UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

THESE

présentée à

L'UNIVERSITÉ DES SCIENCES ET TECHNIQUES DE LILLE

pour obtenir

LE TITRE DE DOCTEUR DE 3ème CYCLE

(Informatique)

par

Jean-Claude VAN DORPE

INFORMATION GENERALISEE ET VALUATION DANS LES TREILLIS

APPLICATION A UN PROBLEME DE TAXINOMIE NUMERIQUE

Thèse soutenue le 14 décembre 1976, devant la Commission d'Examen

MEMBRES DU JURY	M. P. POUZET,	Président
	M. C.F. PICARD,	Rapporteur
	MM. G. JACOB	
	J. LOSFELD,	Examinateurs

Je tiens à exprimer mes remerciements à Monsieur le Professeur P. POUZET, qui a bien voulu me faire l'honneur de présider le Jury de cette Thèse.

Que Monsieur C.F. PICARD, Directeur de Recherche au C.N.R.S., rapporteur de ce travail, trouve ici l'expression de ma gratitude pour l'intérêt qu'il a porté à mes recherches, les conseils et remarques qu'il m'a prodigués.

Je remercie également Monsieur le Professeur G. JACOB d'avoir accepté de participer au Jury et de m'avoir permis, par ses remarques pertinentes, d'améliorer le texte que je lui avais confié.

Que J. LOSFELD trouve ici l'expression de ma profonde reconnaissance pour les conseils, les suggestions et l'aide constante qu'il m'a apportés.

Monsieur le Professeur J. KAMPE DE FERIET a toujours suivi avec intérêt l'évolution de mes travaux, son aide et son intérêt me furent très précieux dans l'élaboration de cette Thèse; qu'il trouve ici l'expression de ma très sincère reconnaissance.

Je remercie également mon collègue Gérard COMYN pour sa précieuse collaboration.

Je tiens enfin à remercier Madame DESCARPENTRIES, du Secrétariat du Département Informatique de l'I.U.T. de LILLE, pour le soin constant qu'elle a apporté dans la préparation matérielle de cette thèse et des travaux que j'ai pu faire ces dernières années.

TABLE DES MATIERES

INTRODUCTION

Dans un treillis T quelconque, muni d'une mesure d'information généralisée J, nous définissons des séquences d'éléments comparables. Il en découle naturellement les notions de perte et gain d'information, notions auxquelles on associe des quantités notées δ^+, δ^- dont on déduit une distance en information δ ainsi qu'une application notée J_R qui généralise à un treillis T quelconque la notion d'information mutuelle. L'interprétation de ces grandeurs nous amène à définir les notions d'information "parasite" et d'information "non expliquée".

L'introduction d'hypothèses particulières, dites de type H, sur les mesures utilisées, permet d'obtenir des expressions simples des différentes quantités introduites précédemment.

Dès le début des travaux sur le théorie généralisée de l'information, deux directions axiomatiques sont essentiellement apparues :

- L'étude de la mesure d'information, notée $I(A)$, fournie par la réalisation d'un événement A ; l'espace de travail est (Ω,S), où S est une classe de parties de Ω (souvent il s'agit de l'algèbre de Boole des parties de Ω); cette axiomatique a été développée essentiellement par J. KAMPE DE FERIET et B. FORTE [17].

- L'étude de la mesure d'information, notée $H(\Pi)$ fournie par une expérience (ou partition Π) ; l'espace de référence est alors le treillis des partitions d'un ensemble Ω donné. Cette axiomatique a fait l'objet de nombreux travaux (en particulier de la part de B. FORTE et N. PINTACUDA [8]). C'est dans ce contexte que sont le plus souvent utilisées les informations habituelles intervenant en Théorie des Questionnaires. [23]

Nous avons été amené , aux chapitres II et III, à étudier l'application des résultats du chapitre I dans le cadre des deux treillis particuliers cités ci-dessus. Nous définissons entre autres des distances entre événements ou entre partitions qui ont été utilisées pour la réalisation d'algorithmes en taxinomie numérique, notamment pour la détermination de "classes homogènes" d'individus et la construction de questionnaires sur ces classes [15], ainsi que pour la recherche de "bonnes synthèses" de tableaux de données réelles (chapitre V).

L'ensemble des résultats concernant les quantités δ^+ , δ^- et δ établis au chapitre I est une conséquence directe de l'axiome de monotonicité de J . Ceci nous a amenés à étudier les propriétés intrinsèques des applications monotones définies sur T et à valeurs dans $\mathbb{R} \cup \{+\infty\}$. Nous définissons alors une quasi-métrique sur T dont nous déduisons une métrique lorsque l'application v est strictement monotone. L'utilisation d'hypothèses sur v analogues à celles définies au chapitre I nous permet de donner des expressions simples de la métrique. Nous retrouvons notamment la distance habituelle sur les treillis modulaires lorsque v est une valuation strictement croissante. Nous montrons que les hypothèses H1 et H2 n'entraînent aucune structure particulière sur T mais que si le treillis est gradué, nous réobtenons la caractérisation classique de la semi-modularité [4].

Nous proposons, dans la deuxième partie du chapitre IV, une extension aux cas des <u>demi-treillis</u>

- des hypothèses H1 et H2
- de la notion de valuation
- de la caractérisation de la semi-modularité dans les treillis gradués énoncée par G. BIRKHOF [4].

C H A P I T R E   I

-=-=-=-=-=-=-=-=-=-

TREILLIS ET INFORMATION GENERALISEE

-=-=-=-=-=-=-=-=-=-

I.- TREILLIS ET INFORMATION GENERALISEE

Sur un treillis T muni d'une information généralisée J , ne nous intéressant qu'aux propriétés intrinsèques de l'axiome de monotonicité, nous définissons les notions d'informations absorbée, transmise et apportée par une séquence d'éléments comparables du treillis. Des notions introduites résulte naturellement la définition d'un écart sur T dont nous déduisons une métrique lorsque J est strictement monotone.

Cette démarche nous conduit à définir une application J_R de $T \times T$ dans $\mathbb{R}$ possédant les propriétés essentielles de l'information mutuelle de SHANNON définie sur le treillis des partitions d'un ensemble Ω .

L'introduction d'hypothèses particulières sur J nous permet d'obtenir des formules explicitant la métrique et l'application J_R .

Enfin, l'interprétation des quantités ainsi définies nous permet de mesurer le degré de ressemblance entre deux éléments quelconques d'un treillis.

Ce chapitre résulte d'un travail effectué en collaboration avec Georges GRIMONPREZ.

I.1.- RAPPELS

Nous rappelons, dans ce paragraphe, les définitions d'informations généralisées introduites par J. LOSFELD [20].

I.1.1.- INFORMATION SUR UN TREILLIS.

Soit T un treillis dont la relation d'ordre est notée $\leqslant$; une *mesure d'information généralisée* définie sur T est :

AXIOME 1.- une application J de T dans $\bar{\mathbb{R}}^+$

AXIOME 2.- monotone (nous la supposons décroissante)

$$\forall\ (x,y) \in T \times T \quad , \quad x \leqslant y \Longrightarrow J(x) \geq J(y)$$

I.1.2.- INFORMATIONS CONDITIONNELLE ET MUTUELLE.

. L'information *conditionnelle*, fournie par l'élément $y \in T$, connaissant l'information fournie par $x \in T$, est définie par :

$$J(x/y) = J(x \wedge y) - J(y) \quad \text{si} \quad J(y) < +\infty$$

$$J(x/y) = J(x) \qquad \text{si} \quad J(y) = +\infty$$

Notons que, pour tout $y \in T$ fixé, $J(./y)$ est une information généralisée, ce qui justifie la terminologie employée. Dans tout ce qui suit, cette définition servira uniquement de notation, commode pour les valeurs infinies.

. L'information *mutuelle* entre deux éléments x et y de T est définie par :

$$J(x;y) = J(x) + J(y) - J(x \wedge y) \quad \text{si} \quad J(x) < +\infty \quad \text{et} \quad J(y) < +\infty$$

$$J(x;y) = 0 \quad \text{sinon}$$

Notons qu'en général, l'information mutuelle peut prendre des valeurs négatives et qu'elle n'a aucune propriété de monotonicité. La terminologie employée, consacrée par l'usage, peut donc prêter à confusion.

I.2.- PERTE ET GAIN D'INFORMATION DANS UN TREILLIS.

Soit $(T,\leqslant,J)$ un *treillis* T muni d'une *information généralisée* J. Sur l'exemple de la figure 1, les lettres représentent les éléments de T, les nombres entre parenthèses représentent les valeurs de J associées aux éléments de T.

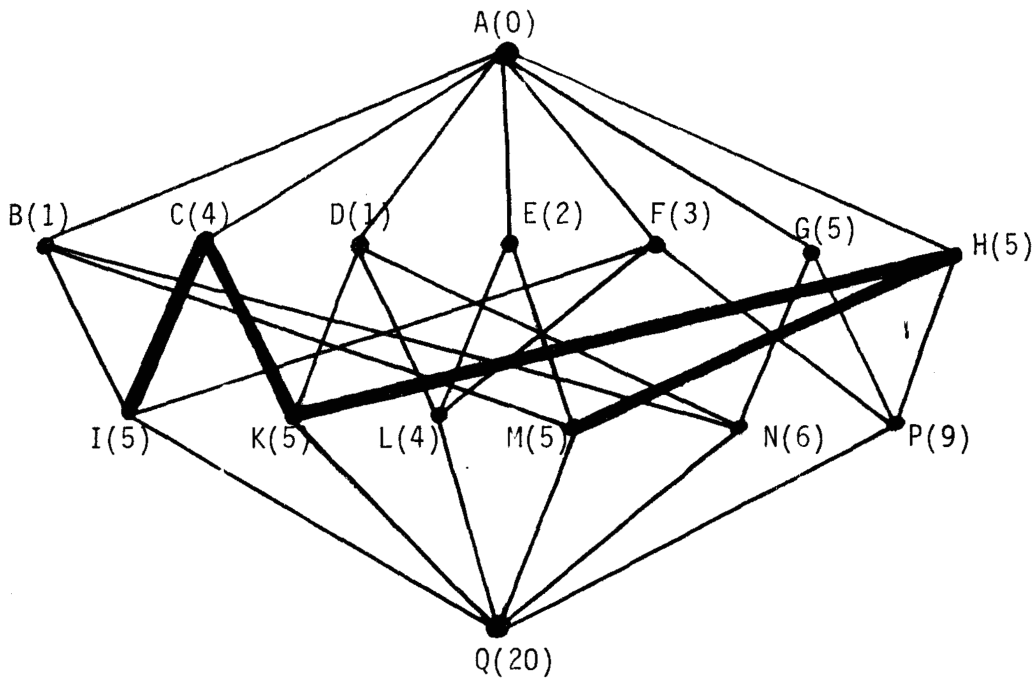


FIGURE 1

Considérons le treillis T comme représentant l'ensemble des états possibles d'un Système; l'information du système à l'état x est égale à la quantité $J(x)$. La suite (I, C, K, H, M) définit une séquence d'éléments comparables deux à deux de T ; cette suite formalise une évolution du système qui passe de l'état I à l'état M par une suite d'états intermédiaires.

I.2.1.- DEFINITION :

Une *séquence d'éléments comparables, de longueur* $n \in \mathbb{N}$, de T (en abrégé s.e.c.) est une séquence notée $[x_i]_0^n$ de $n+1$ éléments de T telle que, pour tout indice $i = 0,1,\dots,n-1$, on ait :

$$x_i \leqslant x_{i+1} \qquad \text{ou} \qquad x_i \geqslant x_{i+1}$$

Si la propriété de transitivité est vérifiée pour tous les éléments d'une s.e.c., alors, conformément à la terminologie utilisée en Théorie des Treillis [4] , nous appelons *chaîne* une s.e.c. $[x_i]_0^n$ telle que :

$$\forall\ i = 0,\dots,n-1 : x_i \lessdot x_{i+1}$$

ou

$$\forall\ i = 0,\dots,n-1 : x_i \gtrdot x_{i+1}$$

Lorsque, dans la s.e.c. $[x_i]_0^n$, nous avons $x_i \lessdot x_{i+1}$, l'information du système diminue. Nous considérons alors que le système perd, c'est-à-dire que la s.e.c. *absorbe* une certaine quantité d'information égale à $J(x_i) - J(x_{i+1}) = J(x_i/x_{i+1})$.

Inversement, si $x_i \gtrdot x_{i+1}$, l'information du système augmente. Nous considérons alors que le système gagne une certaine quantité d'information égale à $J(x_{i+1}) - J(x_i) = J(x_{i+1}/x_i)$, c'est-à-dire que la s.e.c. *apporte* de l'information au système.

Le passage de l'état x_0 à l'état x_n s'accompagne d'une variation globale de l'information du système égale à $J(x_n) - J(x_0)$, variation ne dépendant que des extrémités de la s.e.c.

Soit I l'ensemble des indices d'une s.e.c. $[x_i]_0^n$ donnée :

Notons $I^+ = \{i \in I | x_{i+1} \lessdot x_i\}$

$$I^- = \{i \in I | x_I \lessdot x_{i+1}\} .$$

Nous associons à une s.e.c. $[x_i]_0^n$ les quantités suivantes :

I.2.2.- DEFINITIONS.

. L'information *gagnée* par le système, c'est-à-dire *apportée* par la s.e.c. :

$$\Delta^+([x_I]_0^n) = \sum_{i \in I^+} J(x_{i+1}/x_i).$$

. L'information *perdue* par le système, c'est-à-dire *absorbée* par la s.e.c. :

$$\Delta^-([x_i]_0^n) = \sum_{i \in I^-} J(x_i/x_{i+1}) \ .$$

. La quantité

$$\Delta([x_i]_0^n) = \sum_{i=0}^{n-1} |J(x_{i+1}) - J(x_i)| = \Delta^+([x_i]_0^n) + \Delta^-([x_i]_0^n)$$

Δ^+ , Δ^- et Δ sont des quantités finies ou non, positives ou nulles. Si Δ^+ et Δ^- ne sont pas simultanément des valeurs infinies, nous définissons :

. La *variation d'information* du système, c'est-à-dire l'information *transmise* par la s.e.c., par :

$$V([x_i]_0^n) = J(x_n) - J(x_0) = \Delta^+([x_i]_0^n) - \Delta^-([x_i]_0^n) \ .$$

Cette quantité, finie ou non, peut être positive, négative ou nulle.

EXEMPLE : Les quantités d'information associées à la s.e.c. $C = (I, C, K, H, M)$ de la figure 1 , sont égales à :

$$V(C) = 0 \quad , \quad \Delta^+(C) = 1 \quad , \quad \Delta^-(C) = 1 \quad , \quad \Delta(C) = 2 \ .$$

Comme nous le montrons par ailleurs [10], la terminologie ainsi introduite (informations absorbée, apportée, transmise), généralise celle utilisée en Théorie des Questionnaires [23] .

I.2.3.- ETUDE DES PROPRIETES DE Δ^+ , Δ^- ET Δ.

Pour tout couple fixé (x,y) de $T \times T$, notons $\mathcal{C}(x,y)$ l'ensemble de toutes les s.e.c. reliant x à y .

Etant donnés deux éléments x et y de T , il existe une bijection entre $\mathcal{C}(x,y)$ et $\mathcal{C}(y,x)$:

A toute s.e.c. $C = (x_0 = x \ , x_1,\ldots,x_n = y)$ correspond une s.e.c. $\bar{C}$ appelée s.e.c. *inverse* de C reliant y à x :

$$\bar{C} = (y_0 = y \ , y_1 = x_{n-1},\ldots,y_n = x_0 = x),$$

il vient $\Delta^+(C) = \Delta^-(\bar{C})$, $\Delta^-(C) = \Delta^+(\bar{C})$, $\Delta(C) = \Delta(\bar{C})$

I.2.3.1.- PROPOSITION

Etant donnés deux éléments x *et* y *de* T *tels que* $J(x) \,.\, J(y) < +\infty$, C_1 *et* C_2 *deux s.e.c. reliant* x *à* y *, nous avons les équivalences :*

$$\Delta(C_1) \leq \Delta(C_2) \iff \Delta^+(C_1) \leq \Delta^+(C_2) \iff \Delta^-(C_1) \leq \Delta^-(C_2)$$

DEMONSTRATION.- Par hypothèse, $J(x)$ et $J(y)$ prennent des valeurs finies, Δ^+ , Δ^- et Δ prennent alors simultanément des valeurs finies ou infinies. Le cas infini étant trivial, considérons que Δ^+ , Δ^- et Δ prennent des valeurs finies :

Supposons que $\Delta(C_1) \leq \Delta(C_2)$

par définition $\Delta(C_1) = \Delta^+(C_1) + \Delta^-(C_1)$

$$\Delta(C_2) = \Delta^+(C_2) + \Delta^-(C_2)$$

On en déduit :

$$V(C_1) = \Delta^+(C_1) - \Delta^-(C_1) = J(y) - J(x) = \Delta^+(C_2) - \Delta^-(C_2) = V(C_2)$$

$$\left.\begin{array}{l}\Delta(C_1) \leq \Delta(C_2)\\ \\ V(C_1) = V(C_2)\end{array}\right\} \Longrightarrow \begin{array}{l}\Delta^+(C_1) + \Delta^-(C_1) + \Delta^+(C_1) - \Delta^-(C_1) \leq \ldots\\ \ldots \leq \Delta^+(C_2) + \Delta^-(C_2) + \Delta^+(C_2) - \Delta^-(C_2)\end{array}$$

soit $\Delta^+(C_1) \leq \Delta^+(C_2)$.

On montre, de manière analogue :

$$\Delta^+(C_1) \leq \Delta^+(C_2) \Longrightarrow \Delta^-(C_1) \leq \Delta^-(C_2) \Longrightarrow \Delta(C_1) \leq \Delta(C_2)$$

I.2.3.2.- PROPOSITION.

Soient x *et* y *deux éléments comparables de* T*,* C *une chaîne quelconque reliant* x *à* y *, alors :*

(i) Si $x \lesssim y$: $\Delta^+(C) = 0$, $\Delta^-(C) = \Delta(C) = J(x/y)$

(ii) Si $x \gtrsim y$: $\Delta^-(C) = 0$, $\Delta^+(C) = \Delta(C) = J(y/x)$

La démonstration de cette proposition est immédiate.

Nous définissons :

. $\delta^+(x,y) = \inf_{C \in \mathcal{C}(x,y)} \Delta^+(C)$

. $\delta^-(x,y) = \inf_{C \in \mathcal{C}(x,y)} \Delta^-(C)$

. $\delta(x,y) = \inf_{C \in \mathcal{C}(x,y)} \Delta(C)$

δ^+ , δ^- et δ sont des quantités finies ou non, positives ou nulles.

$\delta^+(x,y)$ représente la quantité minimale d'information que doit gagner le système pour passer de l'état x à l'état y .

$\delta^-(x,y)$ représente la quantité minimale d'information que doit perdre le système pour passer de l'état x à l'état y .

Des relations de définition de Δ , Δ^+ et Δ^- on tire :

$$\Delta([x_i]_0^n) = \Delta^+([x_i]_0^n) + \Delta^-([x_i]_0^n)$$

$$J(x_n) - J(x_0) = \Delta^+([x_i]_0^n) - \Delta^-([x_i]_0^n) \;;$$

et de la proposition I.2.3.1., il vient immédiatement :

I.2.3.3.- THEOREME.

(i) $\delta(x,y) = \delta^+(x,y) + \delta^-(x,y)$

(ii) $J(y) - J(x) = \delta^+(x,y) - \delta^-(x,y)$.

La relation (ii) de ce théorème nous autorise à définir des applications J_R et J_T de $T \times T$ dans $\bar{\mathbb{R}}^+$ par :

. $J_T(x,y) = J(x) + \delta^+(x,y) = J(y) + \delta^-(x,y)$ pour tout couple (x,y) de $T \times T$.

. $J_R(x,y) = J(x) - \delta^-(x,y) = J(y) - \delta^+(x,y)$ si $J(x) \,.\, J(y) < +\infty$

nous précisons, au paragraphe I-4, la définition pour les cas limites (valeur infinie pour $J(x)$ ou $J(y)$).

Nous donnons par la suite (paragraphe I-6), une interprétation de $J_T(x,y)$ et $J_R(x,y)$ ainsi que des quantités $\delta^+(x,y)$ et $\delta^-(x,y)$ introduites précédemment.

I.2.4.- REMARQUES.

Nous précisons, dans ce paragraphe, quelques résultats élémentaires concernant l'existence de s.e.c. particulières; enfin, nous étudions les cas limites liés aux valeurs infinies par les définitions des quantités de type δ et Δ .

. Etant donnée une s.e.c. $[x_i]_0^n \in C(x,y)$, il est toujours possible de construire une s.e.c. $[y_i]_0^m \in C(x,y)$ à partir de $[x_i]_0^n$ telle que :

- la s.e.c. $[y_i]_0^m$ soit *sans cycle*, c'est-à-dire :

$$\forall\ (i,j) \in \{0,1,\dots,m\} \times \{0,1,\dots,m\} \text{ l'on ait } y_i \neq y_j$$

- la s.e.c. $[y_i]_0^m$ soit *alternée*, c'est-à-dire :

$$\forall\ i = 0,\dots,m-2 : \quad (y_i \lneq y_{i+1} \Longrightarrow y_{i+1} \gneq y_{i+2}) \quad \text{ou}$$

$$(y_i \gneq y_{i+1} \Longrightarrow y_{i+1} \lneq y_{i+2}) \ .$$

La construction de la s.e.c. alternée s'effectue par utilisation de la transitivité de la relation d'ordre.

- alors, nous obtenons :

$$\Delta([y_i]_0^m) \leq \Delta([x_i]_0^n)\ ,\ \Delta^+([y_i]_0^m) \leq \Delta^+([x_i]_0^n)\ ,$$

$$\Delta^-([y_i]_0^m) \leq \Delta^-([x_i]_0^n)\ .$$

. Si le treillis T est *fini*, il existe donc au moins une s.e.c. C_0 de $C(x,y)$ telle que :

$$\delta(x,y) = \operatorname*{Inf}_{C \in C(x,y)} \Delta(C) = \Delta(C_0)$$

. Si $J(x) \,.\, J(y) < +\infty$ alors :

- Δ^+ , Δ^- et Δ prennent simultanément des valeurs finies ou infinies.

- δ^+ , δ^- et δ ont des valeurs finies.

. Si $J(x) = +\infty$ et $J(y) < +\infty$ alors :

- $\Delta^- = \Delta = +\infty$, $\Delta^+ \leq +\infty$

- $\delta^- = \delta = +\infty$, $\delta^+ < +\infty$

. Si $J(x) < +\infty$ et $J(y) < +\infty$ alors :

- $\Delta^+ = \Delta = +\infty$, $\Delta^- \leq +\infty$

- $\delta^+ = \delta = +\infty$, $\delta^- < +\infty$

. Si $J(x) = J(y) = +\infty$ alors :

- $\Delta^+ = \Delta^- = \Delta + \infty$

- $\delta^+ = \delta^- = \delta = +\infty$

I.3.- ÉTUDE DES PROPRIÉTÉS DE δ^+ , δ^- ET δ .

Les propriétés de δ étudiées dans ce paragraphe permettent d'étendre la notion de distance à un treillis quelconque.

I.3.1.- PROPOSITION.

L'application δ *de* $T \times T$ *dans* $\bar{\mathbb{R}}^+$ *définie par* $\delta(x,y) = \inf\limits_{C \in C(x,y)} \Delta(C)$ *est un écart sur* T , *c'est-à-dire une application de* $T \times T$ *dans* $\bar{\mathbb{R}}^+$ *vérifiant :*

(i) $\delta(x,x) = 0$ *pour tout* $(x,x) \in T \times T$

(ii) $\delta(x,y) = \delta(y,x)$ *pour tout* $(x,y) \in T \times T$

(iii) $\delta(x,y) \leq \delta(x,z) + \delta(z,y)$ *pour tout* $(x,y,z) \in T \times T \times T$ *(inégalité triangulaire)*

de plus

(iv) : s'il existe un minorant universel m *et un majorant*

universel M *dans* T *, alors :*

$$\forall\ (x,y) \in T \times T \ : \ \delta(x,y) \leq \delta(m,M)$$

DEMONSTRATION.

(i) $\delta(x,x) = 0$. Il suffit de considérer la s.e.c. de longueur 0 formée de l'élément x .

(ii) $\delta(x,y) = \delta(y,x)$. Cette égalité résulte immédiatement de la remarque suivant la définition de la s.e.c. inverse (paragraphe I-2-3)

(iii) $\delta(x,y) \leq \delta(x,z) + \delta(z,y)$.

Soient $[x_i]_0^n$ une s.e.c. quelconque reliant x à z , $[y_i]_0^m$ une s.e.c. quelconque reliant z à y .

La s.e.c. $C = (x_0,\ldots,x_n = y_0,\ldots,y_m)$ relie x à y et l'on a :

$$\delta(x,y) \leq \Delta(C)$$

Il est clair que $\Delta(C) = \Delta([x_i]_0^n) + \Delta([y_i]_0^m)$

Nous avons donc $\delta(x,y) \leq \Delta([x_i]_0^n) + \Delta([y_i]_0^m)$

Cette inégalité étant vérifiée par toute s.e.c. $[x_i]_0^n$ reliant x à z et toute s.e.c. $[y_i]_0^m$ reliant z à y , nous en déduisons :

$$\delta(x,y) \leq \delta(x,z) + \delta(z,y)$$

(iv) . Si $J(m) = +\infty$, alors $\delta(m,M) = J(m) - J(M) = +\infty$.

. Si $J(m) < +\infty$,

$$\delta(x,y) \leq \delta(x,m) + \delta(m,y) = 2\ J(m) - J(x) - J(y)$$

$$\delta(x,y) \leq \delta(x,M) + \delta(M,y) = J(x) + J(y) - 2\ J(M)$$

ce qui donne

$$2\ \delta(x,y) \leq 2\ J(m) - 2\ J(M) = 2\ \delta(m,M).$$

I.3.2.- PROPOSITION.

Les applications δ^+ *et* δ^- *de* $T \times T$ *dans* $\bar{\mathbb{R}}^+$ *vérifient :*

(i) $\delta^+(x,x) = \delta^-(x,x) = 0$ *pour tout* $x \in T$

(ii) $\delta^+(x,y) = \delta^-(y,x)$ *pour tout* $(x,y) \in T \times T$

(iii) δ^+ *et* δ^- *vérifient l'inégalité triangulaire.*

La démonstration de cette proposition est analogue à celle de la proposition précédente.

I.3.3.- PROPOSITION.

Etant donné un treillis T *fini*, $T_0 = \{x \in T | J(x) < +\infty\}$, *une condition nécessaire et suffisante pour que la restriction de l'application* δ *à l'ensemble des couples* (x,y) *de* $T_0 \times T_0$ *définisse une distance est que l'application* J *soit strictement monotone.*

DEMONSTRATION.- En raison de la proposition I.3.1., il suffit de montrer :

$$J \text{ strictement monotone} \Longleftrightarrow (\delta(x,y) = 0 \Longrightarrow x = y)$$

Démontrons la négation de cette équivalence :

- Hypothèse : $\delta(x,y) = 0 \Longrightarrow x = y$

  Supposons J non strictement monotone :

  $\exists\ (x,y) \in T \times T \ , \ x \prec y \Longrightarrow J(x) = J(y)$

  soit C la s.e.c. de longueur 1 définie par (x,y) :

  $\Delta(C) = J(x/y) = J(x) - J(y) = 0$

  $\Longrightarrow \delta(x,y) = 0 \Longrightarrow x = y$

- Hypothèse : J strictement monotone.

  Supposons qu'il existe un couple $(x,y) \in T \times T$ tel que :

$$\delta(x,y) = 0 \text{ et } x \neq y$$

$$\left.\begin{array}{l}\delta(x,y) = 0 \\ \\ T \text{ fini (rem. I.2.4.)}\end{array}\right\} \Longrightarrow \exists\ [x_i]_0^n \in C(x,y) \text{ telle que}$$

$$\Delta([x_i]_0^n) = 0$$

d'oũ $J(x_0) = J(x_n) = J(x_i) \quad \forall\ i = 1,\dots,n-1$

ce qui est impossible car J est strictement monotone et $x \neq y$ ·
On en déduit :

$$x \neq y \ : \ \forall\ C \in C(x,y)\ ,\ \Delta(C) > 0$$

T étant supposé fini, $\delta(x,y)$ est strictement positif, c.q.f.d.

I.4.- PROPRIÉTÉS DES APPLICATIONS J_R ET J_T.

Rappelons les définitions des applications J_R et J_T de $T \times T$ dans $\bar{\mathbb{R}}^+$:

. $J_T(x,y) = J(x) + \delta^+(x,y) = J(y) + \delta^-(x,y)$ pour tout couple (x,y) de $T \times T$.

. $J_R(x,y) = J(x) - \delta^-(x,y) = J(y) - \delta^+(x,y)$ si $J(x) \,.\, J(y) < +\infty$

$= J(x) - \delta^-(x,y)$ si $J(y) = +\infty$ et $J(x) < +\infty$

$= J(y) - \delta^+(x,y)$ si $J(x) = +\infty$ et $J(y) < +\infty$

$J_R(x,y)$ n'est pas définie si $J(x) = +\infty$ et $J(y) = +\infty$

Remarquons que $J_R(x,y)$ prend toujours une valeur finie, alors que $J_T(x,y) = +\infty$ si et seulement si $J(x) \,.\, J(y) = +\infty$.

I.4.1.- PROPOSITION.

Les applications J_R *et* J_T *sont symétriques en* x *et* y *:*

$J_R(x,y) = J_R(y,x)$, $J_T(x,y) = J_T(y,x)$

et vérifient l'égalité :

$J(x) + J(y) = J_R(x,y) + J_T(x,y)$.

La symétrie résulte de l'égalité $\delta^+(x,y) = \delta^-(y,x)$ (prop. I.3.2, (ii)). L'égalité se déduit immédiatement du théorème I.2.3.3. (ii).

I.4.2.- THEOREME.

Les applications J_R *et* J_T *quand elles sont définies vérifient :*

(i) $$0 \leq J(x \vee y) \leq J_R(x,y) \leq \left\{\begin{matrix} J(x) \\ \\ J(y) \end{matrix}\right\} \leq J_T(x,y) \leq J(x \wedge y)$$

(ii) pour tout y *fixé de* T *tel que* $J(y) < +\infty$, *les applications* $J_R(.,y)$ *et* $J_T(.,y)$ *sont décroissantes.*

(iii) Si l'application J *est strictement décroissante, alors*

- $J_R(x,y) = J(y) \Longleftrightarrow x \preccurlyeq y$
- $J_T(x,y) = J(y) \Longleftrightarrow x \succcurlyeq y$

DEMONSTRATION.-

(i) a) $J_T(x,y) \leq J(x \wedge y)$

Si $J(x) = +\infty$ ou $J(y) = +\infty$ alors $J(x \wedge y) = +\infty$

$J_T(x,y) = +\infty$, donc $J_T(x,y) = J(x \wedge y)$

Si $J(x) . J(y) < +\infty$, soit la s.e.c. $C = (x, x \wedge y, y)$

$\delta^+(x,y) \leq \Delta^+(C) = J(x \wedge y) - J(x)$

$\Longrightarrow J_T(x,y) = J(x) + \delta^+(x,y) \leq J(x \wedge y)$

b) $\left.\begin{array}{l} J(x) \\ \\ J(y) \end{array}\right\} \leq J_T(x,y)$

Si $J(x) = +\infty$ ou $J(y) = +\infty$ alors $J_T(x,y) = +\infty$
d'où l'inégalité (b) .

Si $J(x) . J(y) < +\infty$, l'inégalité (b) résulte du fait que $\delta^+(x,y)$ et $\delta^-(x,y)$ sont des quantités toujours ≥ 0 .

c) $J_R(x,y) \leq \begin{cases} J(x) \\ J(y) \end{cases}$

Cette inégalité se déduit immédiatement de la définition de J_R

d) $J(x \vee y) \leq J_R(x,y)$

Si $J(x) < +\infty$ et $J(y) = +\infty$ alors $J(x \vee y) < +\infty$

$J_R(x,y) = J(x) - \delta^-(x,y)$.

Soit la s.e.c. $C = (x, x \vee y, y)$

$\Delta^-(C) = J(x) - J(x \vee y)$

$\delta^-(x,y) \leq \Delta^-(C) \Longrightarrow J_R(x,y) \geq J(x) - \Delta^-(C) = J(x \vee y)$.

La démonstration est analogue lorsque $J(x) = +\infty$ et $J(y) < +\infty$ et lorsque $J(x)\, J(y) < +\infty$.

(ii) Soient deux éléments x et x' de T tels que $x \lesssim x'$

Supposons que $J(x) = J(x') = +\infty$

alors $J_R(x,y) = J(y) - \delta^+(x,y)$

$J_R(x',y) = J(y) - \delta^+(x',y)$

$\delta^+(x,y) \leq \delta^+(x,x') + \delta^+(x',y)$ (proposition I.3.2. (iii))

$\delta^+(x,x) = 0$ (conséquence immédiate de la proposition I.2.3.2)

d'où

$\delta^+(x,y) \leq \delta^+(x',y) \Longrightarrow J_R(x,y) \geq J_R(x',y)$

On montre de manière identique que J_R est décroissante lorsque

$J(x) = +\infty$ et $J(x') < +\infty$, $J(x) < +\infty$ et $J(x') = +\infty$

et enfin lorsque $J(x)\,J(x') < +\infty$

La démonstration pour l'application J_T est analogue à celle de J_R .

(iii) Supposons $x \lessapprox y$.

. Si $J(x)\,J(y) < +\infty$ alors $\delta^+(x,y) = 0 \Longrightarrow J_R(x,y) = J(y)$

. Si $J(x) = +\infty$ et $J(y) < +\infty$ alors

$$J_R(x,y) = J(y) - \delta^+(x,y) = J(y)$$

. Si $J(x) < +\infty$ et $J(y) = +\infty$ alors

$$J_R(x,y) = J(x) - \delta^-(x,y)$$

$$x \lessapprox y \Longrightarrow \delta^-(x,y) = J(x) - J(y) \Longrightarrow J_R(x,y) = J(y)$$

Supposons maintenant que x et y soient tels que $J_R(x,y) = J(y)$.
$J_R(x,y)$ étant toujours fini, nous avons $J(y) < +\infty$. On en déduit :

$$J_R(x,y) = J(y) = J(x) - \delta^+(x,y) \Longrightarrow \delta^-(x,y) = 0$$

J étant strictement décroissante, on obtient $x \lessapprox y$.

La démonstration est analogue pour J_T .

<u>I.4.3.- PROPOSITION.</u>

Les applications J_R *et* J_T *vérifient :*

$J_R(x,y) \geq J(x) + J(y) - J(x \wedge y)$ *si* $J(x)\,J(y) < +\infty$

$J_T(x,y) \leq J(x) + J(y) - J(x \vee y)$ *pour tout couple* (x,y) *de* $T \times T$.

<u>DEMONSTRATION.</u>

. $J_R(x,y) = J(x) - \delta^-(x,y)$

soit la s.e.c. $C = (x, x \wedge y, y)$: $\Delta^-(C) = J(x \wedge y) - J(y)$

$$\delta^-(x,y) \leq \Delta^-(C) \Longrightarrow J_R(x,y) \geq J(x) - \Delta^-(C) \text{ , soit :}$$

$$J_R(x,y) \geq J(x) + J(y) - J(x \wedge y) \text{ .}$$

. La démonstration est analogue pour J_T .

I.5.- EXPRESSION DES QUANTITÉS δ , J_R ET J_T .

I.5.1.- INTRODUCTION.

Nous proposons, dans ce paragraphe, des résultats permettant de fournir des formulations aisément calculables des quantités δ , J_R et J_T ...

Il résulte de la proposition I.2.3.1 qu'étudier les s.e.c. de $\mathcal{C}(x,y)$ minimisant les quantités Δ^+ , Δ^- ou Δ , revient à faire cette étude pour l'une d'entre elles, par exemple Δ^- . Nous disons alors qu'une *s.e.c. est J-minimale si et seulement si elle minimise l'une quelconque des quantités* Δ^+ *,* Δ^- *ou* Δ .

La première remarque du paragraphe I.2.4 nous permet, pour la suite, de ne plus considérer que des s.e.c. sans cycle et alternées.

Lorsque le treillis T *est fini,* il existe au moins une s.e.c. J-minimale entre deux éléments x et y donnés de T , ce qui n'est pas le cas si T est quelconque. D'autre part, l'ensemble des s.e.c. sans cycle et alternées reliant x et y est fini. Il est alors possible de mettre en oeuvre un algorithme du type Branch et Bound [9] permettant d'obtenir toutes les s.e.c. J-minimales reliant x à y , ce qui donne la possibilité de calculer les quantités $\delta(x,y)$, $J_R(x,y)$ et $J_T(x,y)$.

Cet algorithme n'offre cependant qu'un intérêt relatif, étant donné qu'il nécessite des temps d'exécution importants, aussi avons-nous été amenés [11] , *lorsque le treillis* T *est quelconque (fini ou infini),* à introduire des hypothèses particulières sur J permettant d'obtenir des expressions explicites des quantités δ , J_R et J_T .

I.5.2.- HYPOTHESES H

Il est clair que la valeur minimale de la quantité $\Delta^-(C)$, pour toutes les s.e.c. de longueur 2 reliant x à y , deux éléments donnés de T , est égale à :

(1) $\mathrm{Inf}(J(x/y), J(x/x \vee y))$, ce qui s'écrit, dans le cas fini :

(2) $\mathrm{Inf}(J(x \wedge y) - J(y) , J(x) - J(x \vee y))$

c'est-à-dire qu'il existe une s.e.c. J-minimale de longueur 2 de la forme $(x, x \vee y, y)$ notée $C_\vee(x,y)$, ou de la forme $(x, x \wedge y, y)$ notée $C_\wedge(x,y)$, ce qui nous amène à énoncer le théorème suivant :

I.5.2.1.- THEOREME.

Une condition nécessaire et suffisante pour que, pour tout couple (x,y) *de* $T \times T$ *, la s.e.c.* $C_\wedge(x,y) = (x, x \wedge y, y)$ *soit* J*-minimale est que* J *vérifie l'hypothèse* H1 *suivante :*

(H1) $\forall (x,y) \in T \times T \quad , \quad J(x) + J(y) \geq J(x \wedge y) + J(x \vee y)$

DEMONSTRATION.

. Condition nécessaire

$\forall\ (x,y) \in T \times T \ , \ \forall\ C \in \mathcal{C}(x,y) : \Delta^-(C_\wedge(x,y)) \leq \Delta^-(C)$

soit $\quad C = (x, x \vee y, y) : \Delta^-(C) = J(x) - J(x \vee y)$

$\Delta^-(C_\wedge(x,y)) = J(x \wedge y) - J(y)$

d'où $\quad J(x \wedge y) - J(y) \leq J(x) - J(x \vee y) \qquad$ c.q.f.d.

. Condition suffisante

Si x est égal à y le problème est résolu.
Supposons x différent de y :
La démonstration s'effectue par récurrence sur la longueur de la s.e.c. reliant x à y .
Il est facile de montrer que $C_\wedge(x,y)$ est J-minimale pour l'ensemble des s.e.c. de longueur 1 et 2 reliant x à y .

Soit C une s.e.c. de longueur 3 reliant x à y :

$$C = (x,z,t,y) \quad \text{avec} \quad (x \lesssim z \ , \ z \gtrsim t \ , \ t \lesssim y)$$

$$\text{ou} \quad (x \gtrsim z \ , \ z \lesssim t \ , \ t \gtrsim y)$$

Supposons que $(x \lesssim z \ , \ z \gtrsim t \ , \ t \lesssim y)$

$$\left.\begin{array}{l} t \lesssim y \Longrightarrow x \wedge t \lesssim x \wedge y \\ J \text{ décroissante} \end{array}\right\} \Longrightarrow J(x \wedge t) \geq J(x \wedge y)$$

$$\Delta^{-}(C_{\wedge}(x,y)) = J(x \wedge y) - J(y) \leq J(x \wedge t) - J(y)$$

$$\Delta^{-}(C_{\wedge}(x,y)) \leq J(x \wedge t) - J(t) + J(t) - J(y)$$

HYPOTHESE H1 $\Longrightarrow J(x \wedge t) - J(t) \leq J(x) - J(x \vee t)$

$$\Delta^{-}(C_{\wedge}(x,y)) \leq J(x) - J(x \vee t) + J(t) - J(y)$$

$$z \gtrsim x \vee t \Longrightarrow J(z) \leq J(x \vee t)$$

ce qui donne :

$$\Delta^{-}(C_{\wedge}(x,y)) \leq J(x) - J(z) + J(t) - J(y) = \Delta^{-}(C)$$

On montre, de la même manière, que si C est de la forme :

$$(x \gtrsim z \ , \ z \lesssim t \ , \ t \gtrsim y) \quad \text{alors} \quad \Delta^{-}(C_{\wedge}(x,y) \leq \Delta^{-}(C)$$

Ainsi, la s.e.c. $C_{\wedge}(x,y)$ est J-minimale pour toute s.e.c. de longueur 3.
Supposons, maintenant, que, pour tout couple (x,y) de $T \times T$, la s.e.c. $C_{\wedge}(x,y)$ soit J-minimale pour toute s.e.c. de longueur $n-1$ reliant x à y. Soit $[x_i]_0^n$ une s.e.c. de longueur n reliant x à y.
L'hypothèse de récurrence entraîne :

$$\Delta^-(C_\wedge(x,x_{n-1})) \leq \Delta^-([x_i]_0^{n-1})$$

$$\Delta^-([x_i]_0^n) = \Delta^-([x_i]_0^{n-1}) + \Delta^-(C_1)$$

où C_1 est définie par la séquence (x_{n-1},x_n) , d'où :

$$\Delta^-([x_i]_0^n) \geq \Delta^-(C_\wedge(x,x_{n-1})) + \Delta^-(C_1) = \Delta^-(C_2)$$

où C_2 est définie par la séquence $(x,x \wedge x_{n-1},x_{n-1},y)$.

Le résultat sur les s.e.c. de longueur 3 entraîne :

$$\Delta^-(C_2) \geq \Delta^-(C_\wedge(x,y))$$

ce qui donne :

$$\Delta^-(C_\wedge(x,y)) \leq \Delta^-([x_i]_0^n) \qquad \text{c.q.f.d.}$$

On démontre de la même manière les résultats suivants :

I.5.2.2.- THEOREME.

Une condition nécessaire et suffisante pour que, pour tout couple, (x,y) *de* $T \times T$ *, la s.e.c.* $C_\vee(x,y) = (x,x \vee y,y)$ *soit* J*-minimale est que* J *vérifie l'hypothèse H2 :*

(H2) $\quad \forall\ (x,y) \in T \times T \quad , \quad J(x) + J(y) \leq J(x \wedge y) + J(x \vee y)$

I.5.2.3.- COROLLAIRE

Une condition nécessaire et suffisante pour que, pour tout couple (x,y) *de* $T \times T$ *, les s.e.c.* $C_\wedge(x,y)$ *et* $C_\vee(x,y)$ *soient* J*-minimales est que* J *vérifie l'hypothèse HO :*

(HO) $\quad \forall\ (x,y) \in T \times T \quad , \quad J(x) + J(y) = J(x \wedge y) + J(x \vee y)$.

Ces résultats permettent de généraliser immédiatement la proposition I.3.3.

I.5.2.4.- PROPOSITION.

Etant donné un treillis T *quelconque, si* J *vérifie l'une quelconque des hypothèses H0 , H1, ou H2, une condition nécessaire et suffisante pour que la restriction de* δ *à l'ensemble des couples* (x,y) *de* $T_0 \times T_0$ *définisse une distance est que l'application* J *soit strictement monotone.*

Ces résultats nous permettent également de donner une expression explicite des quantités δ , J_R et J_T . Nous donnons ces expressions dans le tableau I.5.2.5. Le tableau I.5.2.6. donne les formes particulières obtenues lorsque les éléments x et y de T sont tels que $J(x) < +\infty$ et $J(y) < +\infty$.

	H1	H2
$\delta^+(x,y)$	$J(y/x)$	$J(y/x \vee y)$
$\delta^-(x,y)$	$J(x/y)$	$J(x/x \vee y)$
$\delta(x,y)$	$J(y/x) + J(x/y)$	$J(y/x \vee y) + J(x/x \vee y)$
$J_R(x,y)$	$J(x) + J(y) - J(x \wedge y)$	$J(x \vee y)$
$J_T(x,y)$	$J(y) + J(y/x)$ $=$ $J(x) + J(x/y)$	$J(x) + J(y/x \vee y)$ $=$ $J(y) + J(x/x \vee y)$

TABLEAU I.5.2.5. (cas général)

.../...

TABLEAU I.5.2.6. (Lorsque $J(x)\,J(y) < +\infty$)

	H1	H2	H0
$\delta^+(x,y)$	$J(x \wedge y) - J(x)$	$J(y) - J(x \vee y)$	$J(x \wedge y) - J(x)$ = $J(y) - J(x \vee y)$
$\delta^-(x,y)$	$J(x \wedge y) - J(y)$	$J(x) - J(x \vee y)$	$J(x \wedge y) - J(y)$ = $J(x) - J(x \vee y)$
$\delta(x,y)$	$2J(x \wedge y) - J(x) - J(y)$	$J(x) + J(y) - 2J(x \vee y)$	$J(x \wedge y) - J(x \vee y)$
$J_R(x,y)$	$J(x) + J(y) - J(x \wedge y)$	$J(x \vee y)$	$J(x) + J(y) - J(x \wedge y)$ = $J(x \vee y)$
$J_T(x,y)$	$J(x \wedge y)$	$J(x) + J(y) - J(x \vee y)$	$J(x) + J(y) - J(x \vee y)$ = $J(x \wedge y)$

Pour l'hypothèse H0, nous pouvons prendre indifféremment les expressions de ces quantités relatives à la condition H1 ou à la condition H2.

REMARQUES : L'information J définie sur la figure 1 (page I-4), ne vérifie ni H1, ni H2.

En effet : $J(I) + J(K) < J(I \vee K) + J(I \wedge K)$

$$J(C) + J(H) > J(C \vee H) + J(C \wedge H)$$

Dans cet exemple, la s.e.c. J-minimale reliant I à M est la s.e.c. I C K H M de longueur 4 soulignée en gras sur la figure.

I.6.- APPLICATION : RÉSUMÉS ET INFORMATIONS [13]

I.6.1.- Notons H l'information de SHANNON sur les partitions finies Π d'un espace probabilisé (Ω,S,P) définie par :

$$H(\Pi) = \sum_{A\in\Pi} P(A) \text{ Log } 1/P(A)$$

Nous savons [20] que l'information mutuelle $H(\Pi;\Pi')$ apportée par les partitions Π et Π' est définie par :

$$H(\Pi;\Pi') = \sum_{(A,B)\in\Pi\times\Pi'} P(A \cap B) \text{ Log } \frac{P(A \cap B)}{P(A)P(B)}$$

et possède les propriétés :

$\forall\, \Pi$, Π' et Π'' partitions de Ω :

. $H(\Pi;\Pi') \geq 0$

. $\Pi' \leq \Pi'' \Longrightarrow H(\Pi;\Pi') \geq H(\Pi;\Pi'')$

. $H(\Pi;\Pi') \leq H(\Pi)$ avec égalité si et seulement si $\Pi' \leq \Pi$.

La partition Π étant supposée fixée, l'interprétation de l'application $H(.;\Pi)$ est la suivante : $H(.;\Pi)$ est une information généralisée mesurant ce que la connaissance d'une partition Π' apporte sur Π . En effet : $H(\Pi;\Pi')$ est une quantité positive, inférieure à $H(\Pi)$ mais d'autant plus grande que la partition Π' se "rapproche" de la partition Π.

On a montré (chap. II) que l'information de SHANNON sur les partitions vérifie l'hypothèse H1 et donc $H_R(.,\Pi)$ est identique à $H(.,\Pi)$ pour cet exemple.

I.6.2.- De manière générale, soit H une mesure d'information généralisée quelconque définie sur le treillis des partitions d'un ensemble fini, non vide. Le Théorème I.4.2. montre que l'application $H_R(.,\Pi)$ possède les propriétés intéressantes de l'information mutuelle. Cette application, égale à $H(.;\Pi)$ lorsque H vérifie l'hypothèse H1 se présente donc bien comme une extension de l'information mutuelle.

Considérons une partition Π comme ensemble de référence; lorsque l'on cherche à expliquer Π par Π' ; c'est-à-dire à considérer Π' comme un résumé (modèle) de Π alors :

. Nous interprétons $H_R(\Pi,\Pi')$ comme la quantité d'information en commun entre Π et Π' et $H_T(\Pi,\Pi')$ comme la quantité totale d'information apportée par Π et Π' .

. $\delta^-(\Pi,\Pi') = H(\Pi) - H_R(\Pi,\Pi') = H_T(\Pi,\Pi') - H(\Pi')$

définit la quantité d'information apportée par Π et non expliquée par Π' .

. $\delta^+(\Pi,\Pi') = H(\Pi') - H_R(\Pi,\Pi') = H_T(\Pi,\Pi') - H(\Pi)$

définit la quantité d'information "parasite" apportée par Π' et ne servant pas à la connaissance de Π.

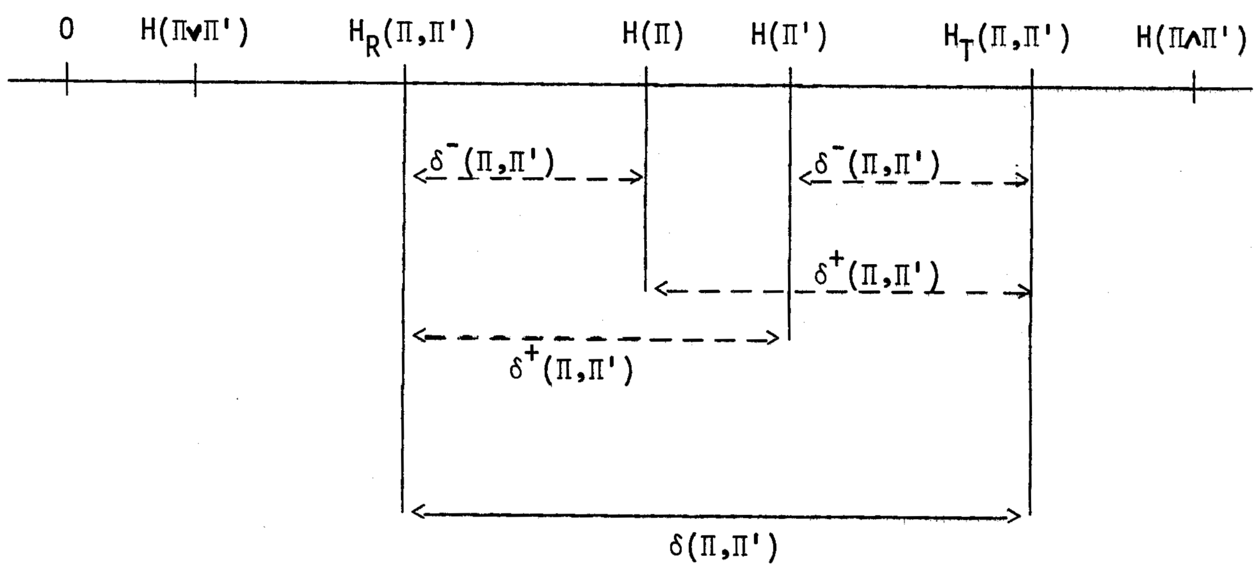


FIGURE I.6.1.

Certains algorithmes utilisés en analyse des données [3] (par exemple les méthodes de classification hiérarchique), se ramènent à la donnée d'une partition Π et à la construction d'une séquence $(\Pi_i | i = 1,\dots,)$ de partitions de plus en plus fines, cherchant à approcher Π de mieux en mieux.

Les inégalités $\Pi_1 \succcurlyeq \Pi_2 \succcurlyeq \cdots \succcurlyeq \Pi_i \succcurlyeq \Pi_{i+1} \succcurlyeq \cdots$ impliquent :

$$H_R(\Pi,\Pi_1) \leq \ldots \leq H_R(\Pi,\Pi_i) \leq H_R(\Pi,\Pi_{i+1}) \leq \ldots$$

$$H_T(\Pi,\Pi_1) \leq \ldots \leq H_T(\Pi,\Pi_i) \leq H_T(\Pi,\Pi_{i+1}) \leq \ldots$$

$$\delta^-(\Pi,\Pi_1) \geq \ldots \geq \delta^-(\Pi,\Pi_i) \geq \delta^-(\Pi,\Pi_{i+1}) \geq \ldots$$

$$\delta^+(\Pi,\Pi_2) \leq \ldots \leq \delta^+(\Pi,\Pi_i) \leq \delta^+(\Pi,\Pi_{i+1}) \leq \ldots$$

ce qui formalise le fait que la suite de partitions $\{\Pi_i\}$ approche de mieux en mieux l'information de Π , mais elle introduit aussi de plus en plus d'information parasite (inutile dans la connaissance de Π).

Nous remarquons que $\delta(\Pi,\Pi') = H_T(\Pi,\Pi') - H_R(\Pi,\Pi')$ *; la fonction* $\delta(.,\Pi)$ *n'est pas monotone en général. Les minima locaux de* δ *représentent alors de "bons" résumés de* Π *dans le sens où ils correspondent à des positions d'équilibre concernant les variations respectives de* H_R *et* H_T .

Les remarques et interprétations précédentes ont été présentées dans le cadre du treillis des partitions d'un ensemble Ω et sont utilisées dans le chapitre V pour les applications (algorithme RESUM2) . Rappelons que les résultats et propriétés fondamentaux permettant ces interprétations ont été établis ici dans le cadre général d'un treillis T quelconque muni d'une information généralisée J :

Soient $(x,y) \in T \times T$.
$J(x)$ *(respectivement* $J(y)$ *) représente la quantité d'information apportée par la connaissance de la réalisation de l'état* x *(resp.* y*) du système :*

- $J_R(x,y)$ *représente la quantité d'information commune à* x *et* y *;*

- $\delta^-(x,y) = J(x) - J_R(x,y)$ *la quantité d'information spécifique à* x *(dans le contexte* y*)*

. $\delta^+(x,y) = J(y) - J_R(x,y)$ *la quantité d'information spécifique à* y *(dans le contexte* x*)* .

Ou encore, si x *est l'état du système, si* y *est l'explication qu'on en donne (résumé ou modèle).*

. $J_T(x,y)$ *est la quantité totale d'information apportée par le résumé* y, *comprenant, en particulier, une certaine quantité inutile, dans la connaissance de* x .

. $J_R(x,y)$ *est la quantité d'information qu'apporte le résumé* y *sur la connaissance de* x .

. $\delta^-(x,y)$ *la quantité d'information de* x *non "expliquée" par le résumé* y .

. $\delta^+(x,y)$ *une quantité d'information "parasite" apportée par le résumé* y *sans rapport avec l'information de* x .

C H A P I T R E II

-=-=-=-=-=-=-=-=-=-=

INFORMATION GENERALISEE SUR LES PARTITIONS

-=-

II.- INFORMATION GENERALISEE SUR LES PARTITIONS.

Dans le paragraphe II-1, nous rappelons les définitions et propriétés de l'information généralisée sur les partitions.

Nous étudions ensuite, de manière systématique, les principales mesures d'information sur les partitions. Nous montrons que parmi celles-ci, l'information de SHANNON :

$$H_N^1(\Pi) = \sum_{A_i \in \Pi} P(A_i) \operatorname{Log} \frac{1}{P(A_i)}$$

vérifie l'hypothèse H1 , que l'information :

$$H_0(\Pi) = \operatorname{Card}(\Pi) - 1$$

vérifie l'hypothèse H2, et que les autres mesures d'information habituellement utilisées ne vérifient ni H1 ni H2 .

Nous donnons, quand c'est possible, une formulation simple des quantités δ^+ , δ^- , δ , H_R et H_T . Nous appliquons ces résultats aux informations de type produit que l'on utilise dans les applications.

II.1.- INFORMATION GÉNÉRALISÉE SUR LE TREILLIS DES PARTITIONS.

II-1.1.- TREILLIS DES PARTITIONS - RAPPELS.

Soit Ω un ensemble fini, non vide, quelconque, $S = \mathcal{P}(\Omega)$.

Une partition Π de Ω est un ensemble fini de parties de Ω tel que :

- $A \in \Pi \Longrightarrow A \neq \emptyset$
- $(A,A') \in \Pi \times \Pi \Longrightarrow A = A'$ ou $A \cap A' = \emptyset$
- $\bigcup_{A \in \Pi} A = \Omega$

Soit T l'ensemble des partitions de Ω . Une partition Π_1 est dite "plus fine" qu'une partition Π_2 si :

$$\forall A_1 \in \Pi_1 \quad , \quad \exists A_2 \in \Pi_2 \text{ tel que } A_1 \subseteq A_2$$

et nous noterons : $\Pi_1 \leqslant \Pi_2$

Cette relation est un ordre partiel, réflexif sur T ; tout couple de partitions (Π_1,Π_2) possède un infimum $\Pi_1 \wedge \Pi_2$ et un supremum $\Pi_1 \vee \Pi_2$:

$$\Pi_1 \wedge \Pi_2 = \{A \cap B | A \in \Pi_1 \text{ , } B \in \Pi_2\}$$

$$\Pi_1 \vee \Pi_2 = \text{Inf } \{\Pi \in T | \Pi \geqslant \Pi_1 \text{ et } \Pi \geqslant \Pi_2\}$$

De plus T possède un majorant universel $\Pi_M = \{\Omega\}$ et un minorant universel Π_m

$$\Pi_m = \{|\omega| \ |\omega \in \Omega\}$$

Ces propriétés suffisent pour pouvoir définir une information généralisée au sens de J. LOSFELD, sur T . Dans le cadre du treillis des partitions, les définitions énoncées au paragraphe I-1, s'écrivent :

II.1.2.- INFORMATION SUR T.

- Une information généralisée sur l'ensemble T des partitions de Ω est une application H définie sur T et qui vérifie les axiomes :

AXIOME 1 : H est une application de T dans $\bar{\mathbb{R}}^+$

AXIOME 2 : L'application H est monotone décroissante pour la relation d'ordre $\precsim$ sur T .

On peut prendre comme valeur universelle $H(\{\Omega\}) = 0$ (nous trouverons par la suite une mesure d'information en vérifiant pas cette condition). Nous ne ferons aucune hypothèse sur la valeur de $H(\Pi_m)$, en particulier, nous ne supposerons pas que $H(\Pi_m) = +\infty$ (nous obtenons par la suite des mesures d'information vérifiant $H(\Pi_m) < +\infty$).

II.1.3.- INFORMATION MOYENNE SUR T.

Etant données une probabilité P et une information I au sens de J. KAMPE DE FERIET (voir paragraphe III-1-1.), soit Π une partition finie S-mesurable de Ω [20] Nous supposons que le résultat d'une expérience est la réalisation d'un événement élémentaire $A_i \in \Pi$ sans que l'on puisse dire lequel.

L'information moyenne qui est apportée par cette expérience, quand on la renouvelle un grand nombre de fois est égale à :

$$H(\Pi) = \sum_{A_i \in \Pi} P(A_i)\, I(A_i)$$

Cette formule n'est qu'une possibilité parmi d'autres de relier les axiomes de l'information apportée par un événement à ceux de l'information apportée par une expérience ou une partition.

C'est une *information généralisée* au sens précédent (II.1.2.).

II.2.- PRINCIPALES MESURES D'INFORMATION SUR LES PARTITIONS ET HYPOTHÈSES H.

Soit Q un questionnaire de type probabiliste.

On applique Q sur un ensemble Ω , l'ensemble E des réponses de Q définissant une partition Π de Ω en N événements disjoints, la distribution de probabilités de ces réponses étant définie par

$$P = (p_1, p_2, \dots, p_N) \quad \text{tels que} \quad \sum_{i=1}^{N} p_i = 1 \quad , \quad p_i \in [0,1]$$

Dans les notions informationnelles intervenant en Théorie des Questionnaires apparaissent deux aspects de l'information :

- l'information à caractère local
- l'information à caractère global

qui ont été mis en évidence par S. PETOLLA et C.F. PICARD [22]:

"L'information considérée dans les questionnaires porte sur les aspects locaux et globaux. *Localement*, un chemin conduisant de la racine à un sommet i apporte une certaine information liée à la connaissance de l'*événement* associé à i par rapport à l'incertitude initiale correspondant à la seule connaissance, à la racine, que l'événement sera un sous-ensemble de E , comportant un ou plusieurs éléments. Les aspects *globaux* de l'information étant liés au *canal*".

Il est commun de distinguer, dans les théories de l'information généralisée, l'information I(A) relative à un événement de l'information H(Π) relative à une expérience réalisant une partition Π de l'espace des événements Ω en N éléments $(A_1, \dots, A_N)$.

En Théorie des Questionnaires, les mesures d'information sont le plus souvent utilisées dans le contexte information sur une partition (information transmise); FORTE et PINTACUDA, par une extension de type questionnaire, ont étendu aux expériences les informations généralisées de KAMPE DE FERIET.

Rappelons rapidement la définition des mesures les plus courantes [22] :

. SHANNON (type 1 et d'ordre 1)

$$H_N^1(\Pi) = \sum_{A_i \in \Pi} P(A_i) \text{ Log } \frac{1}{P(A_i)}$$

.

$$H_0(\Pi) = \sum_{A_i \in \Pi} P(A_i) \left(\frac{1}{P(A_i)} - 1\right)$$

. HYPERBOLIQUE

$$Y_N^0(\Pi) = \left(\sum_{A_i \in \Pi} \frac{1}{P(A_i)}\right) - 1$$

. RENYI (ordre α)

En prenant $\alpha > 0$, $\alpha \neq 1$, a entier au moins égal à 2 :

$$_\alpha H_N(\Pi) = \frac{1}{1-\alpha} \text{ Log}_a \sum_{A_i \in \Pi} P^\alpha(A_i)$$

. HAVRDA-CHARVAT (type α)

$$H_N^\alpha(\Pi) = \frac{\left(\sum_{A_i \in \Pi} P^\alpha(A_i)\right) - 1}{a^{1-\alpha} - 1} \quad \text{avec } \alpha > 0 \ , \ \alpha \neq 1$$

. AGGARWAL - CESARI - PICARD

En prenant $\alpha > 0$, $\alpha \neq 1$

$$J_N^\alpha(\Pi) = \sum_{A_i \in \Pi} P^\alpha(A_i)$$

. BELIS - GUIASCU

$$G_N(\Pi, U) = \sum_{A_i \in \Pi} \omega(A_i) \text{ Log } \frac{1}{P(A_i)}$$

En dehors de l'axiomatique de l'information généralisée de J. KAMPE DE FERIET, définie pour des espaces mesurables, deux types principaux de modèle ont permis la construction de ces mesures d'information :

. Certains auteurs définissent des *fonctions d'information,* déterminées éventuellement comme solutions d'équations fonctionnelles [24]
C.F. PICARD étudie notamment une équation basée sur la *propriété de somme,* définissant une mesure d'information comme somme de fonctions continues ne dépendant que des probabilités P_i

$$(I_N(P) = \sum_{i=1}^{N} f(P_i))$$

La solution de cette équation lui permet d'obtenir l'information de type α , quelque soit α positif [24]. Ce modèle est essentiellement lié au *caractère global* de l'information.

. D'autres auteurs utilisent *"localement"* une mesure d'information, puis pondèrent les réponses de l'expérience; c'est le cas, notamment, des informations de WATANABE

$$(\sum_{i=1}^{N} k_i \text{ Log } \frac{1}{p_i})$$

dont l'information de BELIS et GUIASCU est un cas particulier.

C'est également la voie qu'a choisie J. LOSFELD en proposant une formulation du *type moyenne,*

$$H(\Pi) = \sum_{A_i \in \Pi} p(A_i)\ I(A_i) \ ,$$

où $I(A_i)$ est une information généralisée sur les événements au sens de J. KAMPE DE FERIET.

En 1975, N.L. AGGARWAL a défini une mesure d'information notée

$$H_\emptyset(\Pi) = \sum_{i=1}^{N} p(A_i)\ I_\emptyset(A_i) \qquad [1]$$

dont la formulation et les propriétés sont proches de celles de l'information moyenne de J. LOSFELD. Il a pu établir les liens existant entre $H_{\emptyset}(\Pi)$ et certaines des mesures définies plus haut (SHANNON, type α , hyperbolique) [1].

Nous définissons localement des mesures d'information I qui nous permettent, en utilisant le modèle information moyenne, de retrouver les expressions de certaines des mesures utilisées dans les questionnaires.

Nous étudierons plus précisément ces informations généralisées sur les événements au chapitre III .

II.2.1.- INFORMATION DE SHANNON SUR LES PARTITIONS.

. L'information apportée par une partition au sens de SHANNON, s'écrit à un coefficient multiplicatif près (strictement positif) :

$$H_N^1(\Pi) = \sum_{A_i \in \Pi} p(A_i) \operatorname{Log} \frac{1}{p(A_i)}$$

où $\operatorname{Log} \frac{1}{p(A_i)}$ est l'information apportée par l'événement A_i quelconque dans (Ω,S,P).

. C'est une information moyenne sur le treillis T des partitions de l'espace des événements Ω , vérifiant les propriétés suivantes :

(i) $H_N^1(\Pi_1;\Pi_2) = \sum_{(A,B)\in\Pi_1 x\Pi_2} P(A \cap B) \operatorname{Log} \frac{P(A \cap B)}{P(A).P(B)}$ positif ou nul

(ii) Quelque soit Π_1 fixé, l'application $H_N^1(.;\Pi_1)$ est une *information généralisée* sur $(T,\leqslant)$.

II.2.1.1.- PROPOSITION.

La mesure d'information de SHANNON sur les partitions vérifie l'hypothèse H1 , soit :

$$H_N^1(\Pi_1) + H_N^1(\Pi_2) \geq H_N^1(\Pi_1 \wedge \Pi_2) + H_N^1(\Pi_1 \vee \Pi_2)$$

Il vient alors :

$$\delta^+(\Pi_1,\Pi_2) = \delta^-(\Pi_2,\Pi_1) = \sum_{(A,B)\in\Pi_1 x\Pi_2} P(A \cap B) \text{ Log } \frac{P(A)}{P(A \cap B)}$$

$$\delta(\Pi_1,\Pi_2) = \sum_{(A,B)\in\Pi_1 x\Pi_2} P(A \cap B) \text{ Log } \frac{P(A).P(B)}{P^2(A \cap B)}$$

$$H_R(\Pi_1,\Pi_2) = \sum_{(A,B)\in\Pi_1 x\Pi_2} P(A \cap B) \text{ Log } \frac{P(A \cap B)}{P(A)\ P(B)}$$

$$H_T(\Pi_1,\Pi_2) = \sum_{(A,B)\in\Pi_1 x\Pi_2} P(A \cap B) \text{ Log } \frac{1}{P(A \cap B)}$$

<u>DEMONSTRATION</u>.-

Démontrer que $H_N^1(\Pi)$ vérifie l'hypothèse H1 revient à prouver l'inégalité :

$$H_N^1(\Pi_1;\Pi_2) \geq H_N^1(\Pi_1 \vee \Pi_2)$$

soit encore

$$\sum_{A\in\Pi_1} \sum_{B\in\Pi_2} P(A \cap B) \text{ Log } \frac{P(A \cap B)}{P(A)\ .\ P(B)} \geq \sum_{C\in\Pi_1 \vee\Pi_2} P(C) \text{ Log } \frac{1}{P(C)}$$

Il suffit de démontrer l'inégalité pour un élément quelconque C de $\Pi_1 \vee \Pi_2$, soit :

$$\alpha_C = \Big(\sum_{A\in\Pi_1(C)} \sum_{B\in\Pi_2(C)} P(A \cap B) \text{ Log } \frac{P(A \cap B)}{P(A)\ .\ P(B)}\Big) - P(C) \text{ Log } \frac{1}{P(C)} \geq 0$$

avec :

$$\Pi_1(C) = \{A_i \mid A_i \in \Pi_1 \quad \text{et} \quad A_i \subset C\}$$

$$\Pi_2(C) = \{B_j \mid B_j \in \Pi_2 \quad \text{et} \quad B_j \subset C\}$$

De $P(C) = \sum_{A \in \Pi_1(C)} \; \sum_{B \in \Pi_2(C)} P(A \cap B)$, il vient :

$$\frac{\alpha_C}{P(C)} = \sum_{A \in \Pi_1(C)} \; \sum_{B \in \Pi_2(C)} P(A \cap B/C) \text{ Log } \frac{P(A \cap B/C)}{P(A/C) \cdot P(B/C)}$$

$$= H_N^1(\Pi_1(C) \; ; \; \Pi_2(C)/C)$$

En notant $H_N^1(\Pi/\Pi')$ l'information conditionnelle fournie par la partition connaissant l'information apportée par Π' [20].
Donc la quantité α_C est positive ou nulle (propriété (i)).

II.2.2.- INFORMATION $H_0(\Pi)$

Soit l'information hyperbolique sur l'espace des événements (Ω, S, P) introduite par J. KAMPE DE FERIET

$$I(A) = \frac{1}{P(A)} - 1$$

L'information *moyenne* définie à partir de cette mesure s'écrit :

$$H_0(\Pi) = \sum_{A_i \in \Pi} P(A_i) \left(\frac{1}{P(A_i)} - 1\right)$$

$$= \text{card}(\Pi) - 1$$

II.2.2.1.- PROPOSITION.

La mesure d'information définie sur le treillis des partitions par $H_0(\Pi) = \text{card}(\Pi) - 1$ *vérifie l'hypothèse H2.*
Il vient alors :

$$\delta^+(\Pi_1, \Pi_2) = \delta^-(\Pi_2, \Pi_1) = \text{card}(\Pi_2) - \text{card}(\Pi_1 \vee \Pi_2)$$

$$\delta(\Pi_1, \Pi_2) = \text{card}(\Pi_1) + \text{card}(\Pi_2) - 2 \, \text{card}(\Pi_1 \vee \Pi_2)$$

$$H_R(\Pi_1,\Pi_2) = \mathrm{card}(\Pi_1 \vee \Pi_2) - 1$$

$$H_T(\Pi_1,\Pi_2) = \mathrm{card}(\Pi_1) + \mathrm{card}(\Pi_2) - \mathrm{card}(\Pi_1 \vee \Pi_2) - 1$$

La démonstration résulte directement de l'inégalité bien connue sur le treillis des partitions :

$$\mathrm{card}(\Pi_1) + \mathrm{card}(\Pi_2) \leq \mathrm{card}(\Pi_1 \vee \Pi_2) + \mathrm{card}(\Pi_1 \wedge \Pi_2)$$

II.2.3.- MESURES D'INFORMATION DE :

- . AGGARWAL - CESARI - PICARD
- . HAVRDA - CHARVAT (type α)
- . BELIS et GUIASCU
- . RENYI (ordre α)

Une étude détaillée des informations définies ci-dessus (voir introduction du paragraphe II-2) nous a permis d'établir les résultats présentés dans le tableau II-4-1.

Toutes ces quantités sont des informations généralisées au sens du paragraphe II-1; les trois premières mesures peuvent se formaliser en information de type moyenne, ce qui n'est pas le cas pour l'information de RENYI ; les informations généralisées sur les événements associées étant respectivement:

. $I(A) = P^{\alpha-1}(A)$

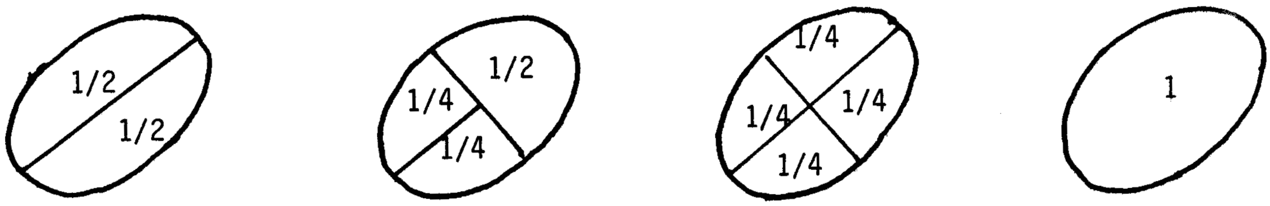
. $I(A) = \dfrac{P^{\alpha-1}(A) - 1}{a^{1-\alpha} - 1}$

. $I(A) = \mathrm{Log}\,\dfrac{1}{P(A)}$

Il est facile de déterminer des contre-exemples prouvant qu'aucune de ces mesures ne vérifie l'une des hypothèses H1 ou H2 . Par exemple, étudions l'information d'AGGARWAL - CESARI - PICARD :

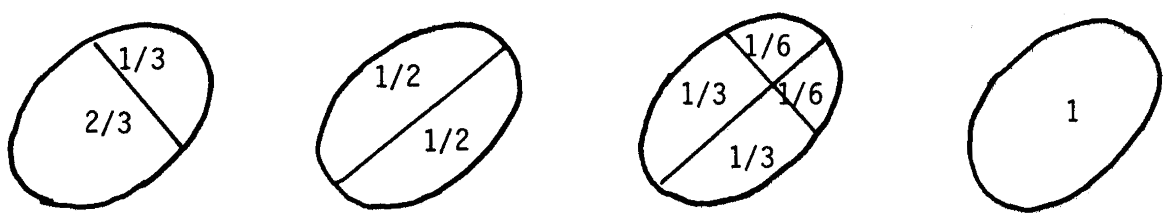
. Pour les deux partitions Π_1 et Π_2 représentées ci-dessous,

$J_N^{1/2}(\Pi)$ vérifie l'hypothèse H1 (les chiffres indiquent les probabilités des éléments de partitions correspondants)



Π_1 Π_2 $\Pi_1 \wedge \Pi_2$ $\Pi_1 \vee \Pi_2$

. Avec les deux partitions ci-dessous, $J_N^{1/2}(\Pi)$ vérifie l'hypothèse H2



Π_1 Π_2 $\Pi_1 \wedge \Pi_2$ $\Pi_1 \vee \Pi_2$

II.3.- MESURES D'INFORMATION DE TYPE PRODUIT

La formalisation des exemples présentés dans l'introduction de sa thèse [20] a amené J. LOSFELD à orienter son étude vers les informations généralisées définies sur un espace produit.

Soit O un ensemble fini (d'observateurs) et une famille finie $\{H_\xi | \xi \in O\}$ de mesures d'informations généralisées sur le treillis T des partitions de l'ensemble Ω fini, et λ une mesure de probabilités sur $(O, P(O))$. On définit dans [20] diverses mesures d'informations généralisées H sur le treillis T par des formules du type :

$$H(\Pi) = \sum_{\xi \in O} \lambda(\xi) \cdot H_\xi(\Pi)$$

Il est clair que si toutes les mesures d'informations H_ξ vérifient l'hypothèse H1 (resp. H2), alors H vérifie aussi l'hypothèse H1 (resp. H2).

En particulier, si nous utilisons pour H_ξ , et ceci pour tout $\xi \in \mathcal{O}$ la mesure d'information de SHANNON sur les partitions (resp. $H_0(\Pi)$) alors $H(\Pi)$ vérifiera l'hypothèse H1 (resp. H2).

Nous résumons dans le tableau II-4-2 , les expressions des quantités δ^+ , δ^- δ , H_R et H_T pour les mesures d'information vérifiant l'une des hypothèses de type H .

voir tableau II-4-1 page suivante

	Mesure	Information généralisée au sens de J. LOSFELD	Information moyenne / I(A)	Hypothèse H
SHANNON	$\sum_{A\in\Pi} P(A) \operatorname{Log} \frac{1}{P(A)}$	OUI	OUI / $\operatorname{Log} \frac{1}{P(A)}$	H1
H_0	$\sum_{A\in\Pi} P(A) \left(\frac{1}{P(A)} - 1\right)$	OUI	OUI / $\frac{1}{P(A)} - 1$	H2
AGGARWAL CESARI PICARD J_N^α	$\sum_{A\in\Pi} P^\alpha(A)$ $0 < \alpha < 1$	OUI	OUI / $P^{\alpha-1}(A)$	ni H1 ni H2 pour tout $0 < \alpha < 1$
HAVRDA CHARVAT (type α)	$\frac{\left(\sum_{A\in\Pi} P^\alpha(A)\right) - 1}{a^{1-\alpha} - 1}$ $\alpha > 0$	OUI	OUI / $\frac{P^{\alpha-1}(A) - 1}{a^{1-\alpha} - 1}$	ni H1 ni H2 pour tout $\alpha > 0$
BELIS	$\sum_{A\in\Pi} \omega(A) \operatorname{Log} \frac{1}{P(A)}$	OUI	OUI / $\operatorname{Log} \frac{1}{P(A)}$	ni H1 ni H2
RENYI (Ordre α)	$\frac{1}{1-\alpha} \operatorname{Log} \sum_{A\in\Pi} p^\alpha(A)$ $\alpha > 0$	OUI	NON	ni H1 ni H2 pour tout $\alpha > 0$
Type PRODUIT [20]	$\sum_{\xi\in O} \lambda(\xi)\, H_\xi(\Pi)$	OUI	OUI si toutes les H_ξ sont moyennes	H1 (resp. H2) si toutes les H_ξ vérifient H1 (resp. H2

INFORMATION GENERALISEE SUR LE TREILLIS DES PARTITIONS

TABLEAU RECAPITULATIF

TABLEAU II-4-1

	$\delta^+(\Pi_1,\Pi_2) = \delta^-(\Pi_2,\Pi_1)$	$\delta(\Pi_1,\Pi_2)$	$H_R(\Pi_1,\Pi_2)$	$H_T(\Pi_1,\Pi_2)$
SHANNON	$\sum_{(A,B)\in\Pi_1 x\Pi_2} P(A\cap B) \text{ Log } \frac{P(A)}{P(A\cap B)}$	$\sum_{(A,B)\in\Pi_1 x\Pi_2} P(A\cap B) \text{ Log } \frac{P(A)P(B)}{P^2(A\cap B)}$	$\sum_{(A,B)\in\Pi_1 x\Pi_2} P(A\cap B) \text{ Log } \frac{P(A\cap B)}{P(A)P(B)}$	$\sum_{(A,B)\in\Pi_1 x\Pi_2} P(A\cap B) \text{ Log } \frac{1}{P(A\cap B)}$
H_0	$\text{Card}(\Pi_2) - \text{Card}(\Pi_1 \vee \Pi_2)$	$\text{Card}(\Pi_1) + \text{Card}(\Pi_2) - 2\ \text{Card}(\Pi_1 \vee \Pi_2)$	$\text{Card}(\Pi_1 \vee \Pi_2) - 1$	$\text{Card}(\Pi_1) + \text{Card}(\Pi_2) - \text{Card}(\Pi_1 \vee \Pi_2) - 1$
type PRODUIT 6 vérifiant H1 ou H2	$\sum_{\xi\in O} \lambda(\xi) \cdot \delta^+_\xi(\Pi_1,\Pi_2) = \sum_{\xi\in O} \lambda(\xi) \cdot \delta^-_\xi(\Pi_2,\Pi_1)$	$\sum_{\xi\in O} \lambda(\xi) \cdot \delta_\xi(\Pi_1,\Pi_2)$	$\sum_{\xi\in O} \lambda(\xi) \cdot H_{\xi R}(\Pi_1,\Pi_2)$	$\sum_{\xi\in O} \lambda(\xi) \cdot H_{\xi T}(\Pi_1,\Pi_2)$

EXPRESSIONS DES δ^+ , δ^- , δ , H_R , H_T

TABLEAU RECAPITULATIF

TABLEAU II-4-2

C H A P I T R E   III

-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

INFORMATION GENERALISEE SUR LES EVENEMENTS

-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=-=

III.- INFORMATION GENERALISEE SUR LES EVENEMENTS.

Après avoir rappelé les définitions et propriétés des mesures d'information généralisée au sens de J. KAMPE DE FERIET sur les événements, nous étudions les informations composables.

Nous donnons une condition nécessaire et suffisante pour qu'une information composable vérifie l'hypothèse H2 et étudions ce que devient cette condition dans le cas d'une loi de composition régulière et dérivable. Nous appliquons ensuite ces résultats pour les mesures d'informations sur les événements les plus utilisées dans les applications (Théorie des Questionnaires, Analyse des Données).

Nous terminons ce chapitre par un tableau récapitulatif des expressions des quantités δ^+ , δ^- , δ , I_R et I_T , obtenues pour les principales mesures d'information généralisée définies sur une algèbre de Boole de parties d'un ensemble Ω et qui vérifient l'hypothèse H2.

III.1.- RAPPELS.

III.1.1.- INFORMATION GENERALISEE SUR LES EVENEMENTS.

Soient Ω un ensemble quelconque, non vide, d'événements élémentaires Ω , S une σ-algèbre de Boole de parties de Ω . Une mesure d'information généralisée sur les événements de (Ω,S) au sens de J. KAMPE DE FERIET ou plus brièvement une "information sur les événements de (Ω,S)" est une fonction d'ensemble I définie sur S et satisfaisant aux axiomes suivants :

Axiome 1 : I est une application de S dans $\bar{\mathbb{R}}^+$

Axiome 2 : I est monotone pour l'inclusion :

$$(A,B) \in S \times S \text{ et } A \subseteq B \Longrightarrow I(A) \geq I(B)$$

Nous prendrons comme valeurs universelles $I(\emptyset) = +\infty$ et $I(\Omega) = 0$, sauf pour certains cas particuliers étudiés par la suite où, soit $I(\Omega) \neq 0$, soit $I(\emptyset) \neq +\infty$ et pour lesquels nous donnerons les résultats correspondants.

III.1.2.- LOI DE COMPOSITION

Une information I , définie sur une classe S fermée par rapport à l'union est dite composable s'il existe une application F de $\bar{\mathbb{R}}^+ \times \bar{\mathbb{R}}^+$ dans $\bar{\mathbb{R}}^+$ telle que, pour tout couple $(A,B) \in S \times S$, $A \cap B = \emptyset$, on ait [7] :

$$I(A \cup B) = F(I(A) , I(B))$$

où F a les propriétés d'une opération de composition :

$$F : \bar{\mathbb{R}}^+ \times \bar{\mathbb{R}}^+ \longrightarrow \bar{\mathbb{R}}^+$$

- $F(x,y) = F(y,x)$
- $F(x,F(y,z)) = F(F(x,y),z)$
- $F(x,+\infty) = x$
- $x' < x'' \Longrightarrow F(x',y) \leq F(x'',y)$

On note qu'alors, pour toute opération de composition F :

$$F(x,y) \leq \mathrm{Inf}(x,y)$$

Notons :

$\Gamma = \{(x,y) \in \bar{\mathbb{R}}^+ \times \bar{\mathbb{R}}^+ : x = I(A) \ , \ y = I(B) \ , A \in S \ , \ B \in S \ , A \cap B = \emptyset\}$

$\mathcal{F} = \{(A,B) \in S \times S \ , \ A \cap B = \emptyset\}$

III-2.- INFORMATIONS COMPOSABLES.

Avec la valeur universelle $I(\emptyset) = +\infty$ introduite en III-1-1, il est clair que seule l'hypothèse H2 peut être éventuellement vérifiée par I .

HYPOTHESE 2 : Pour tout $(A,B) \in S \times S$;

$$I(A) + I(B) \leq I(A \cup B) + I(A \cap B)$$

Supposons, de plus, que la mesure d'information I soit composable. Il vient alors :

II.2.1.- THEOREME.

Une condition nécessaire et suffisante pour que la mesure d'information I *, composable de loi* F *, définie sur* (Ω,S) *vérifie l'hypothèse H2 est que l'application* $x - F(x,y)$ *soit sur* Γ *, non décroissante en* x *, pour tout* y *fixé.*

On a :

$$\delta(A,B) = I(A) + I(B) - 2\, F(I(A)\ ,\ I(B))$$

$$\delta^+(A,B) = \delta^-(B,A) = I(B) - F(I(A)\ ,\ I(B))$$

$$I_R(A,B) = F(I(A)\ ,\ I(B))$$

$$I_T(A,B) = I(A) + I(B) - F(I(A)\ ,\ I(B))$$

DEMONSTRATION.-

a) Supposons que I vérifie l'hypothèse H2 :

Pour tout $(A,B) \in S \times S$, $I(A) + I(B) \leq I(A \cup B) + I(A \cap B)$

$$I(B) = F(I(A \cap B) , I(B-A))$$

$$I(A \cup B) = F(I(A) , I(B-A))$$

ce qui donne, en posant $x = I(A)$, $y = I(B-A)$, $z = I(A \cap B)$

$$x - F(y,x) \leq z - F(y,z)$$

avec $x \leq z$, quelque soit y positif ou nul.

Il en résulte que l'application $x - F(x,y)$ est monotone non décroissante en x , quelque soit y positif ou nul.

b) La démonstration de la condition suffisante est évidente.

III.2.2.- INFORMATION GENERALISEE MUNIE D'UNE LOI DE COMPOSITION REGULIERE.

Etudiant plus particulièrement les mesures d'information possédant une opération de composition continue, J. KAMPE DE FERIET et P. BENVENUTI [18] appellent régulière une opération de composition vérifiant, en plus des axiomes énoncés en III.1.2, l'axiome suivant :

$$F \in C[\bar{\mathbb{R}}^+ \times \bar{\mathbb{R}}^+]$$

Ces auteurs déterminent toutes les opérations régulières et l'expression générale de ces opérations : il existe un fermé Λ de $\bar{\mathbb{R}}^+$ contenant 0 et $+\infty$. Le complémentaire de Λ dans $\bar{\mathbb{R}}^+$ est réunion dénombrable d'intervalles ouverts disjoints non vides $]a_i\ b_i[$ avec $i \in I$ (card $I \leq \aleph_0$) .

A chaque $i \in I$ associons une application Θ_i continue, strictement décroissante de $[0,\bar{\mu}_i]$ sur $[a_i\ b_i]$ avec $\bar{\mu}_i = \Theta_i^{-1}(a_i)$.

Soit Φ_i la pseudo inverse de Θ_i^{-1} .

La forme générale d'une opération de composition régulière est donnée par :

$$F(x,y) = \Phi_i[\Theta_i^{-1}(x) + \Theta_i^{-1}(y)] \quad \text{si} \quad (x,y) \in [a_i\ b_i] \times [a_i\ b_i]$$

$$F(x,y) = \mathrm{Inf}(x,y) \quad \text{si} \quad (x,y) \in \bar{\mathbb{R}}^+ \times \bar{\mathbb{R}}^+ - \bigcup_{i \in I} [a_i\ b_i] \times [a_i\ b_i]$$

III.2.2.1.- THEOREME.

Soit I *une mesure d'information généralisée sur* (Ω,S) *composable de loi* F . *Si* F *est une opération de composition régulière prenant la forme générale définie par J. KAMPE DE FERIET, telle que toutes les applications* Θ_i *soient continuement dérivables sur* $[0,\bar{\mu}_i]$ *, une condition nécessaire et suffisante pour que* I *vérifie l'hypothèse H2 est que, pour tout* i *, les dérivées* $\Theta_i'(x)$ *soient des applications monotones non décroissantes en* x .

DEMONSTRATION.

D'après le théorème III-2-1 , une condition nécessaire et suffisante pour que I vérifie H2 est que l'application x - F(x,y) soit non décroissante en x , pour tout y fixé. Il est facile de voir que cette condition équivaut à montrer que les applications

$$\Psi_i(x) = x - \Phi_i[\Theta_i^{-1}(x) + \Theta_i^{-1}(y)]$$

sont non décroissantes, pour tout y fixé appartenant à $[a_i\ b_i]$.

On montre que si $x < \Theta_i[\bar{\mu}_i - \Theta_i^{-1}(y)]$, alors

$$\Psi_i(x) = x - a_i$$

Ψ_i est alors une application non décroissante.

Si $x \geq \Theta_i[\bar{\mu}_i - \Theta_i^{-1}(y)]$ et si l'on suppose que Θ_i est une application dérivable, on montre facilement que Ψ_i est une application non décroissante si et seulement si Θ_i' est une application non décroissante, soit encore, si elle existe, Θ_i'' positive ou nulle.

III.2.3.- EXEMPLE.

. J. KAMPE DE FERIET et B. FORTE ont introduit une classe particulière d'informations composables dites de type M [17]:

Soit (Ω,S,P) un espace de mesure quelconque $(\mu(\Omega) = \bar{\mu})$;
$0 < \bar{\mu} \leq +\infty$, pour tout $A \in S$, posons :

$$I(A) = \Theta(\mu(A))$$

où Θ est une application vérifiant les propriétés :

- $\Theta : \mu(S) \to \mathbb{R}^+$
- Θ est strictement décroissante
- $\Theta(0) = +\infty,\ \Theta(\bar{\mu}) = 0$

$$\forall\ (x,y) \in \Gamma \quad , \quad F(x,y) = \Theta(\Theta^{-1}(x) + \Theta^{-1}(y))$$

. Lorsqu'une mesure d'information de type M sur les événements vérifie l'hypothèse H2 , il vient alors :

$$\delta^+(A,B) = \delta^-(B,A) = \Theta(\mu(B)) - \Theta(\mu(A \cup B))$$

$$\delta(A,B) = \Theta(\mu(A)) + \Theta(\mu(B)) - 2\,\Theta(\mu(A \cup B))$$

$$I_R(A,B) = \Theta(\mu(A \cup B))$$

$$I_T(A,B) = \Theta(\mu(A)) + \Theta(\mu(B)) - \Theta(\mu(A \cup B))$$

. Toutes les informations de type M ne vérifient pas l'hypothèse H2.

EXEMPLES.-

a) Cas de l'information de SHANNON

En prenant $\Lambda = (0,+\infty)$, $\bar{\mu} = 1$ et Θ telle que :

$$\left|\begin{array}{l} \Theta(x) = c \operatorname{Log} \frac{1}{x} \quad \text{si} \quad x \in \mathbb{R}_+^* \\ \Theta(0) = +\infty \\ \Theta(1) = 0 \end{array}\right.$$

où c est une constante positive.

On obtient l'opération de composition régulière

$$F(x,y) = \operatorname{Sup}(0, -c \operatorname{Log}(e^{-\frac{x}{c}} + e^{-\frac{y}{c}}))$$

correspondant à l'information de WIENER-SHANNON

$$I(A) = \operatorname{Log} \frac{1}{P(A)} \quad \text{si} \quad P(A) \neq 0$$

$$I(A) = +\infty \quad \text{si} \quad P(A) = 0$$

définie sur l'espace probabilisé (Ω,S,P)

L'application du théorème III.2.2.1. montre que cette mesure d'information vérifie l'hypothèse H2 .

b) En prenant $\Lambda = (0, +\infty)$, $\bar{\mu} = 1$ et $\Theta(x) = \operatorname{cotg} \frac{3\Pi x}{4} + 1$ avec $\Theta(0) = +\infty$ et $\Theta(1) = 0$, on obtient l'opération de composition régulière :

pour tout $(x,y) \in \Gamma$, $F(x,y) = \frac{xy - 2}{x+y-2}$

L'application du théorème III.2.2.1. montre que la mesure d'information de type-M sur les événements définie, pour tout $A \in S$, par $I(A) = \operatorname{cotg} \frac{3\Pi\, P(A)}{4} + 1$, ne vérifie pas l'hypothèse H2.

III-3.- APPLICATIONS AUX PRINCIPALES INFORMATIONS COMPOSABLES.

En reprenant les notations de III.2.2., nous présentons dans ce paragraphe les résultats concernant les principales mesures d'information composables.

En application du théorème III.2.2.1, nous établissons le tableau suivant en indiquant, pour chaque mesure d'information, l'expression de $I(A)$, la loi de composition associée, si l'hypothèse H2 est vérifiée ou non, ainsi que les expressions des quantités δ^+ , δ^- , δ , I_R et I_T (quand c'est possible).

REMARQUE.-

Si α est supérieur à 2 , la mesure associée à l'information de type α sur les partitions, définie par

$$I(A) = \frac{P^{\alpha-1}(A) - 1}{a^{1-\alpha} - 1}$$

vérifie l'hypothèse H1 (ceci est dû au fait que $I(\Phi)$ n'est pas égal à $+\infty$).

Il vient alors :

$$\delta^+(A,B) = \delta^-(B,A) = \frac{P^{\alpha-1}(A \cap B) - P^{\alpha-1}(A)}{a^{1-\alpha} - 1}$$

$$\delta(A,B) = \frac{2\, P^{\alpha-1}(A \cap B) - P^{\alpha-1}(A) - P^{\alpha-1}(B)}{a^{1-\alpha} - 1}$$

$$I_R(A,B) = \frac{P^{\alpha-1}(A) + P^{\alpha-1}(B) - P^{\alpha-1}(A \cap B) - 1}{a^{1-\alpha} - 1}$$

$$I_T(A,B) = \frac{P^{\alpha-1}(A \cap B) - 1}{a^{1-\alpha} - 1}$$

	$I(A)$	$F(x,y)$	HYPOTHESE H	$\delta^{+}(A,B) = \delta^{-}(B,A)$	$\delta(A,B)$	$I_R(A,B)$	$I_T(A,B)$
type M vérifiant l'hypothèse H2	$\Theta(\mu(A))$	$\Theta(\Theta^{-1}(x)+\Theta^{-}(y))$	H2	$\Theta(\mu(B))-\Theta(\mu(A\cup B))$	$\Theta(\mu(A))+\Theta(\mu(B)) - 2\Theta(\mu(A\cup B))$	$\Theta(\mu(A\cup B))$	$\Theta(\mu(A))+\Theta(\mu(B)) - \Theta(\mu(A\cup B))$
SHANNON	$\mathrm{Log}\,\frac{1}{P(A)}$	$\mathrm{Sup}(0,-c\,\mathrm{Log}(e^{-\frac{x}{c}} + e^{-\frac{y}{c}}))$	H2	$\mathrm{Log}\,\frac{P(A\cup B)}{P(B)}$	$\mathrm{Log}\,\frac{P^2(A\cup B)}{P(A)\,P(B)}$	$\mathrm{Log}\,\frac{1}{P(A\cup B)}$	$\mathrm{Log}\,\frac{P(A\cup B)}{P(A).P(B)}$
HYPERBOLIQUE	$\frac{1}{P(A)}-1$	$\frac{1}{\frac{1}{x+1}+\frac{1}{y+1}}-1$	H2	$\frac{1}{P(B)}-\frac{1}{P(A\cup B)}$	$\frac{1}{P(A)}+\frac{1}{P(B)}-\frac{2}{P(A\cup B)}$	$\frac{1}{P(A\cup B)}-1$	$\frac{1}{P(A)}+\frac{1}{P(B)}-\frac{1}{P(A\cup B)}-1$
INFORMATION associée à la mesure J_N^{α} sur les partitions	$P^{\alpha-1}(A)$ $0<\alpha<1$	$(x^{\frac{1}{\alpha-1}}+y^{\frac{1}{\alpha-1}})^{\alpha-1}$	H2	$P^{\alpha-1}(B)-P^{\alpha-1}(A\cup B)$	$P^{\alpha-1}(A)+P^{\alpha-1}(B) - 2P^{\alpha-1}(A\cup B)$	$P^{\alpha-1}(A\cup B)$	$P^{\alpha-1}(A)+P^{\alpha-1}(B)-P^{\alpha-1}(A\cup B)$
INFORMATION associée à l'information de type α sur les partitions	$\frac{P^{\alpha-1}(A)-1}{a^{1-\alpha}-1}$ $\alpha>0$	$\frac{[(x(a^{1-\alpha}-1)+1)^{1/\alpha-1}+(y(a^{\alpha-1}-1)+1)^{1/\alpha-1}]^{\alpha-1}}{a^{1-\alpha}-1}$	. H2 si $0<\alpha\leq 2$. H1 si $\alpha>2$	$\frac{P^{\alpha-1}(B)-P^{\alpha-1}(A\cup B)}{a^{1-\alpha}-1}$ si $0<\alpha\leq 2$	$\frac{P^{\alpha-1}(A)+P^{\alpha-1}(B) - 2P^{\alpha-1}(A\cup B)}{a^{1-\alpha}-1}$ si $0<\alpha\leq 2$	$\frac{P^{\alpha-1}(A\cup B)-1}{a^{1-\alpha}-1}$ si $0<\alpha\leq 2$	$\frac{P^{\alpha-1}(A)+P^{\alpha-1}(B)-P^{\alpha-1}(A\cup B)-1}{a^{1-\alpha}-1}$ si $0<\alpha\leq 2$
Type Inf		$\mathrm{Inf}(x,y)$	H2	$\mathrm{Sup}(0,I(B)-I(A))$	$\lvert I(A)-I(B)\rvert$	$I(A\cup B) = \mathrm{Inf}(I(A),I(B))$	$\mathrm{Sup}(I(A),I(B))$

INFORMATION GENERALISEE SUR LES EVENEMENTS [12]

TABLEAU RECAPITULATIF

C H A P I T R E   IV

-=-=-=-=-=-=-=-=-=-=-=

TREILLIS ET APPLICATIONS MONOTONES

-=-=-=-=-=-=-=-=-=-=-=

IV - TREILLIS ET APPLICATIONS MONOTONES.

Au chapitre I, nous avons tiré un certain nombre de conséquences de l'existence d'une fonction monotone définie sur un treillis T et à valeurs dans $\bar{\mathbb{R}}^+$.

Dans le présent chapitre, nous avons pour but d'étudier les propriétés des applications monotones sur un ensemble partiellement ordonné qui n'est pas forcément un treillis.

Comme au chapitre I, nous associons à toute fonction monotone v d'un treillis T dans $\mathbb{R} \cup \{+\infty\}$ une quasi métrique δ. Cette quasi métrique devient une métrique lorsque v est strictement monotone.

L'introduction d'hypothèses particulières sur v permet de donner des expressions simples de cette métrique. Si v est une valuation strictement croissante, on retrouve ainsi la distance habituelle sur les treillis modulaires. Si v est une mesure de probabilités, on retrouve la distance définie à partir de la différence symétrique.

Dans la première partie, nous reprenons brièvement les résultats du chapitre I dans le cadre d'une application monotone définie sur T à valeurs dans $\mathbb{R} \cup \{+\infty\}$. Nous énonçons sans démonstration ces résultats, les démonstrations étant semblables à celles effectuées au chapitre I.

Dans la deuxième partie, nous cherchons à généraliser ces résultats aux demi-treillis.

Cela nous a amené à caractériser les applications v monotones de T dans R telles que, pour tout couple (x,y) d'éléments de T, la variation de v sur la s.e.c. $(x, x \vee y, y)$ (dans le cas d'un sup-demi-treillis) est v-minimale. On en déduit [6] une généralisation de la notion de valuation aux demi-treillis ainsi qu'une caractérisation métrique de la semi-modularité dans les demi-treillis. On généralise ainsi une caractérisation bien connue (BIRKHOF [4]).

IV-1.- NOTATIONS ET RÉSULTATS GÉNÉRAUX DANS LES TREILLIS.

Soit T un treillis dont la relation d'ordre est notée $\leqslant$.

Soit v une application monotone de T dans $\mathbb{R} \cup \{+\infty\}$ que nous supposerons décroissante pour la relation d'ordre $\leqslant$.

Les résultats obtenus au chapitre I pour une mesure d'information généralisée J restent valables pour toute application monotone à valeurs dans $\mathbb{R} \cup \{+\infty\}$; ils ont été développés dans [14].

IV.1.1.- On a donc les résultats :

IV.1.1.1.- PROPOSITION.

- *L'application* δ *de* $T \times T$ *dans* $\bar{\mathbb{R}}^+$ *est une quasi distance sur* T
- *Si* $x \leqslant y$, $\delta(x,y) = v(x) - v(y)$
- *Les quantités* δ^+ *et* δ^- *sont des applications de* $T \times T$ *dans* $\bar{\mathbb{R}}^+$ *vérifiant les propriétés suivantes :*

  - $\delta^+(x,x) = \delta^-(x,x) = 0$ *pour tout* $x \in T$
  - $\delta^+(x,y) = \delta^-(y,x)$ *pour tout* $(x,y) \in T \times T$
  - δ^+ *et* δ^- *vérifient l'inégalité triangulaire.*

De plus δ^+ *et* δ^- *vérifient :*

(i) $v(y) - v(x) = \delta^+(x,y) - \delta^-(x,y)$

pour tout (x,y) *de* $T \times T$ *tel que* $v(x) \, . \, v(y) < +\infty$

(ii) $\delta(x,y) = \delta^+(x,y) + \delta^-(x,y)$

IV-1-1-2.- COROLLAIRE.

Etant donné un treillis T *fini,* $T_0 = \{x \in T | v(x) < +\infty\}$, *une condition nécessaire et suffisante pour que la restriction de l'application* δ *à l'ensemble des couples* (x,y) *de* $T_0 \times T_0$ *définisse une distance est que l'application* v *soit strictement monotone*

IV.1.2.- HYPOTHESES H.

H0 : $\forall\ (x,y) \in T \times T\ ,\quad v(x) + v(y) = v(x \wedge y) + v(x \vee y)$

H1 : $\forall\ (x,y) \in T \times T\ ,\quad v(x) + v(y) \geq v(x \wedge y) + v(x \vee y)$

H2 : $\forall\ (x,y) \in T \times T\ ,\quad v(x) + v(y) \leq v(x \wedge y) + v(x \vee y)$

IV.1.2.1.- THEOREME.

Les énoncés suivants sont équivalents :

(i) v *vérifie* H1

(ii) δ^+ *vérifie les égalités*
$\delta^+(x,y) = \Delta^+(C_\wedge(x,y)) = v(x \wedge y) - v(x)$
pour tout couple (x,y) *de* $T \times T$

(iii) δ^- *vérifie les égalités*
$\delta^-(x,y) = \Delta^-(C_\wedge(x,y)) = v(x \wedge y) - v(y)$
pour tout couple (x,y) *de* $T \times T$

(iv) δ *vérifie l'égalité* $\delta(x,y) = 2\ v(x \wedge y) - v(x) - v(y)$
pour tout couple (x,y) *de* $T \times T$.

En d'autres termes v vérifie H1 si et seulement si la s.e.c. $C_\wedge(x,y)$ est v-minimale.

On en déduit, dans le cas d'un treillis non nécessairement fini, la généralisation suivante du corollaire IV.1.1.2.

IV.1.2.2.- COROLLAIRE.

Etant donné un treillis T *quelconque (fini ou infini)*
$T_0 = \{x \in T \mid v(x) < +\infty\}$
si v *vérifie l'une des hypothèses* H0 , H1 , H2 , *une condition nécessaire et suffisante pour que la restriction de l'application* δ *à l'ensemble des couples* (x,y) *de* $T_0 \times T_0$ *définisse une distance est que l'application* v *soit strictement monotone.*

IV.1.3.- DISTANCES SUR T - FORMULATIONS EXPLICITES.

Chacune des hypothèses H nous permet de calculer les valeurs de δ^+, δ^- et δ, et d'en donner des formulations explicites très simples que nous présentons dans le tableau qui suit. Nous en ferons un usage permanent dans la construction des algorithmes du chapitre V.

Nous en donnons le tableau ci-dessous dans le cas d'une application v décroissante. Les formulations dans le cas où v est croissante s'obtiennent par dualité (on remplace le supremum (resp. l'infimum) par l'infimum (resp. le supremum), dans chaque formule du Tableau)

voir tableau page suivante

En application de ce tableau, il faut noter que toute mesure de probabilités P définie sur une algèbre de Boole S de parties d'un ensemble Ω, est monotone croissante et vérifie HO . On obtient donc

pour tout $(A,B) \in S \times S$

$$\delta^+(A,B) = \delta^-(B,A) = P(B) - P(A \cap B) = P(A \cup B) - P(A)$$

$$\delta(A,B) = P(A \cup B) - P(A \cap B) = P(A \Delta B)$$

La distance que nous avons ainsi construite coïncide donc, dans ce cas, avec la distance bien connue en Théorie des Probabilités, définie à partir de la différence symétrique.

v	HYPOTHESE H	δ^+	δ^-	δ
décroissante	H0	$v(x \wedge y) - v(x)$ = $v(y) - v(x \vee y)$	$v(x \wedge y) - v(y)$ = $v(x) - v(x \vee y)$	$v(x \wedge y) - v(x \vee y)$
décroissante	H1	$v(x \wedge y) - v(x)$	$v(x \wedge y) - v(y)$	$2\,v(x \wedge y) -$ $v(x) - v(y)$
décroissante	H2	$v(y) - v(x \vee y)$	$v(x) - v(x \vee y)$	$v(x) + v(y)$ $- 2\,v(x \vee y)$

TABLEAU DES DISTANCES SUR T

IV - 2.- EXTENSION AUX DEMI-TREILLIS

Soit T un treillis, et v une application monotone de T dans $\mathbb{R} \cup \{+\infty\}$.

Supposons que v soit une valuation. En d'autres termes, v est croissante et vérifie H0 . D'après le tableau précédent, on a :

$$\delta(x,y) = v(x \vee y) - v(x \wedge y)$$

La quasi-distance δ n'est autre que celle définie classiquement sur les treillis.

On sait [2] [4] qu'il y a équivalence entre les treillis modulaires et les treillis à valuation strictement croissante. On connaît [2] aussi la caractérisation suivante de la semi-modularité des treillis :

"Une condition nécessaire et suffisante pour qu'un treillis T soit semi-modulaire supérieurement (resp. inférieurement) est qu'il soit gradué, sa fonction de rang vérifiant :

$$r(x) + r(y) \geq r(x \vee y) + r(x \wedge y)$$

$$(\text{resp. } r(x) + r(y) \leq r(x \vee y) + r(x \wedge y))\text{ "}$$

Autrement dit, un treillis T est semi-modulaire supérieurement (resp. inférieurement) si et seulement s'il est muni d'une graduation qui vérifie H1 (resp. H2).

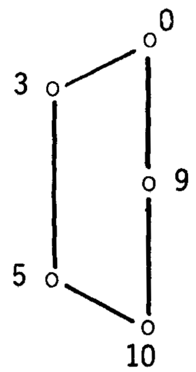
Il était donc naturel de se poser les questions suivantes :

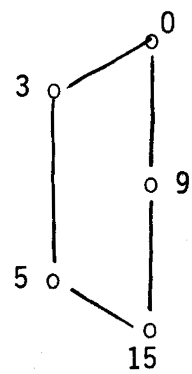
1.- Les hypothèses H1 et H2 permettent-elles, plus généralement, d'obtenir les mêmes résultats sur des treillis munis d'une application strictement monotone. ?

La réponse est NON. Si v n'est pas une valuation, les hypothèses H1 et H2 n'entraînent aucune structure particulière sur le treillis T , ni la modularité, ni la semi-modularité.
On le constatera sur les contre-exemples suivants :

Nous y présentons :

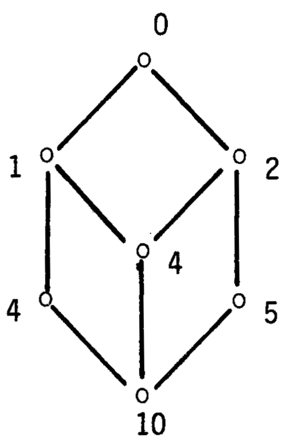
. sur le plus petit treillis non modulaire, des applications strictement monotones vérifiant l'une des hypothèses H1 ou H2 (les valeurs indiquées sont celles que prend v en chaque point du treillis).

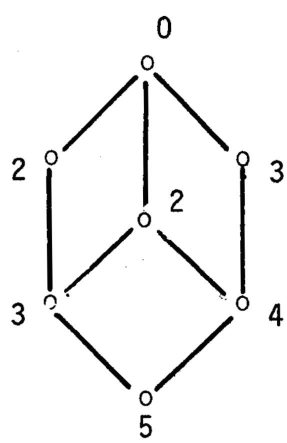


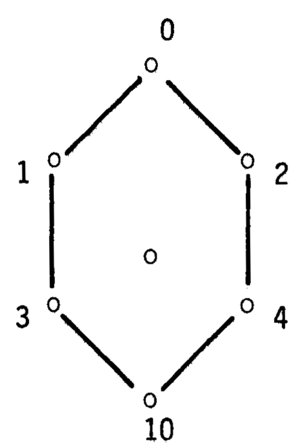


HYPOTHESE vérifiée : <u>H1</u> <u>H2</u>

. sur des treillis respectivement semi-modulaire inférieurement, semi-modulaire supérieurement, non semi-modulaires, des applications strictement monotones vérifiant également l'une quelconque des hypothèses H1 ou H2 .







HYPOTHESE vérifiée <u>H2</u> <u>H1</u> <u>H2</u>

2.- Les caractérisations définies plus haut peuvent-elles être obtenues avec des ensembles ordonnés quelconques ?

Nous apportons un élément de réponse en donnant, dans le cas des demi-treillis gradués, une caractérisation métrique de la semi-modularité.

La preuve de ce résultat repose sur le Théorème IV.2.2.1.
On y caractérise les applications v strcitement croissantes de T dans $\mathbb{R}$ telles que, pour tout couple (x,y) de $T \times T$, la variation de v suivant la s.e.c. $(x, x \vee y, y)$ est minimale.

Ces résultats nous ont suggéré de définir sur les sup-demi-treillis (resp. inf-demi-treillis) une notion de sup-valuation (resp. inf-valuation)

IV.2.1.- NOTATIONS.

- Soit T un sup-demi-treillis (de longueur finie), c'est-à-dire un ordre tel que deux éléments quelconques x et y de T admettent un supremum noté $x \vee y$.
- Soit v une application strictement croissante de T dans l'ensemble des nombres réels $\mathbb{R}$

$$x \prec y \text{ implique } v(x) < v(y)$$

En reprenant les notations de IV-1, pour toute s.e.c., nous définissons :

$$\Delta^+([x_i]_0^n) = \sum_{i \in I^-} [v(x_{i+1}) - v(x_i)]$$

$$\Delta^-([x_i]_0^n) = \sum_{i \in I^+} [v(x_i) - v(x_{i+1})]$$

$$\Delta([x_i]_0^n) = \Delta^+([x_i]_0^n) + \Delta^-([x_i]_0^n)$$

Notons :

$$\delta(x,y) = \operatorname*{Inf}_{C \in C(x,y)} \Delta(C)$$

Il a été démontré, au paragraphe IV-1-2, que, lorsqu'elle est définir à partir d'une application v strictement croissante sur T, δ est une distance sur T.

Nous étudions, par la suite, les propriétés de δ et des conditions pour qu'elle se confonde avec la distance habituelle notée d , définie sur le graphe de couverture G correspondant au sup-demi-treillis T par $d(x,y)$ = longueur du plus court chemin entre x et y , pour tout (x,y) de $T \times T$.

IV.2.2.- CARACTERISATION DE s.e.c. v-MINIMALE DANS T.

IV.2.2.1.- THEOREME.

Soit T *un sup-demi-treillis, soit* v *une application strictement croissante, définie sur* T , *à valeurs dans* $\mathbb{R}$.
Une condition nécessaire et suffisante pour que la s.e.c. $C_v(x,y) = (x, x \vee y, y)$ *soit* v-*minimale est que, pour tout triplet* (x,y,z) *d'éléments de* T *tels que :*

$$z \lneq x \quad \text{et} \quad z \lneq y$$

la condition suivante soit vérifiée.:

$$v(x \vee y) + v(z) \leq v(x) + v(y)$$

DEMONSTRATION.

. La condition est nécessaire

En effet, $C_v(x,y)$ est v-minimale

$\iff \forall (x,y) \in T \times T \quad , \quad \Delta^-(C_v(x,y)) \leq \Delta^-(C)$

pour tout $C \in \mathcal{C}(x,y)$

$\iff \quad v(x \vee y) - v(y) \leq \Delta^-(C)$

En particulier, pour la séquence (x,z,y) , pour tout z vérifiant la condition $z \lneq x$, $z \lneq y$,

$$v(x \vee y) - v(y) \leq v(x) - v(z)$$

soit

$$v(x \vee y) + v(z) \leq v(x) + v(y) \qquad \text{c.q.f.d.}$$

. <u>La condition est suffisante</u>

Soit, en effet, $(x,y) \in T \times T$, un couple quelconque d'élements de T ;

a) Supposons qu'il existe un élément z de T vérifiant $z \not\leq x$ et $z \not\leq y$ (nous étudierons ultérieurement le cas où un tel z n'existe pas).

On montre que $C_v(x,y)$ est v-minimale par récurrence sur la longueur des séquences reliant x à y.

La propriété est évidente pour des séquences de longueur égale à 1 (x et y comparables), ainsi que pour des séquences de longueur égale à 2 . En effet, soit t un élément de T, soit C la séquence (x,t,y)

- Si $t \not\leq x$, $t \not\leq y \Longrightarrow v(x \vee y) - v(y) \leq v(x) - v(t)$

$$\Longrightarrow \Delta^-(C_v(x,y)) \leq \Delta^-(C)$$

- Si $t \not\geq x$, $t \not\geq y$, alors $t \not\geq x \vee y$

$$\Longrightarrow v(t) \geq v(x \vee y)$$

$$\Longrightarrow v(t) - v(y) \geq v(x \vee y) - v(y)$$

$$\Longrightarrow \Delta^-(C) \geq \Delta^-(C_v(x,y))$$

Montrons que la propriété reste vraie pour des séquences de longueur 3 . Soit $C = (x,t,u,y)$, $t \not\geq x$, $u \not\leq t$, $u \not\leq y$ une séquence de longueur 3 (la démonstration serait analogue si $t \not\leq x$, $t \not\leq u$, $u \not\geq y$).

$$t \not\geq x \Longrightarrow y \vee t \not\geq x \vee y \Longrightarrow v(t \vee y) \geq v(x \vee y)$$

$$\Longrightarrow \Delta^-(C_v(x,y)) = v(x \vee y) - v(y) \leq v(t \vee y) - v(y)$$

$$\Longrightarrow \Delta^-(C_v(x,y)) \leq v(t \vee y) - v(t) + v(t) - v(y)$$

Or : $\qquad v(t \vee y) - v(t) \leq v(y) - v(u)$ par hypothèse

$$(u \lneq t \quad \text{et} \quad u \not\leq y)$$

$$\Longrightarrow \quad \Delta^-(C_v(x,y)) \leq v(y) - v(u) + v(t) - v(y)$$

$$\Longrightarrow \quad \Delta^-(C_v(x,y)) \leq v(t) - v(u) = \Delta^-(C)$$

La propriété est donc vérifiée par les séquences de longueur 3.

Supposons, maintenant que, pour tout couple $(x,y) \in T \times T$, la séquence $C_v(x,y)$ soit v-minimale par rapport à toute séquence de longueur n-1 reliant x à y . Soit $C(x,y) = [x_i]_0^n$, avec $x_0 = x$ et $x_n = y$, une séquence de longueur n :

$$\Delta^-(C(x,y)) = \Delta^-([x_i]_0^{n-1}) + \Delta^-((x_{n-1},y))$$

$$\geq \Delta^-(C_v(x,x_{n-1})) + \Delta^-((x_{n-1},y))$$

En appliquant l'hypothèse de récurrence :

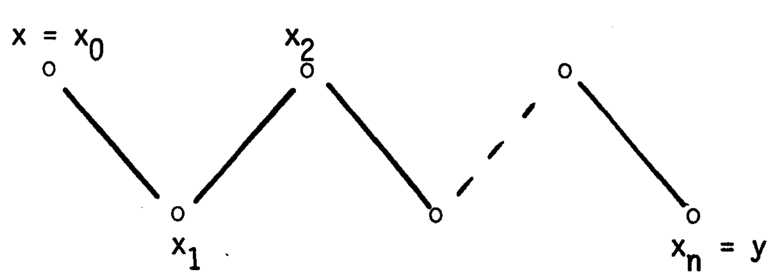
$$\Longrightarrow \quad \Delta^-(C(x,y)) \geq \Delta^-((x, x \vee x_{n-1}, x_{n-1}, y))$$

$$\geq \Delta^-((x, x \vee y, y)) = \Delta^-(C_v(x,y))$$

en appliquant le résultat obtenu pour les séquences de longueur 3 .

<u>c.q.f.d.</u>

b) S'il n'existe pas de z tel que $z \leqslant x$ et $z \leqslant y$, la s.e.c. $C_v(x,y)$ est encore v-minimale pour tout $(x,y) \in T \times T$.
En effet, si $[x_i]_0^n$ est une séquence de longueur n reliant x à y :



On peut appliquer au triplet (x_0, x_1, x_2) le raisonnement précédent qui permet d'obtenir :

$$\Delta^-((x_0,x_1,x_2)) \geq \Delta^-(x_0,x_0 \vee x_2,x_2)$$

En réitérant ce raisonnement, on est ramené au cas des s.e.c. de longueur 2 pour lesquelles le résultat est évident.

Par dualité, on obtient un théorème analogue pour les inf-demi-treillis.

IV.2.3.- VALUATION DANS LES DEMI-TREILLIS.

Notons d la distance définie dans le graphe de couverture associée à T par :

$d(x,y)$ = longueur du plus court chemin entre x et y.

Nous pouvons alors généraliser comme suit la notion de valuation dans les treillis au cas de demi-treillis.

IV.2.3.1.- PROPOSITION.

Soit v *une application strictement croissante d'un sup-demi-treillis* T *dans* $\mathbb{R}$ *; les propriétés suivantes sont équivalentes :*

(1) $d(x,y) = d(x,x \vee y) + d(x \vee y,y)$ *pour tout* $(x,y) \in T \times T$

(2) $d(x,y) = 2v(x \vee y) - v(x) - v(y)$ *pour tout* $(x,y) \in T \times T$
($C_v(x,y)$ *est v-minimale).*

(3) $2v(x \vee y) - v(x) - v(y)$ *est une distance sur* T.

(4) *pour tout* $x,y,z \in T$, $v(x \vee y) + v(z) \leq v(x \vee z) + v(z \vee y)$

(5) *pour tout* $x,y,z \in T$ *avec* $z \lessdot x$

$$v(x \vee y) + v(z) \leq v(x) + v(z \vee y)$$

(6) *pour tout* $x, y, z \in T$ *avec* $z \lessgtr x$, $z \lessgtr y$

$$v(x) + v(y) \geq v(z) + v(x \vee y)$$

Si, de plus T *est un treillis, ces conditions sont équivalentes à*

(7) *pour tout* $x, y \in T$, $v(x \vee y) + v(x \wedge y) \leq v(x) + v(y)$

DEMONSTRATION.-

. (1) $\Longrightarrow$ (2) $\Longrightarrow$ (3) .

Les démonstrations ont été réalisées par ailleurs lorsque T est un treillis. Elles se transposent facilement au cas d'un sup-demi-treillis.

. (3) $\Longrightarrow$ (4)

En appliquant l'inégalité triangulaire à $2v(x \vee y) - v(x) - v(y)$, on obtient :

$$2v(x \vee y) - v(x) - v(y) \leq 2v(x \vee y)\ )\ v(x) - v(z) + 2v(y \vee z) - v(y) - v(z)$$

soit encore :

$$v(x \vee y) + v(z) \leq v(x \vee z) + v(y \vee z)$$

. (4) $\Longrightarrow$ (5) et (5) $\Longrightarrow$ (6) sont évidents .

- L'équivalence des propriétés (1) et (6) a fait l'objet du théorème IV.2.2.1.

- (7) $\Longrightarrow$ (2) La démonstration a été réalisée au chapitre I.

Si l'on suppose, maintenant, que T est un treillis, il résulte immédiatement des résultats énoncés en IV.1, que les propriétés (1) et (7) sont équivalentes.

REMARQUE :

G. BORDES [5] a montré par ailleurs que la relation (5) impliquait (3). On obtient, par dualité, les propriétés relatives aux applications strictement croissantes sur les inf-demi-treillis.

La proposition précédente permet de généraliser de la façon suivante à des demi-treillis, la notion de valuation définie sur des treillis :

IV.2.3.2.- DEFINITION.

Une application strictement croissante d'un sup-demi-treillis (resp. inf-demi-treillis) dans $\mathbb{R}$ *sera dite Sup-valuation (resp. inf-valuation) si elle vérifie l'une quelconque des propriétés figurant dans la proposition IV.2.3.1. (resp. la proposition duale).*

Il en résulte :

IV.2.3.3.- COROLLAIRE.

Dans un treillis T *une application* v *strictement croissante est une sup-valuation (resp. inf-valuation) si et seulement si elle vérifie H1 (resp. H2).*

IV.2.3.4.- COROLLAIRE.

Dans un treillis T *, une application* v *strictement croissante est une valuation si et seulement c'est à la fois une sup-valuation et une inf-valuation.*

IV.2.4.- GRADUATION ET SEMI-MODULARITE DANS LES DEMI-TREILLIS.

Soit T un sup-demi-treillis.

Nous rappelons que T est semi-modulaire supérieurement [2] [4] si la condition suivante est vérifiée pour tout (x,y,z) de $T \times T \times T$

$$z \prec x \text{ et } z \prec y \Longrightarrow x \prec x \vee y \text{ et } y \prec x \vee y$$

en notant $\prec$ la relation, $x \prec y \Longleftrightarrow x$ précède y

Nous rappelons aussi qu'une application r de T dans $\mathbb{N}$ est appelée graduation ou fonction de rang si elle vérifie la condition suivante :

$$x \prec y \Longrightarrow r(y) = r(x) + 1$$

Les résultats que nous présentons dans ce paragraphe sont obtenus par application de ceux obtenus en IV-1 aux demi-treillis gradués.

LEMME 1.- [2] *Un sup-demi-treillis semi modulaire supérieurement est gradué.*

LEMME 2.- [2] [4] *Un sup-demi-treillis* T *est semi-modulaire supérieurement si et seulement si, pour tout* $x,y,z \in T$, $x \prec y$, $x \prec z$ *et* $y \| z$ *implique* $z \prec y \vee z$ *(*$y \| z$ *signifie* y *non comparable à* z*).*

IV.2.4.1.- THEOREME.

Un sup-demi-treillis T *est semi-modulaire supérieurement si et seulement s'il est gradué et si la fonction de rang est une sup-valuation, i.e. vérifie, pour tout* $x,y,z \in T$ *avec*

$$z \preceq x$$
$$z \preceq y$$

la condition :

$$(1) \qquad r(x \vee y) + r(z) \leq r(x) + r(y)$$

DEMONSTRATION.-

- Condition suffisante :

Soient $x,y,z \in T$ tels que :

$$z \prec x \qquad \text{et} \qquad z \prec y$$

$\Longrightarrow$ $r(x) = r(y) = r(z) + 1$

$r(x \vee y) - r(y) \leq r(x) - r(z) = 1$

$\Longrightarrow$ $r(x \vee y) - r(y) = 1$ (le cas $r(x \vee y) - r(y) = 0$ est impossible

$\Longrightarrow$ $y \prec x \vee y$

On montre de même que $x \prec x \vee y$

$\Longrightarrow$ T est semi-modulaire supérieurement.

c.q.f.d.

- Condition nécessaire :

Soient $x,y,z \in T$ tels que :

$$z \not\leqslant x \qquad \text{et} \qquad z \leqslant y$$

La démonstration se fait par récurrence sur la longueur de la s.e.c. reliant z à x.

En vertu du lemme 1 , T est gradué par une fonction de rang r.

1°) Supposons que $z \prec x$:

$\Longrightarrow$ $r(x) - r(z) = 1$

$z \prec x \Longrightarrow x \not\geqslant z \vee y = y$

. Si $x \leqslant z \vee y \Longrightarrow x \vee y = z \vee y = y \Longrightarrow x \leqslant y$

$\Longrightarrow$ $r(x \vee y) - r(y) = 0 < r(x) - r(z) = 1$

. Si $x \,||\, z \vee y \Longrightarrow x \vee y \succ z \vee y = y$, en vertu du lemme 2

$\Longrightarrow$ $r(x \vee y) - r(y) = 1$

$\Longrightarrow$ $r(x \vee y) - r(y) = r(x) - r(z)$

$\Longrightarrow$ l'inégalité (1) est bien vérifiée dans les 2 cas.

2°) Supposons alors qu'il existe une s.e.c. $x_1,\dots,x_n$ telle que :

$$x \succ x_1 \succ x_2 \succ \dots \succ x_n \succ z$$

et supposons l'inégalité (1) vérifiée à l'ordre $n-1$, c'est-à-dire que le plus court chemin pour aller de x_1 à y est celui passant par $x_1 \vee y$, soit

$$r(x_1 \vee y) - r(y) \leq r(x_1) - r(z)$$

$x_1 \prec x$ implique $x \not\succ x_1 \vee y$

. soit $x \not\prec x_1 \vee y \Longrightarrow x \vee y = x_1 \vee y$

$$\Longrightarrow r(x \vee y) - r(y) = r(x_1 \vee y) - r(y)$$
$$\leq r(x_1) - r(z) < r(x) - r(z)$$

. soit $x \| x_1 \vee y \Longrightarrow x_1 \vee y \prec x \vee y$, en vertu du lemme 2

$$\Longrightarrow r(x \vee y) - r(y) = r(x_1 \vee y) + 1 - r(y)$$
$$\leq r(x_1) + 1 - r(z) = r(x) - r(z)$$

c.q.f.d.

IV.2.4.2.- PROPOSITION (CARACTERISATION METRIQUE DE LA SEMI-MODULARITE).

Soit T *un sup-demi-treillis; soit* r *une fonction de graduation sur* T. *Une condition nécessaire et suffisante pour que* T *soit semi-modulaire supérieurement est que* $C_v(x,y)$ *soit* r*-minimale pour tout couple* (x,y) *de* $T \times T$.

DEMONSTRATION.-

. Condition nécessaire.

Supposons que $C_v(x,y)$ soit r-minimale pour tout couple $(x,y) \in T \times T$.

Soit z tel que $\left| \begin{array}{l} z \prec x \\ z \prec y \end{array} \right.$

$$\Delta^-(C_v(x,y)) \leq \Delta^-(C) \qquad \forall\, C \in \mathcal{C}(x,y)$$

$$\Longrightarrow r(x \vee y) - r(y) \leq r(x) - r(z)$$

$$\Longrightarrow r(y) < r(x \vee y) \leq r(y) + 1$$

$$\Longrightarrow \quad r(x \vee y) = r(y) + 1$$

$$\Longrightarrow \quad y \prec x \vee y$$

Le même raisonnement avec $\Delta^+(C)$ montre que $x \prec x \vee y$.

$\Longrightarrow$ T est semi-modulaire supérieurement. <u>c.q.f.d.</u>

. <u>Condition suffisante</u>

Réciproquement, supposons que T soit semi-modulaire supérieurement. Soit (x,y) un couple quelconque d'éléments de T .

a) S'il existe $z \in T$ tel que $z \prec x$ et $z \prec y$

$\Longrightarrow$ x et $y \prec x \vee y$ (par hypothèse)

$\Longrightarrow$ $\Delta^-(C_\vee(x,y)) = 1 \leq \Delta^-(C) \quad \forall\, C \in \mathcal{C}(x,y)$

$\Longrightarrow$ $C_\vee(x,y)$ est r-minimale. <u>c.q.f.d.</u>

b) S'il n'existe pas de z précédant x et y , les théorèmes IV.2.2.1. et IV.2.4.1. permettent de démontrer que ces résultats restent vrais.

Par dualité, on obtient des énoncés comparables pour les inf-demi-treillis.

Il résulte du corollaire IV.2.4.2. et du résultat dual :

<u>IV.2.4.5.- COROLLAIRE.</u>

Soit T *un treillis; soit* r *une graduation sur* T . *Une condition nécessaire et suffisante pour que* T *soit modulaire est que, pour tout couple* (x,y) *de* $T \times T$, *les séquences* $C_\vee(x,y)$ *et* $C_\wedge(x,y)$ *soient r-minimales.*

On retrouve donc la caractérisation des treillis modulaires : Un treillis T est modulaire si et seulement s'il est gradué et si sa fonction de rang vérifie, pour tout (x,y) de $T \times T$:

$$r(x \vee y) + r(x \wedge y) = r(x) + r(y)$$

Signalons que la caractérisation métrique de la semi-modularité introduite à la proposition IV.2.4.2., vient d'être trouvée de manière indépendante par B. MONJARDET [21] , à la suite d'une étude faite sur les ensembles ordonnés filtrant supérieurement.

Signalons aussi que les ensembles ordonnés connexes semi-modulaires ont été étudiés par HASKIN et GUDDER [16] qui donnent d'autres caractérisations des ensembles ordonnés modulaires.

C H A P I T R E V

-=-=-=-=-=-=-=-=-=-

APPLICATION AUX PROBLEMES D'ANALYSE DES DONNEES

-=-=-=-=-=-=-=-=-=-

V.- APPLICATION AUX PROBLEMES D'ANALYSE DES DONNEES.

Soit T un tableau à double entrée ($\mathcal{O} \times \Omega$) fournissant les réponses d'un ensemble Ω d'individus à un ensemble $\mathcal{O}$ d'observateurs [20]. Nous considérons que chaque observateur définit une fonction caractéristique d'un sous-ensemble de Ω. Le tableau de données étant modélisé comme représentant le produit cartésien de ces sous-ensembles, cette hypothèse suppose que la matrice de données est codée à l'aide de deux symboles, en général 0 et 1.

Un tableau R de même nature, c'est-à-dire de la forme ($\mathcal{O} \times \Omega$), peut être considéré comme représentant un modèle du tableau T. Il convient alors de pouvoir mesurer le degré de ressemblance du tableau R par rapport à T : l'évaluation des quantités δ, δ^+, δ^-, J_R et J_T définies au chapitre I permet de résoudre ce problème (programmes DISTABLE et DISTARES).

Ces tableaux modèles de T peuvent être fournis par les algorithmes classiques [3] d'analyse factorielle et de classification hiérarchique : ceux-ci permettent de définir des classes homogènes de Ω : l'obtention de ces classes nous permet (algorithme RESUM1) de construire un tableau R associé à celles-ci et de mesurer le degré de ressemblance entre T et R.

Enfin, nous proposons (programme RESUM2) un algorithme original de construction d'une suite R_i de tableaux modèles de T dont la complexité va croissante, et dont nous extrayons les modèles intéressants.

V.1.- PROGRAMMES DISTABLE ET DISTARES

a) Pour mesurer le degré de ressemblance d'un modèle R et d'un tableau de données T , nous établissons par simulation une échelle pour la distance et les quantités qui lui sont associées (programme DISTABLE).

Soient t et t' deux observateurs quelconques de $\mathcal{O}$, Ω_1^t et $\Omega_1^{t'}$ les sous-ensembles des individus de Ω ayant répondu 1 respectivement à t et t' .

Avec l'information de SHANNON sur les événements $I(\Omega_1^t) = \text{Log} \dfrac{1}{P(\Omega_1^t)}$, la distance δ entre Ω_1^t et $\Omega_1^{t'}$ s'écrit :

$$\delta(\Omega_1^t , \Omega_1^{t'}) = \text{Log}_2 \frac{P^2(\Omega_1^t \cup \Omega_1^{t'})}{P(\Omega_1^t)\ P(\Omega_1^{t'})}$$

De même :

$$\delta^+(\Omega_1^t , \Omega_1^{t'}) = \text{Log}_2 \frac{P(\Omega_1^t \cup \Omega_1^{t'})}{P(\Omega_1^{t'})}$$

Il vient alors :

$$d(T,R) = \sum_{(t,t')\in \mathcal{O}\times\mathcal{O}} \lambda(t,t')\ \delta(\Omega_1^t , \Omega_1^{t'})$$

$$d^+(T,R) = \sum_{(t,t')\in \mathcal{O}\times\mathcal{O}} \lambda(t,t')\ \delta^+(\Omega_1^t , \Omega_1^{t'})$$

$$d^-(T,R) = d(T,R) - d^+(T,R)$$

$$I(T) = \sum_{t\in\mathcal{O}} \lambda(t)\ I(\Omega_1^t)$$

$$I_R(T,R) = I(T) - d^-(T,R)$$

$$I_T(T,R) = I(T) + d^+(T,R)$$

$$I(R) = I_R(T,R) + d^+(T,R)$$

où λ est un système de pondération sur les observateurs de $\mathcal{O}$.

b) On construit un modèle R de T en introduisant un pourcentage fixé k de modifications des réponses des individus de Ω aux observateurs de $\mathcal{O}$, à l'aide d'une fonction aléatoire.
On peut alors calculer $d(T,R)$, $d^+(T,R)$ et toutes les autres quantités énumérées plus haut.

Nous opérons ainsi pour différentes valeurs de k, ce qui nous permet d'établir une échelle pour toutes les quantités citées ci-dessus.

Etant donné alors un tableau T et un modèle R de T, le programme DISTARES nous permet de déterminer toutes les quantités utiles à la comparaison de T et R.

c) Structure du programme DISTABLE.

α) Lecture de T

β) On se fixe une valeur de k

γ) Pour chaque observateur t appartenant à $\mathcal{O}$, on simule un nouvel observateur t' et par suite un nouveau tableau R

η) On calcule $d(T,R)$ et toutes les autres quantités.

V.2.- PROGRAMME RESUM1.

a) Etant donné un tableau de données T du type $(\mathcal{O} \times \Omega)$ codé en $(0,1)$, soit une partition Π de Ω pour laquelle chaque classe est "homogène", c'est-à-dire "correspond à un sous-ensemble d'observateurs ayant le même comportement pour l'ensemble des individus constituant cette classe", cette partition ayant été obtenue par une méthode quelconque [3] [15].

Nous construisons un modèle du tableau initial, associé à ces classes, dans la mesure où nous recherchons les observateurs caractéristiques de la classe.

b) . En reprenant les définitions de [15] , le principe de la méthode est le suivant :

Etant donnée une classe homogène N , soient t un observateur quelconque de $\mathcal{O}$, N_1^t et N_0^t l'ensemble des éléments de N répondant respectivement 1 à t , et 0 à t .

L'observateur t sera dit <u>caractéristique de N</u> si

$$\mathrm{Min}(\delta(N_1^t , t_1^N) \quad , \quad \delta(N_0^t , t_0^N) \quad , \quad \delta(N_1^t , N_1^{t_{N_S}}))$$

$$= \delta(N_1^t , t_1^N) \quad \text{ou} \quad \delta(N_0^t , t_0^N)$$

où t_1^N et t_0^N sont deux observateurs fictifs auxquels tous les individus de N répondraient respectivement 1 et 0, $N_1^{t_{N_S}}$ est l'ensemble des éléments de N répondant 1 à un observateur non significatif associé à l'observateur t et à la classe N .

La distance que l'on utilise est celle associée à l'information de SHANNON sur les événements (voir chap. III).

Une étude faite dans [15] montre que δ mesure le degré de ressemblance entre N_1^t et t_1^N (ou entre N_0^t et t_0^N) .

. Si nous considérons donc un observateur quelconque $t \in \mathcal{O}$, si t est caractéristique de N par t_1^N (resp. par t_0^N) , alors nous remplaçons dans le tableau de données T , l'ensemble $N_1^t \cup N_0^t$ par t_1^N (resp. t_0^N) .

Si t n'est pas caractéristique de N , alors nous n'effectuons aucune modification.

Nous opérons ainsi successivement pour tous les observateurs de $\mathcal{O}$

. Nous construisons donc un modèle simplifié du tableau initial associé aux classes.

L'application du programme DISTARES permet alors de calculer toutes les quantités utiles à la comparaison des deux tableaux.

c) Structure du programme RESUM1

α) Lecture des données

β) Pour l'ensemble des observateurs, déterminer ceux qui sont caractéristiques d'une classe fixée, modifier éventuellement les données.

γ) S'il reste une classe à traiter, retourner en β) , sinon c'est terminé.

V.3.- CONSTRUCTION D'UNE SUITE DE RÉSUMÉS (PROGRAMME RESUM2).

Etant donné un tableau T du type $(\mathcal{O} \times \Omega)$ codé en (0,1) , nous construisons à partir de T , une suite de modèles (n modèles s'il y a n observateurs dans $\mathcal{O}$) , parmi lesquels il sera possible de déterminer les "bonnes" synthèses de T .

a) Le principe de construction du modèle R_{k+1} à partir du modèle R_k et du tableau T , est le suivant :

. Prenons un observateur t_{i_0} encore disponible. On lui associe un tableau $T_{t_{i_0}}$ du type $(\mathcal{O} \times \Omega)$ de la manière suivante :

Pour chaque observateur t_i appartenant à $\mathcal{O}$, on calcule :

$$\mathrm{Min}[\delta(\Omega_1^{t_{i_0}}(T) , \Omega_1^{t_i}(T)) , \delta(\Omega_1^{t_i}(T) , \Omega_1^{t_i}(R_k))]$$

où $\Omega_1^{t_{i_0}}(T)$ est l'ensemble des individus de Ω ayant répondu 1 à l'observateur t_{i_0} dans le tableau T , la distance utilisée étant toujours celle construite à partir de l'information de SHANNON sur les événements .

Si ce minimum est $\delta(\Omega_1^{t_{i_0}}(T) , \Omega_1^{t_i}(T))$, alors nous remplaçons dans le tableau R_k les réponses à t_i par les réponses réelles à t_{i_0} se trouvant dans le tableau T ; sinon, nous ne changeons rien.

Nous construisons ainsi un tableau $T_{t_{i_0}}$ associé à t_{i_0} ; nous calculons $d(T,T_{t_{i_0}})$.

. Toujours à partir des tableaux T et R_k , nous recommençons les opérations précédentes pour tous les observateurs t encore disponibles.

. Soit alors t_k l'observateur tel que :

$$d(T,T_{t_k}) = \underset{\substack{t \in \mathcal{O} \\ t \text{ encore disponible}}}{\text{Min}} d(T,T_t)$$

Nous recherchons les modifications (changements de 0 en 1, ou de 1 en 0) à effectuer sur les réponses à l'observateur t_k des individus de Ω , de manière à diminuer au maximum la distance $d(T,T_{t_k})$.

Ceci nous permet de déterminer le modèle R_{k+1} (le meilleur T_{t_k} modifié).

. On élimine alors l'observateur qui a permis de construire le tableau R_{k+1} .

On opère ainsi jusqu'à l'épuisement des observateurs disponibles.

. L'initialisation de l'algorithme se fait en prenant pour R_0 un tableau du type $\mathcal{O} \times \Omega$ tel que, quelque soit l'observateur t appartenant à $\mathcal{O}$, tous les individus de Ω répondent 1 .

b) <u>Structure du programme RESUM2</u>

α) Lecture des données

β) Pour chaque observateur t de $\mathcal{O}$ encore disponible, construction "fictive" d'un tableau T_t de type $(\mathcal{O} \times \Omega)$, puis construction réelle du modèle en prenant le meilleur (modifié) au sens de d :
il correspond à un observateur t_{i_1} .

γ) Suppression de l'observateur t_{i_1}

λ) S'il reste encore un observateur disponible on retourne en β), sinon on s'arrête.

V.4.- APPLICATION.

Nous présentons ici les résultats obtenus pour les différents programmes, sur le tableau des données $(O \times \Omega)$ qui suit et qui comporte 40 individus et 40 observateurs.

Il est codé en (0,1) , les 0 étant représentés par des . , les 1 par des I .

```
40 CARACTERES   40 INDIVIDUS
 1 A1   ........II...............IIIIIIIIIIII..I
 2 A2   .......III.................IIIIIIIIIIIII
 3 A3   ..........II.............III..IIIIIIIIII
 4 A4   ..............I....I.....IIIIIIIII.IIIII
 5 A5   ....I.....I....II........IIII.II.IIIIIII
 6 A6   .I.I.....II...............IIIIIIIIIIIIII
 7 A7   .......II....I......I...IIII..II.IIIII.I
 8 A8   .I.I.I...........II......I.I.IIII.IIIIII
 9 B1   .....II....III.II..IIIIII.....IIII..I...
10 B2   ....III.....I.I.IIIIIIII.II.I.I.........
11 B3   ....II......IIII..IIII.IIII....I...I..I.
12 B4   ....II....I.IIIIII.III.IIII.....II......
13 B5   ........I...IIIIIIIIIIIIIII......II.....
14 B6   ..II...I....II.II.II.IIIIII........I.I..
15 B7   .I.....I.......IIIIIIIIIIII............I
16 C1   ..............III........III.IIIIIIII...
17 C2   ..................II..I..IIIII.IIIIII...
18 C3   .I.I.I.I.................IIII.IIIIIII...
19 C4   ...I.....I....I..........IIIIII.IIII.I..
20 C5   ............II.....I.I...IIIIIIII.III...
21 C6   ........I.I.I.I..........IIIIIII.IIII...
22 C7   ....II...................IIIIII.IIIII..I
23 D1   III.IIIII...........IIII...III......III.
24 D2   I.IIIIIIII...........II......I..........
25 D3   IIII.II.II.I.I.I........................
26 D4   I.IIIIIIII...................II.I.I.....
27 D5   IIIIIIIII..............................I
28 D6   IIIIIII.II.............II...............
29 E1   II.I.IIIIIII........I...................
30 E2   I.IIIIIIIIII.......II..................I
31 E3   III.IIII.III.......IIIII......II........
32 E4   IIIIIIIII..II...........................
33 E5   IIIII.III.II....II......I....I..........
34 E6   IIIIIIIIIII........I......I.......I.....
35 F1   .........................II....I....I..I
36 F2   .........................I.....I........
37 F3   ..I....I.................II.....I.......
38 F4   ...........I....I.........I..........I.I
39 F5   ....I....................II........I....
40 F6   .........................II...I.........
```

<u>TABLEAU T</u>

a) La simulation de ce tableau par DISTABLE permet d'obtenir le tableau suivant :

k	d	d^+	d^-	I_R	I_T
0,05	4,34798	2,884	1,46399	63,40862	67,75659
0,1	8,78619	5,75605	3,03015	61,78455	70,57074
0,15	13,09	8,83222	4,25777	60,58588	73,67587
0,2	17,46832	11,41085	6,05747	58,37976	75,84808
0,25	21,77652	14,47472	7,3018	57,13423	78,91075
0,3	25,87427	16,94859	8,92567	55,43672	81,31099
0,35	30,52184	19,8726	10,64923	53,43758	83,95941
0,4	34,57280	22,57031	12,00249	51,63051	86,20331
0,45	39,2887	25,52055	13,76822	49,71535	89,00412
0,5	43,07722	28,2905	14,78673	48,2533	91,33052

TABLEAU V.4.1.

Nous remarquons que :

- d^+ croit pratiquement de manière linéaire avec k
- d^+ croit plus rapidement que d et d^-

b) Le programme RESUM1 permet d'obtenir le modèle R ci-dessous du tableau de départ T , à partir des classes suivantes :

C_1 : A1 , A2 , A3 , A4 , A5 , A6 , A7 , A8 , B1 , B2

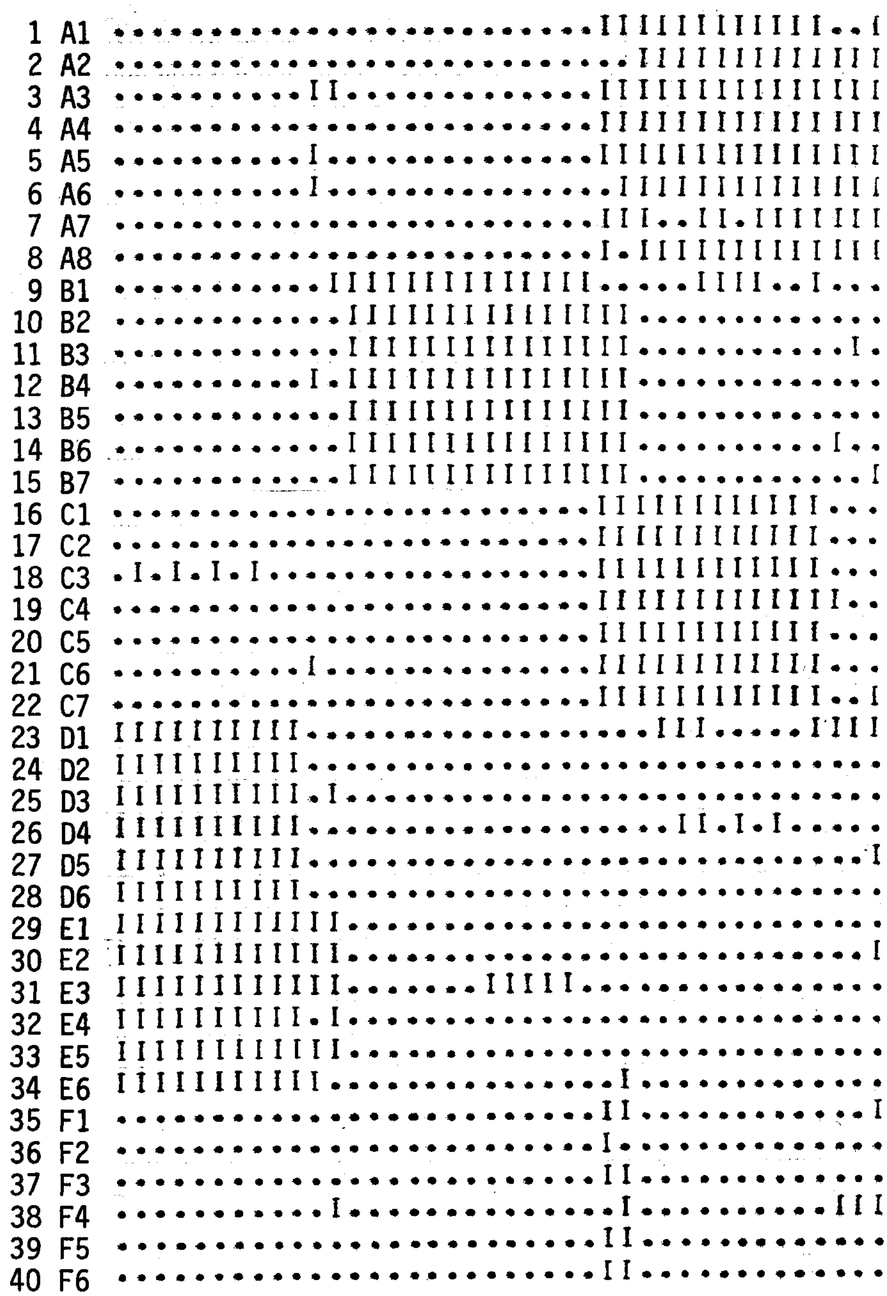
C_2 : B3 , B4

C_3 : B5 , B6 , B7 , C1 , C2 , C3 , C4 , C5 , C6 , C7 , D1 , D2 , D3

C_4 : D4 , D5

C_5 : D6 , E1 , E2 , E3 , E4 , E5 , E6 , F1 , F2 , F3

C_6 : F4 , F5 , F6



MODELE R

c) Le programme DISTARES nous donne les valeurs suivantes pour d , d^+ , d^- , I_R et I_T :

$d(T,R)$ = 8,07473

$d^+(T,R)$ = 5,29221

$d^-(T,R)$ = 2,78252

$I_R(T,R)$ = 62,21353

$I_T(T,R)$ = 70,28826

INTERPRETATION.

. $d^+(T,R)$ représente une quantité d'information "parasite" apportée par le modèle R et qui ne sert à rien dans la connaissance du tableau T .

. $d^-(T,R)$ est la quantité d'information du tableau T non expliquée par le modèle R.

. $I_R(T,R)$ est la quantité d'information qu'apporte le modèle R sur la connaissance du tableau T .

. $I_T(T,R)$ est la quantité totale d'information apportée par le modèle R et qui comprend une certaine quantité inutile dans la connaissance de T .

Les valeurs numériques de ces quantités, comparées à celles données par le tableau de simulation obtenu par DISTABLE (Tableau V.4.1) montrent que ce modèle R correspond à une modification à moins de 10 % des réponses des individus de Ω aux observateurs de $\mathcal{O}$ (changements de 0 en 1 , ou de 1 en 0).

d) α) Le programme RESUM2 a également été appliqué sur ce tableau T de type (40 x 40).
Nous établissons la courbe représentant les variations de l'information de ressemblance I_R en fonction du modèle R_i ; nous repérons sur cette courbe les modèles intéressants, car correspondant à des positions d'équilibre entre les variations de l'information parasite et de l'information non expliquée, c'est-à-dire encore entre la complexité du modèle et la connaissance qu'il nous apporte sur le tableau de départ T .

Nous présentons également deux de ces modèles établis à deux étapes très différentes de l'algorithme.

β) NOTATIONS

- Dans le modèle R4 , la ligne :

D1 ++++

signifie que tous les individus de Ω répondent 1 à l'observateur D1.

- Nous faisons intervenir en tête des tableaux, deux colonnes :
  . la première colonne nous donne la liste des observateurs de $\mathcal{O}$, des permutations sur les observateurs étant effectuées au cours de l'algorithme.
  . la deuxième colonne nous fournit la liste des observateurs ayant permis la construction du modèle.

γ)

```
D1   ++++
A2   A1   ..........................IIIIIIIIIIIIIII
A3   A1   ..........................IIIIIIIIIIIIIII
A4   A1   ..........................IIIIIIIIIIIIIII
A5   A1   ..........................IIIIIIIIIIIIIII
A6   A1   ..........................IIIIIIIIIIIIIII
A7   A1   ..........................IIIIIIIIIIIIIII
A8   A1   ..........................IIIIIIIIIIIIIII
C1   A1   ..........................IIIIIIIIIIIIIII
C2   A1   ..........................IIIIIIIIIIIIIII
C3   A1   ..........................IIIIIIIIIIIIIII
C4   A1   ..........................IIIIIIIIIIIIIII
C5   A1   ..........................IIIIIIIIIIIIIII
C6   A1   ..........................IIIIIIIIIIIIIII
C7   A1   ..........................IIIIIIIIIIIIIII
A1   A1   ..........................IIIIIIIIIIIIIII
B2   B5   ..........IIIIIIIIIIIIIIIIII.............I
B3   B5   ..........IIIIIIIIIIIIIIIIII.............I
B4   B5   ..........IIIIIIIIIIIIIIIIII.............I
B5   B5   ..........IIIIIIIIIIIIIIIIII.............I
B6   B5   ..........IIIIIIIIIIIIIIIIII.............I
B7   B5   ..........IIIIIIIIIIIIIIIIII.............I
B1   B5   ..........IIIIIIIIIIIIIIIIII.............I
D2   E4   IIIIIIIIIIII..............................
D3   E4   IIIIIIIIIIII..............................
D4   E4   IIIIIIIIIIII..............................
D5   E4   IIIIIIIIIIII..............................
D6   E4   IIIIIIIIIIII..............................
E1   E4   IIIIIIIIIIII..............................
E2   E4   IIIIIIIIIIII..............................
E3   E4   IIIIIIIIIIII..............................
E4   E4   IIIIIIIIIIII..............................
E5   E4   IIIIIIIIIIII..............................
E6   E4   IIIIIIIIIIII..............................
F1   F1   ..........................II............II.
F2   F1   ..........................II............II.
F3   F1   ..........................II............II.
F4   F1   ..........................II............II.
F5   F1   ..........................II............II.
F6   F1   ..........................II............II.
```

<u>MODELE R4</u>

$d(T,R4) = 25,37244$

$I_R(T,R4) = 56,18454$

D'après le tableau V.4.1., R4 est un modèle de T où le pourcentage de modifications est compris entre 25 % et 30 %.

Nous retrouvons pratiquement sur R4 la séparation des observateurs en 4 grandes classes :

- C'_1 = (A1, A2, A3, A4, A5, A6, A7, A8, C1, C2, C3, C4, C5, C6, C7)
- C'_2 = (B1, B2, B3, B4, B5, B6, B7)
- C'_3 = (D1, D2, D3, D4, D5, D6, E1, E2, E3, E4, E5, E6)
- C'_4 = (F1, F2, F3, F4, F5, F6)

n)

```
A1   A1   .........I...............IIIIIIIIIIIIIII
A2   A1   .........I...............IIIIIIIIIIIIIII
A3   A1   .........I...............IIIIIIIIIIIIIII
A4   A1   .........I...............IIIIIIIIIIIIIII
A6   A1   .........I...............IIIIIIIIIIIIIII
A5   A5   ....I.....I....II........IIII.II.IIIIIII
A7   A7   .......II....I......I...IIII..II.IIIII.I
A8   A8   .I.I.I...........II......I.I.IIII.IIIIII
B1   B1   .....II....III.II..IIIIII.....IIII..I...
B2   B2   ....III.....I.I.IIIIIIII.II.I.I.........
B3   B3   ....II......IIII..IIII.IIII....I...I..I.
B4   B5   ....II......IIIIIIIIII.IIII......I......
B5   B5   ....II......IIIIIIIIII.IIII......I......
B6   B6   ..II...I....II.II.II.IIIIII........I.I..
B7   B7   .I.....I.......IIIIIIIIIIII............I
C1   C1   ..............III........III.IIIIIIII...
C2   C2   ..................II..I..IIIII.IIIIII...
C3   C3   .I.I.I.I.................IIII.IIIIIII...
C4   C4   ...I.....I....I..........IIIIII.IIII.I..
C5   C5   ............II.....I.I...IIIIIIII.III...
C6   C6   ........I.I.I.I..........IIIIIII.IIII...
C7   C7   ....II...................IIIIII.IIIII..I
D1   D1   III.IIIII............IIII...III.....III.
D2   D2   I.IIIIIIII...........II......II.I.......
C4   D2   I.IIIIIIII...........II......II.I.......
C3   D3   IIII.II.II.I.I.I........................
D6   D6   IIIIIII.II.............II...............
E1   E1   II.I.IIIIIII........I...................
E2   E2   I.IIIIIIIIII.......II..................I
E3   E3   III.IIII.III.......IIIII......II........
D5   E4   IIIIIIIII..I...........................I
E4   E4   IIIIIIIII..I...........................I
E5   E5   IIIII.III.II....II......I....I..........
E6   F6   IIIIIIIIIII........I......I.......I.....
F1   F1   .........................II....I........
F2   F2   .........................I.....I........
F3   F3   ..I....I.................II.....I.......
F4   F4   ...........I....I.........I..........I.I
F5   F5   ....I....................II........I....
F6   F6   .........................II...I.........
```

<u>MODELE R33</u>

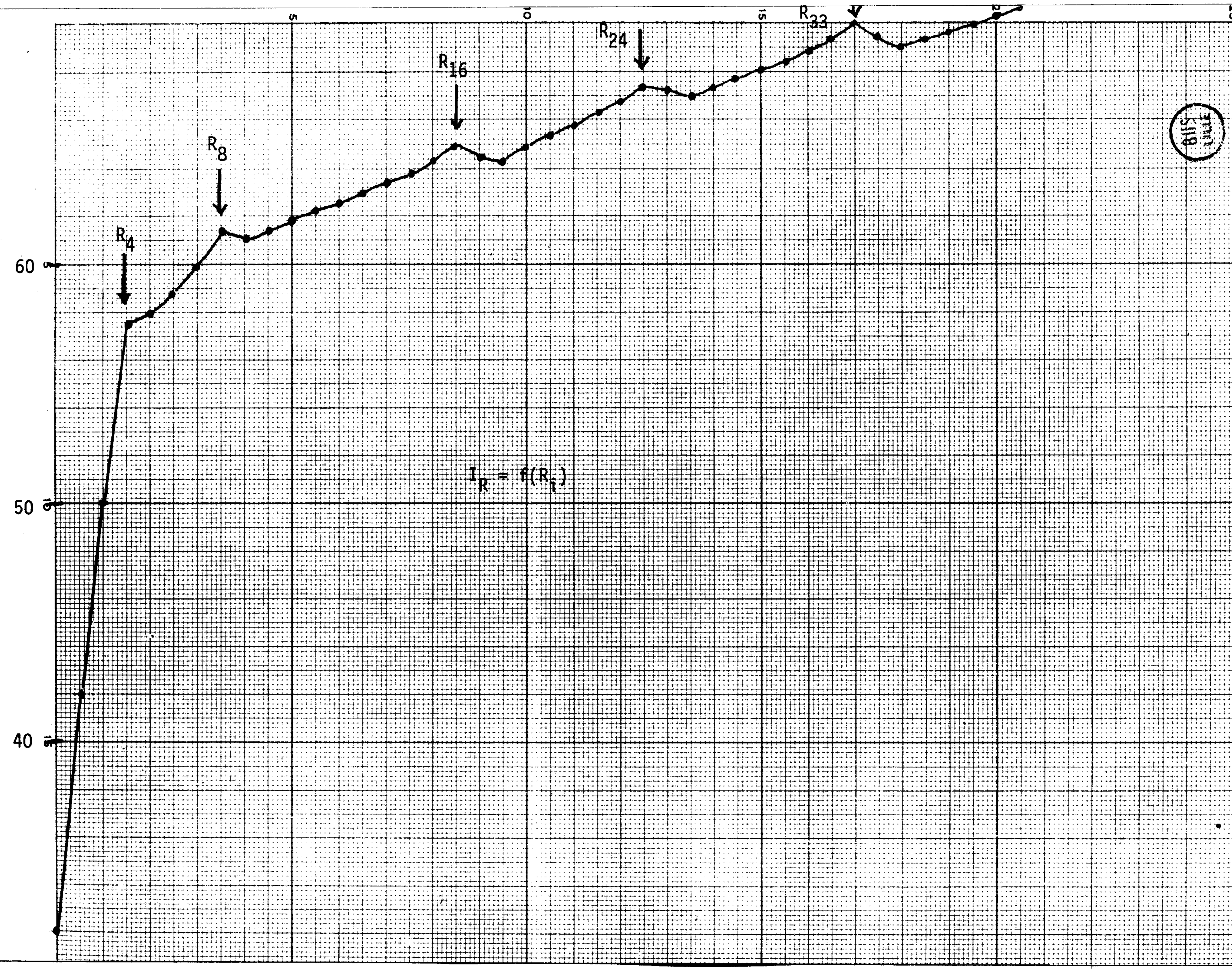
$d(T,R_{33}) = 3,98546$

$I_R(T,R_{33}) = 67,63055$

R_{33} est un modèle de T à moins de 5 %

Ce modèle est évidemment plus proche de T que R_4 , mais il est aussi plus compliqué.

e) Signalons que ces différents programmes ont été appliqués sur un tableau de données fournies par les biologistes, et correspondant aux réponses codées en (0,1) , de 175 souches de bactéries (ENTEROBACTER) à 199 tests.



60
50
40
5
10
15
20
25
R_4
R_8
R_{16}
R_{24}
R_{33}
$I_R = f(R_i)$

A N N E X E

Nous avons montré, au chapitre IV, que l'hypothèse H1 (resp. H2) est une condition nécessaire et suffisante pour que la s.e.c. $C_{\wedge}(x,y) = (x, x \wedge y, y)$ (resp. $C_{\vee}(x,y) = (x, x \vee y, y)$) soit v-minimale. Ce résultat permet, à partir d'hypothèses claires sur l'application v, de calculer de manière simple les valeurs des quantités δ, δ^+, δ^- (et J_R et J_T dans le chapitre I). En fait, les hypothèses H1 et H2 sont dans certains cas trop restrictives, et nous nous sommes intéressés à définir d'autres hypothèses sur v, plus larges que H1 et H2, mais qui permettent toujours d'expliciter facilement les quantités δ, δ^+, δ^-, J_R et J_T.

Nous nous posons la question :

Quelle condition doit vérifier une application v monotone décroissante de $(T,\leq)$ dans $\mathbb{R} \cup \{+\infty\}$ telle que, pour tout couple (x,y) de $T \times T$, l'une ou l'autre des s.e.c. $C_{\wedge}(x,y)$ ou $C_{\vee}(x,y)$ soit v-minimale ?

Explicitons cette condition pour l'ensemble des s.e.c. de longueur inférieure ou égale à 3, c'est-à-dire que, pour tout couple (x,y) de $T \times T$, nous traduisons l'énoncé en termes de Δ :

l'infimum de $\Delta(C_{\wedge}(x,y))$ et de $\Delta(C_{\vee}(x,y))$ est inférieur au Δ d'une s.e.c. quelconque de longueur 3.

On en déduit l'hypothèse notée H3.

Nous en déduisons de même les hypothèses H4 et H5 à partir de Δ^+ et Δ^-.

Enonçons les conditions utilisées par la suite :

Pour tout $(x,y,z) \in T \times T \times T$:

$$H3 : \operatorname{Inf}(v(x) + v(y) - 2\, v(x \vee y) \;,\; 2\, v(x \wedge y) - v(x) - v(y)) \leq \ldots$$

$$\ldots \leq \operatorname{Inf}(2\, v(y \wedge z) - 2\, v(x \vee z) + v(x) - v(y) \;,$$

$$2\, v(x \wedge z) - 2\, v(y \vee z) - v(x) + v(y))$$

$$H4 : \operatorname{Inf}(v(x) - v(x \vee y) \;,\; v(x \wedge y) - v(x)) \leq \ldots$$

$$\ldots \leq \operatorname{Inf}(v(y \wedge z) - v(x \vee z) \;,\; v(x \wedge z) - v(x) + v(y) - v(y \vee z))$$

$$H5 : \operatorname{Inf}(v(x) - v(x \vee y) \;,\; v(x \wedge y) - v(y)) \leq \ldots$$

$$\ldots \leq \operatorname{Inf}(v(x) - v(x \vee z) + v(y \wedge z) - v(y) \;,\; v(x \wedge z) - v(y \vee z))$$

Il est facile de montrer, à partir de la proposition I.2.3.1., que les *trois hypothèses H3, H4 et H5 sont équivalentes.*

PROPOSITION.-

Une condition nécessaire et suffisante pour que, pour tout couple (x,y) *de* $T \times T$ *, l'une ou l'autre des s.e.c.* $C_\wedge(x,y)$ *ou* $C_\vee(x,y)$ *soit v-minimale est que l'hypothèse H3 (ou H4 ou H5) soit vérifiée pour tout triplet* (x,y,z) *de* $T \times T \times T$

DEMONSTRATION.-

. La condition nécessaire est immédiate.

. La condition suffisante est vraie pour les s.e.c. de longueur 2 ou 3, la démonstration s'effectuant par récurrence sur la longueur de la s.e.c. suivant la technique du théorème I.5.2.1.

On peut affaiblir la condition en supposant seulement que z appartient à l'intervalle $[x,y]$, c'est-à-dire :

$$x \wedge y \leqslant z \leqslant x \vee y$$

Il vient alors :

PROPOSITION.-

Une condition nécessaire et suffisante pour que, pour tout couple (x,y) *de* $T \times T$ *, l'une ou l'autre des s.e.c.* $C_{\vee}(x,y)$ *ou* $C_{\wedge}(x,y)$ *soit v-minimale est que l'hypothèse H3 (ou H4 ou H5) soit vérifiée pour tout* $(x,y) \in T \times T$ *, et tout* $z \in [x,y]$

DEMONSTRATION.-

. La condition nécessaire est immédiate.

. La démonstration de la condition suffisante s'effectue en associant à un élément t quelconque de T , un élément $u = [t \wedge (x \vee y)] \vee (x \wedge y)$ de l'intervalle $[x,y]$ et en montrant que si H3 est vraie pour (x,y,u) , H3 est vraie pour (x,y,t)

B I B L I O G R A P H I E

[1] N.L. AGGARWAL, C.F. PICARD *"DES MESURES DE L'INFORMATION ET DE LEURS APPLICATIONS"*
Journées Lyonnaises des Questionnaires (1975), PARIS.
Publications du C.N.R.S., Groupe de Recherche 22, Structure de l'information pp. 31-44

[2] M. BARBUT, B. MONJARDET *"ORDRE ET CLASSIFICATION"*
Tomes 1 et 2, PARIS, HACHETTE, 1970.

[3] J.P. BENZECRI et collaborateurs. *"L'ANALYSE DES DONNEES"*
Tomes 1 et 2, PARIS, DUNOD, 1973

[4] G. BIRKHOF *"LATTICE THEORY"*
PROVIDENCE, Amer. Math. Society, Colloquium Publications, Vol. XXV, 1967

[5] G. BORDES *"METRIQUES BORNEES DEFINIES PAR DES VALUATIONS SUR UN DEMI-TREILLIS"*
Math. Sci. hum., 56, 1976

[6] G. COMYN, J. Cl. VAN DORPE *"VALUATION ET SEMI-MODULARITE DANS LES DEMI-TREILLIS"*
Math. Sci. hum., 56, 1976

[7] G. COQUET *"SEMI-GROUPES TOPOLOGIQUES ET APPLICATION A LA THEORIE DE L'INFORMATION"*
Publication de l'U.E.R. de Maths Pures et Appliquées de l'Université des Sciences et Techniques de LILLE-I

[8] B. FORTE et N. PINTACUDA *"INFORMATION FOURNIE PAR UNE EXPERIENCE"*
C.R.Acad. Sci., 266, Série A, 1968, pp. 242-244

[9] H. GREENBERG *"INTEGER PROGRAMMING"*
p. 177, Vol. 76, Math. in Science and Engineering, 1971, MONTEREY, Academic Press

[10] G. GRIMONPREZ
J. Cl. VAN DORPE
"INFORMATION GENERALISEE ET QUESTIONNAIRES"
Séminaire sur les Questionnaires,
Université de PARIS VI, 23 Janvier 1975

[11] G. GRIMONPREZ, J. LOSFELD
J. Cl. VAN DORPE
"PERTE ET GAIN D'INFORMATION SUR UNE SEQUENCE D'ELEMENTS COMPARABLES D'UN TREILLIS ET METRIQUE ASSOCIEE"
C.R. Acad. Sci. PARIS, 280, Série A
(14-4-1975), p. 957

[12] G. GRIMONPREZ,
J. Cl. VAN DORPE
"PERTE ET GAIN D'INFORMATION SUR UNE SEQUENCE D'ELEMENTS COMPARABLES D'UN TREILLIS ET METRIQUE ASSOCIEE - (APPLICATIONS)"
C.R. Ac. Sci, PARIS, 280, Série A,
(26-5-75-, p. 1381

[13] G. GRIMONPREZ
J. Cl. VAN DORPE
"CONSEQUENCES DE L'AXIOME DE MONOTONICITE. NOTION DE RESUME"
Journées Lyonnaises des Questionnaires
(1975), Publications du C.N.R.S.,
Groupe de Recherche 22, Structure de
l'information, pp. 71-106

[14] G. GRIMONPREZ
J. Cl. VAN DORPE
"DISTANCE DEFINIE PAR UNE APPLICATION MONOTONE SUR UN TREILLIS".
Math. Sci. Hum., 56, 1976

[15] G. GRIMONPREZ
THESE DE DOCTORAT DE 3° CYCLE.
LILLE, DECEMBRE 1976

[16] L. HASKINS, S. GUDDER
"HEIGHTON POSETS AND GRAPHS"
Discret Mathematics 2, 1972,
pp. 357-382

[17] J. KAMPE DE FERIET
B. FORTE
"INFORMATION ET PROBABILITE"
C.R. Acad. Sci., PARIS, 265, Série A,
1967, pp. 110-114, 142-146, 350-353

[18] J. KAMPE DE FERIET,
P. BENVENUTI,
"FORME GENERALE DE L'OPERATION DE COMPOSITION CONTINUE D'UNE INFORMATION"
C.R. Acad. Sci. PARIS, 269, Série A,
1969, pp. 529-534

[19] J. LOSFELD
"INFORMATION GENERALISEE ET RELATION D'ORDRE"
MARSEILLE-LUMINY, 1973,
Lect. Notes in Maths, 398, pp. 49-61.

[20] J. LOSFELD — *"INFORMATION FOURNIE PAR UN ENSEMBLE D'OBSERVATEURS ET APPLICATIONS AUX QUESTIONNAIRES ET A L'ANALYSE DES DONNEES"*

Thèse de Doctorat d'Etat, LILLE, Juin 1974

[21] B. MONJARDET — *"CARACTERISATIONS METRIQUES DES ENSEMBLES ORDONNES SEMI-MODULAIRES"*

Math. Sci. hum., 56, 1976

[22] S. PETOLLA, C.F. PICARD — *"AXIOMATIQUE ET SYSTEMATIQUE DES QUESTIONNAIRES"*

Actes des rencontres MARSEILLE-LUMINY, Lect. Notes in Maths, 398, 1973, pp. 126-151

[23] C.F. PICARD — *"GRAPHES ET QUESTIONNAIRES"*

Tomes 1 et 2, 1972, GAUTHIER VILLARS

[24] C.F. PICARD — *"PROPRIETE D'ADDITIVITE DE L'INFORMATION DE TYPE β"*

C.R. Acad. Sci., 282, Série A, (26-4-76), pp. 915