

Analysis of women Merovingian graves on the basis of Beads by Aid of a Symbolic Clustering Algorithm

Rosanna Verde

SUN - Seconda Università di Napoli
ITALY

Yves Lechevallier

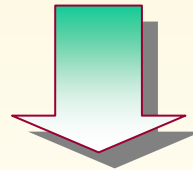
INRIA - Rocquencourt
FRANCE

AIM

- + *Analyse typologies of archeological artifacts* (Merovingian beads)
through

SYMBOLIC DATA ANALYSIS

(Dynamical Clustering technique)



New efficient statistical techniques recently developed to analyze structured and complex data (symbolic objects)

Study of grave goods

of the Merovingian cemetery of Weingarten

- *Necklaces **beads** from women graves*
(450 -- 720 A.C.)



Font of the data: Sasse, (2001)

In a previous study (Workshop on SDA, Munich 2001)
we presented a symbolic classification approach on

*Small iron or bronze belt **Buckles** of the same*
Merovingian period

like that:



Font of the data: Leredde, (1979)

*The coding of the "beads" data
arise from an expert classification*

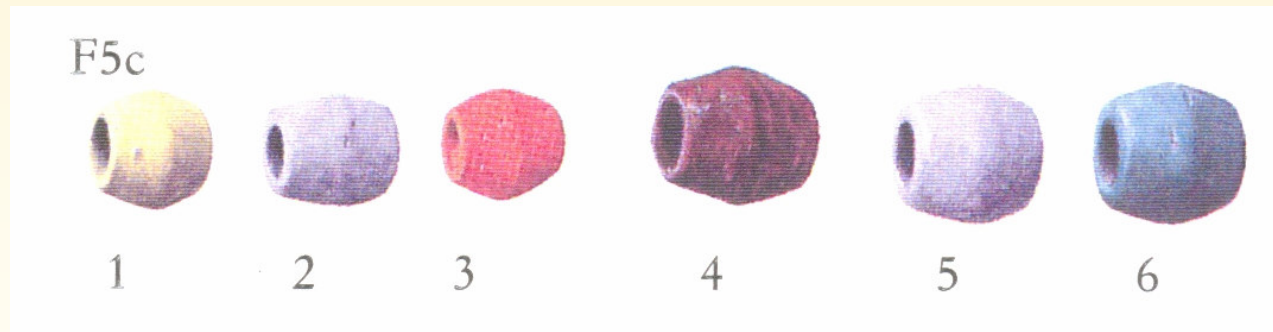
Expert Classification of the beads according to the shape, color and decoration

F5 biconical

F5 a - small



F5 b - long-streck



F5 c - normal

Shapes



Classification of the Beads for shape distinguishes also beads in

monochromatic -- polychromatic

F

F1 - ring-form

F2 - barrel form

F3 - circular

F4 - ...

...

F17 -



V



V1 - Points in the middle

V2 - spots one color

...

V14 - double swag

...

***V19 - combed one color
spirales***

...

V23 - millefiori

Color

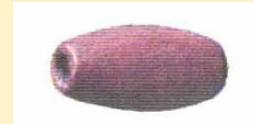
simple color classification

Monochromatic



Black

Crimson



Dark blue

Orange



Green-blue

Yellow



Dark green

Translucid
glass



Olive green

White milk



Brown

White



Polychromic Decoration



Combed spirale



flowers



Reticella



dots




Little colored spots

Association of the Beads to the Graves

F10:

a1 grave 108; b1 grave 87; c1-2 grave 270; c3-4 grave 233



F1d4	T103	Black	Monochromic	oval
V17a1	T103	Brown	spirales	barrel
F10a1	T108	Dark blue	Monochromic	Cylindrical
F1g1	T108	Dark blue	Monochromic	oval
V23a3	T108	Translucid glass	flowers	prismatic
F9b6	T109	Olive green	Monochromic	melon
V16a9	T109	Black	spirales	biconical
V21a1	T109	Dark blue	combed spirale	Cylindrical

Symbolic
data
description

Symbolic object for the beads description

Symbolic objects is a new way of modelling

(statistical) information

(Diday, 1987;
Bock & Diday, 2000)

based on "concepts representation" of:

typologies - expert knowledge
homogeneous classes

these concepts are characterised by a set of descriptors

$$Y_1 \ Y_2 \ \dots \ Y_p$$

That - differently from the classical variables - can assume more than one value for each object:

a sequence of categories - intervals - distributions

Symbolic Object definition

Bock & Diday, 2000

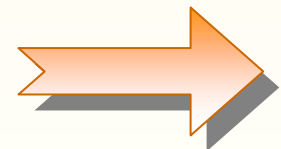
A symbolic object s is defined as a triple :

$$(a, R, d)$$

where d is the description of a native concept, a category, a class, etc.

which doesn't depend from the observations and R is a set of relations

The SO definition contains the mapping function a that allows to recognize individuals as belonging to its Extent



- Let Ω be a set of elementary elements ω

The Extent of s is computed by means of a Mapping function

which compares the values $y(\omega)$ with the description d of the SO

$$a(\omega) = \bigwedge_{i=1, \dots, p} [y_i(\omega) R_i d_i] \in \{true, false\}$$

$$Extent(s) = \{\omega \in \Omega \mid a(\omega) = true\}$$

Set of the elements ω of Ω satisfying the properties of the SO s

Symbolic object contains in its definition a classification rule !

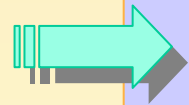
Relationships among the several descriptors
can be also considered in symbolic objects
modeling - expressed by:

- logical rules, *for example:*

if Monochromic then *decoration = inapplicable*

- taxonomical structures





Classification of the Beads according to a taxonomic structure of theirs descriptors

(according to Theune-Vogt, 1991)

A - Monochromatic

A1. Method of production (wound or drawn glass)

A2. Shape (oval, prismatic, biconical, cilindrical,..)

A3. Color (brown, white, green-blue, ...)

B - Polychromatic

B1. Decoration (spots, eyes, waves, double swag, ...)

B2. Shape (oval, prismatic, biconical, cilindrical,..)

A3. Color of the ground and decoration glass

Bead symbolic description

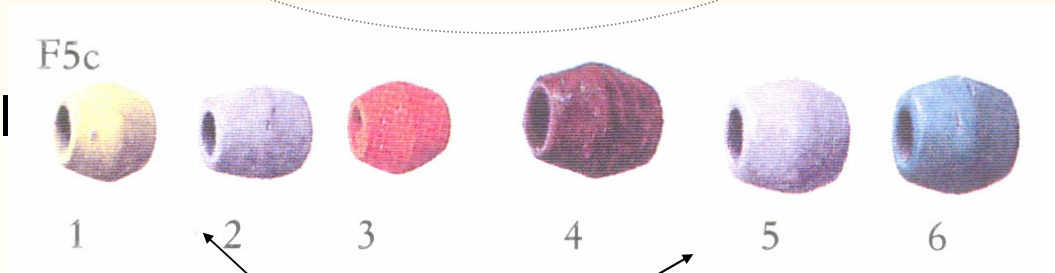
[Color = {"green-blue"}]^
[Decoration = {"monochromatic"}]^
[Shape = {"biconical"}]



from expert knowledge

Representation of a
"typology" of beads
through a symbolic description

Obtained by a conceptual
generalization process



by shape :

biconical

Classification by graves

Symbolic description of the BEADS discovered in the grave 81:

[Color = {"Brown"}]^
 [Decoration = { "Monochromic",
 "colored waves",
 "little colored spot"}]^
 [Shape = {"barrel", "Cylindrical"}]

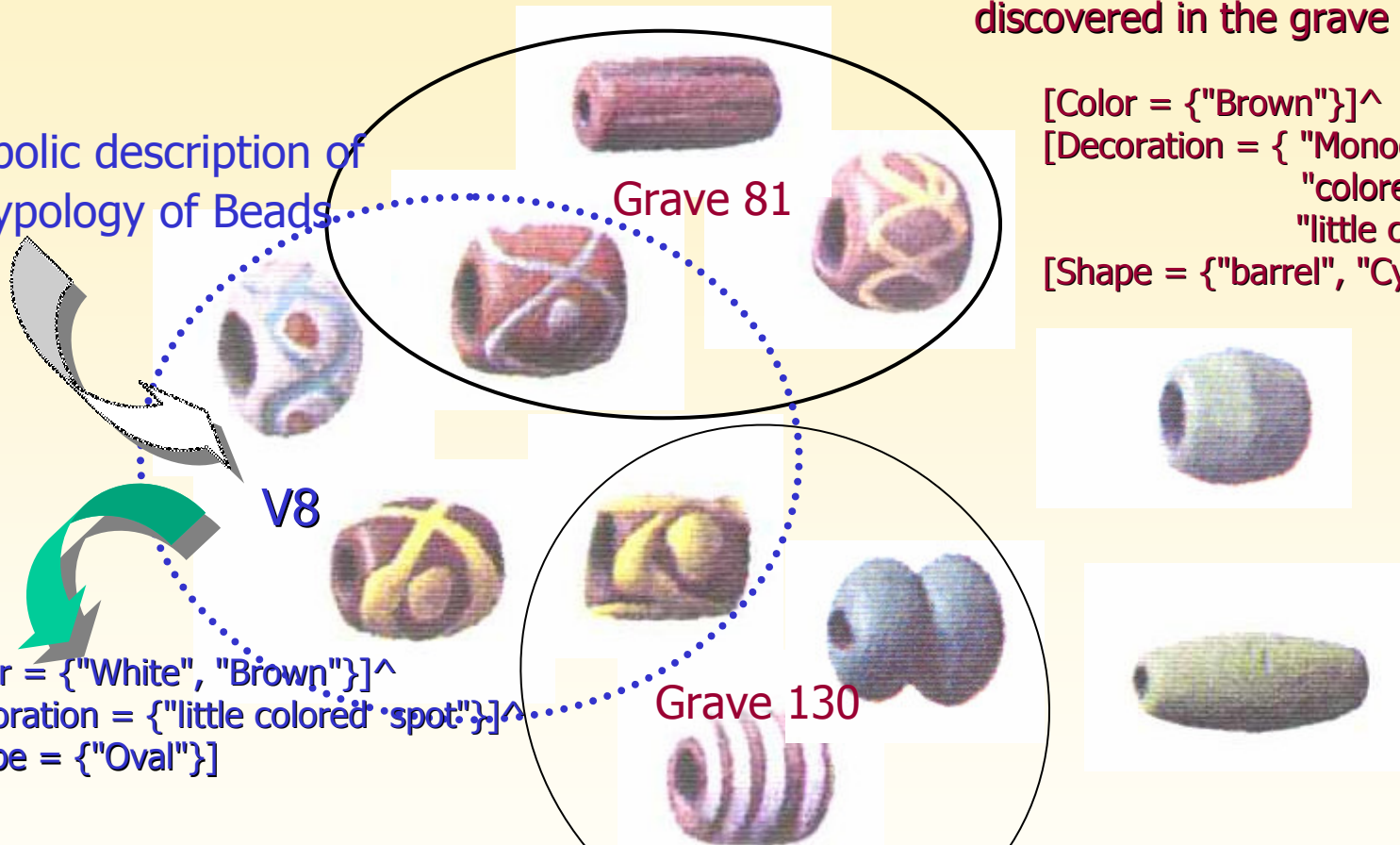
Symbolic description of a typology of Beads...

V8

[Color = {"White", "Brown"}]^
 [Decoration = {"little colored spot"}]^
 [Shape = {"Oval"}]

Symbolic description of the BEADS discovered in the Grave 130:

[Color = {"Brown", "Green-blue"}]^
 [Decoration = { "Monochromic",
 "spiraes", "little colored spot"}]^
 [Shape = {"barrel", "Cylindrical", "segmented"}]



Merovingian beads CODING IN symbolic data table

	<i>Color</i>	<i>Decoration</i>	<i>Contours</i>	<i>Shape</i>	
T222	Green-blue; Yellow; Dark blue Translucid glass; White milk Brown	Yes, No	Monochromic colored waves dots	barrel Cylindrical segmented Oval; prismatic	
T244	Translucid glass White milk	No	Monochromic	oval Segmented	
T53	Brown Crimson	Yes	spiraes Reticella combed spirale	Cylindrical oval	

The **Classification** plays a main role in the framework of

Symbolic Data Analysis

SCA provides to:



- connect similar concepts
(**classes** of Symbolic Objects) or



- generalize some concepts
(classes, **second order objects**)

The basic choices of the dynamic clustering algorithm

E : set of the elements to be partitioned

→ E : set of the Symbolic Objects

y : set of the variables which describe the elements of E

→ y : set of the SO's descriptors

ψ : proximity measure

→ ψ : proximity function able to compare couples of Symbolic Objects

Symbolic Objects Classes

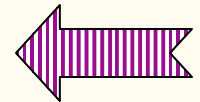
Whereas the elements of the set E to be partitioned are *Symbolic Objects*:

a class C_i of the partition P of E assumes a
« *more complex meaning* » than
the simple union of the elements belonging to it
Conceptual classification – (Michalski, Stepp, Diday, '81,...)

The Classification « *looks for a classes structure* » :

Knowledge discovery –

capability to represent the data (*classes*)



Dynamic clustering algorithm

» *nuées dynamiques* « (Diday, 1972)

Dynamic clustering algorithm optimises a **criterion Δ** which measures **the best fitting** between a partition P of the set of objects E in k classes and the way to represent the classes $\{C_1, \dots, C_i, \dots, C_k\} \in P$

$$\Delta(P^*, L^*) = \text{Min}\{\Delta(P, L) \mid P \in P_k, L \in L_k\}$$

Representation Function

For the best description and interpretation of the classes

$$g: P_k \rightarrow L_k \quad \Rightarrow \quad g(P) = L$$

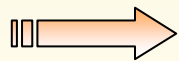
with $P = (C_1, \dots, C_k)$ and $L = (G_1, \dots, G_k)$

The most representative element/s of the class
such that :

$$\min_{G_j \in L_k} \sum_{s \in C_j} \Delta_j(s, G_j)$$

... in the dynamic clustering algorithm

"nuées dynamiques"



G_j

can be :

"seed"

- the **center of gravity (average)**
- a **group of elements**
- a **linear function**
- a **probability function ...**

... in the symbolic context :

G_j



*prototype of the class:
Symbolic Object*

Allocation function $\psi: L_k \rightarrow P_k \Rightarrow \psi(L)=P$

with $C_i = \{s \in E \mid \psi(s, G_i) < \psi(s, G_j)\}$ (for $i \neq j$)

... in the symbolic clustering analysis

ψ is a suitable *Proximity Measure*

defined on Symbolic Objects

*L_2 norm, Φ^2 distance, Hausdroff, Ichino's distance,
two components distance (De Carvalho)*

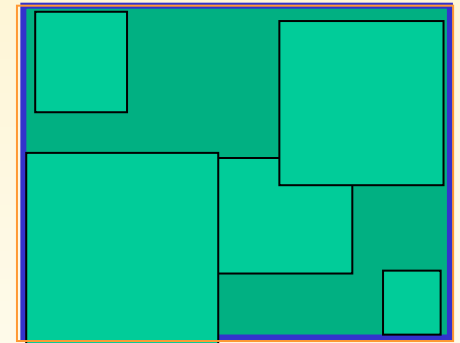
First step

Choice of the best representation of a cluster according to Proximity function



A SO class is bestly represented by its elements

- *Class of Symbolic Objects - CSO -*
Proximity constraints between the elements (symbolic objects) which belong to a class
- *Symbolic Object Class - OSC -*
Generalization of the characteristics of the elements (symbolic objects) belonging to a class



Moreover

In order to take into account the different **distribution of the characteristics** presented by the SO's belonging to the class :

$G_i \Rightarrow \text{modal SO}$

Prototype : Modal Object

The representation of the prototype of each class is a SO **which generalises the characteristics of the SO's belonging to the class**

$$G_i = (s_1 \oplus \dots \oplus s_{ni})$$

with description:

$$d(G_i) = \wedge_{j=1..p} [y_j(G_i), q_j(G_i)]$$

Taking into account the **distributions of the values of its** descriptors

$$y_j(G_i) = y_j(s_1) \oplus \dots \oplus y_j(s_k)$$

Empirical distribution associates to the multi-nominal descriptors y_j of the G_i

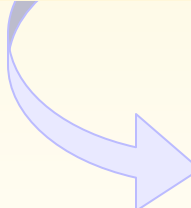
Let's D_j the domain of a multi-nominal descriptors y_j

$$D_j \equiv \{m_1, \dots, m_H\}$$

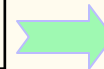
$d_j(s)$ is the set of values in the space of representation of s

for example

$$d_j(s) = \{m_1, m_2, m_H\}$$

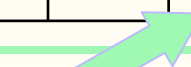


y_j	m_1	m_2	m_3	...	m_H
s	X	X			X



y_j	q_1	q_2	q_3	...	q_H
s	1/3	1/3			1/3

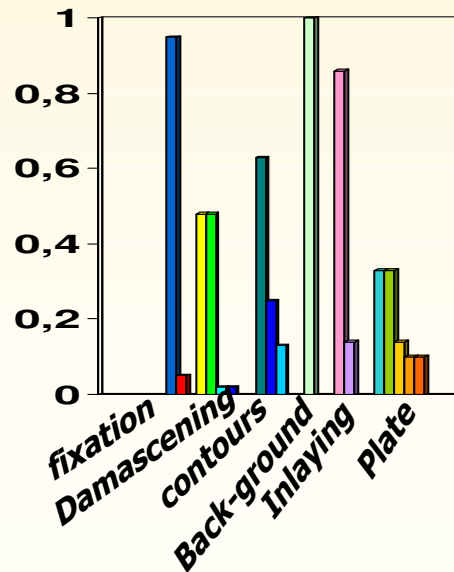
Diagram of the frequencies (modes)



$$q_{j,m_h}(s) = \begin{cases} \frac{1}{|d_{j,m_h}|} & \text{if } m_h \in d_{j,s} \\ 0 & \text{otherwise} \end{cases}$$

Empirical distribution associated to the set-valued variables of $d(G_i)$

The representation of C_i is given by the distributions of the multi-nominal variables which described the objects belonging to the class :



$$g_j(G_i) = f(q_j(s) \mid \{s\} \in C_i)$$

Diagram of the frequencies (modes)

$$g_{j,h}(G_i) = \frac{1}{n_i} \sum_{s \in C_i} q_{j,h}(s)$$

Results of the classification of the
Merovingian women graves

Symbolic Data Set description

The set of symbolic data is constituted by **39 burials**
described with respect the characteristics of 222 beads
characterized by **3 multi-nominal** descriptors with the **31 categories**:

Color \in {Dark blue ; Olive green ; Brown ;White milk; Dark green; Crimson; Dark green; Crimson; Green-blue; translucent glass ; Black; White; Yellow; Orange};

Decoration \in {Monochromatic; colored waves; spirales; Reticella; combed spirale, dots; followers, little colored spot};:

Shape \in {Cylindrical; almond; biconical ; segmented; Prismatic; oval; barrel; melon}.

It is possible to consider also a logical rule between descriptors:

If color = monochromatic then decoration = inapplicable

Partition of the 39 graves in 2 classes according to the beads descriptions

Class 1 #20 Graves



Index of spread from the prototype

Graves

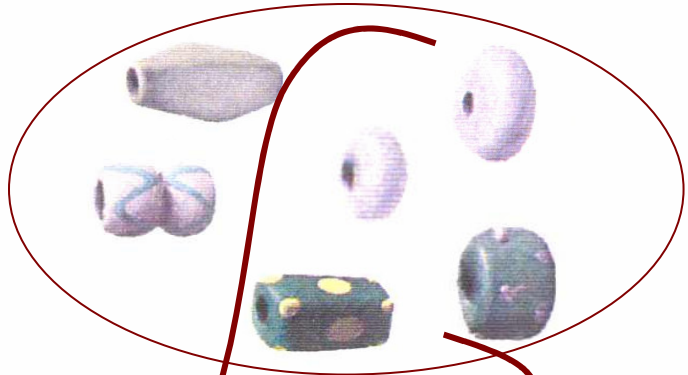
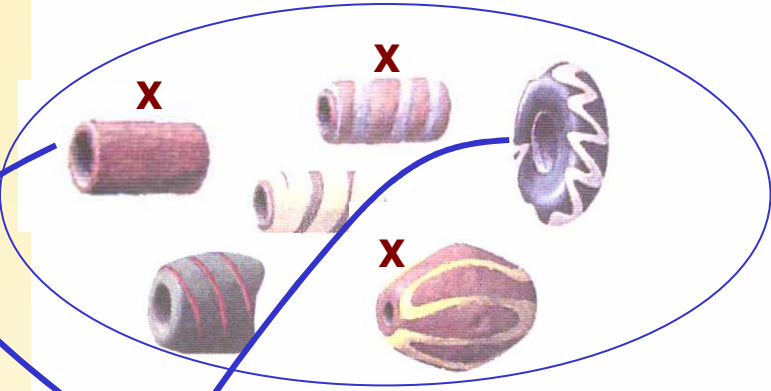
Class 2 #19 Graves

"G197" [0.6]	"G 90" [0.9]	"G 94" [1.2]
"G100" [0.7]	"G 97" [0.9]	"G 84" [1.2]
"G 58" [0.8]	"G130" [0.9]	"G 78" [1.3]
"G175" [0.8]	"G103" [0.9]	"G101" [1.3]
"G202" [0.8]	"G109" [1.0]	"G 61" [1.4]
"G239" [0.8]	"G 81" [1.0]	"G188" [1.5]
"G 53" [0.8]		"G155" [1.5]
"G156" [0.3]	"G85" [0.9]	"G108" [1.2]
"G 89" [0.4]	"G 8" [1.0]	"G 87" [1.2]
"G270" [0.6]		"G131" [1.3]
"G222" [0.6]		"G118" [1.3]
"G233" [0.7]		"G244" [1.3]
"G191" [0.7]		"G148" [1.5]
"G132" [0.7]		"G111" [1.6]
"G 80" [0.8]		"G 1" [1.2]
		"G 56" [1.4]

Partition of the SO set in 2 classes: Prototypes description

	<i>Color</i>	<i>Decoration</i>	<i>Shape</i>		
G₁  (B)	Dark blue 0,04; Olive green 0,02; Brown 0,43 ; White milk 0,03; Dark green 0,02; Crimson 0,09; Green-blue 0,04; Translucid glass 0,03; Black 0,2 ; White 0,01; Yellow 0,05; Orange 0,03.	Monochromic 0,27 ; colored waves 0,15 ; spirales 0,21 ; Reticella 0,01; combed spirale 0,2; dots 0,08; flowers 0,02. little colored spot 0,05	Cylindrical 0,29 ; Almond 0,02; Biconical 0,18 ; segmented 0,04; prismatic 0,02; oval 0,21 ; Barrel 0,21 ; melon 0,03.		
G₂  (A)	Dark blue 0,09; Olive green 0,07; Brown 0,09; White milk 0,21 ; Dark green 0,04; Crimson 0,07; Green-blue 0,13 ; Translucid glass 0,1; Black 0,07; White 0,03; Yellow 0,07; Orange 0,02.	Monochromic 0,46 ; colored waves 0,17 ; spirales 0; dots 0,19 ; Reticella 0,04; combed spirale 0,02; flowers 0,05; little colored spot 0,08.	Cylindrical 0,11; almond 0,09; Biconical 0,14 ; Barrel 0,06; Segmented 0,09; Prismatic 0,17 ; Oval 0,29 ; Melon 0,05.		

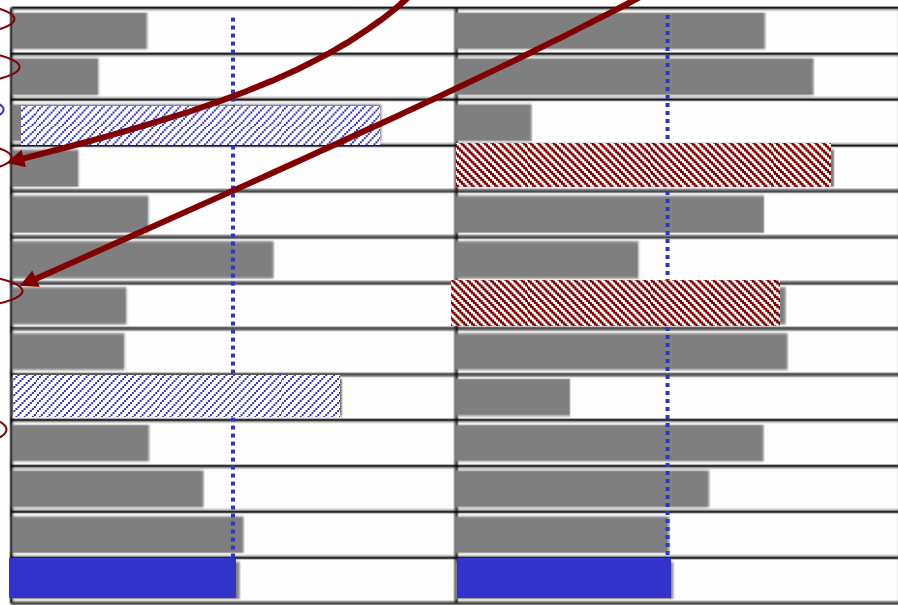
***The procedure has been run 30 times and it stopped at
the best value of the global criterion: $\Delta=58,54$***



Prismatic 0,17; Oval 0,29;

Color

- Dark blue
- Olive green
- Brown
- White milk
- Dark green
- Crimson
- Green-blue
- Translucid glass
- Black
- White
- Yellow
- Orange
- Set



Class 1

Class 2

Prototypes representation of the class $C_1 \Rightarrow B$ and $C_2 \Rightarrow A$ according to the characteristics of the beads

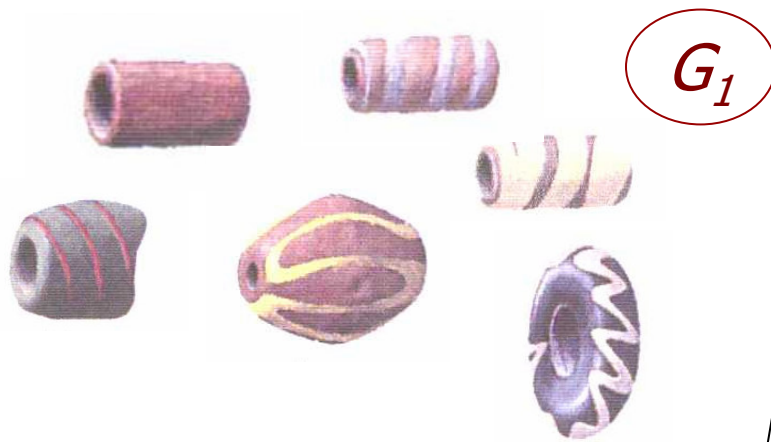
Expert classification

{Brown 0,43; Black 0,2; Monochromatic 0,27} ^
{colored waves 0,15; spirales 0,21} ^
{cylindrical 0,29; biconical 0,18; oval 0,21; barrel 0,21}

End of 5th until 1st half of 6th century:

*Oval monochromatic; Polycromatic types,
colored waves; prismatic*

Reticella beads



G_1

*2nd half of 6th until the middle of 7th
century:*

*Biconical monochromatic; cylindrical;
multicolored Polycromatic types, black with
many colored spots*



G_2

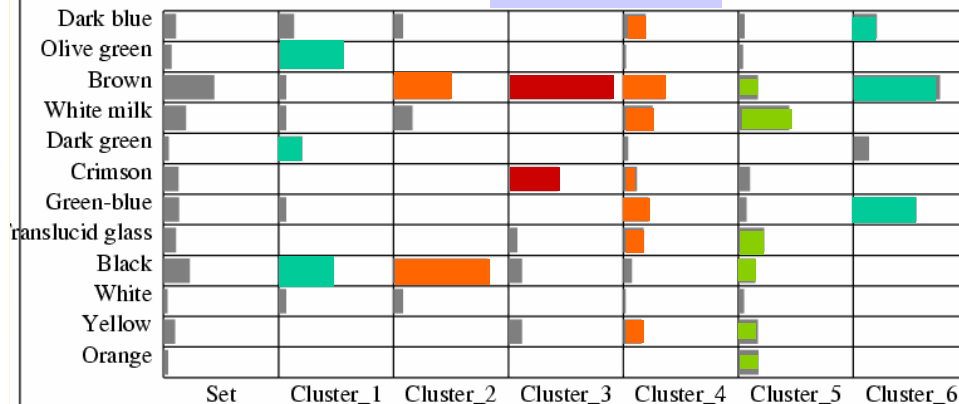
{White milk 0,21; Green-blue 0,13} ^
{Monochromatic 0,46; colored waves 0,17; dots 0,19} ^
{Biconical 0,14; Prismatic 0,17; Oval 0,29; }

Partition of the 39 graves in 6 classes according to the symbolic descriptions of the beads retrieved inside

Class 1 #4 Graves	"T233" [0.7]	"T118" [1.0]	"T109" [1.1] "T87" [1.2]
Class 2 #5 Graves	"T197" [0.2] "T100" [0.5]	"T103" [1.0] "T97" [1.1]	"T188" [2.2]
Class 3 #8 Graves	"T53" [0.4] "T58" [0.7] "T202" [0.7] "T132" [0.8]	"T84" [1.1]	"T78" [1.4] "T155" [1.5] "T101" [1.3]
Class 4 #9 Graves	"T89" [0.3] "T191" [0.6] "T222" [0.4]	"T85" [0.9] "T80" [1.0]	"T108" [1.4] "T 1" [1.4] "T56" [1.3] "T61" [1.9]
Class 5 #6 Graves	"T270" [0.4] "T156" [0.2]		"T244" [1.5] "T155" [1.5] "T148" [1.6] "T94" [1.3]
Class 6 #7 Graves	"T90" [0.6] "T130" [0.7]	"T131" [1.1] "T239" [0.9] "T 8" [1.0]	"T111" [1.7] "T81" [1.1]

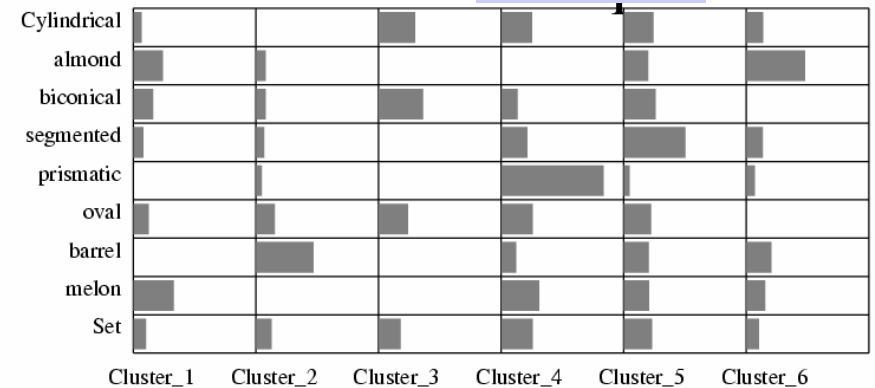
Partition of the set in 6 classes

Color

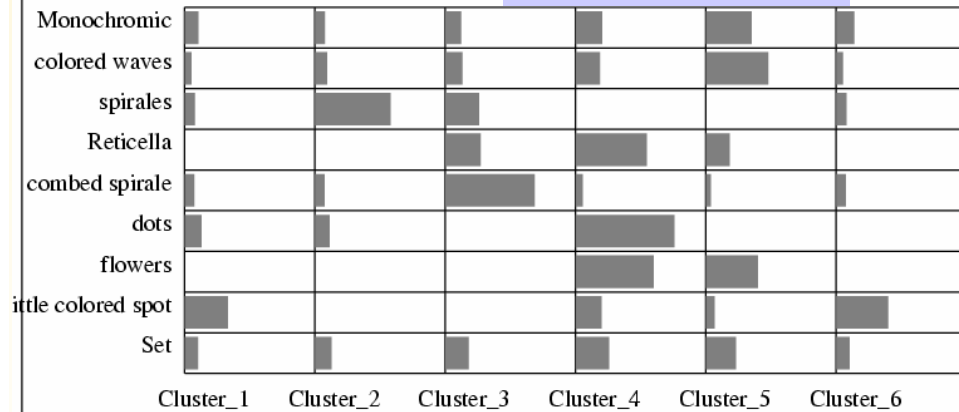


Prototype representations according to characteristics of the beads retrieved in the graves of the different classes

Shape

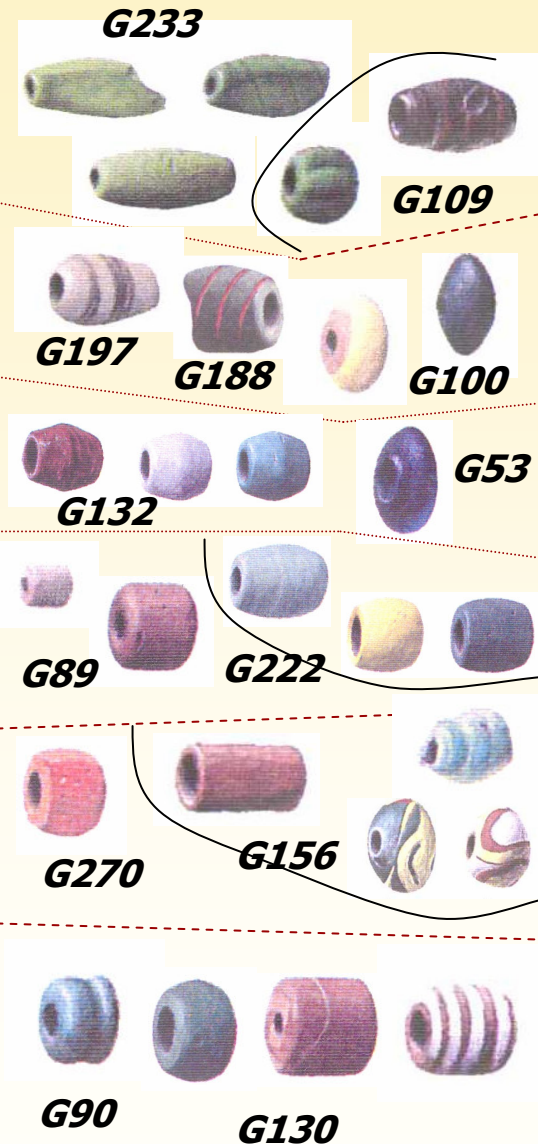


Decoration

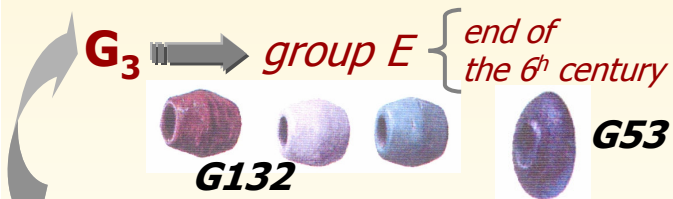


Partition in 6 classes: prototypes description

	Color	Decoration	Shape
G₁	Olive green (0.33); Black (0.29); Dark green (0.13)	Monochromatic (0.37); dots (0.17); little colored spot (0.21)	Cylindrical (0.13); almond (0.13); barrel (0.26); oval (0.3); melon (0.13)
G₂	Brown (0.3); Black (0.5)	Monochromatic (0.22); spiraes (0.48); dots (0.12)	oval (0.3); barrel (0.5)
G₃	Brown (0.55); Crimson (0.26)	Monochromatic (0.22); spiraes (0.15); combed spirale (0.44)	Cylindrical (0.33); biconical (0.33); oval (0.33)
G₄	Dark blue (0.12); Brown (0.22); White milk (0.16); Green Blue (0.13)	Monochromatic (0.29); colored waves (0.12); dots (0.39)	Cylindrical (0.2); prismatic (0.3); oval (0.25)
G₅	White milk (0.27); Translucid grass (0.14)	Monochromatic (0.56); colored waves (0.33)	Cylindrical (0.21); Biconical (0.19); segmented (0.14); oval (0.24); barrel (0.12)
G₆	Brown (0.46); Green blue (0.33)	Monochromatic (0.5); little colored spot (0.25)	Cylindrical (0.27); almond (0.25); barrel (0.27)

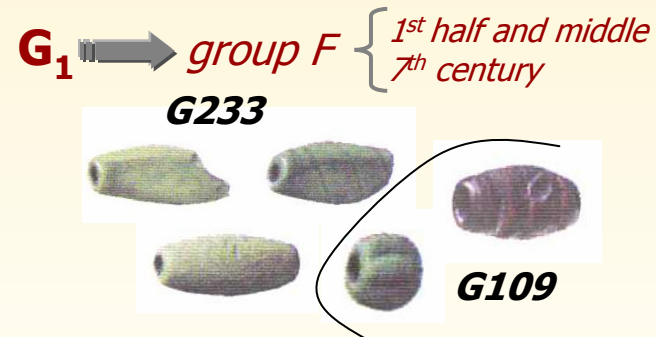
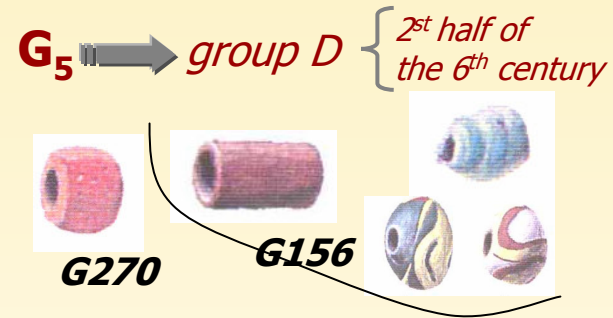


***Relationships between the classes descriptions
and the graves chronology (Theune-Vogt, 91)***

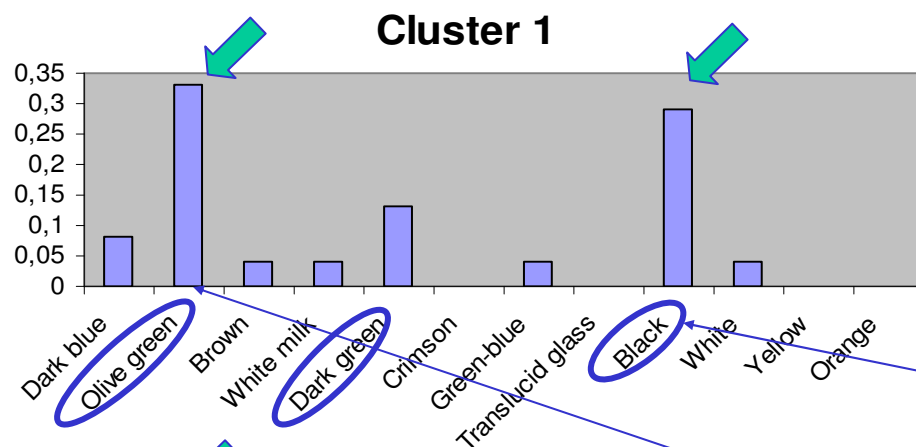


The appearance of the necklaces is characterised by many multicolored polychromic beads and by bicone beads in various colors.

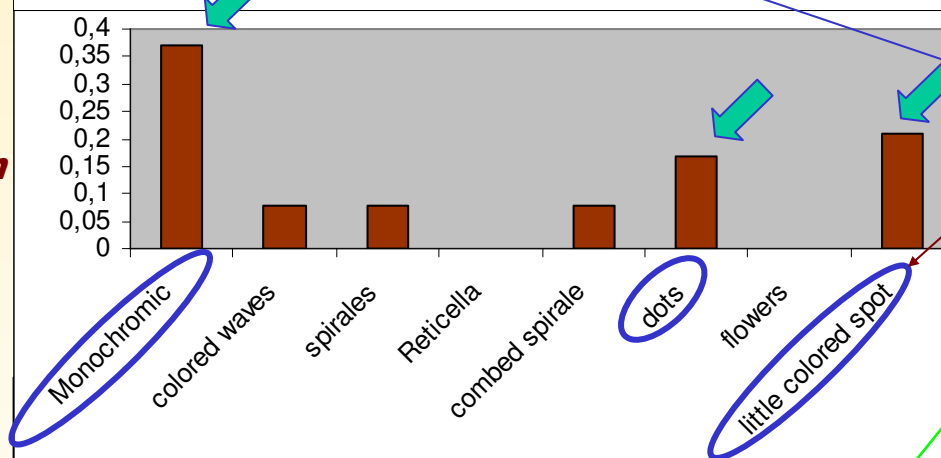
Among polychromic beads those with double swang, spiral linen and short cylindrical beads are notable.



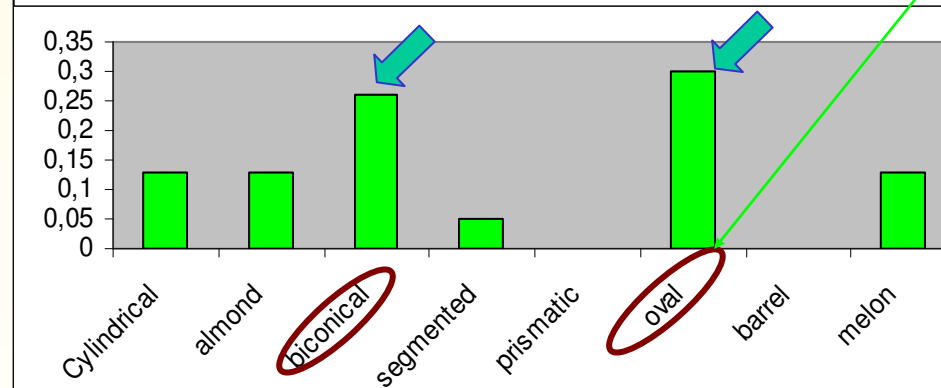
COLOR



Decoration



SHAPE



*PROTOTYPE REPRESENTATION of the
CLUSTER 1*

Corresponding to the Group F

Dated : 1st half of the middle of the 7th century

*Small black beads with many
colored small spots*

and

**Symbolic
classification**

*Oval olive-green beads typical
of the the **GROUP 5***

*Regular
spots*



*Monocromatics beads, mostly
biconical in dark color*

or

**First beads of the late
7th century A.D.**

Spots decoration

Conclusions

- ✱ *Numerical optimization in the interpretation context of the conceptual classification*
- ✱ *Preserving of the flexibility and generality of the Dynamical Partitioning in the knowledge discovery*
- ✱ *Cluster characterization by considering the distributions of the symbolic descriptors*
- ✱ *Criteria : Generality and specificity*

Working in progress

- Extension of the clustering algorithm to the partition of a set of SO's described by **any kind of variables** (real, at intervals, nominal, multi-nominal and modal)
- Clustering of SO's characterized by **logical relationships** among their descriptors (expressed through **Boolean and/or probabilistic rules**)
- Determination of the **optimal number** of SO clusters