IFAC/IFORS/IIASA/TIMS

The International Federation of Automatic Control
The International Federation of Operational Research Societies
The International Institute for Applied Systems Analysis
The Institute of Management Sciences

SUPPORT SYSTEMS FOR DECISION AND NEGOTIATION PROCESSES

Preprints of the IFAC/IFORS/IIASA/TIMS Workshop

Warsaw, Poland

June 24-26, 1992

Editors:

Roman Kulikowski
Zbigniew Nahorski
Jan W.Owsiński
Andrzej Straszak

Systems Research Institute
Polish Academy of Sciences
Warsaw, Poland

VOLUME 1:

Names of first authors: **A-K**

SYSTEMS RESEARCH INSTITUTE, POLISH ACADEMY OF SCIENCES

SUPPORT SYSTEMS FOR DECISION AND NEGOTIATION PROCESSES
Preprints, IFAC/IFORS/IIASA/TIMS Workshop, June 24-26, 1992, Warsaw, Poland

FACTORIAL AXIS INTERPRETATION BY SYMBOLIC OBJECTS

Edwin DIDAY
University Paris IX Dauphine /INRIA - Rocquencourt - FRANCE

Mireille GETTLER-SUMMA
Université Paris IX Dauphine - FRANCE

Abstract

The main aim of the symbolic approach in data analysis is to extend problems, methods and algorithms used on classical data to more complex data called "symbolic objects"which are well adapted to representing knowledge and which can "unify" unlike usual observations which characterize "individual things". We focus here on boolean and probabilist objects and we briefly present some of their qualities and properties. We finally develop in the context of symbolic analysis of a classical data table, a factorial axis characterisation as a probabilist object; it completes the usual vectorial representation which is not so explicit for the standard user. We particularly show the application of learning algorithms to explain multiple correspondence analysis axis which are so useful for enquiry treatments.

Key-words : Knowledge analysis, symbolic data analysis, uncertainty logic, factorial analysis, interpretation aids

Introduction

If we wish to describe the fruits produced by a village, by the fact that "The weight is between 300 and 400 grammes and the color is white or red and if the color is white then the weight is lower than 350 grammes", it is not possible to put this kind of information in a classical data table where rows represent villages and columns descriptors of the fruits. This is because there will not be a single value in each cell of the table (for instance, for the weight) and also because it will not be easy to represent rules (if..., then...) in this table. It is much easier to represent this kind of information by a logical expression such as :

a_i = [weight = [300,400]] $\wedge$ [color = {red, white}] $\wedge$ [if [color = white] then [weight $\leq$ 350]],

where a_i, associated to represents the ith village, is a mapping defined on the set of fruits such that for a given fruit w, a_i (w) = true if the weight of w belongs to the interval [300,400], its color is red or white and if it is white then its weight is less than 350 gr. Following the terminology of this paper a_i is a kind of symbolic object. If we have a set of 1000 villages represented by a set of 1000 symbolic objects $a_1,..., a_{1000}$, an important problem is to know how to apply statistical methods to statistics on it. For instance, what is a histogram or a probability law for such a set of objects ? The aim of symbolic data analysis (Diday 1990,1991) is to provide tools for answering this problem.

In some fields a boolean representation of the knowledge (a_i (w) = true or false) is sufficient to get the main information, but in many cases we need to include uncertainty to represent the real world with more accuracy. For instance, if we say that in the ith village "the color of the fruits is often red and seldom white" we may represent this information by a_i = [color = often red, seldom white]. More generally, in the case of boolean objects or objects where frequency appears, we may write a_i = [color = q_i] where q_i is a characteristic function in the boolean case, and a probability measure in the second case. More precisely, in the boolean case, if a_i = [color = red, white] we have q_i (red) = q_i (white) = 1 and q_i = 0, for the other colors ;in the probabilist case, if a_i = [color = 0.9 red, 0.1 white] we have q_i (red) = 0.9, q_i (white) = 0.1.

[color = red, white] we have q_i (red) = q_i (white) = 1 and q_i = 0, for the other colors ;in the probabilist case, if a_i = [color = 0.9 red, 0.1 white] we have q_i (red) = 0.9, q_i (white) = 0.1. If an expert says that the fruits are red we may represent this information by a symbolic object a_i = [color = q_i] where q_i is a "possibilist" function in the sense of Dubois and Prade (1986) ; we will have for instance q_i (white) = 0, q_i (pink) =0.5 and q_i (red) = 1. If an expert who has to study a representative sample of fruits from the ith village, says that 60% are red, 30% are white and the color is unknown for 10% which were too rotten, we may represent this information by a_i = [color = q_i] where q_i is a belief function such that q_i (red) = 0.6, q_i (white) = 0.3 and q_i (O) = 1,where O is the set of possible colors. Depending on the kind of the mapping q_i used, a_i has been called a boolean, probabilist, possibilist or belief object. In all these cases a_i is a mapping from Ω in [0,1]. Now, the problem is to know how to compute a_i (w) ; if there is doubt about the color of a given fruit w, for instance, if the expert says that "the color of w, is red or pink" then, w may be described by a charateristic function r and represented by a symbolic object w^s = [color = r] such that r (red) = r (pink) = 1 and r = 0 for the other colors. Depending on the kind of knowledge that the user wishes to represent, r may be a probability, possibility or belief function. Having a_i =[color = q_i] and w^s = [color = r] to compute a_i (w) we introduce a comparison function g such that a_i (w)= g (q_i,r) measures the fit between q_i and r. What is the meaning of a_i (w) ? May we say that a_i (w) measures a kind of probability, possibility or belief that w belongs to the class of fruits described by a_i when q_i and r are respectively charateristic, probability, possibility or belief functions ? To answer this question we have extended a_i to a "dual" mapping a_i^* (such that a_i (w) = a_i^* (w^s)) defined on the set of symbolic objects of the a_i kind denoted $\mathfrak{a}_x$ and an extension of the union, intersection and complementary operators of classical sets denoted $OP_x = \{\cup_x, \cap_x, c_x\}$ where x depends upon the kind of knowledge used; then, we have shown that when x represents probability , then a_i^* satisfies the axioms of probability measures by using OP_{pr}(x= probability) and in the case of possibilist objects that a_i^* satisfies the axioms of possibility functions by using some given operators denoted OP_{pos} (see Diday (1991) for more details).

In probability theory , very little is said about events which are generally identified as parts of the sample set Ω. In computer science, object oriented languages consider more general events called objects or "frames" defined by intention. In data analysis (multidimensional scaling, clustering, exploratory data analysis etc.) more importance is given to the elementary objects which belong to the sample Ω than in classical statistics where attention is focused on the probability laws of Ω ; however, objects of data analysis are generally identified to points of $\mathbb{R}^p$ and hence are unable to treat complex objects coming for instance from large data bases, and knowledge bases. Our aim is to define complex objects called "symbolic objects" inspired by those of oriented object languages in such a way that data analysis becomes generalized in knowledge analysis. Objects will be defined by intention by the properties of their extension. More precisely, we distinguish objects which "unify" rather than elementary observed objects which characterize "individual things" (their extension) : for instance "the customers of my shop" instead of "a customer of my shop", "a species of mushroom" instead "the mushroom that I have in my hand".
We have not used the notion of "predicates" from classical logic, firstly, because by using only functions, things seem more understandable, especially to statisticians ; secondly, because they cannot be used simply in the case of probabilist, possibilist and belief objects where uncertainty is present.

1. Boolean symbolic objects

We consider Ω a set of individual things called "elementary objects" and a set of descriptor functions $y_i : \Omega \rightarrow O_i$.

A basic kind of symbolic object are "*events* ". An event denoted $e_i = [y_i=V_i]$ where $V_i \subseteq O_i$ is a function $\Omega \rightarrow \{true, false\}$ such that $e_i(w)$ = true iff $y_i(w) \in V_i$. For instance, if e_i = [color=red,white], then $e_i(w)$ = true iff the color of w is red or white. When $y_i(w)$ is meaningless (the kind of computer used by a company without computer) $V_i = \phi$ and when it has a meaning but this is not known $V_i = O_i$. The extension of e_i in Ω denoted by ext (e_i/Ω) is the set of elements $w \in \Omega$ such that $e_i(w)$ = true.

An assertion is a conjunction of events $a = \underset{i}{\wedge} [y_i=V_i]$; the extension of a denoted $ext(a/\Omega)$ is the set of elements of Ω such that $\forall i \; y_i(w) \in V_i$.

A "*horde* " is a symbolic object which appears, for instance, when we need to express relations between parts of a picture that we wish to describe. More generally a horde is a function h from Ω^p in {true, false} such that $h(u) = \underset{i}{\wedge} [y_i(u_i) = V_i]$ if $u = (u_1,....,u_p)$. For example : $h = [y_1(u_1) = 1] \wedge [y_2(u_2) = \{3,5\}] \wedge [y_3(u_1) = [30,35]] \wedge$ [neighbour (u_1,u_2) = yes].

A synthesis object is a conjunction or a semantic link between hordes denoted in the case of conjunction by $s = \underset{i}{\wedge} h_i$ where each horde may be defined on a different set Ω_i by different descriptors. For instance Ω_1 may be individuals, Ω_2 location, Ω_3 kind of job etc. All these objects are detailed in Diday (1991).

2. External modal objects

Suppose that we wish to use a symbolic object to represent individuals of a set satisfying the following sentence : "It is possible that their weight be between 300 and 500 grammes and their color is often red or seldom white" ; this sentence contains two events e_1 = [color = {red, white}] which lack the modes *possible* , *often* and *seldom* , a new kind of event, denoted f_1 and f_2, is needed if we wish to introduce them f_1 = *possible* [height = 300,500]] and f_2 = [color = {*often* red , *seldom* white}] ; we can see that f_1 contains an *external* mode *possible* affecting e_1 whereas f_2 contains *internal* modes affecting the values contained in e_2. Hence, it is possible to describe informally the sentence by a modal assertion object denoted $a = f_1 \wedge_x f_2$ where $\wedge_x$ represents a kind of conjunction related to the background knowledge of the domain. The case of modal assertions of the kind $a = \underset{i}{\wedge} f_i$ where all the f_i are events with external modes has been studied, for instance, in Diday (1990).

3. Internal modal objects

3.1. A formal definition of internal modal objects

Let x be the background knowledge and

. M^x a set of modes, for instance M^x = {often, sometimes, seldom, never} or $M^x = [0,1]$.

. $Q_i = \{ q_i^j \}_j$ a set of mappings q_i^j from O_i in M^x, for instance O_i = {red, yellow, green},

$M^x = [0,1]$ and q_i^j (red) = 0.1 ; q_i^j (yellow) = 0.3 ; q_i^j (green) = 1, where the meaning of the values 0.1, 0.3, 1 depends on the background knowledge (for instance q_i^j may express a possibility)

. y_i is a descriptor (the *color* for instance) ; it is a mapping from Ω in Q_i. Notice that in the case of boolean objects y_i was a mapping from Ω in O_i, and not Q_i.

Example : if O_i and M^x are chosen as in the previous example and the color of w is red then $y_i(w) = r$ means that $r \in Q_i$ be defined by r (red) = 1, r(yellow) = 0, r (green) = 0.

. $OP_x = \{ \cup_x, \cap_x, c_x \}$ where $\cup_x, \cap_x$ expresses a kind of union and intersection between subsets of Q_i and $c_x(q_i)$ (sometimes denoted $\bar{q}_i$, the complementary of $q_i \in Q_i$).

Example : if $q_i^j \in Q_i$ and $Q_i^j \subseteq Q_i$

$$q_i^1 \cup_x q_i^2 = q_i^1 + q_i^2 - q_i^1\, q_i^2$$

$$q_i^1 \cap_x q_i^2 = q_i^1\, q_i^2 \text{ where } q_i^1\, q_i^2\,(v) = q_i^1\,(v)\ \ q_i^2\,(v)\ ;\ c_x(q_i) = 1 - q_i$$

$$Q_i^1 *_x Q_i^2 = b(Q_i^1) *_x b(Q_i^2) \text{ where } *_x \in \{ \cup_x, \cap_x \} \text{ and}$$

$$b(Q_i^j) = \{ \cup_x q_i \,/\, q_i \in Q_i^j \} \text{ and } c_x(Q_i^j) = 1 - c_x(b(Q_i^j)).$$

This choice of OP_x is "archimedian" because it satisfies a family of properties studied by Shweizer and Sklar (1960) and recalled by Dubois and Prade (1988).

. g_x is a "comparison" mapping from $Q_i \times Q_i$ in an ordered space L^x.

Example : $L^x = M^x = [0,1]$ and $g_x(q_i^1, q_i^2) = \langle q_i^1, q_i^2 \rangle$ the scalar product

. f_x is an "aggregation" mapping from $P(L^x)$ the power set of L^x in L^x . For instance, $f_x(\{L_1,\ldots,L_n\}) = \text{Max } L_i$.

Let $Y = \{y_i\}$ be a set of descriptors and $V = \{V_i\}$ a set of subsets of Q_i such that $V_i = \{q_i^j\} \subseteq Q$. Now we are able to give the formal definition of an internal object (called "im" object).

Definition of an im assertion

Given OP_x, g_x and f_x, an im assertion is a mapping a_{YV} from Ω in an ordered space L^x denoted $a = \wedge_i [y_i = \{q_i^j\}_j]$ such that $w \in \Omega$ is described for any i by $y_i(w) = \{r_i^j\}_j$ then

$$a_{yv}(w) = f_x(\{g_x(\cup_{x\,j}\, q_i^j, \cup_{x\,j}\, r_i^j)\}_i).$$

We denote by $\mathfrak{a}_x$ the set of im objects associated to background knowledge x and φ the mapping from Ω in $\mathfrak{a}_x$ such that $\varphi(w) = w^s = \wedge_{x\,i} [y_i = y_i(w)]$.

Notice that more complex objects may occur when instead of only one, as in the preceding definition, several events concern the same variable ; if we notice $a = \wedge_i e_i$ with $e_i = [y_i = \{q_i\}]$ for instance, for the ith variable, instead of only $e_i = [y_i = \{q_i\}]$, we may have the event

$a_i = \wedge_x [y_i = \hat{q}_i]$; in which case, it is necessary to introduce a third mapping h from $P(L^x)$ in L^x such that $a_i(w) = h\ \{g(\hat{q}_i, r_i)\}^\wedge)$; hence, more generally if $a = \wedge_{i x}\ a_i = \wedge_{i x} \wedge_x [y_i = \hat{q}_i]$ then

$a(w) = f_x(\{a_i(w)\}_i) = f_x(\{h_x(\{g_x(\hat{q}_i, r_i)\}^\wedge)\}_i)$.

Example : Let $M_i^X = [0,1]$, $O_i = \{v_1, v_2\}$, and Q_i be the set of probability measures $P(O_i) \rightarrow [0,1]$; y is a mapping from a set Ω in Q_i and $w^s = [y_i = r]$ is such that $r(v_1) = r(v_2) = \frac{1}{2}$; the set of im assertions $e_i = [y = q_i]$ such that $a_i(w) \geq \frac{1}{2}$ is defined by the set of probability measures q_i which satisfy the inequality $e_i(w) = f_x(g_x(q_i, r)) \geq \frac{1}{2}$; if f_x is the mean and g_x is the scalar product we get $e_i(w) = \text{Mean}(\{<q_i,r>\}) = <q_i,r>$ as there is only one variable. Hence q_i has to satisfy the following inequality :

$e_i(w) = <q_i, r> = q_i(v_1)\, r(v_1) + q_i(v_2)\, r(v_2) \geq \frac{1}{2}$ which is equivalent to $\frac{1}{2}\ q_i(v_1) + \frac{1}{2}\ q_i(v_2) \geq \frac{1}{2}$ which is satisfied by any assertion a, as $q(v_1)+q(v_2) = 1$ for any measure of probability defined on O_i. Let be $a_i = \wedge_{e x} \{\hat{e}_i\ /\ \{\hat{e}_i(w) \geq \frac{1}{2}\}\}$ then $a_i(w) = h_x(\{\hat{e}_i(w)\}^\wedge)$; if $h_x = \min$ then $a_i(w) = \text{Min}(\{\hat{e}_i(w)\}^\wedge) = \frac{1}{2}$.

3.2. Extension of im objects

There are at least two ways to define the extension of an im object a. The first consists in considering that each element $w \in \Omega$ is more or less in the extension of a according to its weight given by a(w); in this case the extension of a denoted Ext (a/Ω) will be the set of couples $\{(w, a(w))\ /\ w \in \Omega\}$. The second requires a given threshold α and then, the extension of a will be Ext $(a/\Omega, \alpha) = \{(w, a(w))\ /\ w \in \Omega, a(w) \geq \alpha\}$.

3.3. Semantic of im objects

In addition to the modes, several other notions may be expressed by an im object a :

a) Certitude : a(w) is not true or false as for boolean objects but expresses a degree of certitude.

b) Variation : this appears at two levels in an im object denoted $a = \wedge_{i x} [y_i = \{q_i^j\}_j]$; first in each q_i^j, for instance if y_i is the color and q_i^1 (red) = 0.5, q_i^1 (green) = 0.3 it means that a variation exists between the individual objects which belong to the extension of a (for instance a species of mushrooms) where some are red and others are green; second, for given description y_i between the q_i^j (each q_i^j expresses for instance the variation in a different kind of species).

c) Doubt : if we say that the color of a species of mushroom is red "or" green, it is an "or" of variation, but if we say that the color of the mushroom which is in my hand is red "or" green, it is an "or" of doubt.

Hence, if we describe $w \in \Omega$ by $\varphi(w) = w^s = \wedge_i \; [y_i = y_i \; (w)]$ where $y_i \; (w) = \{r_i^j\}_j$ we express a doubt in each r_i^j and among the r_i^j provided, for instance, by several experts.

3.4. An example of background knowledge expressing "intensity".

Here the background knowledge x is denoted i, for intensity. Each individual object $w \in \Omega$ is a manufactered object described by two features y_1 which expresses the degree of "roundness" and "flatness" and y_2 the "heaviness" : O_1 = {flat, round}, O_2 = {heavy}; M^i = {*very, quite, a little, very little, nil*}

Let a and w^s be defined by :

a = [y_1 = *a little* flat, *quite* rounded] $\wedge_i$ [y_2 = *a little* heavy]

w^s = [y_1 = *quite* rounded] $\wedge_i$ [y_2 = *very* heavy, *quite* heavy].

(The user has a doubt for w between *very* and *quite* heavy).
The problem is to know if it is acceptable to say that w belongs to the class of manufactured objects described by a.

Hence q_1^1 (flat) = *a little* ; q_1^1 (rounded) = *quite* ; q_2^1 (heavy) = *a little* , r_1^1 (flat) = *nil* ;

r_1^1 (rounded) = *quite* ; r_2^1 (heavy) = *very* , r_2^2 (heavy) = *quite* .

A given taxonomy Tax which expresses the background knowledge on the values of M^i makes it possible to say that Tax (*very, quite*) = *somewhat* ; hence if we settle that

$r_2^1 \cup_i r_2^2$ (v) = Tax (r_2^1 (v), r_2^2 (v)) we have $r_2^1 \cup_i r_2^2$ (heavy) = Tax (*very, quite*) = *somewhat.*

We define L^i by L_1 = not acceptable, L_2 = acceptable, L_3 = completely acceptable and we suppose that the comparison mapping g_i is given by a table T_{gi} such that

g_i (q_1^1, r_1^1) = T_{gi} ((*a little* flat, *quite* rounded), (*nil* flat, *quite* rounded)) = acceptable and

g_i (q_2^1, $r_2^1 \cup_i r_2^2$) = T_{gi} (*a little* heavy, *somewhat* heavy) = not acceptable.

Finally if we settle f({L_i}) = Min L_i and $L_1 < L_2 < L_3$ we obtain
a(w) =f_i (g_i (q_1^1, r_1^1) , g_i (q_2^1, $r_2^1 \cup_i r_2^2$)) = f_i (not acceptable, acceptable) = not acceptable.

4. Probabilist objects

4.1. The probabilist approach

First we recall the well known axioms of Kolmogorov :

If $C(\Omega)$ is a σ-algebra on Ω (i.e. a set of subsets stable for numerable intersection or union and for complementary).We say that p is a measure of probability on (Ω, $C(\Omega)$) if

i) $p(\Omega) = 1$

ii) $p(\cup_i A_i) = \Sigma\, p(A_i)$ if $A_i \in C(\Omega)$ and $A_i \cap A_j = \phi$.

There are several semantics which follow these axioms : for instance luck in games, frequencies, some kind of uncertainty by subjective probability. Let Q_i be a set of measures of probabilities defined on $(O_i, C(O_i))$.

Definition

A probabilist assertion is an im assertion which takes its values in $L^{pr} = [0,1]$

$OP_{pr}: \forall q_i^1, q_i^2 \in Q_i \;\; q_i^1 \cup_{pr} q_i^2 = q_i^1 + q_i^2 - q_i^1 q_i^2; q_i^1 \cap_{pr} q_i^2 = q_i^1 q_i^2$ *which is the mapping which associate to* $v \in O_i$, $q_i^1(v)\, q_i^2(v)$; $c_{pr}(q) = \bar{q} = 1 - q$.

$g_{pr}: \forall q_i^1, q_i^2 \in Q_i \;\; g_{pr}(q_i^1, q_i^2) = < q_i^1, q_i^2 > = \Sigma\{q_i^1\, v)\, q_i^2(v) / v \in O_i\}$.

$f_{pr}: f_{pr}(\{L_i\})$ = *mean of the* L_i.

Notice that it may happen that if there are some characteristic dependancy between variables, $[y_i = q_i]$ may represent them;for instance,if the expert wishes to describe the dependencies between y_1, y_3, y_7,then,this information may be represented by the event denoted $[y_{137} = pr(y_1, y_3, y_7)]$ where $pr(y_1, y_3, y_7)$ represents the conjoint probability of y_1, y_3, y_7;this event is of the form $[y_i = q_i]$ where $y_i = y_{137}$ and $q_i = pr(y_1, y_3, y_7)$.In the case where the same dependencies do not appear in the probabilist assertion a and in w^s (because they are not given by the expert), to compute a(w) it is needed to use propagation technics in a belief network may be used (see J. Pearl (1988) or D.J.Speegelhalter & al (1989)) for finding the missing one.

To give an intuitive idea of the notion of union of measures of probabilities it is easy to see that if q_i^1 and q_i^2 are the measure of probabilities associated to two dice, $q_i^1 \cup_{pr} q_i^2$ (V) is the probability that the event V occurs, for one dice **or** (not exclusive) for the other ,when the two dices are thrown independently. Notice that $q_i^1 \cup_{pr} q_i^2$ is not a measure of probability because even if $q_i^1 \cup_{pr} q_i^2(v) \in [0,1]$ the sum of the $q_i^1 \cup_{pr} q_i^2(v)$ on O_i is larger then 1. Also, $q_i^1 \cap_{pr} q_i^2$ is not necessarily a measure of probability because the sum of the $q_i^1 \cap_{pr} q_i^2(v)$ on O_i may be lower than 1.

4.2. Example

An object w is described by its color $y_1(w)$ which may be red or blue and its roundness $y_2(w)$ which may be round or flat.

Let $a = [y_1 = q_1^1, q_1^2] \wedge_{pr} [y_2 = q_2]$ and $w^s = [y_1 = r_1] \wedge_{pr} [y_2 = r_2]$ where q_1^1 (red) = 0.9; q_1^1 (blue) = 0.1; q_1^2 (red) = 0.5; q_1^2 (blue) = 0.5; q2 (round) = 0.2; q2 (flat) = 0.8. It results that a is described by two kind of objects : either often red and rarely blue, or red or blue with equal probability.

By using $q_1^3 = q_1^1 \cup_{pr} q_1^2 = q_1^1 + q_1^2 - q_1^1 q_1^2$ we obtain

q_1^3 (red) = 0.9 + 0.5 - 0.9 x 0.5 = 0.95

q_1^3 (blue) = 0.1 +0.5 - 0.1 x 0.5 = 0.55

If r_1 and r_2 are defined as follows :

r_1 (red) = 1, r_1 (blue) = 0; r_2 (round) = 1, r_2 (flat) = 0, it results that

$a(w) = g_{pr}(q_1^3, r1) \wedge_{pr} g_{pr}(q_2, r_2)$

$= (0.95 \times 1 + 0.55 \times 0) \wedge_{pr} (0.2 \times 1 + 0.8 \times 0)$

$= 0.95 \wedge_{pr} 0.20 = \frac{1}{2}(0.95 + 0.20) = 0.57$, which represents a membership degree for w to the im object defined by a.

5. The particular case of boolean objects

A boolean object $a = \hat{\wedge}_i [y_i = V_i]$ is an im object $a_b = \hat{\wedge}_i [y_i = q_i]$ where q_i is the characteristic mapping of V_i in O_i, $OP_b = \{\cup_b, \cap_b, c_b\}$ is such that $q_1 \cup_b q_2 = \mathrm{Max}(q_1, q_2)$, $q_1 \cap_b q_2 = \min(q_1, q_2)$ and $c_b(q) = 1 - q$;if $w = \hat{\wedge}_i [y_i = r_i]$ where r_i is the characteristic mapping of $y_i(w)$ in O_i, $g_b(q_i, r_i) = \langle q_i, r_i \rangle$ and $f_b = \min$; it results that if there exists only a single $v \in O_i$ such that $r_i(v) \neq 0$ then $a_b(w) = 1$ (thus $r_i \leq q_i$) $\Leftrightarrow a(w) = \text{true}$ and then $a_b(w) = 0 \Leftrightarrow a(w) = \text{false}$.
If we denote $|a|_\Omega$ the set of elements of Ω such that a(w) = true, we have $|a|_\Omega = \mathrm{Ext}(a_b / \Omega, \alpha)$ $\forall \alpha \in]0,1]$.

6. Some qualities and properties of symbolic objects

6.1. Order, union and intersection between im objects

It is possible to define a partial preorder $\leq_\alpha$ on the im objects by setting that : $a_1 \leq_\alpha a_2$ iff $\forall w \in \Omega \;\; \alpha \leq a_1(w) \leq a_2(w)$.

We deduce from this preorder an equivalence relation R by $a_1 R a_2$ iff $\mathrm{Ext}(a_1 / \Omega, \alpha) = \mathrm{Ext}(a_2 / \Omega, \alpha)$ and a partial order denoted $\leq_\alpha$ and called "symbolic order" on the equivalence classes induced from R.
We say that a_1 inherits from a_2 or that a_2 is more general than a_1, at the level α, iff $a_1 \leq_\alpha a_2$ (which implies $\mathrm{Ext}_\alpha(a_1 / \Omega, \alpha) \subseteq \mathrm{Ext}_\alpha(a_2 / \Omega, \alpha)$).
We call intention at the level α of a subset $\Omega_1 \times \Omega$ the symbolic object b defined by the conjunction of events whose extension at the level α contains Ω_1.
The symbolic union $a_1 \cup_{x,\alpha} a_2$ (resp. intersection $a_1 \cap_{x,\alpha} a_2$) at the level α is the intention of $\mathrm{Ext}(a_1 / \Omega, \alpha) \cup \mathrm{Ext}(a_2 / \Omega, \alpha)$ (resp. $\mathrm{Ext}(a_2 / \Omega, \alpha) \cap \mathrm{Ext}(b / \Omega, \alpha)$).

6.2. Some qualities of symbolic objects

As in the boolean case, see Brito, Diday (1989), it is possible to define different kinds of qualities of symbolic objects (refinement, simplicity, completeness etc.).

For instance, we say that a symbolic object s is complete iff the properties which characterize its extension are exactly those whose conjunction defines the object; in other words s is a complete symbolic object if it is the intention of its extension. More intuitively if I can see some white dogs and I state "I can see some dogs", my statement doesn't describe the dogs in a complete way, since I am not saying that they are white.

On the other hand,the simplicity at level α of an im object is the smallest number of elementary events whose extension at level α coincides with the extension of s at the same level.

6.3. Some properties of im objects : lattice and completeness

It may be shown, see Diday (1992) for instance, that given a level α the set of im objects is a lattice for the symbolic order and that the symbolic union and intersection define the supremum and infimum of any couple. To do so, f_x, g_x and h_x (see § 3.1) have to be well chosen and we introduce a "full" and an "empty" symbolic object denoted Ω^s and ϕ^s such that $\forall w \in \Omega$, $\Omega^s(w) = 1$ and $\phi^s(w) = 0$.

It may also be shown that the symbolic union and intersection of complete im objects are complete im objects and hence that the set of complete im objects is also a lattice.

7. Statistics and data analysis of symbolic objects

a) Four kinds of data analysis problems

Several studies have recently been carried out in this field : for histograms of symbolic objects, see De Carvalho & al (1990) and (1991); for generating rules by decision graph on im objects in the case of possibilist objects with typicalities as modes see Lebbe and Vignes (1991); for generating overlapping clusters by pyramids on symbolic objects see Brito, Diday (1990).

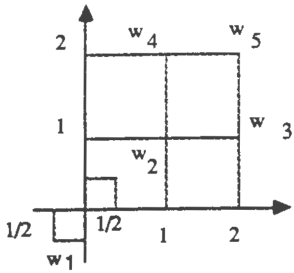
More generally, four kinds of data analysis may roughly be defined depending on the input and output : a) numerical analysis of classical data tables b) symbolic analysis of classical data tables, (for instance obtaining a factor analysis or a clustering automatically interpreted by symbolic objects) c) numerical analysis of symbolic objects (for instance by defining distances between objects) d) symbolic analysis of symbolic objects where the input and output of the methods are symbolic objects. We shall here illustrate only the second approach which is the point of view we need for factorial axis symbolic interpretation.

b) Symbolic analysis of classical data table.

Let T be the following data table where the set of individual objects is $\Omega = \{w_1,...,w_5\}$ which are five companies described by two variables y_1 : the employment rate and y_2 : the profit.

	w_1	w_2	w_3	w_4	w_5
y_1	-1/2	1/2	2	1	2
y_2	-1/2	1/2	1	2	2

Table T



Graphical representation of table T

. Principal component analysis of Table T : From the covariance matrix $V = \begin{pmatrix} 0.9 & 0.7 \\ 0.7 & 09 \end{pmatrix}$ we deduce the eigen values $\lambda_1 = 1.6$ and $\lambda_2 = 0.2$ and the eigen vectors $u_1^T = \frac{1}{\sqrt{2}}(1\ 1)$, $u_2^T = \frac{1}{\sqrt{2}}(1\ -\ 1)$. Finally we get the principal component representation given in figure 1, where the projection of w_j on the axis i is given by $F_i(w_j) = u_i^T x_j$ where $x_j^T = (y_1(w_1)-Y_1, y_2(w_2)-Y_2)$ where $Y_i=1$, is the mean of y_i ;for instance, $F_1(w_1) = \frac{1}{\sqrt{2}}(1\ 1)\begin{pmatrix} -3/2 \\ -3/2 \end{pmatrix}$.

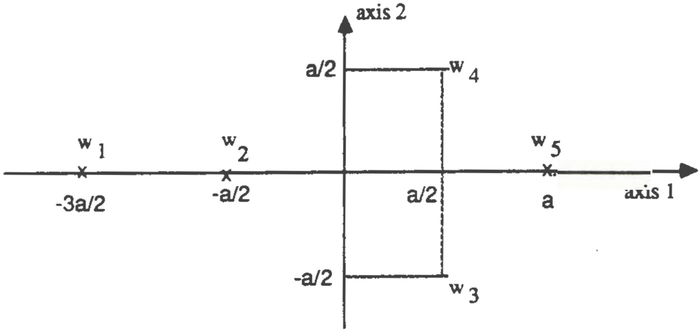


Figure 1.Principal component analysis of table T whith a= √2.

The correlation between (w_1, ..., w_5) with the first axis of the principal component analysis is respectively (-1, -0.707, 0.707, 0.707, 1); if we associate to each side of the first axis the objects whose correlation is higher than 0.707 or lower than -0.707, we obtain two classes of objects; the first class, $C_1 = \{w_1, w_2\}$, explains the left side of the axis and the second one $C_2 = \{w_3, w_4, w_5\}$ explains the right side. By using these classes, we get two kinds of symbolic interpretation of the first axis, by using assertion we may say that the left side is explained by : $a_1 = [y_1 = -1/2, 1/2] \wedge [y_2 = -1/2, 1/2]$; the right side is explained by $a_2 = [y_1 = 1,2] \wedge [y_2 = 1,2]$. If at the input we have a taxonomy saying that the rate of employment and the profit are low when they are lower than $\frac{1}{2}$ and high when they are higher than 1, we may use the assertions a_1 and a_2 to get the following explanation of the first axis : it is explained by two opposite assertions which characterize two classes of companies :

a_1=[Rate of employment = low] $\wedge$ [Profit = low]

a_2=[Rate of employment = high] $\wedge$ [Profit = high]

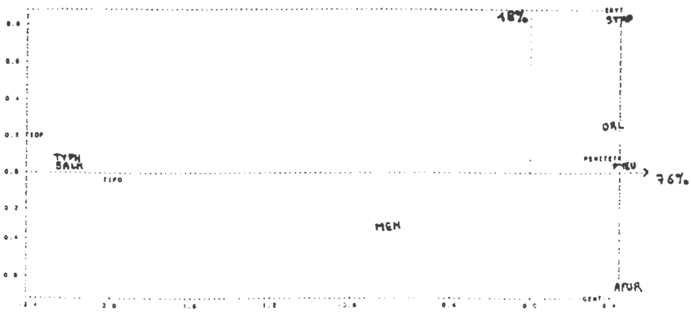
Of course, in real examples things become much more complicated; for instance, to get more accuracy when the two classes contain numerous objects, each side of the axis may be explained by a disjunction of assertions obtained by a symbolic interpretation of a clustering done on each class. We may also enrich the interpretation by adding certain properties; for instance, we may add to a_1 the following rules : [if $y_1 = \frac{1}{2}$ then $y_2 = -\frac{1}{2}$] $\wedge$ [if $y_1 = \frac{1}{2}$ then $y_2 = \frac{1}{2}$] and to a_2 the rule [if $y_1 = 1$ then $y_2 = 2$].

We may also give an interpretation of the first axis by a horde object h : $h = a_1(u_1) \wedge a_2(u_2)$ = [Rate of employment (u_1) = low] $\wedge$ [Profit (u_1) = low] $\wedge$ [Rate of employment (u_2) = high] $\wedge$ [Profit (u_2) = high] whose extension is composed of couples of companies (w_i, w_j) the first element of the couple w_i ,being of low rate of employment and profit and the second one w_j, of high rate of employment and profit. If an external variable gives the age of the companies the horde object h may become : $h = a_1(u_1) \wedge a_2(u_2) \wedge$ [age(u_1) <age(u_2)].

Lets consider Françoise Benzecri's example (Benzécri F. 1980) which was proposed for Tenon Hospital conference on factorial and clustering methods by P. and M. Curie University statistical laboratory in June 1980.
Rows are diseases and columns treatments.
Each data represents the number of cases in which a treatment has been applied to a disease.
In the case of a correspondence analysis, coordinates, absolute and relative contributions, and related representation on the two first axis B_1 and B2 are the following :

			COORD B1	B2					ABS CONTR B1	B2					REL CON B1	B2
TYPH	0.087	5.39	-2.31	0.08	0.16	0.03	0.00	0.00	53.0	0.3	4.3	0.4	0.1	0.0	0.99	0.00
SALM	0.043	5.39	-2.31	0.08	0.16	0.03	0.00	0.00	26.5	0.1	2.1	0.2	0.0	0.0	0.99	0.00
ORL	0.275	0.22	0.44	0.14	0.06	-0.08	-0.01	0.00	6.1	2.8	1.8	6.7	60.5	0.0	0.86	0.09
PNEU	0.246	0.23	0.44	0.03	0.11	-0.16	0.01	0.00	5.4	0.1	6.1	22.1	37.4	0.0	0.83	0.01
MENI	0.072	1.05	-0.82	-0.31	-0.73	-0.17	0.00	0.00	3.2	3.5	79.1	6.9	0.0	0.0	0.37	0.09
AFUR	0.174	0.79	0.41	-0.78	0.04	0.22	0.00	0.00	3.3	49.1	0.6	29.6	0.0	0.0	0.21	0.72
STAP	0.101	1.22	0.46	0.94	-0.17	0.31	0.00	0.00	2.4	44.1	5.9	34.2	2.5	0.0	0.17	0.72

			COORD B1	B2					ABS CONTR B1	B2					REL CON B1	B2
PENI	0.348	0.18	0.37	0.05	-0.11	-0.07	-0.01	0.00	5.5	0.5	8.4	6.8	44.1	0.0	0.87	0.02
TIFO	0.116	4.18	-2.02	-0.04	-0.30	-0.09	0.01	0.00	53.8	0.1	21.1	3.5	9.9	0.0	0.98	0.00
TETR	0.188	0.33	0.47	0.05	0.26	-0.21	0.01	0.00	4.6	0.3	25.9	29.4	20.9	0.0	0.65	0.01
ERYT	0.116	1.15	0.48	0.92	-0.07	0.27	0.01	0.00	3.0	48.2	1.0	29.0	7.1	0.0	0.20	0.74
TIOP	0.043	6.87	-2.47	0.18	0.70	0.20	-0.01	0.00	30.2	0.7	43.6	6.2	14.9	0.0	0.91	0.01
GENT	0.188	0.72	0.37	-0.74	0.00	0.19	0.00	0.00	2.9	50.2	0.0	25.0	3.0	0.0	0.19	0.76



Among disease data, "typhoide" and "salmonellose" may be chosen, as representative of the negative part of B_1 and with a contribution threshold of 25% .
Those diseases can be resumed in terms of original variables as those who are never treated with "penicilline" :

["penicilline " treatment = never]

and this description perfectly discriminates them from the other diseases. In fact they are the only diseases which have zero in correspondance with "penicilline "

We shall now focuse on multiple correspondence analysis axis interpretation.

3 . Characteristic assertion generator for a factorial axis in correspondence analysis

Let a table of a classical nominal data set T on two finite sets I and J; let $\{ m_s \}$ be the r different levels of the q J-variables and $\{ w_i \}$ the N units of I

3.1. Factorial axis interpretational aid summary

a) Barycentric interpretation

In two-way correspondence analysis, relations between elements of I and J can be made explicit by *Transition Formulas* . Let $F_A (w_i)$ and $G_A (y_j)$ be the coordinates on A axis associated to the eigen value λ_A (not equal to zero) of a unit w_i and a variable y_j . Let k_{ij}, $k(i)$, $k(j)$, f_{ij}, f_i and f_j the w_i and y_j associated values, weights and profiles in classical Benzecri's notation. The following relations hold :

$$F_A (w_i) = \Sigma_j f_{ij} G_A (y_j) / (\lambda_A)^{1/2} f_i$$

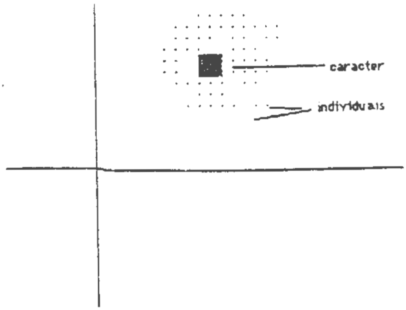
$$G_A (y_j) = \Sigma_i f_{ij} F_A (w_i) / (\lambda_A)^{1/2} f_j$$

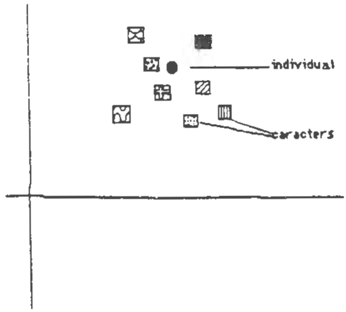
The coordinate of an element w_i of I is the centroid of the coordinates of the elements y_j of J with masses having for values the coordinates of the profile $f_J{}^i$ (close to the multiplicative factor).
This point of view is to take in account when interpreting the factorial planes : it gives an indication on unit and variable associations which may be allowed from the mapping observation .
In the case of multiple correspondence, analysis is generally carried on the Burt table B, which is built from the complete disjunctive form D of the original data set T. D lays as a supplementary table, by B. Let $GB_A (m_t)$ be the coordinates on Burt table axis of the modality m_t from variable y_j .
The transition relation shows that each modality is the centroid of the individuals which have that modality :

$$GB_A (m_t) = \Sigma \left\{ F_A (w_i) , w_i \in I , y_j (w_i) = m_t \right\} / k (m_t)$$

On a similar way, one can demonstrate that w_i coordinate on D is the mean value of the associated modality coordinates on A normalized eigen vector.

Finally, by analogy with active modalities, similar results can be established for a supplementary attribute m_s from variable y_{j+} .





We may say that a supplementary response in a survey is on each factorial plane a quasi barycenter of the respondants who have chosen that modality of response.

For example :

$$G_A (m_s) = \Sigma \left\{ F_A (w_i), w_i \in I, \; y_{j+} (w_i) = m_s \right\} \Big/ (\lambda_A)^{1/2} \, k (m_s)$$

b) Coordinates, absolute and relative contributions

Units and modality projections can be placed all along every factorial axis (which is not associated to an eigen value equal to zero) by their coordinates; they are computed from the original data set on the new basis vectors which are the correspondence analysis normalized eigen vectors.

More extreme is the place of an element on a factorial axis, more important is that element generally considered for the axis interpretation.

The percentage of absolute contribution of a point to the moment of inertia λ_A is computed as follows :

$$CTR (m_k) = f_j \; G_A^2 (m_k) \; / \; \lambda_A$$

$$CTR (y_j) = \Sigma \; \{ \; CTR (m_k), \; m_k \in y_j \; \}$$

$$CTR (w_i) = f_i \; F_A^2 (w_i) \; / \; \lambda_A$$

The relative contribution of the factor A to the point w_i is computed as follows :

$$\cos^2(w_i) = F_A^2(w_i) \left[\Sigma_j \ (f_j^i - f_j)^2 / f_j \right]^{-1}$$

The above numbers are the principal interpretational aids for a factor :

. a factor is dependant on the elements which contribute the most to its dispersion. The CTR will be therefore examined in priority in order to identify or name the factor

. the $\cos^2$ numbers are similar to correlation coefficients; when they are summarized on the l first axis, they give a percentage of the quality of the explanation of the element w_i in the factorial space of dimension l. In order to study factorial axes with high rank, which generally express localized effects, $\cos^2$ are more useful than CTR

REMARK

All these coefficients may be computed on I elements as on J active elements; but the absolute contribution of a supplementary element has no meaning as it does not take part in the construction of the factorial axis.

c) Test- value notions for supplementary modalities

It is often interesting for enquiry results to characterize the respondants by descriptions such as sex, age, etc. But generally, they are only supplementary elements for a factorial analysis which is much more concerned by the problem concepts as active variables.
So, it is consequently difficult to appreciate supplentary element importance as they have no CTR on factorial axis as previous remark (8.2) mentioned.
To have nevertheless a quantitative information on such an element position, A. Morineau [Morineau (Mars 1986)] proposed a test on the hypothesis H_0 of an hypergeometric law as a theoretical model for the coordinate distribution. Expected means and standard deviation can so be computed on each factorial axis. One can demonstrate that the variance then should be :

$$V_{H0}[\, G_A(m)\,] = \frac{N-n_m}{N-1} \; \frac{1}{n_m}$$

Because of central limit theorem, $\sqrt[2]{\frac{N-n_m}{N-1} \; \frac{1}{n_m}} \; G_A(m)$ will follow a

centered reduced normal law. The following quantity is called *test value* for the modality m :

$$t_{m,A} = \left[\frac{N-n_m}{N-1} \; \frac{1}{n_m} \right]^{1/2} G_A(m)$$

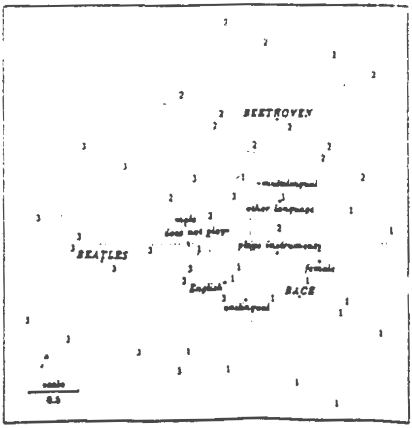
These computations are also meaningful for active modalities where they usually take high values, but they are essentially used for supplementary variables.

d) Principal interpretation difficulties

In spite of the numerous numerical coefficients which are listed as analysis results, and in spite of the classical factorial mappings which are usually displayed, correspondence interpretation is a delicate phase for different reasons (see Escoffier -Pagès [1990]) :

. thresholds are necessary to select "good" contributions, correlations and so on

. an abuse of graphical proximity leads the standard user to state conjunctions or even rules between elements on the mapping when the method gives no justification for them ; in the following example [Greenacre 1991], Greenacre demonstrates that the statment, from dimensional interpretation, of an association between "*male* " and " *does not play* " , or between " *Bach* " and " *female* " would be erroneous :



. a factorial axis is a vectorial element the componants of which are not explicit in the terms of initial data

. to appreciate unit subset densities all along factorial axis, cluster center projections are often represented on factorial planes. But they lack of a direct explanation as monothetic classes have : in fact, their descriptors are quantified by statistical tests which represent tendencies in the group so that they are not so easily understandable and cannot be easily managed by the user

An experienced analyst and an expert of the data domain are both important to extract correct knowledge from the analysis proceeding.
The following descriptions of factorial axis as true conjunctions of initial variables, and finally as disjunctions of modal assertions on initial data, will be a real aid to understand the factorial analysis results.

3.2 Assertion generator for axis extremities

The main idea consists in producing the best adapted symbolic objects (see Diday, 1989) for a partition of monothetic classes, created at an axis extremity (that is classes of respective individuals such as all of them have the same common modalities).
Let $c_A{}^1$ a threshold of " good contribution " on one of the two A extremities.

Let E the set of I units of the concerned extremity (see coordinate signs) which CTR (see 8.1 b)) are equal at least to $c_A{}^1$.

Let CE be the I-complement of E.
E will be called the set of examples and CE, set of counter examples for the concept of "good contribution " on A extremity.

We shall now use a supervised rule generator on E and CE to characterize by attribute conjunctions subsets of E.

We propose for example an adaptation of the learning algorithm CABRO to the context of multiple correspondence analysis.

The principle steps of the proceeding will be the followings :

Let α be a threshold of discrimination for E from CE

Let β be a threshold of generalization for assertion on E
Let CTR (m_k) or Test-value (m_k) be the elements of an ordered list L associated to the modalities of the data table (active and supplementary)

Remark : for every I-unit the conjunction of all its modalities represent an assertion which is true on that unit

In search of a more general assertion than the original one for each w_i of I

. one starts from an empty conjunction, which is obviously very general and non discriminant

. then one ties the best axis extremity related modality, m_1, thanks to the L list information . The generality of the conjunction diminishes but it may still remain non discriminant.
The ratio

$$R_1 = \frac{\text{ext}(m_1 / E)}{\text{ext}(m_1 / E) + \text{ext}(m_1 / CE)}$$

is a measurement of m_1 discriminating power.

. This phase is repeated with the remaining modalities of w_i until one finds a conjunction such as the associated ratio R is at least equal to α , and which extension contains a percentage of examples greater than β

Find a set of assertions characterizing the classes of a partition on A extremity

.one determines with the previous approach an assertion from each example

. one keeps from the previous research the assertion of maximal extension in E, a_{max} , as an element of the final result.

. one repeats these phases on $E \setminus \text{ext}(a_{max} / E)$ until there are no more example

Find assertions quickly for a large data set

. one determines an ordered set of fictitious objects in the form of a tree the root and nodes of which are defined as in CABRO's algorithm, but replacing the frequencies by the L scores

. one finds for each fictitious object its nearest neighbour in the real data set (for example with Hamming punctual distance)

. one proceeds as in the general case to find the final characteristic assertions, but the assertions are computed only on the set of fictitious objects nearest neighbours.

One must point out , after this brief recall of CABRO's approach, that this adaptation has important peculiarities :

- as the user's demand is a great preoccupation, there is no systematic attempt to optimize a generalization criteria as in original CABRO. On the contrary, an on-line implementation should be able to follow user's constraints on variable choices. For example, the age of the respondant may be requested and forced in the result , even if it is not the most efficient variable for the axis extremity, in order to garantee a more explicit description.

- another reason not to optimize a generalization criteria is that the modality choice depends here on their L scores (contributions etc.) and not on their frequency, which would have been more related with a generality notion

4 . Disjunction of modal assertions to interpret factorial axis

Cabro's approach is not the only possible one for multiple correspondence analysis, to build assertions. For example, any supervised decision tree for qualitative data will give a response, but it is necessary to make it flexible to the user's point of view (for example allowing priority to some variables he requests). It is particularly important in some case to abandon eventual probability tests on misclassification and contingency thresholds in order to go on with the dichotomies to obtain enough detail on the extreme classes which are really the most interesting for the axis interpretation.

4.1 Discrimination, generalization, contributions levels

Generally a problem one may go through, consists in balancing the different parameters : contribution, discrimination, and generalization thresholds. For example, as factorial axis show the extreme points of the data spatial disposition , sometimes there are very few individuals in these regions so that the generated assertions may have very weak extensions if no generalization level is requested.

One can also choose a lower contribution level to increase the extensions.

4.2 Union, intersection, background knowledge

One can also enhance the generality of the assertions using the two operators intersection or union , but still taking into account the discriminating and generalization constraints. In the particular case of preexisting taxonomies, either on the modalities of the same variable, or on different variables, one can merge a disjunction of attributes obtained by union, rewriting it with the related level in the taxonomy.

Example :

Original assertions

[age = [13, 19]] $\wedge$ [practice = with friends] $\wedge$ ...

[age = [18, 25]]   ∧ [practice = with the family] ∧ ...
[member of an association = yes] ∧ ...

A background knowledge simply consists in the following taxonomy :
13 - 14 _ 15 _ 16 _ 17 _ 18 _ 19 _ [20 , 25] ___________ young
with friends _ with the family _______________________ not alone

union operator
[age = [13, 19] , [18 , 25]]   ∧ [practice = with friends, with the family] ∧ ...
[member of an association = yes] ∧ ...

rewriting
[[age = young]   ∧ [practice = not alone] ∧ ...] V [member of an association = yes]∧.
...
(required condition : discrimination level verification)

4 . 3 Imperfect discrimination

The most frequent situation one has to front is that of an imperfect discrimination : for example, one may find that left side of A axis is represented by " joung and athletic people, who use mountain bikes for competition " , but some exceptional " young, athletic and competiting person " may have a projection near the gravity center or even on the other side of the axis, as he is also a very good swimmer. In that case, the strategy consists in decreasing the discrimination level in order to find a sufficient generalization level for the assertions (because of course each original example considered as an assertion is 100% discriminating but really too specific !) .One can anyway save the information on misclassified elements and "misdescriptions" by introducing previous ratio R as an external mode on the assertion.

Example :

O.9 [age = [13, 19]]

means that 90% of the teenagers of the data are at that axis extremity
the remaining 10% are on the remaining part of the axis

4 . 4 Modal symbolic object for factorial axis interpretation

Assertion extensions may be considered in two different ways :

. on subsets that constitutes one axis extremity
. on subsets that all well represented on one axis extremity

Example .

. teenagers represent 20% of the axis extremity
. 90 % of the teenagers are at that axis extremity

External modes may preserve these informations

More, when merging assertions by union or intersection, internal modes may be necessary to express those types of information :

[age = 0.8 [13 , 19] , 0.2 [18 , 25]]

One can also prefer the contribution semantic and save the global contribution of the obtained subsets (this global contribution is the individual contribution sum), to the axis :

Finally, a factorial axis interpretation can be written as a disjunction of modal assertions, which semantics are to be precised , but which are essentially of probabilistic type as modes often come from relative frequencies.

5 . CONCLUSION

Symbolic descriptions for factorial axis fulfill much more than any other interpretation aid the 3rd Yule's condition : a statistical index should have a concrete meaning; it is better to choose a real value than a characteristic which is none of the possible values.

But, on the other hand, their welcome flexibility to the user's requires put them very far from any optimality, validation or robustness preoccupation. These are some of the main directions to improve that approach.

Other developments will be an extension of symbolic interpretation to factorial planes and also to any kind of factorial analysis; one can for example use a segmentation algorithm on continue variables to characterize by conjunctions of interval disjunctions the classes of the required partition on "well contributing" elements.

Answers to threshold management will be obtained by applying the method to the greatest number of possible different domains.

Thanks to the comparison operators on modal symbolic objects [Diday, 1991] , symbolic axis description can also be used for example to study the evolution of the principal axis of a given situation on different periods by comparing them directly with their symbolic formulations. In that type of development, one could think the whole proceeding appears to be referred to probabilistic induction, that is numeric one. In fact, the symbolic definition A^s of a factorial axis A is true on a certain subset of the original data, which is precisely the extent of the related symbolic object; that subset can be considered as statistically meaningful for this axis in terms, for example, of summarized contributions (see 3 - 4) .
But generally, on real data, $\neg A^s$ has a non empty extension (examples 2 are too simple !) so that A^S and $\neg A^s$ are to some extent simultaneously true; we may so consider that A^s gives an incertain information on principal direction A for the original data, and that we have to handle with contradiction. These last considerations and the large use of background knowledge both argue for symbolic rather for mere numeric approach.

More generally, one should think on the following two aspects :

- . numerically, a " principal axis " has no incertainty : it is one of the eigen vector of a given matrix computed from the original data
- . semantically, " principal " is not a perfectly defined concept, and it brings incertainty in the user's interpretation

The symbolic expression of a factorial axis transforms it from a vectorial nature to a nature similar to other symbolic objects that statistics will be able to compute, data bases to manage and data analysis to provide with numerous treatments (see for example De Caravalho FAT, 1991).

References

. Benzecri F. , (1980) *"Introduction à l'analyse des correspondances d'après un exemple de données médicales* "Les Cahiers de l'Analyse de Données- Vol V 1980 -n° 3

. Benzecri J.P. ,(1973) *"L'analyse des données, tome 1 et tome 2"*, DUNOD

. Cazes P. (1983), *"Analyse des correspondances multiples ; application à l'étude des questionnaires"*, Bulletin de l'ADDAD n° 12, laboratoire de statistiques, Paris VI

. Calot ,(1984), " *Statistique descriptive* " DUNOD

. De Carvalho F.A.T. (1991), "Histogramme en Analyse des Données Symboliques", Rapport de Recherche INRIA (à paraître).

. Diday E. (1989), *"Introduction symbolique à l'Analyse des Données"* Recherche Opérationnelle, vol 23, n°2, p.193-236

. Diday E., (1990), *"Knowledge representation and symbolic data analysis"*, in NATO ASI Series, Vol. F 61, Knowledge Data and computer-assisted Decisions edited by Schader and W. Gaul. Springer Verlag.

. Diday E., (1992), *"Objets modaux pour l'analyse des connaissances"*, Rapport INRIA, Rocquencourt, 78150, France.

. Diday E., Kodratoff, (1991) *"Des objets de l'analyse des données à ceux de l'analyse des connaissances"* Cepadues - Editions, 1991

. Escofier B., Pagès G.,(1990), " *Analyse factorielle simple et multiple*" Dunod

. M.J.Greenacre (1991),*"Interpreting multiple correspondance analysis* " Applied stochastic models and data analysis, vol 7

.Ho Tu Bao, Diday E., Gettler-Summa M., (1988) - *"Generating rules for expert system from observations"* , paru dans Pattern Recognition letters 7 pp 265-271 (North-Holland).

. M.Gettler-Summa, H.Ralambondrainy, E. Diday (1988) -*"Data Analysis and Expert Systems: Generating Rules from Data"*. Nato Advanced Research Workshop - Data, Expert Knowledge and Decisions, Hamburg .

. Lebart L.,Morineau A. Fenelon J.P., (1982), " *Traitement statistique des données* " DUNOD

. Morineau A., (1986) ," *Un exemple d'analyse des données* " Bulletin du CEPREMAP de Mars 1986

. Pearl J., (1988), "Probabilist reasoning in intelligent systems" Morgan Kaufman, San Mateo.

.Spiegelhalter, D.J. and Lauritzen, S.L. (1989), "Sequential updating of conditional probabilities on directed graphical structures", Technical report R-89-10, Aalborg University, Aalborg, UK.