

On Finding the Bidimension of a Relation

M. G. M. KOPPEN

University of Utrecht, Utrecht, The Netherlands

A method is presented for evaluating the bidimension of a finite binary relation, i.e., the number of biorders (Guttman relations) needed to yield the relation as their intersection. In case the relation is induced by a binary data matrix, the bidimension equals the minimal number of dimensions needed for a representation of the data matrix according to the conjunctive model of C. H. Coombs and R. C. Kao (*Nonmetric factor analysis*, Engineering Research Bulletin No. 38, Univ. of Michigan Press, Ann Arbor, 1955). Central to the evaluation of the bidimension is its characterization, provided by J.-P. Doignon, A. Ducamp, and J.-C. Falmagne (*Journal of Mathematical Psychology*, **28**, 73-109, 1984), as the chromatic number of some associated hypergraph. A procedure is described to reduce hypergraphs of this kind to subhypergraphs with the same chromatic number. This reduction can be used throughout in applying a recurrence relation that expresses the chromatic number of a hypergraph in terms of the chromatic numbers of some of its subhypergraphs. © 1987 Academic Press, Inc.

1. INTRODUCTION

In many settings of psychological research and testing the situation is so complex that it is highly unrealistic to expect that the behaviour of the subjects can be explained satisfactorily by one single dimension along which they and the experimental stimuli ("items") vary. In such cases, a multidimensional model is called for. An all-important aspect of such a model is its composition rule, by which positions on the separate, unobserved dimensions combine to give a "score" that is directly related to the observed behaviour. The most widely used composition rule consists in mapping positions on the separate dimensions into the real number line and expressing the observed score as a linear combination (weighted sum) of these scores on the separate dimensions (e.g., factor analysis, analysis of variance, regression analysis). A model with this kind of composition rule can be considered as a compensatory one: a deficiency on one dimension can be compensated for (and without limit) by a surplus on another dimension. In many situations such a rule is, again, not very realistic, especially in the case of binary data that are instances of the general type: subject solves/fails item. In such cases an essentially different composition rule is conceivable in which "a person solves an item" (these are all used as generic terms) if and only if on each of the separate dimensions the position of the

This research was supported by the Netherlands Organisation for the Advancement of Pure Research (Grant No. 560-670-006). The author is grateful to J.-P. Doignon for his useful comment on a previous draft and to a referee for supplying the data set of Section 4.

person dominates the position of the item. In this multidimensional generalization of Guttman's (1944) idea there is no compensating mechanism involved; rather, an item corresponds to a threshold on each dimension that has to be surpassed by a person in order to solve the item. The observed score, then, is not seen as an arithmetical sum of the scores on the separate dimensions, but rather as the logical product of these scores. This explains the name conjunctive model (Coombs and Kao, 1955; Coombs, 1964) for models with this composition rule.

The conjunctive composition rule immediately suggests another one, in which a person solves an item if and only if the position of the item on at least one dimension. With this rule the observed score corresponds to the logical sum of the scores on the separate dimensions, so this model is called disjunctive. Thanks to the logical equivalence

$$(A_1 \text{ or } \cdots \text{ or } A_n) \text{ iff not } ((\text{not } A_1) \text{ and } \cdots \text{ and } (\text{not } A_n)),$$

the disjunctive model needs no separate consideration. By flipping the binary data and reversing the direction of each dimension it is directly translated into the conjunctive form.

An informal discussion of how to construct for some given data matrix a representation according to the conjunctive model can be found in Coombs and Kao (1955) and Coombs (1964). In order to develop an explicit general algorithm, however, one has to face two important problems. The first one is that of finding the minimum dimensionality needed for such a representation (note that for the present we are considering deterministic models only) and the second has to do with the uniqueness of an obtained representation in the minimum dimensionality. This paper will be concerned with the problem having logical priority, finding the minimum dimensionality.

That this is not a trivial problem at all is shown by Doignon, Ducamp, and Falmagne (1984). In their fundamental paper they cast the problem in terms of representing the relation between the set of persons and the set of items defined by the binary data matrix as the intersection of a minimal number of biorders (Guttman relations), this number being its so-called bidimension. Some of their results were, for the finite case, independently obtained by Cogis (1980, 1982). Regarding computational complexity (see, e.g., Garey and Johnson, 1979), Cogis (1982) showed that the problem of finding the bidimension of a relation is polynomially equivalent to that of determining the dimension of a partial order (i.e., the minimal number of linear orders to yield the partial order as their intersection; see Dushnik and Miller, 1941), and Yannakakis (1982) showed the latter problem to be NP-complete for a dimension greater than 2 (the two-dimensional case is polynomially decidable). Doignon *et al.* give a characterization of this bidimension as the chromatic number of a certain hypergraph, generalizing a similar approach by Trotter (1983) for the usual order dimension. It is this equivalence that we will use for the computation of the bidimension. So first we summarize in the next section the relevant part of the theory developed in Doignon

et al. (1984). Next, in Section 3, we describe some principles enabling us to reduce a hypergraph of the considered type to a subhypergraph having the same chromatic number. This reduction will appear to include as a special case the collapsing of the data matrix into a submatrix, its so-called core, as is obtained by Chubb (1986) in a somewhat more general context (i.e., in principle not restricted to biorder representations). Furthermore, it generalizes the restriction in the case of a partial order to the so-called "non-forced" pairs (Trotter, 1983). In Section 4 this reduction process is illustrated, using an empirical data set. In Section 5 we give a recursive formula for the chromatic number of a hypergraph in terms of the chromatic number of some of its subhypergraphs and in Section 6 we show how the reduction and recursion we developed can be combined to compute the bidimension of another empirical data set. In the last section we discuss the prospects for integrating the findings of the preceding sections in a really explicit and reasonably efficient algorithm. In this context our second problem, that of uniqueness, also comes into view: Is it possible to combine our approach of determining the bidimension with a computation or a characterization in some sense of all possible representations in that dimensionality?

2. BASIC THEORY

First we will fix some general notation and definitions. For two sets A and D , $R \subseteq A \times D$ is called a relation between A and D ; if $R \subseteq A \times A$ it is a relation on A . We will write ad for the ordered pair (a, d) and $ad \in R$ or aRd , equivalently. $\bar{R} = (A \times D) - R$ denotes the complement of R . The cardinality of a set A is denoted $|A|$. Since we aim to applying their theory to real data, we will, in contrast to Doignon *et al.*, throughout assume A and D to be finite. So $A = \{a_1, \dots, a_N\}$ and $D = \{d_1, \dots, d_K\}$ for some natural numbers N and K . As a consequence, a relation R between A and D , as well as its complement \bar{R} , will always be finite.

For interpretative purposes the elements of A can be thought of as persons, those of D as items, and R can be regarded as a dominance relation, $a_i R d_j$ meaning: person a_i solves correctly item d_j . Because of finiteness we can represent a relation R in an $N \times K$ (0, 1)-matrix, called the data matrix, having a 1 in cell (i, j) if $a_i R d_j$ and a 0 if $a_i \bar{R} d_j$. We will denote this matrix as $[R]$; $[R]_{.i}$ is its i th row and is sometimes called the pattern of a_i , $[R]_{.j}$ is its j th column or the pattern of d_j , and $[R]_{ij}$ is the value in cell (i, j) .

There are partial orders on the rows and columns of $[R]$ (corresponding directly to the quasi-orders R_A and R_D of Doignon *et al.*) defined by

$$[R]_{.i} \leq [R]_{.j} \quad \text{iff} \quad (\text{for } k = 1, \dots, K, [R]_{ik} \leq [R]_{jk})$$

and

$$[R]_{.i} \leq [R]_{.j} \quad \text{iff} \quad (\text{for } n = 1, \dots, N, [R]_{ni} \leq [R]_{nj}).$$

Now we will summarize the findings of Doignon *et al.* (1984) as far as they are relevant to our more practical goal. Where useful we will specialize their statements

to the finite case and give a translation in terms of the data matrix. In the sequel we may assume that A and D are disjoint: if they are not, we pass to disjoint copies A' and D' of A and D , respectively, and for any relation $R \subseteq A \times D$ we consider its isomorphic image $R' \subseteq A' \times D'$; results there can be directly translated back to the original $R \subseteq A \times D$ (see Doignon *et al.*).

The central concept, already defined in Ducamp and Falmagne (1969), is that of a biorder:

$R \subseteq A \times D$ is called a *biorder* between A and D iff for all $a, b \in A, d, e \in D$ we have: if $(aRd$ and $bRe)$ then $(aRe$ or $bRd)$.

An equivalent, more symmetrical formulation of this condition reads:

not $(aRd$ and bRe and $a\bar{R}e$ and $b\bar{R}d)$,

from which it is immediately clear that \bar{R} is a biorder iff R is one. For the matrix $[R]$ this definition means that R is biorder iff $[R]$ has no 2×2 submatrix (permutations of rows and columns allowed) of the form

$$\begin{array}{cc} 0 & 1 \\ 1 & 0 \end{array} \quad (2.1)$$

which implies there are permutations of the rows and the columns of $[R]$ that bring the matrix in triangular form (the matrix is then said to have *triangular structure*).

Of course not every relation is a biorder, but Doignon *et al.* show that any relation R between A and D is, in a trivial way, the intersection of $|\bar{R}|$ biorders, which in our case is a finite cardinal. This leads to the concept of bidimension, which for the finite case reads:

The *bidimension* of a relation $R \subseteq A \times D$, denoted *Bidim* R , is the smallest number q for which there is a collection of q biorders $B_i \subseteq A \times D, i = 1, \dots, q$, with $R = \bigcap_{i=1}^q B_i$.

The relevance of the concept of bidimension for a representation of R according to the conjunctive model lies in the following proposition (where \mathbb{R} denotes the set of real numbers):

Bidim R is the smallest number q for which there are two mappings $f = (f_1, \dots, f_q): A \rightarrow \mathbb{R}^q$ and $g = (g_1, \dots, g_q): D \rightarrow \mathbb{R}^q$ such that for all $a \in A, d \in D: aRd$ iff $f_i(a) \leq g_i(d)$ for $i = 1, \dots, q$.

So we have the practical problem of computing the bidimension of a relation. To that end we will use the equivalence, derived by Doignon *et al.*, between the bidimension and the chromatic number of a certain hypergraph (for hypergraphs, see Berge, 1973).

A *hypergraph* $H = \langle V, E \rangle$ is a set V of elements called *vertices* together with a collection E of subsets of V , called *edges*. A subset of V is called

stable in H iff it includes no edge of H and the *chromatic number* of H , denoted $\text{Chrom } H$, is the smallest number q for which there is a q -colouring of H , that is, a partition of V into q stable sets, the *colour classes*; in other words, in a colouring no edge is "monochromatic."

We will not consider hypergraphs having singleton edges; in this case the trivial partition into one-element classes is a colouring, so the chromatic number is well-defined and is finite whenever the number of vertices is.

The definition of the hypergraph in question is based on the following generalization of the violating case for a biorder:

An n -alternating cycle of \bar{R} is a sequence $(a_i d_i)_{i=0}^{n-1}$ of elements in \bar{R} such that $a_{i+1} R d_i$ for all i taken modulo n .

This definition is ours; Doignon *et al.* define the corresponding notion as a sequence of elements of A and D , alternately, that "induces" the sequences in \bar{R} and R of the definition above. (The sequence $(a_{i+1} d_i)$ in R is in turn an n -alternating cycle of R , in reverse order.) Since it is the sequence in \bar{R} that is important in the sequel, we think our definition is more direct.

So by definition R is a biorder iff \bar{R} has no 2-alternating cycle, but it follows more generally:

R is a biorder iff \bar{R} has no alternating cycle.

In terms of the matrix $[R]$ we see that the existence of, say, a 4-alternating cycle of \bar{R} implies the existence of a 4×4 submatrix of $[R]$ (permutations of rows and columns allowed) of the form

$$\begin{array}{cccc} 0 & x & x & 1 \\ 1 & 0 & x & x \\ x & 1 & 0 & x \\ x & x & 1 & 0 \end{array} \quad (2.2)$$

where the x 's are arbitrary.

Now the *hypergraph associated with R* , which we will, in a slight departure from Doignon *et al.*, denote $H(\bar{R})$, is defined as follows:

$H(\bar{R})$ is the hypergraph whose vertices are the elements of \bar{R} and whose edges are the alternating cycles of \bar{R} , interpreted as sets.

Then we have the promised equivalence

$$\text{Bidim } R = \text{Chrom } H(\bar{R}),$$

on which our method of finding the bidimension of a relation will be based.

3. REDUCTION OF THE HYPERGRAPH

The significance of the equivalence between the bidimension of a relation and the chromatic number of the associated hypergraph for the practical purpose of computing the bidimension depends on the extent to which this approach gives way to more efficient algorithms. By the computational equivalence, established by Cogis (1982), of the bidimension problem for an arbitrary relation and the order dimension problem for a partial order, combined with a result of Yannakakis (1982) for the complexity of the latter problem, we know that for $k > 2$ deciding whether k is an upper bound for the bidimension of a given relation is an NP-complete problem. So we cannot expect to find a theoretically efficient, i.e., polynomial time-bounded algorithm. The efficiency will, in practice, strongly depend on the size of the input. Hence, when we translate the problem in terms of finding the chromatic number of a hypergraph, it will be of importance to take care that this hypergraph be as small as possible. To that end we will search for conditions under which we can reduce the hypergraph associated with a relation to one having fewer vertices and edges but the same, yet unknown, chromatic number.

In order to formulate our principle of reducing the hypergraph, we first need some definitions.

3.1. DEFINITION. Let $\langle V, E \rangle$ be a hypergraph. For $V^* \subseteq V$ we define $E(V^*) = \{U \in E : U \subseteq V^*\}$. So $\langle V^*, E(V^*) \rangle$ is the subhypergraph obtained from $\langle V, E \rangle$ by restricting the set of vertices to V^* and the set of edges to those included in V^* .

Obviously, if $V_1 \subseteq V_2 \subseteq V$, then $\text{Chrom } \langle V_1, E(V_1) \rangle \leq \text{Chrom } \langle V_2, E(V_2) \rangle$.

3.2. DEFINITION. For a hypergraph $\langle V, E \rangle$ and subsets $V_1 \subseteq V_2 \subseteq V$, let φ_1 (φ_2) be a colouring of $\langle V_1, E(V_1) \rangle$ ($\langle V_2, E(V_2) \rangle$). Then φ_2 is said to be an *extension* of φ_1 iff for all $C \in \varphi_1$ there is $C' \in \varphi_2$ with $C \subseteq C'$.

So extensions to V_2 do not break up the classes already formed in V_1 . Note that we may obtain non-trivial extensions while $V_2 = V_1$ by merging different colour classes into one new.

3.3. DEFINITION. For a hypergraph $H = \langle V, E \rangle$ and some $V^* \subseteq V$, let φ be a colouring of $\langle V^*, E(V^*) \rangle$. Then $\text{Chrom}_\varphi H$ is defined to be the chromatic number of H under the restriction to colourings that are extensions of φ .

As can be checked easily from Definition 3.2, the extension relation on colourings of subhypergraphs of H is transitive (in fact, it is a partial order): an extension of an extension of some colouring φ is again an extension of φ . In this way we see that for two colouring φ_1 and φ_2 of subhypergraphs of H , φ_2 being an extension of φ_1 implies

$$\text{Chrom } H \leq \text{Chrom}_{\varphi_1} H \leq \text{Chrom}_{\varphi_2} H,$$

simply because, going from left to right in the above inequalities, the minimum is taken over a decreasing collection of colourings of H . Clearly both inequalities may be replaced by equalities iff φ_2 (and thus φ_1) can be extended to a minimal colouring of H .

The reduction of the hypergraph we are going to consider here is based on the following property of a vertex:

3.4. DEFINITION. Let $H = \langle V, E \rangle$ be a hypergraph and $v, w \in V$. Then v is said to be *dominated in H by w* iff for any edge U of H that contains v the set $(U - \{v\}) \cup \{w\}$ is non-stable in H .

More generally, for $v \in V$ and $T \subseteq V$, stable in H , we will say that v is *dominated in H by T* iff for any edge U of H that contains v the set $(U - \{v\}) \cup T$ is non-stable in H .

This definition of being dominated is equivalent to saying that whenever adding a dominated vertex to a stable set in H would turn it into a non-stable set, adding the dominating vertex (subset) instead would have the same effect.

Now we are ready to formulate, for hypergraphs in general, the following

3.5. REDUCTION PRINCIPLE. Let $H = \langle V, E \rangle$ be a hypergraph and let v be an element of V that is dominated in H by another vertex $w \in V$. Then any colouring of $H^* = \langle V - \{v\}, E(V - \{v\}) \rangle$ can be extended to a colouring of H by giving v the colour of w .

In particular, $\text{Chrom } H = \text{Chrom } H^* = \text{Chrom}_{\{v, w\}} H$.

The proof of this reduction principle really is immediate from Definition 3.4, but we will deduce it from the following more general version 3.5' that makes some assumption on the colouring of the subhypergraph. The proof of 3.5' will again appear to be an easy consequence of Definition 3.4.

3.5'. REDUCTION PRINCIPLE (generalized, conditional version). Let $H = \langle V, E \rangle$ be a hypergraph and let v be an element of V that is dominated in H by some stable $T \subseteq V - \{v\}$. Then any colouring φ^* of $H^* = \langle V - \{v\}, E(V - \{v\}) \rangle$ that is an extension of $\{T\}$, i.e., in which T is monochromatic, can be extended to a colouring φ of H by adding v to the colour class of T .

In particular, $\text{Chrom}_{\varphi^*} H = \text{Chrom}_{\varphi^*} H^* = \text{Chrom}_{\varphi} H$.

Proof. Any monochromatic edge U of H induced by adding v to the colour class of T in φ^* must contain v and so the set $(U - \{v\}) \cup T$, which is a subset of $V - \{v\}$, is monochromatic in φ^* . However, since v is dominated by T , this set is non-stable and this contradicts the assumption that φ^* is a colouring of H^* . ■

Proof of 3.5. Put $T = \{w\}$ and apply 3.5', noting that any colouring of H^* is an extension of $\{\{w\}\}$. ■

Note that the reduction principle (in both versions) is constructive in the sense

that it gives a way to construct a minimal colouring of the greater hypergraph from a minimal colouring of the smaller one.

By reduction principle 3.5', a vertex that is dominated by a (disjoint) stable subset of vertices needs no separate consideration in the colouring process of a hypergraph, *provided that a dominating subset is monochromatic*. In that case any colour that is good for the dominating subset will do for the dominated vertex as well. As for the chromatic number of the hypergraph, we can split the collection of colourings into those in which the dominating subset is monochromatic and those in which it is not. In finding the chromatic number over the former subcollection we may discard the dominated vertex from the hypergraph. The version 3.5 is just the special case where there is a singleton dominating subset, being monochromatic in any colouring. In this case the dominated vertex can be discarded unconditionally. In the rest of this paper we will in fact use only this special version; in Section 5 a procedure is sketched that allows this restriction, but we may think of alternatives for that procedure that use the more general, conditional version 3.5'.

All this is very fine, but from a practical point of view two intrinsically related questions naturally arise: Will there, in the type of hypergraph defined in the preceding section, be any dominated vertices and how are we going to find them? In order to establish that vertex v is dominated by vertex w , we have, by Definition 3.4, to enumerate all edges containing v , replace v by w in each such edge, and determine whether the resulting set contains an edge. This seems like a lot of work, even when we use the observation that we need not really check *all* edges, but only those that are minimal (i.e., that do not strictly include another edge), which observation by the way introduces the problem of checking whether an edge is minimal. In our type of hypergraph a vertex will generally be part of many (minimal) edges of varying sizes. Moreover, this kind of hypergraph is only implicitly given in the data matrix: its edges are not directly visible, but must be detected by completing alternating cycles.

To get an idea of the intricacy of the hypergraphs considered here, suppose we have tracked down in our data matrix a minimal 4-edge. It must be contained in a submatrix as in (2.2) in which all x 's are zeros (for otherwise the 4-edge is not minimal, as can be checked easily). This means, however, that this submatrix contains not just one minimal 4-edge, but six of them. In general, a minimal n -edge implies an $n \times n$ submatrix with just one 1-entry in each row and each column. In such a matrix there are in fact $(n-1)!$ minimal n -edges; more generally, it contains, for $k = 2, \dots, n$, $n!/(n-k)!/k$ minimal k -edges. These counts can be justified by noting that each such k -edge is determined by the sequence of k 1-entries, used as "stepping-stones" in the alternating cycle. Clearly there are $n!/(n-k)!$ sequences of k out of n 1-entries. We have to divide by a factor k because a k -edge is invariant under cyclic permutations of the k 1-entries involved. The reader may verify that the matrix of (2.2) (with x 's equal to zero) contains six 4-edges, eight 3-edges, and six 2-edges.

In view of this, prospects for applying the reduction principle 3.5 (or 3.5') seem rather poor. Therefore the following proposition is crucial, which states that for this

special type of hypergraph, dominated vertices can be found by inspecting 2-edges only. We first need one more definition.

3.6. DEFINITION. We call $ad, a^*d^* \in \bar{R}$ *enemies* in \bar{R} iff $ad^* \in R$ and $a^*d \in R$. The (simple, unoriented) graph of the so defined symmetric *enemy relation* on \bar{R} is denoted as $G(\bar{R})$.

We see that two elements of \bar{R} are enemies iff they constitute a 2-edge of $H(\bar{R})$, which means that they can never be in one colour class of $H(\bar{R})$. In the matrix two zeros are enemies iff they are the zeros of a 2×2 submatrix as in (2.1). The graph $G(\bar{R})$ is the partial (hyper)graph of $H(\bar{R})$, obtained by discarding all edges with more than two elements.

For $V \subseteq \bar{R}$ denoting by $H(V)$ the subhypergraph $\langle V, E(V) \rangle$ of $H(\bar{R})$, we can now state

3.7. PROPOSITION. Let $R \subseteq A \times D$, $V \subseteq \bar{R}$, $ad, a'd' \in V$. Then ad is dominated in $H(V)$ by $a'd'$ if it is dominated in $G(\bar{R})$ by $a'd'$.

We will again deduce Proposition 3.7 from a more general version 3.7', which is in terms of being dominated by a subset of vertices.

3.7'. PROPOSITION. Let $R \subseteq A \times D$, $V \subseteq \bar{R}$, $ad \in V$, and $T \subseteq V$, stable in $H(V)$. Then ad is dominated in $H(V)$ by T if it is dominated in $G(\bar{R})$ by T .

Proof. Let U be an edge of $H(V)$ containing ad ; $U = \{ad, a_1d_1, \dots, a_nd_n\}$, say, where the ordering of the elements corresponds to the underlying alternating cycle. In order to show that $(U - \{ad\}) \cup T$ contains an edge of $H(V)$ if T dominates ad in $G(\bar{R})$, we consider the pair a_1d_n . If it is in R , the subset $\{a_1d_1, \dots, a_nd_n\} = U - \{ad\}$ is an edge of $H(V)$ and we are finished. If a_1d_n is in \bar{R} , then ad and a_1d_n are enemies in \bar{R} . Since T dominates ad in $G(\bar{R})$, $T \cup \{a_1d_n\}$ contains an edge of $G(\bar{R})$, which means that a_1d_n has some enemy a^*d^* in T . This, however, implies that $\{a^*d^*, a_1d_1, \dots, a_nd_n\} \subseteq (U - \{ad\}) \cup T$ is an edge of $H(V)$. ■

Proof of 3.7. Put $T = \{a'd'\}$ and apply 3.7'. ■

By Propositions 3.7 and 3.7' we may, when applying 3.5, resp. 3.5', to (sub)hypergraphs of our special type, replace the phrase "dominated in H " by "dominated in $G(\bar{R})$." The fact that a vertex ad is dominated in $G(\bar{R})$ by $a'd'$ simply means that any enemy of ad in \bar{R} is an enemy of $a'd'$ as well; likewise, ad being dominated in $G(\bar{R})$ by a subset T means that any enemy of ad in \bar{R} has an enemy in T . So questions concerning being dominated in $G(\bar{R})$ can easily be settled from inspection of the tableau of the binary enemy relation on \bar{R} . In this way, finding dominated vertices is in some sense reduced to inspecting 2-edges only. Notice, however, the appearance of $G(\bar{R})$ in 3.7 and 3.7' instead of the perhaps more expected $G(V)$ (with the obvious meaning for this notation). While it is clear that being dominated in $G(V)$ is necessary for being dominated in $H(V)$ ($G(V)$ being a partial hypergraph of $H(V)$), it is not sufficient. The reason for this can be seen in the

proof of 3.7': there we call on a pair $a_1 d_n$ that, in the non-trivial case, is in \bar{R} , but in no way needs to be in V . On the other hand, being dominated in $G(\bar{R})$, while sufficient, is not necessary for being dominated in $H(V)$: there is no "only if" in 3.7 or 3.7'. This means that by applying 3.7 or 3.7' we may not detect all dominated vertices in $H(V)$. The danger of missing some dominated vertices certainly depends on the discrepancy between the sufficient $G(\bar{R})$ and the necessary $G(V)$, that is, the discrepancy between V and \bar{R} . (In case $V = \bar{R}$ we trivially do have "only if's" in 3.7 and 3.7'.) In this respect it is important to realize that whenever $V \subseteq A' \times D'$ for some $A' \subseteq A$, $D' \subseteq D$, then $V \subseteq \bar{R}'$, where R' is the restriction of R to $A' \times D'$, and we may use \bar{R}' instead of \bar{R} in 3.7 and 3.7'.

There is an important special application of Proposition 3.7 in which dominated vertices can be detected from inspection of the data matrix itself, that is, without even explicitly considering the enemy relation on the vertices (the zeros in the matrix). For this special kind of dominated vertices there is, moreover, some intuitive justification that indeed they have no bearing on the dimensionality of a representation of a relation R according to the conjunctive model. The subset of vertices in question can best be characterized in terms of the data matrix $[R]$, in which we need to find a colouring for all zeros.

Consider two ordered patterns, of persons a_i and a_j say; assume $[R]_j \leq [R]_i$. Because of this ordering, in any column in which $[R]_i$ has a zero, $[R]_j$ has one, too. Let k be such a column and let $a_j d_k$ and $a^* d^*$ be enemies. Then we know that $a^* d_k \in R$ and $a_j d^* \in R$; by the ordering the latter implies $a_i d^* \in R$ and together with the assumption $a_i d_k \in \bar{R}$ it follows that $a_i d_k$ and $a^* d^*$ are enemies. So, by Proposition 3.7, the vertex $a_j d_k$ is dominated in $H(\bar{R})$ by $a_i d_k$.

There is a completely analogous version for the case of two ordered column patterns. From this we see that in a pattern we have left to be coloured only those zeros for which there is no higher-ordered pattern with a zero in the same position. A direct consequence is that in determining the bidimension, a pattern in which all occurring zeros can be "explained" in this way, that is, a pattern that is the conjunction of a number of higher-ordered patterns, can be discarded altogether. Intuitively one might feel that such a pattern does not offer any new information regarding the multidimensional representation. Suppose, for instance, we have a representation for two persons, one of which failed on item 1 only, the other on item 2 only. Then, precisely because we work in the conjunctive model, there must already be in that representation a place for a person failing on just the items 1 and 2. According to the conjunctive model the (possible) occurrence of the last pattern is, actually, implied by the occurrence of the first two. Again there is an analogous intuitive argument with the roles of persons and items reversed. Patterns, rows or columns, that for this reason can be removed from the data matrix will be called *implied patterns*. Separate zeros that, in the sense defined above, can be explained by a zero a in a higher-ordered pattern will be called *implied zeros*. Note that repetitions in the data matrix of one and the same pattern constitute a special, trivial case of implied patterns. So, such repetitions can be discarded without changing the bidimension, a fact that is intuitively self-evident.

Although it is quite easy to present data matrices in which reduction of the hypergraph according to 3.5 and 3.7 has no effect at all, the special case of implied zeros shows that under rather mild conditions (occurrence of ordered rows or columns in the data matrix) the associated hypergraph is bound to contain dominated vertices. On the other hand, because the conjunctive model predicts in some sense, their occurrence, as we have seen above, the amount of implied zeros in a data matrix may be considered as some sort of measure of confirming evidence for the conjunctive model as the operative one in producing the data.

There exist, as was pointed out by Doignon (personal communication), close connections between the notions of implied patterns and implied zeros and the work of Chubb (1986) and Trotter (1983), respectively.

In fact, the special application of Proposition 3.7 consisting of removing implied rows and columns gives exactly the restriction of R as described by Chubb. Our non-implied rows (columns) constitute his minimal row (column) \cap -generating set and the corresponding restriction of R is its so-called \cap -core. We see that by the present reduction we can, in addition, "remove" separate zeros and by fully using Proposition 3.7 this may well be more zeros than just the implied ones. In particular this means that we may, possibly, remove more rows or columns, thereby obtaining a still smaller "core" of R .

In the special case where $D = A$ and the relation R is a (reflexive) partial order on A , its bidimension equals its order dimension (Doignon *et al.*, 1984) and then the non-implied zeros correspond exactly to the "non-forced pairs," the subset of incomparable pairs of R on which Trotter (1983) defines his hypergraph. It can be shown that in this case all dominated zeros are implied zeros and in this way Proposition 3.7 (together with 3.5) is equivalent to a result of Maurer, Rabinovitch, and Trotter (1980), implying that for the dimension of a partial order consideration can be confined to the set of non-forced pairs. In the description of the implied zeros above we have seen a way to find this set.

4. AN ILLUSTRATION OF THE REDUCTION PROCESS

Here we will demonstrate the potential significance of the reduction of a hypergraph according to Proposition 3.7. We will use a data set from Chubb (1986) that originates, in turn, from data of Stouffer *et al.* (1950) on six polytome items administered to a set of American World War II GI's and intended to assess their "readiness to enter into battle." By appropriately dichotomizing the responses Chubb obtained the data matrix of Table 4.1(a), where capitals A to F denote the six items and the rows represent the observed response patterns. Of course, the response patterns occurred with varying frequencies, but for a deterministic analysis of the data these frequencies are irrelevant and it suffices to consider the reduced data matrix without duplicate rows or columns. (When searching for approximate representations of the data in a lower dimensionality it is natural to take the frequencies into account.)

TABLE 4.1

(a) Data Matrix from Chubb (* Indicates a Row-Implied Zero, Non-Implied Rows Are Numbered);
 (b) Submatrix of Non-Implied Patterns (* for an Implied Zero)

(a)	A	B	C	D	E	F	(b)	A	B	C	D	E	F
	1	1	1	1	1	1	1	1	1	0	1	1	1
(1)	1	1	0	1	1	1	2	1	0	1	1	1	1
(2)	1	0	1	1	1	1	3	0	1	1	1	1	1
(3)	0	1	1	1	1	1	4	*	1	1	1	1	0
(4)	*	1	1	1	1	0	5	*	1	1	0	1	1
(5)	*	1	1	0	1	1	6	1	1	*	0	0	1
	*	*	1	1	1	1	7	*	*	*	1	0	*
	1	*	*	1	1	1							
	*	*	1	*	1	1							
	*	*	*	1	1	1							
	*	1	*	*	1	1							
(6)	1	1	*	0	0	1							
	*	*	1	*	1	*							
	*	*	*	*	1	1							
	*	1	*	*	*	1							
	*	*	*	*	*	1							
(7)	*	*	*	1	0	*							
	*	*	*	*	*	*							

In Table 4.1(a) we have already indicated which zeros are row-implied: a "*" denotes a zero for which there is a non-implied zero, denoted as "0," in the same column in some higher-ordered pattern (i.e., higher in the partial order on the rows of the matrix). By the special application of Proposition 3.7 discussed in the preceding section we may confine our attention to the subset of non-implied rows. These are numbered in Table 4.1(a) and the corresponding submatrix is given separately in Table 4.1(b). There are no extra column-implied zeros: only one row (6) has more than one "living" zero, but the corresponding columns (*D* and *E*) are unordered. So here ends the special application of Proposition 3.7 consisting of removing implied zeros from the hypergraph, and the submatrix of Table 4.1(b) (without the distinction between implied and non-implied zeros) is the "core," the restriction that is obtained by Chubb.

In order to fully use Proposition 3.7 we now consider the enemy relation between the vertices of the reduced hypergraph (the non-implied zeros) and those of the full hypergraph (all zeros in Table 4.1(b)). This enemy relation is represented in the 8×15 matrix of Table 4.2(a), where 1-entries indicate that the vertices in the corresponding row and column constitute a 2-edge. We see that the vertex 3*A* is dominated by 4*F*; 6*E* is dominated by 1*C*, and 7*E* by 4*F* (for instance). In Table 4.2(a) the rows of dominated vertices are starred and by Proposition 3.7 the problem reduces to the situation in Table 4.2(b). There we see that this reduction has turned the vertex 6*D* into a dominated one and its removal leads us to the

TABLE 4.2

- (a) Enemy Relation for the Hypergraph of Table 4.1(b) (Starred Rows Indicate Dominated Vertices);
 (b) Resulting Submatrix and Enemy Relation after Removing Dominated Vertices;
 (c) Final Reduction

(a)	1C	2B	3A	4F	5D	6D	6E	7E	4A	5A	6C	7A	7B	7C	7F
1C	.	1	1	1	1	.	.	.	1	1
2B	1	.	1	1	1	1	1	.	1	1	1
*3A	1	1	.	.	.	1	1	.	.	.	1
4F	1	1	.	.	1	1	1	.	.	.	1
5D	1	1	.	1	.	.	.	1	1	1	1
6D	.	1	1	1	1	.	.	1	1	.	1
*6E	.	1	1	1	1	1
*7E	1

(b)	B	C	D	F		1C	2B	4F	5D	6D	6C
1	1	0	1	1		1C	.	1	1	1	.
2	0	1	1	1		2B	1	.	1	1	1
4	1	1	1	0		4F	1	1	.	1	1
5	1	1	0	1		5D	1	1	1	.	.
6	1	*	0	1		*6D	.	1	1	.	.

(c)	B	C	D	F		1C	2B	4F	5D
1	1	0	1	1		1C	.	1	1
2	0	1	1	1		2B	1	.	1
4	1	1	1	0		4F	1	1	.
5	1	1	0	1		5D	1	1	1

hypergraph consisting of four vertices and the 4×4 submatrix of Table 4.2(c). In fact, here reduced and full hypergraphs are the same and after inspecting the enemy relation on its vertices we see that now Proposition 3.7 has spent itself: there are no more dominated vertices. But in this case, there is no more problem, either! For we have shown that the chromatic number of the full hypergraph corresponding to Table 4.1(a) is equal to the chromatic number of the subhypergraph consisting of the four vertices that are present in Table 4.2(c). But from Table 4.2(c) it is clear that no two of these vertices can have the same colour: they are all mutual enemies. So we will need exactly four colours and we may conclude that the bidimension of the original data matrix equals 4.

This means that the relation underlying the data matrix can be represented as the intersection of four biorders. But these biorders are by no means unique: many collections of four biorders have the given relation as their intersection. In other words, there are many different collections of four dimensions (Guttman scales) that are, under the conjunctive composition rule, compatible with the observed data matrix. By inspecting the content of the four items present in Table 4.2(c),

Chubb manages to attach a verbal label to each dimension on which one of these items scores highest and to construct four dimensions that reflect reasonably well the interpretation suggested by these labels. There is, however, definitely some arbitrariness involved here, which seems inevitable when embedding six items in 4-dimensional space on the basis of binary data. Generally, representations in the minimum dimensionality will be far from unique and we clearly see the need for "best approximate" representations in some lower dimensionality with a higher degree of uniqueness. (Of course, of greatest help would be a strong psychological theory about the data that allows focusing on specific aspects in the collection of representations.)

Returning to our more basic problem of finding the dimensionality of the observed data, we see that by applying Proposition 3.7 we were able to reduce from a hypergraph with 51 vertices based on a 18×6 matrix to one with four vertices based on a 4×4 matrix. This shows the potential power of the reduction principle described in the preceding section. Here we were very lucky indeed; in general, however, there will be a non-trivial problem left when no more reduction is possible. In the next section we discuss a possible approach in that case. This approach will then be illustrated in Section 6, using another data set. There, again, we will make use of Proposition 3.7 whenever we can. Its effects will not be as dramatic as was the case here, but it still will turn out to be very useful.

5. A RECURSION FORMULA FOR THE CHROMATIC NUMBER

With all possible reductions carried out, there will come a moment when the real work has got to start. The problem's being NP-complete suggests that at such a point some sort of exhaustive search will be inevitable. We will describe here some such search in which we use the notion of a *maximal stable set* in a hypergraph, that is, a stable set that is not contained in any other stable set.

For maximal stable sets in an arbitrary hypergraph we can establish the following properties:

5.1. PROPOSITION. (i) *For any vertex v of a hypergraph H there is a minimal colouring of H in which the colour class of v is maximal.*

(ii) *For any maximal stable set M in a hypergraph $\langle V, E \rangle$ we have $\text{Chrom}_{\{M\}} \langle V, E \rangle = 1 + \text{Chrom} \langle V - M, E(V - M) \rangle$.*

Proof. (i) Let C be the colour class of v in a minimal colouring of H . If there is a vertex w , not in C , such that $C \cup \{w\}$ still is stable in H , then the transfer of w to C clearly gives a colouring of H without introducing new colours. This process can be repeated until there are no more candidates, that is, until the class of v is maximal.

(ii) Because M is stable, any $(q-1)$ -colouring of $\langle V - M, E(V - M) \rangle$ can be extended to a q -colouring of $\langle V, E \rangle$ by adding M as a colour class. This proves $\text{Chrom}_{\{M\}} \langle V, E \rangle \leq 1 + \text{Chrom} \langle V - M, E(V - M) \rangle$. For the reverse inequality,

consider a q -colouring of $\langle V, E \rangle$ in which M is monochromatic. Then M , being maximal, must itself be a colour class and we are in fact given a $(q-1)$ -colouring of $\langle V-M, E(V-M) \rangle$. ■

From these properties we easily obtain

5.2. COROLLARY. For any vertex v of a hypergraph $\langle V, E \rangle$ we have

$$\text{Chrom } \langle V, E \rangle = 1 + \min_{M \in \text{MAX}(v)} \{ \text{Chrom } \langle V-M, E(V-M) \rangle \},$$

where $\text{MAX}(v)$ denotes the collection of maximal stable sets in $\langle V, E \rangle$ containing the vertex v .

Proof. By Proposition 5.1(i), $\text{Chrom } \langle V, E \rangle = \min_{M \in \text{MAX}(v)} \{ \text{Chrom}_{\{M\}} \langle V, E \rangle \}$ and applying Proposition 5.1(ii) completes the proof. ■

We see that Corollary 5.2 gives a recursive formula for the chromatic number of a hypergraph in terms of the chromatic numbers of a collection of strictly smaller hypergraphs. So, when applied to the finite, special type of hypergraph that turns up in the bidimension problem, this recursion will give the chromatic number and thereby the bidimension in a finite number of steps. The big question, of course, is whether computation of the bidimension according to this recursion will be feasible in practice. In this context it is worth noting some additional properties of maximal stable sets in a hypergraph associated with a relation:

5.3. PROPOSITION. For $R \subseteq A \times D$ the following two properties hold:

- (i) Any maximal stable set in $H(\bar{R})$ is a biorder between A and D .
- (ii) If M is a maximal stable set in $H(\bar{R})$, then for some $a_0 \in A$, $\bar{R} \cap (\{a_0\} \times D) \subseteq M$ and for some $d_0 \in D$, $\bar{R} \cap (A \times \{d_0\}) \subseteq M$.

Proof. (i) (Also in Doignon *et al.*, 1984, p. 95.) Let M be a maximal stable set in $H(\bar{R})$ and suppose M has a violation of the biorder property: $ad, be \in M$ and $ae, bd \in \bar{M}$. We are going to derive a contradiction by showing that M must contain an edge of $H(\bar{R})$. If both ae and bd are in R , then M contains the 2-edge $\{ad, be\}$. If one, say ae , is in R and the other, bd , is in \bar{R} , then by maximality of M there is an edge included in $M \cup \{bd\}$, which of course contains the vertex bd . Let $bd, a_1d_1, \dots, a_nd_n$ be the corresponding alternating cycle, then the sequence $be, ad, a_1d_1, \dots, a_nd_n$ is an alternating cycle in M . If, eventually, both ae and bd are in \bar{R} , then, by the same argument, there are alternating cycles $bd, a_1d_1, \dots, a_nd_n$ and $ae, a'_1d'_1, \dots, a'_md'_m$ in $M \cup \{bd\}$ and $M \cup \{ae\}$, respectively. But then the sequence $ad, a_1d_1, \dots, a_nd_n, be, a'_1d'_1, \dots, a'_md'_m$ is an alternating cycle in M .

(ii) Consider two elements a_i, a_j in A . If there are $d_1, d_2 \in D$ such that $a_id_1 \in M$ and $a_id_2 \in \bar{M}$, while $a_jd_2 \in \bar{M}$ and $a_jd_2 \in M$, then M certainly is no biorder. So if M is a maximal stable set in $H(\bar{R})$ it is, by part (i), a biorder between A and D , consequently for any pair a_i, a_j we have either $a_id \in M$ implies $a_jd \in M$ for all

$d \in D$, or $a_i d \in M$ implies $a_i d \in M$ for all $d \in D$. In this way M induces a weak ordering on A (transitivity is easily checked) and since A is finite we can find an element a_0 in A that is maximal in this ordering, i.e., for which for any $a' \in A$ and any $d \in D$, $a'd \in M$ implies $a_0 d \in M$. We will show $\bar{R} \cap (\{a_0\} \times D) \subseteq M$ for this a_0 by showing that the set $M \cup (\bar{R} \cap (\{a_0\} \times D))$ is a biorder, which is an equivalent assertion since M is maximal. If $M \cup (\bar{R} \cap (\{a_0\} \times D))$ is not a biorder, then, since M itself is, any violation of the biorder property must involve the element a_0 . In particular there must then be $a' \in A$ and $d' \in D$ such that $a'd'$ is in $M \cup (\bar{R} \cap (\{a_0\} \times D))$ and $a_0 d'$ is not. The former, however, implies $a'd' \in M$ (a_0 and a' clearly being distinct) and thus, by our choice of a_0 , $a_0 d' \in M \subseteq M \cup (\bar{R} \cap (\{a_0\} \times D))$, a contradiction. The proof of the other half of (ii), with the roles of A and D reversed, is completely analogous. ■

In terms of the data matrix, Proposition 5.3(i) states that a maximal stable set of zeros has triangular structure in the matrix. In this perspective 5.3(ii) really is obvious: it asserts that in a row (column) that contains, compared to other rows (columns), a maximal number of elements of the set in question, *all* zeros belong to the set. If not, they could be added without disturbing the triangular structure (i.e., biorderhood) of the set, which consequently would not have been maximal.

Corollary 5.2 poses the problem of finding the relevant collection of maximal stable sets. We want to use the corollary for subhypergraphs $H(V)$, obtained after maximally reducing a hypergraph $H(\bar{R})$. If \bar{R}' is the smallest restriction of \bar{R} that contains V , we may apply Corollary 5.2 to the hypergraph $H(\bar{R}')$, in which case Proposition 5.3(i) gives a translation in terms of maximal biorders contained in \bar{R}' and containing some specified vertex. When applying the corollary to the reduced hypergraph $H(V)$ itself, this equivalence may still be useful if we notice that any maximal stable set in $H(V)$ is the restriction to V of some maximal stable set in $H(\bar{R}')$.

Obviously, the amount of work implied by the recursion formula in 5.2 will depend on the reduction obtained before invoking it, but we must realize that at that moment the reduction process of Section 3 is not simply set aside. Rather, at each step of the recursion, by deleting a maximal stable set from the hypergraph under consideration, we get not only a strictly smaller hypergraph, but also one that is again susceptible of reduction. The removal of a maximal stable set of $H(V)$ leads to the removal of at least one row and one column from the underlying matrix: any maximal stable set of $H(V)$ is the restriction to V of some maximal stable set of $H(\bar{R}')$, and by Proposition 5.3(ii) the latter "includes" at least one column and one row of $[R']$. So, in general, before applying the next step in the recursion, we can further shrink the hypergraph and possibly also the submatrix that it is based on.

Further remarks on the question of turning the results of this section and Section 3 into a reasonably practical algorithm will be made in the last section.

6. EXAMPLE OF FINDING THE BIDIMENSION ON EMPIRICAL DATA

Here we will show how the findings of Sections 3 and 5 can be combined to compute the bidimension of an empirical data matrix. The data set we use is borrowed from Marcovici (1981, p. 122). The context is a signal-detection experiment with several conditions of induced colour-blindness. The obtained data matrix is reproduced in Table 6.1(a). Columns refer to figures to be detected and rows of the matrix correspond to subject \times condition combinations (there were 3 subjects and 7 conditions). Since our only objective here is to compute the dimensionality of this data matrix, we will in the sequel pay no attention to its construction and interpretation; it will simply be considered a given 21×10 (0, 1)-matrix.

We will compute the bidimension of the data matrix as the chromatic number of its associated hypergraph and we will start therefore by reducing this hypergraph as much as possible according to Proposition 3.7. In Table 6.1(a) we have already

TABLE 6.1

(a) Data Matrix from Marcovici (1981) (* Indicates a Row-Implied Zero, Non-Implied Row Patterns are Numbered); (b) Matrix of Non-Implied Row Patterns (Extra Column-Implied Zeros Found, Here Marked by an Underscore); (c) Further Reduction by Removal of Implied Column Patterns (All Implied Zeros Are Detected and Displayed as *)

(a)	A	B	C	D	E	F	G	H	I	J
	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1
	1	1	1	1	1	1	1	1	1	1
(1)	1	1	1	1	1	0	1	1	1	1
(2)	1	1	1	0	1	1	1	1	1	1
(3)	0	1	1	*	1	1	1	1	1	1
	*	1	1	*	1	1	1	1	1	1
	* *	1 1	1 1	* *	1 1	1 1	1 1	1 1	1 1	1 1
(4)	1	1	1	1	0	*	1	1	1	0
(5)	1	1	0	*	0	*	1	1	1	1
(6)	1	1	1	*	0	*	1	1	0	1
(7)	0	1	0	1	*	*	0	1	1	*
(8)	1	1	0	1	*	*	0	0	0	*
(9)	1	1	1	*	*	*	0	0	*	*
	1	1	*	*	*	*	*	*	*	*
(10)	*	0	*	*	*	*	*	1	1	*
	*	*	*	*	*	*	*	1	*	*
(11)	*	*	*	*	*	*	*	0	1	*
	*	1	*	*	*	*	*	*	*	*
	*	*	*	*	*	*	*	*	*	*

(b)	A	B	C	D	E	F	G	H	I	J
1	1	1	1	1	1	0	1	1	1	1
2	1	1	1	0	1	1	1	1	1	1
3	0	1	1	*	1	1	1	1	1	1
4	1	1	1	1	<u>-</u>	*	*	1	1	0
5	1	1	0	*	<u>-</u>	*	*	1	1	1
6	1	1	1	*	<u>-</u>	*	*	1	1	0
7	0	1	0	1	*	*	0	1	1	*
8	1	1	0	1	*	*	<u>-</u>	0	0	*
9	1	1	1	*	*	*	<u>-</u>	0	*	*
10	*	0	*	*	*	*	*	1	1	*
11	*	*	*	*	*	*	*	0	1	*

(c)	A	B	C	D	F	G	H	I	J
1	1	1	1	1	0	1	1	1	1
2	1	1	1	0	1	1	1	1	1
3	0	1	1	*	1	1	1	1	1
4	1	1	1	1	*	1	1	1	0
5	1	1	0	*	*	1	1	1	1
6	1	1	1	*	*	1	1	0	1
7	0	1	0	1	*	0	1	1	*
8	1	1	0	1	*	*	0	0	*
9	1	1	1	*	*	*	0	*	*
10	*	0	*	*	*	*	1	1	*
11	*	*	*	*	*	*	0	1	*

indicated which zeros in the matrix are row-implied and we see that we can restrict our attention to the 11×10 submatrix in Table 6.1(b). Finding here some new column-implied zeros we obtain the 11×9 matrix of Table 6.1(c).

Next, to further exploit Proposition 3.7, we consider the enemy relation in this submatrix for the 15 vertices of the reduced hypergraph. This relation is represented in Table 6.2(a), where, in order to save space, enemies that are no longer contained in the hypergraph are given in listed form. Inspecting the partial order on the rows we find 5 dominated vertices and the data matrix reduces to that of Table 6.2(b). The tableau of the corresponding enemy relation reveals another 3 dominated vertices and we can further reduce to the situation of Table 6.2(c). Inspection shows that now there are no more dominated vertices and so the reduction process stops here.

Proposition 3.7 has enabled us to reduce the original hypergraph containing 82 vertices and based on a 16×10 matrix (not counting perfect one- or zero-patterns) to one having 7 vertices and based on a 7×6 matrix. But in this case we are still left with a problem. From inspection of Table 6.2(c) it is not at all clear which value the chromatic number of the resulting hypergraph has. So we have to invoke Corollary 5.2; that is, we have to choose a vertex of the reduced hypergraph, generate all maximal stable sets in the reduced hypergraph that contain the chosen vertex, and compute successively the chromatic number of the hypergraphs obtained by removing each such maximal stable set from the reduced hypergraph. The chromatic number we are searching for will be the minimum over this set of numbers, increased by one.

We choose, for instance, the vertex $3A$. From Table 6.2(c) we see that there are only two vertices, $2D$ and $7A$, which are not enemies of $3A$. Since $2D$ and $7A$, on their turn, are mutual enemies, it is clear that there are just two maximal stable sets in the reduced hypergraph that contain the vertex $3A$, $\{3A, 2D\}$ and $\{3A, 7A\}$. So, first we will remove the vertices $3A$ and $2D$ from the reduced hypergraph. The resultant submatrix and its enemy relation are displayed in Table 6.3(a).

The important point to be noticed is that, by removing a maximal stable set from the reduced hypergraph, at least one row and one column can be removed from the data matrix and we get a hypergraph in which, again, some vertices may be dominated ones. So we see in Table 6.3(a) that we may remove vertex $8H$, next in Table 6.3(b) $7A$ may be discarded, and as a result we are left with the three vertices in the 3×3 matrix of Table 6.3(c). These vertices, all being mutual enemies, clearly need three different colours. An alternative way of putting this is that any vertex in Table 6.3(c) constitutes a maximal stable set by itself and so, by repeatedly applying Corollary 5.2, we can remove the vertices one after the other, each time increasing the chromatic number by one, until there are no more vertices left. Anyway, the conclusion is the same: if the vertices $3A$ and $2D$ are to have the same colour, we need four colours for the hypergraph represented in Table 6.2(c).

Now we can do the same, taking $\{3A, 7A\}$ as the maximal stable set to start with. The results of this choice are given in Table 6.4.

Apparently we must conclude: if the vertices $3A$ and $7A$ are to have the same

TABLE 6.2

(a) Enemy Relation for the Hypergraph of Table 6.1(c) (Removed Enemies Are Listed, Dominated Rows Are Starred); (b), (c) Further Reductions until No More Dominated Vertices Appear

(a)	1F	2D	3A	4J	5C	6I	7A	7C	7G	8C	8H	8I	9H	10B	11H
* 1F	.	1	1	3D
2D	1	.	.	1	.	.	1	1	1	1	1	1	.	.	4F, 7F, 7J, 8F, 8G, 8J
3A	1	.	.	1	1	1	.	.	.	1	1	1	1	.	4F, 5F, 6F, 8F, 8G, 8J, 9F, 9G, 9I, 9J
4J	.	1	1	.	1	1	3D, 5D, 6D
5C	.	.	1	1	.	1	1	.	9G, 9I, 9J
6I	.	.	1	1	1	.	1	1	1	.	.	.	1	1	7J, 10A, 10C, 10G, 10J, 11A, 11B, 11C, 11G, 11J
7A	.	1	.	.	.	1	1	1	1	.	5D, 6D, 9D, 9I
7C	.	1	.	.	.	1	1	.	3D, 6D, 9D, 9I
* 7G	.	1	.	.	.	1	3D, 5D, 6D
8C	.	1	1	3D, 6D, 9D
8H	.	1	1	.	.	.	1	1	.	3D, 5D, 6D, 10A, 10D
8I	.	1	1	.	.	.	1	1	.	3D, 5D, 6D, 10A, 10D, 11A, 11B, 11D
* 9H	.	.	1	.	1	.	1	1	1	10A, 10C
*10B	1	1	1	1	.	9I
*11H	1	—

(b)	A	C	D	H	I	J	2D	3A	4J	5C	6I	7A	7C	8C	8H	8I
2	1	1	0	1	1	1	2D	.	.	1	.	1	1	1	1	7J, 8J
3	0	1	*	1	1	1	3A	.	.	1	1	1	.	1	1	8J
4	1	1	1	1	1	0	4J	1	1	.	1	1	.	.	.	3D, 5D, 6D
5	1	0	*	1	1	1	5C	.	1	1	.	1	.	.	.	—
6	1	1	*	1	0	1	6I	.	1	1	1	1	1	.	.	7J
7	0	0	1	1	1	*	7A	1	.	.	.	1	.	.	1	5D, 6D
8	1	0	1	0	0	*	*7C	1	.	.	.	1	.	.	.	3D, 6D
							*8C	1	1	3D, 6D
							8H	1	1	.	.	.	1	.	.	3D, 5D, 6D
							*8I	1	1	.	.	.	1	.	.	3D, 5D

(c)	A	C	D	H	I	J	2D	3A	4J	5C	6I	7A	8H
2	1	1	0	1	1	1	2D	.	.	1	.	1	7C, 7J, 8C, 8I, 8J
3	0	1	*	1	1	1	3A	.	.	1	1	1	8C, 8I, 8J
4	1	1	1	1	1	0	4J	1	1	.	1	1	3D, 5D, 6D
5	1	0	*	1	1	1	5C	.	1	1	.	1	—
6	1	1	*	1	0	1	6I	.	1	1	1	1	7C, 7J
7	0	*	1	1	1	*	7A	1	.	.	.	1	5D, 6D, 8I
8	1	*	1	0	*	*	8H	1	1	.	.	1	3D, 5D, 6D

TABLE 6.3

(a) Data Matrix and Corresponding Enemy Relation after Removing Vertices 3*A* and 2*D* from Hypergraph of Table 6.2(c); (b), (c) Reductions by Removing Dominated Vertices

(a)	<i>A</i>	<i>C</i>	<i>H</i>	<i>I</i>	<i>J</i>		4 <i>J</i>	5 <i>C</i>	6 <i>I</i>	7 <i>A</i>	8 <i>H</i>		
	4	1	1	1	1	0	4 <i>J</i>	.	1	1	.	.	—
	5	1	0	1	1	1	5 <i>C</i>	1	.	1	.	.	—
	6	1	1	1	0	1	6 <i>I</i>	1	1	.	1	.	7 <i>C</i> , 7 <i>J</i>
	7	0	*	1	1	*	7 <i>A</i>	.	.	1	.	1	8 <i>I</i>
	8	1	*	0	*	*	*8 <i>H</i>	.	.	.	1	.	—

(b)	<i>A</i>	<i>C</i>	<i>I</i>	<i>J</i>		4 <i>J</i>	5 <i>C</i>	6 <i>I</i>	7 <i>A</i>		
	4	1	1	1	0	4 <i>J</i>	.	1	1	.	—
	5	1	0	1	1	5 <i>C</i>	1	.	1	.	—
	6	1	1	0	1	6 <i>I</i>	1	1	.	1	7 <i>C</i> , 7 <i>J</i>
	7	0	*	1	*	*7 <i>A</i>	.	.	1	.	—

(c)	<i>C</i>	<i>I</i>	<i>J</i>		4 <i>J</i>	5 <i>C</i>	6 <i>I</i>		
	4	1	1	0	4 <i>J</i>	.	1	1	—
	5	0	1	1	5 <i>C</i>	1	.	1	—
	6	1	0	1	6 <i>I</i>	1	1	.	—

TABLE 6.4

(a) Data Matrix and Corresponding Enemy Relation after Removing Vertices 3*A* and 7*A* from Hypergraph of Table 6.2(c); (b), (c) Reductions by Removing Dominated Vertices

(a)	<i>C</i>	<i>D</i>	<i>H</i>	<i>I</i>	<i>J</i>		2 <i>D</i>	4 <i>J</i>	5 <i>C</i>	6 <i>I</i>	8 <i>H</i>	
	2	1	0	1	1	1	2 <i>D</i>	.	1	.	.	1 8 <i>C</i> , 8 <i>I</i> , 8 <i>J</i>
	4	1	1	1	1	0	4 <i>J</i>	1	.	1	1	5 <i>D</i> , 6 <i>D</i>
	5	0	*	1	1	1	5 <i>C</i>	.	1	.	1	—
	6	1	*	1	0	1	6 <i>I</i>	.	1	1	.	—
	8	*	1	0	*	*	*8 <i>H</i>	1	.	.	.	5 <i>D</i> , 6 <i>D</i>

(b)	<i>C</i>	<i>D</i>	<i>I</i>	<i>J</i>		2 <i>D</i>	4 <i>J</i>	5 <i>C</i>	6 <i>I</i>	
	2	1	0	1	1	*2 <i>D</i>	.	1	.	—
	4	1	1	1	0	4 <i>J</i>	1	.	1	5 <i>D</i> , 6 <i>D</i>
	5	0	*	1	1	5 <i>C</i>	.	1	.	—
	6	1	*	0	1	6 <i>I</i>	.	1	1	—

(c)	<i>C</i>	<i>I</i>	<i>J</i>		4 <i>J</i>	5 <i>C</i>	6 <i>I</i>		
	4	1	1	0	4 <i>J</i>	.	1	1	—
	5	0	1	1	5 <i>C</i>	1	.	1	—
	6	1	0	1	6 <i>I</i>	1	1	.	—

colour, we need four colours for the hypergraph of Table 6.2(c). By Corollary 5.2, the chromatic number of this hypergraph is the minimum of the values obtained from these two computations and, by Proposition 3.7 and the reduction principle 3.5, this number equals the chromatic number of the hypergraph in Table 6.1(a) we started with. By the fundamental theorem of Doignon *et al.*, then, we may conclude that the bidimension of the original data matrix equals 4.

7. DISCUSSION

We have seen that determining the dimensionality needed for the representation of a data matrix according to the conjunctive model is, in general, a hard, indeed an NP-hard problem, whereas from a casual inspection of Coombs' (1964) example it first appeared to be an automatically obtained by-product of a constructive scaling procedure. On the basis of the equivalence, given by Doignon *et al.*, between this bidimension and the chromatic number of a certain hypergraph, we have gathered in Sections 3 and 5 some results which seem to make computation of the bidimension feasible in practice, at least for data sets of moderate size. (As we have seen, the computations of Sections 4 and 6 were rather easily carried out with paper and pencil.) So the first thing to do is to combine these findings in an explicit, reasonably efficient algorithm. As for the reduction process, this does not seem to offer many problems: it is already rather explicitly described and illustrated in Section 3 and in the examples of Sections 4 and 6.

The recursion formula of Corollary 5.2 will need more consideration. It poses the problem of computing the collection of all maximal stable sets containing a certain vertex. In the example of the preceding section this problem was easily solved "by inspection," but the general case will be NP-hard. (Without appealing to any reduction mechanism, the recursion still solves the NP-hard problem of finding the bidimension.) So it will be important to devise a practical algorithm for this subproblem and Proposition 5.3(i) may turn out to be useful in this context. At least as important, however, will be trying to avoid needless exhaustive execution of this algorithm. Returning, for instance, to our example in the preceding section, we can see that consideration of the second computation, based on the maximal stable set $\{3A, 7A\}$, was in fact pointless. For, as a result of the first computation, we had established 4 as an upper bound for the chromatic number. In this computation, however, the vertices $3A$, $4J$, $5C$, and $6I$ were put in different maximal stable sets and in Table 6.2(c) we can easily check that these four vertices are all mutual enemies (they are said to form a 4-clique in $G(\bar{R})$). So in any colouring they must have different colours and this establishes 4 as a lower bound and thereby solves the problem.

In general, then, we may use the sequence of maximal stable sets formed in a computation for detecting maximal such cliques from the last set backwards. In this way the search tree induced by the recursion formula of Corollary 5.2 may be pruned considerably. Suppose, for instance, we have a "solution" in five maximal

stable sets and suppose we have traced a 3-clique of enemies in the sets at the last three levels. Then we know that, given our present choice of the maximal stable set at level 2, we can, at the next levels, do not better than we did. Hence we need not consider any alternative choices at the levels 5, 4, or 3; instead we may backtrack to the next alternative at the second level. We can, moreover, try to exploit the freedom we have in choosing a vertex in Corollary 5.2 in order to keep the branching of this tree to a minimum in the first place. Heuristically it seems reasonable to choose a vertex with a maximum number of enemies in the subhypergraph in question, expecting such a vertex to have a minimum number of "surrounding" maximal stable sets. We must remark, after all this, that application of Corollary 5.2 is just one possible way of tackling the problem. Alternative and possibly more efficient procedures may exist, waiting to be developed.

Just computing the dimensionality of a representation will not be very useful, however. What we really want are the very representations. This presents the second problem alluded to in the introduction, that of uniqueness of solutions.

It may be noticed that the manner of determining the bidimension of a relation R as sketched in the previous sections is constructive in the sense that it is easy to derive from the computations at least one possible representation of R in the minimum dimensionality, i.e., we can construct at least one minimal collection of biorders having R as their intersection. For the bidimension is found by completing sequences of alternately reducing a hypergraph to a subhypergraph of non-dominated vertices and removing from this subhypergraph a maximal stable set, starting with the original hypergraph $H(\bar{R})$ and ending when there are no more vertices left. At the moment the bidimension is known, we have executed at least one such computation that annihilates $H(\bar{R})$ in a minimal number of steps. In such a computation any removed dominated vertex has left a dominating vertex in the subhypergraph as a representative, so ultimately any vertex has a representative in one of the maximal stable sets of this sequence. Hence by adding each removed dominated vertex to a set containing a representative for it, we obtain a minimal colouring of the original hypergraph. Now, from a q -colouring of $H(\bar{R})$ we can derive at least one representation of R as the intersection of q biorders: consider for each colour class the collection of biorders contained in \bar{R} and containing that class. This collection is non-empty, since any colour class can be expanded to a maximal stable set and these are biorders by Proposition 5.3(i). Clearly any combination of such biorder extensions of the different colour classes covers \bar{R} ; hence the intersection of the complementary biorders is R . In this way one can find some of the generally many distinct representations in the minimum dimensionality. If one should really want to have them all, the only way seems to be an—in principle—exhaustive trial of all combinations of Bidim R biorders containing R (or their complements contained in \bar{R}).

For interpretative purposes we may think of a reasonable reduction of the uniqueness problem in that we do not want all representations of R , but only those in which the biorders are minimal. Suppose $R = \cap B_i$ for biorders B_i ($i = 1, \dots, q$). Now for each i we can choose a biorder $B'_i \subseteq B_i$ that still includes R and that is minimal

in this respect. Then still $R = \cap B'_i$ and we may prefer the latter representation because "each B'_i is more like R than B_i is" ($B'_i - R$ is a subset of $B_i - R$). In terms of matrices, a biorder corresponds to a $(0, 1)$ -matrix having triangular structure that represents the hypothetical data matrix on one of the latent dimensions. Thus, writing R as the intersection of a number of biorders is equivalent to writing $[R]$ as the direct logical product of the same number of matrices having triangular structure. Now the restriction to representations in minimal biorders amounts to the fact that if we have a 0 in the observed data matrix, then we assume a 0 in the corresponding position in each hypothetical factor matrix, unless the triangular structure of that matrix forces a 1. To give an extreme example, if an observed 0 can be "explained" by 0's in all dimensions we are not going to explain it by a 0 in one dimension and 1's in all other dimensions. If both representations are possible, one can argue that the former is more likely. (Another possible restriction, one that would be less severe, is to representations that are "obedient" as defined by Chubb (personal correspondence). Among other things this means restriction to biorders that are compatible with the partial orders on the rows and columns of the data matrix.)

If we restrict the class of solutions for R to the set of all representations in minimal biorders, then, in the procedure sketched above for obtaining some representations from a minimal colouring of $H(\bar{R})$, we need only consider expansions of the colour classes to maximal stable sets. By taking complements we obtain from such a collection a representation of R in minimal biorders.

Even with the above restrictions in mind we may expect that in some cases finding all solutions is not practically feasible. More important, however, are the theoretical problems posed by the occurrence of multiple solutions. Can we single out any one from these as "the right one"? If we can appeal to some psychological theory underlying the data this could be possible. For cases in which this does not apply we can try to find a formal rationale to grade the various solutions, thus obtaining a "best one." Another approach could be trying to capture in some explicit characterization the essence common to large classes of solutions. The last two questions will in particular be important when we do not really want a solution of the deterministic model, but instead our ultimate goal is to find a "best fitting" solution in some lower dimensionality of a probabilistic version of the conjunctive model.

REFERENCES

- BERGE, C. (1973). *Graphs and hypergraphs*. Amsterdam: North-Holland.
- CHUBB, C. (1986). Collapsing binary data for algebraic multidimensional representation. *Journal of Mathematical Psychology*, **30**, 161-187.
- COGIS, O. (1980). *La dimension Ferrers des graphes orientés*. Thèse, Université Pierre et Marie Curie, Paris.
- COGIS, O. (1982). On the Ferrers dimension of a digraph. *Discrete Mathematics*, **38**, 47-52.
- COOMBS, C. H. (1964). *A theory of data*. New York: Wiley.

- COOMBS, C. H., & KAO, R. C. (1955). *Nonmetric factor analysis*. Engineering Research Bulletin No. 38, Ann Arbor: Univ. of Michigan Press.
- DOIGNON, J.-P., DUCAMP, A., & FALMAGNE, J.-C. (1984). On realizable biorders and the biorder dimension of a relation. *Journal of Mathematical Psychology*, **28**, 73–109.
- DUCAMP, A., & FALMAGNE, J.-C. (1969). Composite measurement. *Journal of Mathematical Psychology*, **6**, 359–390.
- DUSHNIK, B., & MILLER, E. W. (1941). Partially ordered sets. *American Journal of Mathematics*, **63**, 600–610.
- GAREY, M. R., & JOHNSON, D. S. (1979). *Computers and intractability: A guide to the theory of NP-completeness*. San Francisco: Freeman.
- GUTTMAN, L. (1944). A basis for scaling qualitative data. *American Sociological Review*, **9**, 139–150.
- MARCOVICI, S. (1981). *The multidimensional analysis of partially ordered data: Theory, method and applications*. Ph.D. thesis, New York University, New York.
- MAURER, S. B., RABINOVITCH, I., & TROTTER, W. T., JR. (1980). Large minimal realizers of a partial order, II. *Discrete Mathematics*, **31**, 297–313.
- STOUFFER, S. A., GUTTMAN, L., SUCHMAN, E. A., LAZARSFELD, P. F., STARR, S. A., & CLAUSEN, J. A. (1950). *Measurement and prediction; Studies in social psychology in World War II*. Princeton, N.J.: Princeton Univ. Press.
- TROTTER, W. T., JR. (1983). Graphs and partially ordered sets. In L. W. Beineke & R. J. Wilson, *Selected topics in graph theory 2*. London/New York: Academic Press.
- YANNAKAKIS, M. (1982). The complexity of the partial order dimension problem. *SIAM Journal on Algebraic and Discrete Methods*, **3**, 351–358.

RECEIVED: February 20, 1986