

P. BERTRAND

**Propriétés et caractérisations topologiques d'une
représentation pyramidale**

Mathématiques et sciences humaines, tome 117 (1992), p. 5-28

http://www.numdam.org/item?id=MSH_1992__117__5_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/>

PROPRIÉTÉS ET CARACTÉRISATIONS TOPOLOGIQUES D'UNE REPRÉSENTATION PYRAMIDALE

P. BERTRAND¹

RÉSUMÉ — *Ce texte présente quelques caractéristiques géométriques des dissimilarités robinsoniennes. Ces dissimilarités constituent un modèle très général de représentation des mesures de proximité entre objets (ou groupes d'objets) lorsque ces entités sont rangées suivant un ordre total. Les propriétés géométriques des dissimilarités robinsoniennes sont exposées en utilisant les notions de segment et de frontière introduites pour une dissimilarité quelconque. Nous considérons ensuite l'ensemble des dissimilarités robinsoniennes ne possédant pas de triplet ultramétrique ; cet ensemble est dense dans l'ensemble des dissimilarités robinsoniennes, et il est facile de le caractériser par des conditions géométriques.*

SUMMARY — *Properties and topological characterizations of a pyramidal representation. This paper presents some geometric characteristics of robinsonian dissimilarities. These dissimilarities make up a very general model for representing proximity measures between objects (or groups of objects) when these entities are ranked according to a total order. Geometric properties of robinsonian dissimilarities are presented with the help of the concepts of segment and boundary which are introduced for an arbitrary dissimilarity. We then focus on the set of robinsonian dissimilarities without any ultrametric triplet ; this set is dense in the set of robinsonian dissimilarities, and can easily be characterized by geometric conditions.*

INTRODUCTION

Les dissimilarités robinsoniennes (Robinson (1951)), encore appelées pyramidales, présentent un grand intérêt pour les problèmes de sériation. En fait, une dissimilarité robinsonienne est caractérisée par l'existence d'au moins un ordre tel que les objets diffèrent entre-eux d'autant plus que l'écart entre leurs rangs pour cet ordre est grand. Une telle condition est évidemment satisfaite par les dissimilarités obtenues à partir de toute méthode de sériation, par exemple toute méthode de multidimensional scaling en dimension un.

L'attrait pour ce type de dissimilarité s'est accru depuis l'introduction récente (Diday (1984)) de la représentation pyramidale, généralisation du modèle hiérarchique, qui permet de visualiser par un graphe plan les dissimilarités robinsoniennes. Jusqu'à présent, plusieurs auteurs se sont posé le problème de trouver une dissimilarité robinsonienne proche d'une dissimilarité quelconque donnée (Robinson (1951), Kendall (1969), Hubert (1974), Diday (1984), Bertrand (1986), Guénoche (1987), Durand (1989), Batbedat (1990), ...). Cependant, la recherche d'une telle dissimilarité robinsonienne afin d'approcher une dissimilarité quelconque, mène en général à des problèmes de complexité combinatoire élevée. Dans ce texte, nous proposons de nouvelles caractérisations des dissimilarités robinsoniennes, ce qui permet de poser d'une nouvelle façon le problème de la recherche d'une approximation robinsonienne.

¹ Université Paris IX - Dauphine et I.N.R.I.A. (Institut National de Recherche en Informatique et en Automatique).

Dans la première partie, nous examinons les propriétés géométriques des dissimilarités robinsoniennes (Critchley et Van Cutsem (1989) ont abordé ce sujet dans le cadre plus spécifique des ultramétriques). Après avoir introduit les notions de *frontière* et de *segment* pour une dissimilarité quelconque, nous présentons des propriétés qui permettent de se faire une meilleure idée des cas où le modèle robinsonien est pertinent. Ces propriétés mettent en évidence l'intérêt des dissimilarités robinsoniennes pour lesquelles il n'existe aucun triangle isocèle ayant sa base plus petite que les deux côtés égaux, que nous avons appelées dissimilarités robinsoniennes *sans triplet ultramétrique*. Cette classe de dissimilarités a pour caractéristique d'être dense dans l'ensemble des dissimilarités robinsoniennes. On montre de plus que chaque dissimilarité de ce type n'est compatible qu'avec deux ordres (ces deux ordres étant inverses l'un de l'autre).

Dans la deuxième partie, nous déterminons diverses caractérisations des dissimilarités robinsoniennes en commençant par le cas des dissimilarités robinsoniennes sans triplet ultramétrique. Nous caractérisons ces dissimilarités à l'aide de la fonction L qui associe à toute paire de points l'intersection des deux boules centrées sur ces deux points et de même rayon égal à la dissimilarité entre ces points. Nous montrons qu'une dissimilarité est robinsonienne et sans triplet ultramétrique si et seulement si la fonction L est une fonction d'intervalle, ou encore si et seulement si l'ensemble des images par L est une pyramide saturée. La première condition, déjà étudiée par plusieurs auteurs (par exemple, M. Sholander (1952)), donne lieu à d'autres caractérisations n'utilisant que la fonction L.

Enfin nous donnons une caractérisation des dissimilarités robinsoniennes quelconques. Seule la deuxième condition énoncée précédemment se généralise à ce cas et l'on obtient une formulation proche de celle adoptée par Durand et Fichet (1988).

1. ORDRES, DISSIMILARITÉS ET PYRAMIDES

Par la suite, $\Omega = \{a, b, c, d, \dots\}$ désigne l'ensemble des objets étudiés, encore appelés points, $\mathbb{P}(\Omega)$ l'ensemble des parties de Ω , $\text{card}(X)$ le nombre d'éléments d'un ensemble fini X , δ une dissimilarité définie sur Ω , et θ (noté aussi \leq_θ) un ordre total défini sur Ω . Donnons les définitions de base que nous utilisons par la suite :

Ordre total, intervalle et fonction d'intervalle

Un ordre θ sur Ω est dit *total* si deux éléments quelconques sont comparables selon cet ordre. Soit a et b appartenant à Ω tels que $a \theta b$. L'*intervalle* fermé de bornes a et b , noté $[a, b]_\theta$, est défini par : $[a, b]_\theta = \{x \in \Omega \mid a \leq_\theta x \leq_\theta b\} = \{x \in \Omega \mid a \theta x \theta b\}$.

Quand il n'y aura pas d'ambiguïté sur le choix de θ , on écrira $[a, b]$ au lieu de $[a, b]_\theta$.

On définit de même les intervalles $]a, b[$, $]a, b]$, $[a, b[$ qui sont respectivement les intervalles ouvert, semi-ouvert à gauche et semi-ouvert à droite, de bornes a et b .

L'ensemble Ω est dit *borné* pour θ , si Ω possède un unique plus petit élément et un unique plus grand élément selon l'ordre θ .

Par la suite, nous n'envisagerons que les ordres totaux vérifiant la propriété suivante :

$$A \text{ est un intervalle de } \theta \Leftrightarrow \forall x, y \in A, (x \theta y \Rightarrow [x, y] \subseteq A)$$

Cette propriété est vraie pour les ordres totaux usuels : i.e. ceux définis sur \mathbb{R} ou sur un ensemble fini.

La *fonction d'intervalle* d'un graphe est une notion classique en théorie des graphes (voir par exemple Mulder (1980)), sa valeur pour deux sommets r et s étant l'ensemble des sommets situés sur les plus courts chemins reliant r et s .

Dans le cas d'un graphe qui est une chaîne (cas d'un ordre total si on oriente la chaîne), la *fonction d'intervalle* d'un ordre total θ , notée I_θ , est définie par :

$$\forall a, b \in \Omega^2, \quad I_\theta(a, b) = \begin{cases} [a, b] & \text{si } a \theta b, \\ [b, a] & \text{sinon.} \end{cases}$$

PROPOSITION 1 : Si Ω est borné pour l'ordre total θ , alors :

- (i) $\forall a, b \in \Omega, \quad I_\theta(a, b) = I_\theta(b, a)$
- (ii) $\exists m$ et $M \in \Omega$ tels que : $\Omega = I_\theta(m, M)$
- (iii) $\forall a, b \in \Omega, \forall c \in I_\theta(a, b), \quad I_\theta(a, c) \cap I_\theta(c, b) = \{c\}$
et $I_\theta(a, c) \cup I_\theta(c, b) = I_\theta(a, b)$
- (iv) $\forall a, b, c \in \Omega$, l'une au moins des conditions suivantes est vraie :
 $c \in I_\theta(a, b), b \in I_\theta(a, c)$ ou $a \in I_\theta(b, c)$
- (v) $\forall a, b, c \in \Omega, \quad a \in I_\theta(c, b) \text{ et } c \in I_\theta(a, b) \Rightarrow a = c$

Preuve. Il s'agit de conséquences directes de la définition des intervalles fermés d'un ordre total.

LEMME 1 : Si θ_1 et θ_2 sont deux ordres totaux distincts, alors θ_1 et θ_2 admettent la même fonction d'intervalle si et seulement si θ_2 est l'ordre inverse de θ_1 .

Preuve. Supposons que θ_1 et θ_2 admettent la même fonction d'intervalle. Comme θ_1 et θ_2 sont deux ordres distincts, il existe deux éléments a et b , tels que :

$$a \theta_1 b \quad \text{et} \quad b \theta_2 a$$

Soit x un élément quelconque de Ω . Les configurations possibles des éléments x , a et b pour l'ordre θ_1 , sont : $x \theta_1 a \theta_1 b$, $a \theta_1 x \theta_1 b$ et $a \theta_1 b \theta_1 x$

Les fonctions d'intervalle de θ_1 et de θ_2 sont identiques, donc en notant I cette fonction d'intervalle commune et en tenant compte de $b \theta_2 a$, nous avons :

$$\begin{array}{llll} x \theta_1 a \theta_1 b & \Rightarrow & a \in I(x, b) & \Rightarrow & b \theta_2 a \theta_2 x \\ a \theta_1 x \theta_1 b & \Rightarrow & x \in I(a, b) & \Rightarrow & b \theta_2 x \theta_2 a \\ a \theta_1 b \theta_1 x & \Rightarrow & b \in I(a, x) & \Rightarrow & x \theta_2 b \theta_2 a \end{array}$$

Ce qui prouve les implications suivantes :

$$a \theta_1 x \Rightarrow x \theta_2 a \quad \text{et} \quad x \theta_1 a \Rightarrow a \theta_2 x \quad (*)$$

Pour montrer que θ_2 est l'ordre inverse de θ_1 , il suffit de prouver que pour tout x, y de Ω , on a :

$$x \theta_1 y \Rightarrow y \theta_2 x$$

Supposons donc $x \theta_1 y$. Il y a alors trois configurations possibles entre x , y et a qui sont :

$$a \theta_1 x \theta_1 y, \quad x \theta_1 a \theta_1 y \quad \text{et} \quad x \theta_1 y \theta_1 a.$$

En tenant compte des implications (*) et du fait que θ_1 et θ_2 ont la même fonction d'intervalle, il vient :

$$\begin{array}{llll} a \theta_1 x \theta_1 y & \Rightarrow & x \in I(a,y), \quad x \theta_2 a \quad \text{et} \quad y \theta_2 a & \Rightarrow & y \theta_2 x \theta_2 a \\ x \theta_1 a \theta_1 y & \Rightarrow & a \theta_2 x \quad \text{et} \quad y \theta_2 a & \Rightarrow & y \theta_2 a \theta_2 x \\ x \theta_1 y \theta_1 a & \Rightarrow & y \in I(a,x), \quad a \theta_2 y \quad \text{et} \quad a \theta_2 x & \Rightarrow & a \theta_2 y \theta_2 x \end{array}$$

Ce qui prouve que $x \theta_1 y$ implique $y \theta_2 x$, et par conséquent que θ_2 est l'ordre inverse de θ_1 . Réciproquement si θ_2 est l'ordre inverse de θ_1 , alors il est clair que θ_1 et θ_2 admettent la même fonction d'intervalle.

Dissimilarités

Rappelons qu'une *dissimilarité* δ sur Ω est une application de Ω^2 dans \mathbf{R}^+ telle que :

$$\forall x,y \in \Omega, \quad \delta(x,y) = \delta(y,x) \geq \delta(x,x) = 0$$

Un ordre θ est dit *δ -compatible* si θ est total et si :

$$\forall x,y,z \in \Omega, \quad x \theta y \theta z \Rightarrow \delta(x,z) \geq \max(\delta(x,y), \delta(y,z))$$

Une dissimilarité δ est dite :

- *propre* si : $\forall x,y \in \Omega, \quad \delta(x,y) = 0 \Rightarrow x = y$
- *bornée* si : $\exists m \text{ et } M \in \Omega \quad \text{tq} \quad \forall x,y \in \Omega, \quad \delta(x,y) \leq \delta(m,M)$
- *ultramétrique* si : $\forall x,y,z \in \Omega, \quad \delta(x,z) \leq \max(\delta(x,y), \delta(y,z))$
- *robinsonienne* (Robinson (1951) ; on dit encore *pyramidale*) si :
 δ est propre et s'il existe un ordre δ -compatible.
- *fortement robinsonienne* (Durand et Fichet (1988)) si :
 δ est robinsonienne et si pour tout ordre θ qui est δ -compatible et pour tout triplet (x,y,z) satisfaisant $x \theta y \theta z$, on a :
 $\delta(x,z) = \delta(y,z) \Rightarrow \forall w \text{ tel que } w \theta z, \text{ on a } \delta(x,w) = \delta(y,w)$
 $\delta(x,z) = \delta(x,y) \Rightarrow \forall w \text{ tel que } x \theta w, \text{ on a } \delta(w,z) = \delta(w,y).$

Les dissimilarités ultramétriques et robinsoniennes peuvent être visualisées par des graphes plans, plus précisément, par des arbres (dans le cas des ultramétriques) et par des pyramides indicées au sens large (dans le cas des dissimilarités robinsoniennes (Diday (1984))).

Pyramides

Une famille \mathbf{P} de parties non vides de Ω , est appelé *pyramide classifiante sur Ω* si :

- (p1) $\Omega \in \mathbf{P}$, et $\forall x \in \Omega, \{x\} \in \mathbf{P}$,
- (p2) $\forall A, B \in \mathbf{P}, \quad A \cap B \in \mathbf{P} \cup \{\emptyset\}$,
- (p3) Il existe un ordre sur Ω tel que tout élément de \mathbf{P} soit un intervalle de cet ordre.

Par la suite, tout ordre vérifiant la condition (p3) sera dit *compatible avec \mathbf{P}* . Les éléments de \mathbf{P} sont appelés *paliers*, et par souci de simplification, nous dirons que \mathbf{P} est une *pyramide* au lieu de *pyramide classifiante*.

Enfin nous dirons qu'un palier B est un *prédécesseur* du palier A si :

$$A \subset B, A \subseteq C \subset B \text{ et } C \in P \Rightarrow A = C$$

où le symbole " \subset " désigne l'inclusion stricte.

Soit f une application définie sur une pyramide P et à valeurs dans \mathbf{R}^+ . Le couple (P, f) est appelé *pyramide indicée au sens large* si pour tout A, B dans P :

$$(p4) \quad f(A) = 0 \quad \Leftrightarrow \quad \text{card}(A) = 1,$$

$$(p5) \quad A \subseteq B \quad \Rightarrow \quad f(A) \leq f(B),$$

$$(p6) \quad A \subset B \text{ et } f(A) = f(B) \quad \Rightarrow \quad A \text{ admet deux prédécesseurs.}$$

La condition (p6) peut ne pas paraître naturelle, sa justification vient a posteriori par plusieurs résultats intéressants, à savoir : les hiérarchies indicées au sens large sont des pyramides indicées au sens large et les ultramétriques sont des dissimilarités robinsoniennes (Diday (1984)) ; de plus, on peut étendre la bijection classique entre hiérarchies indicées au sens large et ultramétriques, en mettant en correspondance bijective les pyramides indicées au sens large et les dissimilarités robinsoniennes (Diday (1984)). Une autre extension de cette bijection classique a été proposée par Durand et Fichet (1988) : cette extension met en correspondance bijective un cas particulier de pyramides indicées au sens large, i.e. les pseudo-hiérarchies (totales) indicées, avec les dissimilarités fortement robinsoniennes.

Boules, cercles

On définit généralement deux types de boules de centre a et de rayon r ($r \geq 0$) :

- la *boule ouverte* de centre a et de rayon r , notée $BO(a, r)$:

$$BO(a, r) = \{x \in \Omega \mid \delta(a, x) < r\}$$

- la *boule fermée* de centre a et de rayon r , notée $B(a, r)$:

$$B(a, r) = \{x \in \Omega \mid \delta(a, x) \leq r\}$$

Le *cercle de centre a et de rayon r* , noté $C(a, r)$, est égal à $B(a, r) - BO(a, r)$, c'est-à-dire :

$$C(a, r) = \{x \in \Omega \mid \delta(a, x) = r\}.$$

Triplets ultramétriques

Un triplet (a, b, c) de Ω est dit *ultramétrique* si :

$$\delta(a, b) = \delta(a, c) \geq \delta(b, c)$$

Une dissimilarité δ est dite *sans triplet ultramétrique* si aucun triplet de points distincts (deux à deux) n'est ultramétrique, autrement dit si tous les triangles isocèles de δ ont une base strictement plus grande que les deux côtés égaux.

Fonctions L et \bar{L}

Nous noterons par L_δ , ou plus simplement L , s'il n'y a pas d'ambiguïté sur le choix de δ , la fonction suivante :

$$L_\delta(a, b) = B(a, \delta(a, b)) \cap B(b, \delta(b, a)) = \{x \in \Omega \mid \delta(a, b) \geq \max \{\delta(a, x), \delta(x, b)\}\}.$$

Jain et Dubes (1988) signalent l'usage de cette fonction (qu'ils nomment *lune*, d'où la notation L) en reconnaissance des formes dans le cadre de la technique RNG (Relative Neighborhood Graph). Durand et Fichet (1988) utilisent aussi cette notion pour établir le théorème de bijection entre dissimilarités robinsoniennes et pyramides indicées au sens large.

De même, nous noterons LO la fonction définie sur Ω^2 par :

$$LO(a,b) = BO(a,\delta(a,b)) \cap BO(b,\delta(b,a)) = \{x \in \Omega \mid \delta(a,b) > \max(\delta(a,x),\delta(x,b))\}$$

Enfin nous noterons \bar{L} la fonction définie sur Ω^2 par :

$$\bar{L}(a,b) = \cap \{L(x,y) \mid a \in L(x,y) \text{ et } b \in L(x,y)\}$$

REMARQUE 1 : une dissimilarité propre δ est robinsonienne si et seulement s'il existe un ordre θ tel que pour tout x et y dans Ω , on a $I_\theta(x,y) \subseteq L(x,y)$.

Illustration géométrique dans le plan euclidien

Dans cet exemple, $L(a,b)$ (resp. $LO(a,b)$) désigne l'ensemble des points situés dans la zone hachurée, la frontière étant comprise (resp. non comprise), la dissimilarité δ étant ici la distance euclidienne usuelle.

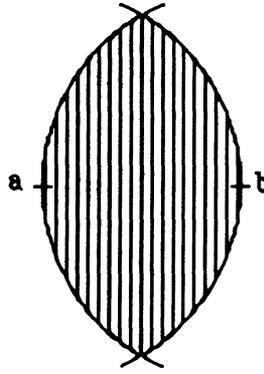


Figure 1

2. PROPRIÉTÉS GÉOMÉTRIQUES DES DISSIMILARITÉS ROBINSONIENNES

Nous introduisons les notions de segment, de frontière et d'intérieur pour une dissimilarité δ quelconque. Les propriétés de ces notions sont étudiées dans le cas où δ est robinsonienne.

2.1. Fonction segment

Nous proposons une définition du *segment* joignant deux points dans le cas d'une dissimilarité δ propre mais non nécessairement euclidienne. Si δ est euclidienne, on prouve que cette définition équivaut à la définition usuelle d'un segment entre deux points.

DÉFINITION : On appelle *segment (fermé) d'extrémités a et b*, la partie de Ω , notée $S_\delta(a,b)$, définie par :

$$S_\delta(a,b) = \{x \in \Omega \mid \forall y \in \Omega, \delta(a,x) \leq \delta(a,y) \text{ ou } \delta(b,x) \leq \delta(b,y)\}.$$

Cette définition de $S_\delta(a,b)$ peut aussi s'énoncer sous la forme :

$$S_\delta(a,b) = \{x \in \Omega \mid \nexists y \in \Omega, \delta(a,y) < \delta(a,x) \text{ et } \delta(b,y) < \delta(b,x)\}.$$

S'il n'y a pas d'ambiguïté sur le choix de la dissimilarité δ , on notera cette partie $S(a,b)$.

Comme les fonctions L , LO et \bar{L} , la fonction S est symétrique.

Exemple :

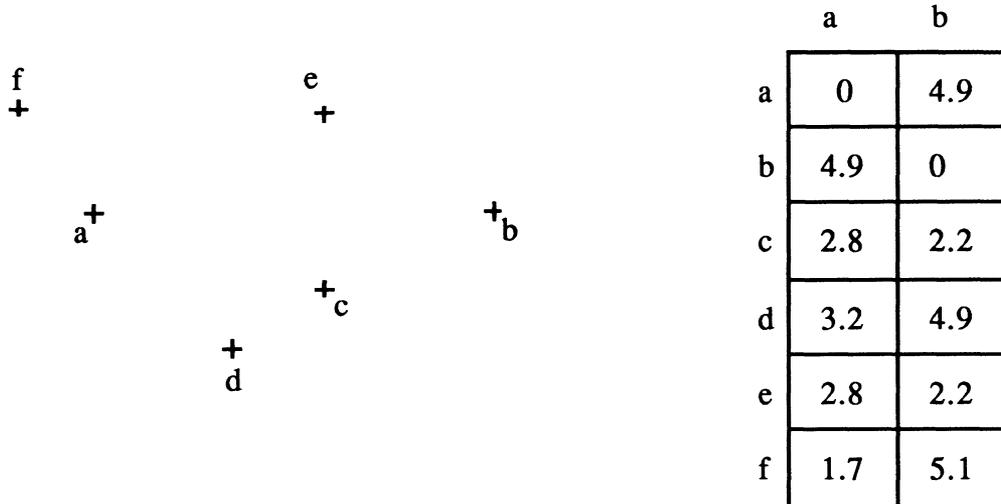


Figure 2

Dans cet exemple, les dissimilarités entre a,b,c,d,e et f sont simplement les distances en centimètres entre ces points dans le plan euclidien usuel.

Remarquons que le contenu d'un segment dépend du référentiel Ω : ainsi, si l'on choisit

$\Omega = \{a,b,c,d,e,f\}$, alors $S(a,b) = \{a,b,c,e\}$ et $L(a,b) = S(a,b) \cup \{d\}$. Par contre, si l'on prend $\Omega = \{a,b,d,f\}$, alors $S(a,b) = \{a,b,d\} = L(a,b)$.

PROPOSITION 2 : Si δ est une dissimilarité propre alors pour tout a et b dans Ω , on a :

- (i) $S(a,a) = L(a,a) = \{a\}$,
- (ii) $\{a,b\} \subseteq S(a,b) \subseteq L(a,b)$.

Preuve.

(i) est évident.

(ii) L'inclusion $\{a,b\} \subseteq S(a,b)$ est évidente. Pour prouver que $S(a,b)$ est inclus dans $L(a,b)$, considérons un point x de $S(a,b)$, avec x distinct de a et de b, et considérons successivement les points a et b, alors d'après la définition de S :

$$(1) \delta(a,x) \leq \delta(a,a) = 0 \quad \text{ou} \quad \delta(b,x) \leq \delta(b,a).$$

$$(2) \delta(a,x) \leq \delta(a,b) \quad \text{ou} \quad \delta(b,x) \leq \delta(b,b) = 0.$$

Comme δ est propre et $x \notin \{a,b\}$, on en déduit que $\delta(b,x) \leq \delta(b,a)$ et $\delta(a,x) \leq \delta(a,b)$. Par conséquent x appartient à $L(a,b)$, et donc $S(a,b) \subseteq L(a,b)$.

Montrons que la notion de segment ainsi introduite généralise la notion usuelle de segment définie dans le cadre euclidien.

PROPOSITION 3 : Si Ω est un espace affine euclidien et si δ est une distance euclidienne sur Ω , alors pour tout a et b de Ω , $S(a,b) = \{\lambda a + (1-\lambda)b \mid 0 \leq \lambda \leq 1\}$

Preuve. Étant donné deux points a et b de Ω , notons ab l'ensemble défini par :

$$ab = \{\lambda a + (1-\lambda)b \mid 0 \leq \lambda \leq 1\}$$

Soit x un élément de $S(a,b)$ et h la projection orthogonale de x sur la droite passant par les points a et b. Deux cas se présentent selon que h appartient ou n'appartient pas à ab (cf. figure 3) :

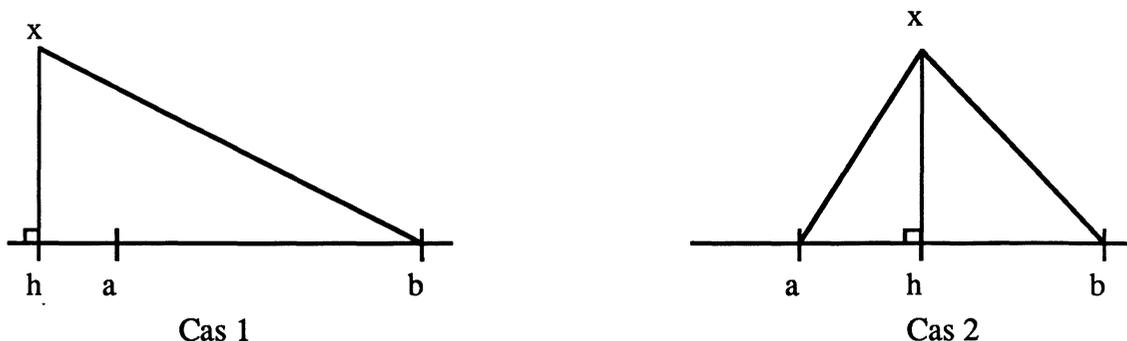


Figure 3

Lorsque h n'appartient pas à ab (cf. cas1), on a :

$$\max(\delta(h,a), \delta(h,b)) > \delta(a,b).$$

D'où d'après le théorème de Pythagore, $\max(\delta(x,a), \delta(x,b)) > \delta(a,b)$, donc x n'appartient pas à $L(a,b)$, ce qui contredit l'hypothèse x appartient à $S(a,b)$ (cf. proposition 2-(ii)).

Lorsque h appartient à ab (cf. cas2), si l'on suppose $x \neq h$, alors d'après le théorème de Pythagore, on a $\delta(h,a) < \delta(x,a)$ et $\delta(h,b) < \delta(x,b)$, ce qui contredit l'hypothèse x appartient à $S(a,b)$. Par conséquent, le seul cas possible est $x = h$ avec h élément de ab . Donc $S(a,b) \subseteq ab$. Montrons que ab est inclus dans $S(a,b)$: soit z un élément de ab , y un élément quelconque de Ω , et h la projection de y sur la droite passant par a et b . Un examen des situations possibles entre h et z montre immédiatement que :

$$\delta(z,a) \leq \delta(h,a) \text{ ou } \delta(z,b) \leq \delta(h,b).$$

D'après le théorème de Pythagore, on en déduit que $\delta(z,a) \leq \delta(y,a)$ ou $\delta(z,b) \leq \delta(y,b)$, ce qui prouve que z appartient à $S(a,b)$. Donc $ab \subseteq S(a,b)$, et par conséquent $ab = S(a,b)$.

2.2. Propriétés des fonctions L , LO et S dans le cas robinsonien

Dans le cas où δ est robinsonienne, C. Durand (1989, p. 57) a montré que L et \bar{L} sont des intervalles de tout ordre δ -compatible. La proposition ci-dessous contient ce résultat ainsi que d'autres, permettant par exemple de comparer \bar{S} avec L et LO .

PROPOSITION 4 : Si δ est robinsonienne et si θ est δ -compatible, alors pour tout a, b et c appartenant à Ω :

- (i) $LO(a,b) \subseteq I_\theta(a,b) \subseteq S(a,b) \subseteq L(a,b)$,
- (ii) les valeurs des fonctions S , L et \bar{L} sont des intervalles de θ ,
- (iii) $LO(a,b) \cap LO(b,c) \cap LO(a,c) = \emptyset$,
- (iv) $\forall c, d \in LO(a,b)$, $a \notin LO(c,d)$ et $b \notin LO(c,d)$.

Preuve.

(i) La relation $S(a,b) \subseteq L(a,b)$ est vraie d'après la proposition 2-(ii). Montrons l'inclusion $LO(a,b) \subseteq I_\theta(a,b)$. Pour cela raisonnons par l'absurde : supposons qu'il existe un point x de $LO(a,b)$ qui n'appartienne pas à $I_\theta(a,b)$. Quitte à permuter a et b , supposons $a \theta b$. Deux cas se présentent : $x \theta a \theta b$ et $a \theta b \theta x$. Dans le premier cas, c'est-à-dire $x \theta a \theta b$, on a, puisque θ est δ -compatible, $\delta(x,b) \geq \delta(a,b)$. Dans le deuxième cas, c'est-à-dire $a \theta b \theta x$, on obtient

$\delta(a,x) \geq \delta(a,b)$. Dans les deux cas, il y a contradiction avec l'hypothèse $x \in LO(a,b)$. Par conséquent, $LO(a,b) \subseteq I_\theta(a,b)$.

Montrons $I_\theta(a,b) \subseteq S(a,b)$. Supposons encore $a \theta b$, ce qui est possible quitte à permuter a et b . Soit x un élément de $I_\theta(a,b)$, c'est-à-dire $a \theta x \theta b$, et soit y un élément de Ω , on a :

- si $y \theta x$, alors $y \theta x \theta b$, donc $\delta(b,x) \leq \delta(b,y)$,

- si $x \theta y$, alors $a \theta x \theta y$, donc $\delta(a,x) \leq \delta(a,y)$.

On en déduit que x appartient à $S(a,b)$ et donc $I_\theta(a,b) \subseteq S(a,b)$.

(ii) Montrons tout d'abord que S prend ses valeurs dans l'ensemble des intervalles de θ .

Notons m_1 (resp. m_2) le plus petit (resp. grand) élément de $S(a,b)$ selon l'ordre θ . Quitte à permuter a et b , supposons $a \theta b$. Par conséquent :

$$m_1 \theta a \theta b \theta m_2$$

D'après (i), $[a,b]$ est inclus dans $S(a,b)$, il suffit donc de prouver que $[m_1,a]$ et $[b,m_2]$ sont inclus dans $S(a,b)$. Considérons un élément x de $[m_1,a]$, on a :

$$m_1 \theta x \theta a \theta b, \text{ donc } \delta(a,x) \leq \delta(a,m_1) \text{ et } \delta(b,x) \leq \delta(b,m_1)$$

Tout élément y de Ω vérifie $\delta(a,m_1) \leq \delta(a,y)$ ou $\delta(b,m_1) \leq \delta(b,y)$, car $m_1 \in S(a,b)$. D'où :

$$\delta(a,x) \leq \delta(a,m_1) \leq \delta(a,y) \text{ ou } \delta(b,x) \leq \delta(b,m_1) \leq \delta(b,y)$$

Ce qui prouve que x appartient à $S(a,b)$, donc $[m_1,a]$ est inclus dans $S(a,b)$. On montre de même que $[b,m_2]$ est inclus dans $S(a,b)$, donc $[m_1,m_2] \subseteq S(a,b)$, et par suite :

$$[m_1,m_2] = S(a,b).$$

Pour montrer que les valeurs de L et \bar{L} sont des intervalles de θ , il suffit de prouver que pour tout u et v dans Ω , $L(u,v)$ est un intervalle de θ . En effet, $\bar{L}(a,b)$ étant l'intersection de parties de la forme $L(u,v)$ ayant toutes au moins un point commun (i.e. a et b), $\bar{L}(a,b)$ sera bien un intervalle de θ si les parties $L(u,v)$ sont des intervalles de θ .

Il suffit donc de prouver que pour tout u et v dans Ω , on a :

$$x \in L(u,v), z \in L(u,v) \text{ et } x \theta y \theta z \Rightarrow y \in L(u,v).$$

A priori, deux cas se présentent : $u \theta y$ et $y \theta u$.

Si $u \theta y$, alors $u \theta y \theta z$, donc $\delta(u,y) \leq \delta(u,z)$, ce qui entraîne $\delta(u,y) \leq \delta(u,v)$ car $z \in L(u,v)$.

Si $y \theta u$, alors $x \theta y \theta u$, donc $\delta(u,y) \leq \delta(u,x)$, ce qui entraîne $\delta(u,y) \leq \delta(u,v)$ car $x \in L(u,v)$.

Dans les deux cas, on a $\delta(u,y) \leq \delta(u,v)$. Un raisonnement analogue montre que $\delta(v,y) \leq \delta(u,v)$. Donc y appartient bien à $L(u,v)$.

(iii) Cette propriété est une conséquence de l'inclusion $LO(a,b) \subseteq I_\theta(a,b)$ et des propriétés des intervalles d'un ordre total (cf. proposition 1).

(iv) La partie $LO(a,b)$ est incluse dans $I_\theta(a,b) - \{a,b\}$, donc pour toute paire $\{c,d\}$ d'éléments de $LO(a,b)$, on a $\{c,d\} \subseteq I_\theta(a,b) - \{a,b\}$. Donc $I_\theta(c,d)$ est inclus dans $I_\theta(a,b) - \{a,b\}$, ce qui implique que $LO(c,d)$ est inclus dans $I_\theta(a,b) - \{a,b\}$, et par suite, ni a ni b n'appartiennent à $LO(c,d)$.

REMARQUE 2 : On ne peut pas comparer $S(a,b)$ et $\bar{L}(a,b)$ au sens de l'inclusion. Les deux exemples décrits ci-dessous (cf. Figure 4) montrent que $S(a,b)$ peut être inclus dans $\bar{L}(a,b)$ (cf. cas 1), mais que l'on peut aussi obtenir des cas où $\bar{L}(a,b)$ est inclus dans $S(a,b)$ (cf. cas 2). Ainsi, dans le premier cas, $S(a,b) = \{a,b,d\}$ et $\bar{L}(a,b) = \{a,b,c,d\}$, alors que dans le deuxième cas, on a $S(a,b) = \{a,b,c,d\}$ et $\bar{L}(a,b) = \{a,b\} = L(c,b) \cap L(a,d)$.

	c	a	d	b
c	0	2	2	3
a		0	1	3
d			0	2
b				0

(cas 1)

	c	a	b	d
c	0	1	3	5
a		0	3	3
b			0	1
d				0

(cas 2)

Figure 4

Quelques cas de figures exclus dans le cadre robinsonien

En utilisant les propriétés de la proposition 4, nous donnons ci-dessous une liste (non exhaustive) de cas de figures exclus lorsque δ est robinsonienne. Dans chacun des exemples suivants, la valeur de la dissimilarité entre deux points est simplement la distance euclidienne dans le plan.

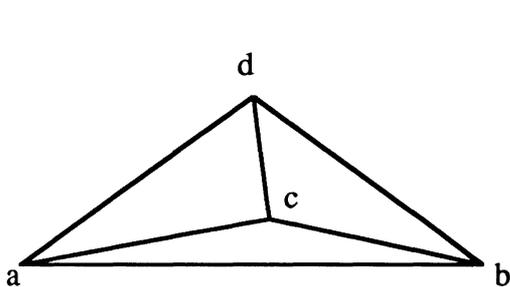


Figure 5a

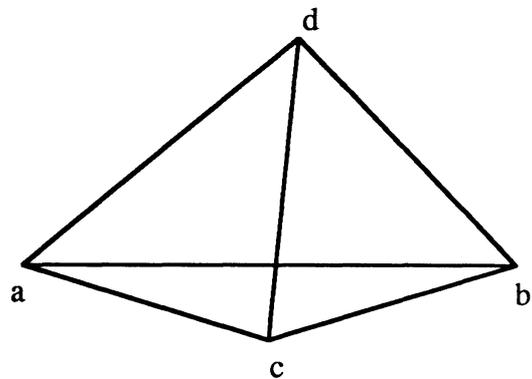


Figure 5b

Dans les deux cas de figures 5a et 5b, la propriété $LO(a,b) \subseteq S(a,b)$ montre que la distance entre les points n'est pas robinsonienne. En effet, on a alors : $LO(a,b) = \{c,d\}$ et $S(a,b) = \{a,b,c\}$.

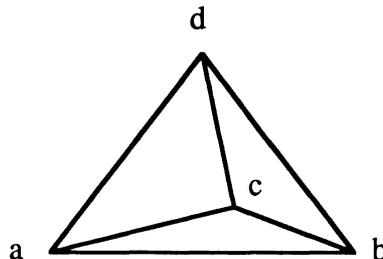


Figure 6

Ce cas de figure n'est pas exclu par la propriété $LO(a,b) \subseteq S(a,b)$ car ici $d \notin LO(a,b)$, mais la condition $LO(a,b) \cap LO(b,c) \cap LO(a,c) = \emptyset$ exclut le cas de la figure 6, car dans ce cas, on a : $LO(a,b) \cap LO(b,d) \cap LO(a,d) = \{c\}$. Cette propriété exclut aussi le cas de la figure (5a), mais pas celui de la figure (5b).

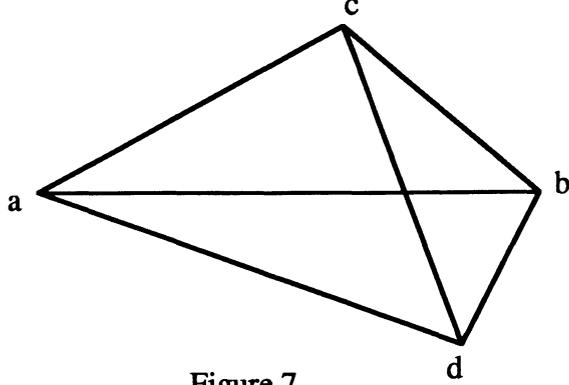


Figure 7

La propriété $\forall c,d \in LO(a,b), a \text{ et } b \notin LO(c,d)$ exclut la situation de la figure 7 : en effet, dans cette figure, $c \text{ et } d \in LO(a,b)$ et $b \in LO(c,d)$. Mais aucune des deux propriétés précédentes n'exclut cette situation.

COROLLAIRE 1 : Si δ est robinsonienne et sans triplet ultramétrique et si θ est δ -compatible, alors :

(i) $S = \bar{L} = L = I_\theta$

(ii) Il n'existe que deux ordres δ -compatibles, et ces deux ordres sont inverses l'un de l'autre.

Preuve.

(i) On constate facilement que :

$$L(a,b) = LO(a,b) \cup \{x \mid (a,x,b) \text{ ou } (b,x,a) \text{ est ultramétrique}\}.$$

Si δ est sans triplet ultramétrique, alors (a,x,b) (respectivement (b,x,a)) est ultramétrique si et seulement si $x = b$ (respectivement $x = a$). Donc $L(a,b) = LO(a,b) \cup \{a,b\}$, s'il n'existe pas de triplet ultramétrique. Si de plus δ est robinsonienne, alors d'après la proposition 4-(i), on a $LO(a,b) \subseteq S(a,b)$, donc $L(a,b) \subseteq S(a,b)$, ce qui prouve que $L(a,b) = S(a,b)$ car d'après la proposition 2-(ii) on a $S(a,b) \subseteq L(a,b)$.

On montre de même que $L(a,b) = I_\theta(a,b)$ en se servant des inclusions suivantes (cf. proposition 4-(i)) :

$$LO(a,b) \subseteq I_\theta(a,b) \subseteq L(a,b).$$

L'égalité $L(a,b) = \bar{L}(a,b)$ se montre de la même manière, en utilisant les relations suivantes déduites de la définition de \bar{L} et de la proposition 4-(i) et (ii) :

$$LO(a,b) \subseteq I_\theta(a,b) \subseteq \bar{L}(a,b) \subseteq L(a,b).$$

(ii) Il s'agit d'une conséquence de la propriété (i) ci-dessus et du lemme 1.

2.3. Frontière et intérieur d'une partie

Nous définissons la frontière et l'intérieur d'une partie de Ω pour une dissimilarité δ quelconque. Puis nous étudions les propriétés de ces notions dans le cas où δ est robinsonienne.

DÉFINITIONS : Soit A une partie de Ω , rappelons que le diamètre de A , noté $\text{diam}(A)$, est la plus grande valeur de δ séparant deux éléments de A :

$$\text{diam}(A) = \max \{ \delta(u,v) \mid u \in A, v \in A \}$$

Pour $x \in A$, on appelle *frontière de A selon x* l'ensemble des y de A tels que $\delta(x,y) = \text{diam}(A)$:

$$F_x(A) = \{ y \in A \mid \delta(x,y) = \text{diam}(A) \}.$$

On appelle *frontière de A*, et on note $F(A)$, l'ensemble défini par :

$$F(A) = \{ x \in A \mid \exists y \in A \text{ tq } \delta(x,y) = \text{diam}(A) \}$$

Autrement dit, $F(A) = \cup \{ F_x(A) \mid x \in A \}$.

L'*intérieur de A*, noté $I(A)$, est défini par :

$$I(A) = A - F(A) = A \cap F(A)^c = \{ x \in A \mid \forall y \in A, \delta(x,y) < \text{diam}(A) \}.$$

Autrement dit, $I(A) = \cup \{ x \in A \mid F_x(A) = \emptyset \}$.

Enfin on note R_A la relation d'équivalence définie sur A par :

$$x R_A y \Leftrightarrow F_x(A) = F_y(A)$$

Par la suite, lorsqu'il n'y aura pas d'ambiguïté possible sur le choix de la partie A , nous noterons \tilde{x} la classe d'équivalence de x selon R_A .

REMARQUE 3

(R1) $I(A)$ est la classe d'équivalence de R_A constituée des éléments x tels que $F_x(A) = \emptyset$. Il en résulte que les classes d'équivalence différentes de $I(A)$ forment une partition de $F(A)$, autrement dit : $F(A) = \cup \{ \tilde{x} \mid x \in F(A) \}$.

(R2) $\delta(x,y) = \text{diam}(A) \Leftrightarrow y \in F_x(A) \Leftrightarrow x \in F_y(A) \Leftrightarrow \tilde{y} \subseteq F_x(A) \Leftrightarrow \tilde{x} \subseteq F_y(A)$.

(R3) $\delta(x,y) < \text{diam}(A) \Leftrightarrow y \notin F_x(A) \Leftrightarrow x \notin F_y(A) \Leftrightarrow \tilde{y} \cap F_x(A) = \emptyset \Leftrightarrow \tilde{x} \cap F_y(A) = \emptyset$.

(R4) D'après les trois remarques précédentes, quel que soit l'élément x de $F(A)$, chaque classe d'équivalence \tilde{y} est soit incluse dans $F_x(A)$, soit disjointe de $F_x(A)$. Par conséquent :

$$F_x(A) = \cup \{ \tilde{y} \mid y \in F_x(A) \}.$$

(R5) si $x R_A y$ alors $\delta(x,y) < \text{diam}(A)$; en effet, sinon y appartiendrait à $F_x(A) = F_y(A)$ et donc nous aurions $\delta(y,y) = \text{diam}(A) > 0$. Il en résulte que $\tilde{x} \cap F_x(A) = \emptyset$.

Exemple (trapèze) : dans cet exemple, Ω est l'ensemble des points v, w, x, y et z disposés selon la figure 8 ci-dessous, $A = \Omega$, et δ est la restriction de la distance euclidienne usuelle à Ω .

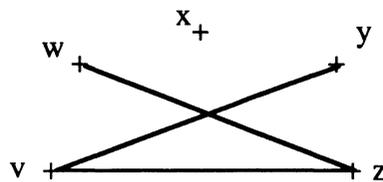


Figure 8.

Sur cette figure, on a indiqué par des traits les dissimilarités entre deux points qui sont égales au diamètre de $A = \Omega$. On vérifie aisément que :

$$\begin{aligned} \text{diam}(A) &= \delta(v,y) = \delta(w,z) = \delta(v,z) \\ F(A) &= \{v, w, y, z\} ; I(A) = \{x\} \\ F_v(A) &= \{y, z\} ; F_w(A) = \{z\} ; F_y(A) = \{v\} ; F_z(A) = \{v, w\} \end{aligned}$$

Ici la classe d'équivalence selon R_A de chaque élément de $F(A)$ est réduite à un singleton, par conséquent $F(A)$ est bien la réunion des classes d'équivalence de ses éléments.

2.4. Propriétés de la frontière et de l'intérieur d'une partie dans le cas robinsonien

PROPOSITION 5 : Si δ est robinsonienne et si θ est δ -compatible, alors toute classe d'équivalence de R_A est un intervalle de θ .

Preuve. Considérons une classe d'équivalence de R_A non réduite à un singleton (sinon le résultat est évident). Soit deux éléments distincts de cette classe : notons x le plus petit et y le plus grand de ces deux éléments selon l'ordre θ , et posons $K = F_x(A) = F_y(A)$.

Il s'agit de montrer que si z appartient à $[x,y]$, alors $K = F_z(A)$. Montrons l'inclusion $K \subseteq F_z(A)$ qui est évidente si $K = \emptyset$. Remarquons que si $K \neq \emptyset$, alors tout élément k de K ne peut être compris entre x et y selon l'ordre θ : en effet, sinon nous aurions :

$$\delta(x,y) \geq \delta(x,k) = \text{diam}(A),$$

ce qui entraînerait $\delta(x,y) = \text{diam}(A)$, et contredirait l'hypothèse $x R_A y$.

Par conséquent $k \notin [x,y]$. En utilisant la symétrie des rôles joués par x et y , il est possible de n'étudier que la situation $k \theta x$ (le cas $y \theta k$ étant symétrique du cas $k \theta x$). Donc supposons que l'on ait $k \theta x$, il en résulte $\delta(k,x) = \text{diam}(A) \leq d(k,z)$, ce qui entraîne que $\delta(k,z) = \text{diam}(A)$, donc $k \in F_z(A)$, d'où l'inclusion de K dans $F_z(A)$.

Montrons $F_z(A) \subseteq K$. Pour cela nous établissons $v \notin K \Rightarrow v \notin F_z(A)$.

Soit donc v un élément de A tel que $v \notin K = F_x(A) = F_y(A)$.

Si $v \theta z$, alors $v \theta z \theta y$ donc $\delta(v,z) \leq \delta(v,y) < \text{diam}(A)$.

Si $z \theta v$, alors $x \theta z \theta v$ donc $\delta(v,z) \leq \delta(v,x) < \text{diam}(A)$.

Dans les deux cas, v n'appartient pas à $F_z(A)$.

De la remarque (R1) et de la proposition précédente, nous déduisons le corollaire suivant :

COROLLAIRE 2 : Si δ est une dissimilarité robinsonienne, alors l'intérieur de toute partie de Ω est un intervalle de tout ordre δ -compatible.

PROPOSITION 6 : Soit n le nombre de classes d'équivalence de R_A différentes de l'intérieur de A , on a :

(1) Si $n \leq 3$, alors : $\forall x,y \in F(A), \tilde{x} \neq \tilde{y} \Rightarrow \delta(x,y) = \text{diam}(A)$

(2) Si $n \geq 3$, si δ est robinsonienne et s'il existe au moins trois éléments de $F(A)$ qui sont séparés deux à deux par une dissimilarité égale au diamètre de A , alors $I(A) = \emptyset$.

Preuve

(1) Nous supposons $n > 1$, car si $n = 1$ alors tous les éléments de la frontière sont équivalents selon R_A . Soit alors x et y deux éléments de $F(A)$ tels que $\tilde{x} \neq \tilde{y}$.

Si $n = 2$ alors \tilde{x} et \tilde{y} sont les seules classes différentes de l'intérieur de A . D'après les remarques (R4) et (R5), $F_x(A) = \tilde{y}$ car $F_x(A)$ ne peut être vide. Donc $\delta(x,y) = \text{diam}(A)$.

Si $n = 3$ alors il existe un élément z de la frontière de A qui n'est ni équivalent à x , ni équivalent à y . D'après les remarques (R4) et (R5), nous obtenons que :

$$F_x(A) \subseteq \tilde{y} \cup \tilde{z} \quad \text{et} \quad F_y(A) \subseteq \tilde{x} \cup \tilde{z}$$

Raisonnons par l'absurde : supposons $\delta(x,y) < \text{diam}(A)$, alors $F_x(A) = \tilde{z}$ puisque $F_x(A)$ n'est pas vide. De même $F_y(A) = \tilde{z}$, ce qui est contradictoire car $F_x(A)$ doit être différent de $F_y(A)$ puisque x et y ne sont pas équivalents. Donc $\delta(x,y) = \text{diam}(A)$.

(2) Soit θ un ordre δ -compatible et trois éléments de $F(A)$ séparés deux à deux par une dissimilarité égale au diamètre de A . Notons x, y , et z ces éléments rangés en ordre croissant selon l'ordre θ . Raisonnons par l'absurde : supposons que $I(A)$ ne soit pas vide. Dans ce cas, il existe un élément u de A qui vérifie :

$$\max(\delta(x,u), \delta(y,u), \delta(z,u)) < \text{diam}(A)$$

D'où $\max(\delta(x,u), \delta(y,u), \delta(z,u)) < \delta(x,y)$, ce qui prouve que u appartient à $]x,y[$. De même u appartient à $]y,z[$, ce qui est en contradiction avec l'hypothèse : x, y, z sont des éléments distincts et rangés en ordre croissant selon l'ordre θ . Par conséquent $I(A) = \emptyset$.

De la proposition 6, nous déduisons le résultat suivant :

COROLLAIRE 3 : Soit δ une dissimilarité robinsonienne sur Ω et A une partie de Ω et soit (i) et (ii), les deux affirmations suivantes :

(i) le nombre de classes d'équivalence de R_A , différentes de $I(A)$, est égal à 3.

(ii) le cardinal de $F(A)$ est supérieur ou égal à 3 et les éléments de $F(A)$ sont situés à "égale dissimilarité" les uns des autres.

Si (i) ou (ii) est vraie, alors $I(A) = \emptyset$.

REMARQUE 4 : le résultat (1) de la proposition 6 est en général faux si $n \geq 4$. L'exemple du trapèze donné plus haut, le prouve pour $n = 4$; en effet, pour ces données, les éléments y et z vérifient :

$$y \in F(A), \quad z \in F(A), \quad \tilde{y} = \tilde{z} \quad \text{et} \quad \delta(y,z) < \text{diam}(A)$$

3. CARACTÉRISATIONS DES DISSIMILARITÉS ROBINSONIENNES

Dans ce paragraphe, nous proposons plusieurs caractérisations des dissimilarités robinsoniennes. Lorsque l'on impose aux dissimilarités robinsoniennes de ne pas posséder de triplet ultramétrique, on obtient un sous-ensemble dense dans l'ensemble des dissimilarités robinsoniennes et pour lequel on peut donner plusieurs caractérisations géométriques (§3.1.).

3.1. Cas des dissimilarités robinsoniennes sans triplet ultramétrique

Nous savons que si δ est robinsonienne et sans triplet ultramétrique, alors la fonction segment L est une fonction d'intervalle (cf. corollaire 1). En établissant la réciproque (§3.1.1), nous obtenons une caractérisation des dissimilarités robinsoniennes sans triplet ultramétrique en termes de fonction d'intervalle. On en déduit que toute caractérisation des fonctions d'intervalle engendre une caractérisation des dissimilarités robinsoniennes sans triplet ultramétrique. Ainsi après avoir indiqué une nouvelle caractérisation des fonctions d'intervalle, nous proposons une caractérisation simple formulée avec la fonction L (§3.1.2). Finalement, nous indiquons une

caractérisation "globale" des dissimilarités robinsoniennes sans triplet ultramétrique lorsque Ω est fini (§3.1.3).

3.1.1. Caractérisation à l'aide de la notion de fonction d'intervalle

PROPOSITION 7 : Si δ est bornée, alors δ est robinsonienne et sans triplet ultramétrique si et seulement si L est une fonction d'intervalle.

Preuve. D'après le corollaire 1, il suffit de montrer que si L est une fonction d'intervalle, alors δ est robinsonienne et sans triplet ultramétrique.

Supposons que L soit une fonction d'intervalle. Dans ce cas, $L(x,x) = \{x\}$, donc δ est propre. De plus, il existe un ordre θ tel que pour tout x,y de Ω , $L(x,y) = I_\theta(x,y)$. Donc, d'après la remarque 1 (§1), δ est robinsonienne.

Il reste à prouver que δ n'admet pas de triplet ultramétrique. Raisonnons par l'absurde : supposons qu'il existe trois points distincts a,b,c tels que $\delta(b,c) \leq \delta(a,b) = \delta(a,c)$. Dans ce cas, $c \in L(a,b)$ et $b \in L(a,c)$, ce qui entraîne $b = c$ d'après la propriété (v) de la proposition 1, puisque L est une fonction d'intervalle, ce qui contredit l'hypothèse "b et c distincts".

REMARQUE 5 : Les dissimilarités robinsoniennes sans triplet ultramétrique constituent un cas particulier de dissimilarité fortement robinsonienne. En effet, par définition, pour toute dissimilarité robinsonienne sans triplet ultramétrique, il ne peut exister de triplet (x,y,z) tels que :

$$(x \theta y \theta z \text{ et } \delta(x,z) = \delta(y,z)) \quad \text{ou} \quad (x \theta y \theta z \text{ et } \delta(x,z) = \delta(x,y))$$

REMARQUE 6 : Le résultat de la proposition 7 n'est plus vrai si l'on remplace la fonction L par la fonction S . En effet, considérons l'exemple suivant où δ est définie par le tableau de la figure 9a et correspond à la configuration donnée sur la figure 9b.

	a	x	b	c	
a	0	1	3	3	a
x		0	1	2	x
b			0	1.8	b
c				0	c

Figure 9a

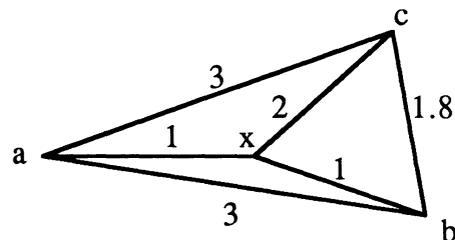


Figure 9b

Dans cet exemple, si l'on choisit $\Omega = \{a,b,c,x\}$, alors S est la fonction d'intervalle de l'ordre (a,x,b,c) , δ est robinsonienne, mais le triplet (a,b,c) est ultramétrique. De plus, S peut très bien ne pas être une fonction d'intervalle et δ être robinsonienne (considérer l'exemple ci-dessus et choisir $\Omega = \{a,b,c\}$).

M. Altwegg (1950) donne la caractérisation suivante des fonctions d'intervalle :

- (a1) $\forall a \in \Omega, \quad f(a,a) = \{a\}$
- (a2) $\forall a,b \in \Omega, \quad f(a,b) = f(b,a)$
- (a3) $b \in f(a,c) \text{ et } c \in f(b,d) \text{ et } b \neq c \Rightarrow b \in f(a,d)$
- (a4) $\forall a,b,c \in \Omega, \quad b \in f(a,c), \text{ ou } c \in f(b,a), \text{ ou } a \in f(c,b).$

M. Sholander (1952) considère les propriétés suivantes :

- (a1) $\forall a \in \Omega, \quad f(a,a) = \{a\}$
 (a5) $b \in f(a,c) \text{ et } d \in f(b,e) \quad \Rightarrow \quad b \in f(c,d) \text{ ou } b \in f(e,a)$
 (a6) $\forall a,b,c \in \Omega, \quad f(a,b) \cap f(b,c) \cap f(c,a) \neq \emptyset$
 (a4) $\forall a,b,c \in \Omega, \quad b \in f(a,c), \text{ ou } c \in f(b,a), \text{ ou } a \in f(c,b).$

M. Sholander montre que les conditions (a1), (a5) et (a6) caractérisent les arbres et que (a1), (a5) et (a4) caractérisent les fonctions d'intervalle.

De ces caractérisations et de la proposition 7, nous déduisons en tenant compte du fait que L vérifie toujours (a1) et (a2) :

THÉORÈME 1 : Si δ est bornée, alors δ est robinsonienne et sans triplet ultramétrique si et seulement si L vérifie la condition (a4) et l'une des deux conditions (a3) ou (a5).

3.1.2. Une nouvelle caractérisation des fonctions d'intervalle

D'après la proposition 1, la fonction d'intervalle f d'un ordre total borné sur Ω vérifie les quatre conditions suivantes:

- (i1) $\forall a,b \in \Omega, \quad f(a,b) = f(b,a)$
 (i2) $\exists m \text{ et } M \in \Omega, \quad \Omega = f(m,M)$
 (i3) $\forall c \in f(a,b), \quad f(a,c) \cap f(c,b) = \{c\}$
 (i4) $\forall c \in f(a,b), \quad f(a,b) = f(a,c) \cup f(c,b)$

Le but de ce paragraphe est de prouver que (i1), (i2), (i3) et (i4) caractérisent les fonctions d'intervalle d'un ordre borné.

Les contre-exemples présentés ci-dessous prouvent que les conditions (i1), (i2), (i3) et (i4) sont "indépendantes", c'est-à-dire qu'aucune de ces quatre conditions n'est impliquée par l'une des trois autres.

Contre-exemple E1

(i2), (i3) et (i4) $\not\Rightarrow$ (i1)

Soit $\Omega = \{a,b,c\}$ et pour tout $(x,y) \neq (a,b)$, $f(x,y) = \{x,y\}$, et $f(a,b) = \Omega$
 f vérifie (i2), (i3) et (i4), mais pas (i1) puisque $f(a,b) = \Omega \neq f(b,a) = \{b,a\}$.

Contre-exemple E2

(i1), (i3) et (i4) $\not\Rightarrow$ (i2)

Soit Ω un ensemble fini de cardinal strictement supérieur à 2 et f la fonction définie par :

$$\forall x,y \in \Omega, \quad f(x,y) = \{x,y\}$$

Trivialement, f vérifie (i1), (i3) et (i4) mais pas (i2).

Contre-exemple E3

(i1), (i2) et (i4) $\not\Rightarrow$ (i3)

Soit $\Omega = \{a,b,c\}$; $\forall x \in \Omega, f(x,x) = \{x\}$; $\forall x,y \in \Omega, f(x,y) = f(y,x)$
 $f(a,b) = f(b,c) = \Omega$; $f(a,c) = \{a,c\}$

Les axiomes (i1), (i2) sont trivialement vérifiés et (i4) l'est car :

$$f(a,b) = \Omega = f(a,c) \cup f(c,b) \text{ et } f(b,c) = \Omega = f(b,a) \cup f(a,c)$$

Par contre (i3) n'est pas vérifié car $a \in f(b,c)$ et $f(b,a) \cap f(a,c) = \{a,c\}$.

Contre-exemple E4

(i1), (i2) et (i3) $\not\Rightarrow$ (i4)

Considérons la fonction f définie par :

$$\Omega = \{a,b,c,d\}; f(x,y) = \{x,y\} \text{ si } \{x,y\} \neq \{a,b\}; f(a,b) = f(b,a) = \Omega$$

Ici la fonction f vérifie trivialement (i1), (i2) ainsi que (i3) car :

$$f(a,c) \cap f(c,b) = \{c\} \text{ et } f(a,d) \cap f(d,b) = \{d\}$$

Mais f ne vérifie pas (i4), puisque $f(a,b) = \Omega \neq f(a,c) \cup f(c,b) = \{a,c,b\}$

PROPOSITION 8 : Si une application f de Ω^2 dans $\mathbb{P}(\Omega)$ vérifie (i1), (i2), (i3) et (i4), alors:

- | | | |
|-----|------------------------------------|--|
| (1) | $\forall a,b \in \Omega,$ | $\{a,b\} \subseteq f(a,b)$ |
| (2) | $\forall a \in \Omega,$ | $f(a,a) = \{a\}$ |
| (3) | la relation θ définie par : | " $a \theta b \Leftrightarrow a \in f(m,b)$ ", est un ordre total sur Ω . |
| (4) | $\forall a, b, c \in \Omega,$ | $c \in f(a,b)$ ou $b \in f(a,c)$ ou $a \in f(b,c)$ |
| (5) | $\forall a, b, c \in \Omega,$ | $a \in f(c,b)$ et $c \in f(a,b) \Rightarrow a = c$ |
| (6) | $\forall a, b, c, d \in \Omega,$ | $d \in f(c,b)$ et $c \in f(a,b) \Rightarrow c \in f(a,d)$ |

Preuve.

(1) Si a et b appartiennent à Ω , alors d'après (i2), il existe m et M dans Ω tels que : $\Omega = f(m,M)$, donc tels que $b \in f(m,M)$, d'où d'après (i4), $\Omega = f(m,b) \cup f(b,M)$.

Deux cas se présentent selon que $a \in f(m,b)$ ou $a \in f(b,M)$:

- si $a \in f(m,b)$, alors d'après (i3), $f(m,a) \cap f(a,b) = \{a\}$, donc $a \in f(a,b)$;
- si $a \in f(b,M)$, alors d'après (i3), $f(b,a) \cap f(a,M) = \{a\}$, donc $a \in f(b,a) = f(a,b)$.

Dans les deux cas $a \in f(a,b)$. Par symétrie, c'est-à-dire en utilisant (i1), $b \in f(a,b)$.

(2) D'après (1) $a \in f(a,a)$, donc d'après (i3) $f(a,a) \cap f(a,a) = \{a\}$, d'où $f(a,a) = \{a\}$.

(3) *Réflexivité* : conséquence de (1).

Anti-symétrie : soit a et b dans Ω tels que $a \theta b$ et $b \theta a$, i.e. $a \in f(m,b)$ et $b \in f(m,a)$. D'après (i3) et (i4) :

$$f(m,b) = f(m,a) \cup f(a,b) \quad \text{avec} \quad f(m,a) \cap f(a,b) = \{a\}$$

Comme $b \in f(m,a)$, $b \in f(m,a) \cap f(a,b)$, ce qui entraîne $a = b$.

Transitivité : Soit a,b,c trois points tels que : $a \theta b$ et $b \theta c$, i.e. $a \in f(m,b)$ et $b \in f(m,c)$.

D'après (i4), $f(m,c) = f(m,b) \cup f(b,c)$, donc $a \in f(m,c)$, c'est-à-dire $a \theta c$.

Relation d'ordre total : Soit a et b deux éléments distincts de Ω , montrons que a et b sont comparables selon l'ordre θ . Comme $b \in \Omega = f(m,M)$, nous avons d'après (i3) et (i4) :

$$\Omega = f(m,b) \cup f(b,M) \quad \text{avec} \quad f(m,b) \cap f(b,M) = \{b\}.$$

Par conséquent soit $a \in f(m,b)$, soit $a \in f(b,M)$.

Si $a \in f(m,b)$, alors $a \theta b$. Si $a \in f(b,M)$, il résulte:

$$f(b,M) = f(b,a) \cup f(a,M) \quad \text{avec} \quad f(b,a) \cap f(a,M) = \{a\}.$$

D'où un autre recouvrement de Ω :

avec

$$\begin{aligned} \Omega &= f(m,b) \cup f(b,a) \cup f(a,M) & (*) \\ f(m,b) \cap (f(b,a) \cup f(a,M)) &= \{b\} & (**) \\ f(b,a) \cap f(a,M) &= \{a\} & (***) \end{aligned}$$

Comme a est distinct de b , nous déduisons de (***) que b n'appartient pas à $f(a,M)$. D'où :

$$f(m,b) \cap f(a,M) = \emptyset$$

En désignant par A^c le complémentaire dans Ω d'une partie A de Ω , on déduit de (*) :

$$(\mathbf{f}(a,M) - \{a\})^c = \mathbf{f}(m,b) \cup \mathbf{f}(b,a).$$

D'autre part, comme $a \in \Omega = \mathbf{f}(m,M)$, nous déduisons :

$$\Omega = \mathbf{f}(m,a) \cup \mathbf{f}(a,M) \quad \text{avec} \quad \mathbf{f}(m,a) \cap \mathbf{f}(a,M) = \{a\}$$

$$\begin{aligned} \text{D'où :} & \quad (\mathbf{f}(a,M) - \{a\})^c = \mathbf{f}(m,a) \\ \text{Donc :} & \quad \mathbf{f}(m,a) = \mathbf{f}(m,b) \cup \mathbf{f}(b,a) \end{aligned}$$

Ce qui prouve que b appartient à $\mathbf{f}(m,a)$, i.e. $b \theta a$.

(4) D'après (3), les éléments a , b et c sont comparables deux à deux pour la relation θ .
Plaçons nous dans le cas particulier où : $a \theta b \theta c$. On a donc :

$$\begin{aligned} a &\in \mathbf{f}(m,b) & (a1) \\ b &\in \mathbf{f}(m,c) & (a2) \\ a &\in \mathbf{f}(m,c) & (a3) \end{aligned}$$

De (a3), nous déduisons :

$$\mathbf{f}(m,c) = \mathbf{f}(m,a) \cup \mathbf{f}(a,c) \quad \text{avec} \quad \mathbf{f}(m,a) \cap \mathbf{f}(a,c) = \{a\}$$

Donc en utilisant (a2), soit $b \in \mathbf{f}(a,c)$, soit $b \in \mathbf{f}(m,a)$.

Si $b \in \mathbf{f}(m,a)$, alors $a \theta b$ et $b \theta a$, donc $b = a$ qui appartient bien à $\mathbf{f}(a,c)$.

Donc l'hypothèse $a \theta b \theta c$ entraîne $b \in \mathbf{f}(a,c)$.

Les autres cas ($a \theta c \theta b$, $b \theta a \theta c$, $b \theta c \theta a$, etc...) s'obtiennent par permutation circulaire sur les lettres a,b,c , donc les conclusions seront :

$$c \in \mathbf{f}(a,b), \quad a \in \mathbf{f}(b,c) \quad \text{ou} \quad b \in \mathbf{f}(a,c).$$

(5) Soit $c \in \mathbf{f}(a,b)$ donc $\mathbf{f}(a,c) \cap \mathbf{f}(c,b) = \{c\}$; or $a \in \mathbf{f}(c,b)$, donc $a \in \mathbf{f}(a,c) \cap \mathbf{f}(c,b) = \{c\}$, c'est-à-dire $a = c$.

(6) D'après (4), trois cas sont possibles : $c \in \mathbf{f}(a,d)$, $a \in \mathbf{f}(c,d)$, $d \in \mathbf{f}(a,c)$. Il s'agit de montrer que les deux derniers cas entraînent le premier à savoir $c \in \mathbf{f}(a,d)$.

Si $d \in \mathbf{f}(a,c)$, alors $d \in \mathbf{f}(a,c) \cap \mathbf{f}(c,b)$ puisque $d \in \mathbf{f}(c,b)$; or on a $\mathbf{f}(a,c) \cap \mathbf{f}(c,b) = \{c\}$ car $c \in \mathbf{f}(a,b)$, d'où $d = c$, donc $c \in \mathbf{f}(a,d)$ est bien vérifié.

Si $a \in \mathbf{f}(c,d)$, alors $a \in \mathbf{f}(c,b) = \mathbf{f}(c,d) \cup \mathbf{f}(d,b)$ car $d \in \mathbf{f}(c,b)$. Par conséquent $a \in \mathbf{f}(c,b)$ et $c \in \mathbf{f}(a,b)$, donc d'après (5) $a = c$, donc $c \in \mathbf{f}(a,d)$ est bien vérifié.

PROPOSITION 9 : Si \mathbf{f} vérifie les conditions (i1), (i2), (i3) et (i4), alors il existe deux ordres définis sur Ω tels que \mathbf{f} soit la fonction d'intervalle de chacun de ces ordres. De plus, ces deux ordres sont inverses l'un de l'autre.

Preuve. D'après le lemme 1, il suffit de montrer qu'il existe au moins un ordre θ tel que \mathbf{f} soit la fonction d'intervalle de θ .

D'après la proposition 8-(3), la relation θ définie par :

$$\forall x,y \in \Omega, \quad x \theta y \Leftrightarrow x \in \mathbf{f}(m,y)$$

est une relation d'ordre total sur Ω . Montrons que f est la fonction d'intervalle de l'ordre θ . Soit x et y dans Ω tels que $x \theta y$; il s'agit de prouver que $f(x,y) = [x,y]$.
Soit $u \in [x,y]$, c'est-à-dire $x \theta u \theta y$. Par définition de θ , on a donc :

$$u \in f(m,y), \quad x \in f(m,u) \quad \text{et} \quad x \in f(m,y).$$

De $x \in f(m,y)$, nous déduisons :

$$f(m,y) = f(m,x) \cup f(x,y) \quad \text{avec} \quad f(m,x) \cap f(x,y) = \{x\}.$$

Puisque $u \in f(m,y)$, deux cas sont possibles : $u \in f(m,x)$ ou $u \in f(x,y)$. Si $u \in f(m,x)$, alors $u \theta x$ donc $x = u$ (car par hypothèse $x \theta u \theta y$). Donc dans les deux cas, $u \in f(x,y)$.

Réciproquement, soit $u \in f(x,y)$. Comme θ est un ordre total, x est comparable à y selon θ . Supposons par exemple, $x \theta y$, c'est-à-dire $x \in f(m,y)$. De plus, $u \in f(x,y)$, donc d'après la proposition 8 : $x \in f(m,u)$, c'est-à-dire $x \theta u$.

Enfin $x \in f(m,y)$ implique $f(m,y) = f(m,x) \cup f(x,y)$, donc $u \in f(m,y)$, c'est-à-dire $u \theta y$, et par conséquent $u \in [x,y]$.

On montre de la même façon que dans le cas $y \theta x$, on a $u \in [x,y]$. Par conséquent dans les deux cas $x \theta y$ et $y \theta x$, l'élément u appartient à $[x,y]$. Il y a donc bien identité entre $f(x,y)$ et $[x,y]$.

Des propositions 1 et 9, nous déduisons :

PROPOSITION 10 : Si f est une fonction de Ω^2 dans $\mathbb{P}(\Omega)$, alors f est la fonction d'intervalle d'un ordre total borné si et seulement si f vérifie les conditions (i1), (i2), (i3) et (i4) définies par :

- | | | |
|------|---|-------------------------------|
| (i1) | $\forall a,b \in \Omega,$ | $f(a,b) = f(b,a)$ |
| (i2) | $\exists m \text{ et } M \in \Omega \text{ tq :}$ | $\Omega = f(m,M)$ |
| (i3) | $\forall c \in f(a,b),$ | $f(a,c) \cap f(c,b) = \{c\}$ |
| (i4) | $\forall c \in f(a,b),$ | $f(a,c) \cup f(c,b) = f(a,b)$ |

D'après ce qui précède, nous pouvons énoncer le résultat suivant :

THÉORÈME 2 : Si δ est propre et bornée, alors δ est robinsonienne sans triplet ultramétrique si et seulement si pour tout a et b dans Ω et tout c dans $L(a,b)$, on a :

$$L(a,b) = L(a,c) \cup L(c,b) \quad \text{et} \quad L(a,c) \cap L(c,b) = \{c\}.$$

Illustration géométrique

Les trois points a , b et c étant placés selon la figure 10, on a indiqué les ensembles $L(a,b)$, $L(a,c)$ et $L(c,b)$ (qui coïncident ici avec les segments $S(a,b)$, $S(a,c)$ et $S(c,b)$). La propriété (iii) indique où placer un quatrième point pour que la distance euclidienne entre ces quatre points soit robinsonienne et sans triplet ultramétrique : les zones interdites pour ce quatrième point sont les zones hachurées.

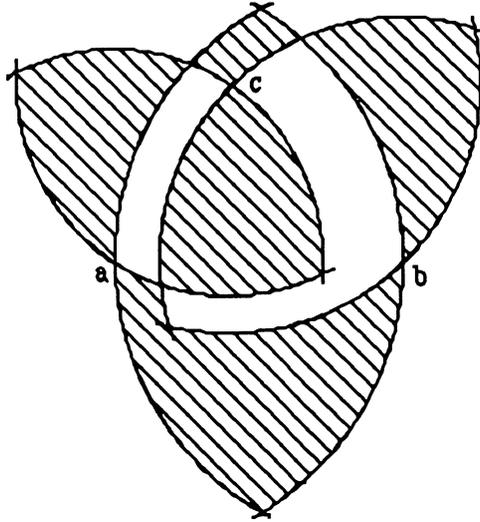


Figure 10

3.1.3. Caractérisation des dissimilarités robinsoniennes sans triplet ultramétrique dans le cas où Ω est fini

Dans le cas où Ω est fini, nous donnons une caractérisation qui est liée à l'existence d'une bijection entre pyramides indicées au sens large et dissimilarités robinsoniennes.

REMARQUE 7 : soit \mathbf{P} une pyramide et θ un ordre compatible avec \mathbf{P} . Notons \mathbf{I}_θ la fonction d'intervalle de θ . L'axiome (p3) de la définition d'une pyramide (cf. §1) peut alors s'écrire sous la forme : $\mathbf{P} \subseteq \mathbf{I}_\theta(\Omega^2)$.

Si l'on suppose θ borné, alors Ω appartient à $\mathbf{I}_\theta(\Omega^2)$, donc $\mathbf{I}_\theta(\Omega^2)$ vérifie l'axiome (p1). De façon évidente, $\mathbf{I}_\theta(\Omega^2)$ satisfait les axiomes (p2) et (p3), donc $\mathbf{I}_\theta(\Omega^2)$ est une pyramide.

Ainsi toute pyramide \mathbf{P} ayant θ pour ordre compatible, est incluse dans la pyramide $\mathbf{I}_\theta(\Omega^2)$.

Dans le cas où Ω est fini, tout ordre total sur Ω est borné, et par conséquent si un ordre θ est compatible avec une pyramide \mathbf{P} , alors cette pyramide \mathbf{P} est incluse dans la pyramide $\mathbf{I}_\theta(\Omega^2)$.

Diday (1984) a introduit la notion de pyramide *saturée* : une pyramide \mathbf{P} sur Ω est dite *saturée* si le nombre de ses paliers est maximum, c'est-à-dire si toutes les pyramides sur Ω ont un cardinal inférieur ou égal au cardinal de \mathbf{P} .

PROPOSITION 11 : Soit Ω un ensemble fini de cardinal égal à n , et \mathbf{P} un ensemble de parties de Ω . Les propriétés suivantes sont équivalentes :

- (i) \mathbf{P} est une pyramide saturée,
- (ii) \mathbf{P} est une pyramide et $\text{card}(\mathbf{P}) = n(n+1)/2$,
- (iii) il existe un ordre total sur Ω , noté θ , tel que $\mathbf{P} = \mathbf{I}_\theta(\Omega^2)$.

Preuve. Remarquons d'abord que si θ est un ordre total défini sur Ω , alors :

$$\begin{aligned} \text{card}(\mathbf{I}_\theta(\Omega^2)) &= \text{nombre d'intervalles non vides de } \theta \\ &= \text{nombre de parties à 1 ou 2 éléments de } \Omega = n(n+1)/2. \end{aligned}$$

D'après ce qui précède ou simplement par définition, les affirmations suivantes sont vraies :

- (1) si θ est compatible avec la pyramide \mathbf{P} , alors $\mathbf{P} \subseteq \mathbf{I}_\theta(\Omega^2)$,
- (2) $\text{card}(\mathbf{I}_\theta(\Omega^2)) = n(n+1)/2$,
- (3) $\mathbf{I}_\theta(\Omega^2)$ est une pyramide.

De ces trois affirmations (1), (2) et (3), on déduit facilement l'équivalence des propriétés (ii) et (iii), puis de (i) et (iii).

Pour établir la proposition 12, nous utilisons le lemme suivant :

LEMME 2 : Soient X et Y deux ensembles finis, f et g deux applications définies sur X et à valeurs dans $\mathbb{P}(Y)$. Si f et g vérifient les trois conditions suivantes :

- (a) $\forall x, f(x) \subseteq g(x)$,
- (b) $f(X) = g(X)$,
- (c) f et g sont injectives,

alors $f = g$.

Preuve. Raisonnons par l'absurde : supposons que l'ensemble E des éléments x de X tels que $f(x) \neq g(x)$ ne soit pas vide. Considérons un élément u de cet ensemble tel que la partie $g(u)$ de Y soit maximale (pour la relation d'inclusion) parmi toutes les parties de la forme $g(x)$ avec $x \in X$.

D'après (a) : $f(u) \subseteq g(u)$,

cette inclusion étant stricte, car $u \in E$. Comme $g(X) = f(X)$, il existe un élément $v \neq u$ tel que $g(u) = f(v)$.

Donc, d'après (a) : $g(u) = f(v) \subseteq g(v)$.

Or cette inclusion est nécessairement stricte car g est injective, ce qui contredit la définition de l'élément u . Donc $f = g$.

PROPOSITION 12 : Si Ω est fini, alors δ est robinsonienne et sans triplet ultramétrique si et seulement si $L(\Omega^2)$ est une pyramide saturée.

Preuve. Supposons que δ soit robinsonienne et sans triplet ultramétrique, ce qui revient à supposer, puisque Ω est fini, que L est une fonction d'intervalle. Dans ce cas, il existe un ordre θ tel que $L = I_\theta$, donc tel que $L(\Omega^2) = I_\theta(\Omega^2)$. Donc $L(\Omega^2)$ est une pyramide saturée, d'après la proposition 11.

Réciproquement, supposons que $L(\Omega^2)$ soit une pyramide saturée, la proposition 11 montre qu'il existe un ordre total sur Ω , noté θ , tel que $L(\Omega^2) = I_\theta(\Omega^2)$. Considérons alors les deux fonctions f et g , à valeurs dans $Y = \mathbb{P}(\Omega^2)$, et définies sur l'ensemble X des paires d'éléments de Ω , par : $\forall \{x,y\} \in X, f(\{x,y\}) = I_\theta(x,y)$ et $g(\{x,y\}) = L(x,y)$.

Les fonctions L et I_θ étant symétriques, les valeurs de f et g sont bien définies.

Comme $f(X) = L(\Omega^2)$ et $g(X) = I_\theta(\Omega^2)$, on en déduit que $f(X) = g(X)$. De $L(\Omega^2) = I_\theta(\Omega^2)$, il résulte que tout ensemble de la forme $L(x,y)$ est un intervalle de θ , donc pour tout x,y :

$$I_\theta(x,y) \subseteq L(x,y),$$

et par suite : $f(\{x,y\}) \subseteq g(\{x,y\})$.

Enfin remarquons que $\text{card}(f(X)) = \text{card}(I_\theta(\Omega^2)) = n(n+1)/2 = \text{card}(X)$, d'où :

$$\text{card}(f(X)) = \text{card}(g(X)) = \text{card}(X),$$

ce qui prouve l'injectivité de f et de g . Donc, d'après le lemme 2, $f = g$, d'où $L = I_\theta$ car L et I_θ sont symétriques. Donc δ est robinsonienne et sans triplet ultramétrique (cf. proposition 7).

3.2. Caractérisation des dissimilarités robinsoniennes quelconques à l'aide de la fonction \bar{L}

Il s'agit ici de caractériser les dissimilarités robinsoniennes quelconques, en généralisant la condition énoncée dans la proposition 12. La caractérisation que nous présentons

(cf. théorème 3) est différente et plus simple que celle proposée par C. Durand (cf. C. Durand (1989), p.57 à 60). En effet, nous montrons que pour que δ soit robinsonienne, il faut et il suffit que $\bar{L}(\Omega^2)$ soit une pyramide, alors que la condition énoncée par C. Durand est, en tenant compte de nos notations : " $(\bar{L}(\Omega^2), \text{diam})$ est une pyramide indicée au sens large".

LEMME 3 : Si Ω est fini, si δ est propre et si θ est δ -compatible, alors pour tout w, x, y et z dans Ω :

$$\bar{L}(x,y) = I_\theta(z,w) \Rightarrow \bar{L}(x,y) = \bar{L}(z,w).$$

Preuve. D'après la proposition 4-(ii), nous savons que $\bar{L}(z,w)$ est un intervalle de θ , donc $I_\theta(z,w) \subseteq \bar{L}(z,w)$, et comme nous supposons $\bar{L}(x,y) = I_\theta(z,w)$, il en résulte :

$$\bar{L}(x,y) \subseteq \bar{L}(z,w).$$

De $\bar{L}(x,y) = I_\theta(z,w)$, nous déduisons que $\bar{L}(x,y)$ est une intersection d'éléments de $L(\Omega^2)$ qui contiennent chacun z et w . Par conséquent, $\bar{L}(z,w) \subseteq \bar{L}(x,y)$, ce qui prouve l'égalité de ces deux ensembles.

THÉORÈME 3 : Si Ω est fini et δ est une dissimilarité propre sur Ω , alors δ est robinsonienne si et seulement si $\bar{L}(\Omega^2)$ est une pyramide.

Preuve. Montrons que la condition " $\bar{L}(\Omega^2)$ est une pyramide" est nécessaire. Par définition de \bar{L} , le singleton $\{x\}$ est inclus dans $\bar{L}(x,x)$ qui lui-même est inclus dans $L(x,x)$. Or $L(x,x)$ est réduit à $\{x\}$ car la dissimilarité δ est supposée propre. Donc $\{x\} = \bar{L}(x,x)$, ce qui prouve que chaque singleton appartient à $\bar{L}(\Omega^2)$.

- Notons θ un ordre tel que la dissimilarité robinsonienne δ soit θ -compatible. L'ensemble Ω est fini, donc il existe un plus petit et un plus grand élément de Ω selon l'ordre total θ , c'est-à-dire il existe deux éléments m et M tels que : $[m,M] = \Omega$.

Or :

$$[m,M] \subseteq \bar{L}(m, M) \subseteq \Omega$$

Ce qui prouve que Ω est égal à $\bar{L}(m, M)$, et donc appartient à $\bar{L}(\Omega^2)$.

- Soient p, q, r et s des points quelconques de Ω , montrons que $A = \bar{L}(p,q) \cap \bar{L}(r,s)$ est soit vide, soit un élément de $\bar{L}(\Omega^2)$.

D'après la proposition 4-(ii) et le lemme 3, il existe quatre points t, u, v, w de Ω qui vérifient :

$$\begin{aligned} \bar{L}(p,q) &= [t,u] = \bar{L}(t,u) && \text{avec } t \theta u \\ \bar{L}(r,s) &= [v,w] = \bar{L}(v,w) && \text{avec } v \theta w. \end{aligned}$$

Donc $A = [t,u] \cap [v,w]$. Comme les positions de t, u, v , et w respectent les relations $t \theta u$ et $v \theta w$, six cas sont possibles : en effet, il suffit de choisir les rangs de t et de u parmi les quatre rangs possibles 1, 2, 3, 4 pour déterminer les positions relatives de t, u, v et w . Ces cas sont :

- CAS 1 : $t \theta u \theta v \theta w$
- CAS 2 : $t \theta v \theta u \theta w$
- CAS 3 : $v \theta t \theta u \theta w$
- CAS 4 : $t \theta v \theta w \theta u$
- CAS 5 : $v \theta t \theta w \theta u$
- CAS 6 : $v \theta w \theta t \theta u$

Examinons chacun de ces cas.

CAS 1 : si u est distinct de v , alors A est vide. Si $u = v$, alors $A = \{u\} = \{v\}$ qui appartient à $\bar{L}(\Omega^2)$.

CAS 6 : analogue au cas 1 car il suffit d'échanger t avec v et u avec w , pour retrouver la même situation.

CAS 3 : $A = [t,u] = \bar{L}(t,u) \in \bar{L}(\Omega^2)$.

CAS 4 : $A = [v,w] = \bar{L}(v,w) \in \bar{L}(\Omega^2)$.

CAS 2 et CAS 5 : ce sont des cas symétriques l'un de l'autre ; en effet, ces deux cas se déduisent l'un de l'autre en échangeant les rôles de t et v d'une part, et de u et w d'autre part.

Traisons le cas 2.

Dans le cas 2, A est égal à $[v,u]$. La partie $\bar{L}(u,v)$ étant un intervalle de θ , $A = [v,u]$ est inclus dans $\bar{L}(u,v)$. Inversement, en tant qu'intersection de $\bar{L}(p,q)$ et de $\bar{L}(r,s)$, A apparaît comme l'intersection d'éléments de $\bar{L}(\Omega^2)$ qui contiennent chacun u et v puisque $A = [v,u]$. Par conséquent, A contient nécessairement $\bar{L}(u,v)$ qui est l'intersection de *tous* les éléments de $L(\Omega^2)$ qui contiennent u et v . Donc finalement $A = \bar{L}(u,v)$.

Ainsi dans tous les cas, l'intersection de deux éléments de $\bar{L}(\Omega^2)$ est soit l'ensemble vide, soit un élément de $\bar{L}(\Omega^2)$.

- D'après la proposition 4, il existe au moins un ordre total sur Ω tel que chaque élément de $\bar{L}(\Omega^2)$ soit un intervalle de cet ordre.

Ce qui permet de conclure, d'après ce qui précède que $\bar{L}(\Omega^2)$ est une pyramide.

Réciproquement, montrons que la condition " $\bar{L}(\Omega^2)$ est une pyramide" est suffisante. Si $\bar{L}(\Omega^2)$ est une pyramide, alors il existe au moins un ordre θ total défini sur Ω tel que chaque élément de $\bar{L}(\Omega^2)$ soit un intervalle de θ , par conséquent :

$$\forall x,y \in \Omega, \quad I_{\theta}(x,y) \subseteq \bar{L}(x,y)$$

Or $\bar{L}(x,y)$ est inclus dans $L(x,y)$, donc $I_{\theta}(x,y)$ est inclus dans $L(x,y)$, ce qui entraîne que δ est robinsonienne, car δ est supposée propre.

BIBLIOGRAPHIE

ALTWEGG M., Zur Axiomatik der teilweise geordneten Mengen, *Comment. Math. Helv.*, 24 (1950), 149-155.

BATBEDAT A., *Les approches pyramidales dans la classification arborée*, Paris, Masson, 1990.

BERTRAND P., *Étude de la représentation pyramidale*, Thèse de 3^o cycle, Université Paris IX-Dauphine, 1986.

BERTRAND P., DIDAY E., A visual representation of the compatibility between an order and a dissimilarity index : the pyramids, *Computational Statistics Quarterly* 2, (1985), 31-44.

CRITCHLEY F., VAN CUTSEM B., *Predissimilarities, prefilters and ultrametrics on an arbitrary set*, Rapport de Recherche, IMAG, Université Grenoble I, 1989.

DIDAY E., *Une représentation visuelle des classes empiétantes : les pyramides*, Rapport I.N.R.I.A. 291, Rocquencourt, 1984, paru aussi dans R.A.I.R.O. APII, 20, (1986), 5, 475-526.

DURAND C., *Ordres et graphes pseudo-hiérarchiques : théorie et optimisation algorithmique*, Thèse, Université de Provence, 1989.

DURAND C., FICHET B., One-to-one correspondences in pyramidal representation : a unified approach, *Classification and Related Methods of Data Analysis*, H.H. Bock (ed.), Amsterdam, North-Holland, 1988, 85-90.

GUENOCHÉ A., Méthodes combinatoires de sériation à partir d'une dissimilarité, *Data Analysis and Informatics 5*, E. Diday et al. (eds.), Amsterdam, North-Holland, 1987.

HUBERT L., Some applications of graph theory to clustering, *Psychometrika*, 39, (1974), 283-309.

JAIN A.K., DUBES R.C., *Algorithms for Clustering Data*, Prentice Hall Advanced Reference Series, Englewood Cliffs, New Jersey, 1988.

KENDALL D.G., Incidence Matrices, Interval Graphs and Seriation in Archaeology, *Pacific Journal of Mathematics*, 28, (1969), 3, 565-570.

LECLERC B., Arbres minimums communs et compatibilité de données de types variés, *Mathématiques et Sciences humaines*, 98, (1987), 41-67.

MULDER H.M., *The interval function of a graph*, Mathematical Centre tracts 132, Amsterdam, 1980.

SHOLANDER M., Trees, Lattices, Order, and Betweenness, *Proceedings of the American Mathematical Society*, 3 (1952), 369-381.

ROBINSON W.S., A method for chronological ordering of archaeological deposits, *American Antiquity*, 16 (1951), 4, 293-301.