)+p(a_2 b_2 c_1)&=0.1\\
p(a_1 b_2 c_2)+p(a_2 b_2 c_2)&=0.1\\
p(.) &\geqq 0.
\end{aligned}$$

(Observe that any such system of equations implies that $\sum_{t \in \mathrm{dom}(V)} p(t)=1.$) □

Each (sub) state $a \in \mathrm{dom}(J)$ in the domain of a (sub) distribution in S with scheme J contributes exactly one equation to the system, and that equation constrains each of the tuples $t \in \mathrm{dom}(V)$ such that $t[J]=a$, according to the definition of projection.

The equations may be used as a test for whether a distribution p (elicited from an expert, for example) is in fact a member of E(S). If any of the equations is violated, p is disqualified. Such an algorithm could be an aid in the assessment of joint probability distributions from marginal probabilities (Moskowitz and Wallenius, 1984).

Via linear programming, bounds on components $p(t)$ as p ranges over E(S) can be determined. To illustrate, the bounds for tuples $a_2 b_1 c_1$ and $a_2 b_1 c_2$ are:

$$0.3 \leqq p(a_2 b_1 c_1) \leqq 0.5$$

and

$$0.1 \leqq p(a_2 b_1 c_1) \leqq 0.3.$$

These two tuples cannot achieve either their maximum or their minimum values simultaneously without violating the third equation,

$$p(a_2 b_1 c_1)+p(a_2 b_1 c_2)=0.6.$$

E(S) is always a subset and usually a proper subset of the set of distributions satisfying each of the bounds. Thus, a test utilizing them, while more efficient, provides merely a necessary condition for membership in E(S).

It is also possible using linear programming techniques to enumerate all of the vertices of E(S) (Pittarelli, 1989*a*). Any $p \in E(S)$ is a convex combination of these distributions. Let $e_1, \ldots, e_k$ denote the (finite) set of vertices of E(S). Then

$$p \in E(S) \text{ implies } p = \lambda_1 e_1 + \ldots + \lambda_k e_k,$$

for some $\lambda_1, \ldots, \lambda_n \geqq 0$ such that $\sum_{i=1}^{k} \lambda_i = 1$. However, it is possible for structure systems involving only a few variables, with small domains, to have hundreds of vertices.

Whereas probabilistic reconstruction families are sets of solutions to systems of equations of the form

$$\sum_{t \in \text{dom}(V),\, t[V_i]=a} f(t) = f_i(a),$$

the equations defining a possibilistic reconstruction family are of the form

$$\max_{t \in \text{dom}(V),\, t[V_i]=a} f(t) = f_i(a).$$

For the data of Example 2.1, $E(\pi_{\{\{A,B\},\{B,C\}\}}(f_{po}))$ is the set of solutions to

$$\max\{f(a_1 b_1 c_1), f(a_1 b_1 c_2)\} = 0.3\overline{3}$$
$$\max\{f(a_1 b_2 c_1), f(a_1 b_2 c_2)\} = 0.1\overline{3}$$
$$\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots$$
$$\max\{f(a_1 b_2 c_2), f(a_2 b_2 c_2)\} = 0.1\overline{3}$$
$$f(.) \geqq 0.$$

In Higashi *et al.* (1984), it is shown that such systems are a special type of fuzzy relation equation, and methods for characterizing solution sets of such equations (Higashi and Klir, 1984) are applied to the possibilistic identification problem. Define a partial ordering $\precsim$ on possibility distributions as

$$f \precsim f' \quad \text{iff} \quad f(t) \leqq f'(t), \quad \text{for all } t \in \text{dom}(V).$$

A possibilistic reconstruction family has a unique maximum element, $\overline{f}$, under this ordering,

$$f \in E(S) \text{ implies } f \leqq \overline{f},$$

and a finite number of minimal elements, $\underline{f}$:

$$\text{for no } f \neq \underline{f} \text{ is } f \precsim \underline{f}.$$

Any component $\overline{f}(t)$ of the maximum element is quickly computed as the minimum of the right-hand-sides of all the equations constraining t. For the example, $\overline{f}$ is the distribution

A	B	C	$\overline{f}(t)$
a_1	b_1	c_1	$0.3\overline{3}$
a_1	b_1	c_2	$0.3\overline{3}$
a_1	b_2	c_1	$0.1\overline{3}$
a_1	b_2	c_2	$0.1\overline{3}$
a_2	b_1	c_1	1.0
a_2	b_1	c_2	0.6
a_2	b_2	c_1	$0.1\overline{3}$
a_2	b_2	c_2	$0.1\overline{3}$

Notice that f_{po} of Example 2.1 is (necessarily) an element of E(S), and $f_{po} \precsim \overline{f}$. Let $\underline{E}(S)$ denote the set of minimal elements of E(S). In (Higashi *et al.*, 1984) it is proved that $f \in E(S)$ iff

$$\underline{f}_i \precsim f \precsim \overline{f}, \text{ for some } \underline{f}_i \in \underline{E}(S).$$

(However, just as generation of all the vertices of a probabilistic reconstruction family is extremely expensive, it is also computationally difficult to enumerate $\underline{E}(S)$.)

Since a relational system represents a special type of possibilistic system, possibilistic techniques may be used to characterize relational reconstruction families. Relational behavior functions take values in the finite set $\{0, 1\}$ and therefore relational reconstruction families are always finite. Cavallo and Klir (1981*b*) first studied the algebraic structure of relational E(S), characterizing an E(S) in terms of a unique maximum element and a finite set of "irredundant" (*i.e*, minimal) elements under the partial ordering $\subseteqq$, the ordinary subset relation between sets of tuples. Clearly, the representation in terms of the partially ordered set $(\mathbf{P}(\text{dom}(V)), \subseteqq)$ is isomorphic to that in terms of $(\{f | f : \text{dom}(V) \to \{0, 1\}\}, \precsim)$.

Measures of the degree to which a single $f \in E(S)$ is identifiable from S have been referred to as identifiability quotients and measures of reconstruction

uncertainty (Cavallo and Klir, 1981 *b*; Pittarelli, 1989 *a*). As the size of an E(S) increases, the value of an identifiability quotient decreases and the value of a measure of reconstruction uncertainty (from which identifiability quotient values are calculated via a particular transformation) increases. Because measures of reconstruction uncertainty are primary, the discussion will be restricted to them.

For any measure, m, of reconstruction uncertainty, a reasonable requirement is that

$$E(S) \subseteq E(S') \text{ implies } m(E(S)) \leqq m(E(S')).$$

This, in fact, is the only formal restriction placed on such measures. The choice of a particular measure (or measures) depends on a number of factors: whether a single $f \in E(S)$ is actually to be estimated, or the entire set E(S) is to be utilized instead (Pittarelli, 1989 *a*, 1990); whether or not interest is concentrated on only a few of the tuples $t \in \text{dom}(V)$, etc.

For relational systems, Cavallo and Klir (1981 *b*) proposed the logarithm of the cardinality of E(S). For possibilistic systems, Higashi *et al.* (1984) essentially use the measure

$$m(E(S)) = \prod_t (1 + \max_{f \in E(S)} f(t) - \min_{f \in E(S)} f(t)).$$

Both of these clearly satisfy the subset monotonicity requirement.

Most of the work done on measures of reconstruction uncertainty has concentrated on probabilistic systems (Klir, 1986; Pittarelli, 1989 *a*).

When a single distribution $p^* \in E(S)$ is estimated, it would be useful to have an indication of the degree to which the estimate may be in error. Let R(E(S)) denote the *estimation radius*,

$$R(E(S)) = \max_{p \in E(S)} d(p, p^*),$$

where d is a suitable distance measure. When d is convex (Euclidean distance, cross-entropy, etc.), there exist mathematical programming methods for calculating R. Similarly, a *diameter* may be calculated as

$$D(E(S)) = \max_{p,\, p' \in E(S)} d(p, p').$$

Clearly, $R(E(S)) \leqq D(E(S))$ and (letting $0/0 = 0$) the ratio R/D takes values between 0 and 1, where larger values represent greater eccentricity of the estimate p^*. (Note that such analyses presume that the $p \in E(S)$ are each

equally likely to be the unknown distribution of which the elements of S are projections.)

Measures concentrating on individual tuples $t \in \text{dom}(V)$ are also useful. For $Q \subseteq \text{dom}(V)$ let

$$r_Q(E(S)) = 1/|Q| \sum_{t \in Q} (\max_{p \in E(S)} p(t) - \min_{p \in E(S)} p(t)),$$

which may be calculated by linear programming. When $Q = \text{dom}(V)$, this quantity is the average uncertainty per tuple for $p \in E(S)$ (Dalkey, 1985; Klir, 1986; Pittarelli, 1989 *a*) and is abbreviated $r(E(S))$. (Variations on this measure and additional measures are discussed in Pittarelli, 1989 *a*.)

Example 3.1. – Let $X = \{\{A\}, \{B, C\}\}$ and $Y = \{\{A, B\}, \{B, C\}\}$. For the data of Example 2.1, $r(E(\pi_X(f_{pr}))) = 0.25$ and $r(E(\pi_Y(f_{pr}))) = 0.1875$. □

This illustrates a situation that holds more generally.

Definition 3.3 (Cavallo and Klir, 1979). – A structure X is a refinement of structure Y iff for every $V_x \in X$ there exists a $V_y \in Y$ such that $V_x \subseteq V_y$. □

Theorem 3.1 (Higashi, 1984; Pittarelli, 1990). – $X \leqq Y$ *implies* $E(\pi_Y(f)) \subseteq E(\pi_X(f))$. □

Corollary 3.1. – *If m is a measure of reconstruction uncertainty and* $X \leqq Y$, *then* $m(E(\pi_Y(f))) \leqq m(E(\pi_X(f)))$. □

Thus, if one is performing an investigation of a system defined in terms of variables V which it is impossible to measure simultaneously, one should attempt, to the degree to which it is feasible, to obtain data over as unrefined a structure as possible.

When a reconstruction family E(S) contains more than one element, the practice has been to select a single $f \in E(S)$ as a best estimate of the actual (but unknown) behavior function of which the elements of S are projections.

For relational systems, the unique maximum element with respect to the subset relation is selected (Cavallo and Klir, 1981 *b*). This relation contains all of the overall tuples t such that for each element F_i of S the projection $t[V_i]$ is present. Elimination of any of these tuples would amount to the introduction of a constraint among the $t \in \text{dom}(V)$ that is not implied by S.

In the reconstructability analysis literature, estimates that do not assume any constraint beyond that implied by the available data are referred to as *unbiased reconstructions*, but not in the statistical sense of the term "unbiased". Any unbiased reconstruction uniquely maximizes within E(S) a measure of uncertainty appropriate to behavior functions of its type. For relational

systems, the unbiased reconstruction maximizes the *Hartley information* (Klir and Folger, 1988):

$$\mathrm{I}(f)=\log_2 \sum_t f(t).$$

The unbiased relational reconstruction r^* may be calculated in various ways (Cavallo and Klir, 1981 *b*). For $\mathrm{S}=\{r_1, \ldots, r_m\}$,

$$r^*(t)=\min_i \{r_i(a) \mid a=t[\mathrm{V}_i]\}.$$

For possibilistic systems, the unique maximum element with respect to the $\gtrsim$ relation, $\bar{f}$, is selected. It is straightforward to show that $\bar{f}$ uniquely maximizes Higashi and Klir's *U-uncertainty* measure for possibility distributions (Higashi and Klir, 1983 *a*):

$$\mathrm{U}(f)=1/l_f \sum_{k=1}^{r-1} (l_{k+1}-l_k)\log_2 |c(f, l_{k+1})|,$$

where $l_1=0, l_2, \ldots, l_r=l_f$ are the distinct values taken by any $f(t)$, augmented if necessary by $l_1=0$, in increasing order by subscript (l_f denotes the largest possibility value, l_r), and $c(f, l_j)$ is the set of all values $f(t)$ such that $f(t)\geqq l_j$. For $\mathrm{S}=\{f_1, \ldots, f_m\}$, $\bar{f}$ may be calculated as

$$\bar{f}(t)=\min_i \{f_i(a) \mid a=t[\mathrm{V}_i]\},$$

which coincides with the relational formula. Further, the U-uncertainty measure reduces to Hartley information in the "crisp possibilistic" case $f(t)\in\{0, 1\}$.

The situation is more complicated for probabilistic systems. In practice, the *maximum entropy element* p^* of $\mathrm{E}(\mathrm{S})$ is selected: $\mathrm{H}(p^*)=\max_{p\in \mathrm{E}(\mathrm{S})} \mathrm{H}(p)$, where H is the (Shannon) entropy measure

$$\mathrm{H}(p)=-\sum_t p(t)\log_2 p(t).$$

Since $\mathrm{E}(\mathrm{S})$ is convex, this element is unique (Jaynes, 1984). The fact that the constraints subject to which entropy is to be maximized are marginal probabilities (projections) permits the use of "probabilistic join" procedures for calculating p^* that, although computationally quite expensive, are more efficient than the optimization techniques required to maximize entropy subject to arbitrary (convex or even linear) constraints. Such procedures were

first studied by Deming and Stephan (1940), Brown (1959) and Lewis (1959). They are first discussed in the context of reconstructability analysis by Cavallo and Klir (1981 *a*). Recent algorithmic advances have been made by Tian (1988).

In the simplest case, S consists of two distributions: $\mathrm{S}=\{p_1, p_2\}$. The *pairwise join* procedure is then defined as

$$p^*(t)=p_1(a)\times p_2(b),$$

where

$$a=t[\mathrm{V}_1] \quad \text{and} \quad b=t[\mathrm{V}_2], \quad \text{if} \quad \mathrm{V}_2\cap\mathrm{V}_1=\varnothing;$$

and

$$p^*(t)=p_1(a)\times p_2(b)/\sum_{c\in \mathrm{dom}(\mathrm{V}_2),\, c[\mathrm{V}_2\cap\mathrm{V}_1]=b[\mathrm{V}_2\cap\mathrm{V}_1]} p_2(c),$$

otherwise.

When the structure $\mathrm{X}=\{\mathrm{V}_1, \ldots, \mathrm{V}_m\}$ of S is *loopless* (*i.e.*, α-acyclic) (Fagin, 1983), an ordering of elements of S can be found such that p^* may be calculated by repeated application of the pairwise join procedure (Krippendorff, 1986; Tian, 1988). A simple test for looplessness is provided by the algorithm below.

ALGORITHM 3.1 (Fagin, 1983). – Apply repeatedly in any order until neither has any effect:

(1) If v appears in only V_i, delete v from V_i.

(2) If $\mathrm{V}_i\subseteqq\mathrm{V}_j$, where $i\neq j$, eliminate V_i. □

The algorithm terminates in the structure $\{\varnothing\}$ if and only if the original (input) structure is loopless. A more expensive iterative procedure must be used for structures that are not loopless (Cavallo and Klir, 1981 *b*). This procedure is also used when only some, not all, of the substate probabilities are given (Jones, 1984). (The iterative procedure still is less expensive than maximizing entropy by mathematical programming.)

Example 3.2. – For the structure system of Example 2.2, the maximum entropy element of E(S) is

A	B	C	$p^*(t)$	
a_1	b_1	c_1	0.125	$=0.2\times0.5/(0.5+0.3)$
a_1	b_1	c_2	0.075	$=0.2\times0.3/(0.5+0.3)$
a_1	b_2	c_1	0.05	$=0.1+0.1/(0.1+0.1)$
a_1	b_2	c_2	0.05	$=0.1\times0.1/(0.1+0.1)$
a_2	b_1	c_1	0.375	$=0.6\times0.5/(0.5+0.3)$
a_2	b_1	c_2	0.225	$=0.6\times0.3/(0.5+0.3)$
a_2	b_2	c_1	0.05	$=0.1\times0.1/(0.1+0.1)$
a_2	b_2	c_2	0.05	$=0.1\times0.1/(0.1+0.1)$

Estimation by means of the various probabilistic join procedures is usually justified by appeal to E. T. Jaynes' *maximum entropy principle:* "When we make inferences based on incomplete information, we should draw them from that probability distribution that has the maximum entropy permitted by the information that we do have (Jaynes, 1982)." With information quantified as negative entropy, selection of p^* amounts to assuming no information beyond what is implied by the given marginal probabilities. Thus, p^* is an "unbiased" estimate of $p\in\mathrm{E(S)}$.

Additional arguments in favor of selection of the maximum entropy element have been made: it is in some sense a *maximum likelihood* estimate (Jaynes, 1979, 1982); failure to maximize entropy in making an estimate violates certain axioms of inductive inference (Shore and Johnson, 1980); the behavior of any man-made probabilistic system coincides with the probabilistic join of its subsystems (Cavallo and Klir, 1981 *b*). These and the maximum entropy principle generally are discussed critically in (Pittarelli, 1989 *a*).

Alternatives to maximum entropy estimation have been proposed and investigated theoretically and empirically (Pittarelli, 1989 *a*), but have not been used in practice. Cavallo and Klir (1982 *a*) proposed what has become known as a "least-risk" or "minimax" estimate: for some distance measure d, select p^* such that

$$\min_{p\in\mathrm{E(S)}}\ \max_{p'\in\mathrm{E(S)}} d(p,p') = \max_{p'\in\mathrm{E(S)}} d(p^*,p').$$

Thus, p^* minimizes the maximum error, measured by d, when it is selected as the distribution p such that $\mathrm{S}=\pi_{\mathrm{X}}(p)$. For certain distance measures, techniques from location theory may be used to calculate such estimates. The *centroid* (center of mass) of E(S) is also a reasonable estimate. This distribution minimizes the expected value of the squared error of estimation $\sum_t (p^*(t)-p(t))^2$. Experimental studies indicate that the maximum entropy element tends to be very close to the centroid of E(S) (Pittarelli, 1989 *a*).

4. Reconstruction

Most of the interest in reconstructability analysis has been generated by the reconstruction problem, and it is this subproblem that is algorithmically the better developed.

For the reconstruction problem, a behavior function f over dom(V) is given and a structure X over V is sought such that the complexity of X is minimized while minimizing $d(f, \mathrm{J}(\pi_{\mathrm{X}}(f)))$, where J is the join operation producing the maximum (Hartley, U-, or Shannon) uncertainty element of the structure system $\pi_{\mathrm{X}}(f)$ and d is a measure of *information distance*, *i.e.*, a measure of the amount of information that is lost when f is replaced by the join of its projections onto X, which is the maximum loss of information possible when f is replaced by $f'\in\mathrm{E(S)}$. In this context, a structure X is referred to as a *reconstruction hypothesis*. The relative validity of two competing reconstruction hypotheses X and Y is determined by comparing the information distance values $d(f, \mathrm{J}(\pi_{\mathrm{X}}(f)))$ and $d(f, \mathrm{J}(\pi_{\mathrm{Y}}(f)))$. Structures may be ordered by complexity in a number of ways. The refinement ordering gives a particularly useful and quite natural complexity ordering: X is more complex than Y if $\mathrm{Y}\leqq\mathrm{X}$. Not all structures over a set of variables V are comparable under the partial ordering $\leqq$. However, it will be seen that the refinement ordering has properties relative to measures of information loss that facilitate the search for acceptable structures.

Following Gaines (1977), Klir (1985) proposes that a solution set for the reconstruction problem be defined in terms of a *joint preference ordering* $\leqq^*$, *where*

$$\mathrm{X}\leqq^*\mathrm{Y} \quad \text{iff} \quad \mathrm{X}\leqq\mathrm{Y} \quad \text{and} \quad d(f, \mathrm{J}(\pi_{\mathrm{X}}(f)))\leqq d(f, \mathrm{J}(\pi_{\mathrm{Y}}(f))).$$

For a given set of reconstruction hypotheses M, the solution set is then the set of minimal elements of M with respect to $\leqq^*$, which are referred to as *admissible solutions*.

Ideally, for a given structure X, no information is lost when f is replaced by its projections onto X. In this case, f is said to be *reconstructable from* X.

Definition 4.1. – f is reconstructable from X iff $f=\mathrm{J}(\pi_{\mathrm{X}}(f))$. □

When a behavior function f is reconstructable from a structure X, X is usually interpreted as indicating something of the pattern of interaction among the variables V. For example, in the case of probabilistic behavior

functions, it is possible to show for loopless structures X that for any pair of elements V_i, $V_j \in X$ such that $V_i \cap V_j \neq \varnothing$, the variables V_i are conditionally independent of the variables V_j, given the variables $V_i \cap V_j$. Further, for any type of structure, loopless or not, Higashi (1984) has shown that if the set of variables of actual interest is some subset $Q \subseteq V$, then if p is reconstructable from X, attention may be restricted just to the set of variables in the neighborhood $n(Q, X)$ of Q, where

$$n(Q, X) = \cup \{ V_i \in X \mid V_i \cap Q \neq \varnothing \}.$$

When $p = J(\pi_X(p))$,

$$H(Q \mid V - Q) = H(Q \mid n(Q, X) - Q),$$

where $H(V_i \mid V_j)$ is the conditional entropy of V_i, given V_j (Klir and Folger, 1988). As yet, no studies have been conducted to determine the degree to which such relations (of independence or irrelevance) hold when p is only *approximately* reconstructable: $0 < d(p, J(\pi_X(p))) < \varepsilon$, for some small ε. Since it is usually the case that some information is lost when a distribution is replaced by any of its projections, investigation of this issue should receive high priority among topics for future research.

When f is relational, it seems reasonable to quantify information loss in terms of the number of tuples on which functions f and $J(\pi_X(f))$ disagree, *i.e.*, as the ordinary Hamming distance:

$$d(f, J(\pi_X(f))) = \sum_t |f(t) - J(\pi_X(f))(t)|.$$

Note that, since $f \gtrsim J(\pi_X(f))$, this may be reexpressed as

$$d(f, J(\pi_X(f))) = \sum_t J(\pi_X(f))(t) - \sum_t f(t).$$

For possibilistic behavior functions, the general measure of information distance derived by Higashi and Klir (1983*b*) degenerates to

$$d(f, J(\pi_X(f))) = U(J(\pi_X(f))) - U(f).$$

Information distance takes a similarly simple form for probabilistic functions. A measure of distance between probability distributions that is theoretically very well justified is the cross-entropy (directed divergence, relative entropy):

$$d(f, f') = \sum_t f(t) \log_2 f(t)/f'(t).$$

This function is not defined for arbitrary distributions f, f' over V unless (using the convention $0 \log_2 0/0 = 0$) $f'(t) = 0$ implies $f(t) = 0$. This condition holds, however, when $f' = J(\pi_X(f))$ (Higashi, 1984). Higashi also proves that

$$d(f, J(\pi_X(f))) = H(J(\pi_X(f))) - H(f).$$

Thus, to compare a set $X_1, \ldots, X_k$ of reconstruction hypotheses for behavior functions of any of these three types, it is not necessary actually to calculate the information distance. It is sufficient merely to calculate and compare the values $\sum_t J(\pi_X(f))$, $U(J(\pi_X(f)))$ or $H(J(\pi_X(f)))$.

Each of these information distance measures is monotonic with respect to the refinement ordering:

$$X \leqq Y \quad \text{implies} \quad d(f, J(\pi_Y(f))) \leqq d(f, J(\pi_X(f))).$$

In other words, the information loss never decreases when a structure is replaced by a more refined structure. Not all measures have this intuitively reasonable property (for example, the Hamming distance when applied to probabilistic behavior functions). It follows from Theorem 3.1 that any function d that reduces to $h(J(\pi_X(f))) - h(f)$, where $h(J(\pi_X(f))) = \max_{f' \in E(S)} h(f')$, will satisfy this requirement.

Without question the most serious challenge facing researchers in reconstructability analysis is the development of algorithms for reconstruction hypothesis search with acceptable performance characteristics when applied to large systems. Even for systems with only a few variables, exhaustive comparison of all hypotheses is out of the question: for $|V| = 7$, there are already over 2×10^{12} different structures covering V. Most of the strategies in use at this time are heuristic; they are not guaranteed to generate all and only members of a particular solution set (say, with $M = \{ X : \max_{V_x \in X} |V_x| \leqq k \}$), but have been shown empirically to perform very well (Conant, 1988; Hai and Klir, 1985; Klir and Parviz, 1986; Klir and Uyttenhove, 1977).

Certain of the more exhaustive procedures, for example, those proposed by Cavallo and Klir (1979), consist of subprocedures that may be classified as either global or local and as involving either aggregation or refinement. Usually, the overall process is one of refinement, involving some stopping criterion: an absolute value of information loss that is not to be exceeded; a value expressed as a percentage of the maximum possible loss, $d(f, J(\pi_X(f)))$, where X is the most refined structure in the lattice; detection of a sharp

increase in information loss from one "level of refinement" to the next, etc. In the case of a threshold criterion, the monotonicity property (Theorem 3.1) guarantees that further refinement will fail to locate acceptable structures. Stopping criteria based on differences in d from one step to the next seem to be very well justified by the results of experimental studies of their accuracy (Hai and Klir, 1985; Klir and Parviz, 1986).

Before looking at these procedures in any greater detail, it is necessary to discuss briefly the algebraic structure of the set of all reconstruction hypotheses over a set of variables V. (To simplify the discussion slightly, attention will be restricted to covers of V.)

Let G denote the set of all structures over a set V. Then the pair $(G, \leqq)$, where $\leqq$ is the refinement relation, is a lattice. Each pair of structures has a least upper bound (least common aggregate) and a greatest lower bound (greatest common refinement), and there is a universal upper bound, $\{V\}$, and a universal lower bound, $\{\{v\} \mid v \in V\}$.

Figure 1 is a Hasse diagram of the lattice of structures over $V = \{v_1, v_2, v_3\}$. The rectangles in a block diagram depicting a structure X represent elements $V_i \in X$. Each line in a block diagram represents a single variable $v_j \in V$. The line for v_j intersects the rectangle for V_i iff $v_j \in V_i$. An arrow connecting the block diagrams for a pair of structures X and Y, where X's diagram is above Y's, indicates that Y is an *immediate refinement* of X (and, conversely, that X is an *immediate coarsening*, or *aggregate*, of Y):

$$Y \leqq X \text{ and there is no Z distinct from X and Y such that } Y \leqq Z \leqq X.$$

Associated with any structure X in G is an undirected graph defined by the binary relation on V:

$$v_i \, R \, v_j \text{ iff there exists a } V_k \in X \text{ such that } v_i, v_j \in V_k.$$

It is possible for two distinct structures to have identical associated graphs. For example, $\{\{v_1, v_2, v_3\}\}$ and $\{\{v_1, v_2\}, \{v_2, v_3\}, \{v_1, v_3\}\}$ are both associated with the complete graph on $\{v_1, v_2, v_3\}$. For any equivalence class of structures with the same associated graph, there is a unique C-*structure* defined as the set of all and only the maximal compatibility classes of the graph. (Note: the problem of generating the C-structure associated with a given graph is NP-hard [Garey and Johnson, 1979].) Figure 2 illustrates two equivalent structures, one of which is the C-structure for the equivalence class of which they are both members, together with their associated graph (self-loops are not drawn).

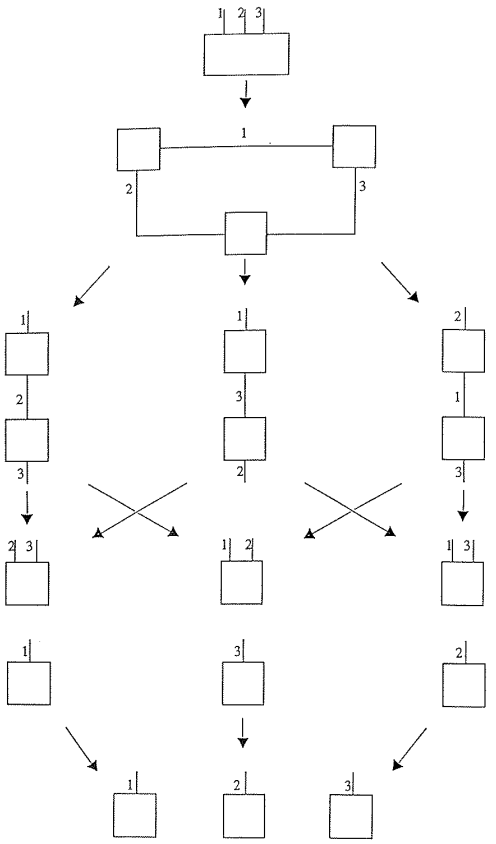

Figure 1.

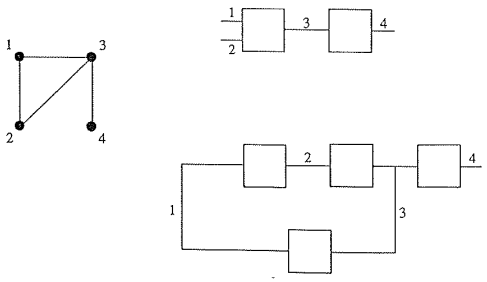

Figure 2.

Let C denote the set of all C-structures over V. Then the pair $(C, \leqq)$ is also a lattice, a sublattice of $(G, \leqq)$. Since $|C| < |G|$, for $|V| > 2$, it might appear reasonable to search the set of C-structures exhaustively and then, for each member of the solution set, explore its *G-structure neighborhood*, the set of G structures that are (immediate or nearly immediate) refinements or aggregates of it. This is not feasible either. The size of C grows exponentially with the size of V, as $|C| = 2^{(|V|^2 - |V|)/2}$. For $|V| = 10$, there are over 3.5×10^{13} C-structures.

Instead, a *heuristic* global search of $(C, \leqq)$ is conducted (by refinement, although global aggregation is also feasible), followed by local G-structure refinement and aggregation. The chief virtue of the C-structure lattice is the (conceptual) simplicity with which refinements may be generated: the (C-structure) refinements of a given C-structure X with associated graph g are those C-structures associated with each of the graphs resulting from removing a single edge from g. Since the maximum number of edges (not counting self-loops, which are not removed) in such a graph is $|V| \times (|V| - 1)/2$, there are this many *levels of refinement* in the C-structure lattice for a set of variables V.

The breadth of such a heuristic search may be either wide or narrow. At the narrowest extreme, only the best structure at any level with respect to information loss is chosen to refine further. The problem with this option, efficient though it may be, is that it is possible that even though structure X is superior to structure Y at the same level of refinement, there may exist an immediate C-structure refinement of Y that is superior to all of the immediate C-structure refinements of X (Cavallo, 1989). At the opposite extreme, all structures at any level are refined, which amounts to exhaustive search of $(C, \leqq)$. In practice, a compromise is made, and the best k structures at any level, or those with associated information loss within a certain range of that associated with the best structure, are refined. [It should be noted that only very recently have algorithms been devised for refining a set of k structures at a given level without generating duplicate structures at the next level (Cavallo and De Voy, 1989). Previously, use of a costly comparison and removal procedure was also required.]

Refinement terminates when some stopping criterion is met, and may be followed by local G-structure refinement and aggregation procedures. Refinement and aggregation procedures are discussed in considerable detail by Cavallo and Klir (1979, 1982*a*) and in forthcoming papers by Cavallo (1989) and Cavallo and De Voy (1989).

More radical heuristics have been proposed. For example, a procedure invented by Conant (1988) for probabilistic systems may be roughly described

as follows. For each variable $v \in V$, a dependency set $D(v) \subseteqq V - \{v\}$ is determined such that, where T is the *transmission* measure

$$T(v : A) = \sum_{x \in \mathrm{dom}\,(v),\ t \in \mathrm{dom}\,(A)} \pi_{\{v\} \cup A}(p)(x, t) \log_2 \pi_{\{v\} \cup A}(p)(x, t)/\pi_{\{v\}}(p)(x)\,\pi_A(p)(t),$$

(i) $Q \subseteqq D(v)$ implies $T(v : Q)$ is significantly smaller than $T(v : D(v))$;

(ii) $D(v) \subseteqq Q$ implies $T(v : Q)$ is not significantly greater than $T(v : D(v))$.

A graph $g = (V, E)$ with edge set E is constructed from the set $Z = \{\{v\} \cup D(v) \mid v \in V\}$ as $(v_i, v_j) \in E$ iff there exists an element $W \in Z$ such that $v_i, v_j \in W$. The C-structure associated with this graph is posited as a reconstruction hypothesis, the structure neighborhood of which may then be searched.

Frequently, only a very restricted subclass of the set of all G-structures may be recognized as meaningful. For example, it may be desired to search only for structures $X = \{V_1, \ldots, V_m\}$ that are partitions, *i.e.*, such that $i \neq j$ implies $V_i \cap V_j = \varnothing$. A very simple heuristic for such structures (illustrating, incidentally, a global aggregation procedure) is given below:

ALGORITHM 4.1 (Let $|V| = n$)

```
j := n;
X := {{v} | v ∈ V};
while j > 1 and the stopping criterion is not met do
 begin
  for i := 1 to j × (j − 1)/2 do
   begin
    construct W_i by merging two as yet unmerged elements V_i, V_j ∈ X;
    record d(f, J(π_{W_i}(f)));
   end;
  identify W* ∈ {W_i} with minimum information loss;
  X := W*;
  j := j − 1
 end; □
```

A possible sequence $(X_n, \ldots, X_1)$ for five variables:

$$X_5 = \{\{v_1\}, \{v_2\}, \{v_3\}, \{v_4\}, \{v_5\}\}$$
$$X_4 = \{\{v_1\}, \{v_2\}, \{v_3, v_5\}, \{v_4\}\}$$
$$X_3 = \{\{v_1, v_4\}, \{v_2\}, \{v_3, v_5\}\}$$
$$X_2 = \{\{v_1, v_4\}, \{v_2, v_3, v_5\}\}$$
$$X_1 = \{\{v_1, v_2, v_3, v_4, v_5\}\}.$$

The procedures of Conant (1988) and Cavallo and Klir (1979) have undergone extensive experimental testing, primarily by simulation studies in which they are applied to behavior functions with known reconstruction properties. The testing reported in (Hai and Klir, 1985; Klir and Parviz, 1986) is particularly noteworthy because it provides evidence for the validity of a *reconstruction principle of inductive reasoning* applicable to the estimation of arbitrary (*i.e.*, reconstructable or otherwise) probability distribution functions. What was discovered is that when a probability distribution p is estimated as a relative frequency distribution $\hat{p}$ from data generated by means of it (a total of N tuples $t \in \mathrm{dom}(\mathrm{V})$ generated randomly with expected frequency of generation of tuple t equal to $p(t) \times \mathrm{N}$), the distribution p^* reconstructed from projections of $\hat{p}$ onto the best structure at low levels of refinement tends to be closer to p than is $\hat{p}$ itself:

$$\sum_t |p(t) - p^*(t)| < \sum_t |p(t) - \hat{p}(t)|.$$

[The Hamming distance is used to ensure applicability to all pairs $(p, \hat{p})$ and (p, p^*).]

This effect is most pronounced for reasonably small values of N, since $\lim_{\mathrm{N} \to \infty} \hat{p} = p$. However, it is precisely when one has few observations that one's confidence in the accuracy of the usual estimate $\hat{p}(t) = n(t)/\mathrm{N}$ is weakest and one might be willing to apply such a technique to the problem of estimating a probability distribution function from frequency data. (*See* Klir, 1990*b*, this issue, for further discussion.)

5. Areas for future research

Since the publication of the previous overview articles (Cavallo and Klir, 1981*a*; Klir, 1984; Klir and Way, 1985; Klir, 1987), progress has been made on a number of fronts. Procedures for generating without duplicates all immediate refinements of a set of G-structures have been developed (Cavallo and De Voy, 1989). The identification problem for probabilistic systems has been studied in considerable detail (Pittarelli, 1989*a*, 1990). More work has been done on reconstructability analysis of systems with dynamic properties (Klir, 1990*a*, this issue). Although no further results have been published on the topic of resolution of inconsistent structure systems, techniques for reconstructability analysis for interval-valued probabilistic systems have recently been devised (Pittarelli, 1989*b*). Such systems are much less likely to be inconsistent than are standard probabilistic systems.

Further development of techniques for reconstructability analysis of interval-valued behavior functions is necessary, as is work on semantics (data dependencies) for interpreting the structure of a reconstruction hypothesis for such a system. Similarly, extension to fuzzy measures (Klir and Folger, 1988) other than possibilistic, crisp possibilistic (relational), and probabilistic measures remains an open problem. Also, the question of exactly what is indicated by a reconstruction hypothesis from which a behavior function is only approximately reconstructable is an issue that has yet to be investigated either theoretically or empirically.

Although many of the existing heuristics have been shown experimentally to work quite well, even for systems with up to 100 variables (Conant, 1988), it is fair to say that there is still a great need for additional work on algorithms. This becomes apparent when one considers that in some fields (e.g., economics or chemical engineering), a single system may be defined on many hundreds, or even thousands, of variables. At the same time, techniques in use by researchers in these fields tend to be based predominantly on analysis of pairwise interactions between individual variables. Thus, some studies have failed to detect higher-order constraints detectable by reconstructability analysis (Cavallo, 1979). In order to provide a usable alternative, *i.e.*, in order more adequately to analyze such large-scale systems, it will be necessary somehow to improve dramatically the efficiency of search procedures for reconstructability analysis.

Acknowledgement

I wish to acknowledge Philip Meguire for his translation into French of the abstract and for discussions and correspondence over several years regarding connections between the methodology of reconstructability analysis and topics in Bayesian econometric methodology.

References

W. R. ASHBY, Constraint analysis of many-dimensional relations, *General Systems Yearbook*, 9, 1964, pp. 99-105.

Y. BARD, Estimation of state probabilities using the maximum entropy principle, *I.B.M. J. Res. and Development*, 24, 1980, pp. 563-569.

Y. BARD, A model of shared D.A.S.D. and multipathing, *Communications of the A.C.M.*, 23, 1980, pp. 564-572.

C. BERGE, *Graphs and Hypergraphs*, North-Holland, Amsterdam, 1973.

D. T. BROWN, A note on approximations to discrete probability distributions, *Information and Control*, 2, 1959, pp. 386-392.

R. CAVALLO, *The role of systems methodology in Social Science Research*, Martinus Nijhoff, Boston, 1979.

R. CAVALLO, Reconstructability and identifiability in the evaluation of structure hypotheses: an issue in the logic of modelling, *Proc. Soc. for Gen. Systems Research*, 1980, pp. 647-654.

R. CAVALLO, Lattices of structure models and database schemes, *Int. J. of General Systems* (submitted), 1989.

R. CAVALLO and J. De VOY, Iterative and recursive algorithms for tree and partition search of the lattice of structure models, *Int. J. of General Systems* (submitted), 1989.

R. CAVALLO and G. KLIR, Reconstructability analysis of multi-dimensional relations: a theoretical basis for computer-aided determination of acceptable systems models, *Int. J. of General Systems*, 5, 1979, pp. 143-171.

R. CAVALLO and G. KLIR, Reconstructability analysis: overview and bibliography, *Int. J. of General Systems*, 7, 1981 *a*, pp. 1-6.

R. CAVALLO and G. KLIR, Reconstructability analysis: evaluation of reconstruction hypotheses, *Int. J. of General Systems*, 7, 1981 *b*, pp. 7-32.

R. CAVALLO and G. KLIR, Decision making in reconstructability analysis, *Int. J. of General Systems*, 8, 1982 *a*, pp. 243-255.

R. CAVALLO and G. KLIR, Reconstruction of possibilistic behavior systems, *Fuzzy Sets and Systems*, 8, 1982 *b*, pp. 175-197.

F. CELLIER and D. YANDELL, SAPS-II: a new implementation of the systems approach problem solver, *Int. J. of General Systems*, 13, 1987, pp. 307-322.

R. CHRISTENSEN, *Entropy Minimax Sourcebook*, vol. 1: *General Description*. Entropy Ltd., Lincoln, Massachusetts, 1981.

R. CONANT, Extended dependency analysis of large systems, parts I and II, *Int. J. General Systems*, 14, 1988, pp. 97-123 and pp. 125-141.

N. DALKEY, Inductive inference and the representation of uncertainty, *Proc. 1st Workshop on Uncertainty in Artificial Intelligence*, 1985, pp. 109-116.

W. DEMING and F. STEPHAN, On the least squares adjustment of a sampled frequency table when the expected marginal totals are known, *Annals of Mathematical Statistics*, 11, 1940, pp. 427-444.

R. FAGIN, Degrees of acyclicity of hypergraphs and relational database schemes, *J. of the A.C.M.*, 30, 1983, pp. 514-550.

B. R. GAINES, System identification, approximation and complexity, *Int. J. of General Systems*, 3, 1977, pp. 145-174.

M. GAREY and D. JOHNSON, *Computers and Intractability*, W. H. Freeman, San Francisco, 1979.

A. HAI and G. KLIR, An empirical investigation of reconstructability analysis: probabilistic systems, *Int. J. Man-Machine Studies*, 22, 1985, pp. 163-192.

M. HIGASHI, A systems modelling methodology: probabilistic and possibilistic approaches, *Ph. D. dissertation*, SUNY-Binghamton, Binghamton, New York, 1984.

M. HIGASHI and G. KLIR, Measures of uncertainty and information based on possibility distributions, *Int. J. of General Systems*, 9, 1983 *a*, pp. 43-58.

M. HIGASHI and G. KLIR, On the notion of distance representing information closeness: possibility and probability distributions, *Int. J. of General Systems*, 9, 1983 *b*, pp. 103-115.

M. HIGASHI and G. KLIR, Resolution of finite fuzzy relation equations, *Fuzzy Sets and Systems*, 13, 1984, pp. 65-82.

M. HIGASHI, G. KLIR and M. PITTARELLI, Reconstruction families of possibilistic structure systems, *Fuzzy Sets and Systems*, 12, 1984, pp. 37-60.

E. T. JAYNES, Where do we stand on maximum entropy?, *in:* R. LEVINE and M. TRIBUS Eds., *The Maximum Entropy Formalism*, M.I.T. Press, Cambridge, Massachusetts, 1979.

E. T. JAYNES, On the rationale of maximum-entropy methods, *Proc. of the I.E.E.E.*, 70, 1982, pp. 939-952.

E. T. JAYNES, Prior information and ambiguity in inverse problems, *in:* D. McLAUGHLIN Ed., *Inverse problems, S.I.A.M.-A.M.S.* Proceedings, 14, American Mathematical Society, Providence, Rhode Island, 1984, pp. 151-166.

B. JONES, Determination of unbiased reconstructions, *Int. J. of General Systems*, 10, 1984, pp. 169-176.

B. JONES, Reconstructability analysis for general functions, *Int. J. of General Systems*, 11, 1985, pp. 133-142.

G. KLIR, Identification of generative structures in empirical data, *Int. J. of General Systems*, 3, 1976, pp. 89-104.

G. KLIR, Reconstructability analysis: an overview, *in:* T. OREN *et al.* Eds., *Simulation and Model-Basel Methodologies*, Springer-Verlag, New York, 1984.

G. KLIR, *Architecture of Systems Problem Solving*, Plenum Press, New York, 1985.

G. KLIR, The role of reconstructability analysis in social science research, *Mathematical Social Sciences*, 12, 1986, pp. 205-225.

G. KLIR, Reconstructability analysis, *Systems and Control Encyclopedia*, Pergamon, Oxford, 1987, pp. 3983-3990.

G. KLIR, A principle of uncertainty invariance, *Int. J. of Approximate Reasoning*, 1989 (to appear).

G. KLIR, Dynamic aspects in reconstructability analysis: the role of minimum uncertainty principles, *Revue Int. de Systemique*, this issue, 1990 *a*.

G. KLIR, Reconstruction principle of inductive reasoning, *Revue Int. de Systemique*, this issue, 1990 *b*.

G. KLIR and T. FOLGER, *Fuzzy Sets, Uncertainty, and Information*, Prentice-Hall, Englewood Cliffs, New Jersey, 1988.

G. KLIR, M. MARIANO, M. PITTARELLI and K. KORNWACHS, The potential of reconstructability analysis for production research, *Int. J. of Prod. Research*, 26, 1988, pp. 629-645.

G. KLIR and B. PARVIZ, General reconstruction characteristics of probabilistic and possibilistic systems, *Int. J. of Man-Machine Studies*, 25, 1986, pp. 367-397.

G. KLIR, B. PARVIZ and M. HIGASHI, Relationship between true and estimated possibilistic systems and their reconstructions, *Int. J. of General Systems*, 12, 1986, pp. 319-331.

G. KLIR and E. WAY, Reconstructability analysis: aims, results, open problems, *Systems Research*, 2, 1985, pp. 141-163.

G. KLIR and H. UYTTENHOVE, On the problem of computer-aided structure identification: some experimental observations and resulting guidelines, *Int. J. of Man-Machine Studies*, 9, 1977, pp. 593-628.

K. KRIPPENDORFF, *Information theory: structural models for qualitative data*, Sage Publications, Beverly Hills, California, 1986.

P. M. LEWIS, Approximating probability distributions to reduce storage requirements, *Information and Control*, 2, 1959, pp. 214-225.

R. MADDEN and W. R. ASHBY, The identification of many-dimensional relations, *Int. J. of Systems Science*, 3, 1972, pp. 343-356.

M. MARIANO, Aspects of inconsistency in reconstructability analysis, *Ph. D. dissertation*, SUNY-Binghamton, Binghamton, NY, 1987.

H. MOSKOWITZ and J. WALLENIUS, *Conditional versus joint probability assessments*, Paper # 851, Krannert Graduate School of Management, Purdue University, W. Lafayette, Indiana, 1984.

M. PITTARELLI, Uncertainty and estimation in reconstructability analysis, *Int. J. of General Systems*, 15, 1989 *a*, pp. 1-58.

M. PITTARELLI, Reconstructability analysis using probability intervals, *Int. J. of General Systems*, 1989 *b* (to appear).

M. PITTARELLI, Data synthesis from probabilistic structure systems, *Revue Int. de Systemique*, this issue, 1990.

G. SHAFFER and P. CAHOON, Extracting information from ecological data containing high spatial and temporal variability: benthic microfloral production, *Int. J. of General Systems*, 13, 1987, pp. 107-123.

J. SHORE and R. JOHNSON, Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy, *I.E.E.E. Trans. Information Theory*, 26, 1980, pp. 26-37.

Y. TIAN, Probabilistic databases over acyclic schemes, *M.S. thesis*, SUNY Institute of Technology, Utica, New York, 1988.

H. UYTTENHOVE, Systems approach problem solver and the open heart surgery patient, *in:* R. TRAPPL Ed., *Cybernetics and Systems Research*, North-Holland, Amsterdam, 1982, pp. 655-661.

R. VALDES-PERES and R. CONANT, Information loss due to data quantization in reconstructability analysis, *Int. J. of General Systems*, 9, 1983, pp. 235-247.

REVUE INTERNATIONALE DE SYSTEMIQUE
Vol. 4, N° 1, 1990, pp. 33 à 43

DYNAMIC ASPECTS IN RECONSTRUCTABILITY ANALYSIS: THE ROLE OF MINIMUM UNCERTAINTY PRINCIPLES *

George J. KLIR

State University of New York [1]

Abstract
The role of principles of minimum uncertainty in dealing with the reconstruction problem of systems with dynamic properties is discussed. The aim of the reconstruction problem, one of two problems addressed by reconstructability analysis, is to determine the smallest possible sub-systems by which a given overall system can be adequately represented.

Résumé
Nous traitons du rôle des principes de moindre incertitude en analysant le problème de reconstruction de systèmes à caractère dynamique. Le but du problème de reconstruction, l'un des deux problèmes concernés par l'analyse de la reconstructibilité, est de déterminer les sous-systèmes minimaux capables de représenter de manière adéquate un système global donné.

This little paper is dedicated to Wyllis Bandler. In my opinion, Wyllis is a rather unusual mathematician in the sense that his research work has almost always focused on important but underdeveloped areas of mathematics. One area that has considerably been advanced by Wyllis' research is the area of mathmatical relations (Bandler and Kohout, 1980 *a*, *b*, 1986, 1987 *a*, *b*). It was primarily this area of research, which was of interest to both of us, that brought us together some 15 years ago. The aim of this paper is to illustrate

* This paper was supported in part by the National Science Foundation under Grant IST 86-44676; its original version was prepared for Wyllis Bandler's Festschrift.

1. Binghamton, New York 13901 U.S.A., Department of Systems Science, Thomas J. Watson School of Engineering, Applied Science, and Technology.

Rev. intern. systémique. 0980-1472 Vol. 4/90/01/ 33 /11/$ 3.10/© Afcet Gauthier-Villars