

**Notice du programme  
CHAVL**



Nom du programme : CHAVL  
Date de mise à jour : Mars 1992  
Numéro de version : 1.2  
Langage : FORTRAN 77  
Auteurs du programme : P. PETER (IRESTE)  
H. LEREDDE (Univ. PARIS NORD),  
I.C. LERMAN (IRISA),

IRESTE  
CP 3003 - La Chantrerie  
44087 NANTES CEDEX 03 - Tél. (+33) 40 68 30 00  
e.mail : ppeter@ireste.fr

UNIVERSITE PARIS NORD  
Av. J.B. Clément  
93430 VILLETANEUSE

IRISA  
Campus de Beaulieu  
Av. du Gal Leclerc  
35042 RENNES CEDEX - Tél. (+33) 99 84 71 00  
e.mail : lerman@irisa.fr

## **A - PRESENTATION DU PRODUIT**

### **1. PRESENTATION GENERALE**

Le programme CHAVL propose une chaîne complète de classification ascendante hiérarchique permettant de classer aussi bien les individus que les variables descriptives. La méthodologie mise en oeuvre est celle de la Vraisemblance du Lien mise au point par I.C. LERMAN.

Le traitement s'articule en plusieurs étapes :

- **Calcul des similarités entre objets :**  
étape SIMOB pour les individus ou  
étape ASVAR pour les variables,
- **Construction de l'arbre :** étape AVLVR,
- **Calcul d'indices d'aide à l'interprétation :** étape INTRP,
- **Dessin de l'arbre :** étape DESAB.

Ces étapes communiquent par l'intermédiaire de fichiers formatés. Un traitement peut n'utiliser que certaines étapes.

### **2. REFERENCES BIBLIOGRAPHIQUES**

I.C. LERMAN : Classification et analyse ordinale des données - DUNOD 1981.

I.C. LERMAN et P. PETER : Elaboration et logiciel d'un indice de similarité entre objets d'un type quelconque. Application au problème de consensus en classification - Rapport de recherche INRIA, Rennes juillet 1985.

### **3. FICHIERS UTILISES**

Les numéros et noms de fichiers sont gérés dans le module principal CHAVL. Le contenu des différents fichiers sera explicité dans la documentation des étapes concernées. Un résumé se trouve à la fin de cette documentation (juste avant les jeux d'essai).

a) Le fichier des paramètres : *chavl.par*

Nom dans le programme : IFENT ou IFICH(1), étiquette Numéro 10, fichier formaté en entrée, utilisé par le module principal et par les étapes SIMOB, ASVAR et DESAB.

b) Le fichier des résultats : *chavl.lis*

Nom dans le programme : IFSORT ou IFICH(2), étiquette numéro 11, fichier formaté en sortie, utilisé par le module principal et par toutes les étapes.

c) Le fichier des données : *chavl.don*

Nom dans le programme : IFDON ou IFICH(3), étiquette numéro 12, fichier formaté en entrée, utilisé par l'étape SIMOB ou ASVAR.

d) Le fichier de travail : *chavl.do1*

Nom dans le programme : IFDO1 ou IFICH(4), étiquette numéro 13, fichier formaté créé par l'étape SIMOB ou ASVAR et utilisé par les étapes AVLVR et INTRP.

e) Le fichier de travail : *chavl.do2*

Nom dans le programme : IFDO2 ou IFICH(5), étiquette numéro 14, fichier formaté créé par l'étape AVLVR, utilisé et complété par l'étape INTRP et utilisé par l'étape DESAB.

f) Le fichier dictionnaire : *chavl.dic*

Nom dans le programme : IFDIC ou IFICH(6), étiquette numéro 15, fichier formaté en entrée, utilisé par les étapes INTRP et DESAB.

g) Les fichiers temporaires : *chavl.tp1* et *chavl.tp2*

Noms respectifs dans le programme : IFTAB (ou IFICH(7)) et IFTABT (ou IFICH(8)), étiquettes 16 et 17, fichiers temporaires non formatés, créés et utilisés par l'étape SIMOB.

h) Les fichiers de dessin de l'arbre : *chavl.ar1..9*

Noms respectifs dans le programme : IFARB ou IFICH(9) à IFICH(17), étiquettes numéros 21 à 29, fichiers formatés en sortie, utilisés par l'étape DESAB.

#### 4. NOTICE D'UTILISATION DU MODULE PRINCIPAL

Ce module ne sert qu'à gérer les différentes étapes et n'utilise que le fichier des paramètres (*chavl.par*). En fonction des données initiales et des résultats souhaités, certaines étapes peuvent être facultatives.

Le fichier des paramètres devra contenir :

<b>LIGNE 1</b>	<b>FORMAT (A80)</b>
<b>TITRE</b>	Titre de l'étude

<b>LIGNE 2</b>	<b>FORMAT (4I1)</b>
la liste des étapes à exécuter dans le vecteur ETAPE	
ETAPE(1)	= 1 étape SIMOB, = 2 étape ASVAR, = 0 pas de calcul de la matrice des similarités.
ETAPE(2)	= 1 étape AVLVR, = 0 pas de construction de l'arbre.
ETAPE(3)	= 1 étape INTRP, = 0 pas de calcul d'aide à l'interprétation.
ETAPE(4)	= 1 étape DESAB, = 0 pas de représentation graphique de l'arbre.

#### **LIGNES SUIVANTES**

les paramètres des différentes étapes à exécuter (dans leur ordre d'exécution).

## B - PRESENTATION DES ETAPES

### I - ETAPE SIMOB

#### 1 - DESCRIPTION DE LA METHODE

Cette étape assure le calcul des indices de similarité entre individus décrits par différents types de variables. Les variables descriptives traitées sont :

- \* les variables quantitatives,
- \* les attributs logiques,
- \* les variables qualitatives nominales,
- \* les variables qualitatives ordinales,
- \* les variables qualitatives préordonnances,
- \* les tableaux de contingence.

Ces indices ont été élaborés avec la méthodologie de *l'Algorithme de la Vraisemblance du Lien*. Une réduction statistique globale est réalisée sur les similarités.

**ATTENTION** : les tableaux Individus x Variables qualitatives et les tableaux de contingence sont des tableaux de nombres entiers, ils doivent donc être lus avec un format en Ix.

#### 2 - PLACE MEMOIRE OCCUPEE

Elle dépend directement du type des variables descriptives utilisées.

Type des variables	Tableaux d'entiers	Tableaux de réels
var. numériques	0	$\max(N(N+7)/2, 12N+P)$
attributs logiques	$11N+P$	$N(N+5)/2$
var. qual. nominales	$11N+2P$	$N(N+1)/2 + QMAX$
var. qual. ordinales	$11N+2P$	$N(N+1)/2 + QMAX$
val. qual. preord.	$\max(11N+2P, P+QMAX)$	$N(N+1)/2 + NRGT$
tab. de contingence	$11N+2P+NBST$	$N(N+4NBST+3)/2+2NBST+P$

avec :

N nombre de lignes (d'individus),

P : nombre de colonnes (de variables),

QMAX : nombre maximum de modalités (variables qualitatives),

NBST : nombre de sous-tableaux (tableau de contingence),

NRGT :  $\{\sum Q_i(Q_{i+1})/2, i=1, P\}$  où  $Q_i$  est le nombre de modalités de la variable  $i$ .

### 3 - NOTICE D'EMPLOI DE L'ETAPE

#### 3.1 Liste des fichiers utilisés

**chavl.par** (IFENT), étiquette No 10, fichier formaté en entrée : fichier des paramètres,

**chavl.lis** (IFSORT), étiquette No 11, fichier formaté en sortie : fichier des résultats,

**chavl.don** (IFDON), étiquette No 12, fichier formaté en entrée : fichier des données,

**chavl.do1** (IFDO1), étiquette No 13, fichier de travail formaté en sortie,

**chavl.tp1** (IFTAB), étiquette No 16, fichier temporaire non formaté : contient le tableau des données,

**chavl.tp2** (IFTABT), étiquette No 17, fichier temporaire non formaté : contient le fichier des données transposé.



## 3.2 Entrées

### 3.2.1 Fichier des paramètres chavl.par

LIGNE 1	FORMAT (4I5)
NOBS	Nombre d'individus
NVAR	Nombre de variables
INDI	Type des variables : = 1 : variables quantitatives = 2 : attributs logiques = 3 : variables qualitatives nominales = 4 : variables qualitatives ordinales = 5 : variables qualitatives préordonnances = 6 : tableau de contingence
PAR4	* Variables qualitatives : code de la première modalité * Tableau de contingence : nombre de sous-tableaux * Autres types de variables : 0

LIGNE 2	FORMAT (A80)
FMTD	Format de lecture des données

LIGNE 3	FORMAT (20I4)
NBMOD	* Variables qualitatives : nombre de modalités par variable * Tableau de contingence : nombre de modalités par sous-tableau * Autres types de variables : néant

Les lignes suivantes (4 et 5) ne concernent que les variables qualitatives préordonnances.

LIGNE 4	FORMAT (A80)
FMTRG	Format de lecture de la table des rangs.

LIGNE 5	FORMAT FMTRG
Table des rangs (demi-matrice triangulaire inférieure) : Les rangs doivent être donnés variable par variable, et pour une variable donnée dans l'ordre suivant : $r(1,1)$ , $r(1,2)$ , $r(2,2)$ , ..., $r(1,i)$ , $r(2,i)$ , ..., $r(i,i)$ , ..., $r(q,q)$ si la variable possède $q$ modalités (on retourne à a ligne pour la variable suivante).	

### 3.2.2 Fichier des données chavl.don

Il contient les données dans le format spécifié dans le fichier des paramètres.

## 3.3 Sorties

### 3.3.1 Fichier des résultats chavl.lis

Il contient un compte-rendu de l'exécution de l'étape avec le rappel des paramètres et les éventuelles erreurs détectées.

### 3.3.2 Fichier de travail chavl.do1

Il contient :

**Sur la ligne 1** en format (I5) : le nombre d'individus,

**Sur la ligne 2** : le format des lignes suivantes : (8F10.5),

**Sur les lignes suivantes** : la demi-matrice inférieure des similarités entre individus rangée ligne par ligne dans le format spécifié sur la ligne 2.

## 3.4 Gestion des erreurs

Le programme contrôle la dimension des super-tableaux. De plus, il s'assure lors des réductions qu'aucune variance ne soit nulle. En outre, l'existence et la validité des fichiers d'entrée sont également vérifiées.

## II - ETAPE ASVAR

### 1 - DESCRIPTION DE LA METHODE

Cette étape assure le calcul des indices de similarité entre variables. Les variables descriptives traitées sont :

- \* les variables quantitatives
- \* les attributs logiques
- \* les variables qualitatives nominales
- \* les variables qualitatives ordinales

Ces indices ont été élaborés avec la méthodologie de l'Algorithme de la Vraisemblance du Lien. Une réduction statistique globale est réalisée sur les similarités.

**ATTENTION** : les tableaux Individus x Variables qualitatives sont des tableaux de nombres entiers, ils doivent donc être lus avec un format en Ix.

## 2 - PLACE MEMOIRE OCCUPEE

Elle dépend directement du type des variables descriptives traitées.

Type des variables	Tableaux d'entiers	Tableaux de réels
var. numériques	0	$NVAR(NVAR+7)/2$
attributs logiques	NVAR	$NVAR(NVAR+5)/2$
var. qual. nominales	$NOBS*NVAR$	$NVAR(NVAR+5)/2+MODMX(MODMX+1)$
var. qual. ordinales	$NOBS*NVAR$	$(NVAR+9)/2+MODMX(MODMX+1)$

où NOBS est le nombre d'individus,  
 NVAR est le nombre de variables,  
 MODMX est le nombre maximum de modalités.

## 3 - NOTICE D'EMPLOI DE L'ETAPE

### 3.1 Liste des fichiers utilisés

**chavl.par** (IFENT), étiquette No 10, fichier formaté en entrée : fichier des paramètres,

**chavl.lis** (IFSORT), étiquette No 11, fichier formaté en sortie : fichier des résultats,

**chavl.don** (IFDON), étiquette No 12, fichier formaté en entrée : fichier des données,

**chavl.do1** (IFDO1), étiquette No 13, fichier de travail formaté en sortie,

## 3.2 Entrées

### 3.2.1 Fichier des paramètres chavl.par

LIGNE 1	FORMAT (4I5)
NOBS	Nombre d'individus
NVAR	Nombre de variables
INDI	Type des variables : = 1 : variables quantitatives = 2 : attributs logiques = 3 : variables qualitatives nominales = 4 : variables qualitatives ordinales
PAR4	* Variables qualitatives : code de la première modalité * Autres types de variables : 0

LIGNE 2	FORMAT (A80)
FMTD	Format de lecture des données

LIGNE 3	FORMAT (20I4)
Variables qualitatives uniquement : nombre de modalités par variable	

### 3.2.2 Fichier des données chavl.don

Il contient les données dans le format spécifié dans le fichier des paramètres.

## 3.3 Sorties

### 3.3.1 Fichier des résultats chavl.lis

Il contient un compte-rendu de l'exécution de l'étape avec le rappel des paramètres et les éventuelles erreurs détectées.

### 3.3.2 Fichier de travail chavl.do1

Il contient :

**Sur la ligne 1** en format (I5) : le nombre d'individus,

**Sur la ligne 2** : le format des lignes suivantes : (8F10.5),

**Sur les lignes suivantes** : la demi-matrice inférieure des similarités entre individus rangée ligne par ligne dans le format spécifié sur la ligne 2.

### 3.4 Gestion des erreurs

Le programme contrôle la dimension des super-tableaux. De plus, il s'assure lors des réductions qu'aucune variance ne soit nulle. En outre, l'existence et la validité des fichiers d'entrée sont également vérifiées.

## III - ETAPE AVLVR

### 1 - DESCRIPTION DE LA METHODE

Cette étape assure la classification ascendante hiérarchique d'un ensemble d'éléments. L'algorithme employé est celui des voisins réciproques, et le critère d'agrégation utilisé est celui de l'Algorithme de la Vraisemblance du Lien. Cette étape nécessite le calcul préalable de la demi-matrice triangulaire inférieure des indices de similarité (ce calcul peut être réalisé par l'étape SIMOB ou ASVAR).

L'arbre de classification est établi en représentation polonaise préfixée. On peut ensuite enchaîner les étapes INTRP et DESAB afin d'obtenir des indices d'aide à l'interprétation des résultats ainsi qu'une représentation semi-graphique de l'arbre de classification.

### 2 - PLACE MEMOIRE OCCUPEE

tableaux d'entiers	tableaux de réels	tableaux de booléens
13 N	$N(N+5)/2$	N

où N est le nombre d'éléments à classier.

## 3 - NOTICE D'EMPLOI DE L'ETAPE

### 3.1 Liste des fichiers utilisés

**chavl.lis** (IFSORT), étiquette No 11, fichier formaté en sortie : fichier des résultats,

**chavl.do1** (IFDO1), étiquette No 13, fichier de travail formaté en entrée,

**chavl.do2** (IFDO2), étiquette No 14, fichier formaté en sortie.

## 3.2 Entrées

### 3.2.1 Fichier des données chavl.do1

Ce fichier a été généré automatiquement si l'étape SIMOB ou l'étape ASVAR a été utilisée auparavant. Sinon, il doit être créé et doit contenir :

LIGNE 1	FORMAT (I5)
NOBS	Nombre d'éléments à classifier

LIGNE 2	FORMAT (A80)
FMTDO1	Format de lecture des similarités

LIGNE 3	FORMAT FMTDO1
Demi-matrice inférieure des similarités entre individus rangée ligne par ligne	

## 3.3 Sorties

### 3.3.1 Fichier des résultats chavl.lis

Il contient :

- \* la représentation polonaise préfixée de l'arbre de classification en format (16I5),
- \* dans le cas où une erreur aurait été détectée, le message d'erreur correspondant.

### 3.3.2 Fichier de travail chavl.do2

Il contient :

**Sur la ligne 1** en format (2I5) : le nombre d'individus NOBS et le nombre de niveaux de l'arbre NBNIV,

**Sur la ligne 2** : le format des lignes suivantes,

**Sur les lignes suivantes** : la représentation polonaise de l'arbre dans le format spécifié sur la ligne 2.

### 3.4 Gestion des erreurs

Le programme contrôle la dimension des super-tableaux. En outre, l'existence et la validité des fichiers d'entrée sont également vérifiées.

## IV - ETAPE INTRP

### 1 - DESCRIPTION DE LA METHODE

Cette étape calcule des indices d'aide à l'interprétation des résultats d'une classification ascendante hiérarchique réalisée avec l'Algorithme de la Vraisemblance du Lien (elle fait suite aux modules SIMOB/ASVAR puis AVLVR). Elle nécessite en entrée la demi-matrice des indices de similarité et la représentation polonaise préfixée de l'arbre de classification.

Les aides à l'interprétation calculées sont :

- \* les individus les plus influents lors de la formation des classes, conformément aux indices de dispersion,

- \* les statistiques locales et globales des niveaux de l'arbre.

### 2 - PLACE MEMOIRE OCCUPEE

Tableaux d'entiers	Tableaux de réels	Tableaux de chaînes (24 car.)
N	$N(N+5)/2$	N

### 3 - NOTICE D'EMPLOI DE L'ETAPE

#### 3.1 Liste des fichiers utilisés

**chavl.dic** (IFDIC), étiquette No 15, fichier formaté en entrée : fichier des libellés des éléments à classier,

**chavl.lis** (IFSORT), étiquette No 11, fichier formaté en sortie : fichier des résultats,

**chavl.do1** (IFDO1), étiquette No 13, fichier de travail formaté en entrée,

chavl.do2 (IFDO2), étiquette No 14, fichier formaté en entrée puis en sortie.

### 3.2 Entrées

#### 3.2.1 Fichier dictionnaire chavl.dic

Ce fichier doit contenir le libellé des éléments à classifier en format (A24).

#### 3.2.2 Fichier des données chavl.do1

Ce fichier a été généré automatiquement si l'étape SIMOB ou l'étape ASVAR a été utilisée auparavant. Sinon, il doit être créé et doit contenir :

LIGNE 1	FORMAT (I5)
NOBS	Nombre d'éléments à classifier

LIGNE 2	FORMAT (A80)
FMTDO1	Format de lecture des similarités

LIGNE 3	FORMAT FMTDO1
Demi-matrice inférieure des similarités entre individus rangée ligne par ligne	

#### 3.2.3 Fichier de travail chavl.do2

Ce fichier a été généré automatiquement si l'étape AVLVR a été utilisée auparavant (il sera de plus complété à la fin de cette étape) ; sinon il doit contenir :

LIGNE 1	FORMAT (2I5)
NOBS	Nombre d'éléments à classifier
NBNIV	Nombre de niveaux de l'arbre

LIGNE 2	FORMAT (A80)
FMTDO2	Format de lecture de la représentation polonaise

LIGNE 3	FORMAT FMTDO2
Représentation polonaise de l'arbre	



### 3.3 Sorties

#### 3.3.1 Fichier des résultats *chavl.lis*

Il contient :

- \* le classement des individus par valeurs de dispersion décroissantes,
- \* les statistiques locales et globales des niveaux de l'arbre avec indication des maxima locaux de la statistique locale,
- \* dans le cas où une erreur aurait été détectée, le message d'erreur correspondant.

#### 3.3.2 Fichier de travail *chavl.do2*

A la fin de cette étape, il contient :

**Sur la ligne 1** en format (3I5) : le nombre d'individus NOBS, le nombre de niveaux de l'arbre NBNIV et le nombre NBNVSG de niveaux significatifs,

**Sur la ligne 2** : le format de la représentation polonaise de l'arbre,

**Sur les lignes suivantes** : la représentation polonaise de l'arbre dans le format spécifié sur la ligne 2.

**Sur la ligne suivante** : le format de la liste des niveaux significatifs,

**Sur les lignes suivantes** : la liste des niveaux significatifs (en ordre croissant) dans le format précisé sur la ligne précédente.

### 3.4 Gestion des erreurs

Le programme contrôle la dimension des super-tableaux. En outre, l'existence et la validité des fichiers d'entrée sont également vérifiées.

## V - ETAPE DESAB

### 1 - DESCRIPTION DE L'ETAPE

Cette étape construit une représentation graphique d'un arbre codé en représentation polonaise préfixée. L'arbre est condensé sur des niveaux sélectionnés au préalable ; les niveaux non sélectionnés sont automatiquement rehaussés sur le dessin au niveau retenu immédiatement supérieur (le dernier niveau est automatiquement retenu). Ces niveaux peuvent être les niveaux significatifs de l'arbre de classification.

### 2 - PLACE MEMOIRE OCCUPEE

Soient NOBS le nombre d'individus et NBNV le nombre de niveaux retenus.

Tableaux de booléens :  $NBNV+1$ ,

Tableaux de caractères : NOBS chaînes de 24 caractères,

Tableaux d'entiers :  $2*NOBS*NBNV + 10*NOBS + 3*NBNV + 2$ .

### 3 - NOTICE D'EMPLOI DE L'ETAPE

#### 3.1 Liste des fichiers utilisés

**chavl.par** (IFENT), étiquette No 10, fichier formaté en entrée : fichier des paramètres,

**chavl.dic** (IFDIC), étiquette No 15, fichier formaté en entrée : fichier des libellés des éléments à classifier,

**chavl.lis** (IFSORT), étiquette No 11, fichier formaté en sortie : fichier des résultats,

**chavl.do2** (IFDO2), étiquette No 14, fichier formaté en entrée,

**chavl.ar1** à **chavl.ar9** (IFARB), étiquettes No 21 à 29, fichiers formatés en sortie destinés à recevoir le dessin de l'arbre.

## 3.2 Entrées

### 3.2.1 Fichier des paramètres *chavl.par*

LIGNE 1	FORMAT (2I4)
LG	= 0 : dessin sur 132 caractères par ligne = 1 : dessin sur 80 caractères par ligne
IFREQ	= 0 : condensation de l'arbre aux niveaux significatifs > 0 : un niveau sur IFREQ est sélectionné pour condenser l'arbre

### 3.2.2 Fichier dictionnaire *chavl.dic*

Ce fichier doit contenir le libellé des éléments à classifier en format (A24).

### 3.2.3 Fichier de travail *chavl.do2*

Ce fichier a été généré automatiquement si l'étape AVLVR a été utilisée auparavant. Dans le cas où l'on souhaite condenser l'arbre à des niveaux sélectionnés (et seulement dans ce cas), les deux dernières lignes sont obligatoires. Lorsque INTRP a été employé auparavant, ce fichier est créé et les niveaux sélectionnés sont les niveaux significatifs de l'arbre. Si ce fichier n'existe pas ou est incomplet il doit être créé (ou complété) et contenir les informations suivantes :

LIGNE 1	FORMAT (3I5)
NOBS	Nombre d'éléments à classifier Nombre de niveaux de l'arbre Nombre de niveaux sélectionnés pour condenser l'arbre, (facultatif quand IFREQ > 0)
NBNIV	
NBNVSG	

LIGNE 2	FORMAT (A80)
FMTDO2	Format de lecture de la représentation polonaise

LIGNE 3	FORMAT FMTDO2
Représentation polonaise de l'arbre	

Les lignes suivantes ne sont nécessaires que si IFREQ = 0

LIGNE 4	FORMAT (A80)
FMTDO3	Format de lecture des niveaux sélectionnés

LIGNE 5	FORMAT FMTDO3
Liste des niveaux sélectionnés pour la condensation de l'arbre	

### 3.3 Sorties

#### 3.3.1 Fichier des résultats *chavl.lis*

On y ajoute les paramètres de l'étape, de plus, si une erreur a été détectée, ce fichier contient le message correspondant.

#### 3.3.2 Fichiers de dessin *chavl.ar1* à *chavl.ar9*

C'est dans ces fichiers que se trouve le dessin de l'arbre. Si l'arbre est trop haut pour contenir dans un seul fichier (*chavl.ar1*), son dessin se poursuit dans les fichiers de dessin suivants (jusqu'à *chavl.ar9*)

### 3.4 Gestion des erreurs

Le programme contrôle la dimension des super-tableaux. En outre, l'existence et la validité des fichiers d'entrée sont également vérifiées.

## VI- RESUME

### 1- FICHIERS UTILISES

Nom	Unité	Fonctions
<i>chavl.par</i>	IFCENT=10	paramètres (F)
<i>chavl.don</i>	IFDON=12	données (F)
<i>chavl.dic</i>	IFDIC=15	dictionnaire des libellés (F)
<i>chavl.lis</i>	IFSORT=11	listing (F)
<i>chavl.ar1-9</i>	IFARB=21 à 29	dessin de l'arbre (F)
<i>chavl.do1</i>	IFDO1=13	travail (demi-matrice) (F)
<i>chavl.do2</i>	IFDO2=14	travail (F)
<i>chavl.tp1</i>	IFTAB=16	travail (NF)
<i>chavl.tp2</i>	IFTAB=17	travail (NF)

(F) : fichier formaté

(NF) : fichier non formaté

## 2- ENTREES

CHAVL : RESUME DES ENTRES			
Fichier	Ligne	Format	
Paramètres	1	A80	Titre
	2	4I1	vecteur ETAPE
	3	4I5	NOBS, NVAR, INDI, PAR4
	4	A80	FMTD
	5	20I4	NBMOD (*)
	6	A80	FMTRG (*)
	7	FMTRG	RANGS (*)
	8	2I4	LG, IFREQ
Dictionnaire	1	A24	Libellés des éléments à classier (un par ligne)
Données	1	FMTD	Tableau des observations

Ce tableau correspond à un traitement complet (du calcul des indices au dessin de l'arbre). Pour un traitement partiel, certaines lignes peuvent être à éliminer et les fichiers chavl.do1 et chavl.do2 devront être construits (voir documentation détaillée)

(\*) Ligne présente uniquement pour certains types de données.

## C - JEUX D'ESSAI

### I - CLASSIFICATION DES VARIABLES

Source des données :

M. JAMBU

"Exploration informatique et statistique des données", DUNOD 1989.

#### 1 - Fichier chavl.don

HAUS	610	140	60	10	120	95	115	760	175	315
FAUS	475	90	250	30	140	120	100	775	115	305
FNAU	10	0	495	110	170	110	130	785	160	430
HMUS	615	140	65	10	115	90	115	765	180	305
FMUS	179	29	421	87	161	112	119	776	143	373
HCUS	585	115	50	0	150	105	100	760	150	385
FCUS	482	94	196	18	141	130	96	775	132	336
HAWE	653	100	95	7	57	85	150	808	115	330
FAWE	511	70	307	30	80	95	142	816	87	262
FNAW	20	7	568	87	112	90	180	843	125	368
HMWE	656	97	97	10	52	85	152	808	122	321
FMWE	168	22	528	69	102	83	174	824	119	311
HCWE	643	105	72	0	62	77	140	813	100	388
FCWE	429	34	262	14	92	97	147	849	84	392
HAYO	650	140	120	15	85	90	105	760	70	365
FAYO	560	105	375	45	90	90	95	745	60	235
FNAY	10	10	710	55	145	85	130	815	60	380
HMYO	650	145	112	15	85	90	105	760	80	358
FMYO	260	52	576	59	116	85	117	775	65	295
HCYO	615	125	95	0	115	90	85	760	40	475
FCYO	433	89	318	23	112	96	102	774	45	408
HAES	650	142	122	22	76	94	100	764	96	334
FAES	578	106	338	42	106	94	92	752	64	228
FNAE	24	8	594	72	158	92	128	840	86	398
HMES	652	133	134	22	68	94	102	763	122	310
FMES	436	79	433	60	119	90	107	772	73	231
HCES	627	148	68	0	88	92	86	770	58	463
FCES	434	86	297	21	129	102	94	799	58	380

#### 2 - Fichier chavl.par

```

DONNEES DE JAMBU
2111
  28  10  1  0
(4X,10F5.0)
  0  1

```

#### 3 - Fichier chavl.dic

```

TRAVAIL
TRANSPORT
MENAGE
ENFANTS
COURSES
TOILETTE
REPAS
SOMMEIL
TELEVISION
LOISIRS AUTRES QUE TELE

```

#### 4 - Fichier de travail chavl.do1 (après ASVAR)

```

10
(8F10.5)
1.00000
2.11459 1.00000
-1.90493 -1.82114 1.00000
-1.82245 -1.68972 1.95769 1.00000
-1.35982 -1.02161 1.17094 1.26164 1.00000
-0.16849 -0.09617 0.00034 0.34743 1.37206 1.00000
-0.91642 -1.26130 0.86445 0.87855 -0.32545 -0.70844 1.00000
-1.09768 -1.45558 1.02242 0.68120 0.01171 -0.39790 1.86094
1.00002
-0.05147 -0.01885 -0.37319 0.34241 0.54755 0.77919 0.76781
0.11573
1.00000
-0.33709 -0.15317 -0.17028 -0.16160 0.59083 0.23641 -0.01000
0.53053
-0.13164 1.00000

```

#### 5 - Fichier de travail chavl.do2

##### 5.1 - après AVLVR

```

10 9
(16i5)
-9 -1 1 2 -8 -5 -2 3 4 -4 5 6 -7 -6
-3 7 8 9 10 0

```

##### 5.2 - après INTRP

```

10 9 1
(16i5)
-9 -1 1 2 -8 -5 -2 3 4 -4 5 6 -7 -6
-3 7 8 9 10 0
(16i5)
8

```

#### 6 - Fichier chavl.lis

PROGRAMME DE CLASSIFICATION HIERARCHIQUE CHAVL  
\*\*\*\*\*

TITRE DE L'ETUDE

DONNEES DE JAMBU

1 \*\*\* ETAPE ASVAR \*\*\*

PARAMETRES DE L'ETAPE

Nombre d'individus : 28

Nombre de variables : 10  
Ce sont des variables numeriques

Format de lecture des donnees :  
(4X,10F5.0)

--- FIN DE L'ETAPE ASVAR ---

1 \*\*\* ETAPE AVLVR \*\*\*

## REPRESENTATION POLONAISE DE L'ARBRE

-9 -1 1 2 -8 -5 -2 3 4 -4 5 6 -7 -6 -3 7  
8 9 10 0

--- FIN DE L'ETAPE AVLVR ---

1 \*\*\* ETAPE INTRP \*\*\*

RANGEMENT DES ELEMENTS PAR VALEURS DE DISPERSIONS CROISSANTES  
\*\*\*\*\*

ELEMENT	10	: LOISIRS AUTRES QUE TELE	DISPERSION :	0.08805
ELEMENT	9	: TELEVISION	DISPERSION :	0.13957
ELEMENT	6	: TOILETTE	DISPERSION :	0.31780
ELEMENT	5	: COURSES	DISPERSION :	0.79878
ELEMENT	7	: REPAS	DISPERSION :	0.84781
ELEMENT	8	: SOMMEIL	DISPERSION :	0.85878
ELEMENT	1	: TRAVAIL	DISPERSION :	1.23868
ELEMENT	2	: TRANSPORT	DISPERSION :	1.25104
ELEMENT	4	: ENFANTS	DISPERSION :	1.27785
ELEMENT	3	: MENAGE	DISPERSION :	1.40541

STATISTIQUES DES NIVEAUX  
\*\*\*\*\*

	NIVEAU	STATISTIQUE GLOBALE	STATISTIQUE LOCALE
	1	2.1146	2.1146
	2	2.9176	0.8030
	3	3.5178	0.6003
	4	3.8033	0.2855
	5	3.8794	0.0761
	6	3.8951	0.0157
	7	3.7088	-0.1863
1 MAXIMUM	8	4.5967	0.8878
	9	0.0005	-4.5962

--- FIN DE L'ETAPE INTRP ---

1 \*\*\* ETAPE DESAB \*\*\*

## PARAMETRES DE L'ETAPE

Nombre d'elements : 10  
 Nombre de niveaux reels de l'arbre : 9  
 Nombre de niveaux representes : 9  
 Un niveau sur 1 est represente  
 Longueur d'une ligne du dessin : 132

--- FIN DE L'ETAPE DESAB ---

FIN DU PROGRAMME CHAVL  
\*\*\*\*\*



## 7 - Fichier chavl.ar1

TRAVAIL	>---*								
TRANSPORT	> *1-----*								I
MENAGE	>-----*								I
ENFANTS	> *2-----*								I
COURSES	>-----*								I
TOILETTE	>-----*								I
REPAS	>-----*								I
SOMMEIL	> *3-----*								I
TELEVISION	>-----*								I
LOISIRS AUTRES QUE TELE	>-----*								I

## II - CLASSIFICATION DES INDIVIDUS AVEC V.Q. PREORDONNANCES

### 1 - Fichier chavl.don

123  
312  
211  
222  
111  
321  
113  
323

### 2 - Fichier chavl.par

```

jeu test : Individus x Variables Qualitatives Preordonnances
1111
  8   3   5   1
(3I1)
  3   2   3
(10F8.2)
  5.00  1.50  5.00  3.00  1.50  5.00
  2.00  1.00  2.00
  5.00  2.00  5.00  1.00  2.00  5.00
  1   0

```

### 3 - Fichier chavl.dic

INDIVIDU 1 : \*123\*  
INDIVIDU 2 : \*312\*  
INDIVIDU 3 : \*211\*  
INDIVIDU 4 : \*222\*  
INDIVIDU 5 : \*111\*  
INDIVIDU 6 : \*321\*  
INDIVIDU 7 : \*113\*  
INDIVIDU 8 : \*323\*

#### 4 - Fichier de travail chavl.do1 (après SIMOB)

```

      8
(8F10.5)
  0.00000
-0.74462  0.00000
-1.89196  0.00641  0.00000
  0.00641 -0.19913  0.22088  0.00000
-0.20232  0.73054  1.27603 -1.46875  0.00000
  0.30734  0.22088 -0.19913  0.00641  0.52500  0.00000
  1.49051  0.73054 -0.41680 -1.46875  1.27284 -1.16783  0.00000
  2.00016  0.22088 -1.89196  0.00641 -1.16783  1.27284  0.52500  0.00000

```

#### 5 - Fichier de travail chavl.do2

##### 5.1 - après AVLVR

```

      8      7
(16i5)
  -7  -6  -4  -3  -1  1  8  7  6  -5  2  -2  3  5  4  0

```

##### 5.2 - après INTRP

```

      8      7      2
(16i5)
  -7  -6  -4  -3  -1  1  8  7  6  -5  2  -2  3  5  4  0
(16i5)
  3      5

```

#### 6 - Fichier chavl.lis

PROGRAMME DE CLASSIFICATION HIERARCHIQUE CHAVL  
 \*\*\*\*\*

TITRE DE L'ETUDE

jeu test : Individus x Variables Qualitatives Preordonnances

1 \*\*\* ETAPE SIMOB \*\*\*

RAPPEL DES PARAMETRES  
 -----

NOMBRE D'INDIVIDUS : 8

NOMBRE DE VARIABLES : 3

VARIABLES QUALITATIVES PREORDONNANCES

Premiere modalite codee a : 1

FORMAT DES DONNEES : (311)

--- FIN DE L'ETAPE SIMOB ---

1 \*\*\* ETAPE AVLVR \*\*\*

REPRESENTATION POLONAISE DE L'ARBRE

```

-7  -6  -4  -3  -1  1  8  7  6  -5  2  -2  3  5  4  0

```

--- FIN DE L'ETAPE AVLVR ---

1 \*\*\* ETAPE INTRP \*\*\*

RANGEMENT DES ELEMENTS PAR VALEURS DE DISPERSIONS CROISSANTES  
\*\*\*\*\*

ELEMENT	2	: INDIVIDU 2	: "312"	DISPERSION :	0.20532
ELEMENT	6	: INDIVIDU 6	: "321"	DISPERSION :	0.41575
ELEMENT	4	: INDIVIDU 4	: "222"	DISPERSION :	0.41929
ELEMENT	5	: INDIVIDU 5	: "111"	DISPERSION :	0.93789
ELEMENT	3	: INDIVIDU 3	: "211"	DISPERSION :	1.00009
ELEMENT	7	: INDIVIDU 7	: "113"	DISPERSION :	1.02866
ELEMENT	1	: INDIVIDU 1	: "123"	DISPERSION :	1.29689
ELEMENT	8	: INDIVIDU 8	: "323"	DISPERSION :	1.34650

STATISTIQUES DES NIVEAUX  
\*\*\*\*\*

	NIVEAU	STATISTIQUE GLOBALE	STATISTIQUE LOCALE
	1	2.0002	2.0002
	2	2.3013	0.3011
1 MAXIMUM	3	2.8367	0.5354
	4	2.5930	-0.2437
2 MAXIMUM	5	2.5147	-0.0782
	6	1.7320	-0.7827
	7	-0.0017	-1.7337

--- FIN DE L'ETAPE INTRP ---

1 \*\*\* ETAPE DESAB \*\*\*

PARAMETRES DE L'ETAPE

Nombre d'elements : 8  
 Nombre de niveaux réels de l'arbre : 7  
 Nombre de niveaux représentés : 3  
 Les niveaux représentés sont les niveaux significatifs  
 Longueur d'une ligne du dessin : 80

--- FIN DE L'ETAPE DESAB ---

FIN DU PROGRAMME CHAVL  
\*\*\*\*\*

## 7 - Fichier chavl.ar1

INDIVIDU 1 : "123"	>---*
	> 1
INDIVIDU 8 : "323"	>---I---*
	> *3 I
INDIVIDU 7 : "113"	>---* 4---*
	> I I
INDIVIDU 6 : "321"	>-----* 6
	> I
INDIVIDU 2 : "312"	>-----* I
	> *5---I---
INDIVIDU 3 : "211"	>---* I I
	> 2---* 7
INDIVIDU 5 : "111"	>---* I
	> I
INDIVIDU 4 : "222"	>-----*

### III - CLASSIFICATION DES INDIVIDUS AVEC VARIABLES NUMERIQUES

Source des données :

"Donnees TODD" MODULAD (supplément à la documentation V2.1 p. 45)

#### 1 - Fichier chavl.don

AUTR	-0.220	-0.006	0.108
FRAN	-0.210	-0.003	-0.110
PORT	-0.369	-0.257	-0.065
ALLF	-0.245	0.017	0.149
BELG	-0.007	0.095	-0.037
FINL	0.258	0.270	0.178
SUED	0.054	0.214	0.058
SUIS	-0.015	0.212	0.211
ITAL	-0.484	-0.287	-0.090
IRLN	0.727	-0.691	0.048
DANM	-0.021	0.289	0.334
ISLA	0.328	0.283	-0.241
ECOS	0.215	0.109	-0.203
ESPA	-0.392	-0.278	-0.083
NORV	0.234	0.250	-0.176
IRLS	0.242	0.100	-0.379
PABA	0.133	0.142	-0.068
ANGL	0.200	0.141	-0.065
USA	0.253	-0.447	0.195

#### 2 - Fichier chavl.par

"Donnees TODD" MODULAD (supplément à la documentation V2.1 p. 45)

```
1111
 19      3      1      0
(5X,3F10.3)
 1      2
```

#### 3 - Fichier chavl.dic

```
AUTRICHE
FRANCE
PORTUGAL
ALLEMAGNE FEDERALE
BELGIQUE
FINLANDE
SUEDE
SUISSE
ITALIE
IRLANDE DU NORD
DANEMARK
ISLANDE
ECOSSE
ESPAGNE
NORVEGE
IRLANDE DU SUD
PAYS-BAS
ANGLETERRE
USA
```

#### 4 - Fichier de travail chavl.do1 (après SIMOB)

```

19
(8F10.5)
0.00000
0.69843 0.00000
1.04736 1.33559 0.00000
1.70925 0.50871 0.84750 0.00000
-0.30362 0.52232 -0.57563 -0.29293 0.00000
-0.20450 -1.16982 -1.30607 -0.03619 0.55315 0.00000
0.04724 -0.51916 -1.09891 0.19796 1.18590 1.51720 0.00000
0.70864 -0.88797 -0.92092 0.89839 0.25679 1.32414 1.37074 0.00000
1.08154 1.40415 1.71967 0.88524 -0.45373 -1.27684 -1.01431 -0.00000
0.00000
-0.59320 -0.71628 0.09012 -0.68152 -1.09519 0.03020 -0.69930 -0.84694
-0.01045 0.00000
0.73290 -0.97996 -0.92154 0.92516 0.09501 1.29085 1.28778 1.71838
-0.87099 -0.80679 0.00000
-1.27183 -0.16030 -0.95436 -1.29227 1.24113 0.66492 0.80393 -0.40688
-0.90133 0.03818 -0.54933 0.00000
-1.53157 -0.06590 -0.77689 -1.61323 0.96294 0.19620 0.28699 -1.02357
-0.74233 0.22840 -1.16564 1.65017 0.00000
1.00175 1.35154 1.72310 0.79645 -0.56443 -1.33900 -1.12627 -0.98042
1.71883 0.09419 -0.98407 -0.92394 -0.73384 0.00000
-1.17858 -0.13212 -0.98073 -1.18555 1.34163 0.73690 0.92837 -0.25923
-0.91797 -0.10090 -0.40531 1.71595 1.59951 -0.95346 0.00000
-1.67666 0.18895 -0.56197 -1.82066 0.79789 -0.40891 -0.21962 -1.62219
-0.52631 0.06239 -1.77585 1.46498 1.65380 -0.50757 1.39571 0.00000
-1.03506 -0.27563 -1.05492 -1.00932 1.33410 0.99833 1.13316 0.05953
-0.99465 -0.04533 -0.07381 1.68166 1.50095 -1.03617 1.70155 1.22380
0.00000
-1.06917 -0.47611 -1.06007 -1.03898 1.10189 1.10642 1.07118 0.07565
-1.02342 0.30826 -0.03587 1.64696 1.49096 -1.04372 1.64445 1.17920
1.68881 0.00000
-0.03732 -0.86336 0.28028 -0.09868 -1.64921 -0.09923 -0.88797 -0.53970
0.16833 1.54633 -0.44995 -0.79301 -0.68446 0.26144 -0.90726 -0.90571
-0.77531 -0.38850 0.00000

```

#### 5 - Fichier de travail chavl.do2

##### 5.1 - après AVLVR

```

19 18
(16i5)
-18 -17 -15 -4 1 4 -10 -9 2 9 -1 3 14 -7 10 19
-16 -14 -12 5 -8 6 7 -2 8 11 -13 -11 -3 12 15 -5
17 18 -6 13 16 0

```

##### 5.2 - après INTRP

```

19 18 3
(16i5)
-18 -17 -15 -4 1 4 -10 -9 2 9 -1 3 14 -7 10 19
-16 -14 -12 5 -8 6 7 -2 8 11 -13 -11 -3 12 15 -5
17 18 -6 13 16 0
(16i5)
11 13 15

```

## 6 - Fichier chavl.lis

PROGRAMME DE CLASSIFICATION HIERARCHIQUE CHAVL  
\*\*\*\*\*

## TITRE DE L'ETUDE

\*Donnees TODD\* MODULAD (supplement a la documentation V2.1 p. 45)

1 \*\*\* ETAPE SIMOB \*\*\*

RAPPEL DES PARAMETRES  
-----

NOMBRE D'INDIVIDUS : 19

NOMBRE DE VARIABLES : 3

VARIABLES NUMERIQUES

FORMAT DES DONNEES : (5X,3F10.3)

---- FIN DE L'ETAPE SIMOB ----

1 \*\*\* ETAPE AVLVR \*\*\*

## REPRESENTATION POLONAISE DE L'ARBRE

-18	-17	-15	-4	1	4	-10	-9	2	9	-1	3	14	-7	10	19
-16	-14	-12	5	-8	6	7	-2	8	11	-13	-11	-3	12	15	-5
17	18	-6	13	16	0										

--- FIN DE L'ETAPE AVLVR ---

1 \*\*\* ETAPE INTRP \*\*\*

RANGEMENT DES ELEMENTS PAR VALEURS DE DISPERSIONS CROISSANTES  
\*\*\*\*\*

ELEMENT	10	: IRLANDE DU NORD	DISPERSION	:	0.33799
ELEMENT	19	: USA	DISPERSION	:	0.44091
ELEMENT	2	: FRANCE	DISPERSION	:	0.61209
ELEMENT	5	: BELGIQUE	DISPERSION	:	0.72976
ELEMENT	7	: SUEDE	DISPERSION	:	0.81325
ELEMENT	6	: FINLANDE	DISPERSION	:	0.82320
ELEMENT	8	: SUISSE	DISPERSION	:	0.84392
ELEMENT	11	: DANEMARK	DISPERSION	:	0.87320
ELEMENT	9	: ITALIE	DISPERSION	:	0.96802
ELEMENT	4	: ALLEMAGNE FEDERALE	DISPERSION	:	0.96803
ELEMENT	1	: AUTRICHE	DISPERSION	:	0.96850
ELEMENT	14	: ESPAGNE	DISPERSION	:	0.99696
ELEMENT	3	: PORTUGAL	DISPERSION	:	1.00286
ELEMENT	18	: ANGLETERRE	DISPERSION	:	1.05366
ELEMENT	17	: PAYS-BAS	DISPERSION	:	1.10596
ELEMENT	15	: NORVEGE	DISPERSION	:	1.15713
ELEMENT	12	: ISLANDE	DISPERSION	:	1.17199
ELEMENT	13	: ECOSSE	DISPERSION	:	1.21228
ELEMENT	16	: IRLANDE DU SUD	DISPERSION	:	1.26147

STATISTIQUES DES NIVEAUX  
 \*\*\*\*\*

	NIVEAU	STATISTIQUE GLOBALE	STATISTIQUE LOCALE	
	1	1.7231	1.7231	
	2	2.4370	0.7139	
	3	2.9862	0.5492	
	4	3.4482	0.4620	
	5	3.8483	0.4001	
	6	4.1982	0.3499	
	7	4.4820	0.2838	
	8	4.7406	0.2586	
	9	4.9496	0.2090	
	10	5.9580	1.0085	
1	MAXIMUM	11	6.9801	1.0220
		12	7.0437	0.0636
2	MAXIMUM	13	8.5420	1.4984
		14	8.9366	0.3946
3	MAXIMUM	15	9.4403	0.5037
		16	9.5017	-0.0613
		17	8.7949	-0.7068
		18	-0.0777	-8.8726

--- FIN DE L'ETAPE INTRP ---

1 \*\*\* ETAPE DESAB \*\*\*

## PARAMETRES DE L'ETAPE

Nombre d'elements : 19  
 Nombre de niveaux reels de l'arbre : 18  
 Nombre de niveaux representes : 9  
 Un niveau sur 2 est represente  
 Longueur d'une ligne du dessin : 80

--- FIN DE L'ETAPE DESAB ---

 FIN DU PROGRAMME CHAVL  
 \*\*\*\*\*

## 7 - Fichier chavl.ar1

AUTRICHE	>-----*			
	> *4-----*			
ALLEMAGNE FEDERALE	>-----*		I	
			15-----*	
FRANCE	>-----*		I	I
		9	I	I
ITALIE	>-----*		I	I
		*10	I	I
PORTUGAL	>-----*			17
	> 1-----*			I
ESPAGNE	>-----*			I
				I
IRLANDE DU NORD	>-----*			I
		7-----*		I
USA	>-----*			I-----
				I
BELGIQUE	>-----*			I
		*12-----*		I
FINLANDE	>-----*		I	I
		*8-----*	I	18
SUEDE	>-----*		*14-----*	I
			I	I
SUISSE	>-----*		I	I
	> *2-----*		I	I
DANEMARK	>-----*		I	I
			I	I
ISLANDE	>-----*		*16-----*	
	> 3-----*		I	I
NORVEGE	>-----*		I	I
		11-----*	I	I
PAYS-BAS	>-----*		I	I
	> 5-----*		I	I
ANGLETERRE	>-----*		13-----*	
			I	
ECOSSE	>-----*		I	
	> *6-----*			
IRLANDE DU SUD	>-----*			