Available online at www.sciencedirect.com
SCIENCE DIRECT®

Discrete Applied Mathematics 147 (2005) 91 – 100

DISCRETE APPLIED MATHEMATICS

www.elsevier.com/locate/dam

Conjunctive prediction of an ordinal criterion variable on the basis of binary predictors

L. Lombardi[a, b, *], I. Van Mechelen[a]

[a] *Katholieke Universiteit Leuven, Department of Psychology, Tiensestraat 102, B-3000 Leuven, Belgium*
[b] *Università di Trento, Department of Cognitive Sciences and Education, via Matteo del Ben 5, I-38068 Rovereto (Trento), Italy*

Received 19 November 2001; received in revised form 15 July 2003; accepted 21 June 2004

Abstract

In this paper we propose an empirical prediction method to retrieve, for a given ordinal criterion and a set of binary predictors, a series of nested sets of predictors, each set containing all singly necessary (and, if feasible, jointly sufficient) predictors for a particular criterion value. The method extends a previously developed approach to construct approximate Galois lattice models of binary data. After sketching an outline of the new model and associated algorithm we illustrate our method with an application to real psychological data on the experience of anger.

Keywords: Ordinal criterion prediction; Maximal nested conjunctions; Approximate Galois lattice; HICLAS modeling

1. Introduction

In several empirical applications, the aim is to look for logical relations between a criterion variable and a set of predictor variables. In case both the criterion and the predictors are binary (that is, each variable can take the value of either 0 or 1) different models have been proposed within the context of classification analysis where the problem comes

* Corresponding author. Dipartimento di Scienze della Cognizione e della Formazione, via Matteo del Ben 5, I-38068 Rovereto (Trento), Italy.

E-mail addresses: luigi.lombardi@unitn.it (L. Lombardi), iven.vanmechelen@psy.kuleuven.ac.be (I. Van Mechelen).

0166-218X/$ - see front matter
doi:10.1016/j.dam.2004.06.023

down to finding appropriate classification rules. Examples of models dealing with perfect connections between predictors and criterion have been proposed by several authors [10,21] in the context of Boolean function analysis. In empirical prediction problems, however the condition of perfect equivalence of a criterion with a logical combination of predictors is seldom met because of error-perturbed data and/or incomplete sets of predictor variables; for such cases, various models that allow for imperfect connections between predictors and criterion have been developed (e.g., [14,18,19,15]).

A particular case where logical relations are of use is the situation in which two different types of variables (e.g., binary vs. ordinal, or binary vs. interval) are involved in the prediction problem. In this paper we propose an new empirical method within the domain of combinatorial data analysis [1,11] for the prediction of an ordinal criterion by means of a logical combination of binary predictors. As such, combinatorial data analysis (CDA) is concerned with "the location of arrangements of objects that are optimal for the representation of a given data set and is usually operationalized using a specific loss-function that guides the combinatorial optimization process" [1, p. 5]. Usually, CDA does not postulate any stochastic model as underlying the data set. Typically, CDA is based on a deterministic structure that is fitted to the data while not including an explicit error model. Useful applications of CDA models are within the context of exploratory data analysis where, mainly for descriptive purposes, a researcher may adopt a deterministic model to suggest hypotheses or explanations about the data structure. In the present paper, we will take such a descriptive approach to deal with the prediction problem as mentioned above (although a stochastic extension of our approach would be considered as well—see, e.g. [17]).

We introduce the prediction problem in a more formal way. Let c be an ordinal criterion that takes values on an ordered set $V = \{v_0, v_1, \dots, v_t\}$ (with $v_j \prec v_{j+1}$), and P a set of dichotomous potential predictors $p_1, p_2, \dots, p_m$. For example, in a psychiatric diagnosis scenario the symptom *depressive mood* might be characterized by a 0–6 severity rating scale. On the other hand, symptoms like *motor retardation*, *guilty feelings* and *somatic concern* might be considered as potential binary predictors of the depressive mood-state.

Our method attempts to find a set $\mathscr{R}$ of *nested conjunctive combinations*, where R_j in $\mathscr{R}$ is equivalent to $\bigwedge P_j$, where P_j constitutes a maximal set of singly necessary predictors for v_j (or, if it turns out feasible, a maximal set of singly necessary and jointly sufficient predictors for v_j). More formally, each $R_j \in \mathscr{R}$ ($j = 0, \dots, t$) takes the form $\bigwedge P_j$ with $P_i \subseteq P_{i+1} \subseteq P$ ($i = 0, \dots, t-1$). Note that in this way the order relation $\prec$ on V is also reflected by the hierarchical structure of the prediction rules.

The nested conjunctive combinations looked for have particular substantive relevance in several contexts where the intensity of a particular attribute may be conjectured to be theoretically related to a sequence of nested sets of more basic features [9,22,20,6]. For example, a quantitative psychological dimension, such as psychiatric sickness, often may be conjectured to correspond to a sequence of nested sets of symptoms which is such that the symptoms associated with *b* include the symptoms associated with *a* whenever *b* represents a more severe level of sickness than *a*.

It is straightforward to show that Galois lattice techniques [2,8,27,5] can be used in the prediction problem as sketched above, provided an appropriate recoding of the ordinal criterion under study into a set of dummy variables (see Section 3). Traditional Galois lattices, however, yield exact representations of data sets. The latter may be troublesome in empirical prediction problems for two types of reasons: First, in case of data sets that are not very small, the high complexity of the associated lattices can be very difficult to manage. Second, in case of error perturbed data, exact representations may be simply undesirable.

A possible way out consists of approximating the data matrix with a model matrix that has a much simpler lattice representation. A method to construct such *approximate Galois lattices*, and with which the present paper immediately links up, has been proposed by Van Mechelen [24], and Van Mechelen et al. [25]. Unfortunately, classical approximate Galois lattices also suffer from some limitations when applied to the ordinal prediction problems sketched above. In particular: (1) there is no guarantee that they correctly represent the natural order relation implied by the ordinal target criterion, (2) in the construction of classical approximate lattices of [criterion + predictor] variable data, criteria and predictors equally contribute to the loss function; such a symmetric treatment, however, may be undesirable given that, whereas the criterion is always important, the predictor set may include both relevant and irrelevant predictors. Note in this respect that, within a Boolean prediction context, a predictor is defined irrelevant for a given criterion whenever all four possible logical combinations: (1) "false positive": $p = 1$ and $c = 0$ (2) "false negative": $p = 0$ and $c = 1$ (3) "true positive": $p = 1$ and $c = 1$ and (4) "true negative": $p = 0$ and $c = 0$, show up in a non-negligible number of cases. For example, in a psychiatric diagnosis scenario the symptom *suspiciousness* might be an irrelevant predictor for a diagnosis of simple schizophrenia in that this symptom could be partially present in schizophrenic as well as in non-schizophrenic patients.

The method to be presented in this paper, which is called OPHICLAS (hiclas model for prediction of an ordinal criterion) is a novel extension of the approximate Galois approach to deal with the problem of predicting an ordinal criterion variable on the basis of binary predictors. Interestingly this new method has two important features:

(1) it guarantees that the order relation in V is well represented in the approximate Galois lattice;
(2) in the data analysis, it allows one to obtain a good representation of the criterion in the lattice in the presence of both relevant and irrelevant predictors in the data.

To provide a self-contained exposition, the next section (Section 2) briefly recapitulates the main aspects of the approximate Galois lattice method. Section 3 proceeds with showing in which way the novel OPHICLAS method can be adopted to solve the prediction problem. Finally, Section 4 illustrates the new method with an application to real psychological data.

2. Approximate Galois lattices

Let S be a binary relation over the Cartesian product $O \times A$, where O and A are a finite set of objects and a finite set of binary attributes, respectively. Moreover, let f and g be two mappings: $f(X) = \{y \in A : \forall x \in X, (x, y) \in S\}$ and $g(Y) = \{x \in O : \forall y \in Y, (x, y) \in S\}$. Both mappings constitute a Galois connection between 2^O and 2^A (resp. 2^A and 2^O). If we define $\gamma_O = g \circ f$ and $\gamma_A = f \circ g$ then the set $\mathfrak{S}$ of all maximal rectangles (also called formal concepts) of S is defined as

$$\mathfrak{S} = \{(\gamma_O(X), f(X)) : X \subseteq O\} = \{(g(Y), \gamma_A(Y)) : Y \subseteq A\}.$$

$\mathfrak{S}$ constitutes a lattice $\mathscr{L} = \langle \mathfrak{S}, \preccurlyeq \rangle$, where $(X_1, Y_1) \preccurlyeq (X_2, Y_2) \Leftrightarrow X_1 \subseteq X_2 (\Leftrightarrow Y_2 \subseteq Y_1)$, with *join* and *meet* represented by

$$(X_1, Y_1) \smile (X_2, Y_2) \equiv (\gamma_O(X_1 \cup X_2), Y_1 \cap Y_2),$$

$$(X_1, Y_1) \frown (X_2, Y_2) \equiv (X_1 \cap X_2, \gamma_A(Y_1 \cup Y_2)),$$

respectively. The lattice $\mathscr{L}$ is called *Galois lattice* [2] (or *formal concept lattice* [27,5]) and $\mathscr{L}_O = \langle \mathfrak{S}_O, \subseteq \rangle$ (with $\mathfrak{S}_O = \{\gamma_O(X) : X \subseteq O\}$) and $\mathscr{L}_A = \langle \mathfrak{S}_A, \subseteq \rangle$ (with $\mathfrak{S}_A = \{\gamma_A(Y) : Y \subseteq A\}$) are called its object projection lattice and attribute projection lattice, respectively.

$\mathscr{L}_A$ represents implication relations between an attribute and conjunctive combinations of other attributes. In general, if $\gamma_A(\{a_1\}) \supseteq \gamma_A(\{a_2, \ldots a_m\})$ then a_1 implies the conjunction of $a_2, \ldots, a_m$. Moreover, if $\gamma_A(\{a_1\}) = \gamma_A(\{a_2, \ldots a_m\})$ then the conjunction of $a_2, \ldots, a_m$ is equivalent with a_1. We denote these two types of relations by *simple implication* ($a_1 \Rightarrow a_2 \wedge \cdots \wedge a_m$) and *perfect equivalence* ($a_1 \Leftrightarrow a_2 \wedge \cdots \wedge a_m$), respectively.

Galois lattices of empirical binary relations are typically very complex, even with data sets of moderate size. Moreover in case of error perturbed data an exact representation of the lattice could be troublesome. In this latter case, small changes in the original data set may lead to considerable changes in the set of maximal rectangles (resp. formal concepts) and therefore in the lattice structure. Within a lattice context several authors have proposed non-stochastic solutions to deal with error perturbed data [4,7,23]. In particular, one method to deal with both the complexity problem and the problem of error-perturbed data, and with which the present paper immediately links up, has been proposed by Van Mechelen [24], and Van Mechelen et al. [25]. This method is based on the idea of *approximate Galois lattices* and does not rely on external information to retrieve more robust lattice structure.

An approximate Galois lattice with restricted size and length of a given binary relation $S \subseteq O \times A$ can be obtained by performing a conjunctive HICLAS analysis on the incidence matrix $\mathbf{D}$ of S [24,25]. Such an analysis [25] approximates $\mathbf{D}$ by a model matrix $\mathbf{M}_r$ that can be decomposed into two binary matrices $\{\mathbf{S}_r, \mathbf{P}_r\}$ by the following rule:

$$\mathbf{D} \approx \mathbf{M}_r = [\mathbf{S}_r^C \otimes \mathbf{P}_r']^C, \tag{1}$$

where $\otimes$, ^{C}, $\mathbf{P}_r'$ and r denote the Boolean matrix product [12], complement, transposed of matrix $\mathbf{P}_r$ and rank of the model decomposition (that is, the number of columns of $\mathbf{S}_r$ and $\mathbf{P}_r$), respectively. In particular, $\mathbf{S}_r$ (resp. $\mathbf{P}_r$) defines r possibly overlapping clusters of objects (resp. attributes). The approximation is done such that for a fixed rank r the loss function

$$E = \sum_i \sum_j |d_{ij} - m_{ij}| \tag{2}$$

is minimized. It is easy to show that the Galois lattice of $\mathbf{M}_r$ as defined in (1) has size $\leqslant 2^r$ and length $\leqslant r$ [25].

With regard to the problem of error-perturbed data it may be worthwhile to note that Leenen and Van Mechelen [16] did an extensive simulation study with 'true' matrices $\mathbf{T}$ to which error was systematically added, resulting in error-perturbed data matrices $\mathbf{D}$. In this study, it was found that the HICLAS algorithm that optimized loss function (2) had an excellent goodness of recovery, in that $\mathbf{M}_r$ was usually very close to $\mathbf{T}$. Therefore, the approximate Galois lattices resulting from an HICLAS procedure may at least in part, be looked at as an intended recovery of underlying true structures.

3. Conjunctive prediction of an ordinal criterion variable

3.1. Standard approximate Galois lattice approach

We introduce some other basic notation. Let O and A be a set of objects and a set of attributes, respectively. Moreover, we assume that the set of attributes is partitioned into two distinct sets P and C, where, in particular, P denotes a set $\{p_1, p_2, \ldots, p_m\}$ of Boolean predictors and C a singleton $\{c\}$ containing an ordinal criterion variable that takes values on an ordered set $V = \{v_0, v_1, \ldots, v_t\}$, respectively. Let $\mathbf{D}_P$ be the $n \times m$ incidence matrix of the relation $S \subseteq O \times P$ (where $(x, y) \in S$ iff Object x has positive value on Predictor y) and $\mathbf{d}_c$ the $n \times 1$ ordinal criterion vector.

The final goal of an OPHICLAS analysis is to find a collection $\mathscr{R}$ of nested maximal conjunctive combinations, where each R_j in $\mathscr{R}$ is equivalent to $\bigwedge P_j$, where P_j constitutes a maximal set of singly necessary predictors for v_j (or, if it turns out feasible, a maximal set of singly necessary and jointly sufficient predictors for v_j).

In order to extract the set of prediction rules from the n (objects) $\times(m+1)$ (attributes) data matrix $\mathbf{D} = \mathbf{D}_P \cup \mathbf{d}_c$ we first recode $\mathbf{D}$ into an $n \times (m + t)$ Boolean data matrix $\mathbf{D}^*$ which is defined as follows:

$$d^*_{il} = \begin{cases} d_{il} & \text{if } l \leqslant m, \\ 1 & \text{if } l > m \text{ and } d_{i(m+1)} \geqslant v_l, \\ 0 & \text{if } l > m \text{ and } d_{i(m+1)} < v_l, \end{cases} \tag{3}$$

for all $i = 1, \ldots, n$ and for all $l = 1, \ldots, m + t$. Next, the matrix $\mathbf{D}^*$ is subjected to a rank-r approximate Galois lattice analysis. This implies the search of a set of bundles $\{\mathbf{S}^*, \mathbf{P}^*\}$ that yield a model matrix $\mathbf{M}^*$ via the association rule (1) which is such that the loss function $E^* = \sum_i \sum_l |d^*_{il} - m^*_{il}|$ is minimized. Finally, the Galois lattice $\mathscr{L}^*_A$ of the model matrix $\mathbf{M}^*$ has to be constructed.

The collection $\mathscr{R} = \{R_1, \ldots, R_t\}$ of nested maximal conjunctive combinations can be derived in the following way: Let C^* be the set $\{c_0, \ldots, c_t\}$ of recoded dummy variables for the ordinal criterion c, $A^* = P \cup C^*$ the new set of binary attributes and $S^* \subseteq O \times A^*$ the new relation as defined by (3), respectively. Each $R_j \in \mathscr{R}$ is equal to $\bigwedge P_j$, where $P_j = P \cap \gamma_{A^*}(c_j)$. Notice that, given that $\gamma_{A^*}(P_j) \subseteq \gamma_{A^*}(c_j)$, the implication $v_j \Rightarrow \bigwedge P_j$ always holds for all j. Furthermore, if $c_j \in \gamma_{A^*}(P_j)$ then a perfect equivalence $v_j \Leftrightarrow \bigwedge P_j$ holds, that is, P_j then constitutes the maximal set of singly necessary and jointly sufficient predictors for v_j.

3.2. New OPHICLAS approach

In case of the conjunctive prediction of an ordinal criterion variable the standard approximate lattice approach suffers from two important limitations:

(1) It does not necessarily correctly represent the natural order $v_i \preccurlyeq v_{i'}$ $(\forall i' \leqslant i)$.
(2) It does not imply any kind of tool in order to discriminate between relevant and irrelevant predictors.

In this section we propose an algorithmic strategy which solves the limitations of the original conjunctive HICLAS algorithm and provides a procedure to perform a conjunctive prediction of an ordinal criterion variable on the basis of binary predictors.

The OPHICLAS algorithm generalizes the conjunctive HICLAS algorithm as described by Van Mechelen et al. [25] and can be split into four distinct and consecutive steps:

S1: For a fixed $\theta \geqslant 1$ derive a new adjusted loss function

$$E^+ = \overbrace{\sum_{i=1}^{n}\sum_{j=1}^{m} |d^*_{ij} - m^*_{ij}|}^{\text{predictors component}} + \theta \overbrace{\sum_{i=1}^{n}\sum_{j=m+1}^{m+t} |d^*_{ij} - m^*_{ij}|}^{\text{criterion component}}. \tag{4}$$

In this way, we provide more weight to the criterion component, which implies a better model representation for the criterion itself, with less influence on the loss function E^+ by possibly irrelevant predictors.

S2: Apply a *constrained* r-rank conjunctive HICLAS analysis to $\mathbf{D}^*$ in order to obtain the set of bundles $\{\mathbf{S}^*, \mathbf{P}^*\}$ that reconstructs the model matrix $\mathbf{M}^* \approx \mathbf{D}^*$ via the association rule (1) in such a way that the loss function E^+ is minimized. With a *constrained* conjunctive HICLAS analysis we mean a modified version of the original conjunctive HICLAS algorithm [25,16] that satisfies the property

$$\mathbf{p}^*_j \geqslant \mathbf{p}^*_{j'} \quad \forall j \forall j' \in \{m+1, \ldots, m+t\} \quad \text{s.t. } j > j'. \tag{5}$$

Note that the addition of a constraint in Step 2 is necessary because an unconstrained conjunctive HICLAS analysis of a recoded data matrix $\mathbf{D}^*$ may yield bundle matrices $\{\mathbf{S}^*, \mathbf{P}^*\}$ with the submatrix $\mathbf{P}^*_c = (\mathbf{p}^*_j : j = m+1, \ldots, m+t)$ of $\mathbf{P}^*$ not inversely representing the natural order on V.

The routine which looks for matrices $\{\mathbf{S}^*, \mathbf{P}^*\}$, such that (4) is minimal and (5) is satisfied, is a constrained version of the alternating greedy procedure for HICLAS analyses (see [16]). In particular, the submatrix $\mathbf{P}^*_c$ is estimated via a greedy procedure which, given $\mathbf{S}^*$, first looks for the value bundle pattern $\mathbf{p}^*_{m+t}$ that minimizes (4). In the next steps $\mathbf{p}^*_{j-1}$ is estimated conditionally upon $\mathbf{p}^*_j$, that is, $\mathbf{p}^*_{j-1}$ is chosen such as to minimize

$$\sum_{i=1}^{n} |d_{i(j-1)} - (\mathbf{s}^{*C}_i \mathbf{p}^*_{j-1})^C|$$

subject to the constraint that $\mathbf{p}^*_{j-1} \leqslant \mathbf{p}^*_j$ ($j = m+t-1, \ldots, m+2$).

S3: Derive the attribute Galois lattice $\mathscr{L}^*_A$ from $\mathbf{M}^*$.

S4: Extract the set $\mathscr{R}$ of nested maximal conjunctive combinations from $\mathscr{L}^*_A$.

4. An empirical application

In this section we illustrate the new approach with an example in the field of emotion concepts. According to several authors in this field, emotion concepts can be defined by a set of singly necessary and jointly sufficient *semantic primitives*, which are “terms of words which are intuitively understandable (nontechnical), and which themselves are not names of specific emotions or emotional states" (see [26, p. 541]). Table 1 lists some of the semantic primitives that have been investigated in a pilot study on irritation and anger [13]. If different levels of an emotion correspond to conjunctive combinations of semantic primitives, it may be appropriate to apply an OPHICLAS analysis to data on the applicability of semantic primitives (as predictors) and the intensity of a particular emotion (as the ordinal criterion) in a particular situation.

67 students from the University of Leuven were each asked to judge one selected negative social situation. In particular, the subjects were asked: (1) to specify whether or not each of the 21 semantic primitives in Table 1 was true in the situation and (2) to rate the self-experienced anger-induced by the situation on a 4-point scale (v_0: I feel no anger at all; v_1: I feel a bit of anger; v_2: I feel fairly strong anger; v_3: I feel very strong anger). This resulted in a $67 \times (21+1)$ subject by predictor+criterion matrix $\mathbf{D}$. The application of the recoding rule (2) to $\mathbf{D}$ provided a $67 \times (21+4)$ Boolean data matrix $\mathbf{D}^*$. Next, this new matrix was analyzed by means of the OPHICLAS algorithm in ranks 1–7 and $\theta = 1$–3. Taking into account the complexity of the resulting structures as well as the rank by weighted percentage of discrepancies plots, the rank 5 solution with $\theta=2$ and a weighted percentage of discrepancies of 7.5% was retained. Fig. 1 shows the Galois lattice $\mathscr{L}_A^*$ of the $(r=5, \theta=2)$ model, with the empty boxes denoting the conjunctions of the respective superordinate nodes. It is important to note that the exact Galois lattice of $\mathbf{D}^*$ yielded a much more complex lattice representation with 836 maximal rectangles (formal concepts).[1]

Table 1
List of the predictors (P) for the OPHICLAS analysis in the application

P	Class	Semantic primitives
p_1	$[a]$	I want to change something
p_2	$[a]$	The situation is unpleasant
p_3	$[a]$	The situation is unexpected
p_4	$[b]$	I could give up
p_5	$[b]$	I could lose control
p_6	$[b]$	I feel uncertain
p_7	$[d]$	I feel frustrated
p_8	$[d]$	I feel dissatisfied
p_9	$[d]$	It is a possibly negative situation
p_{10}	$[e]$	I would act aggressively
p_{11}	$[f]$	I could find obstacles to my goals
p_{12}	$[g]$	I feel injustice
p_{13}	$[i]$	My position could be threatened
p_{14}	$[i]$	I feel a limitation of freedom
p_{15}	$[j]$	My self-esteem could be threatened
p_{16}	$[k]$	Some norms have been violated
p_{17}	$[k]$	Someone is acting on purpose
p_{18}	$[k]$	Someone is not acting on purpose
p_{19}	$[k]$	The situation is due to other persons
p_{20}	$[k]$	Someone is acting arrogant
p_{21}	$[k]$	Someone is over affording

[1] We used the software ConImp, written by Burmeister [3], in order to get the size of the exact Galois lattice (or formal concept lattice) of $\mathbf{D}^*$.

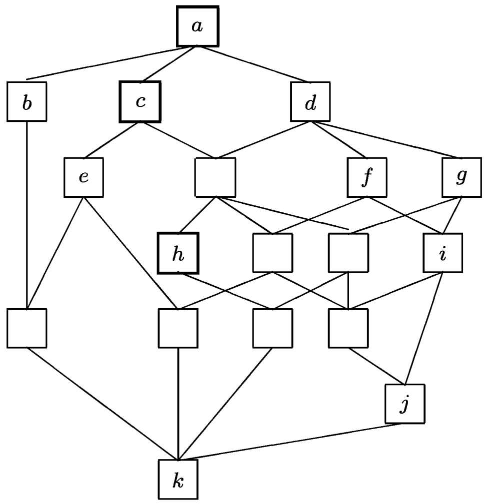

Fig. 1. Approximate Galois lattice $\mathscr{L}_A^*$ of emotion concept data. Labeled classes contain the predictors as indicated in Table 1; regarding the criterion values $v_0, v_1 \in [a]$, $v_2 \in [c]$ and $v_3 \in [h]$ (marked boxes).

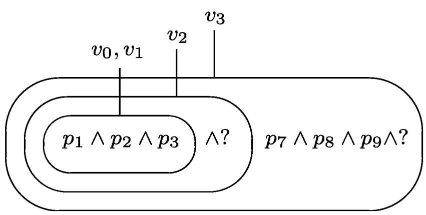

Fig. 2. Maximal nested conjunctions: $R_0 = R_1 \subseteq R_2 \subseteq R_3$.

$\mathscr{L}_A^*$ contains 11 labeled classes (see Table 1 and Fig. 1). Fig. 2 illustrates the nested predictive combinations of $\mathscr{L}_A^*$. $\mathscr{R}$ contains two sets $R_0 = R_1$ of singly necessary and jointly sufficient predictors and two sets R_2, R_3 of simply singly necessary predictors. In substantive terms the rules read as follows: A person reports (s)he experiences a bit of anger (or no anger) ($R_0 = R_1$) in a given situation iff "the situation is unpleasant *and* unexpected *and* (s)he wants to change something". For fairly strong anger (R_2) these conditions are only singly necessary and no longer jointly sufficient (this is indicated in Fig. 2 with a question mark "?"). Finally, for a person reporting very strong anger (R_3) in the given situation, in addition to the previous conditions, it is also necessary that (s)he feels frustrated *and* unsatisfied *and* that (s)he thinks that the situation could have negative consequences.

4.1. Concluding remark

In several empirical contexts like, for example, clinical and social psychology [9,6], it is important to look for maximal sets of singly necessary predictors of a given target criterion.

However, in other situations researchers could be interested in looking for sufficient conditions. Notice that maximal sets of singly sufficient predictors can also be easily derived from $\mathscr{L}_A^*$. However, while for the necessity condition $v_j \Rightarrow \bigwedge P_j$ the uniqueness of P_j always holds, this is not necessarily true for the sufficiency condition $v_j \Leftrightarrow \bigwedge P_j$. In fact, in this latter case, several sets $P_j^{(1)}, \ldots, P_j^{(u)}$ of predictors could be distinctively sufficient for v_j.

Acknowledgements

The research reported in this paper was supported in part by Grant GOA/00/02 of the Research Council of K.U. Leuven. We also are grateful to Peter Kuppens for providing us with the data set on contextualized negative emotions in personality.

References

[1] P. Arabie, L.J. Hubert, An overview of combinatorial data analysis, in: P. Arabie, L.J. Hubert, G. De Soete (Eds.), Clustering and Classification, World Scientific, River Edge, NJ, 1996, pp. 5–63.
[2] M. Barbut, B. Monjardet, Ordre et Classification: Algèbre et Combinatoire 2 Vols, Hachette, Paris, 1970.
[3] P. Burmeister, ConImpf—ein programm zur formalen begriffsanalyse, in: G. Stumme, R. Wille (Eds.), Begriffliche Wissensverarbeitung: Methoden und Anwendungen, Springer, Berlin, 2000, pp. 25–56.
[4] V. Duquenne, The core of finite lattices, Discrete Math. 88 (1991) 133–147.
[5] B. Ganter, R. Wille, Formal Concept Analysis: Mathematical Foundations, Springer, Berlin, 1999.
[6] I. Gati, A. Tversky, Representation of qualitative and quantitative dimensions, J. Experiment. Psychol.: Human Perception Perform. 8 (1982) 325–340.
[7] A. Guenoche, I. Van Mechelen, Galois approach to the induction of concepts, in: I. Van Mechelen, J. Hampton, R. Michalski, P. Theuns (Eds.), Categories and Concepts: Theoretical Views and Inductive Data Analysis, Academic Press, London, 1992, pp. 287–308.
[8] J.-L. Guigues, V. Duquenne, Families minimales d'implications informatives resultant d'un tableau de données binaires, Math. Inform. Sci. Humaines 95 (1986) 5–18.
[9] L. Guttman, A new approach to factor analysis: the radex, in: P.F. Lazarsfeld (Ed.), Mathematical Thinking in the Social Sciences, Free Press, New York, 1954.
[10] A.K. Halder, Grouping table for the minimisation of *n*-variable Boolean functions, Proc. Institut. Electric Eng. London 125 (1978) 474–482.
[11] L. Hubert, P. Arabie, J. Meulman, Combinatorial Data Analysis: Optimization by Dynamic Programming, SIAM, Philadelphia, 2001.
[12] K.H. Kim, Boolean Matrix Theory, Marcel Dekker, New York, 1982.
[13] P. Kuppens, I. Van Mechelen, D.J.M. Smits, P. De Boeck, The appraisal basis of anger and irritation: Specificity, necessity and sufficiency, Emotion 3 (2003) 254–269.
[14] G.S. Lbov, Logical functions in the problem of empirical prediction, in: P.R. Krishnaiah, L.N. Kanal (Eds.), Handbook of Statistics, vol. 2, North-Holland, Amsterdam, 1975, pp. 479–491.
[15] I. Leenen, I. Van Mechelen, A branch-and-bound algorithm for Boolean regression, in: I. Balderjahn, R. Mathar, M. Schader (Eds.), Data Highways and Information Flooding, a Challenge for Classification and Data Analysis, Springer, Berlin, 1998, pp. 164–171.
[16] I. Leenen, I. Van Mechelen, An evaluation of two algorithms for hierarchical classes analysis, J. Classification 18 (2001) 57–80.
[17] I. Leenen, I. Van Mechelen, A. Gelman, Bayesian probabilistic extensions of a deterministic classification model, Comput. Statist. 15 (2000) 355–371.
[18] D.P. McKenzie, D.M. Clarke, L.H. Low, A method of constructing parsimonious diagnostic and screening tests, Int. J. Methods Psychiatric Res. 2 (1992) 71–79.

[19] C.C. Ragin, S.E. Mayer, K.A. Drass, Assessing discrimination: a Boolean approach, Am. Sociol. Rev. 49 (1984) 221–234.
[20] F. Restle, Psychology of Judgment and Choice, Wiley, New York, 1961.
[21] M. Sen, Minimization of Boolean functions of any number of variables using decimal labels, Inform. Sci. 30 (1983) 37–45.
[22] S.S. Stevens, On the psychophysical law, Psychol. Rev. 64 (1957) 153–181.
[23] B. Stoehr, R. Wille, Formal concept analysis of data with tolerances, Preprint 1401, Technische Hochschule, Darmstadt, 1991.
[24] I. Van Mechelen, Approximate Galois lattices of formal concepts, in: O. Opitz, B. Lausen, R. Klar (Eds.), Information and Classification: Concepts, Methods and Applications, Springer, Berlin, 1992, pp. 108–112.
[25] I. Van Mechelen, P. De Boeck, S. Rosenberg, The conjunctive model of hierarchical classes, Psychometrika 60 (1995) 505–521.
[26] A. Wierzbicka, Defining emotion concepts, Cognitive Sci. 16 (1992) 539–581.
[27] R. Wille, Restructuring lattice theory: an approach based on hierarchies of concepts, in: O. Rival (Ed.), Ordered Sets, Reidel, Boston, 1982, pp. 445–470.