

ABDELWAHEB REBAI

Coalescence floue fondée sur des \mathcal{R} -regroupements maximaux

Mathématiques et sciences humaines, tome 117 (1992), p. 29-47

http://www.numdam.org/item?id=MSH_1992__117__29_0

© Centre d'analyse et de mathématiques sociales de l'EHESS, 1992, tous droits réservés.

L'accès aux archives de la revue « Mathématiques et sciences humaines » (<http://msh.revues.org/>) implique l'accord avec les conditions générales d'utilisation (<http://www.numdam.org/conditions>). Toute utilisation commerciale ou impression systématique est constitutive d'une infraction pénale. Toute copie ou impression de ce fichier doit contenir la présente mention de copyright.

NUMDAM

Article numérisé dans le cadre du programme
Numérisation de documents anciens mathématiques
<http://www.numdam.org/>

COALESCENCE FLOUE FONDEE SUR DES \mathfrak{R} -REGROUPEMENTS MAXIMAUX

Abdelwaheb REBAI¹

RÉSUMÉ - *Les concepts d'éléments \mathfrak{R} -ressemblants à un prototype X et de \mathfrak{R} -regroupement d'objets introduits dans cet article, sont basés sur la notion de relation de S -comparaison \mathfrak{R} définie au moyen d'un indice scalaire de similarité défini entre sous-ensembles flous. Cette relation tient compte du fait que la similarité et la non-dissimilarité des sous-ensembles flous ne sont pas en général des synonymes. Une technique de coalescence floue basée sur des \mathfrak{R} -regroupements maximaux est également introduite.*

SUMMARY - *Fuzzy clustering based on maximal \mathfrak{R} -bunches*
The concepts of \mathfrak{R} -likes of a prototype X and \mathfrak{R} -bunch of objects introduced in this paper are based on the notion of a so-called S -comparison relation \mathfrak{R} whose definition is based on the distinction between fuzzy sets similarity and nondissimilarity. In this paper, the author introduces the basic concepts together with a fuzzy clustering technique based on maximal \mathfrak{R} -bunches.

1. INTRODUCTION

Pour analyser un phénomène d'une manière classificatoire, nous recourons à un regroupement des différentes entités de base qui contribuent à l'étude de ce phénomène dans des classes significatives et moins hétérogènes. Or, quand il s'agit de regrouper des objets donnés dans des classes appropriées de similarité, nous nous trouvons, inévitablement, amené à adopter l'une des trois attitudes possibles, suivantes :

- 1 - une attitude classique qui consiste à affecter chaque objet à une classe et une seule, les agglomérats ou «clusters» formés sont, alors, disjoints et l'agglomération ou coalescence est dite classique ou conventionnelle ;
- 2 - une attitude qui tolère l'appartenance des objets à plus d'une classe, les agglomérats formés sont, alors, empiétants ; et
- 3 - une attitude qui tolère l'appartenance des objets à plusieurs classes possibles, tout en spécifiant leurs degrés d'appartenance à chaque classe, les agglomérats formés sont, alors, flous et la coalescence est dite floue.

Dans le présent article, nous introduisons une méthode de coalescence (cf. figure 1) qui tient compte des différentes considérations suivantes :

- 1 - la similarité et la non-dissimilarité des sous-ensembles flous ne sont pas tout à fait des synonymes (cf. [3]) ; l'homogénéité des agglomérats à former va donc dépendre de deux seuils donnés à l'avance, un seuil pour la similarité et un seuil pour la non-dissimilarité ;
- 2 - cette méthode va permettre de construire des agglomérats non nécessairement disjoints et d'induire des agglomérats flous ;

¹Département des Méthodes Quantitatives, faculté des Sciences Economiques et de Gestion, B.P. n° 69 Sfax 3028 Tunisie.

3 - le nombre d'agglomérats à former n'est pas fixé à l'avance, mais au contraire il est déterminé par la méthode utilisée.

2. PRÉTRAITEMENT DES DONNÉES

Dans toute la suite, Ω va désigner une partie finie d'un univers donné d'objets, $\mathbf{D} = \{d_1, d_2, \dots, d_k\}$ sera un ensemble de k descripteurs classifiants, et \mathbf{I} va indiquer l'intervalle unité $[0,1]$.

$\text{Des}(X) = (d_1(X), d_2(X), \dots, d_k(X))$, où $d_r(X)$ désigne le score de l'objet X pour le descripteur d_r , va être la description de l'objet X , i.e., un vecteur de données cardinales relatives à X .

$\underline{X} = \{\mu_1(d_1(X))/d_1(X), \mu_2(d_2(X))/d_2(X), \dots, \mu_k(d_k(X))/d_k(X)\}$, où $\mu_r(d_r(X))$ est le degré de rapprochement de $d_r(X)$ du score le plus désiré du descripteur d_r pour $r = 1, 2, \dots, k$, va être le profil flou de X . Le degré de rapprochement $\mu_r(d_r(X))$ peut être déterminé en utilisant, par exemple, l'une des fonctions standardisantes suivantes proposées par Zeleny [7].

$$\mu_r(d_r(X)) = \frac{d_r(X)}{d_r^{**}}, \mu_r(d_r(X)) = \frac{d_r^*}{d_r(X)} \quad (\text{échelle rapport}),$$

$$\mu_r(d_r(X)) = \frac{d_r(X) - d_r^*}{d_r^{**} - d_r^*}, \mu_r(d_r(X)) = \frac{d_r^{**} - d_r(X)}{d_r^{**} - d_r^*} \quad (\text{échelle intervalle}),$$

$$\mu_r(d_r(X)) = \left(\frac{1}{2} \left[\frac{d_r(X)}{d_r^c} + \frac{d_r^c}{d_r(X)} \right] \right)^{-1}$$

où d_r^{**} (resp. d_r^*) est le maximum (resp. le minimum) des scores atteints par le descripteur d_r ; et d_r^c une valeur idéale de Coombs pour d_r .

Exemple. Supposons que les scores de trois objets X, Y et Z par rapport à trois descripteurs d_1, d_2 et d_3 sont donnés par la matrice 1. Le niveau le plus désiré pour d_1 est le maximum 20, pour d_2 le minimum 5 et pour d_3 une valeur idéale de Coombs égale à 10. La matrice 2 donne alors les profils flous de ces trois objets.

	d₁	d₂	d₃
X	[20*	10	5]
Y	[10	5*	20]
Z	[5	20	10*]

Matrice 1 : Description des objets

	20	5	10
\underline{X}	[1	.5	.8]
\underline{Y}	[.5	1	.44]
\underline{Z}	[.25	.25	1]

Matrice 2 : Profils flous des objets

3. S-COMPARAISON D'OBJETS

Indice scalaire de similarité entre sous-ensembles flous

Rappelons maintenant la définition d'un indice scalaire de similarité entre sous-ensembles flous telle qu'elle est présentée dans [3].

DÉFINITION Soit \mathcal{A} un ensemble donné et $\mathcal{P}(\mathcal{A})$ l'ensemble des parties floues de \mathcal{A} . Un indice scalaire de similarité entre deux sous-ensembles flous \underline{X} et \underline{Y} définis sur \mathcal{A} est une application $S^* : \mathcal{P}(\mathcal{A}) \times \mathcal{P}(\mathcal{A}) \rightarrow I$, vérifiant :

- i) $S^*(\underline{X}, \underline{Y}) = 1$ si et seulement si $\underline{X} \Delta \underline{Y} = \emptyset$;
- ii) Si \underline{X} et \underline{Y} ont des supports disjoints alors $S^*(X, Y) = 0$;
- iii) $S^*(\underline{X}, \underline{Y}) = S^*(\underline{Y}, \underline{X})$,
- iv) $S^*(\underline{X}, \underline{Y})$ dépend d'une évaluation scalaire de $\overline{\underline{X}} \Delta \overline{\underline{Y}}$ ou d'une évaluation scalaire de $\underline{X} \cup \underline{Y}$ et $\overline{\underline{X}} \cup \overline{\underline{Y}}$.

Dans la définition précédente, Δ désigne une différence symétrique et $\overline{\underline{X}}$ désigne le complémentaire flou de \underline{X} . Un évaluateur scalaire de sous-ensembles flous est une mesure floue dans le sens de Sugeno [5], i.e., une application $g : \mathcal{P}(\mathcal{A}) \rightarrow I$ vérifiant :

- i) $g(\emptyset) = 0$;
- ii) $g(\mathcal{A}) = 1$; et
- iii) Si $\underline{A} \subseteq \underline{B}$, alors $g(\underline{A}) \leq g(\underline{B})$.

A présent, il nous est possible de définir un indice scalaire S de similarité entre objets en utilisant un indice scalaire S^* de similarité entre leurs profils flous, en posant $S(X, Y) = S^*(\underline{X}, \underline{Y})$. La valeur de $S(X, Y)$ va dépendre de la façon de construire les profils flous \underline{X} et \underline{Y} ainsi que de l'indice scalaire S^* de similarité choisi entre sous-ensembles flous. Nous aurons, en particulier :

- i) $S(X, X) = 1$, et
- ii) $S(X, Y) = S(Y, X)$.

Relations de S -comparaison d'objets

A partir de l'indice scalaire S de similarité, il est possible de construire, entre autres, un indice \tilde{S} dit de non-dissimilarité entre X et Y , en posant : $\tilde{S}(X, Y) = \tilde{S}^*(\overline{\underline{X}}, \overline{\underline{Y}})$ et il convient de signaler que $S(X, Y) = \tilde{S}^*(X, Y)$ ne fournissent pas, en général, la même information, ce qui rend leur usage simultané plus significatif pour toute coalescence d'objets. Si nous nous donnons un seuillage $\alpha = (\alpha_1, \alpha_2)$ pris dans $I \times I$, nous pouvons définir une relation binaire Σ_α qui véhicule et synthétise une certaine information sur la satisfaction des seuils α_1 et α_2 respectivement par la similarité et la non-dissimilarité des objets X et Y de la façon suivante :

$$\Sigma_\alpha(X, Y) = \begin{cases} 1 & \text{si } S(X, Y) \geq \alpha_1 \quad \text{et} \quad \tilde{S}(X, Y) \geq \alpha_2 \\ 0 & \text{sinon} \end{cases}$$

La relation Σ_α que nous venons de définir sera appelée relation de S -comparaison construite à partir de l'indice scalaire S de similarité, «plus fine» que les deux relations $\Sigma_{\alpha_1}^1$ et $\Sigma_{\alpha_2}^2$ définies par :

$$\Sigma_{\alpha_1}^1(X, Y) = \begin{cases} 1 & \text{si } S(X, Y) \geq \alpha_1 \\ 0 & \text{sinon} \end{cases} \quad \text{et} \quad \Sigma_{\alpha_2}^2(X, Y) = \begin{cases} 1 & \text{si } \tilde{S}(X, Y) \geq \alpha_2 \\ 0 & \text{sinon} \end{cases} .$$

Remarquons que les relations $\Sigma_{\alpha_1}^1$ et $\Sigma_{\alpha_2}^2$ peuvent être obtenues à partir de Σ_α en posant respectivement $\alpha = (\alpha_1, 0)$ et $\alpha = (0, \alpha_2)$.

Nous pouvons également définir une relation θ_α de la façon suivante :

$$\theta_\alpha(X, Y) = \begin{cases} 1 & \text{si } S(X, Y) \geq \alpha_1 \quad \text{et} \quad \tilde{S}(X, Y) \geq \alpha_2 \\ 0 & \text{sinon} \end{cases} .$$

La relation θ_α que nous venons de définir sera dite relation de S -comparaison « moins fine » que les relations $\Sigma_{\alpha_1}^1$ et $\Sigma_{\alpha_2}^2$.

L'introduction d'un seuillage α et de deux types de relations de S -comparaison « plus fine » et « moins fine » va permettre d'enrichir l'analyse. En effet l'utilisateur peut essayer plusieurs seuillages et voir ce que peut donner une coalescence basée sur une S -comparaison « plus fine » ou « moins fine ». Une telle analyse n'est certainement pas possible en recourant à une méthode de coalescence fondée sur un indice unique et surtout ne nécessitant pas l'usage d'un seuillage.

4. \mathcal{R} -REGROUPEMENTS DE Ω

Dans toute la suite \mathcal{R} va être une relation de S -comparaison et $\mathcal{R}(\Omega)$ la matrice de ressemblance des éléments de Ω qui lui est associée.

DÉFINITION. Soit $\mathcal{M} = [m_{ij}]$ une $n \times m$ -matrice, alors \mathcal{M} sera dite une $\{0,1\}$ -matrice si et seulement si $m_{ij} = 1$ ou 0 pour tout $(i,j) \in \{1, 2, \dots, n\} \times \{1, 2, \dots, m\}$.

DÉFINITION. Etant donnée une $\{0,1\}$ -matrice \mathcal{M} , alors la ligne i et la colonne j de \mathcal{M} seront dites incidentes si et seulement si $m_{ij} = 1$. En particulier, une ligne i incidente avec toutes les colonnes de \mathcal{M} sera dite une ligne complète

Remarque. $\mathcal{R}(\Omega)$ est une $\{0,1\}$ -matrice carrée, symétrique dont toute ligne i est incidente avec la colonne i correspondante.

DÉFINITION. Nous appelons \mathcal{L} -matrice toute $\{0,1\}$ -matrice carrée, symétrique qui possède au moins une ligne complète et dont toute ligne est incidente avec la colonne correspondante.

DÉFINITION. Soit \mathcal{M} une $\{0,1\}$ -matrice et \mathcal{M}' une sous-matrice extraite de \mathcal{M} , alors \mathcal{M}' sera dite complète si et seulement si toutes ses lignes sont complètes.

DÉFINITION. Soit \mathcal{M} une $\{0,1\}$ -matrice et \mathcal{M}' une sous-matrice extraite de \mathcal{M} , alors \mathcal{M}' sera dite première si et seulement si \mathcal{M}' est complète et maximale.

DÉFINITION. Soit $X \in \Omega$, on définit l'ensemble $\mathcal{L}(X)$ des \mathcal{R} -ressemblants de X dans Ω par $\mathcal{L}(X) = \{Y \in \Omega / \mathcal{R}(X,Y) = 1\}$.

Remarque. $\mathcal{R}(\mathcal{L}(X))$ est soit une matrice complète, soit une \mathcal{L} -matrice

DÉFINITION. On définit l'ensemble des objets \mathcal{R} -isolés de Ω par :

$$\text{ISOL}(\Omega) = \{X \in \Omega / \mathcal{L}(X) = X\}.$$

Un objet X appartient donc à $\text{ISOL}(\Omega)$ si et seulement si l'ensemble de ses \mathcal{R} -ressemblants dans Ω est trivial (i.e. un singleton).

DÉFINITION. Soit \mathcal{M} la $\{0,1\}$ -matrice associée à une S -comparaison \mathcal{R} donnant lieu au moins à un élément non isolé et $\mathcal{C}(\mathcal{M})$ l'ensemble des sous-matrices carrées, complètes, extraites de \mathcal{M} , alors nous désignerons par $\mathbf{M}_{\max}(\mathcal{M})$ l'ensemble des éléments de $\mathcal{C}(\mathcal{M})$ possédant le plus grand nombre possible de lignes et de colonnes.

DÉFINITION. Nous appelons \mathcal{R} -regroupement d'objets de Ω toute partie \mathcal{B} de Ω telle que la matrice $\mathcal{R}(\mathcal{B})$ soit complète.

DÉFINITION. Un \mathcal{R} -regroupement \mathcal{B} sera dit maximal dans Ω si la matrice $\mathcal{R}(\mathcal{B})$ est une sous-matrice première de $\mathcal{R}(\Omega)$.

DÉFINITION. Nous appelons **P.G.R** (plus grand \mathcal{R} -regroupement) de Ω tout \mathcal{R} -regroupement \mathcal{B} dont la matrice $\mathcal{R}(\mathcal{B})$ appartient à $\mathbf{M}_{\max}(\mathcal{R}(\Omega))$.

PROPOSITION 1. Tout **P.G.R.** de $\mathcal{L}(X)$ est maximal dans Ω .

Preuve. Montrons que si \mathcal{B} est un **P.G.R** de $\mathcal{L}(X)$ alors il n'existe pas de \mathcal{R} -regroupement \mathcal{B}' maximal dans Ω tel que $\mathcal{B}' \supset \mathcal{B}$. Soit donc \mathcal{B} un **P.G.R** extrait de $\mathcal{L}(X)$, alors \mathcal{B} contient nécessairement X , autrement $\mathcal{B} \cup \{X\}$ sera un \mathcal{R} -regroupement extrait de $\mathcal{L}(X)$ dont la matrice $\mathcal{R}(\mathcal{B} \cup \{X\})$ contient la matrice $\mathcal{R}(\mathcal{B})$ donc $\mathcal{R}(\mathcal{B})$ ne peut pas appartenir à $\mathbf{M}_{\max}(\mathcal{R}(\mathcal{L}(X)))$ ce qui contredit le fait que \mathcal{B} est un **P.G.R** de $\mathcal{L}(X)$.

Supposons maintenant que \mathcal{B} ne soit pas maximal dans Ω , alors il existe un \mathcal{R} -regroupement \mathcal{B}' maximal dans Ω tel que $\mathcal{B}' \supset \mathcal{B}$. Deux cas peuvent, alors, avoir lieu

(a) $(\mathcal{B}' \setminus \mathcal{B}) \cap \mathcal{L}(X) \neq \emptyset$ ou (b) $\mathcal{B}' \setminus \mathcal{B} \subseteq \overline{\mathcal{L}(X)}$, où $\overline{\mathcal{L}(X)}$ est le complémentaire de $\mathcal{L}(X)$ dans Ω .

Si $(\mathcal{B}' \setminus \mathcal{B}) \cap \mathcal{L}(X) \neq \emptyset$ alors $\mathcal{B}'' = \mathcal{B} \cup [(\mathcal{B}' \setminus \mathcal{B}) \cap \mathcal{L}(X)]$ sera un \mathcal{R} -regroupement extrait de $\mathcal{L}(X)$ dont la matrice $\mathcal{R}(\mathcal{B}'')$ contient $\mathcal{R}(\mathcal{B})$ donc $\mathcal{R}(\mathcal{B})$ ne peut pas appartenir à $\mathbf{M}_{\max}(\mathcal{R}(\mathcal{L}(X)))$ ce qui contredit le fait que \mathcal{B} est un **P.G.R** de $\mathcal{L}(X)$ d'où le cas (a) ne peut pas avoir lieu.

Si $\mathcal{B}' \setminus \mathcal{B} \subseteq \overline{\mathcal{L}(X)}$, alors $\mathcal{R}(\mathcal{B}')$ ne peut pas être complète car les colonnes des objets $Y \in \mathcal{B}' / \mathcal{B}$ ne sont pas incidentes avec la ligne de X d'où \mathcal{B}' n'est même pas un \mathcal{R} -regroupement et donc \mathcal{B}' ne peut pas être maximal dans Ω . En conclusion \mathcal{B} est maximal dans Ω .

5. PARTAGE DE $\#_0(\mathfrak{M})$

Notons :

$\#_0(l)$ le nombre de zéros figurant sur la ligne l ;

$\#_0(c)$ le nombre de zéros figurant sur la colonne c ; et

$\#_0(\mathfrak{M})$ le nombre d'entrées nulles de la matrice \mathfrak{M} .

DÉFINITION. Soient l et l' deux lignes quelconques de \mathfrak{M} alors nous poserons $l <_{\mathcal{L}} l'$ si et seulement si $\#_0(l) > \#_0(l')$. En particulier, nous désignerons par $L_{\inf}(\mathfrak{M})$ l'ensemble des lignes de \mathfrak{M} telles que $l <_{\mathcal{L}} l'$ pour toute ligne $l' \notin L_{\inf}(\mathfrak{M})$.

Notons l_x (resp. c_x) la ligne (resp. colonne) de \mathfrak{M} associée à l'objet x , alors nous aurons :

$$1 - \#_0(l_x) = \#_0(c_x)$$

2 - Le nombre de zéros associé à l'objet x est

$$\#_0(\{l_x, c_x\}) = \#_0(l_x) + \#_0(c_x) = 2 \#_0(l_x) = 2\#_0(c_x)$$

$$3 - \#_0(\mathfrak{M}) = \sum_{x \in \Omega} \#_0(l_x) = \sum_{x \in \Omega} \#_0(c_x)$$

DÉFINITION. Soit $(n, n_0) \in \mathbb{N}_0^* \times \mathbb{N}_0^*$ vérifiant $n_0 \leq n$, où \mathbb{N}_0^* est l'ensemble des entiers naturels pairs strictement positifs, alors on dit que (n_1, \dots, n_k) est un k -partage ordonné de n en entiers naturels pairs strictement positifs, bornés par n_0 si et seulement si on a :

$$n_0 \geq n_1 \geq \dots \geq n_k \text{ et } \sum_{j=1}^{j=k} n_j = n.$$

Si la condition $n_1 \geq \dots \geq n_k$ n'est pas réalisée on dit qu'on a affaire à un k -partage de n en entiers naturels pairs strictement positifs, bornés par n_0

Notons $\text{Part}_{n_0}^2(n)$ l'ensemble des partages ordonnés de n en entiers naturels pairs strictement positifs, bornés par n_0

PROPOSITION 2. Pour tout triplet $(n, m_0, n_0) \in \mathbb{N}_0^* \times \mathbb{N}_0^* \times \mathbb{N}_0^*$ vérifiant $m_0 \leq n_0 \leq n$ on a :

$$\text{Part}_{m_0}^2(n) \subseteq \text{Part}_{n_0}^2(n)$$

Preuve. Tout partage ordonné de n en entiers naturels pairs strictement positifs, bornés par m_0 est aussi un partage ordonné de n en entiers naturels pairs strictement positifs, bornés par n_0 dès que la condition $m_0 \leq n_0$ est vérifiée.

DÉFINITION. Nous appelons fragment de $\#_0(\mathfrak{M})$ associé à une k -séquence d'éliminations successives x_{i_1}, \dots, x_{i_k} la somme des entiers naturels pairs, notée

$$\text{Frag}_{(x_{i_1}, \dots, x_{i_k})}(\#_0(\mathfrak{M}))$$

et définie par :

$$\text{Frag}_{(x_{i_1}, \dots, x_{i_k})}(\#_0(\mathfrak{M})) = \sum_{j=1}^{j=k} a_{j-1}(n_j)$$

où pour $p \leq j-1$

$$a_p(n_j) = \#_0(\{l_{x_{i_j}}, c_{x_{i_j}}\} / \text{ELIM}(\mathfrak{M}) = \{x_{i_1}, \dots, x_{i_p}\})$$

et pour tout j :

$$n_j = \#_0(\{l_{x_{i_j}}, c_{x_{i_j}}\}).$$

$\text{ELIM}(\mathfrak{M})$ est l'ensemble des éléments dont on a supprimé la ligne et la colonne.

DÉFINITION. Deux k -séquences x_{i_1}, \dots, x_{i_k} et y_{i_1}, \dots, y_{i_k} seront dites équivalentes si et seulement si $\text{Frag}_{(x_{i_1}, \dots, x_{i_k})}(\#_0(\mathfrak{M})) = \text{Frag}_{(y_{i_1}, \dots, y_{i_k})}(\#_0(\mathfrak{M}))$.

PROPOSITION 3. Soit x_{i_1}, \dots, x_{i_k} une k -séquence d'éliminations successives, alors

$$a_p(n_j) = (1 - \chi_1(j)) \cdot \left[n_j + 2 \sum_{m=1}^{m=p} \mathbf{R}(x_{i_m}, x_{i_j}) - 2p \right]$$

où $I = \{1, 2, \dots, k\}$ et χ_I est la fonction caractéristique de I .

Preuve. Si $j \in I$ alors $1 - \chi_I(j) = 0$ et $a_p(n_j) = 0$, donc l'égalité

$$a_p(n_j) = (1 - \chi_I(j)) \cdot \left[n_j + 2 \sum_{m=1}^{m=p} \mathfrak{R}(x_{i_m}, x_{i_j}) - 2p \right] \text{ a lieu.}$$

Si $j \notin I$ alors $1 - \chi_I(j) = 1$ et il est facile de remarquer que

$$a_1(n_j) = \begin{cases} n_j & \text{si } \mathfrak{R}(x_{i_1}, x_{i_j}) = 1 \\ n_j - 2 & \text{si } \mathfrak{R}(x_{i_1}, x_{i_j}) = 0 \end{cases}$$

d'où $a_1(n_j)$ peut s'écrire $a_1(n_j) = n_j + 2 [\mathfrak{R}(x_{i_1}, x_{i_j}) - 1]$ de même

$a_2(n_j)$ peut s'écrire $a_2(n_j) = a_1(n_j) + 2 [\mathfrak{R}(x_{i_2}, x_{i_j}) - 1]$ donc

$$a_2(n_j) = n_j + 2 [\mathfrak{R}(x_{i_1}, x_{i_j}) - 1] + 2 [\mathfrak{R}(x_{i_2}, x_{i_j}) - 1]$$

$a_2(n_j) = n_j + 2 [\mathfrak{R}(x_{i_1}, x_{i_j}) + \mathfrak{R}(x_{i_2}, x_{i_j}) - 2]$ ainsi de suite, jusqu'à l'obtention de la formule

$$a_p(n_j) = (1 - \chi_I(j)) \cdot \left[n_j + 2 \sum_{m=1}^{m=p} \mathfrak{R}(x_{i_m}, x_{i_j}) - 2p \right].$$

Conséquences.

1 - Le fragment de $\#_0(\mathfrak{M})$ associé à la k -séquence x_{i_1}, \dots, x_{i_k} sera

$$\text{Frag}_{(x_{i_1}, \dots, x_{i_k})}(\#_0(\mathfrak{M})) = \sum_{j=1}^{j=k} n_j + \sum_{t=1}^{t=k-1} \sum_{j=t+1}^{j=k} \mathfrak{R}(x_{i_t}, x_{i_j}) - (k-1)k$$

2 - Les k -séquences x_{i_1}, \dots, x_{i_k} et $x_{i_{\sigma(1)}}, \dots, x_{i_{\sigma(k)}}$, où σ est une permutation de $\{1, 2, \dots, k\}$ sont équivalentes.

3 - Si les éléments de la k -séquence x_{i_1}, \dots, x_{i_k} forment un \mathfrak{R} -regroupement, alors le fragment de $\#_0(\mathfrak{M})$ associé à cette k -séquence sera :

$$\text{Frag}_{(x_{i_1}, \dots, x_{i_k})}(\#_0(\mathfrak{M})) = \sum_{j=1}^{j=k} n_j .$$

DÉFINITION. Une k -séquence x_{i_1}, \dots, x_{i_k} sera dite suffisante pour partager $\#_0(\mathfrak{M})$ si et seulement si $\text{Frag}_{(x_{i_1}, \dots, x_{i_k})}(\#_0(\mathfrak{M})) = \#_0(\mathfrak{M})$.

$(n_1, a_1(n_2), \dots, a_{k-1}(n_k))$ sera, alors appelé k -partage de $\#_0(\mathfrak{M})$ réalisé par la séquence x_{i_1}, \dots, x_{i_k} .

N.B. Toutes les matrices \mathfrak{M} qui seront considérées dans les propositions suivantes sont non complètes et sans éléments isolés.

PROPOSITION 4. Notons $\text{Suff}(\#_0(\mathfrak{M}))$ l'ensemble des séquences suffisantes pour partager $\#_0(\mathfrak{M})$, alors $\text{Suff}(\#_0(\mathfrak{M})) \neq \emptyset$.

Preuve. Nous allons faire un raisonnement par récurrence. Soit r l'ordre de la matrice \mathfrak{M} . Pour $r = 3$, les seules matrices de ressemblance non complètes et sans éléments isolés sont les trois \mathcal{L} -matrices suivantes :

$$\mathfrak{M}_1^3 = \begin{matrix} & \begin{matrix} a & b & c \end{matrix} \\ \begin{matrix} a \\ b \\ c \end{matrix} & \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \end{matrix}$$

$$\mathfrak{M}_2^3 = \begin{matrix} & \begin{matrix} a & b & c \end{matrix} \\ \begin{matrix} a \\ b \\ c \end{matrix} & \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 1 \\ 0 & 1 & 1 \end{bmatrix} \end{matrix}$$

$$\mathfrak{M}_3^3 = \begin{matrix} & \begin{matrix} a & b & c \end{matrix} \\ \begin{matrix} a \\ b \\ c \end{matrix} & \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix} \end{matrix}$$

Prenons à titre d'exemple la \mathcal{L} -matrice \mathfrak{M}_1^3 , nous aurons $\#_0(\mathfrak{M}_1^3) = 2$;
 $\#_0(\{l_a, c_a\}) = \#_0(\{l_b, c_b\}) = 2$; et il est clair que
 $\#_0(\mathfrak{M}_1^3) = \#_0(\{l_a, c_a\}) = \#_0(\{l_b, c_b\}) = 2$.

donc $\text{Suff}(\#_0(\mathfrak{M})) \neq \emptyset$. Il en est de même pour \mathfrak{M}_2^3 et \mathfrak{M}_3^3 . Supposons maintenant que pour toute matrice de ressemblance d'ordre $\leq r - 1$, non complète et sans éléments isolés, nous avons $\text{Suff}(\#_0(\mathfrak{M})) \neq \emptyset$.

Soit $X \in \Omega$ tel que $n_x = \#_0(\{l_x, c_x\}) > 0$ et soit $\mathfrak{M}_x = \mathfrak{M} \setminus \{l_x, c_x\}$ alors $\#_0(\mathfrak{M}_x) = \#_0(\mathfrak{M}) - n_x$ et $\text{Suff}(\#_0(\mathfrak{M}_x)) \neq \emptyset$ et d'après l'hypothèse de récurrence il existe une séquence $y_{i_1}, \dots, y_{i_q} \in \text{Suff}(\#_0(\mathfrak{M}_x))$ et donc la séquence $x, y_{i_1}, \dots, y_{i_q} \in \text{Suff}(\#_0(\mathfrak{M}))$ d'où $\text{Suff}(\#_0(\mathfrak{M})) \neq \emptyset$.

PROPOSITION 5. Soit $x_{i_1}, \dots, x_{i_q} \in \text{Suff}(\#_0(\mathfrak{M}))$, alors il existe une q -séquence y_{i_1}, \dots, y_{i_q} telle que pour tout $t \in \{1, 2, \dots, q\}$, nous aurons $y_{i_t} = x_{i_{\sigma(t)}}$, où σ est une permutation de $\{1, 2, \dots, q\}$ et telle que y_{i_1}, \dots, y_{i_q} fournit un partage ordonné de $\#_0(\mathfrak{M})$.

Preuve. Nous avons $n_j = \#_0(\{l_{x_{i_j}}, c_{x_{i_j}}\})$. Désignons, alors, par y_{i_1} un élément $x_{i_{j_0}}$ de l'ensemble $\{x_{i_1}, \dots, x_{i_q}\}$ qui satisfait $n_{j_0} = \max_j n_j$. Posons $\text{ELIM}(\mathfrak{M}) = \{y_{i_1}\}$ et cherchons $m_j = \#_0(\{l_{x_{i_j}}, c_{x_{i_j}}\} / \text{ELIM}(\mathfrak{M}) = \{y_{i_1}\})$. Désignons par y_{i_2} un élément $x_{i_{j_1}}$ de l'ensemble $\{x_{i_1}, \dots, x_{i_q}\} \setminus \{x_{i_{j_0}}\}$ qui satisfait $m_{j_1} = \max_j m_j$, etc. . Nous pouvons ainsi déterminer une q -séquence y_{i_1}, \dots, y_{i_q} équivalente à x_{i_1}, \dots, x_{i_q} et réalisant un partage ordonné de $\#_0(\mathfrak{M})$.

PROPOSITION 6. Notons $\text{Ord}(\#_0(\mathfrak{M}))$ l'ensemble des séquences suffisantes réalisant un partage ordonné de $\#_0(\mathfrak{M})$, alors $\text{Ord}(\#_0(\mathfrak{M})) \neq \emptyset$.

Preuve. Cette proposition est une conséquence directe des propositions 4 et 5.

DÉFINITION. Une k -séquence suffisante x_{i_1}, \dots, x_{i_k} sera dite minimale dans $\text{Suff}(\#_0(\mathfrak{M}))$ si et seulement si toute q -séquence y_{i_1}, \dots, y_{i_q} (où $q < k$) n'est pas suffisante.

DÉFINITION. Un partage quelconque de $\#_0(\mathfrak{M})$ sera dit réalisable si et seulement s'il correspond à un partage de $\#_0(\mathfrak{M})$ fourni par une séquence suffisante.

DÉFINITION. Un p -partage de $\#_0(\mathfrak{M})$ sera dit réalisable minimal si et seulement si tout q -partage de $\#_0(\mathfrak{M})$ (où $q < p$) n'est pas réalisable.

Exemple illustrant les définitions précédentes : Soit $\Omega = \{a, b, c, d, e\}$ et soit $\mathfrak{M} = \mathbf{R}(\Omega)$ la matrice suivante :

$$\begin{array}{c} \begin{array}{ccccc} & a & b & c & d & e \\ a & \left[\begin{array}{ccccc} 1 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 \end{array} \right] \\ b \\ c \\ d \\ e \end{array} \end{array}$$

alors nous avons

- 1 - $\#_0(\mathfrak{M}) = 12$;
- 2 - $L_{\text{inf}}(\mathfrak{M}) = \{c, e\}$;
- 3 - $K^*(\mathfrak{M}) = \{c\}$ ou $K^*(\mathfrak{M}) = \{e\}$;
- 4 - Les séquences a,c,e et b,c,e sont suffisantes pour partager $\#_0(\mathfrak{M})$.
- 5 - Les séquences b,c,d et c,d,e ne sont pas suffisantes.
- 6 - Le 2-partage (6,6) n'est pas réalisable.
- 7 - Le 3-partage (4,4,4) est réalisable par la séquence b,c,e ; et le 4-partage (4,4,2,2) est réalisable par a,d,b,e.
- 8 - $\text{Frag}_{(a,d)}(\#_0(\mathfrak{M})) = 10$ et $\text{Frag}_{(b,c,d)}(\#_0(\mathfrak{M})) = 10$.

Soit $K_0^j(\mathfrak{M})$ pour $j = 1, \dots, s$ les P.G.R de $L_{\text{inf}}(\mathfrak{M})$ et $\mathfrak{M}_j = \mathfrak{M} \setminus \bigcup_{x \in K_0^j} \{l_x, c_x\}$.

Notons $K^*(\mathfrak{M})$ le $K_0^{j_0}(\mathfrak{M})$ tel qu'on ait $\#_0(l_x) > \#_0(l_y)$ pour $x \in L_{\text{inf}}(\mathfrak{M}_{j_0})$ et $y \in L_{\text{inf}}(\mathfrak{M}_j)$, où $j \in \{1, \dots, s\} \setminus \{j_0\}$.

PROPOSITION 7. Soit K un \mathbf{R} -regroupement de $K^*(\mathfrak{M})$.

Posons $\mathfrak{M}_{r_1} = \mathfrak{M} \setminus \bigcup_{y \in K} \{l_y, c_y\}$, alors nous avons $K^*(\mathfrak{M}_{r_1}) = K^*(\mathfrak{M}) \setminus K$.

Preuve. Supposons que $K^*(\mathfrak{M}) = \{x_{i_1}, \dots, x_{i_{r_0}}\}$ et $\text{ELIM}(\mathfrak{M}) = \{x_{i_1}, \dots, x_{i_{r_0}}\}$ en prenant $I = \{1, \dots, r_1\}$ et en posant $p_0 = \#_0(\{l_y, c_y\})$, où $y \in L_{\text{inf}}(\mathfrak{M})$ nous aurons d'après la proposition 3 :

$$a_{r_1}(n_j) = \begin{cases} 0 & \text{si } j \leq r_1 \\ p_0 & \text{si } r_1 + 1 \leq j \leq r_0 \\ n_j + 2 \left[\sum_{m=1}^{m=r_1} \mathbf{R}(x_{i_m}, x_{i_j}) - r_1 \right] & \text{si } r_0 < j \end{cases}$$

Soit N le nombre d'éléments de $L_{\text{inf}}(\mathfrak{M})$ qui pourraient former un \mathfrak{R} -regroupement avec les éléments $\{x_{i_1}, \dots, x_{i_{r_1}}\}$, alors nous avons $N \leq \inf(\#(L_{\text{inf}}(\mathfrak{M})) - r_0, r_0 - r_1)$. Or, parmi ces N éléments aucun ne peut être \mathfrak{R} -ressemblant aux éléments $\{x_{i_{r_1}+1}, \dots, x_{i_{r_0}}\}$, autrement nous aurons $r_1 + (r_0 - r_1) + 1 = r_0 + 1$ éléments formant un \mathfrak{R} -regroupement de $L_{\text{inf}}(\mathfrak{M})$ ce qui est impossible car $\#(K^*(\mathfrak{M})) = r_0$, donc $\#(K_0^j(\mathfrak{M}, r_1)) = r_0 - r_1$ pour tout **P.G.R.**, $K_0^j(\mathfrak{M}, r_1)$, de $L_{\text{inf}}(\mathfrak{M}, r_1)$.

Si $\{x_{i_{r_1}+1}, \dots, x_{i_{r_0}}\}$ est le seul **P.G.R.** de $L_{\text{inf}}(\mathfrak{M}, r_1)$, alors nous aurons $K^*(\mathfrak{M}, r_1) = \{x_{i_{r_1}+1}, \dots, x_{i_{r_0}}\}$. Si $\{x_{i_{r_1}+1}, \dots, x_{i_{r_0}}\}$ n'est pas unique, le simple fait que $\{x_{i_1}, \dots, x_{i_{r_1}}\} \cup K_0^j(\mathfrak{M}, r_1) \neq K^*(\mathfrak{M})$ impliquera immédiatement que $K_0^j(\mathfrak{M}, r_1) \neq K^*(\mathfrak{M}, r_1)$.

DÉFINITION. On dit qu'une k -séquence x_{i_1}, \dots, x_{i_k} est une L_{inf} -séquence si et seulement si
 1 - les q_0 premiers éléments appartiennent à $L_{\text{inf}}(\mathfrak{M}_0)$, où $\mathfrak{M}_0 := \mathfrak{M}$, les q_1 éléments suivants appartiennent à $L_{\text{inf}}(\mathfrak{M}_1), \dots$, les q_r derniers éléments appartiennent à $L_{\text{inf}}(\mathfrak{M}_r)$, et
 2 - les conditions suivantes sont satisfaites

$$\sum_{i=0}^{i=r} q_i = k,$$

$$\forall p \in (0, 1, 2, \dots, r) : \mathfrak{M}_p := \mathfrak{M} \setminus \bigcup_{j \in I(p)} \{l_{x_j}, c_{x_j}\}$$

où $I(p) = \left\{ 1, \dots, \sum_{i=0}^{i=p-1} q_i \right\}$ avec la valeur initiale $I(0) = \emptyset$.

DÉFINITION. On dit qu'une L_{inf} -séquence x_{i_1}, \dots, x_{i_k} est une L_{inf} -séquence optimisée si et seulement si nous avons $\forall p \in \{0, 1, 2, \dots, r\} : q_p = \#(K^*(\mathfrak{M}_p))$ et $\forall j \in I(p+1) \setminus I(p) : x_j \in K^*(\mathfrak{M}_p)$. Désignons, alors par $\text{ELIM}^*(\mathfrak{M})$ une telle séquence.

PROPOSITION 8. Soit K un \mathfrak{R} -regroupement de $K^*(\mathfrak{M})$ et soit $\mathfrak{M}' = \mathfrak{M} \setminus \bigcup_{y \in K} \{l_y, c_y\}$, alors nous avons $\text{ELIM}^*(\mathfrak{M}') = \text{ELIM}^*(\mathfrak{M}) \setminus K$

Preuve. Supposons que $\text{ELIM}^*(\mathfrak{M}) = \{x_1^*, \dots, x_s^*\}$ et $K = \{x_1^*, \dots, x_r^*\}$, où $r \leq q_0$, alors par définition de $\text{ELIM}^*(\mathfrak{M})$ nous aurons : les q_0 premiers éléments appartiennent à $K^*(\mathfrak{M})$ avec $\#(K^*(\mathfrak{M})) = q_0$, les q_1 éléments suivants appartiennent à $K^*(\mathfrak{M}_1)$, avec $\#(K^*(\mathfrak{M}_1)) = q_1, \dots$, et les q_r derniers éléments appartiennent à $K^*(\mathfrak{M}_r)$, avec $\#(K^*(\mathfrak{M}_r)) = q_r$.

Or, d'après la proposition 7 nous avons $K^*(\mathfrak{M}') = K^*(\mathfrak{M}) \setminus K = \{x_{r+1}^*, \dots, x_{q_0}^*\}$ et comme nous pouvons récrire \mathfrak{M}_1 sous la forme

$$\mathfrak{M}_1 = \mathfrak{M}' \setminus \bigcup_{y \in K^*(\mathfrak{M}')} \{l_y, c_y\},$$

il s'en suit que $\{x_{r+1}^*, \dots, x_s^*\} = \text{ELIM}^*(\mathfrak{M}')$.

THEOREME 1. Soit \mathfrak{M} une matrice de ressemblance non complète et sans éléments isolés, alors $\text{ELIM}^*(\mathfrak{M})$ fournit un partage ordonné réalisable minimal de $\#_0(\mathfrak{M})$.

Preuve. Soit \mathfrak{M} une matrice de ressemblance d'ordre r et montrons $\text{ELIM}^*(\mathfrak{M})$ fournit un partage ordonné réalisable minimal de $\#_0(\mathfrak{M})$. Pour cela nous allons faire un raisonnement par récurrence. Pour $r = 3$, les seules matrices de ressemblance non complètes et sans éléments isolés sont les trois L-matrices suivantes :

$$\mathfrak{M}_1^3 = \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 1 & 0 & 1 \\ b & 0 & 1 & 1 \\ c & 1 & 1 & 1 \end{array} \end{array} \quad \mathfrak{M}_2^3 = \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 1 & 1 & 0 \\ b & 1 & 1 & 1 \\ c & 0 & 1 & 1 \end{array} \end{array} \quad \mathfrak{M}_3^3 = \begin{array}{c} \begin{array}{ccc} & a & b & c \\ a & 1 & 1 & 1 \\ b & 1 & 1 & 0 \\ c & 1 & 0 & 1 \end{array} \end{array}$$

Prenons à titre d'exemple la L-matrice \mathfrak{M}_1^3 , nous aurons $\#_0(\mathfrak{M}_1^3) = 2$;

$L_{\text{inf}}(\mathfrak{M}_1^3) = \{a, b\}$, $K^*(\mathfrak{M}_1^3) := \{a\}$ ou $K^*(\mathfrak{M}_1^3) := \{b\}$; $\text{ELIM}^*(\mathfrak{M}_1^3) := \{a\}$ ou

$\text{ELIM}^*(\mathfrak{M}_1^3) := \{b\}$;

$\#_0(\{l_a, c_a\}) = \#_0(\{l_b, c_b\}) = 2$; et il est clair que

$\#_0(\mathfrak{M}_1^3) = \#_0(\{l_a, c_a\}) = \#_0(\{l_b, c_b\}) = 2$.

Chacun de ces deux 1-partages de $\#_0(\mathfrak{M}_1^3)$ est un partage réalisable minimal selon une L_{inf} -séquence optimisée a ou b . Il en est de même pour \mathfrak{M}_2^3 et \mathfrak{M}_3^3 . Supposons maintenant que pour toute matrice de ressemblance d'ordre $\leq r - 1$, $\text{ELIM}^*(\mathfrak{M})$ fournit un partage ordonné réalisable minimal de $\#_0(\mathfrak{M})$ et montrons que le résultat reste vrai pour une matrice \mathfrak{M} d'ordre r . Supposons que $\text{ELIM}^*(\mathfrak{M}) = \{x_1^*, \dots, x_s^*\}$ et posons $\mathfrak{M}' = \mathfrak{M} \setminus \{l_{x_1}^*, c_{x_1}^*\}$ nous aurons $\#_0(\mathfrak{M}') = \#_0(\mathfrak{M}) - n_1^*$ et d'après la proposition 8, $\text{ELIM}^*(\mathfrak{M}') = \{x_2^*, x_s^*\}$ donc d'après l'hypothèse de récurrence $\{x_2^*, x_s^*\}$ fournit un partage ordonné réalisable minimal de $\#_0(\mathfrak{M}')$. Soit (n_2^*, \dots, n_s^*) ce partage, d'où (n_2^*, \dots, n_s^*) est un partage ordonné réalisable de $\#_0(\mathfrak{M})$ et il reste à montrer qu'il est réalisable minimal. Pour ce faire, soit x_{i_1}, \dots, x_{i_k} un élément de $\text{Ord}(\#_0(\mathfrak{M}))$. Si $\#(K^*(\mathfrak{M})) = r_0$ et si r_1 est le nombre d'éléments de la séquence x_{i_1}, \dots, x_{i_k} formant un \mathfrak{R} -regroupement de $L_{\text{inf}}(\mathfrak{M})$, alors nous aurons $r_1 \leq \#(K^*(\mathfrak{M})) = r_0$. Soit \mathfrak{M}_{x_1} la sous-matrice de \mathfrak{M} obtenue par élimination des r_1 éléments $x_1^*, \dots, x_{r_1}^*$ et soit $\mathfrak{M}_{x_{i_1}}$ la sous-matrice de \mathfrak{M} obtenue par élimination des r_1 premiers éléments de la séquence x_{i_1}, \dots, x_{i_k} , alors si nous posons

$p_0 = \#_0(\{l_y, c_y\})$, où $y \in L_{\text{inf}}(\mathfrak{M})$, $p_{x_1} = \#_0(\{l_y, c_y\})$, où $y \in L_{\text{inf}}(\mathfrak{M}_{x_1})$ et

$p_{x_{i_1}} = \#_0(\{l_y, c_y\})$, où $y \in L_{\text{inf}}(\mathfrak{M}_{x_{i_1}})$, nous aurons $p_{x_{i_1}} \leq p_{x_1}$. D'après l'hypothèse de

récurrence, $\text{ELIM}^*(\mathfrak{M}') = \{x_{1+r_1}^*, \dots, x_s^*\}$ fournit un partage ordonné réalisable minimal de

$\#_0(\mathfrak{M}) - r_1 p_0$ et d'après la proposition 2, nous avons

$$\text{Part}_{p_{x_{i_1}}}^2(\#_0(\mathfrak{M}) - r_1 p_0) \subseteq \text{Part}_{p_{x_1}}^2(\#_0(\mathfrak{M}) - r_1 p_0)$$

donc $s - r_1 \leq k - r_1$, et par suite $s \leq k$ ce qui signifie que le partage $(n_1^* - n_s^*)$ est réalisable minimal et le théorème 1 est prouvé.

Remarque. La recherche d'une séquence minimale d'éliminations successives basée sur la notion de $K^*(\mathfrak{M}_p)$ présente un grand intérêt lorsque les tailles des $L_{\text{inf}}(\mathfrak{M}_p)$ rencontrés sont réduites, autrement nous aurons à utiliser la même démarche pour détecter $K^*(\mathfrak{M}_p)$ à chaque itération.

Obtention d'un P.G.R de Ω

Pour obtenir un P.G.R de Ω , il suffit de déterminer $\text{ELIM}^*(\mathfrak{M})$, où $\mathfrak{M} = \mathfrak{R}(\Omega)$ et de poser $\text{P.G.R} = \Omega \setminus \text{ELIM}^*(\mathfrak{M})$.

Exemple. Soit $\Omega = \{a, b, c, d, e, f, g\}$ et soit $\mathfrak{M} = \mathfrak{R}(\Omega)$ la matrice suivante :

$$\begin{array}{c} \begin{array}{cccccc} a & b & c & d & e & f & g \\ \begin{array}{l} a \\ b \\ c \\ d \\ e \\ f \\ g \end{array} \end{array} \left[\begin{array}{cccccc} 1 & 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 & 1 & 1 \end{array} \right]$$

Posons $\#_0^k(j) = \#_0 \left(\left\{ l_{x_{ij}}, c_{x_{ij}} \right\} / \#(\text{ELIM}^*(\mathfrak{M})) = k \right)$

Recherche de $\text{ELIM}^*(\mathfrak{M})$

objet j	$\#_0^0(j)$	$\#_0^1(j)$	$\#_0^2(j)$	$\#_0^3(j)$
a	4	2	0	0
b*	6	0	0	0
c	4	2	2	0
d*	4	4	4	0
e	4	4	2	0
f*	4	4	0	0
g	2	0	0	0

* indique les éléments de $\text{ELIM}^*(\mathfrak{M})$

 indique les éléments du partage ordonné réalisable minimal

Dans cet exemple, nous avons $L_{\text{inf}}(\mathfrak{M}) = \{b\}$ et $K^*(\mathfrak{M}) = \{b\}$, donc on élimine b, nous aurons alors $\mathfrak{M}_1 = \mathfrak{M} \setminus \{l_b, c_b\}$, $L_{\text{inf}}(\mathfrak{M}_1) = \{d, e, f\}$ et $K^*(\mathfrak{M}_1) = \{d, f\}$.

Comme $\#_0(\mathfrak{M}_1) = \#_0^1(d) + \#_0^1(f) = 8$, nous aurons à éliminer d et f.

D'où la L_{inf} -séquence optimisée sera $\text{ELIM}^*(\mathfrak{M}) = \{b, d, f\}$ et le P.G.R = $\{a, c, e, g\}$.

6. UNE COALESCENCE FLOUE EN DEUX PHASES

L'ensemble $\Omega' = \Omega \setminus \text{ISOL}(\Omega)$ constituera le support d'allocation, *i.e.*, l'ensemble des objets qui vont former les \mathcal{R} -regroupements non triviaux. Il s'en suit qu'une condition nécessaire et suffisante pour qu'au moins un \mathcal{R} -regroupement non trivial existe, est que : $\text{ISOL}(\Omega) \neq \Omega$.

Nous allons construire tout d'abord une quasi-partition de Ω' [1], *i.e.* un recouvrement de Ω' par des \mathcal{R} -regroupements maximaux non nécessairement disjoints tels que la suppression d'un \mathcal{R} -regroupement quelconque fait qu'il existe au moins un objet qui ne figure pas dans la réunion des \mathcal{R} -regroupements qui restent .

Soit $\mathcal{J} = \{1, \dots, c\}$, un ensemble d'indices, pour tout $i \in \mathcal{J}$, nous posons $\mathcal{J}(i) = \mathcal{J} \setminus \{i\}$, alors dans notre contexte, la définition de quasi-partition peut s'énoncer de la façon suivante :

DÉFINITION. Une famille de \mathcal{R} -regroupements \mathcal{B}_i où $i \in \mathcal{J}$, forme une quasi-partition de Ω' si et seulement si nous avons :

$$\Omega = \bigcup_{j \in \mathcal{J}} \mathcal{B}_j \text{ et } (\forall i \in \mathcal{J}) : \bigcup_{j \in \mathcal{J}(i)} \mathcal{B}_j \subset \Omega'$$

Obtention d'une quasi-partition

La première phase qui va suivre, va nous permettre de construire cette quasi-partition de Ω par des \mathcal{R} -regroupements maximaux et nous fournit, par la même occasion une technique de coalescence où l'appartenance à plusieurs classes est possible .

La première phase

Pas 1 : Former $\mathcal{R}(\Omega)$.

Pas 2 : Identifier $\text{ISOL}(\Omega)$.

Pas 3 : Former $\Omega' = \Omega \setminus \text{ISOL}(\Omega)$.

Pas 4 : Extraire $\mathfrak{M} = \mathcal{R}(\Omega')$. Si $\#_0(\mathcal{R}(\Omega')) = 0$ alors STOP, sinon aller au pas 5 fin si.

Pas 5 : $\Omega_0 := \Omega'$; $\mathcal{B}_0 := \emptyset$; $r := 0$;

tant que $\Omega_r \setminus \left(\bigcup_{0 \leq k \leq r} \mathcal{B}_k \right) \neq \emptyset$ faire poser $\Omega_{r+1} := \Omega_r \setminus \left(\bigcup_{0 \leq k \leq r} \mathcal{B}_k \right)$

1 - si $r+1$ est impair alors choisir dans Ω_{r+1} un objet $X_{i_{r+1}}$ ayant le plus petit nombre de \mathcal{R} -ressemblants dans Ω sinon choisir un objet $X_{i_{r+1}}$ ayant le plus grand nombre de \mathcal{R} -ressemblants fin si

2 - extraire un P.G.R, \mathcal{B}_{r+1} , de $\mathcal{L}(X_{i_{r+1}})$

$r := r + 1$;

fin tant que

Obtention des \mathcal{R} -regroupements flous induits.

A partir de la quasi-partition déjà obtenue à la première phase, il nous est possible de construire une partition floue induite en utilisant la deuxième phase qui va suivre. Nous obtenons ainsi une technique de coalescence floue. La définition de partition floue (Cf. [2,6]) peut s'énoncer comme suit :

DÉFINITION. Une famille de c sous-ensembles flous $A_i = \{\mu_i(X)/X, X \in \Omega\}$, $i = 1, 2, \dots, c$ forme une c -partition floue de Ω si et seulement si pour tout objet X de Ω nous avons :

$$\sum_{i=1}^{i=c} \mu_i(X) = 1$$

Deuxième phase

Soit $\{\mathbf{B}_1, \mathbf{B}_2, \dots, \mathbf{B}_c\}$ une quasi-partition de Ω obtenue à la première phase, nous abordons la deuxième phase suivante, *i.e.*, obtenir une partition floue $\{\mathbf{F}_1, \mathbf{F}_2, \dots, \mathbf{F}_c\}$ de Ω en sorte que la fonction d'appartenance de chaque \mathbf{F}_j pour $j = 1, 2, \dots, c$, soit déterminée à partir des \mathbf{B}_j de la façon suivante : ($\forall X \in \Omega$) et ($\forall j \in \{1, 2, \dots, c\}$), nous poserons :

$$\mu_{\mathbf{F}_j}(X) = \begin{cases} 0 & \text{si } X \notin \mathbf{B}_j \\ \frac{\sum_{Y \in \mathbf{B}_j} \{S(X, Y) * \tilde{S}(X, Y)\}}{\sum_{\mathbf{B}_r \in \mathbf{B}(X)} \sum_{Z \in \mathbf{B}_r} \{S(X, Y) * \tilde{S}(X, Y)\}} & \text{si } X \in \mathbf{B}_j \end{cases}$$

où $\mathbf{B}(X)$ est l'ensemble des \mathbf{R} -regroupements \mathbf{B}_r contenant X et $*$ est l'opérateur «min» si la relation de S -comparaison choisie est «plus fine» ou l'opérateur «max» si la relation de S -comparaison choisie est «moins fine». L'avantage de cette démarche est qu'elle nous assure que pour tout $j = 1, 2, \dots, c$ et pour tout ξ de $[0, 1]$ la ξ -coupe de \mathbf{F}_j reste un \mathbf{R} -regroupement.

PROPOSITION 9

Soit $\alpha = (\alpha_1, \alpha_2)$ un seuillage, X un objet et p le nombre de \mathbf{R} -regroupements contenant X . Notons n_i le cardinal du $i^{\text{ème}}$ \mathbf{R} -regroupement ($i = 1, \dots, p$). Nous aurons alors dans le cas d'une S -comparaison «plus fine» et pour $k = 1, \dots, p$ les inégalités suivantes :

$$\frac{n_k \cdot \inf(\alpha_1, \alpha_2)}{\sum_{1 \leq i \leq p} n_i} \leq \mu_{\mathbf{F}_k}(X) \leq \inf \left[1, \frac{n_k}{\inf(\alpha_1, \alpha_2) \sum_{1 \leq i \leq p} n_i} \right]$$

Dans le cas d'une S -comparaison «moins fine» il faut remplacer $\inf(\alpha_1, \alpha_2)$ par $\sup(\alpha_1, \alpha_2)$.

Preuve. Supposons que la S -comparaison considérée soit «moins fine», nous aurons alors pour tout $k \in \{1 \dots p\}$ et pour tout $X \in \mathbf{B}_k$

$$\inf(\alpha_1, \alpha_2) n_k \leq \sum_{Y \in \mathbf{B}_k} \{S(X, Y) * \tilde{S}(X, Y)\} \leq n_k \quad (1)$$

et

$$\inf (\alpha_1, \alpha_2) \cdot \sum_{1 \leq i \leq p} n_i \leq \sum_{\mathbf{B}_r \in \mathbf{B}(X)} \sum_{Z \in \mathbf{B}_r} \{S(X, Z) * \tilde{S}(X, Z)\} \leq \sum_{1 \leq i \leq p} n_i$$

donc

$$\frac{1}{\inf (\alpha_1, \alpha_2) \cdot \sum_{1 \leq i \leq p} n_i} \geq \frac{1}{\sum_{\mathbf{B}_r \in \mathbf{B}(X)} \sum_{Z \in \mathbf{B}_r} \{S(X, Z) * \tilde{S}(X, Z)\}} \geq \frac{1}{\sum_{1 \leq i \leq p} n_i} \quad (2)$$

d'où le résultat en utilisant simultanément les encadrements (1) et (2) et le fait que $\mu_{\mathcal{F}_k}(X) \leq 1$. La proposition précédente nous montre l'impact du choix d'un seuillage sur les degrés d'appartenance

7. EXEMPLE ILLUSTRATIF

Supposons pour illustrer la méthode de coalescence à deux phases décrite ci-dessus, que nous désirons établir une typologie des minorités américaines décrites par les taux de chômages à différentes dates . Onze groupes ethniques vont constituer les objets et les taux de chômage des années 1973, 1975 et 1980 vont constituer les descripteurs classifiants. Les données relatives à cet exemple figurent à la Table 1, où mb veut dire masculin blanc, mn masculin noir, mh masculin d'origine hispanique, fb féminin blanc, fn féminin noir, fh féminin d'origine hispanique. Les classes d'âge vont de 16 à 19 ans et de 20 à 24 ans. Dans cet exemple, l'ensemble des objets sera l'ensemble $\Omega = \{mb16-19, mn16-19, fn16-19, mh16-19, fh16-19, mb20-24, fb20-24, mn20-24, fn20-24, mh20-24, fh20-24\}$.

	chôm 73	chôm 75	chôm 80
mb 6 - 19	12.30	18.30	16.20
mn 16 - 19	27.70	38.10	37.40
fn 16 - 19	35.90	41.00	39.90
mh 16 - 19	19.00	27.60	21.70
fh 16 - 19	20.70	27.90	23.70
mb 20 - 24	6.50*	3.20	11.10
fb 20 - 24	7.00	11.20*	8.50*
mn 20 - 24	12.80	24.70	23.80
fn 20 - 24	18.30	24.30	23.40
mh 20 - 24	8.20	16.30	12.20
fh 20 - 24	9.00	17.20	11.20

Table 1 - Taux de chômage des minorités américaine (en %)

Source : *Le Travail dans le Monde 1*, Bureau International du Travail, Genève, 1984 , page 48.

Dans ce cas, les scores les plus bas sont les meilleurs car les descripteurs classifiants sont des taux de chômage, donc les fonctions standardisantes seront de la forme : $f(x) = \frac{x^*}{x}$, où x^* désigne le minimum observé pour chaque descripteur; et si nous choisissons l'indice de similarité entre objets défini par :

$$S(X, Y) = \frac{\sum_{d \in \mathbf{D}} \min(\mu_X(d), \mu_Y(d))}{\sum_{d \in \mathbf{D}} \max(\mu_X(d), \mu_Y(d))}$$

et si nous optons pour une S -comparaison « plus fine », nous obtenons, pour un seuillage $\alpha = (0.70, 0.70)$ la matrice de S -comparaison $\mathbf{R}(\Omega)$ représentée à la Table 2 (Cf. annexe). L'ensemble $\text{ISOL}(\Omega) = \{\text{mb20-24}, \text{fb20-24}\}$ est obtenu. Si nous choisissons fh20-24 parmi les prototypes ayant le moins de \mathbf{R} -semblables dans Ω' , nous obtenons la sous-matrice complète $\mathbf{R}(\mathbf{L}(\text{fh20-24}))$, d'où $\mathbf{B}_1 = \{\text{mh20-24}, \text{fh20-24}\}$ constitue un premier \mathbf{R} -regroupement. Maintenant, mn20-24 est le prototype ayant le plus de \mathbf{R} -semblables dans $\Omega_1 = \Omega' \setminus \mathbf{B}_1$ nous obtenons la sous-matrice $\mathbf{R}(\mathbf{L}(\text{mn20-24}))$ suivante :

	mb16-19	mh16-19	fh16-19	mn20-24	fn20-24
mb16-1	1	0	0	1	1
mh16-1	0	1	1	1	1
fh16-1	0	1	1	1	1
mn20-2	1	1	1	1	1
fn20-2	1	1	1	1	1

où l'élimination de mb16-19 donne lieu à une sous-matrice complète et donc à un deuxième \mathbf{R} -regroupement $\mathbf{B}_2 = \{\text{mh16-19}, \text{fh16-19}, \text{mn20-24}, \text{fn20-24}\}$. Mais Ω' n'est pas encore couvert par \mathbf{B}_1 et \mathbf{B}_2 , Si nous choisissons fn16-19 comme prototype appartenant à $\Omega_2 = \Omega' \setminus (\mathbf{B}_1 \cup \mathbf{B}_2)$ et ayant le moins de \mathbf{R} -semblables dans Ω' , nous obtenons la sous-matrice complète $\mathbf{R}(\mathbf{L}(\text{fn16-19}))$, et donc nous aurons un troisième \mathbf{R} -regroupement $\mathbf{B}_3 = \{\text{mn16-19}, \text{fn16-19}\}$. Ω' n'étant pas encore couvert par \mathbf{B}_1 , \mathbf{B}_2 et \mathbf{B}_3 , $\Omega_3 = \Omega' \setminus (\mathbf{B}_1 \cup \mathbf{B}_2 \cup \mathbf{B}_3)$ est réduit à l'unique prototype mb16-19 qui donne lieu à la sous-matrice complète $\mathbf{R}(\mathbf{L}(\text{mb16-19}))$ et donc au quatrième \mathbf{R} -regroupement $\mathbf{B}_4 = \{\text{mb16-19}, \text{mn20-24}, \text{fn20-24}\}$. Ceci achève le recouvrement de Ω' . La méthode de coalescence décrite ci-dessus, nous a donc donné à la première phase pour un seuillage $\alpha = (0.70, 0.70)$, $\text{ISOL}(\Omega) = \{\text{fb20-24}, \text{mb20-24}\}$ et les \mathbf{R} -regroupements empiétants suivants :

$$\mathbf{B}_1 = \{mh20-24, fh20-24\};$$

$$\mathbf{B}_2 = \{mh16-19, fh16-19, mn20-24, fn20 - 24\};$$

$$\mathbf{B}_3 = \{mn16-19, fn16-19\}; \text{ et}$$

$$\mathbf{B}_4 = \{mb16-19, mn20-24, fn20 - 24\};$$

et à la seconde phase : $ISOL(\Omega) = \{fb20-24, mb20-24\}$ et les \mathbf{R} -regroupements flous suivants :

$$\mathbf{F}_1 = 1/mh20-24 + 1/fh20-24;$$

$$\mathbf{F}_2 = 1/mh16-19 + 1/fh16-19 + 0.57/mn20-24 + 0.59/fn20-24$$

$$\mathbf{F}_3 = 1/mn16-19 + 1/fn16-19; \text{ et}$$

$$\mathbf{F}_4 = 1/mb16-19 + 0.43/mn20-24 + 0.41/fn20 - 24 .$$

Pour un seuillage $\alpha = (0.60,0.60)$, nous obtenons à la première phase : $ISOL(\Omega) = \{fb20-24, mb20-24\}$ et les \mathbf{R} -regroupements empiétants suivants :

$$\mathbf{B}_1 = \{mb16-19, mh20-24, fh20-24\};$$

$$\mathbf{B}_2 = \{mn16-19, fn16-19, fh16-19\}; \text{ et}$$

$$\mathbf{B}_3 = \{mb16-19, mh16-19, fh16-19, mn20-24, fn20 - 24\};$$

et à la seconde phase : $ISOL(\Omega) = \{fb20-24, mb20-24\}$ et les \mathbf{R} -regroupements flous suivants :

$$\mathbf{F}_1 = 0.39/mb16-19 + 1/mh20-24 + 1/fh20-24;$$

$$\mathbf{F}_2 = 1/mn16-19 + 1/fn16-19 + 0.35/fh16-19; \text{ et}$$

$$\mathbf{F}_3 = 0.61/mb16-19 + 1/mh16-19 + 0.65/fh16-19 + 1/mn20-24 + 1/fn20 - 24;$$

Pour un seuillage $\alpha = (0.50,0.50)$, nous obtenons à la première phase : $ISOL(\Omega) = \{fb20-24, mb20-24\}$ et les \mathbf{R} -regroupements empiétants suivants :

$$\mathbf{B}_1 = \{mb16-19, mh20-24, fh20-24\}; \text{ et}$$

$$\mathbf{B}_2 = \{mn16-19, fn16-19, mh16-19, fh16-19, mn20-24, fn20 - 24\};$$

et à la seconde phase : $ISOL(\Omega) = \{fb20-24, mb20-24\}$ et les \mathbf{R} -regroupements flous suivants :

$$\mathbf{F}_1 = 1/mb16-19 + 1/mh20-24 + 1/fh20-24; \text{ et}$$

$$\mathbf{F}_2 = 1/mn16-19 + 1/fn16-19 + 1/mh16-19 + 1/fh16-19 + 0.57/mn20-24 + 0.59/fn20-24.$$

8. CONCLUSION

La littérature sur la classification abonde de techniques dont la plupart sont des techniques d'agglomération dites conventionnelles. Or, pour de nombreuses raisons mathématiques et pratiques, il est devenu plus avantageux de considérer des méthodes de coalescence donnant lieu à des agglomérats flous ou à la limite empiétants. Les concepts et la méthode de coalescence décrits dans le présent article nous permettent :

a) de construire des classes de similarité dites

1 - empiétantes,

2 - μ - différenciées, *i.e.*, les degrés d'appartenance des différents objets à ces classes ne sont pas tous égaux à 1;

3 - \mathbf{R} - homogènes, *i.e.*, $\mathbf{R}(X,Y) = 1$, dès que X et Y sont pris dans la même classe.

b) de mettre en évidence les éléments isolés, les éléments centraux (*i.e.*, dont l'appartenance à une classe est totale) ainsi que les éléments périphériques (*i.e.*, appartenant à plusieurs classes).

De tels concepts et une telle méthode de coalescence trouvent facilement en l'analyse économique spatiale, entre autres, un terrain d'application, étant donné que l'espace géographique flou est un espace différencié et recouvrable (cf.[4]).

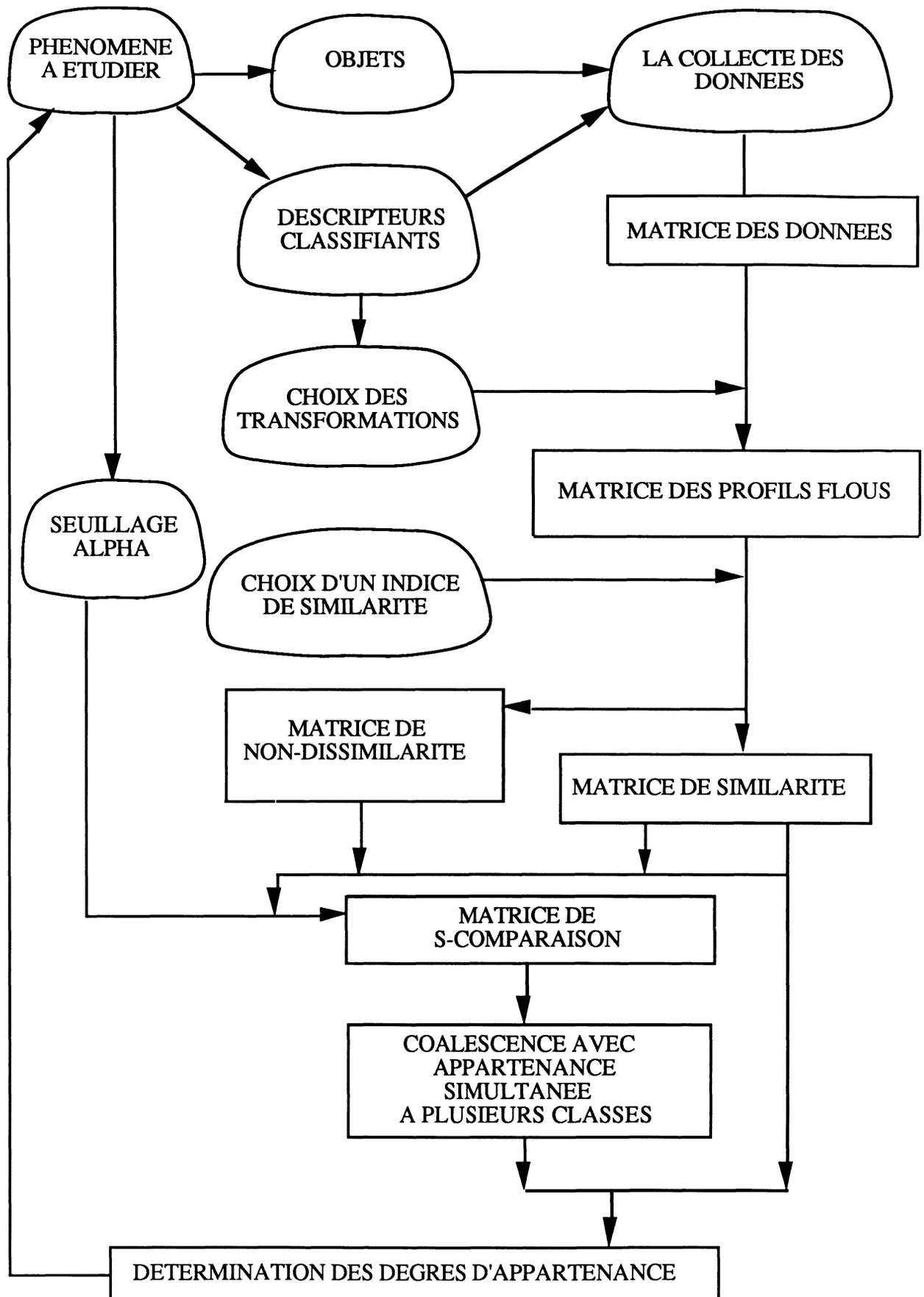


Figure 1 - Organigramme de la méthode de coalescence floue à deux phases

ANNEXE

		16-19					20-24					
		mb	mn	fn	mh	fh	mb	fb	mn	fn	mh	fh
16-19	mb	1							1	1		
	mn		1	1		1						
	fn		1	1								
	mh				1	1		1	1			
	fh		1		1	1		1	1			
20-24	mb						1					
	fb							1				
	mn	1			1	1			1	1		
	fn	1			1	1			1	1		
	mh										1	1
	fh										1	1

Table 2 - Matrice d'une S-comparaison plus fine pour $\alpha = (.7, .7)$

BIBLIOGRAPHIE

- [1] ABID Z., *Contribution à l'analyse structurale des systèmes complexes au moyen des notions de recouvrements et de semi-valuation*, Thèse d'Etat, Université Claude Bernard Lyon-I, mai 1986, pp. 31-39.
- [2] BEZDEK J. C., *Pattern recognition with fuzzy objective function algorithms*, New York, Plenum Press, 1981.
- [3] DUBOIS D. et PRADE H., "A unifying view of comparison indices in a fuzzy set-theoretic framework", *Recent developments in fuzzy set and possibility theory*, ed. Ronald R. Yager, New York, Pergamon Press, 1982, pp.3-13.
- [4] ROLLAND-MAY C., "La théorie des ensembles flous et son intérêt en géographie", *L'espace géographique*, n° 1, 1987, pp. 42 - 50.
- [5] SUGENO M., *Theory of fuzzy integrals and its applications*, Ph.D Thesis, Tokyo, Tokyo Institute of Technology, 1974 .
- [6] ZADEH L. A., "Fuzzy sets and their application to pattern classification and clustering analysis", *Classification and clustering*, ed. Van Ryzin, New York, Academic Press, 1977, pp. 251-299.
- [7] ZELENY M., *Multiple criteria decision making*, New York, Mc Graw-Hill, 1985.