Un modèle mathématique de processus d'interrogation : les pseudoquestionnaires

Michel Terrenoire

N° d'ordre

THESES

présentées à

LA FACULTE DES SCIENCES DE GRENOBLE

pour obtenir

LE GRADE DE DOCTEUR ES SCIENCES MATHEMATIQUES

par

Michel Terrenoire

1ère thèse :

Un modèle mathématique de processus d'interrogation : les pseudoquestionnaires

2ème thèse :

PROPOSITIONS DONNEES PAR LA FACULTE

Thèses soutenues le 20 Octobre 1970 devant la commission d'examen

MM. J. KUNTZMANN	Président
C. F. PICARD	Examinateur
N. GASTINEL	Examinateur
J. C. BOUSSARD	Examinateur
J. FAURE	Invité

L I S T E D E S P R O F E S S E U R S

Doyen honoraire	:	Monsieur M. MORET
Doyen	:	Monsieur E. BONNIER

PROFESSEURS TITULAIRES

MM.	NEEL Louis	Physique Expérimentale
	KRAVTCHENKO Julien	Mécanique Rationnelle
	CHABAUTY Claude	Calcul différentiel et intégral
	BENOIT Jean	Radioélectricité
	CHENE Marcel	Chimie Papetière
	FELICI Noel	Electrostatique
	KUNTZMANN Jean	Mathématiques Appliquées
	BARBIER Reynold	Géologie Appliquée
	SANTON Lucien	Mécanique des Fluides
	OZENDA Paul	Botanique
	FALLOT Maurice	Physique Industrielle
	KOSZUL Jean-Louis	Mathématiques
	GALVANI Octave	Mathématiques
	MOUSSA André	Chimie Nucléaire
	TRAYNARD Philippe	Chimie Générale
	SOUTIF Michel	Physique Générale
	CRAYA Antoine	Hydrodynamique
	REULOS René	Théorie des Champs
	BESSON Jean	Chimie Minérale
	AYANT Yves	Physique Approfondie
	GALLISSOT François	Mathématiques
Melle.	LUTZ Elisabeth	Mathématiques
MM.	BLAMBERT Maurice	Mathématiques
	BOUCHEZ Robert	Physique Nucléaire
	LLIBOUTRY Louis	Géophysique
	MICHEL Robert	Minéralogie et pétrographie
	BONNIER Etienne	Electrochimie et Electrométallurgie
	DESSAUX Georges	Physiologie animale
	PILLET Emile	Physique Industrielle-Electrotechnique
	YOCCOZ Jean	Physique Nucléaire théorique
	DEBELMAS Jacques	Géologie Générale
	GERBER Robert	Mathématiques
	PAUTHENET René	Electrotechnique
	MALGRANGE Bernard	Mathématiques Pures
	VAUQUOIS Bernard	Calcul Electronique
	BARJON Robert	Physique Nucléaire

MM.	BARBIER Jean-Claude	Physique
	SILBER Robert	Mécanique des Fluides
	BUYLE-BODIN Maurice	Electronique
	DREYFUS Bernard	Thermodynamique
	KLEIN Joseph	Mathématiques
	VAILLANT François	Zoologie et Hydrobiologie
	ARNAUD Paul	Chimie
	SENGEL Philippe	Zoologie
	BARNOUD Fernand	Biosynthèse de la Cellulose
	BRISSONNEAU Pierre	Physique
	GAGNAIRE Didier	Chimie Physique
Mme.	KOFLER Lucie	Botanique
MM.	DEGRANGE Charles	Zoologie
	PEBAY-PEROULA Jean-Claude	Physique
	RASSAT André	Chimie Systématique
	DUCROS Pierre	Cristallographie Physique
	DODU Jacques	Mécanique Appliquée I. U. T.
	ANGLES D'AURIAC Paul	Mécanique des Fluides
	LACAZE Albert	Thermodynamique
	GASTINEL Noel	Analyse numérique
	GIRAUD Pierre	Géologie
	PERRET René	Servo-mécanisme
	PAYAN Jean-Jacques	Mathématiques Pures

PROFESSEURS SANS CHAIRE

MM.	GIDON Paul	Géologie
Mme.	BARBIER Marie-Jeanne	Electrochimie
Mme.	SOUTIF Jeanne	Physique
	COHEN Joseph	Electrotechnique
	DEPASSEL R.	Mécanique des Fluides
	GLENAT René	Chimie
	BARRA Jean	Mathématiques Appliquées
	COUMES André	Electronique
	PERRIAUX Jacques	Géologie et Minéralogie
	ROBERT André	Chimie Papetière
	BIARREZ Jean	Mécanique Physique
	BONNET Georges	Electronique
	CAUQUIS Georges	Chimie Générale
	BONNETAIN Lucien	Chimie Minérale
	DEPOMIER Pierre	Physique Nucléaire-Génie Atomique
	HACQUES Gérard	Calcul numérique
	POLOUJADOFF Michel	Electrotechnique
Mme.	KAHANE Josette	Physique
Mme.	BONNIER Jane	Chimie
MM.	VALENTIN Jacques	Physique
	REBECQ Jacques	Biologie
	DEPORTES Charles	Chimie
	SARROT-REYNAULD Jean	Géologie
	BERTRANDIAS Jean-Paul	Mathématiques Appliquées
	AUBERT Guy	Physique

PROFESSEURS ASSOCIES

MM.	RODRIGUES Alexandre	Mathématiques Pures
	MORITA Susumu	Physique Nucléaire
	RADHAKRISHNA	Thermodynamique

MAITRES DE CONFERENCES

MM.	LANCIA Roland	Physique Atomique
Mme.	BOUCHE Liane	Mathématiques
MM.	KAHANE André	Physique Générale
	DOLIQUE Jean Michel	Electronique
	BRIERE Georges	Physique
	DESRE Georges	Chimie
	LAJZEHOWICZ Joseph	Physique
	LAURENT Pierre	Mathématiques Appliquées
Mme.	BERTRANDIAS Françoise	Mathématiques Pures
MM.	LONGEQUEUE Jean-Pierre	Physique
	SOHM Jean-Claude	Electrochimie
	ZADWORNY François	Electronique
	DURAND Francis	Chimie Physique
	CARLIER Georges	Biologie végétale
	PFISTER Jean-Claude	Physique
	CHIBON Pierre	Biologie animale
	IDELMAN Simon	Physiologie animale
	BLOCH Daniel	Electrotechnique I. P.
	MARTIN-BOUYER Michel	Chimie (C. S. U. Chambéry)
	SIBILLE Robert	Construction mécanique (I. U. T.)
	BRUGEL Lucien	Energétique I. U. T.
	BOUVARD Maurice	HYdrologie
	RICHARD Lucien	Botanique
	PELMONT Jean	Physiologie animale
	BOUSSARD Jean-Claude	Mathématiques Appliquées (I. P. G.)
	MOREAU René	Hydraulique I. P. G.
	ARMAND Yves	Chimie I. U. T.
	BOLLIET Louis	Informatique I. U. T.
	KUHN Gérard	Energétique I. U. T.
	PEFFEN René	Chimie I. U. T.
	GERMAIN Jean-Pierre	Mécanique
	JOLY Jean-René	Mathématiques Pures
Melle.	PIERY Yvette	Biologie animale
	BERNARD Alain	Mathématiques Pures
	MOHSEN Tahsin	Biologie (C. S. U. Chambéry)
	CONTE René	Mesures Physiques I. U. T.
	LE JUNTER Noel	Génie Electrique Electronique I. U. T.
	LE ROY Philippe	Génie Mécanique I. U. T.
	ROMIER Guy	Techniques Statistiques quantitatives I. U. T.
	VIALON Pierre	Géologie
	BENZAKEN Claude	Mathématiques Appliquées
	MAYNARD Roger	Physique

MM.	DUSSAUD René	Mathématiques (C. S. U. Chambéry)
	BELORIZKY Elie	Physique (C. S. U. Chambéry)
Mme.	LAJZEROWICZ Jeannine	Physique (C. S. U. Chambéry)
M.	JULLIEN Pierre	Mathématiques Pures
Mme.	RINAUDO Marguerite	Chimie
MM.	BLIMAN Samuel	E. I. E.
	BEGUIN Claude	Chimie Organique
	NEGRE Robert	I. U. T.

MAITRES DE CONFERENCES ASSOCIES

MM.	YAMADA Osamu	Physique du Solide
	NAGAO Makoto	Mathématiques Appliquées
	MAREZIO Massimo	Physique du Solide
	CHEECKE John	Thermodynamique
	BOUDOURIS Georges	Radioélectricité
	ROZMARIN Georges	Chimie Papetière

UN MODELE MATHEMATIQUE DE PROCESSUS D'INTERROGATION :

LES PSEUDOQUESTIONNAIRES

*"Si on ne goûte pas ces caractères,
je m'en étonne. Et si on les goûte,
je m'en étonne de même".*

La Bruyère

Je tiens à exprimer toute ma reconnaissance à Monsieur le Professeur J. KUNTZMANN qui a bien voulu me faire l'honneur de présider le Jury de cette thèse.

Monsieur le Professeur C.F. PICARD a inspiré et dirigé l'étude qui constitue cette thèse. Je le prie de croire à l'expression de ma sincère reconnaissance pour l'aide qu'il m'a constamment apportée par ses conseils et ses critiques.

Je remercie vivement Monsieur le Professeur N. GASTINEL pour l'attention et les suggestions apportées à ce travail.

Je remercie Monsieur J. FAURE pour l'heureuse collaboration qu'il a su développer entre le Service de Toxicologie du C.H.U. de Grenoble et l'I.M.A.G.

Je remercie Monsieur J.C. BOUSSARD d'avoir accepté de participer au Jury et de m'avoir proposé le sujet d'une deuxième thèse.

En outre, je témoigne de l'amitié qui me relie aux membres du Laboratoire de l'Institut de Mathématiques Appliquées de Grenoble, et de l'Equipe de Recherche Structures de l'Information de Paris.

J'adresse enfin mes vifs remerciements aux personnes du Secrétariat et du service de Tirage, pour le soin qu'elles ont apporté à la réalisation matérielle de cette thèse.

TERMINOLOGIE

* $|E|$ désignera le cardinal de l'ensemble E

* Etant donné une arborescence $\mathcal{A} = (X,\Gamma)$, on notera :

  - x_o sa racine
  - (x,y) l'arc joignant le sommet x au sommet y
  - $[x,y]$ le chemin allant de x à y
  - $\hat{\Gamma}_x$ l'ensemble des descendants de x dans $\mathcal{A}$
  - Γ_x l'ensemble des successeurs (descendants directs) de x dans $\mathcal{A}$
  - $A(x)$ l'ensemble des ascendants de x dans $\mathcal{A}$

  on notera que : $\begin{cases} x \in \hat{\Gamma}_x \\ x \notin A(x) \end{cases}$

  - Y l'ensemble des sommets non terminaux de $\mathcal{A}$
  - Z l'ensemble des sommets terminaux de $\mathcal{A}$.

* - Une matrice (m,n) aura m lignes et n colonnes.
  - $A = ((a_i^j))$: matrice de terme général a_i^j, i indice de ligne, j indice de colonne.
  - $\mathcal{S}(a,n)$: ensemble des matrices stochastique (a,n).

* $\mathbb{P}_n = \{\pi \in \mathbb{R}^n : \sum_{j=1}^{n} \pi_j = 1 \ , \quad \pi_j \geq 0 \quad \forall j \in \{1,\ldots,n\}\}$

  $\tilde{\mathbb{P}}_n = \{\pi \in \mathbb{R}^n : \sum_{j=1}^{n} \pi_j = 1 \ , \quad \pi_j > 0 \quad \forall j \in \{1,\ldots,n\}\}$

* Etant donné un détecteur D, on notera :

  - $\mathcal{K}(D)$ l'ensemble des pseudoquestionnaires construits sur D dont les sommets terminaux sont les sommets interprétables

  - $\mathcal{S}(D)$ l'ensemble des points essentiellement atteints par D.

* Etant donné $K \in \mathcal{K}(D)$, $\pi \in \mathbb{P}_n$, $\varepsilon > 0$, $R \in \mathbb{N}$, on notera :

  - $K(R)$ la restriction de K de hauteur R

  - $L[K,\pi]$ la longueur de cheminement de K opérant sur π

  - $K(\pi,\varepsilon)$ la restriction de K dans laquelle un sommet est terminal si et seulement si :

    . il est interprétable à ε près pour π

    . aucun de ses ascendants ne l'est.

  - $K(R\ ;\ \pi,\varepsilon)$ la restriction de hauteur R de $K(\pi,\varepsilon)$

  - $E(R\ ;\ \pi,\varepsilon)$ l'ensemble des sommets de $K(R\ ;\ \pi,\varepsilon)$, de rang R et non interprétables à ε près pour π.

INTRODUCTION

Etant donné un processus aléatoire T pouvant prendre un nombre fini d'états T_j sur une population Ω, nous considérons une famille D de processus aléatoires définis sur Ω. Chaque élément q de D a un nombre fini d'états q_i possibles, et nous supposons connues les probabilités conditionnelles $p(q_i | T_j)$.

L'objet de ce travail est l'étude de processus de décision arborescents, utilisant les éléments de D comme question, et visant à déterminer l'état du système T lors d'une épreuve ω quelconque de Ω.

Les questionnaires de Picard traitent du cas où les probabilités conditionnelles sont toutes égales à 0 ou à 1 ; nous introduisons les pseudoquestionnaires pour traiter le cas général. A différentes exigences quant à la qualité de la décision, nous faisons correspondre différentes notions de convergence pour les pseudoquestionnaires. Nous nous sommes efforcés d'établir des conditions asymptotiques sur D, nécessaires et suffisantes pour assurer ces différentes convergences.

L'interprétation de ces résultats dans le cadre du codage avec bruit fournit des résultats intéressants. Par ailleurs, ce formalisme est bien adapté à certains problèmes de reconnaissance de formes ; toutefois de nombreux problèmes restent ouverts quant à la valorisation des heuristiques utilisées.

CHAPITRE - I

PSEUDOQUESTIONNAIRES

INTRODUCTION

Nous définissons les pseudoquestionnaires et les éléments qui leur sont attachés ; par ailleurs, au paragraphe 5 nous présentons les différents modes de convergence, avec l'exposé de leurs motivations.

I - DEFINITIONS - NOTATIONS

1) Pseudoquestionnaires

Soient un ensemble Ω et une tribu $\mathcal{a}$ d'évènements dans Ω. Nous supposons donnés :

- un système complet d'évènements T sur $(\Omega,\mathcal{a})$, fini, soit : $T = (T_1,\ldots,T_n)$, les T_j formant une partition de Ω.

- n lois de probabilité ℓ_j définies respectivement sur $(T_j,\mathcal{a}_j)$, où $\mathcal{a}_j$ est la trace de $\mathcal{a}$ sur T_j.

Etant donné par ailleurs un entier a ($a \geq 2$), nous considérons l'ensemble Q(a) des systèmes complets d'évènements définis sur $(\Omega,\mathcal{a})$ ayant au plus a éléments. En notant

$$q_1,\ldots,q_{a(q)}$$

les éléments d'un système q, nous dirons que q est une *question*, de *base* a(q), d'issues $(q_1,\ldots,q_{a(q)})$.

Par ailleurs, en notant $S(a,n)$ l'ensemble des matrices stochastiques de a lignes et n colonnes, nous considérons l'application α de Q(a) dans $S(a,n)$ définie de la façon suivante :

$$\alpha(q) = ((\alpha_i^j(q))) \quad , \quad i \in \{1,\ldots,a\},\ j \in \{1,\ldots,n\}$$

avec :

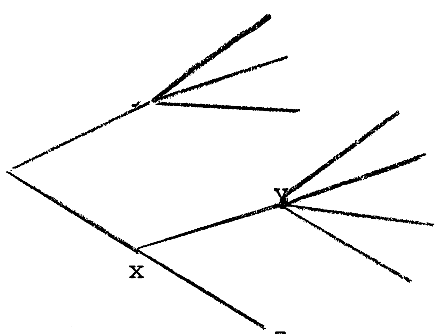
$$\begin{cases} \alpha_i^j(q) = \ell_j(q_i \cap T_j) & i \le a(q) \quad , \quad j \in \{1,\ldots,n\} \\ \alpha_i^j(q) = 0 & i > a(q) \quad , \quad j \in \{1,\ldots,n\}. \end{cases}$$

Remarque

Pour tout q, nous supposerons, sans restriction de généralité que :

$$\forall i \in \{1,\ldots,a(q)\},\ \exists j \in \{1,\ldots,n\} \ : \alpha_i^j(q) \neq 0$$

(s'il n'en était pas ainsi, il nous suffirait de réduire le nombre d'issues de la question q).



$$\begin{cases} u(x) = q, \text{question d'issues } q_1, q_2 \\ b_x((x,y)) = q_1 \\ b_x((x,y)) = q_2 \end{cases}$$

Figure I-1

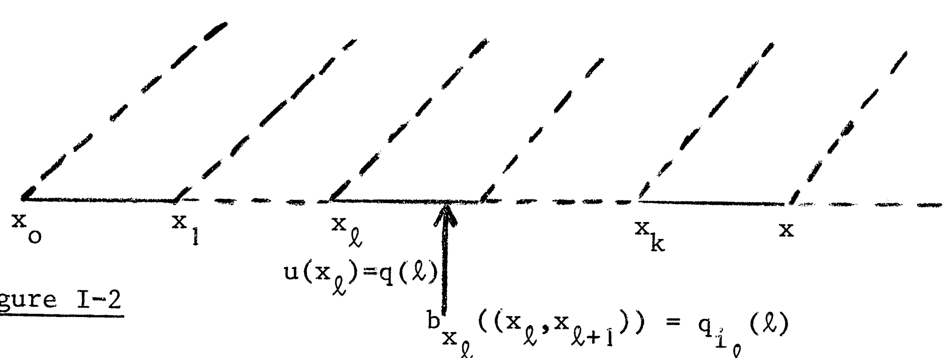


Figure I-2

DEFINITION I-1

Nous appellerons <u>*détecteur*</u> *opérant sur* T *un sous-ensemble* D *de* q(a) *dont tous les éléments sont stochastiquement indépendants pour les lois* ℓ_j, $j \in \{1,\dots,n\}$.

DEFINITION I-2

Etant donné un détecteur D *nous appellerons* <u>*pseudoquestionnaire*</u> *construit sur* D, *un triplet* (A,u,B) *où*

- *A est un graphe arborescent (fini ou infini* [20]) (X,Γ) *dont tous les sommets ont au plus* a *successeurs.*

- u *est une application de l'ensemble* Y *des sommets non terminaux de* A *dans* D *vérifiant :*

  i) $|\Gamma_y| = a(u(y)) \quad \forall y \in Y$

  ii) la restriction de u à l'ensemble des sommets de tout chemin de *A* est injective.

- $B = \{b_y \mid y \in Y\}$, *où* b_y *est une bijection entre l'ensemble des arcs sortant de* y, *et celui des issues de la question* u(y).

(cf. Figure I-1).

<u>Remarque</u>

La condition ii), ci-dessus s'interprète comme suit :
on s'interdit d'utiliser plus d'une fois une même question sur un chemin de *A*.

Nous considèrerons un pseudoquestionnaire K comme un graphe, et utiliserons le langage habituel des graphes [3], et même plus précisemment celui des questionnaires [17]. Nous dirons ainsi qu'un sommet x de A est un sommet de K, et nous noterons r(x) son rang, c'est-à-dire le nombre de ses ascendants dans A. Nous dirons que K est de hauteur R, si tous ses sommets sont de rang au plus égal à R, l'un d'eux étant de rang égal à R.

2) Prolongements, restrictions d'un pseudoquestionnaire - Sous-pseudoquestionnaires

Considérons deux pseudoquestionnaires $K = (A, u, B)$, $K' = (A', u', B')$ construits sur un même détecteur.

Nous dirons que K est un *prolongement* de K' (resp. que K' est une *restriction* de K) si :

- A est un prolongement de A', c'est-à-dire, en notant $A = (X,\Gamma)$ et $A' = (X',\Gamma')$:
  - $X' \subset X$, $\Gamma' \subset \Gamma$
  - A et A' ont même racine
  - $\Gamma_y = \Gamma'_y$ $\forall y \in Y'$ (Y' désignant l'ensemble des sommets non terminaux de A').
- u prolonge u'
- $b'_y = b_y$ $\forall y \in Y'$.

Nous dirons que K' est un *sous-pseudoquestionnaire* de K si :

- A' est une sous-arborescence de A.
- u prolonge u'
- $b'_y = b_y$ $\forall y \in Y'$.

Notation

Etant donné un pseudoquestionnaire K de hauteur R_o, nous noterons K(R) la restriction de K de hauteur R $(R \leq R_o)$.

3) Construction récurrente d'un pseudoquestionnaire

Il est souvent commode de construire un pseudoquestionnaire K par récurrence sur ses restrictions de hauteur R. Il suffit ainsi de définir un procédé permettant de construire K(R+1) à partir de K(R).

Pour ce faire, étant donné $K(R) = (\mathcal{A}_R, u_R, B_R)$ et un sommet x de rang R dans K(R), il suffit, soit de décider que x est terminal dans K(R+1), soit de définir :

- $u_{R+1}(x)$. L'arborescence $\mathcal{A}_{R+1}$ est alors déterminée en raison de la condition : $a(u_{R+1}(x)) = |\Gamma_x|$

- une bijection entre les arcs sortant de x ainsi construits et les issues de la question u(x).

II - PROBABILITES ATTACHES AUX DIFFERENTS ELEMENTS D'UN PSEUDOQUESTIONNAIRE.

Supposons donné un élément π de l'ensemble $\mathbb{P}_n$ défini par :

$$\mathbb{P}_n = \{\pi \in \mathbb{R}^n : \sum_{j=1}^{n} \pi_j = 1 \quad , \quad \pi_j \geq 0 \quad \forall j \in \{1,\dots,n\}\}$$

π engendre sur $(\Omega,\mathcal{a})$ une loi de probabilité p définie par :

$$p(A) = \sum_{j=1}^{n} \pi_j \ell_j (A \cap T_j) \qquad \forall A \in \mathcal{a}.$$

Il est clair que pour cette loi

$$\pi_j = p(T_j)$$

et que quel que soit $\pi \in \mathbb{P}_n$:

$$\ell_j(A \cap T_j) = p(A|T_j) = \frac{p(A \cap T_j)}{p(T_j)} \qquad (\text{si } \pi_j \neq 0) .$$

1) Probabilité d'un sommet d'un pseudoquestionnaire

Considérons un pseudoquestionnaire $K = (A,u,B)$ et un élément π de $\mathbb{P}_n$; soit p la mesure engendrée par π sur (Ω,a). Etant donné un sommet x de K, soit $(x_o,\ldots x_k)$ la suite ordonnée de ses ascendants dans K.
On pose : (cf. Fig. I-2)

$$\begin{cases} q(\ell) = u(x_\ell) & \ell \in \{0,\ldots,k\} \\ b_{x_\ell}((x_\ell,x_{\ell+1})) = q_{i_\ell}(\ell) & \end{cases}$$

$$\mathcal{E}(x) = \bigcap_{\ell=0}^{k} q_{i_\ell}(\ell) .$$

On appelle *probabilité de* x *dans* K *opérant sur* π la quantité $p(\mathcal{E}(x))$, que nous noterons par abus de langage $p(x)$.
En raison de l'indépendance stochastique des questions, nous avons :

$$p(x) = \sum_{j=1}^{n} \pi_j \Lambda_j(x)$$

où

$$\Lambda_j(x) = \prod_{\ell=0}^{k} \alpha_{i_\ell}^{j}(q(\ell)) .$$

On remarque que quel que soit $\pi \in \mathbb{P}_n$, on a :

$$\Lambda_j(x) = p(\mathcal{E}(x)|T_j).$$

Dans la suite, nous désignons sous le nom de *caractéristique du sommet* x *dans* K, l'élément $\Lambda(x)$ de $[0,1]^n$ de composantes $\Lambda_j(x)$; cet élément est indépendant de π.

Probabilité d'une section, ou d'une partie de section

Etant donné une arborescence A, on considère l'ensemble C des chemins sortant de la racine x_o de A.

On dit qu'un ensemble σ de sommets de A est une *section* de A [20], si tout chemin de C passe par un élément et un seul de σ.

Considérons alors un pseudoquestionnaire $K = (A,u,B)$, et une section σ de A. Les évènements $\mathcal{E}(x)$ associés aux éléments x de σ forment une partition de Ω ; donc :

$$\sum_{x \in \sigma} p(x) = 1.$$

Si σ' est une partie de section, nous avons alors

$$\sum_{x \in \sigma'} p(x) \leq 1$$

et nous noterons $p(\sigma')$ la quantité $\sum_{x \in \sigma'} p(x)$.

2) Longueur de cheminement d'un pseudoquestionnaire

La notion de longueur de cheminement introduite par Picard à l'occasion des questionnaires [17], se transpose naturellement pour des pseudoquestionnaires de hauteur finie.

Considérons un élément π de $\mathbb{P}_n$, et désignons par Z et Y les ensembles de sommets respectivement terminaux et non terminaux d'un pseudoquestionnaire K de hauteur finie.

La longueur de cheminement de K opérant sur π est :

$$L[K,\pi] = \sum_{x \in Z} p(x)\ r(x).$$

On montre [17] la propriété :

$$L[K,\pi] = \sum_{x \in Y} p(y).$$

3) Vecteur de probabilité associé à un sommet de probabilité non nulle

Nous considérons un pseudoquestionnaire $K = (\mathcal{A},u,B)$ et la mesure p engendrée par un élément π de $\mathbb{P}_n$.

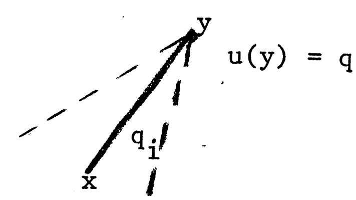
Nous appellerons vecteur de probabilité associé à un sommet x de K opérant sur π, l'élément $P(T|x)$ de $\mathbb{P}_n$ ayant pour composantes :

$$p(T_j|\mathcal{E}(x)) \quad , \quad j \in \{1,\ldots,n\}$$

que nous noterons par abus de langage $p(T_j|x)$.
Il est clair que ceci n'a de sens, que si $p(x) \neq 0$.
Nous avons alors :

$$p(T_j|x) = \frac{\pi_j \Lambda_j(x)}{p(x)} .$$

Remarquons que si $p(T_j|x_o) = \pi_j$ (x_o désignant la racine de K), et si y est l'antécédent de x dans K, nous avons les relations :



$$\left\{\begin{array}{l} p(T_j|x) = \dfrac{p(T_j|y)\ \alpha_i^j(q)}{\sum\limits_{j=1}^{n} p(T_j|y)\ \alpha_i^j(q)} \\[2em] p(x) = \left(\sum\limits_{j=1}^{n} p(T_j|y)\ \alpha_i^j(q)\right)p(y). \end{array}\right. \qquad \text{avec :} \left\{\begin{array}{l} q = u(y) \\[1em] q_i = b_y((y,x)) \end{array}\right.$$

Ainsi, les vecteurs associés aux sommets de K, et les probabilités de ces sommets peuvent se calculer par récurrence sur le rang.

Remarque

Etant donné un sommet x de K, et une section σ du sous-pseudoquestionnaire de racine x de K, nous avons la relation :

$$p(T_j|x) = \sum_{y \in \sigma} \frac{p(y)}{p(x)} p(T_j|y).$$

Ceci peut être considéré comme une propriété de "conservation de flux" sur un pseudoquestionnaire.

III - PSEUDOQUESTIONNAIRES GENERALISES

Toutes les notions définies au paragraphe précédent ne font intervenir les questions q que par l'intermédiaire de leur matrice stochastique $\alpha(q)$, d'où l'idée de généraliser la notion de pseudoquestionnaire.

DEFINITION I-3

Etant donné deux entiers a *et* n $(a \geq 2,\ n \geq 2)$, *et un sous-ensemble* S *de* $\mathcal{S}(a,n)$ *nous appellerons pseudoquestionnaire généralisé construits sur* S *le triplet* (A, v, B) *où :*

- *A est un graphe arborescent* (X,Γ) *dont tous les sommets ont au plus* a *successeurs.*

- v *est une application de l'ensemble des sommets non terminaux* Y *de A dans* S *vérifiant :*

  $a(v(y)) = |\Gamma_y|\quad \forall y \in Y$ (a(v(y)) désignant lenombre de lignes à éléments non tous nuls de la matrice v(y))

- $B = \{b_y | y \in Y\}$, b_y *étant une bijection entre l'ensemble des arcs sortant de* y *et l'ensemble des lignes à éléments non tous nuls de* v(y).

Il est clair qu'à un pseudoquestionnaire $(\mathcal{A},u,B)$, on peut faire correspondre le pseudoquestionnaire généralisé $(\mathcal{A},v,B)$ où :

$$v = \alpha \circ u.$$

Par ailleurs, les prolongements, restrictions, et sous-pseudoquestionnaires d'un pseudoquestionnaire généralisé se définissent de la même façon que pour les pseudoquestionnaires.
On définit de façon analogue les probabilités associées aux sommets d'un pseudoquestionnaire généralisé K opérant sur un élément π de $\mathbb{P}_n$:
soit x un sommet de K, et $(x_o,\ldots,x_k)$ la suite ordonnée de ses ascendants dans K. Nous posons :

$$s(\ell) = v(x_\ell) \qquad \ell \in \{0,\ldots,k\}$$

$$b_{x_\ell}((x_\ell,x_{\ell+1})) = (s^1_{i_\ell}(\ell),\ldots,s^n_{i_\ell}(\ell)) \qquad \ell \in \{0,\ldots,k\}.$$

En prenant alors pour caractéristique $\Lambda(x)$ du sommet x, l'élément de $[0,1]^n$ de composantes :

$$\Lambda_j(x) = \prod_{\ell=0}^{k} s^j_{i_\ell}(\ell)$$

nous définissons la probabilité p(x) du sommet x

$$p(x) = \sum_{j=1}^{n} \pi_j \Lambda_j(x)$$

et, si $p(x) \neq 0$, nous définissons le vecteur de probabilité $P(T|x)$, élément de $\mathbb{P}_n$, de composantes

$$p(T_j|x) = \frac{\pi_j \Lambda_j(x)}{p(x)} .$$

Remarque de continuité :

La notion de pseudoquestionnaire généralisé nous sera très utile par la suite, en raison des propriétés suivantes de continuité.

Considérons les applications f et g_j, $j \in \{1,\ldots,n\}$, suivantes :

$$\begin{cases} (z,t) \xrightarrow[o]{f} \sum_{j=1}^{n} z_j t_j & \text{définie sur } [0,1]^n \times [0,1]^n \\[2ex] (z,t) \xrightarrow[o]{g_j} \dfrac{z_j t_j}{f(z,t)} & \text{définie sur } [0,1]^n \times [0,1]^n \cap \{(z,t) : f(z,t) \neq 0\}. \end{cases}$$

Nous remarquons tout d'abord, qu'étant donné un pseudoquestionnaire (généralisé ou non) K opérant sur π, si x est un sommet de K de caractéristique $\Lambda(x)$, nous avons :

$$\begin{cases} p(x) = f(\pi,\Lambda(x)) \\[2ex] p(T_j|x) = g_j(\pi,\Lambda(x)) \qquad j \in \{1,\ldots,n\}. \end{cases}$$

Ainsi la probabilité d'un sommet x d'un pseudoquestionnaire est une fonction continue de π et de $\Lambda(x)$, et les composantes du vecteur de probabilité d'un sommet x, de probabilité non nulle, sont des fonctions continues de π et de $\Lambda(x)$.

IV - PSEUDOQUESTIONNAIRES ET QUESTIONNAIRES

Nous dirons qu'un *pseudoquestionnaire est trivial dans son opération sur* $\pi \in \mathbb{P}_n$, s'il admet une section C telle que :

$$\forall x \in C,\ P(T|x) \text{ comporte un } 1 \text{ et } (n-1) \text{ zéros.}$$

Il est clair que si un pseudoquestionnaire est trivial pour un élément $\pi°$ de $\mathbb{P}_n$ de composantes toutes non nulles, il l'est pour tout $\pi \in \mathbb{P}_n$. Un tel pseudoquestionnaire sera dit *trivial*. On montre facilement la propriété suivante :

PROPRIETE I-1

Une condition nécessaire et suffisante pour qu'il existe un pseudoquestionnaire trivial construit sur un détecteur D *est que* D *vérifie :*

$$\forall j\in\{1,\dots,n\},\forall j'\in\{1,\dots,n\},j\neq j',\exists q\in D : (\alpha_i^j(q)\neq 0 \Rightarrow \alpha_i^{j'}(q)=0 \quad \forall i\in \{1,\dots,a(q)\}).$$

Lorsqu'un détecteur D admet un pseudoquestionnaire trivial, nous dirons que D est trivial.

On peut vérifier, comme l'illustre l'exemple suivant, qu'à un pseudoquestionnaire trivial correspond un questionnaire latticiel [18] :

Nous supposons n = 3, et considérons le détecteur D constitué des trois éléments q(1), q(2), q(3) dichotomiques suivants :

$$\left\{\begin{array}{lll}
\alpha_1^1(q(1)) = 1 & \alpha_1^2(q(1)) = b \neq 0 & \alpha_1^3(q(1)) = 0 \\
\alpha_1^1(q(2)) = 1 & \alpha_1^2(q(2)) = 0 & \alpha_1^3(q(2)) = c \neq 0 \\
\alpha_1^1(q(3)) = d \neq 0 & \alpha_1^2(q(3)) = 1 & \alpha_1^3(q(3)) = 0.
\end{array}\right.$$

On peut alors construire le pseudoquestionnaire trivial de la figure I-3, et lui associer le questionnaire latticiel de la figure I-4.

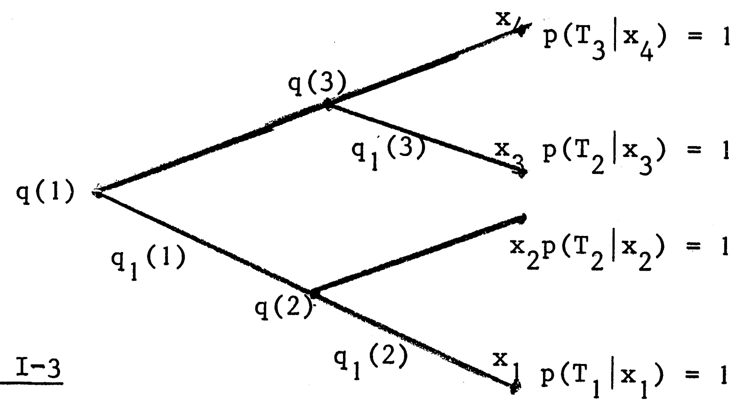


Figure I-3

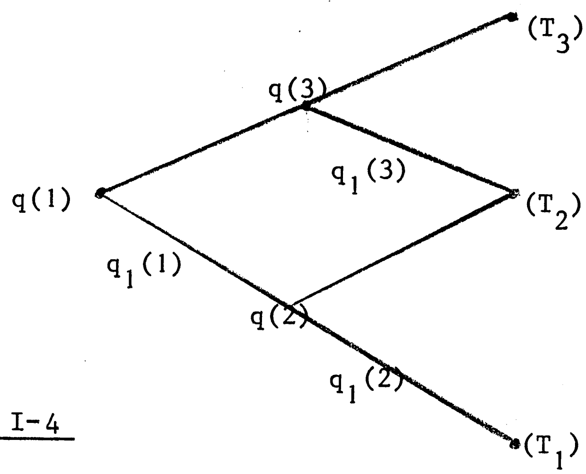


Figure I-4

V - CONVERGENCE DE PSEUDOQUESTIONNAIRES

Un pseudoquestionnaire sera utilisé comme un processus d'interrogation visant à déterminer l'évènement T_j qui s'est réalisé dans une épreuve $\omega \in \Omega$. En particulier, si nous considérons un sommet x d'un pseudoquestionnaire correspondant à la suite de réponses

$$q_{i_1}(1),\ldots,q_{i_k}(k)$$

et tel qu'il existe $j_o \in \{1,\ldots,n\}$ vérifiant $p(T_{j_o} \mid x) = 1$, alors si tous les évènements de cette suite se sont réalisés, nous concluons à la réalisation de T_{j_o} .

Toutefois, les seuls pseudoquestionnaires conduisant sur tous leurs chemins, et en un nombre fini de questions, à une telle situation, sont les pseudoquestionnaires triviaux. Ceci nous conduit à choisir des critères de décision moins exigeants, et à travailler sur les notions suivantes :

1) Sommets interprétables

Etant donné un pseudoquestionnaire K, généralisé ou non, nous dirons que

- un sommet x est *interprétable pour* π, si dans K opérant sur π, il existe $j_o \in \{1,\ldots,n\}$ tel que :

$$p(T_{j_o}|x) = 1$$

- un sommet x est *interprétable*, si il est interprétable pour tout élément π de $\mathbb{P}_n$.
  (Il est clair qu'il suffit pour ceci que x soit interprétable pour un élément $\pi°$ n'ayant aucune composante nulle).

- étant donné un nombre $\varepsilon \in]0,1[$, un sommet x est *interprétable à* ε *près pour* π si dans K opérant sur π, il existe $j_o \in \{1,\ldots,n\}$ tel que :

$$p(T_{j_o}|x) \geq 1-\varepsilon.$$

2) Définition de $\mathcal{H}$(D) et $\mathcal{H}$(S)

Etant donné un détecteur D, *nous désignons par* $\mathcal{H}$(D), *l'ensemble des pseudoquestionnaires construits sur* D *dont les sommets terminaux sont les sommets interprétables.*

Il est clair que cette définition est indépendante de l'élément π sur lequel opère un pseudoquestionnaire, elle est équivalente à :

$$\text{x terminal dans } K \in \mathcal{H}(D) \iff \begin{cases} \exists j_o \in \{1,\ldots,n\} : \Lambda_{j_o}(x) \neq 0 \\ \\ \forall j \in \{1,\ldots,n\},\ j \neq j_o : \Lambda_j(x) = 0. \end{cases}$$

Etant donné une partie S de S(a,n) on définit de même l'ensemble $\mathcal{H}$(S).

<u>Restrictions particulières pour des éléments de $\mathcal{H}$(D). Notations</u>

★ Etant donné un élément K de $\mathcal{H}$ (D) (resp. $\mathcal{H}$(S)), un élément π de $\mathbb{P}_n$, et $\varepsilon \in \left]0,1\right[$, nous noterons

$$K(\pi,\varepsilon)$$

la restriction de K dans laquelle un sommet est terminal si et seulement si :
- il est interprétable à ε près pour π
- aucun de ses ascendants ne l'est.

★ Etant donné un entier positif R, nous noterons

$$K(R\ ;\ \pi,\varepsilon)$$

la restriction de hauteur R de $K(\pi,\varepsilon)$, et

$$L\left[K(R\ ;\ \pi,\varepsilon)\right]$$

sa longueur de cheminement.

★ Par ailleurs, nous noterons

$$E(R\ ;\ \pi,\varepsilon)$$

l'ensemble des sommets de $K(R\ ;\ \pi,\varepsilon)$, de rang R, et non interprétables à ε près pour π.

3) Convergences

DEFINITION I-4

Etant donné un détecteur D *(resp. une partie de* S *de* $S(a,n)$*), un élément* K *de* $\mathcal{H}$ (D) *(resp.* $\mathcal{H}$ (S)*) sera dit :*

1 – *Fortement convergent (F- Convergent) pour un élément* π *de* $\mathbb{P}_n$ *si et seulement si :*

$$\forall \varepsilon > 0 \quad K(\pi,\varepsilon) \text{ est de hauteur finie.}$$

2 – L-*convergent pour un élément* π *de* $\mathbb{P}_n$*, si et seulement si :*

$$\forall \varepsilon > 0 \quad \lim_{R\to\infty} L[K(R\ ;\ \pi,\varepsilon)] < +\infty$$

3 – H-*convergent pour un élément* π *de* $\mathbb{P}_n$*, si et seulement si :*

$$\forall \varepsilon > 0 \quad \lim_{R\to\infty} p(E(R\ ;\ \pi,\varepsilon)) = 0.$$

PROPRIETE I-2

Etant donné un élément K *de* $\mathcal{H}$(D) *(resp.* $\mathcal{H}$(S) *) et un élément* π *de* $\mathbb{P}_n$*, on a :*

K F-converge pour π => K L-converge pour π => K H-converge pour π.

Démonstration

La première implication est tout-à-fait évidente ; montrons la deuxième. On a :

$$L[K(R\ ;\ \pi,\varepsilon)] = \sum_{r=0}^{R-1} p(E(r\ ;\ \pi,\varepsilon))$$

par conséquent, si K L-converge pour π, la série de terme général

$$W_r(\varepsilon) = p(E(r\ ;\ \pi,\varepsilon))$$

converge pour tout $\varepsilon > 0$, et son terme général tend vers zéro pour tout $\varepsilon > 0$.

~~X~~

Interprétation

- La forte convergence d'un pseudoquestionnaire K signifie que quel que soit $\varepsilon > 0$, K permet de décider à ε près le T_j réalisé dans toute épreuve en un nombre fini de questions.

- La L-convergence de K signifie que pour tout $\varepsilon > 0$, K permet de décider à ε près le T_j réalisé dans toute épreuveen un nombre de questions fini en moyenne.

- La H-convergence signifie que pour tout $\varepsilon > 0$, K permet de décider à ε près presque sûrement sur Ω.

- Dans la pratique, on arrêtera évidemment la construction d'un pseudoquestionnaire dès que celui-ci aura une certaine hauteur R, on prendra donc deux risques :

  . le risque ρ_1 de conclure à la réalisation de T_{j_o}, alors qu'en fait c'est l'évènement T_j, $j \neq j_o$, qui s'est réalisé.

  . le risque ρ_2 d'aboutir à un sommet non interprétable à εprès.

On montre facilement que quels que soient K et R, on a :

$$\rho_1 \leq \varepsilon.$$

Par ailleurs, il est clair que :

$$\rho_2 = p(E(R\ ;\ \pi,\varepsilon)).$$

Par conséquent, la H-convergence d'un pseudoquestionnaire signifie que quelle que soit l'exigence que l'on ait vis à vis du premier risque d'erreur ρ_1, on peut rendre le deuxième aussi petit que l'on veut.

Remarque

Etant donné $J \subset \{1,\dots,n\}$, considérons l'ensemble $\mathbb{P}_n(J)$ défini par :

$$\mathbb{P}_n(J) = \{\pi \in \mathbb{P}_n : \pi_j \neq 0 \quad j \in J, \quad \pi_j = 0 \quad j \notin J\}.$$

Etant donné un détecteur D (resp. $S \subset \mathcal{S}(a,n)$) nous définissons l'ensemble $\mathcal{K}(D,J)$ (resp. $\mathcal{K}(S,J)$) des pseudoquestionnaires construits sur D (resp. sur S) dont les sommets terminaux sont les sommets interprétables pour tout élément π de $\mathbb{P}_n(J)$.

Il est clair que pour tout J, tout élément K de $\mathcal{K}(D)$ admet une restriction dans $\mathcal{K}(D,J)$. Toutefois, J étant fixé ($J \neq \{1,\dots,n\}$), et étant donné un élément $K_J \in \mathcal{K}(D,J)$, pour obtenir par prolongement de K_J un élément K de $\mathcal{K}(D)$, on pourra être amené à prolonger indéfiniment certains chemins de K_J.
Les définitions de convergence peuvent s'étendre aux éléments de $\mathcal{K}(D,J)$ (resp. $\mathcal{K}(S,J)$) de la façon suivante :

Un élément K de $\mathcal{K}(D,J)$ sera dit F-convergent pour $\pi \in \mathbb{P}_n(J)$ si et seulement si :

$$\forall \varepsilon > 0 \qquad K(\pi,\varepsilon) \quad \text{est de hauteur finie.}$$

On étendrait de même les définitions de la L-convergence et de la H-convergence.
Par ailleurs, si $K_J \in \mathcal{K}(D,J)$ est F-convergent (resp. L-convergent, resp. H-convergent) pour un élément π de $\mathbb{P}_n(J)$, alors tout élément K de $\mathcal{K}(D)$ obtenu par prolongement à partir de K_J sera F-convergent, (resp. L-convergent, resp. H-convergent) pour π. Par conséquent, étant donné $\pi \in \mathbb{P}_n(J)$, pour construire un élément K de $\mathcal{K}(D)$ F-convergent (resp. L-convergent, resp. H-convergent) pour π, il suffira de construire un élément de $\mathcal{K}(D,J)$.

4) Détecteurs infinis

Il est clair que si un détecteur D est fini, les trois notions de convergence sont équivalentes. Seuls alors les pseudoquestionnaires triviaux sont convergents.

Nous éliminerons ce cas, et étudierons des détecteurs infinis. Notre propos est de trouver des conditions suffisantes et éventuellement nécessaires sur D, pour qu'il existe des pseudoquestionnaires construits sur D, convergents en chacun des trois sens. Nous aurons besoin pour ceci de la notion de *point essentiellement atteint* par un détecteur, définie comme suit :

étant donné un détecteur infini D, considérons l'ensemble

$$\alpha(D) = \{s \in \mathcal{S}(a,n) \; : \; \exists q \in D \quad s = \alpha(q)\}.$$

En ayant fait choix d'une métrique quelconque sur $\mathbb{R}^{n\times a}$, l'adhérence $\overline{\alpha(D)}$ de $\alpha(D)$ est compacte. Par conséquent, D étant infini, soit $\alpha(D)$ est infini et admet alors un point d'accumulation, soit $\alpha(D)$ est fini, et alors il existe une partie D' infinie de D sur laquelle l'application α est constante.

DEFINITION I-5

Nous dirons qu'un point $s \in \mathcal{S}(a,n)$ *est essentiellement atteint par* D, *si pour tout* $\delta > 0$, *il existe une infinité de questions* q *de* D *telles que* $\alpha(q)$ *soit dans la boule de centre* s *et de rayon* δ.

D'après les remarques faites précédemment, s est essentiellement atteint par D, s'il est point d'accumulation de $\alpha(D)$, ou s'il existe une infinité de questions q de D telles que $\alpha(q) = s$. Par ailleurs, tout détecteur infini possède au moins un point essentiellement atteint ; nous noterons

$$\mathcal{S}(D)$$

l'ensemble des points essentiellement atteints par D.

Remarque

Cette notion peut se traduire topologiquement dans Q(a), ensemble des systèmes complets de a éléments au plus.
En effet, on peut montrer que l'image de Q(a) par l'application α est fermée [21]. Donc tout point essentiellement atteint par D est l'image d'un élément $\bar{q}$ de Q(a), n'appartenant pas nécessairement à D.
Pour la topologie non séparée induite sur D par la pseudo distance

$$d(q,q') = \|\alpha(q)-\alpha(q')\|$$

un tel élément $\bar{q}$ est point d'accumulation de D.
Toutefois, il est plus facile pour la suite de travailler sur les matrices stochastiques des éléments de D.

Un exemple de détecteur infini.

Soit une expérience q, telle que si on la répète, les résultats obtenus sont indépendants des résultats antérieurs. A une telle expérience, on peut alors associer le détecteur infini D formé de la suite $\{q(i)\}$, $i \in \mathbb{N}$, telle que tous les éléments de cette suite aient même matrice stochastique.
Cette remarque nous sera particulièrement utile lors de l'application des pseudoquestionnaires au problème du codage avec bruit.

CHAPITRE - II

PROPRIETES D'UNIFORMITE DES CONVERGENCES DE PSEUDOQUESTIONNAIRES

INTRODUCTION

Nous avons vu que les trois notions de convergence d'un pseudoquestionnaire ont été définies pour un élément π de $\mathbb{P}_n$. En notant :

$$\tilde{\mathbb{P}}_n = \{\pi \in \mathbb{P}_n : \pi_j > 0 \quad \forall j \in \{1,\ldots,n\}\}$$

nous allons montrer que si un pseudoquestionnaire est F-convergent (resp. L-convergent) pour un élément $\pi°$ de $\tilde{\mathbb{P}}_n$, alors il l'est pour tout élément de $\tilde{\mathbb{P}}_n$, mais ne l'est pas nécessairement pour des éléments π de $\mathbb{P}_n$, admettant des composantes nulles.

Par contre, si un pseudoquestionnaire est H-convergent pour un élément $\pi°$ de $\tilde{\mathbb{P}}_n$, alors il l'est pour tout élément de $\mathbb{P}_n$.

I - INFLUENCE DE π SUR LA F-CONVERGENCE ET LA L-CONVERGENCE

PROPOSITION II-1

Etant donné un détecteur D *(resp.* $S \subset \mathcal{S}(a,n)$*) et un élément* K *de* $\mathcal{H}$(D) *(resp.* $K \in \mathcal{H}(S)$*), nous avons :*

i) Si K *est* F-*convergent pour un élément* $\pi°$ *de* $\tilde{\mathbb{P}}_n$, *alors* K *est* F-*convergent pour tout élément* π *de* $\tilde{\mathbb{P}}_n$.

ii) Si K *est* L-*convergent pour un élément* $\pi°$ *de* $\tilde{\mathbb{P}}_n$, *alors* K *est* L-*convergent pour tout élément* π *de* $\tilde{\mathbb{P}}_n$.

Démonstration

Nous noterons p_o la mesure engendrée par $\pi°$.

* *Démonstration de i)*

Pour tout sommet x de K, nous avons :

$$\frac{p_o(T_j|x)}{p_o(T_k|x)} = \frac{p(T_j|x)}{p(T_k|x)} \; \frac{\pi_k^\circ}{\pi_j^\circ} \; \frac{\pi_j}{\pi_k} \qquad \forall j,\ k \in \{1,\ldots,n\}.$$

Posons :

$$\lambda = \mathrm{Min}\ \left[\frac{\pi_k^\circ}{\pi_j^\circ}\,\frac{\pi_j}{\pi_k} \;\middle|\; j,\ k \in \{1,\ldots,n\},\ j \neq k\right].$$

Le pseudoquestionnaire K étant supposé F-convergent pour $\pi°$, pour tout chemin C de K, nous avons :

$$(1)\ \forall B_o,\ \exists x\in C,\ \exists j\in\{1,\ldots,n\} \text{ tels que } \frac{p_o(T_j|x)}{p_o(T_k|x)} \geq B_o \qquad \forall k \in \{1,\ldots,n\},\ k \neq j.$$

Pour que K soit F-convergent pour π, il suffit que pour tout chemin C de K, nous ayons :

$$(2)\ \forall B,\ \exists x\in C,\ \exists j\in\{1,\ldots,n\} \text{ tels que } \frac{p(T_j|x)}{p(T_k|x)} \geq B \qquad \forall k \in \{1,\ldots,n\},\ k \neq j.$$

B étant fixé, pour que (2) soit vérifié , il suffit de prendre dans (1), B_o tel que :

$$B_o = \frac{B}{\lambda}\ .$$

~~X~~

* *Démonstration de ii)*

Sous une forme équivalente, nous avons montré ci-dessus que :

(3) Quel que soit ε, il existe ε_o, tel que si un sommet x de K est interprétable à ε_o près pour $\pi°$, alors x est interprétable à ε près pour π.

Il nous suffit donc de montrer que :

(4) Il existe $\mu > 0$, tel que pour tout sommet x de K, on ait :

$$p(x) < \mu\, p_o(x).$$

En effet, étant donné $\varepsilon > 0$, il existera alors en raison de (3) et (4), $\varepsilon_o > 0$ et $\mu > 0$, tels que :

$$L[K(R\ ;\ \pi,\varepsilon)] \leq \mu\, L[K(R\ ;\ \pi°,\varepsilon°)] \quad \forall R.$$

Démontrons (4) ; pour tout sommet x de K, nous avons :

$$\begin{cases} p(x) = \sum_{j=1}^{n} \pi_j \Lambda_j(x) \\ \\ p_o(x) = \sum_{j=1}^{n} \pi_j^\circ \Lambda_j(x). \end{cases}$$

Par suite :

$$p(x) = \sum_{j=1}^{n} \pi_j^\circ \Lambda_j(x) \left(\frac{\pi_j}{\pi_j^\circ}\right).$$

Si nous posons alors :

$$\mu = \text{Max}\left[\frac{\pi_j}{\pi_j^\circ} \mid j \in \{1,\dots,n\}\right]$$

il vient :

$$p(x) \leq \mu\, p_o(x).$$

~~X~~

Remarque

Signalons, comme le montre l'exemple suivant, que si un pseudoquestionnaire est F-convergent pour $\pi^\circ \in \tilde{\mathbb{P}}_n$ (resp. L-convergent), il ne l'est pas nécessairement pour un élément π ayant des composantes nulles.

Nous prenons n égal à 3, et considérons le détecteur D constitué de la suite de questions dichotomiques $\{q(r)\}$, $r \in \mathbb{N}$, définie comme suit :

$$\begin{cases} \alpha_1^1(q(0)) = 1 \quad , \quad \alpha_1^2(q(0)) = 0 \quad , \quad \alpha_1^3(q(0)) = 1 \\ \alpha_1^1(q(r)) = 1 \quad , \quad \alpha_1^2(q(r)) = \alpha_1^3(q(r)) = \frac{r}{r+1} \qquad r \in \{1,2,\ldots\}. \end{cases}$$

Nous considérons le pseudoquestionnaire $K = (\mathcal{A},u,B)$ représenté sur la figure II-1 ; K est un élément de $\mathcal{H}$ (D) tel que :

Quel que soit r, $(r \geq 2)$, K possède 4 sommets de rang r dont deux sont interprétables.

Il est clair que K est F-convergent (donc L-convergent) pour tout élément π° de $\tilde{\mathbb{P}}_3$; en effet, K a un seul chemin infini C, et sur ce chemin, un calcul élémentaire montre que :

$$\forall B \quad \exists y \in C \text{ tel que : } \frac{p_o(T_1|y)}{p_o(T_j|y)} \geq B \qquad \forall j \in \{2,3\}.$$

Soit maintenant un élément π de $\mathbb{P}_3$, tel que :

$$\pi_1 = 0.$$

En désignant par x_r le sommet de rang r tel que

$$u(x_r) = q(r)$$

nous avons :

$$p(x_r) = \frac{r-1}{r}\, p(x_{r-1})$$

d'où :

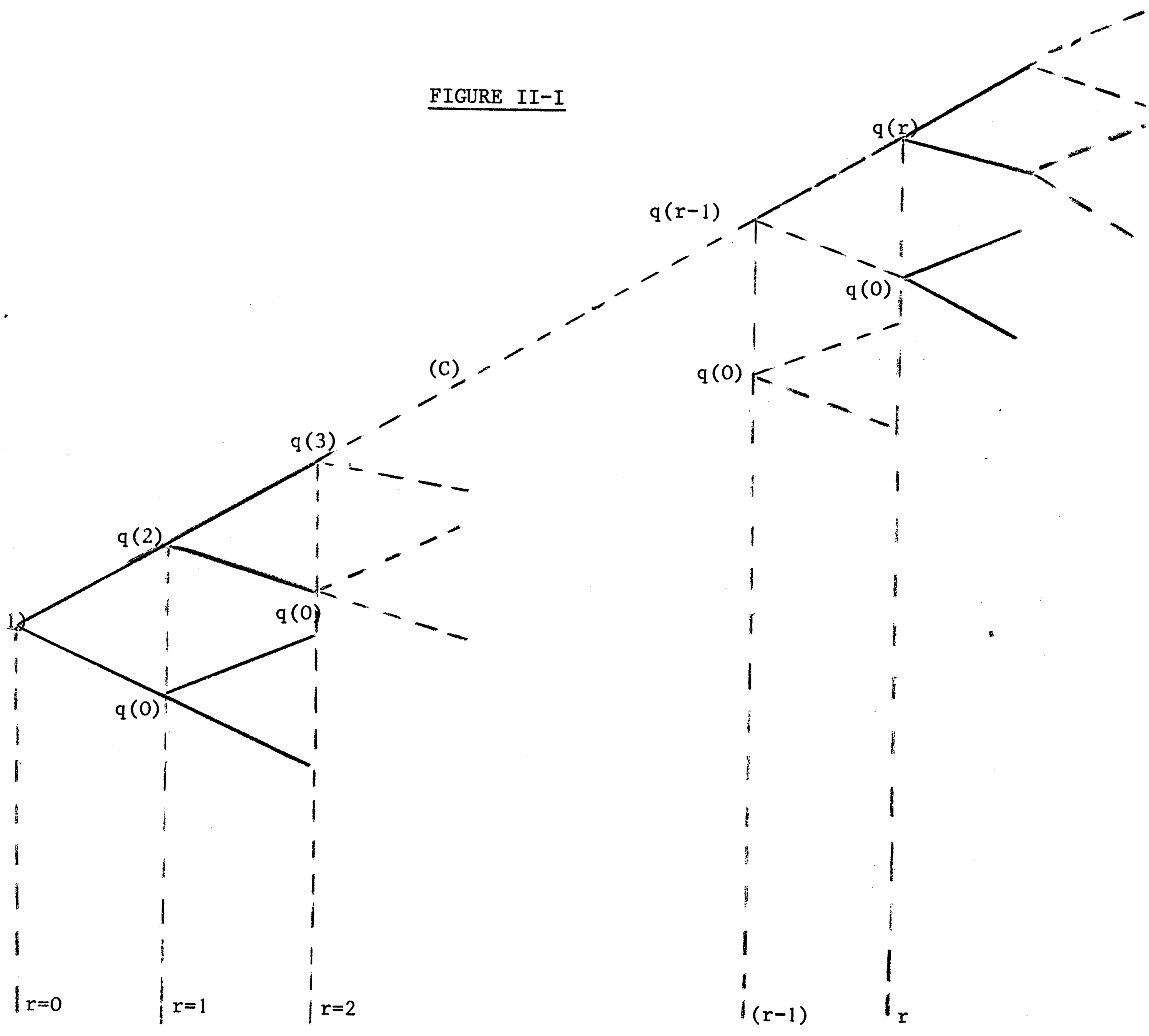
$$p(x_r) = \frac{1}{r}$$

et :

$$L[K(R),\pi] \geq \sum_{r=0}^{R-1} \frac{1}{r} .$$

Donc K n'est pas L-convergent pour π, et à fortiori n'est pas F-convergent pour π.

FIGURE II-I



II - INFLUENCE DE π SUR LA H-CONVERGENCE

PROPOSITION II-2

Etant donné un détecteur D *(resp.* $S \subset S(a,n)$*) et un élément* K *de* $\mathcal{H}$ (D) *(resp.*K $\in \mathcal{H}(S)$*), si* K *est* H-*convergent pour un élément* $\pi°$ *de* $\tilde{P}_n$ *, alors* K *est* H-*convergent pour tout élément* π *de* $\mathbb{P}_n$.

Démonstration

Remarquons tout d'abord qu'en raison de la relation (4), si K est H-convergent pour un élément $\pi°$ de $\tilde{\mathbb{P}}_n$, alors K est H-convergent pour tout élément π de $\check{\mathbb{P}}_n$.

Soit donc $\pi \in \mathbb{P}_n$, tel que l'ensemble

$$J = \{j \in \{1,\ldots,n\} \; : \; \pi_j \neq 0\}$$

soit inclus strictement dans $\{1,\ldots,n\}$, et montrons que K est H-convergent pour π, c'est-à-dire que :

$\forall\varepsilon > 0$, $\forall\eta > 0$, $\exists\bar{R}$ tel que : $R \geq \bar{R} \Rightarrow p(E(R\;;\;\pi,\varepsilon)) \leq \eta$.

Pour ceci, ε étant fixé, nous étudions l'ensemble $E(R\;;\;\pi,\varepsilon)$ par rapport à l'ensemble $E(R\;;\;\pi°,\varepsilon_o)$ pour des valeurs de ε_o bien choisies.

La démonstration de la proposition II-1 permet d'affirmer qu'il existe $\varepsilon_1 > 0$ tel que :

(5) | Si pour un sommet x de K, il existe un élément j de J tel que :

$$p_o(T_j|x) \geq 1-\varepsilon_1$$

alors :

$$p(T_j|x) \geq 1-\varepsilon.$$

Etant donné $\varepsilon_o > \varepsilon_1$, nous considérons l'ensemble $Z(R,\varepsilon_o)$ des sommets x de rang R dans K tels que

$$\left| \begin{array}{l} x \text{ n'est pas interprétable à } \varepsilon \text{ près pour } \pi \\ \\ x \text{ possède un ascendant (ou éventuellement lui-même) interprétable à } \varepsilon_o \\ \text{près pour } \pi^\circ. \end{array} \right.$$

Nous avons :

$$E(R\ ;\ \pi,\varepsilon) \subset E(R\ ;\ \pi^\circ,\varepsilon_o) \cup Z(R,\varepsilon_o).$$

1) Pour η fixé, nous allons montrer que pour ε_o bien choisi, on a :

$$p(Z(R,\varepsilon_o)) \leq \frac{\eta}{2} \qquad \forall R.$$

En raison de (5), tout élément x de $Z(R,\varepsilon_o)$ possède un ascendant y (ou éventuellement lui-même) tel que :

$$\exists j \notin J \text{ tel que } p_o(T_j|y) \geq 1-\varepsilon_o\ .$$

Donc il existe une famille C de sommets de K formant une partie de section telle que :

$$(6) \qquad \left\{ \begin{array}{l} -\forall y \in C \quad \exists j \notin J \text{ tel que } p_o(T_j|y) \geq 1-\varepsilon_o \\ \\ - \text{Tout élément x de } Z(R,\varepsilon_o) \text{ possède un ascendant (éventuellement} \\ \quad \text{lui-même) dans C.} \end{array} \right.$$

Par ailleurs, en posant :

$$\lambda = \text{Max } [\frac{\pi_j}{\pi_j^\circ} \mid j \in J]$$

il vient :

$$p(x) = \sum_{j\in J} \pi_j \Lambda_j(x) \leq \lambda \sum_{j\in J} \pi_j^\circ \Lambda_j(x) \qquad \forall x \in Z(R,\varepsilon_o).$$

En remarquant que

$$\pi_j^\circ \Lambda_j(x) = p_o(T_j|x)p_o(x)$$

nous déduisons :

$$p(x) \leq \lambda p_o(x) \sum_{j\in J} p_o(T_j|x) \qquad \forall x \in Z(R,\varepsilon_o).$$

D'où, en tenant compte de (6) :

$$p(Z(R,\varepsilon_o)) \leq \lambda \sum_{y\in C} \sum_{x\in \hat{\Gamma}_y \cap Z(R,\varepsilon_o)} p_o(x) \sum_{j\in J} p_o(T_j|x)$$

soit :

$$p(Z(R,\varepsilon_o)) \leq \lambda \sum_{y\in C} p_o(y) \sum_{x\in \hat{\Gamma}_y \cap Z(R,\varepsilon_o)} \frac{p_o(x)}{p_o(y)} \sum_{j\in J} p_o(T_j|x).$$

Or, d'après la propriété de "conservation du flux", on a :

$$\sum_{x\in \hat{\Gamma}_y \cap Z(R,\varepsilon_o)} \frac{p_o(x)}{p_o(y)} p_o(T_j|x) = p_o(T_j|y)$$

et d'après (6) :

$$\sum_{j\in J} p_o(T_j|y) \leq \varepsilon_o \qquad \forall y \in C.$$

Par conséquent

$$p(Z(R,\varepsilon_o)) \leq \lambda\, \varepsilon_o.$$

En posant alors

$$\varepsilon_2 = \frac{\eta}{2\lambda}$$

nous avons :

$$p(Z(R,\varepsilon_o)) \leq \frac{\eta}{2} \quad , \forall\varepsilon_o < \mathrm{Min}(\varepsilon_1,\varepsilon_2), \quad \forall R.$$

2) ε_o étant ainsi choisi, montrons que :

$$\exists\bar{R} \quad \text{tel que } R \geq \bar{R} \Rightarrow p(E(R\ ;\ \pi^\circ,\varepsilon_o)) \leq \frac{\eta}{2}.$$

Par hypothèse, K est H-convergent pour π°, donc :

$$p_o(E(R\ ;\ \pi^\circ,\varepsilon_o)) \xrightarrow[R\to\infty]{} 0.$$

Par ailleurs, en raison de la relation (4), nous savons qu'il existe $\mu > 0$ tel que :

$$p(x) \leq \mu \quad p_o(x) \qquad\qquad \forall x \in E(R\ ;\ \pi^\circ,\varepsilon_o)$$

donc :

$$p(E(R\ ;\ \pi^\circ,\varepsilon_o)) \xrightarrow[R\to\infty]{} 0.$$

$\cancel{X}$

CHAPITRE - III

ETUDE DE LA F-CONVERGENCE

INTRODUCTION

La F-convergence apparait comme la généralisation la plus naturelle des processus de décision arborescents, tels que les questionnaires. Etant donné un détecteur D, nous démontrons une condition suffisante (§ I) et une condition nécessaire (§ II), pour que pour tout élément π de $\mathbb{P}_n$, il existe un pseudoquestionnaire construit sur D qui soit F-convergent pour π. Bien que ces conditions ne soient pas identiques, elles sont toutefois suffisamment proches pour mettre en évidence l'exigence de ce mode de convergence. Nous illustrons encore ce fait par diverses remarques au paragraphe III.

I - UNE CONDITION SUFFISANTE D'EXISTENCE DE PSEUDOQUESTIONNAIRES F-CONVERGENTS

L'objet de ce paragraphe est de démontrer le théorème suivant :

THEOREME III-1

Etant donné un détecteur D*, une condition suffisante pour que, quel que soit* $\pi \in \mathbb{P}_n$*, il existe* $K \in \mathcal{K}(D)$ *qui soit* F*-convergent pour* π*, est que* D *vérifie la condition suivante :*

pour tout couple (j,j'), $j \in \{1,\ldots,n\}$, $j' \in \{1,\ldots,n\}$, $j \neq j'$

D *admet un point essentiellement atteint* s *qui vérifie l'hypothèse* $H^1_{j,j'}$ *suivante :*

$H^1_{j,j'}$

pour tout $i \in \{1,\ldots,a(s)\}$, on a :

$$s_i^j \neq s_i^{j'}$$

et il existe au plus une ligne i_o vérifiant :

$$s_{i_o}^j \; s_{i_o}^{j'} \neq 0.$$

Notation

Pour un détecteur D vérifiant les hypothèses du théorème III-1, on désignera par S un sous-ensemble de $\mathcal{S}$(D) tel que :

$\forall (j,j')$, $j \in \{1,\ldots,n\}$, $j' \in \{1,\ldots,n\}$, $j \neq j'$, $\exists s \in S$ vérifiant $H^1_{j,j'}$.

Pour tout élément σ de S, si $S-\{\sigma\}$ n'est pas vide, alors il existe un couple (j,j') tel que :

$$\forall s \in S-\{\sigma\},\ s \text{ ne vérifie pas } H^1_{j,j'}.$$

La démonstration du théorème III-1 comprend 4 parties ; nous avons ainsi dégagé 3 propositions intermédiaires, la dernière partie étant consacrée à la démonstration proprement dite du théorème III-1.

1) <u>Proposition III-1</u>

Si pour tout couple (j,j') *l'hypothèse* $H^1_{j,j'}$ *est vérifiée, alors pour tout élément* π *de* $\mathbb{P}_n$*, il existe un pseudoquestionnaire généralisé* χ *construit sur* S*, tel que :*

i) χ *possède au plus un chemin infini* $(x_0,\ldots,x_R,\ldots)$ *qui vérifie alors :*

$$\forall\varepsilon > 0,\ \exists R \text{ tel que } x_R \text{ soit interprétable à } \varepsilon \text{ près pour } \pi.$$

ii) En tout sommet terminal x *de* χ*, il existe* $j_0 \in \{1,\ldots,n\}$ *tel que :*

$$p(T_{j_0}|x) = 0.$$

<u>Démonstration</u>

Si π a une de ses composantes nulle, le résultat est évident, nous considérons donc un élément π de $\breve{\mathbb{P}}_n$.

Remarquons tout d'abord, que s'il existe $s \in S$, tel que s ait au moins un élément nul par ligne, alors la propriété est triviale ; tout pseudoquestionnaire $\chi = (A,v,B)$ de hauteur égale à 1, de racine x_0, tel que

$$v(x_0) = s$$

répond à la question.

Nous supposons donc que tout élément s de S possède une ligne d'éléments tous non nuls, et en raison de l'hypothèse faite sur D, tout élément de S ne possède qu'une telle ligne.

En posant

$$S = \{s(0),\ldots,s(m-1)\} \qquad (m \leq \frac{n(n-1)}{2})$$

nous supposerons, sans restriction de généralité, que pour tout ℓ, la ligne d'éléments tous non nuls de $s(\ell)$ est la ligne d'indice 1.

La matrice A d'éléments

$$a_\ell^j = s_1^j(\ell) \quad , \quad \ell \in \{0,\dots,m-1\}, \quad j \in \{1,\dots,n\}$$

vérifie donc les hypothèses de la proposition 1 de l'Annexe, par conséquent il existe

- $k \in \{1,\dots,n\}$

- des entiers $r_o,\dots,r_{m-1}$, positifs non tous nuls

- $\delta > 0$

tels qu'en posant

$$\rho = \sum_{\ell=0}^{m-1} r_\ell$$

on ait :

$$(1) \qquad \prod_{\ell=0}^{m-1} \left(\frac{s_1^k(\ell)}{s_1^j(\ell)}\right)^{r_\ell/\rho} \geqslant 1+\delta \qquad \forall j \in \{1,\dots,n\},\ j \neq k.$$

Considérons alors un pseudoquestionnaire généralisé $\chi = (A,v,B)$ construit sur S par récurrence de la façon suivante :

Etant donné un sommet x de χ, de rang $r(x)$

- x est terminal si et seulement si, il existe $j_o \in \{1,\dots,n\}$, tel que :

$$p(T_{j_o}|x) = 0$$

- si x n'est pas terminal, soit ℓ tel que :

$$\lambda\rho + \sum_{i=0}^{\ell-1} r_i \leq r(x) < \lambda\rho + \sum_{i=0}^{\ell} r_i$$

où λ est un entier positif ou nul ; on posera alors :

$$v(x) = s(\ell-1) \text{ ceci pour } \ell \in \{1,\ldots,m\}.$$

Il est clair que χ possède un seul chemin infini $(x_o,\ldots,x_R,\ldots)$, et en raison de (1), on a quel que soit $R = \lambda\rho$:

$$\frac{p(T_k|x_R)}{p(T_j|x_R)} \geq \frac{\pi_k}{\pi_j} (1+\delta)^{\lambda\rho} \qquad \forall j \in \{1,\ldots,n\},\ j \neq k$$

soit :

$$p(T_k|x_{\lambda\rho}) \xrightarrow[\lambda\to\infty]{} 1$$

donc χ répond à la question.

$\cancel{X}$

2) Proposition III-2

Si pour tout couple (j,j') *, l'hypothèse* $H^1_{j,j'}$ *est vérifiée, alors pour tout élément* π *de* $\mathbb{P}_n$*, il existe un élément* K *de* $\mathcal{H}(S)$ *qui est F-convergent pour* π.

Démonstration

Nous allons montrer ce résultat par récurrence sur le nombre t de coordonnées de π qui sont non nulles.

- <u>Cas où t = 2</u>

Soit, par exemple : $\pi_1 \neq 0$, $\pi_2 \neq 0$.

Considérons un élément s de S qui vérifie $H^1_{1,2}$; nous supposons qu'il existe $i_o \in \{1,\ldots,a(s)\}$ tel que :

$$s^1_{i_o} \; s^2_{i_o} \neq 0$$

sinon, la propriété est triviale.

Considérons alors un pseudoquestionnaire généralisé $K = (A,v,B) \in \mathcal{H}(S,J)$, où $J = \{1,2\}$ (cf. Remarque de I-5-3), tel que :

pour tout sommet y non terminal de K, on ait :

$$v(y) = s.$$

Il est clair que K possède un seul chemin infini $(x_o,\ldots,x_R,\ldots)$ et que, sans restriction de généralité, si l'on suppose :

$$s^2_{i_o} > s^1_{i_o}$$

on a :

$$p(T_1 | x_R) \xrightarrow[R\to\infty]{} 0.$$

Donc K est F-convergent pour π.

- <u>Cas où t = n</u>

Nous faisons l'hypothèse de récurrence suivante :

Pour tout élément π *ayant au plus* (n-1) *composantes non nulles, il existe un élément de* $\mathcal{H}(S)$ *qui soit* F-*convergent pour* π.

Soit alors π un élément de $\tilde{\mathbb{P}}_n$; le pseudoquestionnaire généralisé cherché va être obtenu par prolongement du pseudoquestionnaire χ construit dans la proposition III-1, de la façon suivante :

Pour tout sommet y terminal de χ, $P(T|y)$ a au moins une composante nulle, donc, d'après l'hypothèse de récurrence, il existe

$$K_y \in \mathcal{K}(S)$$

qui soit F-convergent pour $P(T|y)$.

Le pseudoquestionnaire généralisé $K \in \mathcal{K}(S)$ cherché est alors le prolongement de χ, tel que pour tout sommet y terminal de χ, le sous-pseudoquestionnaire de racine y dans K soit K_y.

~~X~~

3) Proposition III-3

Si pour tout couple (j,j') , l'hypothèse $H^1_{j,j'}$ *est vérifiée, alors : quel que soit* $D' \subset D$, *tel que* $|D-D'|$ *soit fini, quel que soit* $\varepsilon > 0$, *et quel que soit* $\pi \in \mathbb{P}_n$ *, il existe un pseudo-questionnaire* K_o *de hauteur finie, construit sur* D', *dont tous les sommets terminaux sont interprétables à* ε *près pour* π.

Démonstration

Etant donné $\pi \in \mathbb{P}_n$, nous considérons un élément K de $\mathcal{K}(S)$ F-convergent pour π:

$$\forall \varepsilon > 0 \qquad K(\pi, \tfrac{\varepsilon}{2}) \text{ est de hauteur finie.}$$

Posons :

$$K(\pi, \tfrac{\varepsilon}{2}) = (A, v, B).$$

Considérons un pseudoquestionnaire généralisé $K' = (A, v', B)$; nous avons par continuité :

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ tel que } \begin{cases} \|v(y)-v'(y)\| \leq \delta \\ \\ \forall y \in Y \end{cases} \Rightarrow \begin{cases} \|\mathcal{P}(T|x)-\mathcal{P}'(T|x)\| \leq \frac{\varepsilon}{2} \\ \\ \forall x \in \mathcal{A} \end{cases}$$

où :

- $\mathcal{P}(T|x)$ et $\mathcal{P}'(T|x)$ désignent les vecteurs de probabilité associés respectivement aux sommets x de K opérant sur π et de K' opérant sur π

- Y désigne l'ensemble des sommets non terminaux de $\mathcal{A}$

- $\|\xi\|$ désigne une norme d'un élément ξ de $\mathbb{R}^k$, k entier positif.

Par conséquent, tous les sommets terminaux d'un tel pseudoquestionnaire K' sont interprétables à ε près pour π.

Par ailleurs, étant donné $D' \subset D$, tel que $|D-D'|$ soit fini, les ensembles $\mathcal{S}(D)$ et $\mathcal{S}(D')$ sont identiques.
Donc étant donné :

$$y \in Y \quad \text{et} \quad v(y) \in S$$

quel que soit $\delta > 0$, il existe une infinité d'éléments q de D' vérifiant :

$$\|\alpha(q)-v(y)\| \leq \delta.$$

Par suite, il existe une application u de Y dans D' telle que :

$$\|\alpha o u(y)-v(y)\| \leq \delta \qquad (\forall y \in Y)$$

et le pseudoquestionnaire $K_o = (\mathcal{A},u,B)$, de hauteur finie, a tous ses sommets terminaux interprétables à ε près pour π.

~~X~~

4) Démonstration du théorème III-1

Etant donné $\pi \in \mathbb{P}_n$, nous allons construire par un procédé de récurrence un élément K de $\mathcal{K}(D)$ qui soit F-convergent pour π.

Etant donné $\varepsilon > 0$, $\pi \in \mathbb{P}_n$, et $D' \subset D$ tel que $|D-D'|$ soit fini, nous désignons par

$$G(D',\pi,\varepsilon)$$

l'ensemble des pseudoquestionnaires K' construits sur D' tels que :

- K' soit de hauteur finie
- les sommets terminaux de K' soient les sommets interprétables à ε près pour π.

D'après la proposition III-3, cet ensemble n'est pas vide.

Nous considérons une suite $\{\varepsilon_k\}$, $k \in \mathbb{N}$, de nombres positifs, décroissante et tendant vers zéro, et construisons un pseudoquestionnaire $K = (\mathcal{A},u,B)$ répondant à la question de la façon suivante :

★ Soit ε_1 le premier élément de la suite $\{\varepsilon_k\}$, et K^1 un élément de $G(D,\pi,\varepsilon_1)$.

K sera un prolongement de K^1. Soit x un sommet terminal de K^1, x est interprétable à ε_1 près pour π, soit alors ε_ℓ le plus grand élément de la suite $\{\varepsilon_k\}$ tel que x ne soit pas interprétable à ε_ℓ près pour π $(\ell \geq 2)$.
Le sous-pseudoquestionnaire de racine x dans K sera un prolongement d'un élément de $G(D_x, P(T|x),\varepsilon_\ell)$ où

$$D_x = \{q \in D : u(y) \neq q \quad , \quad \forall y \in A(x)\}$$

(A(x) désignant l'ensemble des ascendants de x).

★ Plus généralement, étant donné un sommet x de K, interprétable à ε_i près pour π, et dont aucun ascendant ne l'a été, soit $\varepsilon_{\ell'}$ le plus grand élément de la suite $\{\varepsilon_k\}$ tel que x ne soit pas interprétable à $\varepsilon_{\ell'}$ près pour π ($\ell' \geq i+1$).

Le sous-pseudoquestionnaire de racine x dans K sera un prolongement d'un élément de $G(D_x, \mathcal{P}(T|x), \varepsilon_{\ell'})$.

Il est clair que K ainsi construit est bien un élément de $\mathcal{K}(D)$. Par ailleurs, étant donné $\varepsilon > 0$, il existe un élément ε_i de la suite $\{\varepsilon_k\}$ tel que :

$$\varepsilon_i \leq \varepsilon < \varepsilon_{i-1}.$$

Or par construction de K, le pseudoquestionnaire $K(\pi, \varepsilon_i)$ est de hauteur finie, donc $K(\pi, \varepsilon)$ l'est aussi.

~~X~~

II - UNE CONDITION NECESSAIRE D'EXISTENCE DE PSEUDOQUESTIONNAIRES F-CONVERGENTS

Dans tout ce paragraphe, nous étudierons des détecteurs *réguliers* définis comme suit :

DEFINITION III-1

Un détecteur D *est dit régulier, si tout élément* s *de* $\mathcal{S}$(D) *vérifie la condition suivante :*
s'il existe un couple (j,j') *tel que* $s^j = s^{j'}$ *, alors :*

$$s_i^j \neq 0 \qquad \forall i \in \{1,\dots,a(s)\}.$$

THEOREME III-2

Etant donné un détecteur D *régulier, non trivial, une condition nécessaire pour que pour tout* $\pi \in \mathbb{P}_n$, *il existe* $K \in \mathcal{K}(D)$ *qui soit* F*-convergent pour* π, *est que* D *vérifie la condition suivante :*

pour tout couple (j,j') , $j \in \{1,\dots,n\}$, $j' \in \{1,\dots,n\}$, $j \neq j'$, D *admet un point essentiellement atteint* s *vérifiant l'hypothèse* $H^2_{j,j'}$ *suivante :*

$H^2_{j,j'}$

Si $I(s) = \{i \in \{1,\dots,a(s)\} : s_i^j \neq 0 \ , \ s_i^{j'} \neq 0\}$
alors :

$$s_i^j \geq s_i^{j'} \quad \forall i \in I(s)$$

de plus il existe $i_o \in \{1,\dots,a(s)\}$ tel que :

$$s_{i_o}^j > s_{i_o}^{j'}$$

Nous démontrerons ce théorème sous la forme suivante :

THEOREME III-2'

Etant donné un détecteur D *régulier, non trivial, s'il existe un couple* (j_1,j_2) , *tel qu'aucun élément de* $S(D)$ *ne vérifie l'hypothèse* $H^2_{j_1,j_2}$, *alors :*

il existe un élément π *de* $\mathbb{P}_n$, *tel que pour tout élément* K *de* $\mathcal{K}(D)$, *il existe* $\eta > 0$ *et un entier* R_o *tels que :*

pour tout R, $R \geq R_o$, K *admet un sommet* x *de rang* R, *tel que dans* K *opérant sur* π, *on ait :*

$$\eta \leq p(T_{j_1} | x) \leq 1-\eta.$$

1) Propriétés préliminaires

Nous considérons la fonction

$$\Psi(w,t,\lambda) = \frac{\lambda w}{\lambda w+(1-\lambda)t}$$

pour $(w,t,\lambda) \in]0,1] \times]0,1] \times [0,1]$.

LEMME III-1

Etant donné $\delta > 0$, si les éléments (w_1,t_1), (w_2,t_2) de $]0,1]^2$ vérifient :

$$(2) \quad \begin{cases} w_1-t_1 \geq \delta \\ t_2-w_2 \geq \delta \\ \delta \leq t_1 \leq w_1 \leq 1 \\ \delta \leq w_2 \leq t_2 \leq 1 \end{cases}$$

alors il existe $\eta_o > 0$, tel que pour tout η, $\eta \leq \eta_o$, on ait :

$$\lambda \in [\eta,1-\eta] \Rightarrow (\Psi(w_1,t_1,\lambda) \in [\eta,1-\eta] \text{ ou } \Psi(w_2,t_2,\lambda) \in [\eta,1-\eta]).$$

Démonstration

Pour w_1, t_1, w_2, t_2 fixés et $\eta \in]0,1[$, considérons le sous-ensemble $V(\eta)$ de $[0,1]$ défini par :

$$V(\eta) = \{\lambda \in [\eta,1-\eta] \ : \Psi(w_1,t_1,\lambda) > 1-\eta \quad \text{et} \quad \Psi(w_2,t_2,\eta) < \eta\} .$$

En remarquant que si (w_1,t_1), (w_2,t_2) vérifient (2) alors on a :

$$\Psi(w_1,t_1,\lambda) > \lambda > \Psi(w_2,t_2,\lambda)$$

il nous suffit de montrer qu'il existe $\eta_o > 0$, tel que pour tout $\eta \leq \eta_o$, on a :

$V(\eta) = \emptyset$ pour tout quadruplet (w_1, t_1, w_2, t_2) vérifiant (2).

Or :

$$\begin{cases} \Psi(w_1, t_1, \lambda) > 1-\eta \iff \lambda > \dfrac{t_1(1-\eta)}{t_1+\eta(w_1-t_1)} \\ \Psi(w_2, t_2, \eta) < \eta \iff \lambda < \dfrac{t_2\eta}{w_2+\eta(t_2-w_2)} \end{cases}$$

En posant :

$$\varphi(\eta) = \frac{t_1(1-\eta)}{t_1+\eta(w_1-t_1)} \quad , \quad \Phi(\eta) = \frac{t_2\eta}{w_2+\eta(t_2-w_2)}$$

il vient :

$$V(\eta) = \{\lambda \in [\eta, 1-\eta] \ : \varphi(\eta) < \lambda < \Phi(\eta)\}$$

les fonctions φ et Φ étant représentées sur la figure III-1.

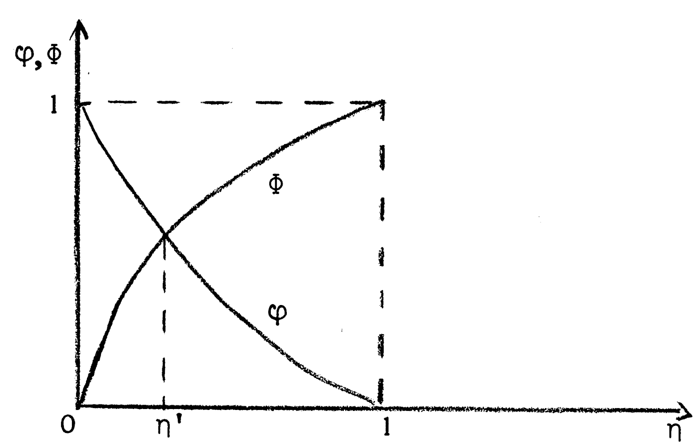


Figure III-1

Soit $\eta'(w_1,t_1,w_2,t_2)$ l'abscisse de l'intersection des courbes représentative de φ et Φ. Pour w_1,t_1,w_2,t_2 fixés, on a :

$$\eta < \eta' \Rightarrow \varphi(\eta) > \Phi(\eta) \Rightarrow V(\eta) = \emptyset$$

et

$$\eta'(w_1,t_1,w_2,t_2) = \frac{-t_1w_2+\sqrt{w_1t_1w_2t_2}}{t_2(w_1-w_2)+(t_2-t_1)w_2}\ .$$

Il suffit alors de remarquer que la fonction η' est continue et non nulle sur le compact défini par (2).

$\cancel{X}$

LEMME III-2

Etant donné $\delta > 0$, *il existe* $\beta_o > 0$, *tel que si les éléments* (w_1,t_1), (w_2,t_2) *de* $[0,1]^2$ *vérifient :*

$$\begin{cases} w_1 \geq t_1 \geq \delta \\ t_2 \geq w_2 \geq \delta \\ |w_i-t_i| \leq \beta_o \qquad i \in \{1,2\} \end{cases}$$

alors pour tout $\eta < \frac{1}{4}$, *on a :*

$$\lambda \in [\eta,1-\eta] \Rightarrow (\Psi(w_1,t_1,\lambda) \in [\eta,1-\eta] \text{ ou } \Psi(w_2,t_2,\lambda) \in [\eta,1-\eta]).$$

Démonstration

Etant donné $\delta \in]0,1[$, et $\beta > 0$, nous considérons le compact $\Delta(\beta)$ de $[0,1]^3$ défini par :

$$\begin{cases} w \in [\delta,1],\ t \in [\delta,1],\ \lambda \in [0,1] \\ |w-t| \leq \beta. \end{cases}$$

La fonction Ψ tend vers λ quand $|w-t|$ tend vers zéro, et elle est uniformément continue par rapport à (w,t,λ) sur $\Delta(\beta)$, par conséquent, il existe un nombre $\beta_o \in]0,\beta]$ tel que :

$$|w-t| \leq \beta_o \Rightarrow |\Psi(w,t,\lambda)-\lambda| < \frac{1}{4} \quad \forall\lambda \in [0,1].$$

Soient alors (w_1,t_1,λ), (w_2,t_2,λ) deux éléments de $\Delta(\beta_o)$ tels que :

$$\begin{cases} w_1 \geq t_1 \\ t_2 \geq w_2. \end{cases}$$

Quel que soit $\lambda \in [0,1]$, on a :

$$\begin{cases} 0 < \Psi(w_1,t_1,\lambda)-\lambda < \frac{1}{4} \\ 0 < \lambda-\Psi(w_2,t_2,\lambda) < \frac{1}{4}. \end{cases}$$

Par conséquent, pour tout $\eta < \frac{1}{4}$, si $\lambda \in [\eta,1-\eta]$ on a :

$$\Psi(w_1,t_1,\lambda) \in [\eta,1-\eta] \text{ ou } \Psi(w_2,t_2,\lambda) \in [\eta,1-\eta].$$

~~X~~

2) Démonstration du théorème III-2'

Nous supposons $j_1 = 1$, $j_2 = 2$. En raison de l'hypothèse de régularité faite sur D, tout élément s de $\mathcal{S}(D)$ vérifie l'une des deux conditions (C_1), (C_2) suivantes :

$$\begin{cases} (C_1) : \exists i,i' \in \{1,\dots,a(s)\} \text{ tels que } \begin{cases} s_i^1 > s_i^2 > 0 \\ s_{i'}^2 > s_{i'}^1 > 0 \end{cases} \\ (C_2) : s_i^1 = s_i^2 \neq 0 \quad \forall i \in \{1,\dots,a(s)\}. \end{cases}$$

Nous considérons une partition de D en D_1, D_2, tels que tous les éléments de $\mathcal{S}(D_1)$ vérifient (C_1), et tous les éléments de $\mathcal{S}(D_2)$ vérifient (C_2).
L'ensemble $\mathcal{S}(D_1)$ étant compact dans $\mathcal{S}(a,n)$, il existe $\delta_1 > 0$, tel que si l'on désigne par $\tilde{D}_1$ l'ensemble des éléments q de D_1 vérifiant

$$(3)\quad \begin{cases} \alpha_1^1(q)-\alpha_1^2(q) \geq \delta_1 \\ \alpha_2^2(q)-\alpha_2^1(q) \geq \delta_1 \\ \delta_1 \leq \alpha_i^j(q) \leq 1 \qquad\qquad i \in \{1,2\},\ j \in \{1,2\} \end{cases}$$

alors l'ensemble $D_1-\tilde{D}_1$ est fini. (Nous avons supposé, sans restriction de généralité que les lignes i et i' concernées par la condition (C_1) étaient les lignes 1 et 2 pour tous les éléments de $\mathcal{S}(D_1)$).

Par ailleurs, pour tout élément q de D_2, il existe deux issues q_{i_1} et q_{i_2} telles que :

$$\begin{cases} \alpha_{i_1}^1(q) \geq \alpha_{i_1}^2(q) \\ \alpha_{i_2}^2(q) \geq \alpha_{i_2}^1(q). \end{cases}$$

Nous supposerons, sans restreindre la généralité,que pour tout élément q de D, on a : $i_1 = 1$, $i_2 = 2$.
Il existe $\delta_2 > 0$, tel que si l'on désigne par D_2' l'ensemble des éléments q de D_2 vérifiant

$$\begin{cases} \alpha_1^1(q) \geq \alpha_1^2(q) \geq \delta_2 \\ \alpha_2^2(q) \geq \alpha_2^1(q) \geq \delta_2 \end{cases}$$

alors l'ensemble D_2-D_2' est fini.

Etant donné δ_2 ainsi fixé, soit β_o qui lui correspond dans le lemme III-2.
L'ensemble $\tilde{D}_2$ des éléments q de D_2' vérifiant :

$$|\alpha_i^1(q)-\alpha_i^2(q)| \leq \beta_o \qquad \forall i \in \{1,2\}$$

est tel que l'ensemble $D_2'-\tilde{D}_2$ soit fini, donc tel que $|D_2-\tilde{D}_2|$ soit fini.

Considérons alors un élément $K = (\mathcal{A}u,B) \in \mathcal{H}(D)$, et un élément π de $\mathbb{P}_n$ tel que :

$$\pi_j \neq 0 \qquad \forall j \in \{1,2\}.$$

En raison de ce qui précède, D étant supposé non trivial, il existe un entier R_o, tel que pour tout sommet x de K de rang supérieur ou égal à R_o, on ait :

$$u(x) \in \tilde{D}_1 \cup \tilde{D}_2.$$

Si $q = u(x) \in \tilde{D}_1$, nous savons en raison de (3) et du lemme III-1 qu'il existe η_o^1 tel que pour tout $\eta \leq \eta_o^1$ si

$$p(T_1|x) \in [\eta, 1-\eta]$$

alors x admet un successeur y tel que

$$p(T_1|y) \in [\eta, 1-\eta].$$

Soit alors un sommet x de rang $r(x) \geq R_o$, tel que $p(T_1|x) > 0$, et posons :

$$\eta_o = \mathrm{Min}(\eta_o^1, \tfrac{1}{4}, p(T_1|x)).$$

En raison des lemmes III-1 et III-2, il est clair qu'il existe un chemin d'origine x dans K, tel que pour tout sommet y de ce chemin, on ait :

$$p(T_1|y) \in [\eta_o, 1-\eta_o].$$

$\cancel{x}$

III - REMARQUES DIVERSES

1) Une conjecture de condition suffisante

Il semble que l'on puisse envisager de démontrer l'existence de pseudo-questionnaires F-convergents sous des conditions sur D moins exigeantes que celles du théorème III-1.
Nous proposons la conjecture suivante :

"Une condition suffisante pour que pour tout $\pi \in \mathbb{P}_n$, il existe un élément de $\mathcal{H}(D)$ qui soit F-convergent pour π, est que D vérifie la condition suivante :

pour tout couple (j,j'), $j \in \{1,\ldots,n\}$, $j' \in \{1,\ldots,n\}$, $j \neq j'$, D admet un point essentiellement atteint s qui vérifie l'hypothèse $\tilde{H}^1_{j,j'}$ suivante :

$$\tilde{H}^1_{j,j'} \left| \begin{array}{ll} s_i^j \neq s_i^{j'} & \forall i \in \{1,\ldots,a(s)\} \\ \\ \text{et si } I(s) = \{i \in \{1,\ldots,a(s)\} : s_i^j \neq 0,\ s_i^{j'} \neq 0\} & \\ \\ \text{alors : } s_i^j > s_i^{j'} \quad \forall i \in I(s). & \end{array} \right.$$ "

2) Etude d'un détecteur non régulier

Considérons le cas où n est égal à 2, et où tous les éléments de D sont des questions dichotomiques. Nous supposons que tous les éléments de $\mathcal{S}(D)$ sont égaux à

$$\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}$$

Dans un tel cas, ni le théorème 3.1, ni même la conjecture ci-dessus ne permettent de conclure à l'existence d'un élément de $\mathcal{H}(D)$ qui soit F-convergent.

PROPOSITION III-4

S'il existe une suite $\{q(\ell)\}$, $\ell \in \mathbb{N}$, *d'éléments de* D, *telle que :*

i) $$w_\ell = \mathrm{Min}\left[\frac{\alpha_2^1(q(\ell))}{\alpha_2^2(q(\ell))}, \frac{\alpha_2^2(q(\ell))}{\alpha_2^1(q(\ell))}\right] \xrightarrow[\ell\to\infty]{} 0$$

ii) la suite $\sum_\ell |\alpha_2^1(q(\ell))-\alpha_2^2(q(\ell))|$ *diverge*

alors il existe $K \in \mathcal{K}(D)$ *qui soit* F-*convergent pour tout élément* π *de* $\mathbb{P}_n$.

Démonstration

Nous supposons, sans restreindre la généralité, que pour tout ℓ

$$\mathrm{Min}\left[\frac{\alpha_2^1(q(\ell))}{\alpha_2^2(q(\ell))}, \frac{\alpha_2^2(q(\ell))}{\alpha_2^1(q(\ell))}\right] = \frac{\alpha_2^1(q(\ell))}{\alpha_2^2(q(\ell))} .$$

Considérons un pseudoquestionnaire $K = (A,u,B) \in \mathcal{K}(D)$, tel que pour tout sommet y non terminal de K et de rang R, on ait :

$$u(y) = q(R).$$

Etant donné un élément π de $\mathbb{P}_2$, et $\varepsilon > 0$, montrons que $K(\pi,\varepsilon)$ est de hauteur finie.

Désignons par $E_1(R)$ (resp. $E_2(R)$) l'ensemble des sommets y_1 (resp. y_2) de rang R tels que :

$$b_x((x,y_1)) = q_1(R-1) \quad (\text{resp. } b_x((x,y_2)) = q_2(R-1))$$

où x désigne l'antécédent de y_1 (resp. y_2).

Montrons tout d'abord, qu'il existe $R_2 > 0$ tel que pour tout $R > R_2$, tous les éléments de $E_2(R)$ soient interprétables à ε près pour π :

$$p(T_1|y_2) = \frac{\dfrac{\alpha_2^1(q(R-1))}{\alpha_2^2(q(R-1))}\, p(T_1|x)}{\dfrac{\alpha_2^1(q(R-1))}{\alpha_2^2(q(R-1))}\, p(T_1|x)+p(T_2|x)}$$

D'après l'hypothèse i), nous avons :

$$\forall\eta > 0 \quad \exists R_2 \text{ tel que } \forall R'>R_2 \quad \frac{\alpha_2^1(q(R'))}{\alpha_2^2(q(R'))} < \eta .$$

Par ailleurs, x n'étant pas terminal dans $K(\pi,\varepsilon)$, on a :

$$p(T_1|x) \leq 1-\varepsilon, \qquad p(T_2|x) \geq \varepsilon.$$

Donc, pour tout $R \geq R_2+1$, on a :

$$p(T_1|y_2) \leq \eta\frac{(1-\varepsilon)}{\varepsilon}.$$

Il suffit donc de choisir η tel que :

$$\eta \leq \frac{\varepsilon^2}{1-\varepsilon}$$

pour que pour tout $R > R_2$, les éléments de $E_2(R)$ soient interprétables à ε près pour π.

Montrons maintenant qu'il existe $R_1 > 0$, tel que pour tout $R \geq R_1$, les sommets de $E_1(R)$ soient interprétables à ε près pour π.

Pour $R > R_2$, considérons un élément y_1 de $E_1(R)$, et soit y son ascendant de rang R_2; on a :

$$\frac{p(T_1|y_1)}{p(T_2|y_1)} = \frac{p(T_1|y)}{p(T_2|y)} \prod_{r=R_2}^{R-1} \frac{\alpha_1^1(q(r))}{\alpha_1^2(q(r))} .$$

En raison de l'hypothèse ii), ce rapport tend vers 0 ou vers l'infini lorsque R tend vers l'infini, ce qui achève la démonstration.

~~X~~

3) Exigence de la F-convergence

★ Les conditions des théorèmes III-1 et III-2 sont des conditions exigeantes ; en particulier, lorsque n = 2, a = 2, nous pouvons en donner une interprétation géométrique.

Nous remarquons tout d'abord que dans ce cas, les conditions nécessaires et suffisantes démontrées sont les mêmes, en raison du fait que :

$$\alpha_1^j(q)+\alpha_2^j(q) = 1 \qquad \forall j \in \{1,2\} \quad , \forall q \in D.$$

Une question q peut être représentée dans $\mathbb{R}^2$ par le point de coordonnées $(\alpha_1^1(q), \alpha_1^2(q))$. Un détecteur sera alors représenté par un ensemble de points du carré $[0,1] \times [0,1]$.

Un détecteur D régulier est tel que, ni le point (0,0), ni le point (1,1) ne sont essentiellement atteints par D.

Alors, étant donné D régulier, une condition nécessaire et suffisante pour que pour tout élément π de $\mathbb{P}_2$, il existe un élément K de $\mathcal{K}(D)$ qui soit F-convergent pour π, est que D admette un point essentiellement atteint sur *la frontière du carré.*

★ Par ailleurs, considérons l'exemple suivant, où n = 2, et où le détecteur D infini est tel que

$$\alpha(q) = \begin{pmatrix} \frac{1}{3} & \frac{1}{2} \\ \frac{2}{3} & \frac{1}{4} \\ 0 & \frac{1}{4} \end{pmatrix} \qquad \forall q \in D.$$

D'après le théorème III-2, étant donné $\pi \in \tilde{\mathbb{P}}_2$, il n'existe pas de pseudo-questionnaire $K \in \mathcal{K}(D)$ F-convergent pour π.

Considérons alors le détecteur D', déduit de D de la façon suivante : à toute question q de D, correspondant à la partition (q_1,q_2,q_3) de Ω, on associe la question q' correspondant à la partition $(q_1 \cup q_2, q_3)$ de Ω ; D' est alors l'ensemble des questions q' ainsi définies.

On a :

$$\alpha(q') = \begin{pmatrix} 1 & \frac{3}{4} \\ 0 & \frac{1}{4} \end{pmatrix} \quad \forall q' \in D'$$

et d'après le théorème III-1, quel que soit $\pi \in \mathbb{P}_2$, il existe un élément de $\mathcal{K}(D')$, qui soit F-convergent pour π.

Pour expliquer ce phénomène, il faut revenir à la définition des risques d'erreur de décision définis en I-5-3.

Etant donné ε, soit $\rho_2(R)$ le risque d'aboutir à un sommet de rang R dont la distribution de probabilité ne soit pas interprétable à ε près.

Pour qu'il y ait F-convergence, il faut que :

(4) $\qquad \forall \varepsilon > 0, \quad \exists R : \quad \rho_2(R) = 0.$

Ainsi un élément $K' \in \mathcal{K}(D')$ aura au plus 2^R sommets de rang R, alors qu'un élément $K \in \mathcal{K}(D)$ en a au plus 3^R ; tandis que (4) est réalisé pour K', il ne l'est plus pour K, car il apparait dans le passage de K' à K des sommets de rang R supplémentaires. Ces sommets sont de "faible probabilité" certes, mais leur distribution de probabilité n'est pas interprétable à ε près.

★ La F-convergence apparaît donc comme un mode de convergence très exigeant, et généralement peu intéressant dans la pratique. Elle se présente comme la généralisation la plus proche des processus finis de décision arborescents, tels que les questionnaires, où les deux risques d'erreur sont nuls.

ANNEXE DU CHAPITRE III

PROPOSITION 1

Si une matrice (m,n), A, *d'éléments réels* a_i^j, *vérifie les hypothèses suivantes :*

i) $a_i^j > 0 \quad \forall i \in \{1,\dots,m\}, \quad \forall j \in \{1,\dots,n\}$

ii) $\forall j \in \{1,\dots,n\}, \forall j' \in \{1,\dots,n\}, j \neq j', \exists i \in \{1,\dots,m\} : a_i^j \neq a_i^{j'}$

alors, en notant k *un élément de* $\{1,\dots,n\}$ *tel que :*

$$\sum_{i=1}^{m} a_i^k = \mathrm{Max}\left[\sum_{i=1}^{m} a_i^j \mid j \in \{1,\dots,n\}\right]$$

il existe $\delta > 0$, *tel que l'ensemble*

$$E(\delta) = \{x \in \mathbb{P}_m : \prod_{i=1}^{m} \left(\frac{a_i^k}{a_i^j}\right)^{x_i} \geq 1+\delta \qquad \forall j \in \{1,\dots,n\},\ j \neq k\}$$

soit non vide et contienne un point de coordonnées toutes rationnelles.

Démonstration

Il nous suffit de montrer que l'ensemble V,

$$V = \{x \in \mathbb{P}_m : \sum_{i=1}^{m} x_i \,\mathrm{Log}\, \frac{a_i^k}{a_i^j} > 0 \qquad \forall j \in \{1,\dots,n\},\ j \neq k\}$$

est non vide et contient un point de coordonnées toutes rationnelles.

a) Montrons que l'élément $\bar{x}$ de composantes

$$\bar{x}_i = \frac{a_i^k}{\sum_{i=1}^{m} a_i^k}$$

appartient à V.

La fonction Log étant concave, nous avons pour tout $x \in \mathbb{P}_m$:

$$\sum_{i=1}^{m} x_i \operatorname{Log} \frac{a_i^k}{a_i^j} \geq \operatorname{Log} \frac{1}{\sum_{i=1}^{m} x_i \frac{a_i^j}{a_i^k}} \qquad \forall j \in \{1,\dots,n\}.$$

Par conséquent :

$$\sum_{i=1}^{m} \bar{x}_i \operatorname{Log} \frac{a_i^k}{a_i^j} \geq 0 \qquad \forall j \in \{1,\dots,n\}.$$

De plus, si j est un indice de $\{1,\dots,n\}$ tel que

$$\sum_{i=1}^{m} a_i^j < \sum_{i=1}^{m} a_i^k$$

alors, nous avons :

$$\sum_{i=1}^{m} \bar{x}_i \operatorname{Log} \frac{a_i^k}{a_i^j} > 0.$$

Soit maintenant un indice j de $\{1,\dots,n\}$ tel que

$$\sum_{i=1}^{m} a_i^j = \sum_{i=1}^{m} a_i^k .$$

Posons :

$$b_i^j = \frac{a_i^j}{\sum_{i=1}^{m} a_i^j}$$

il vient :

$$\sum_{i=1}^{m} \bar{x}_i \operatorname{Log} \frac{a_i^k}{a_i^j} = \sum_{i=1}^{m} b_i^k \operatorname{Log} b_i^k - \sum_{i=1}^{m} b_i^k \operatorname{Log} b_i^j.$$

Or, d'après le lemme utilisé par Ash ([1], page 16), si p et q sont des éléments de $\mathbb{P}_m$, on a :

$$\sum_{i=1}^{m} p_i \operatorname{Log} q_i \leq \sum_{i=1}^{m} p_i \operatorname{Log} p_i$$

avec égalité, si et seulement si : $p_i = q_i \quad \forall i \in \{1,\dots,m\}$.
Par conséquent, en raison de l'hypothèse ii), on a :

$$\sum_{i=1}^{m} \bar{x}_i \operatorname{Log} \frac{a_i^k}{a_i^j} > 0.$$

b) Montrons maintenant que V contient un point de coordonnées toutes rationnelles.
Nous posons

$$\bar{\gamma} = \operatorname{Min}\left(\sum_{i=1}^{m} \bar{x}_i \operatorname{Log} \frac{a_i^k}{a_i^j} \mid j \in \{1,\dots,n\}, \quad j \neq k \right)$$

et considérons l'ensemble W :

$$W = \{x \in \mathbb{R}^m : \sum_{i=1}^{m} x_i \operatorname{Log} \frac{a_i^k}{a_i^j} > \frac{\bar{\gamma}}{2} \qquad \forall j \in \{1,\dots,n\},\ j \neq k\}.$$

W est un ouvert non vide de $\mathbb{R}^m$, il contient donc un point ξ de coordonnées toutes rationnelles.
Soit alors $x \in \mathbb{P}_m$, tel que :

$$\tilde{x}_i = \frac{\xi_i}{\sum_{i=1}^{m} \xi_i}$$

ce point répond à la question.

$\cancel{x}$

CHAPITRE - IV

ETUDE DE LA L-CONVERGENCE

INTRODUCTION

Etant donné un détecteur D, nous démontrons une condition suffisante pour qu'il existe un pseudoquestionnaire construit sur D qui soit L-convergent pour tout élément π de $\mathbb{P}_n$. Si cette condition n'est pas nécessaire, elle concerne toutefois une famille suffisamment large de détecteurs pour mettre en évidence l'intérêt de ce mode de convergence.

L'objet de ce chapitre est de démontrer le théorème suivant [6], [26]:

THEOREME IV-1

Etant donné un détecteur D, *une condition suffisante pour qu'il existe un élément* K *de* $\mathcal{K}$(D) *qui soit* L*-convergent pour tout élément* π *de* $\mathbb{P}_n$, *est que* D *vérifie l'hypothèse* H_L *suivante :*

H_L | *pour tout couple* (j,j') , $j \in \{1,\ldots,n\}$, $j' \in \{1,\ldots,n\}$, $j \neq j'$, D *admet un point essentiellement atteint* s *tel que :*

$$s^j \neq s^{j'}.$$

(où s^j désigne la $j^{\text{ème}}$ colonne de la matrice s).

Remarque

Dans le cas où n est égal à 2, et où D ne comporte que des questions dichotomiques, nous avons vu (III-3-3) que D admet une représentation dans le carré $[0,1] \times [0,1]$. L'hypothèse H_L s'interprète alors comme suit :

D admet un point essentiellement atteint qui n'est pas sur la diagonale du carré.

Notations

* Dans toute la suite, le détecteur D vérifiera l'hypothèse H_L, et nous désignerons par S un sous-ensemble de $\mathcal{S}(D)$ tel que :

$$\begin{cases} - \forall (j,j'),\ j \in \{1,\dots,n\},\ j' \in \{1,\dots,n\},\ j \neq j',\ \exists s \in S \text{ tel que } s^j \neq s^{j'}. \\ - \forall \sigma \in S, \text{ si } S-\{\sigma\} \text{ n'estpas vide, alors il existe un couple } (j,j') \text{ tel que :} \\ \qquad s^j = s^{j'} \qquad \forall s \in S-\{\sigma\}. \end{cases}$$

Nous noterons :

$$S = \{s(0),\dots,s(m-1)\} \qquad (m \leq \frac{n(n-1)}{2}).$$

* Par ailleurs, nous considèrerons un pseudoquestionnaire généralisé $\chi_o = (A_o, v_o, B_o) \in \mathcal{K}(S)$, tel que :

pour tout sommet y non terminal de χ_o, de rang r(y), on a :

$$v_o(y) = s(k)$$

où k est le reste de la division de r(y) par m, nombre d'éléments de S.

Etant donné un élément π de $\mathbb{P}_n$, la probabilité attachée à un élément ξ de χ_o opérant sur π sera notée :

$$p_o(\xi).$$

* La démonstration du théorème IV-1 comprend 3 parties ; nous avons ainsi dégagé 2 propositions intermédiaires, la dernière partie étant consacrée à la démonstration proprement dite du théorème IV-1.

I - PROPOSITION IV-1

Etant donné $\varepsilon > 0$, *il existe un entier* $R_o(\varepsilon)$, *et un réel* $\delta_o(\varepsilon) > 0$ *tels que pour tout* $\pi \in \mathbb{P}_n$, χ_o *possède un sommet* y *de rang au plus égal à* $R_o(\varepsilon)$, *interprétable à* ε *près pour* π, *et de plus :*

$$p_o(y) \geq \delta_o(\varepsilon).$$

Démonstration

Nous ferons cette démonstration par récurrence sur le cardinal de l'ensemble S.

1) Cas où $|S| = 1$

Nous noterons s l'élément de S, qui vérifie, par définition de S :

$$\forall j \in \{1,\ldots,n\}, \quad \forall j' \in \{1,\ldots,n\},\ j \neq j' : s^j \neq s^{j'}.$$

Etant donné $\pi \in \mathbb{P}_n$, nous posons :

$$\pi_{j_1} = \mathrm{Max}\left[\pi_j \mid j \in \{1,\ldots,n\}\right]$$

$$I(j_1) = \{i \in \{1,\ldots,a(s)\} \ : s_i^{j_1} \neq 0\}.$$

Soit alors x_1 un sommet de χ_o de rang égal à $|I(j_1)|$, tel que pour tout $i \in I(j_1)$, le chemin $[x_o,x_1]$ (x_o désignant la racine de χ_o) comporte un arc et un seul correspondant à la $i^{\text{ème}}$ ligne de s.

En notant $\Lambda(x_1) \in \mathbb{R}^n$ la caractéristique de x_1 dans χ_o, on a :

$$\begin{cases} p_o(x_1) = \sum_{j=1}^{n} \pi_j \Lambda_j(x_1) \\ p_o(T_{j_1}|x_1) = \dfrac{\pi_{j_1} \Lambda_{j_1}(x_1)}{p_o(x_1)} . \end{cases}$$

En raison du choix de j_1, il vient :

$$(1) \qquad \begin{cases} p_o(x_1) \geq \frac{1}{n} \Lambda_{j_1}(x_1) \\ p_o(T_{j_1}|x_1) \geq \frac{1}{n} \Lambda_{j_1}(x_1). \end{cases}$$

Montrons que x_1 possède un descendant y dans χ_o, vérifiant les conditions de la proposition.

En notant :

$$J = \{j \in \{1,\ldots,n\} : p_o(T_j|x_1) \neq 0\}$$

nous avons :

$$\begin{cases} \sum_{i \in I(j_1)} s_i^{j_1} = 1 \quad , \quad \sum_{i \in I(j_1)} s_i^j \leq 1 \qquad \forall j \in J \\ \forall j \in J-\{j_1\}, \exists i \in I(j_1) : s_i^j \neq s_i^{j_1}. \end{cases}$$

Par conséquent, la matrice $\sigma = ((s_i^j))$, $i \in I(j_1)$, $j \in J$, remplit les hypothèses de la proposition 1 de l'Annexe du chapitre III, et par suite, il existe

$$\begin{cases} \text{- des entiers } r_i,\ i \in I(j_1), \text{ positifs non tous nuls} \\ - \eta_{j_1} > 0 \end{cases}$$

tels qu'en posant

$$\rho = \sum_{i \in I(j_1)} r_i$$

on ait :

$$(2) \qquad \prod_{i \in I(j_1)} \left(\frac{s_i^{j_1}}{s_i^{j}} \right)^{r_i/\rho} \geq 1+\eta_{j_1} \qquad \forall j \in J-\{j_1\}.$$

Considérons alors, λ étant un entier positif, un descendant y_λ de x_1 dans χ_o défini par :

- y_λ est de rang $r(x_1)+\lambda\rho$
- Sur le chemin $[x_1,y_\lambda]$, pour tout $i \in I(j_1)$, il existe exactement λr_i arcs correspondant à la ième ligne de s.

On a :

$$\frac{p_o(T_{j_1}|y_\lambda)}{p_o(T_j|y_\lambda)} = \frac{p_o(T_{j_1}|x_1)}{p_o(T_j|x_1)} \left[\prod_{i \in I(j_1)} \left(\frac{s_i^{j_1}}{s_i^{j}} \right)^{r_i} \right]^\lambda \qquad \forall j \in J-\{j_1\}.$$

En tenant compte de (1) et (2), il vient :

$$\frac{p_o(T_{j_1}|y_\lambda)}{p_o(T_j|y_\lambda)} \geq \frac{1}{n} \Lambda_{j_1}(x_1) \; (1+\eta_{j_1})^{\lambda\rho} \qquad \forall j \in J-\{j_1\}.$$

Par conséquent, $\varepsilon > 0$ et $\pi \in \mathbb{P}_n$ étant fixés, il existe un entier R_{j_1} de la forme $r(x_1)+\lambda\rho$, tel que χ_o possède un sommet y_λ de rang R_{j_1}, interprétable à ε près pour π.
Le rang R_{j_1} ainsi trouvé ne dépend de π que par l'intermédiaire de l'indice j_1 ; il suffit donc de choisir

$$R_o(\varepsilon) = \mathrm{Max}(R_{j_1} \mid j_1 \in \{1,\ldots,n\}).$$

Il nous reste à montrer qu'il existe $\delta_o(\varepsilon)$ (indépendant de π) tel que :

$$p_o(y) \geq \delta_o(\varepsilon).$$

Or, π étant fixé, on a :

$$(3) \qquad \frac{p_o(y_\lambda)}{p_o(x_1)} = \sum_{j \in J} p_o(T_j | x_1) \frac{\Lambda_j(y_\lambda)}{\Lambda_j(x_1)}$$

et pour tout élément j de J, on a :

$$\frac{\Lambda_j(y_\lambda)}{\Lambda_j(x_1)} \geq \left[\mathrm{Min}(s_i^j \mid i \in I(j_1),\ j \in J\right]^{R_0(\varepsilon)} > 0 .$$

Posons :

$$\delta_{j_1} = \left(\frac{1}{n} \Lambda_{j_1}(x_1)\right) \left[\mathrm{Min}(s_i^j \mid i \in I(j_1),\ j \in J)\right]^{R_0(\varepsilon)} > 0.$$

D'après (1) et (3), on a :

$$p_o(y_\lambda) \geq \delta_{j_1} > 0.$$

Là encore, il suffira de choisir $\delta_o(\varepsilon)$ tel que :

$$\delta_o(\varepsilon) = \mathrm{Min}(\delta_{j_1} \mid j_1 \in \{1,\ldots,n\}).$$

2) Cas où le cardinal de S est quelconque

Nous allons nous ramener au cas précédent en montrant qu'il existe une opération (notée $\star$) dans l'ensemble des matrices stochastiques ayant n colonnes, telle que la matrice

$$\sigma = s(0) \star s(1) \star \ldots \star s(m-1)$$

vérifie :

$$\forall j \in \{1,\ldots,n\},\ \forall j' \in \{1,\ldots,n\},\ j \neq j' : \sigma^j \neq \sigma^{j'}.$$

a) Produit $\star$ de matrices

DEFINITION IV-1

Etant donné deux matrices A *et* B, *ayant même nombre* n *de colonnes, et respectivement* a_1 *et* a_2 *lignes, nous définissons le produit* $\star$ C *de* A *par* B

$$C = A \star B$$

de la façon suivante :
C *est une matrice ayant* n *colonnes et* $(a_1 a_2)$ *lignes, telle que :*

$$C_i^j = A_{i_1}^j \; B_{i_2}^j$$

i *décrivant l'ensemble* $\{1,\dots,a_1 a_2\}$ *lorsque le couple* (i_1,i_2) *décrit l'ensemble* $\{1,\dots,a_1\} \times \{1,\dots,a_2\}$.

Ce produit $\star$ est à rapprocher du produit Kroneckerien [28], la matrice C ainsi définie étant une sous-matrice d'ordre n du produit Kroneckerien de A par B.
Il est clair que cette opération est associative ; par ailleurs si A et B sont des matrices stochastiques, alors C l'est aussi.

PROPRIETE IV-1

La matrice $\sigma = s(0) \star s(1) \star \dots \star s(m-1)$ *vérifie :*

$$\forall j \in \{1,\dots,n\},\ \forall j' \in \{1,\dots,n\},\ j \neq j : \sigma^j \neq \sigma^{j'}.$$

Démonstration

Par définition de l'ensemble S, nous avons :

$\forall j \in \{1,\ldots,n\}$, $\forall j' \in \{1,\ldots,n\}$, $j \neq j'$, $\exists k \in \{0,\ldots,m-1\}$: $s^j(k) \neq s^{j'}(k)$.

Posons :

$$\tilde{\sigma} = s(0) * s(1)$$

Il nous suffit de montrer que pour tout couple (j,j'),

- soit il existe $k \in \{2,\ldots,m-1\}$ tel que : $s^j(k) \neq s^{j'}(k)$
- soit $\tilde{\sigma}^j \neq \tilde{\sigma}^{j'}$.

Considérons un couple (j,j') tel que :

$$\forall k \in \{2,\ldots,m-1\} : s^j(k) = s^{j'}(k).$$

Sans restriction de généralité, nous avons :

$$\begin{cases} \exists i_o \in \{1,\ldots,a(s(0))\} & : s^j_{i_o}(0) \neq s^{j'}_{i_o}(0) \\ s^{j'}_{i_o}(0) \neq 0. & \end{cases}$$

Par conséquent, si l'on avait $\tilde{\sigma}^j = \tilde{\sigma}^{j'}$, on aurait en particulier :

(4) $$s^j_{i_o}(0)\ s^j_i(1) = s^{j'}_{i_o}(0)\ s^{j'}_i(1) \quad \forall i \in \{1,\ldots,a(s(1))\}.$$

En notant :

$$I(j) = \{i \in \{1,\ldots,a(s(1))\} \ : s^j_i(1) = 0\}$$

la relation (4) entraîne :

$$I(j) \subset I(j')$$

soit :

$$\sum_{i \notin I(j)} s_i^j(1) = 1 = \sum_{i \notin I(j)} s_i^{j'}(1).$$

Par ailleurs, (4) entraîne :

$$\frac{s_i^{j'}(1)}{s_i^j(1)} = \frac{s_{i_o}^j(0)}{s_{i_o}^{j'}(0)} \qquad \forall i \notin I(j)$$

et il en résulterait :

$$\frac{s_{i_o}^j(0)}{s_{i_o}^{j'}(0)} = \frac{\sum_{i \notin I(j)} s_i^{j'}(1)}{\sum_{i \notin I(j)} s_i^j(1)} = 1$$

d'où la contradiction.

~~*~~

b) <u>Conséquence</u>

En conservant la notation, $\sigma = s(0) * s(1) * \ldots * s(m-1)$, nous considérons un pseudoquestionnaire généralisé

$$\bar{\chi} = (\bar{A}, \bar{v}, \bar{B}) \in \mathcal{H}(\{\sigma\})$$

tel que :

pour tout sommet y non terminal de $\bar{\chi}$, on ait :

$$\bar{v}(y) = \sigma.$$

D'après ce qui a été démontré en IV-1-1, et la propriété IV-1, nous savons que :

étant donné $\varepsilon > 0$, il existe $\bar{R}(\varepsilon)$ et $\bar{\delta}(\varepsilon) > 0$, tels que pour tout $\pi \in \mathbb{P}_n$, $\bar{\chi}$ possède un sommet de rang au plus égal à $\bar{R}(\varepsilon)$, interprétable à ε près pour π, et de plus :

$$\bar{p}(y) \geq \bar{\delta}(\varepsilon).$$

Or, il est clair qu'il existe une bijection φ entre les sommets y de χ_o de rang λm (λ entier positif) et les sommets de rang λ dans $\bar{\chi}$, telle que pour un élément $\pi \in \mathbb{P}_n$, on ait :

$$\begin{cases} P_o(T|y) = \bar{P}(T|\varphi(y)) \\ p_o(y) = \bar{p}(\varphi(y)) \end{cases}$$

(Etant donné un élément ξ de $\bar{\chi}$ opérant sur π, nous notons $\bar{p}(\xi)$ sa probabilité).

Ceci achève donc la démonstration de la proposition IV-1, et nous remarquons que le rang $R_o(\varepsilon)$ ainsi obtenu est un multiple de m.

$*$

II - EXTENSION PAR CONTINUITE DES RESULTATS DE LA PROPOSITION IV-1

Notations

* Etant donné $\eta > 0$, et un élément s de $S(a,n)$, nous noterons $B(s,\eta)$ la boule de centre s et de rayon η dans $S(a,n)$.

* Nous désignerons par $g(\eta)$ la partie de $\mathcal{H}(S(a,n))$ définie comme suit :

  $\chi = (A,v,B)$ est un élément de $g(\eta)$, si et seulement si, pour tout sommet y non terminal de χ, on a :

$$v(y) \in B(s(k),\eta)$$

  où k est le reste de la division de r(y) par m.

PROPOSITION IV-2

Etant donné $\varepsilon > 0$, *il existe un entier* $R(\varepsilon)$, *et des réels positifs* $\eta(\varepsilon)$, $\delta(\varepsilon)$, *tels que pour tout* $\pi \in \mathbb{P}_n$, *tout élément* χ *de* $g(\eta(\varepsilon))$ *possède un sommet* y *de rang au plus égal à* $R(\varepsilon)$, *interprétable à* ε *près pour* π, *et de plus :*

$$p(y) \geq \delta(\varepsilon).$$

Démonstration

Dans toute la suite, ε est fixé.

D'après la proposition IV-1, pour tout élément π de $\mathbb{P}_n$, le pseudoquestionnaire généralisé $\chi_o = (A_o, v_o, B_o)$ possède un sommet y_π de rang au plus égal à $R_o(\frac{\varepsilon}{2})$, interprétable à $\frac{\varepsilon}{2}$ près, et de plus :

$$p_o(y_\pi) \geq \delta_o(\tfrac{\varepsilon}{2}) > 0.$$

Avec les notations de I-3, nous avons :

$$(5) \qquad p_o(y_\pi) \geq f(\pi, \Lambda_o(y_\pi))$$

où $\Lambda_o(y_\pi)$ désigne la caractéristique de y_π dans χ_o.

Etant donné alors $\delta(\varepsilon) \in \left]0, \delta_o(\frac{\varepsilon}{2})\right[$, en vertu de l'uniforme continuité de f sur $\mathbb{P}_n \times [0,1]^n$, nous allons pouvoir choisir v de telle sorte que dans le pseudoquestionnaire $\chi = (A_o, v, B_o)$, on ait :

$$(6) \qquad p(y_\pi) = f(\pi, \Lambda(y_\pi)) \geq \delta(\varepsilon)$$

où $\Lambda(y_\pi)$ désigne la caractéristique de y_π dans χ.

En effet, il existe η_1 tel que pour tout $\pi \in \mathbb{P}_n$, on ait :

$$\|\Lambda_o(y_\pi) - \Lambda(y_\pi)\| \leq \eta_1 \Rightarrow |f(\pi, \Lambda(y_\pi)) - f(\pi, \Lambda_o(y_\pi))| \leq \delta_o(\tfrac{\varepsilon}{2}) - \delta(\varepsilon).$$

En raison de (5), il suffit donc pour avoir (6) de choisir v de telle façon que :

$$\|\Lambda_o(y_\pi) - \Lambda(y_\pi)\| \leq \eta_1.$$

En désignant alors par Δ le compact défini dans $\mathbb{P}_n \times [0,1]^n$ par :

$$\Delta = \{(z,t) \in \mathbb{P}_n \times [0,1]^n : f(z,t) \geq \delta(\varepsilon)\}$$

nous avons ainsi montré qu'il existe $\eta_1 > 0$ tel que pour tout élément π de $\mathbb{P}_n$, on ait :

$$\{(\pi,t) \ : \ t \in B(\Lambda_o(y_\pi),\eta_1)\} \subset \Delta \ .$$

Par ailleurs, en utilisant toujours les notations de I-3, on a :

$$p_o(T_j|y_\pi) = g_j(\pi,\Lambda_o(\pi)).$$

En vertu de l'uniforme continuité des fonctions g_j sur Δ, nous allons pouvoir choisir v de telle sorte que :

(7) y_π soit interprétable à ε près pour π dans χ.

En effet, il existe $\eta_2 \in \,]0,\eta_1[$, tel que pour tout $\pi \in \mathbb{P}_n$, on ait :

$$\|\Lambda_o(y_\pi)-\Lambda(y_\pi)\| < \eta_2 \Rightarrow |g_j(\pi,\Lambda_o(y_\pi))-g_j(\pi,\Lambda(y_\pi))| < \frac{\varepsilon}{2}, \quad \forall j \in \{1,\dots,n\}.$$

Pour avoir (7), il suffit de remarquer que y_π est interprétable à $\frac{\varepsilon}{2}$ près pour π dans χ_o.

Pour conclure, il suffit donc de prendre

$$R(\varepsilon) = R_o(\frac{\varepsilon}{2})$$

et de remarquer, qu'en notant $Y(R(\varepsilon))$ l'ensemble des sommets non terminaux et de rang inférieur à $R(\varepsilon)$ dans χ_o, il existe trivialement $\eta(\varepsilon) > 0$ tel que :

$$\left\{\begin{array}{l} \|v(y)-v_o(y)\| \leq \eta(\varepsilon) \\ \forall y \in Y(R(\varepsilon)) \end{array}\right. \Rightarrow \left\{\begin{array}{l} \|\Lambda(y)-\Lambda_o(y)\| \leq \mathrm{Min}(\eta_1,\eta_2) \\ \forall y \in Y(R(\varepsilon)) \end{array}\right.$$

COROLLAIRE

Etant donné $\varepsilon > 0$, *il existe* $\eta(\varepsilon) > 0$ *et* $M(\varepsilon)$ *fini, tels que tout élément* χ *de* $g(\eta(\varepsilon))$ *vérifie :*

$$L[\chi(\pi,\varepsilon)] \leq M(\varepsilon) \qquad \forall \pi \in \mathbb{P}_n.$$

Démonstration

Etant donné $\varepsilon > 0$, nous savons d'après la proposition IV-2 qu'il existe $R(\varepsilon)$, $\delta(\varepsilon)$ et $\eta(\varepsilon)$ tels que, étant donné un élément χ de $g(\eta(\varepsilon))$ et $\pi \in \mathbb{P}_n$, tout sous-pseudoquestionnaire de χ possède un sommet de rang (relatif à ce sous-pseudoquestionnaire) inférieur ou égal à $R(\varepsilon)$, interprétable à ε près pour π, et de probabilité (relative) supérieure à $\delta(\varepsilon)$.
Montrons que la suite $L[\chi(R\ ;\ \pi,\varepsilon)]$ converge lorsque R tend vers l'infini.
Cette suite étant croissante, il suffit de montrer que la suite $L[\chi(\lambda R(\varepsilon)\ ;\ \pi,\varepsilon)]$ converge lorsque l'entier λ tend vers l'infini.
Or, nous avons :

$$L[\chi(\lambda R(\varepsilon)\ ;\ \pi,\varepsilon)] \leq L[\chi((\lambda-1)R(\varepsilon)\ ;\ \pi,\varepsilon)]+R(\varepsilon)\ p(E((\lambda-1)R(\varepsilon)\ ;\ \pi,\varepsilon))$$

(rappelons que $E(R\ ;\ \pi,\varepsilon)$ désigne l'ensemble des sommets de rang R d'un pseudoquestionnaire, non interprétables à ε près pour π)
et :

$$p(E((\lambda-1)R(\varepsilon)\ ;\ \pi,\varepsilon)) \leq (1-\delta(\varepsilon))^{(\lambda-1)}.$$

Ces deux inégalités entraînent :

$$L[\chi(\lambda R(\varepsilon)\ ;\ \pi,\varepsilon)] \leq R(\varepsilon) \sum_{k=0}^{\lambda} (1-\delta(\varepsilon))^k$$

Donc :

$$L[\chi(\lambda R(\varepsilon)\ ;\ \pi,\varepsilon)] \leq \frac{R(\varepsilon)}{\delta(\varepsilon)} = M(\varepsilon).$$

$\ast$

III - DEMONSTRATION DU THEOREME IV-1

Considérons une suite $\{\varepsilon_\ell\}$, $\ell \in \mathbb{N}$, de nombres positifs décroissants vers zéro.

Les propositions IV-1 et IV-2 (et son corollaire) associent à chaque ε_ℓ un entier $R(\varepsilon_\ell)$, que nous noterons R_ℓ, un réel $\eta(\varepsilon_\ell)$ que nous noterons η_ℓ, et un nombre $M(\varepsilon_\ell)$.

Considérons la partie g de $\mathcal{K}(S(a,n))$ définie comme suit :

$\chi = (A,v,B)$ est un élément de g, si et seulement si pour tout sommet y non terminal de χ de rang $r(y)$ tel que

$$\sum_{k=1}^{\ell-1} R_k \leq r(y) < \sum_{k=1}^{\ell} R_k$$

on a :

$$v(y) \in B(s(k),\eta_\ell)$$

où k est le reste de la division de $r(y)$ par m.

Montrons qu'un élément χ quelconque de g L-converge pour tout élément π de $\mathbb{P}_n$.

Il suffit de montrer que pour tout élément χ de g, pour tout élément π de $\mathbb{P}_n$, et pour tout $\ell \in \mathbb{N}$, la quantité

$$L[\chi(\pi,\varepsilon_\ell)]$$

est bornée.

Or, tout sous pseudoquestionnaire de χ de racine x telle que

$$r(x) = \sum_{k=1}^{\ell-1} R_k$$

est un élément de $g(\eta_\ell)$; par suite, d'après le corollaire de la proposition IV-2,

on a :

$$L[\chi(\pi,\varepsilon_\ell)] \leq \sum_{k=1}^{\ell-1} R_k + M(\varepsilon_\ell).$$

Pour conclure, il suffit de remarquer, qu'en raison de la définition des points essentiellement atteints, il existe $K = (A,u,B) \in \mathcal{K}(D)$, tel que le pseudoquestionnaire généralisé

$$\chi = (A,\ \alpha \circ u,\ B)$$

appartienne à $\mathcal{g}$.

Par conséquent un tel pseudoquestionnaire K est L-convergent pour tout $\pi \in \mathbb{P}_n$.

IV - UN EXEMPLE D'APPLICATION DE LA L-CONVERGENCE

Considérons le problème suivant, généralisation d'un problème soulevé par Picard [17].

Un voyageur arrive dans une ville comportant trois types d'habitants :

- les habitants du type T_1 qui disent toujours la vérité
- les habitants du type T_2 qui disent le contraire de la vérité
- les habitants du type T_3 qui disent aléatoirement la vérité ou son contraire de façon équiprobable.

Rencontrant un habitant de cette ville, le problème est pour le voyageur, d'identifier le type de cet habitant, en l'interrogeant au moyen de questions dichotomiques.

Quelle que soit la question q, utilisée par le voyageur, la matrice stochastique $\alpha(q)$ de q sera telle que :

$$\alpha(q) = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{pmatrix}$$

Par conséquent, le voyageur, n'étant pas limité dans le nombre de ces questions, il dispose d'un détecteur infini qui admet un point essentiellement atteint s tel que :

$$s = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \end{pmatrix}$$

En raison du théorème IV-1, il lui suffira donc, quelle que soit la répartition en probabilité des types T_1, T_2, T_3 sur la population de la ville, d'un nombre fini de questions en moyenne, pour identifier le type de son interlocuteur.

CHAPITRE - V

PSEUDOQUESTIONNAIRES ET INFORMATION

INTRODUCTION

Après un rappel sur la fonction d'information de Shannon [22], nous donnons quelques résultats précisant des propriétés classiques. Dans un troisième paragraphe, nous introduisons la notion d'information transmise par un pseudoquestionnaire, qui généralise l'information transmise par un questionnaire [19].

I - RAPPELS SUR LA FONCTION D'INFORMATION DE SHANNON [13]

1) Définitions

Etant donné un ensemble Ω, une tribu $\mathcal{a}$ d'évènements dans Ω, et une mesure p sur $(\Omega,\mathcal{a})$, on considère un système complet d'évènements X sur $(\Omega,\mathcal{a})$, et on note $(X_1,\ldots,X_m)$ la partition X de Ω.

On appelle entropie du système X *pour la mesure* p, *la quantité* $H_p(X)$ *définie par :*

$$H_p(X) = \sum_{j=1}^{n} p(X_j) \text{ Log } \frac{1}{p(X_j)} .$$

Cette quantité ne dépend de X et de p que par l'intermédiaire de l'élément π de $\mathbb{P}_n$ de composantes

$$\pi_j = p(X_j).$$

Aussi cette quantité est-elle fréquemment notée

$$H(\pi) \quad , \quad \text{ou } H(\pi_1,\dots,\pi_n).$$

Considérons maintenant un autre système complet d'évènements $Y = (Y_1,\dots,Y_m)$ sur $(\Omega,\mathcal{a})$.

Etant donné un élément A de $\mathcal{a}$, on notera p_A la mesure définie sur $(\Omega,\mathcal{a})$ par :

$$p_A(B) = p(B|A) = \frac{p(A \cap B)}{p(A)} \quad \forall B \in \mathcal{a}.$$

On appelle entropie de X *pour la mesure* p, *conditionnée par* Y *la quantité* $H_p(X|Y)$ *définie par :*

$$H_p(X|Y) = \sum_{i=1}^{m} p(Y_i)\, H_{p_{Y_i}}(X).$$

On appelle information apportée par Y *sur* X *pour la mesure* p *la quantité* $I_p(X,Y)$ *définie par :*

$$I_p(X,Y) = H_p(X) - H_p(X|Y).$$

2) Propriétés fondamentales

a) Propriété d'associativité de l'entropie :

Soit J un sous-ensemble de $\{1,\dots,n\}$, on pose

$$\tilde{X} = \bigcup_{j \in J} X_j$$

et on considère la partition Z de Ω définie par :

$$Z = (\tilde{X}, X_j \; j \notin J)$$

on a alors :

$$H_p(X) = H_p(Z) + p(\tilde{X}) H_{p_{\tilde{X}}}(X).$$

b) Propriété de symétrie de l'information apportée

$$I_p(X,Y) = I_p(Y,X).$$

c) Propriété de positivité de l'information apportée :

$$I_p(X,Y) \geq 0$$

avec égalité, si et seulement si les systèmes X et Y sont stochastiquement indépendants pour la mesure p.

II - PROPRIETES D'ASSOCIATIVITE DE L'INFORMATION APPORTEE PAR Y SUR X

Nous généralisons ici la propriété d'associativité de l'entropie mise en évidence par Fadeev [8]. Avec les notations précédentes, nous avons :

PROPRIETE V-1

$$I_p(X,Y) = I_p(Z,Y) + p(\tilde{X})\, I_{p_{\tilde{X}}}(X,Y)$$

Démonstration

Nous considérons la mesure $p_{\tilde{X}_i}$, définie sur $(\Omega,\mathcal{a})$ comme suit :

$$p_{\tilde{X}_i}(A) = p(A|\tilde{X} \cap Y_i) \qquad \forall A \in \mathcal{a}.$$

D'après la propriété d'associativité de l'entropie, nous avons :

$$\begin{cases} H_p(X) = H_p(Z) + p(\tilde{X})\, H_{p_{\tilde{X}}}(X) \\ H_{p_{Y_i}}(X) = H_{p_{Y_i}}(Z) + p_{Y_i}(\tilde{X})\, H_{p_{\tilde{X}_i}}(X) \end{cases}$$

d'où :

$$I_p(X,Y) = I_p(Z,Y) + p(\tilde{X}) \left[H_{p_{\tilde{X}}}(X) - \sum_{i=1}^{m} p_{\tilde{X}}(Y_i)\, H_{p_{\tilde{X}_i}}(X)\right].$$

~~X~~

PROPRIETE V-2

On a :

$$I_p(X,Y) = \sum_{k=1}^{m-1} p(V_k)\, I_{p_{V_k}}(X,Z_k)$$

où :

$$\begin{cases} - V_k = \bigcup_{i=k}^{m} Y_i & \forall k \in \{1,\dots,m\} \\ - Z_k \ \textit{désigne la partition } (Y_1,\dots,Y_k,V_{k+1}) \textit{ de } \Omega \textit{ pour } k \in \{1,\dots,m-1\}. \end{cases}$$

Démonstration

En raison de l'inclusion des V_k, nous avons :

$$p_{V_k}(A) = p(A \mid \bigcap_{\ell=1}^{k} V_\ell) \qquad \forall A \in \mathcal{a}.$$

En appliquant alors itérativement la propriété V-1, nous obtenons les relations :

$$I_p(X,Y) = I_p(X,Z_1)+p(V_2)I_{p_{V_2}}(X,Z_1)$$

$$I_{p_{V_2}}(X,Z_1) = I_{p_{V_2}}(X,Z_2)+p(V_3)I_{p_{V_3}}(X,Z_2)$$

$$I_{p_{V_k}}(X,Z_{k-1}) = I_{p_{V_k}}(X,Z_k)+p(V_{k+1})I_{p_{V_{k+1}}}(X,Z_k)$$

$$I_{p_{V_{m-1}}}(X,Z_{m-2}) = I_{p_{V_{m-1}}}(X,Z_{m-1})+p(V_m)I_{p_{V_m}}(X,Z_{m-1})$$

En remarquant que $I_{p_{V_m}}(X,Z_{m-1})$ est nul, et en combinant ces égalités, on obtient le résultat.

X

III - INFORMATION TRANSMISE PAR UN PSEUDOQUESTIONNAIRE

1) Information transmise en un sommet non terminal d'un pseudoquestionnaire opérant sur π.

Soient $K = (\mathcal{A},u,B)$ un pseudoquestionnaire construit sur un détecteur D, et y un sommet non terminal de K.

Etant donné un élément π de $\mathbb{P}_n$, le vecteur de probabilité $P(T|y)$, qu'on peut alors associer à y, engendre sur l'espace fondamental une mesure que nous noterons p_y (cf. I-2).

DEFINITION V-1

Nous appellerons information transmise en un sommet y *non terminal d'un pseudoquestionnaire* K *opérant sur* π, *la quantité :*

$$I(y) = I_{p_y}(u(y),T).$$

PROPRIETE V-3

On a la relation :

$$I(y) = H(P(T|y)) - \sum_{z \in \Gamma_y} \frac{p(z)}{p(y)} H(P(T|z)).$$

Démonstration

D'après la définition V-1, on a :

$$I(y) = H_{p_y}(T) - H_{p_y}(T|u(y)).$$

Soit z_i le successeur de y dans K, tel qu'en notant

$$u(y) = q$$

on ait :

$$b_y((y,z_i)) = q_i \qquad \forall i \in \{1,\dots,a(q)\}$$

On a :

$$I(y) = H(\mathcal{P}(T|y)) - \sum_{i=1}^{a(q)} p_y(q_i)\ H(\mathcal{P}(T|z_i))$$

et :

$$p_y(q_i) = \frac{p(z_i)}{p(y)}\ .$$

~~X~~

2) Information transmise par un pseudoquestionnaire

a) Cas d'un pseudoquestionnaire de hauteur finie.

Considérons un pseudoquestionnaire $K = (\mathcal{A},u,B)$ de hauteur finie, opérant sur $\pi \in \mathbb{P}_n$.

On désignera respectivement par Y et Z l'ensemble de ses sommets non terminaux et terminaux.

DEFINITION V-2

Nous appellerons information transmise par un pseudoquestionnaire K *de hauteur finie, opérant sur* $\pi \in \mathbb{P}_n$, *la quantité :*

$$I(K,\pi) = \sum_{y\in Y} p(y)\ I(y).$$

PROPRIETE V-4

On a la relation :

$$I(K,\pi) = H(\pi) - \sum_{x\in Z} p(x)\ H(\mathcal{P}(T|x)).$$

Démonstration

D'après la propriété V-3, il vient :

$$I(K,\pi) = \sum_{y\in Y} p(y)\Big[H(\mathcal{P}(T|y)) - \sum_{z\in\Gamma_y} \frac{p(z)}{p(y)}\ H(\mathcal{P}(T|z))\Big]$$

soit :

$$I(K,\pi) = \sum_{y\in Y} p(y)\ H(\mathcal{P}(T|y)) - \sum_{y\in Y} \sum_{z\in\Gamma_y} p(z)\ H(\mathcal{P}(T|z))$$

d'où :

$$I(K,\pi) = \sum_{y\in Y} p(y)\ H(\mathcal{P}(T|y)) - \sum_{y\in Y-\{x_o\}\cup Z} p(y)\ H(\mathcal{P}(T|y))$$

où x_o désigne la racine de K.

~~X~~

Remarque

La propriété précédente implique que

$$I(K,\pi) \leq H(\pi)$$

avec égalité si et seulement si tous les sommets terminaux de K sont interprétables pour π, c'est-à-dire si K est trivial pour π.

b) Cas d'un pseudoquestionnaire de hauteur infinie

Etant donné un pseudoquestionnaire K de hauteur infinie, nous considérons l'ensemble Y(R) des sommets non terminaux de la restriction K(R) de hauteur R de K.

Etant donné $\pi \in \mathbb{P}_n$, en raison de la propriété V-4, on a pour tout R :

$$I(K(R),\pi) = \sum_{y\in Y(R)} p(y)\ I(y) \leq H(\pi).$$

Par conséquent, la suite $I(K(R),\pi)$ qui est croissante et majorée converge, et Y désignant l'ensemble des sommets non terminaux de K, la famille à termes positifs

$$(p(Y)\ I(y))\ ,\ y \in Y$$

est une famille sommable ; de plus :

$$\sum_{y \in Y} p(y)\ I(y) \leq H(\pi).$$

Ceci nous permet de généraliser la définition V-2.

DEFINITION V-3

Nous appellerons information transmise par un pseudoquestionnaire K *opérant sur* π*, la quantité :*

$$I(K,\pi) = \sum_{y \in Y} p(y)\ I(y).$$

Par ailleurs, nous avons démontré la propriété :

PROPRIETE V-5

Etant donné un pseudoquestionnaire K *opérant sur* π*, on a :*

$$I(K,\pi) \leq H(\pi).$$

CHAPITRE - VI

ETUDE INFORMATIONNELLE DE LA L-CONVERGENCE

INTRODUCTION

Ce chapitre n'apporte pas de résultats sensiblement différents de ceux obtenus au chapitre IV. Toutefois la méthode de démonstration du théorème IV-1 présentée au paragraphe 2, et utilisant la théorie de l'information, nous a paru intéressante par sa rapidité.

Par ailleurs, nous avons été conduits (§ 3) à considérer une famille particulière de pseudoquestionnaires susceptible de présenter un intérêt algorithmique.

I - PROPRIETE PRELIMINAIRE

La proposition VI-1, qui suit, donne quelques précisions sur la propriété de positivité de l'information apportée.

Nous utilisons les mêmes notations qu'en V-1.

PROPOSITION VI-1

Soient deux systèmes complets d'évènements X *et* Y *sur* (Ω, a, p) , $X = (X_1,\ldots,X_n)$, $Y = (Y_1,\ldots,Y_m)$, *et deux éléments* ξ *et* δ *de* $]0,1[$, *tels que deux éléments* j_1 *et* j_2 *de* $\{1,\ldots,n\}$ *vérifient :*

$$\begin{cases} \mathrm{Max}\left[|p(X_{j_1}|Y_i)-p(X_{j_2}|Y_i)| \mid i \in \{1,\ldots,m\}\right] \geq \delta \\ p(X_j) \geq \xi \qquad \forall j \in \{j_1,j_2\} . \end{cases}$$

Alors il existe un nombre $\gamma(\delta,\xi)$ *strictement positif, tel que :*

$$I_p(X,Y) \geq \gamma(\delta,\xi).$$

Démonstration

Nous poserons, sans restriction de généralité :

$$\begin{cases} j_1 = 1 \quad , \quad j_2 = 2 \\ |p(X_1|Y_1)-p(X_2|Y_1)| \geq \delta \end{cases}$$

et noterons :

$$\tilde{X} = X_1 \cup X_2.$$

Considérons la partition Z de Ω définie par :

$$Z = (Y_1, \bigcup_{i=2}^{m} Y_i).$$

D'après les propriétés V-1 et V-2 de l'information apportée, on a :

$$(1) \qquad I_p(X,Y) \geq p(\tilde{X})\ I_{p_{\tilde{X}}}(X,Y) \geq p(\tilde{X})\ I_{p_{\tilde{X}}}(X,Z).$$

En raison des hypothèses, on a :

$$(2)\quad \begin{cases} p(\tilde{X}) \geq 2\,\xi \\ p_{\tilde{X}}(X_j) = \dfrac{p(X_j)}{p(\tilde{X})} \geq \xi \qquad \forall j \in \{1,2\}. \end{cases}$$

Par ailleurs, considérons la fonction $\varphi(w,t,\lambda)$ définie sur $[0,1]^3$ comme suit :

$$\varphi(w,t,\lambda) = H[\lambda w+(1-\lambda)t, 1-\lambda w-(1-\lambda)t] - \lambda H(w,1-w)-(1-\lambda)H(t,1-t)$$

où :

$$H(w,1-w) = w \text{ Log } \frac{1}{w} + (1-w) \text{ Log } \frac{1}{1-w}$$

$$H(0,1) = H(1,0) = 0.$$

On a :

$$I_{p_{\tilde{X}}}(X,Z) = \varphi(p(X_1|Y_1),\ p(X_2|Y_1),\ p_{\tilde{X}}(X_1)).$$

Cette fonction φ est continue, et n'est nulle que si :

$$\lambda = 0, \quad \text{ou} \quad w = t.$$

Considérons alors le compact $\Delta(\delta,\xi)$ défini dans $[0,1]^3$ par :

$$\Delta(\delta,\xi) = \{(w,t,\lambda) \in [0,1]^3 : |w-t| \geq \delta,\ \ \lambda \in [\xi,1-\xi]\}.$$

Il existe un nombre $\bar{\gamma}(\delta,\xi) > 0$, tel que :

$$\varphi(w,t,\lambda) \geq \bar{\gamma}(\delta,\xi) \qquad \forall (w,t,\lambda) \in \Delta(\delta,\xi).$$

Or nous avons :

$$(p(X_1|Y_1),\ p(X_2|Y_1),\ p_{\tilde{X}}(X_1)) \in \Delta(\delta,\xi)$$

donc :

$$I_{p_{\tilde{X}}}(X,Z) \geq \bar{\gamma}(\delta,\xi).$$

Pour conclure, il suffit alors, en raison de (1) et (2), de poser :

$$\gamma(\delta,\xi) = 2\ \xi\ \ \bar{\gamma}(\delta,\xi).$$

~~X~~

II - DEMONSTRATION INFORMATIONNELLE DU THEOREME IV-1

* Nous considérons un détecteur D vérifiant l'hypothèse H_L, et comme au chapitre IV, un sous-ensemble S de $\mathcal{S}(D)$,

$$S = \{s(0),\ldots,s(m-1)\}$$

tel que :

$$\begin{cases} - \forall j \in \{1,\ldots,n\},\ \forall j' \in \{1,\ldots,n\},\ j \neq j'\ ,\quad \exists s \in S : s^j \neq s^{j'}. \\ - \text{Pour tout élément } \sigma \text{ de S, si } S-\{\sigma\} \text{ n'est pas vide, alors il existe un} \\ \quad \text{couple } (j,j') \text{ tel que :} \\ \qquad s^j = s^{j'} \qquad \forall s \in S-\{\sigma\}. \end{cases}$$

* Etant donné $\eta > 0$, nous considérons le sous-ensemble $\mathcal{K}(D,\eta)$ de $\mathcal{K}(D)$ défini de la façon suivante :

K = $(\mathcal{A},u,B)$ est un élément de $\mathcal{K}(D,\eta)$ si et seulement si pour tout sommet non terminal y de K, on a :

$$u(y) \in B(s(k),\eta)$$

où k est le reste de la division du rang r(y) de y par m, et $B(s(k),\eta)$ est la boule de $\mathcal{S}(a,n)$ de centre s(k) et de rayon η.

Nous démontrons alors la proposition suivante :

PROPOSITION VI-2

Il existe $\eta_o > 0$, *tel que tout élément de* $\mathcal{H}(D,\eta_o)$ *soit* L-*convergent pour tout élément* π *de* $\mathbb{P}_n$.

Démonstration

Par définition de S, il existe $\delta > 0$, tel que, pour tout couple (j,j'), il existe $s(k) \in S$ vérifiant :

$$\|s^j(k)-s^{j'}(k)\| \geq \delta.$$

Donc il existe $\eta_o > 0$, tel que pour tout couple (j,j') il existe $k \in \{0,\ldots,m-1\}$ vérifiant :

$$\|s^j-s^{j'}\| \geq \delta \qquad \forall s \in B(s(k),\eta_o).$$

Soient alors $K = (A,u,B) \in \mathcal{H}(D,\eta_o)$, π un élément de $\mathbb{P}_n$, et $\varepsilon > 0$, montrons que :

$$L[K(\pi,\varepsilon)] < +\infty.$$

Pour ceci, nous considérons la restriction K' de K définie comme suit :

un sommet x de K' est terminal si et seulement si x est de rang multiple de m et si x ou l'un de ses ascendants est interprétable à ε près pour π.

Il est clair que K' est un prolongement de $K(\pi,\varepsilon)$; il nous suffit donc de montrer que :

$$L[K',\pi] < +\infty.$$

Considérons un sommet y de rang λm (λ entier positif) non terminal dans K' ; et posons :

$$\begin{cases} p(T_{j_1}|y) = \mathrm{Max}(p(T_j|y) \mid j \in \{1,\ldots,n\}) \\ p(T_{j_2}|y) = \mathrm{Max}(p(T_j|y) \mid j \in \{1,\ldots,n\}-\{j_1\}). \end{cases}$$

Le sommet y n'étant pas interprétable à ε près pour π, on a :

$$p(T_j|y) \geq \frac{\varepsilon}{n-1} \qquad \forall j \in \{j_1,j_2\} .$$

Soit alors $s(k) \in S$ tel que :

$$\|s^{j_1}(k)-s^{j_2}(k)\| > \delta.$$

Nous pouvons supposer, sans restreindre la généralité, que

$$u(y) \in B(s(k),\eta_o).$$

(En effet, c'est vrai à une commutativité près sur les questions du sous-pseudoquestionnaire de racine y, commutativité qui n'altère pas la convergence).

Par conséquent, en raison de la proposition VI-1, il existe $\gamma(\delta,\varepsilon) > 0$, tel que :

$$I(y) \geq \gamma(\delta,\varepsilon).$$

En notant E'(R) l'ensemble des sommets de K', non terminaux et de rang R, on a :

$$\gamma(\delta,\varepsilon) \sum_{\lambda=0}^{\infty} p(E'(\lambda m)) \leq \sum_{\lambda=0}^{\infty} \sum_{y\in E'(\lambda m)} p(y)\, I(y) \leq H(\pi)$$

soit :

$$\sum_{\lambda=0}^{\infty} p(E'(\lambda m)) \leq \frac{H(\pi)}{\gamma(\delta,\varepsilon)} .$$

La suite $p(E'(R))$ étant stationnaire pour R tel que :

$$\lambda m \leq R < (\lambda+1)m$$

on a :

$$\sum_{R=0}^{\infty} p(E'(R)) = L[K',\pi] < +\infty.$$

~~X~~

III - PSEUDOQUESTIONNAIRES PARTICULIERS [26]

Nous allons considérer ici une nouvelle famille de pseudoquestionnaires effectivement constructibles.

Notation

Etant donné un détecteur D, un élément π de $\mathbb{P}_n$ et un nombre $\rho \in]0,1[$, nous considérons l'ensemble

$$\mathcal{H}(D,\pi,\rho)$$

des pseudoquestionnaires construits sur D, définis comme suit :

$K^* = (A,u,B)$ est un élément de $\mathcal{H}(D,\pi,\rho)$ si et seulement si :

- $K^* \in \mathcal{K}(D,J)$, où $J = \{j \in \{1,\dots,n\} : \pi_j \neq 0\}$
  (cf. Remarque de I-5-3).

- soit y un sommet non interprétable pour π ; posons :

$$\begin{cases} D(y) = \{q \in D : u(x) \neq q \quad \forall x \in A(y)\} \\ \bar{m}(y) = \mathrm{Sup}\left[I_{p_y}(q,T) \mid q \in D(y)\right] \\ D'(y) = \{q \in D(y) : (1-\rho)\,\bar{m}(y) \leq I_{p_y}(q,T) \leq \bar{m}(y)\} \end{cases}$$

alors :

$$u(y) \in D'(y).$$

PROPOSITION VI-3

Etant donné un détecteur D *vérifiant l'hypothèse* H_L *, un élément* π *de* $\mathbb{P}_n$*, et* $\rho \in \left]0,1\right[$*, alors tout élément de* $\mathcal{K}(D,\pi,\rho)$ *est* L*-convergent pour* π.

Démonstration

Soit $\varepsilon > 0$ donné, et $K^* = (A,u,B) \in \mathcal{K}(D,\pi,\rho)$; montrons que :

$$L\left[K^*(\pi,\varepsilon)\right] < +\infty.$$

Remarquons tout d'abord que, D vérifiant l'hypothèse H_L, il existe $\delta > 0$, tel que pour tout couple (j,j'), et toute partie D' de D telle que $|D-D'|$ soit fini, il existe un élément q de D' vérifiant :

$$\|\alpha^j(q)-\alpha^{j'}(q)\| \geq \delta.$$

Par ailleurs, de même qu'au paragraphe précédent, si y est un sommet non terminal de K^*, il existe j_1 et j_2 vérifiant :

$$p(T_j|y) \geq \frac{\varepsilon}{n-1} \qquad \forall j \in \{j_1,j_2\}.$$

Donc, en raison de la proposition VI-1, il existe $\gamma(\delta,\varepsilon) > 0$ et un élément q de D(y), tels que :

$$I_{p_y}(q,T) \geq \gamma(\delta,\varepsilon).$$

Par conséquent, pour tout sommet y non terminal de K^*, on a :

$$\begin{cases} \bar{m}(y) \geq \gamma(\delta,\varepsilon) \\ I(y) \geq (1-\rho)\ \gamma(\delta,\varepsilon). \end{cases}$$

En désignant par Y l'ensemble des sommets non terminaux de K^*, il s'en suit que :

$$H(\pi) \geq \sum_{y \in Y} p(y)\ I(y) \geq (1-\rho)\ \gamma(\delta,\varepsilon)\ L[K^*(\pi,\varepsilon)].$$

Intérêt algorithmique

Dans le cas où D est fini, nous avons avec les notations précédentes :

$$\begin{cases} \bar{m}(y) = \mathrm{Max}[I_{p_y}(q,T) \mid q \in D(y)] \\ D'(y) = \{q \in D(y) : I_{p_y}(q,T) = \bar{m}(y)\}. \end{cases}$$

Donc, les pseudoquestionnaires K^* sont faciles à engendrer par un algorithme. Nous étudierons plus précisémment au chapitre VIII l'intérêt de ces pseudoquestionnaires K^*.

CHAPITRE - VII

ETUDE DE LA H-CONVERGENCE

INTRODUCTION

Etant donné un détecteur D, nous démontrons une condition nécessaire et suffisante pour qu'il existe un pseudoquestionnaire construit sur D qui soit H-convergent pour tout élément π de $\mathbb{P}_n$.

Cette démonstration utilise des propriétés aux limites de la fonction gain d'information établies dans la 1ère partie.

Il semble que pour la caractérisation de ce type de décision, la méthode informationnelle soit particulièrement bien adaptée.

I - PROPRIETES PRELIMINAIRES

I - UNE CARACTERISATION DE LA H-CONVERGENCE

1) Proposition VII-1

Etant donné un détecteur D *, et un élément* π *de* $\mathbb{P}_n$*, une condition nécessaire et suffisante pour qu'un élément* K *de* $\mathcal{K}$(D) *soit* H*-convergent pour* π*, est que :*

$$I(K,\pi) = H(\pi)$$

Démonstration

Nous démontrons tout d'abord que la condition est nécessaire.
Pour ceci, considérons un élément K de $\mathcal{K}$(D), H-convergent pour π. Par définition de l'information transmise, nous avons :

$$I(K,\pi) = \lim_{R\to\infty} \left(\sum_{y\in Y(R)} p(y)\, I(y)\right)$$

où Y(R) désigne l'ensemble des sommets non terminaux de K(R).
D'après la propriété V-4, nous avons, pour tout R et tout $\varepsilon > 0$:

$$\sum_{y\in Y(R)} p(y)I(y) \geq H(\pi) - \sum_{y\in E(R;\pi,\varepsilon)} p(y)\; H(P(T|y)) - \sum_{x\in Z(R;\pi,\varepsilon)} p(x)\; H(P(T|x))$$

où Z(R ; π,ε) désigne l'ensemble des sommets de K(R) interprétables à ε près pour π.

Soit x un élément de Z(R ; π,ε) ; en supposant que

$$p(T_1|x) \geq 1-\varepsilon$$

on a :

$$H(\mathcal{P}(T|x)) \leq H(p(T_1|x),\ 1-p(T_1|x)) + \varepsilon \log (n-1)$$

$$H(p(T_1|x),\ 1-p(T_1|x)) \leq H(\varepsilon, 1-\varepsilon).$$

Par conséquent :

$$\sum_{y\in Y(R)} p(y)I(y) \geq H(\pi)-\log(n)\, p(E(R\ ;\pi,\varepsilon)) \quad -H(\varepsilon,1-\varepsilon)-\ \varepsilon\log(n-1).$$

Cette inégalité étant valable pour tout R, et tout $\varepsilon > 0$, nous avons à la limite, K étant H-convergent pour π :

$$I(K,\pi) \geq H(\pi).$$

Or, d'après la propriété V-5, nous avons

$$I(K,\pi) \leq H(\pi).$$

L'égalité est donc réalisée pour un élément K de $\mathcal{K}$(D) qui est H-convergent pour π.

Montrons maintenant que la condition est suffisante.

D'après la propriété V-4, nous avons pour tout R et tout $\varepsilon > 0$:

$$\sum_{y\in Y(R)} p(y)I(y) \leq H(\pi) - \sum_{y\in E(R;\pi,\varepsilon)} p(y)\ H(\mathcal{P}(T|y)) \leq H(\pi)-H(\varepsilon,1-\varepsilon)p(E(R;\pi,\varepsilon))$$

Par conséquent, si

$$I(K,\pi) = H(\pi)$$

alors, quel que soit $\varepsilon > 0$, nous avons :

$$\lim_{R\to\infty} (p(E(R\ ;\ \pi,\varepsilon)) = 0$$

et K est H-convergent pour π.

$\cancel{X}$

2) Une conséquence de la proposition VII-1

Soit $K = (A,u,B) \in \mathcal{K}(D)$; considérons le sous-ensemble $\tilde{D}$ de D défini par :

$$q \in \tilde{D} \iff \exists y \in Y \text{ tel que } u(y) = q$$

(Y désignant l'ensemble des sommets non terminaux de K).

L'ensemble des sommets d'une arborescence étant dénombrable, $\tilde{D}$ est dénombrable :

$$\tilde{D} = \{q(\ell)\},\ \ell \in \mathbb{N}.$$

Considérons alors un élément $\tilde{K} = (\tilde{A},\tilde{u},\tilde{B}) \in \mathcal{K}(\tilde{D})$ tel que pour tout sommet y non terminal de rang R de $\tilde{K}$ on ait :

$$\tilde{u}(y) = q(R).$$

PROPOSITION VII-1'

Si K *est* H-*convergent pour* $\pi \in \mathbb{P}_n$, *alors :*

$$I(\tilde{K},\pi) = H(\pi).$$

Démonstration

Il nous suffit de montrer que, quel que soit R, il existe R' ≥ R, tel que :

$$\sum_{y \in \tilde{Y}(R')} \tilde{p}(y)\ \tilde{I}(y) \geq \sum_{y \in Y(R)} p(y)\ I(y)$$

(les quantités attachées à $\tilde{K}$ sont surmontées de "~").
Considérons pour un sommet x de K l'ensemble

$$G(x) = \{q \in D : q \in u(A(x))\}$$

et posons :

$$G(R) = \bigcup_{\substack{x \in K \\ r(x)=R}} G(x).$$

En notant :

$$\tilde{D}(R) = \{q(\ell) : \ell \leq R\}$$

quel que soit R, il existe R' tel que :

$$G(R) \subset \tilde{D}(R').$$

Le pseudoquestionnaire K(R) peut alors être prolongé en un pseudoquestionnaire χ dont tous les chemins sont de longueur R', tel que chaque question de $\tilde{D}(R')$ soit utilisée une fois et une seule sur chaque chemin de χ.

L'information transmise par χ opérant sur π est alors la même que celle transmise par $\tilde{K}(R')$. Donc :

$$\sum_{y \in \tilde{Y}(R')} \tilde{p}(y)\ \tilde{I}(y) \geq \sum_{y \in Y(R)} p(y)\ I(y).$$

~~X~~

II - COMPORTEMENT ASYMPTOTIQUE DE L'INFORMATION TRANSMISE PAR UNE QUESTION

Nous nous intéressons au cas où T est un système complet de deux évènements, $T = (T_1,T_2)$ sur $(\Omega,\mathcal{a})$; considérons alors une suite de questions $\{q(r)\}$, $r \in \mathbb{N}$, de même base a telles que :

avec :
$$\alpha(q(r)) \xrightarrow[r\to\infty]{} s \in \mathcal{S}(a,2)$$
$$s^1 = s^2$$

Soit alors p la mesure engendrée sur $(\Omega,\mathcal{a})$ par un élément π de P_2 ; nous allons étudier le comportement de l'information transmise $I_p(q(r),T)$ quand r tend vers l'infini.

1) Cas d'une suite de questions dichotomiques (a = 2)

Nous utilisons les notations de VI-1, et considérons ainsi la fonction $\varphi(w,t,\lambda)$ définie sur $[0,1]^3$ par :

$$\varphi(w,t,\lambda) = H(\lambda w+(1-\lambda)t,1-\lambda w-(1-\lambda)t)-\lambda H(w,1-w)-(1-\lambda)H(t,1-t).$$

Avec cette notation, on a :

$$I_p(q,T) = \varphi(\alpha_1^1(q),\alpha_1^2(q),\pi_1).$$

En posant $z = (w,t)$, nous étudions le comportement de la fonction $\varphi(w,t,\lambda)$ au voisinage d'un point $\bar{z} = (\sigma,\sigma)$ tel que $\sigma \in \left]0,1\right[$.

PROPOSITION VII-2

Etant donné $\bar{z} = (\sigma,\sigma)$, $\sigma \in]0,1[$, *on a :*

i) $\forall\varepsilon \in]0,\frac{1}{2}[$, *il existe* $\xi_1(\bar{z}) > 0$ et $\delta_1(\bar{z}) > 0$ *tels que :*

$$\left\{\begin{array}{l} \|z-\bar{z}\| \leq \delta_1(\bar{z}) \\ \lambda \in [\varepsilon, 1-\varepsilon] \end{array}\right. \quad \Rightarrow \quad \varphi(w,t,\lambda) \geq \xi_1(\bar{z})\ (w-t)^2$$

ii) Il existe $\xi_2(\bar{z}) > 0$ *et* $\delta_2(\bar{z}) > 0$ *tels que :*

$$\left\{\begin{array}{l} \|z-\bar{z}\| \leq \delta_2(\bar{z}) \\ \lambda \in [0,1]. \end{array}\right. \quad \Rightarrow \quad \varphi(w,t,\lambda) \leq \xi_2(\bar{z})\ (w-t)^2$$

Démonstration

a) Montrons tout d'abord l'équivalence suivante :

(1) $$\varphi(w,t,\lambda) \sim \frac{\lambda(1-\lambda)}{2\sigma(1-\sigma)}\ (w-t)^2 \quad \text{lorsque } z \to \bar{z}.$$

Il est facile de constater que φ vérifie les conditions de Taylor à l'ordre 3 au voisinage de tout point z tel que :

$$(w,t) \neq (0,0) \quad , \quad (w,t) \neq (1,1).$$

Or, on a :

$$\varphi(\sigma,\sigma,\lambda) = 0 \qquad \forall\lambda \in [0,1]$$

$$\frac{\partial\varphi}{\partial w}(\sigma,\sigma,\lambda) = \frac{\partial\varphi}{\partial t}(\sigma,\sigma,\lambda) = 0 \qquad \forall\lambda \in [0,1].$$

De plus lorsque z tend vers $\bar{z}$, les quantités

$$\frac{\partial^2\varphi}{\partial w^2}(w,t,\lambda),\ \frac{\partial^2\varphi}{\partial t^2}(w,t,\lambda),\ \frac{\partial^2\varphi}{\partial w\partial t}(w,t,\lambda)$$

ont pour limite commune

$$\frac{\lambda(1-\lambda)}{\sigma(1-\sigma)}$$

ce qui achève la démonstration de (1).

b) Démonstration de i)

Etant donné $\varepsilon \in \left]0,\frac{1}{2}\right[$, en raison de (1), la fonction $\Psi(w,t,\lambda)$ définie sur $\left]0,1\right[^3$ par :

$$\Psi(w,t,\lambda) = \begin{cases} \dfrac{\varphi(w,t,\lambda)}{\dfrac{\lambda(1-\lambda)}{2w(1-w)}(w-t)^2} & \text{si } w \neq t \\ 1 & \text{si } w = t \end{cases}$$

est continue par rapport à (w,t,λ), z appartenant à un voisinage fermé de $\bar{z}$, λ appartenant à $[\varepsilon,1-\varepsilon]$.

Elle est donc uniformément continue et par conséquent :

quel que soit $\eta > 0$, il existe $\delta_1(\bar{z})$ tel que :

$$\begin{cases} \|z-\bar{z}\| \leq \delta_1(\bar{z}) \\ \lambda \in [\varepsilon,1-\varepsilon] \end{cases} \Rightarrow \varphi(w,t,\lambda) \geq (1-\eta)\frac{\lambda(1-\lambda)}{2\sigma(1-\sigma)}(w-t)^2 .$$

il suffit donc de poser :

$$\xi_1(\bar{z}) = (1-\eta)\frac{\varepsilon(1-\varepsilon)}{2\sigma(1-\sigma)} .$$

c) Démonstration de ii)

De même que dans b), pour tout $\varepsilon \in \left]0,\frac{1}{2}\right[$, nous avons :

$$(2)\quad \forall\eta > 0,\ \exists\delta(\bar{z})\ \text{tel que} \begin{cases} \|z-\bar{z}\| \leq \delta(\bar{z}) \\ \\ \lambda \in [\varepsilon, 1-\varepsilon] \end{cases} \Rightarrow \varphi(w,t,\lambda) \leq (1+\eta)\,\frac{\lambda(1-\lambda)}{2\sigma(1-\sigma)}\,(w-t)^2 .$$

Par ailleurs, en raison des propriétés de la fonction entropie [1], la fonction $\varphi(w,t,\lambda)$ est concave par rapport à λ.

Si nous désignons par $\lambda^*(w,t)$, la valeur de λ qui maximise φ pour (w,t) fixé, $\lambda^*(w,t)$ est solution de l'équation

$$\frac{\partial\varphi(w,t,\lambda)}{\partial\lambda} = 0.$$

Cette équation se résoud simplement, et on obtient :

$$\lambda^*(w,t) = \frac{w(e^{h(w,t)}+1)-1}{(w-t)(e^{h(w,t)}+1)}$$

avec :

$$h(w,t) = \frac{1}{w-t}\,[H(w,1-w)-H(t,1-t)]$$

Il est alors aisé de vérifier que $\lambda^*(w,t)$ est définie continue en tout point (w,t) tel que $w \neq t$, et que :

$$\lambda^*(w,t) \to \frac{1}{2} \text{ lorsque } |w-t| \to 0$$

ce qui permet de la prolonger par continuité.
Par conséquent :

$$\forall\varepsilon \in \left]0,\tfrac{1}{2}\right[,\ \exists\delta'(\bar{z}) > 0 \text{ tel que } \|z-\bar{z}\| \leq \delta'(\bar{z}) \Rightarrow \lambda^*(w,t) \in [\varepsilon, 1-\varepsilon].$$

Etant donné $\eta > 0$, et $\varepsilon \in \left]0,\frac{1}{2}\right[$, nous posons :

$$\delta_2(\bar{z}) = \mathrm{Min}\,(\delta(\bar{z}),\delta'(\bar{z})).$$

Alors pour tout z vérifiant

$$\|z-\bar{z}\| \leq \delta_2(\bar{z})$$

(2) aura lieu en particulier pour $\lambda^*(w,t)$, soit :

$$\|z-\bar{z}\| \leq \delta_2(\bar{z}) \Rightarrow \varphi(w,t,\lambda^*(w,t)) \leq \frac{(1+\eta)}{8\sigma(1-\sigma)}(w-t)^2 .$$

En raison de la concavité de φ par rapport à λ, nous avons par ailleurs :

$$\varphi(w,t,\lambda) \leq \varphi(w,t,\lambda^*(w,t)) \qquad \forall\lambda \in [0,1].$$

Il suffit donc de choisir $\xi_2(\bar{z})$ tel que :

$$\xi_2(\bar{z}) = \frac{(1+\eta)}{8\sigma(1-\sigma)}.$$

~~X~~

2) Cas d'une suite de questions polychotomiques

Nous considérons maintenant un détecteur dénombrable dont toutes les questions sont polychotomiques de base a.

$$D = \{q(r)\ ;\ r \in \mathbb{N}\}.$$

Etant donné $\varepsilon > 0$, nous noterons :

$$\mathbb{P}_n(\varepsilon) = \{\pi \in \mathbb{P}_n : \pi_j \geq \varepsilon \quad \forall j \in \{1,\ldots,n\}\}.$$

PROPOSITION VII-3

Si l'ensemble $\mathcal{S}$(D) *des points essentiellement atteints par* D *vérifie :*

$$s_i^1 = s_i^2 \neq 0 \qquad \forall i \in \{1,\dots,a\}, \qquad \forall s \in \mathcal{S}(D)$$

alors :

i) *quel que soit* $\varepsilon \in \left]0,\frac{1}{2}\right[$ *, il existe* $\xi_1 > 0$ *et un entier* r_o *tels que*

$$\forall r \geq r_o,\ \forall \pi \in \mathbb{P}_2(\varepsilon)\ :\ I_p(q(r),T) \geq \xi_1 \left[\alpha_i^1(q(r))-\alpha_i^2(q(r))\right]^2 \quad \forall i \in \{1,\dots,n\}.$$

ii) *Il existe* $\xi_2 > 0$ *et un entier* r_o *tels que :*

$$\forall r \geq r_o,\ \forall \pi \in \mathbb{P}_2\ :\ I_p(q(r),T) \leq \xi_2 \sum_{i=1}^{a} \left[\alpha_i^1(q(r))-\alpha_i^2(q(r))\right]^2.$$

Démonstration

a) Cas où $|\mathcal{S}(D)| = 1$

Nous noterons s l'élément de $\mathcal{S}$(D).

Etant donné un élément q de D, nous considérons les ensembles

$$V_k = \bigcup_{i=k}^{a} q_i \quad , \quad k \in \{1,\dots,a\}$$

et les questions σ^k, partitions de Ω définies par :

$$\sigma^k = (q_1,\dots,q_k,V_{k+1}) \quad , \quad k \in \{1,\dots,a-1\}.$$

D'après la propriété V-2, nous avons pour toute mesure p définie sur (Ω,a) :

$$(3) \qquad I_p(q,T) = \sum_{k=1}^{a-1} p(V_k)\ I_{p_{V_k}}(\sigma^k,T)$$

et en posant, pour $k \in \{1,\dots,a\}$

$$w_k = \frac{\alpha_k^1(q)}{\sum_{i=k}^{a} \alpha_k^1(q)} \quad , \quad t_k = \frac{\alpha_k^2(q)}{\sum_{i=k}^{a} \alpha_k^2(q)} \quad , \quad \lambda_k = p(T_1|V_k)$$

nous avons :

$$I_{p_{V_k}}(\sigma^k,T) = \varphi(w_k,t_k,\lambda_k).$$

Par ailleurs, il est facile de vérifier que lorsque

$$\alpha(q) \to s$$

pour tout $k \in \{1,\ldots,a\}$, w_k et t_k ont même limite σ telle que $\sigma \in]0,1[$.
Nous pouvons donc appliquer les résultats de la proposition VII-2.

★ Démonstration de i).

En raison de (3), nous avons :

$$I_p(q,T) \geq I_p(\sigma^1,T) = \varphi(\alpha_1^1(q),\alpha_1^2(q),\lambda_1)$$

et d'après la proposition VII-2 :

quel que soit $\varepsilon \in]0,\frac{1}{2}[$, il existe $\xi_1(s) > 0$ et $\delta_1(s) > 0$ tels que :

$$\left\{\begin{array}{l} \|\alpha(q)-s\| \leq \delta_1(s) \\ \\ \pi \in \mathbb{P}_2(\varepsilon) \end{array}\right. \Rightarrow \quad I_p(q,T) \geq \xi_1(s) \ [\alpha_1^1(q)-\alpha_1^2(q)]^2 .$$

Nous avons ainsi démontré i) pour i = 1, il est clair que cette démonstration a valeur de généralité.

★ Démonstration de ii).

En raison de (3), nous avons :

$$I_p(q,T) \leq \sum_{k=1}^{a-1} I_{p_{V_k}}(\sigma^k,T) = \sum_{k=1}^{a-1} \varphi(w_k,t_k,\lambda_k)$$

et d'après la proposition VII-2, il existe $\xi_2'(s) > 0$ et $\delta_2(s) > 0$ tels que :

$$\left\{\begin{array}{l} \|\alpha(q)-s\| \leq \delta_2(s) \\ \\ \pi \in \mathbb{P}_2 \end{array}\right. \Rightarrow \varphi(w_k,t_k,\lambda_k) \leq \xi_2'(s)\ (w_k-t_k)^2 .$$

Par ailleurs, on vérifie aisément que lorsque

$$\alpha(q) \to s$$

on a :

$$(w_k-t_k)^2 \sim \frac{[\alpha_k^1(q)-\alpha_k^2(q)]^2}{(\sum_{i=k}^{a} s_i^1)^2} .$$

Il suffit donc de poser :

$$\xi_2(s) = \frac{\xi_2'(s)}{\mathrm{Min}[(s_i^1)^2 \mid i \in \{1,\ldots,a\}]} > 0.$$

b) Cas général

Nous faisons la démonstration pour ii), la démontration de i) se ferait de la même façon.

Les résultats montrés dans a) nous permettent d'affirmer que pour tout $s \in \mathcal{S}(D)$, il existe $\xi_2(s) > 0$ et $\delta_2(s) > 0$ tels que :

$$\left\{\begin{array}{l} \|\alpha(q)-s\| \leq \delta_2(s) \\ \\ \pi \in \mathbb{P}_2 \end{array}\right. \Rightarrow I_p(q,T) \leq \xi_2(s) \sum_{i=1}^{a} [\alpha_i^1(q)-\alpha_i^2(q)]^2 .$$

Considérons alors l'ensemble

$$B_2 = \overline{\bigcup_{s\in\mathcal{S}(D)} B(s, \frac{\delta_2(s)}{2})}$$

où $B(s, \frac{\delta_2(s)}{2})$ est la boule fermée de centre s, de rayon $\frac{\delta_2(s)}{2}$ dans $\mathcal{S}(a,n)$.

B_2 est un compact, recouvert par la famille de boules ouvertes de centre $s \in \mathcal{S}(D)$, de rayon $\delta_2(s)$. On peut donc en extraire un recouvrement fini.

Par ailleurs, toutes les questions de D sauf peut être un nombre fini sont dans B_2. Donc, il existe une famille finie $s(\ell)$, $\ell \in \{1,\ldots,\rho\}$, d'éléments de $\mathcal{S}(D)$ telle qu'il existe un entier r_o vérifiant :

$$\forall r \geq r_o \quad \exists \ell \in \{1,\ldots,\rho\} \text{ tel que : } \|\alpha(q(r))-s(\ell)\| \leq \delta_2(s(\ell)).$$

Il suffit donc de poser :

$$\xi_2 = \mathrm{Max}(\xi_2(s(\ell)) \mid \ell \in \{1,\ldots,\rho\})$$

pour achever la démonstration.

~~X~~

II - CONDITION NECESSAIRE ET SUFFISANTE D'EXISTENCE DE PSEUDOQUESTIONNAIRES H-CONVERGENTS DANS LE CAS OU n = 2

La condition nécessaire et suffisante obtenue, dans le cas où n = 2, et dans le cas général, ne porte que sur les détecteurs réguliers définis en III-2. Rappelons qu'un détecteur D est régulier si tout élément s de $\mathcal{S}(D)$ vérifie : s'il existe un couple (j,j') tel que $s^j = s^{j'}$ alors $s_i^j \neq 0 \quad \forall i \in \{1,\dots,a(s)\}$.

THEOREME VII-1

n *étant supposé égal à* 2, *soit* D *un détecteur régulier, non trivial. Une condition nécessaire et suffisante pour qu'il existe un élément* K *de* $\mathcal{K}$(D) *qui soit* H-*convergent pour tout élément* π *de* $\mathbb{P}_2$, *est que* D *admette une suite d'éléments*

$$\{q(r)\} \ , \ r \in \mathbb{N},$$

telle que la série

$$\sum_r \sum_{i=1}^{a(q(r))} \left[\alpha_i^1(q(r))-\alpha_i^2(q(r))\right]^2$$

diverge.

III - DEMONSTRATION DE LA CONDITION SUFFISANTE

Nous désignons par D' l'ensemble des éléments de la suite {q(r)}.

Le nombre de bases distinctes étant fini, nous pouvons, sans restreindre la généralité, supposer que tous les éléments de D' ont même base a.

S'il existe s $\in \mathcal{S}(D')$ tel que $s^1 \neq s^2$, nous savons (Théorème IV-1), qu'il existe K $\in \mathcal{H}(D)$ qui soit L-convergent pour tout $\pi \in \mathbb{P}_2$, donc H-convergent pour tout $\pi \in \mathbb{P}_2$.

Nous supposons donc :

$$s^1 = s^2 \qquad \forall s \in \mathcal{S}(D').$$

Considérons alors un pseudoquestionnaire $K = (A,u,B) \in \mathcal{H}(D)$, tel que pour tout sommet y de rang r, non terminal, on ait :

$$u(y) = q(r).$$

Etant donné $\pi \in \mathbb{P}_2$, nous allons montrer que K est H-convergent pour π, c'est-à-dire qu'étant donné $\varepsilon > 0$

$$p(E(R\ ;\ \pi,\varepsilon)) \xrightarrow[R\to\infty]{} 0.$$

Soit Y(R) l'ensemble des sommets non terminaux de $K(R\ ;\ \pi,\varepsilon)$, d'après la propriété V-4, nous savons que :

$$\sum_{y\in Y(R)} p(y)\ I(y) \leq H(\pi) \qquad (\forall R).$$

Or :

$$Y(R) = \bigcup_{r=0}^{R-1} E(r\ ;\ \pi,\varepsilon)$$

et par conséquent, la série de terme général

$$W_r = \sum_{y\in E(r\ ;\pi,\varepsilon)} p(y)\ I(y)$$

converge.

Il nous suffit donc de montrer qu'il existe M(r) tel que :

$$\begin{cases} I(y) \geq M(r) & \forall y \in E(r\ ;\ \pi,\varepsilon) \\ \text{la série } \sum_r M(r) \text{ diverge} \end{cases}$$

pour s'assurer que la suite $p(E(r\ ;\ \pi,\varepsilon))$ décroissante tende vers zéro lorsque r tend vers l'infini.

Or, par hypothèse, la série

$$\sum_r \sum_{i=1}^{a} (\alpha_i^1(q(r))-\alpha_i^2(q(r)))^2$$

diverge ; donc il existe $i_o \in \{1,\dots,a\}$ tel que la série

$$\sum_r (\alpha_{i_o}^1(q(r))-\alpha_{i_o}^2(q(r))^2$$

diverge.

Par ailleurs, d'après la proposition VII-3, nous savons qu'il existe $\xi > 0$ et un entier r_o tels que :

$$\left\{\begin{array}{l} r \geq r_o \\ \\ \pi \in \mathbb{P}_2(\varepsilon) \end{array}\right. \Rightarrow I_p(q(r),T) \geq \xi\left[\alpha_{i_o}^1(q(r))-\alpha_{i_o}^2(q(r))\right]^2 .$$

Il suffit donc de poser :

$$M(r) = \xi\left[\alpha_{i_o}^1(q(r))-\alpha_{i_o}^2(q(r))\right]^2.$$

~~X~~

IV - DEMONSTRATION DE LA CONDITION NECESSAIRE

Nous supposons que quel que soit $D' \subset D$, D dénombrable, la série

$$\sum_{q\in D'} \sum_{i=1}^{a(q)} \left[\alpha_i^1(q)-\alpha_i^2(q)\right]^2$$

est convergente. En supposant que pour un élément π de $\mathbb{P}_2$, il existe un élément K_o de $\mathcal{K}(D)$ qui soit H-convergent pour π, nous allons montrer qu'il y a alors une contradiction.

Soit D', D' = $\{q(r) \mid r \in N\}$, l'ensemble des éléments de D utilisés dans $K_o = (A_o, u_o, B_o)$:

$$q \in D' \iff \exists x \in K_o \text{ tel que } u_o(x) = q.$$

D'après la proposition VII-1', nous savons que tout pseudoquestionnaire $K = (A, u, B) \in \mathcal{K}(D')$ tel que pour tout sommet y de rang r, non terminal on ait

$$u(y) = q(r),$$

vérifie :

$$I(K, \pi) = H(\pi).$$

Par ailleurs, en désignant par Y(r) l'ensemble des sommets non terminaux de K(r), en raison de la proposition VII-3, nous savons qu'il existe $\xi > 0$ et un entier r_o tels que :

$$\left\{\begin{array}{l} r \geq r_o \\ \\ y \in Y(r) \end{array}\right. \Rightarrow I(y) \leq \xi \sum_{i=1}^{a} \left[\alpha_i^1(q(r)) - \alpha_i^2(q(r))\right]^2.$$

En raison des hypothèses faites sur D', nous savons que :

$$(4) \qquad \forall \eta > 0 \quad \exists R \text{ tel que : } \sum_{r=R}^{\infty} \sum_{y \in E(r)} p(y)\, I(y) < \eta$$

où E(r) désigne l'ensemble des sommets non terminaux de rang r de K.

Par ailleurs, d'après le théorème III-2, il n'existe pas de pseudoquestionnaire F-convergent pour π, construit sur D ; par suite, il existe $\delta > 0$, tel que pour tout r, il existe un sommet y de rang r dans K tel que :

$$H(P(T|y)) > \delta.$$

Soit alors y_o, un tel sommet de rang R :

$$H(P(T|y_o)) > \delta.$$

L'information transmise par le sous-pseudoquestionnaire K_{y_o} de K de racine y_o opérant sur $P(T|y_o)$ est telle que :

$$I(K_{y_o}, P(T|y_o)) = \sum_{r=R}^{\infty} \sum_{y \in \hat{\Gamma}_{y_o} \cap E(r)} \frac{p(y)}{p(y_o)} I(y) = H(P(T|y_o)) > \delta .$$

Il y a donc contradiction avec (4), en choisissant :

$$\eta < \delta p(y_o).$$

~~X~~

V - CAS D'UN DETECTEUR NON REGULIER

Considérons le détecteur D non régulier suivant, constitué d'une suite de questions dichotomiques

$$\{q(r)\}, \ r \in \mathbb{N}$$

telle que :

$$\begin{cases} \alpha_1^2(q(r)) \xrightarrow[r\to\infty]{} 1 \\ \alpha_1^1(q(r)) = 1 \quad \forall r . \end{cases}$$

Tout élément K de $\mathcal{K}(D)$ admet au plus un sommet x_r de rang r non terminal. Soit un pseudoquestionnaire $K = (A,u,B) \in \mathcal{K}(D)$, tel que

$$u(x_r) = q(r).$$

Nous avons pour $\pi \in \mathbb{P}_2$ fixé :

$$p(T_1|x_R) = \frac{\pi_1}{\pi_1 + \prod_{r=0}^{R-1} \alpha_1^2(q(r))\pi_2}$$

Par conséquent, K est H-convergent pour π, si et seulement si :

$$p(T_1|x_R) \xrightarrow[R\to\infty]{} 1$$

c'est-à-dire, si et seulement si la série

$$\sum_r |\alpha_1^1(q(r))-\alpha_1^2(q(r))|$$

diverge.

Nous constatons donc, que dans ce cas de non régularité, la condition du théorème VII-1 reste suffisante, mais n'est plus nécessaire.

III - CONDITION NECESSAIRE ET SUFFISANTE D'EXISTENCE DE PSEUDOQUESTIONNAIRES H-CONVERGENTS - CAS GENERAL

Nous allons montrer le théorème suivant :

THEOREME VII-2

Soit D *un détecteur régulier, non trivial. Une condition nécessaire et suffisante pour qu'il existe* $K \in \mathcal{K}_0(D)$ *qui soit* H*-convergent pour tout élément* π *de* $\mathbb{P}_n$*, est que :*

$\forall j \in \{1,\dots,n\}$, $\forall j' \in \{1,\dots,n\}$, $j \neq j'$ *il existe une suite*

$\{q(r)\}$, $r \in \mathbb{N}$,

d'éléments de D*, telle que la série*

$$\sum_r \sum_{i=1}^{a(q(r))} \left[\alpha_i^j(q(r))-\alpha_i^{j'}(q(r))\right]^2$$

diverge.

Remarquons tout d'abord que la condition nécessaire se déduit immédiatement du théorème VII-1.

VI - DEMONSTRATION DE LA CONDITION SUFFISANTE

Nous supposons sans restriction de généralité, qu'il existe une suite $\{q(r)\}$, $r \in \mathbb{N}$, d'éléments de D, telle que pour tout couple (j,j'), la série

$$\sum_r \sum_{i=1}^{a(q(r))} \left[\alpha_i^j(q(r))-\alpha_i^{j'}(q(r))\right]^2$$

diverge.

Par ailleurs, nous supposons qu'il existe deux éléments j_1 et j_2 de $\{1,\dots,n\}$ tels que :

$$s^{j_1} = s^{j_2} \qquad \forall s \in \mathcal{S}(D).$$

Si ce n'était pas le cas, nous savons (Théorème IV-1), qu'il existerait $K \in \mathcal{K}(D)$ qui soit L-convergent pour tout π, donc H-convergent pour tout π.

Nous considérons un élément $K = (A,u,B) \in \mathcal{K}(D)$, tel que pour tout sommet y non terminal, et de rang r dans K on ait :

$$u(y) = q(r).$$

Nous allons montrer par récurrence sur le nombre t de composantes de π non nulles, que K est H-convergent pour tout π de $\mathbb{P}_n$.

La propriété est vraie pour t = 2, en raison du théorème VII-1, et nous faisons l'hypothèse de récurrence suivante :

$$\forall \pi \in \mathbb{P}_{n-1} \quad , \quad K \text{ est H-convergent pour } \pi.$$

1) Principe de la démonstration

Etant donné $\pi \in \mathbb{P}_n$, et $\varepsilon > 0$, nous voulons montrer qu'étant donné $\delta > 0$, il existe R_o tel que :

$$R \geq R_o \Rightarrow p(E(R\ ;\ \pi,\varepsilon)) \leq \delta.$$

Soient deux réels positifs ε_1 et ε_2 tels que :

$$\varepsilon_1 < \varepsilon_2 < \varepsilon.$$

Nous opérons une partition de $E(R\ ;\ \pi,\varepsilon)$ en trois sous-ensembles $E^1(R)$, $E^2(R)$, $E^3(R)$ définis comme suit :

$$E^1(R) = \{e \in E(R\ ;\ \pi,\varepsilon) : \forall y \in A(e) \cup \{e\},\ p(T_j|y) > \varepsilon_1 \quad \forall j \in \{1,\dots,n\}\}$$

$$E^2(R) = \{e \in E(R\ ;\ \pi,\varepsilon) : \begin{cases} 1)\ \exists x \in A(e)\ \exists j \in \{1,\dots,n\}\ p(T_j|x) < \varepsilon_1 \\ 2)\ \forall y \text{ sur le chemin } [x,e]\ p(T_j|y) < \varepsilon_2 \end{cases} \}$$

$$E^3(R) = \{e \in E(R\ ;\ \pi,\varepsilon) : \begin{cases} 1)\ \exists x \in A(e)\ \exists j \in \{1,\dots,n\}\ p(T_j|x) < \varepsilon_1 \\ 2)\ \exists y \text{ sur le chemin } [x,e]\ p(T_j|y) \geq \varepsilon_2 \end{cases} \}$$

Nous remarquons que :

- si $x \in E^1(R)$, ses successeurs appartiennent soit à $E^1(R+1)$ soit à $E^2(R+1)$
- si $x \in E^2(R)$, ses successeurs appartiennent soit à $E^2(R+1)$ soit à $E^3(R+1)$
- si $x \in E^3(R)$ ses successeurs appartiennent à $E^3(R+1)$.

Nous allons montrer que :

a) il existe R_1 tel que :

$$p(E^1(R_1) \leq \frac{\delta}{3} \cdot$$

La suite $\{p(E^1(R))\}$ étant décroissante, nous aurons donc :

$$\forall R \geq R_1 \quad p(E^1(R)) \leq \frac{\delta}{3}$$

b) ε_2 ayant été choisi inférieur à ε, il existe R_2 tel que

$$\forall R \geq R_2 \quad p(E^2(R)) \leq \frac{\delta}{3} \quad \text{(grâce à l'hypothèse de récurrence)}$$

c) $p(E^3(R)) \leq \dfrac{\varepsilon_1}{\varepsilon_2}$ $\forall R$ (par la propriété de "conservation du flux").

Il suffira donc de choisir ε_1 et ε_2 tels que : $\dfrac{\varepsilon_1}{\varepsilon_2} \leq \dfrac{\delta}{3}$.

2) Etude de la suite $\{p(E^3(R))\}$

Nous supposons R fixé, et à tout élément $e \in E^3(R)$, nous associons $x(e)$ et $y(e)$ où :

- $x(e)$ est le sommet de rang minimal appartenant à $A(e)$, tel que :

$$\exists j \in \{1,\ldots,n\} : p(T_j | x(e)) < \varepsilon_1$$

- $y(e)$ est le sommet de rang minimal sur le chemin $[x(e),e]$, tel que :

$$p(T_j | y(e)) \geq \varepsilon_2.$$

Posons :

$$V = \{x(e) \mid e \in E^3(R)\}.$$

En raison du choix des $x(e)$, V est une partie de section de $K(R)$, par conséquent :

(5) $$p(E^3(R)) = p(V).$$

Par ailleurs, pour tout sommet $x \in V$, on définit :

$$V(x) = E^3(R) \cap \hat{\Gamma}_x$$

$$W(x) = \{y(e) \mid e \in V(x)\}.$$

En raison du choix des $y(e)$, $W(x)$ est une partie de section du sous-pseudoquestionnaire K_x, de racine x, de K, et

$$p(V(x)) = p(W(x)).$$

Pour tout élément x de V, nous avons :

$$\sum_{y \in W(x)} \frac{p(y)}{p(x)} p(T_j | y) \leq p(T_j | x) < \varepsilon_1.$$

Or :

$$p(T_j|y) \geq \varepsilon_2 \qquad \forall y \in W(x).$$

Par suite :

$$p(V(x)) = \sum_{y \in W(x)} p(y) < \frac{\varepsilon_1}{\varepsilon_2} p(x).$$

En tenant compte de (5), il vient alors :

$$p(E^3(R)) = \sum_{x \in V} p(V(x)) < \frac{\varepsilon_1}{\varepsilon_2} \sum_{x \in V} p(x) < \frac{\varepsilon_1}{\varepsilon_2} .$$

3) Etude de la suite $\{p(E^1(R))\}$

Par hypothèse, il existe un couple (j_1,j_2) tel que

$$s^{j_1} = s^{j_2} \quad \forall s \in \mathcal{S}(D).$$

Soit alors x un élément de $E^1(R)$, et p_x la mesure engendrée par $P(T|x)$ sur $(\Omega,\mathcal{a})$.

Considérons la mesure $\tilde{p}_x$ engendrée sur $(\Omega,\mathcal{a})$ par $\tilde{\pi} \in \mathbb{P}_n$, défini comme suit :

$$\tilde{\pi}_{j_1} = \frac{p(T_{j_1}|x)}{p(T_{j_1}|x)+p(T_{j_2}|x)}$$

$$\tilde{\pi}_{j_2} = \frac{p(T_{j_2}|x)}{p(T_{j_1}|x)+p(T_{j_2}|x)}$$

$$\tilde{\pi}_j = 0 \qquad \forall j \in \{1,..,n\}-\{j_1,j_2\}.$$

D'après la proposition V-2 :

$$I(x) = I_{p_x}(q(R),T) \geq (p(T_{j_1}|x)+p(T_{j_2}|x))\ I_{\tilde{p}_x}(q(R),T).$$

Or, de la même façon qu'au paragraphe 3, nous pouvons supposer, sans restreind la généralité, que les éléments de la suite $\{q(r)\}$ ont tous même base a, et qu'il existe $i_o \in \{1,\dots,a\}$ tel que la série

$$\sum_r (\alpha_{i_o}^{j_1}(q(r))-\alpha_{i_o}^{j_2}(q(r)))^2$$

diverge.

D'après la proposition VII-3, nous savons qu'il existe $\xi > 0$ et un entier R_o tels que :

$$\left\{\begin{array}{l} R \geq R_o \\ \\ x \in E^1(R) \end{array}\right. \Rightarrow I_{\tilde{p}_x}(q(R),T) \geq \xi\ \left[\alpha_{i_o}^{j_1}(q(R))-\alpha_{i_o}^{j_2}(q(R))\right]^2 .$$

Par suite, la série

$$\sum_{R=0}^{\infty} \sum_{x \in E^1(R)} p(x)\ I(x)$$

étant convergente, sa somme étant majorée par $H(\pi)$, la suite $\{p(E^1(R))\}$ décroissante, tend vers zéro quand R tend vers l'infini.

4) Etude de la suite $\{p(E^2(R))\}$

Soit R fixé, nous considérons $R' > R$.

Nous opérons une partition de $E^2(R')$ en deux sous-ensembles $V(R')$ et $W(R')$ tels que étant donné $e \in E^2(R')$:

- $e \in W(R')$ si e a un ascendant dans $E^1(R)$
- $e \in V(R')$ sinon.

Nous avons trivialement :

$$p(W(R') \leq p(E^1(R)).$$

Il nous suffit donc de montrer que $p(V(R'))$ tend vers zéro quand $(R'-R)$ tend vers l'infini.

Soit donc $e \in V(R')$ et x son ascendant dans $E^2(R)$, nous savons que :

$$\exists j_o \in \{1,\dots,n\} \text{ tel que } \forall y \in [x,e] : p(T_{j_o}|y) < \varepsilon_2 .$$

Pour simplifier les notations nous supposerons $j_o = n$, et considérons $\tilde{\pi} \in \mathbb{P}_{n-1}$, défini comme suit :

$$\tilde{\pi}_j = \frac{p(T_j|x)}{1-p(T_n|x)} \qquad \forall j \in \{1,\dots,n-1\}.$$

Soit $\tilde{p}$ la mesure engendrée sur $(\Omega,\mathcal{a})$ par $\tilde{\pi}$; considérons le sous-pseudoquestionnaire K_x de racine x, de K.

En désignant par $E_x(\rho ; \tilde{\pi},\tilde{\varepsilon})$ les sommets non terminaux de rang $\rho = R'-R$ de $K_x(\tilde{\pi},\tilde{\varepsilon})$, ($\tilde{\varepsilon} > 0$) , nous allons montrer que pour tout ε_2, $0 < \varepsilon_2 < \varepsilon$, il existe $\tilde{\varepsilon} > 0$ tel que :

$$(6) \qquad V(R') \subset E_x(\rho ; \tilde{\pi},\tilde{\varepsilon}).$$

Remarquons tout d'abord que si on note $\tilde{p}$ les probabilités des éléments de K_x opérant sur $\tilde{\pi}$, on a :

$$\tilde{p}(T_j|y) = \frac{p(T_j|y)}{1-p(T_n|y)} \quad \forall y \in K_x.$$

Or, pour tout élément $e \in V(R')$, on a :

$$\left.\begin{array}{ll} p(T_j|y) < 1-\varepsilon & \forall j \in \{1,\dots,n-1\} \\ p(T_n|y) < \varepsilon_2 & \end{array}\right\} \forall y \in (A(e) \cap \hat{\Gamma}_x) \cup \{e\}.$$

Par suite :

$$\tilde{p}(T_j|y) \le \frac{1-\varepsilon}{1-\varepsilon_2} \qquad \forall j \in (A(e) \cap \hat{\Gamma}_x) \cup \{e\}$$

et pour réaliser (6) il suffit donc de choisir $\tilde{\varepsilon}$ tel que :

$$\tilde{\varepsilon} = 1 - \frac{1-\varepsilon}{1-\varepsilon_2} .$$

Montrons maintenant que, $\tilde{\varepsilon}$ étant ainsi choisi, nous avons :

(7) $$p(V(R')) \leq \frac{1}{1-\varepsilon_2} \; \tilde{p}(E_x(\rho,\tilde{\pi},\tilde{\varepsilon})).$$

Pour ce faire, il nous suffit de montrer que :

$$p(e) \leq \frac{1}{1-\varepsilon_2} \, \tilde{p}(e) \qquad \forall e \in V(R').$$

Or :

$$\frac{p(e)}{\tilde{p}(e)} = \frac{\sum_{j=1}^{n} p(T_j|x) \dfrac{\Lambda_j(e)}{\Lambda_j(x)}}{\sum_{j=1}^{n-1} \tilde{\pi}_j \dfrac{\Lambda_j(e)}{\Lambda_j(x)}}$$

soit :

$$\frac{p(e)}{\tilde{p}(e)} = (1-p(T_n|x)) \left[1+ \frac{p(T_n|x) \dfrac{\Lambda_n(e)}{\Lambda_n(x)}}{\sum_{j=1}^{n-1} p(T_j|x) \dfrac{\Lambda_j(e)}{\Lambda_n(e)}}\right].$$

Or :

$$p(T_n|e) = \frac{p(T_n|x) \dfrac{\Lambda_n(e)}{\Lambda_n(x)}}{\sum_{j=1}^{n} p(T_j|x) \dfrac{\Lambda_j(e)}{\Lambda_j(x)}} < \varepsilon_2 \qquad \forall e \in V(R').$$

D'où :

$$p(T_n|x) \frac{\Lambda_n(e)}{\Lambda_n(x)} \leq \frac{\varepsilon_2}{1-\varepsilon_2} \sum_{j=1}^{n-1} p(T_j|x) \frac{\Lambda_j(e)}{\Lambda_j(x)} \qquad \forall e \in V(R')$$

et nous en déduisons (7) directement.

Pour conclure, il suffit de remarquer, qu'en raison de l'hypothèse de récurrence :

$$\tilde{p}(E_x(\rho,\tilde{\pi},\tilde{\varepsilon})) \xrightarrow[\rho \to \infty]{} 0$$

~~X~~

CHAPITRE - VIII

APPLICATIONS

INTRODUCTION

Nous présentons deux applications des pseudoquestionnaires.
La première se situe dans le cadre du codage avec bruit ; nous établissons des liens entre le théorème IV-1, relatif à la L-convergence et le théorème fondamental de la théorie de l'information [1], et présentons une interprétation des résultats obtenus sur la H-convergence.
L'autre application concerne un certain type de problème de reconnaissance de structures que nous illustrons par deux exemples, l'un relatif à l'aide au diagnostic en toxicologie, l'autre à un problème de segmentation.

I - APPLICATION AU PROBLEME DU CODAGE AVEC BRUIT

De même que les résultats sur les questionnaires s'appliquaient au problème du codage sans bruit, nous allons voir que les résultats sur les pseudoquestionnaires s'appliquent au cas du codage avec bruit.

1) Rappel [1], [29], [10].

Nous considérons le cas de canaux de communication donnant lieu à des erreurs de transmission ; plus précisément, nous considérons un canal discret et sans mémoire,

$$C = (\Gamma, \Gamma', \gamma)$$

où :

- $\Gamma = (E_1, \ldots, E_n)$ est un ensemble fini correspondant aux symboles d'entrée.
- $\Gamma' = (S_1, \ldots, S_a)$ est un ensemble fini correspondant aux symboles de sortie.
- γ est une matrice stochastique dont les éléments sont les probabilités conditionnelles

$$\gamma_i^j = p(S_i | E_j) \quad i \in \{1, \ldots, a\}, \quad j \in \{1, \ldots, n\}.$$

Si l'alphabet d'entrée est probabilisé avec :

$$\pi_j = p(E_j) \qquad j \in \{1, \ldots, n\}$$

l'alphabet de sortie est alors probabilisé de la façon suivante :

$$p(S_i) = \sum_{j=1}^{n} \pi_j \gamma_i^j .$$

Soit

$$I_p(\Gamma', \Gamma)$$

l'information apportée par Γ' sur Γ pour la mesure p ; on appelle *capacité* du canal, la quantité $\chi(C)$ définie par :

$$\chi(C) = \mathrm{Max}(I_p(\Gamma', \Gamma) \mid \pi \in \mathbb{P}_n).$$

Soit alors une R-séquence d'entrée x, c'est-à-dire une séquence de R caractères de Γ, il lui correspond après transmission par le canal, une R-séquence de sortie y dont les caractères appartiennent à Γ'.

Un code (m,R) est alors défini par :

i) un ensemble de R-séquences d'entrée $x^1,\dots,x^m$, appelées "mots de code".

ii) un schéma de décodage, correspondant au principe du maximum de vraisemblance, qui à une R-séquence de sortie associe une des m R-séquences d'entrée.

Soit alors $p(e|x^k)$, $k \in \{1,\dots,n\}$, la "probabilité d'erreur" de décodage, sachant que c'est le mot de code x^k qui a été émis ; on appelle maximum de probabilité d'erreur du code, la quantité

$$\bar{p}(e) = \mathrm{Max}(p(e|x^k) \mid k \in \{1,\dots,m\}).$$

Un code $\mathcal{C} = (m,R,\lambda)$ est un code (m,R) dont le maximum de probabilité d'erreur est inférieur ou égal à λ.

Le théorème fondamental de l'information démontre que [1] :

si $\chi(C) > 0$, alors il existe un nombre positif $\rho < \chi(C)$ tel que la suite de codes $\{\mathcal{C}_R\}$,

$$\mathcal{C}_R = (\lceil a^{\rho R}\rceil, R, \lambda_R)$$

vérifie :

$$\lambda_R \to 0 \text{ lorsque } R \to \infty$$

$\lceil a^{\rho R}\rceil$ désignant le plus grand entier inférieur ou égal à $a^{\rho R}$.

2) Pseudoquestionnaire et code

a) Soit donné un canal $C = (\Gamma,\Gamma',\gamma)$, nous assimilons la transmission d'un caractère à une expérience q, ayant a issues possibles $(q_1,\dots,q_a)$ correspondant aux a caractères $(S_1,\dots,S_a)$ de Γ' susceptibles d'avoir été recueillis, et nous posons :

$$p(q_i | E_j) = \gamma_i^j.$$

Ainsi, à la Rème émission, nous associons l'expérience q(R), telle que pour tout R

$$p(q_i(R) | E_j) = \gamma_i^j.$$

Soit alors D le détecteur constitué de la suite $\{q(R)\}$, et soit $K \in \mathcal{K}(D)$, $K = (\mathcal{A},u,B)$, un pseudoquestionnaire tel que pour tout sommet y non terminal de rang R de K, on ait :

$$u(y) = q(R).$$

En raison de la structure des pseudoquestionnaires, K ne permet d'opérer que sur des R-séquences d'entrée du type

(1) $$\underbrace{(E_j, E_j, \ldots, E_j)}_{\text{R fois}} .$$

Nous ne pouvons donc prendre ainsi en compte que n mots de code distincts.

La restriction de hauteur R, K(R) de K, peut être assimilée au code particulier

$$\mathcal{C}_R = (n, R, \lambda_R)$$

où l'ensemble des n R-séquences d'entrée sont du type (1).

La H-convergence de K est équivalente au fait que la suite $\{\mathcal{C}_R\}$ vérifie :

$$\lambda_R \xrightarrow[R\to\infty]{} 0.$$

Or, dans ce cas particulier, où tous les éléments de D ont même matrice stochastique, il est clair que la H-convergence est équivalente à la L-convergence ; et d'après le théorème IV-1, la L-convergence ("uniforme" par rapport à $\pi \in \mathbb{P}_n$) a lieu pour K si et seulement si :

(2) $\qquad \forall j \in \{1,\ldots,n\},\ \forall j' \in \{1,\ldots,n\},\ j \neq j' : \gamma^j \neq \gamma^{j'}$

γ^j désignant l'élément de $\mathbb{R}^a$ de composantes γ_i^j.

Il est facile de vérifier que la condition (2) est équivalente au fait que le canal C a une capacité strictement positive.

Ainsi dans le cas particulier où tous les éléments de D ont même matrice stochastique, le théorème IV-1 coïncide avec le théorème fondamental de l'information pour une famille particulière de codes.

b) Voyons maintenant comment le théorème VII-2 s'interprète en termes de codage.

Nous supposons qu'au cours de l'émission, le support physique du canal se dégrade de telle sorte que sa capacité tende vers zéro, lorsque le temps d'émission tend vers l'infini.

Nous pouvons prendre en compte ce processus, en associant à la Rème émission, une expérience q(R) ayant a issues ; en notant

$$p(q_i(R)|E_j) = \gamma_i^j(R)$$

nous supposons que les éléments $\gamma^j(R)$ de $\mathbb{R}^a$ ont tous mêmes limites σ quand R tend vers l'infini.

Soit D le détecteur constitué de la suite $\{q(R)\}$, et $K = (A,u,B) \in \mathcal{K}(D)$, tel que pour tout sommet y non terminal et de rang R de K, on ait :

$$u(y) = q(R).$$

Nous assimilons comme précédemment K(R) au code particulier

$$\mathcal{C}_R = (n,R,\lambda_R)$$

où les n R-séquences d'entrée sont du type (1).

Le théorème VII-2 peut alors s'interpréter comme suit :

la suite des codes $\mathcal{C}_R$ vérifiera

$$\lambda_R \xrightarrow[R\to\infty]{} 0$$

si et seulement si la capacité du canal ne tend pas "trop vite" vers zéro ; plus exactement, et en supposant $\sigma_i \neq 0 \quad \forall i \in \{1,\dots,a\}$, si et seulement si les séries

$$\sum_R \sum_{i=1}^{a} \left[\gamma_i^j(q(R)) - \gamma_i^{j'}(q(R))\right]^2$$

divergent pour tout couple (j,j').

II - APPLICATION A CERTAINS PROBLEMES DE RECONNAISSANCE DE STRUCTURES

Nous présentons deux applications pratiques, l'une relative à l'aide au diagnostic en toxicologie, l'autre à un problème de segmentation.

Si ces deux problèmes admettent un formalisme analogue en termes de pseudoquestionnaires, l'exploitation des pseudoquestionnaires ainsi construits différera dans les deux cas ; nous avons été ainsi conduits à dégager deux problèmes d'optimisation différents pour des pseudoquestionnaires construits sur un détecteur fini.

1) Problème P1 : Aide au diagnostic en toxicologie

a) Exposé du problème [5], [7].

Afin de déterminer le toxique origine de l'intoxication d'un malade, le médecin opère sur ce malade un certain nombre de tests, et établit son diagnostic en fonction des "réponses" à ces tests. Ses difficultés sont de deux ordres :

- Mémoriser et intégrer un très grand nombre de données. On distingue en effet, environ 200 familles toxiques (pour 200.000 produits toxiques actuellement sur le marché), alors qu'en cardiologie par exemple, le nombre de diagnostics possibles n'excède pas la cinquantaine, et le toxicologue dispose d'environ 150 tests dont le nombre d'issues possibles peut être souvent bien supérieur à 2.

- L'impossibilité, vue l'urgence, de pratiquer une analyse toxicologique complète des liquides biologiques, analyses dont la réalisation demande plusieurs heures. Par ailleurs, le nombre d'analyses réalisables est encore limité par le volume nécessairement réduit de la prise d'essai constituée par des liquides biologiques (sang, urine, etc...).

Le problème consiste donc à déterminer le toxique T de la famille $(T_1,\ldots,T_n)$, qui a intoxiqué un malade. Pour ceci, on est susceptible de faire subir au malade des tests cliniques, on encore de lui "poser des questions" appartenant à un ensemble D donné.
On s'attachera pour cela à établir un processus d'interrogation ayant des qualités de *rapidité* et de *sécurité* quand à l'interprétation.

b) Etude du problème P1

Nous notons

$$q_1,\ldots,q_{a(q)}$$

les issues d'une question q de D.

Une étude statistique permet d'appréhender les probabilités conditionnelles

$$p(q_i|T_j) \quad j \in \{1,\ldots,n\}, \quad i \in \{1,\ldots,a(q)\}, \quad q \in D$$

dans la population Ω des intoxiqués, ainsi que les probabilités à priori, soit

$$\pi_1,\ldots,\pi_n$$

des éléments $T_1,\ldots,T_n$ dans Ω.

Nous sommes donc en mesure de construire sur le détecteur toxicologique D un pseudoquestionnaire K opérant sur π.
Voyons maintenant comment traduire, en termes de pseudoquestionnaires, les exigences de sécurité et de rapidité du diagnostic.

Exigence de sécurité du diagnostic

La sécurité du diagnostic établi à l'aide d'un pseudoquestionnaire K peut être mesurée par le risque d'erreur de ce pseudoquestionnaire.
Comme nous l'avons vu en I-5-3, ce risque global est la somme de deux risques ρ_1 et ρ_2 de nature différente :

- ρ_1 étant le risque de diagnostiquer un toxique qui en fait n'est pas en cause.

- ρ_2 étant le risque de ne pas pouvoir faire le diagnostic.

* Etant donné la gravité du risque ρ_1, nous devons être très exigeants à son égards ; pour ce faire, nous utilisons le critère de décision suivant :

Etant donné $\varepsilon \in \left]0,\frac{1}{n}\right[$, nous faisons le diagnostic T_{j_o} au sommet x de K, si et seulement si :

$$p(T_{j_o}|x) \geq 1-\varepsilon.$$

Il est clair que lorsqu'on aboutit à un tel sommet x, on arrêtera là l'observation du malade.
Par conséquent, pour ε fixé par le praticien, le pseudoquestionnaire utilisé est la restriction $K(\pi,\varepsilon)$ de K.
Nous avons vu en I-5-3 que ρ_1 était majoré par ε ; et par conséquent on est maître de ce risque d'erreur.

* Considérons maintenant le risque ρ_2. Soit $E(\pi,\varepsilon)$ l'ensemble des sommets terminaux de K non interprétables à ε près pour π, on a :

$$\rho_2 = p(E(\pi,\varepsilon)).$$

Or nous avons vu que pour tout ε, le risque ρ_2 est nul, si et seulement si, le pseudoquestionnaire K est H-convergent pour π.

Le détecteur toxicologique D étant fini, et non trivial, il ne sera pas possible de réduire arbitrairement ρ_2.

Il est donc important de pouvoir estimer à priori, au vu de D, π, ε, l'ordre de grandeur de la plus petite erreur ρ_2 que l'on puisse espérer.
Pour ceci, nous considérons les quantités :

$$H_f[K(\pi,\varepsilon)] = \sum_{x\in Z(K(\pi,\varepsilon))} p(x)\ H(P(T|x))$$

où Z(K) désigne l'ensemble des sommets terminaux d'un pseudoquestionnaire K.

Il est facile de montrer que :

$$(3) \qquad \rho_2 \, H(\varepsilon,1-\varepsilon) \le H_f[K(\pi,\varepsilon)] \le H(\varepsilon,1-\varepsilon)+\varepsilon \log(n-1)+\rho_2 \log n.$$

Par ailleurs, soit K_o un pseudoquestionnaire construit sur D, tel que toute question q de D figure une fois et une seule sur chacun de ses chemins ; il est clair que quel que soit K construit sur D et opérant sur π, on a :

$$H_f(K_o) \le H_f(K).$$

Soit alors $\mathcal{K}(D,\pi,\varepsilon,\eta)$ l'ensemble des pseudoquestionnaires construits sur D, opérant sur π, avec interprétation à ε près, et tels que :

$$\rho_2 \le \eta.$$

Une condition nécessaire pour que cet ensemble soit non vide est que la quantité $H_f(K_o)$ vérifie :

$$H_f(K_o) \le \eta \log n + H(\varepsilon,1-\varepsilon)+ \varepsilon \log (n-1).$$

<u>Exigence de la rapidité du diagnostic</u>

L'ensemble $\mathcal{K}(D,\pi,\varepsilon,\eta)$ étant maintenant supposé non vide, nous avons à trouver un élément $\bar{K}$ de cet ensemble dont la longueur de cheminement soit minimale :

$$L[\bar{K},\pi] = \mathrm{Min}(L[K,\pi] \mid K \in \mathcal{K}(D,\pi,\varepsilon,\eta)).$$

Nous n'avons pas résolu ce problème. Cependant, en raison de (3), nous avons été conduits à utiliser le pseudoquestionnaire $K^* = (A,u,B)$ défini de la façon suivante :

- un sommet est terminal dans K^*, si et seulement si il est interprétable à ε près pour π, ou de rang égal à $|D|$.

- pour un sommet y non terminal, on choisit u(y) tel que :

$$I_{P_y}(u(y),T) = \mathrm{Max}\left[I_{P_y}(q,T) \mid q \in D(y)\right]$$

où :

. $D(y) = \{q \in D : u(x) \neq q \quad \forall x \in A(y)\}$

. P_y est la mesure engendrée par $P(T|y)$.

Cette heuristique consiste donc à poser en chaque sommet la question qui apporte le plus d'information sur l'état du système T.

Nous n'avons montré aucun résultat théorique à ce sujet, si ce n'est la proposition VI-3 relative à la L-convergence de K^*, lorsque D est infini.

Il est clair que la difficulté essentielle provient du fait que l'information transmise par K^* n'est pas nécessairement maximum, bien que localement, l'information transmise par chaque question de K^* le soit.

Remarquons toutefois que dans le cas particulier où les pseudoquestionnaires construits sur D sont des questionnaires, cette heuristique coïncide avec l'algorithme de Fano [9], visant à chercher un questionnaire de longueur de cheminement minimale. L'optimum obtenu par l'algorithme de Huffman [12] ne coïncide généralement pas avec le résultat de l'algorithme de Fano, mais leur longueur de cheminement diffère au plus d'une unité.
Ceci nous laisse supposer que K^* a une "bonne" longueur de cheminement, nous n'avons toutefois pas pu estimer sa distance par rapport à l'optimum.

Signalons pour conclure, que Sobel [23], [24] a tenté de valoriser des heuristiques voisines.

c) Mise en oeuvre pratique

Le médecin dispose en fait de deux catégories de tests ; les "tests à vue" (une quarantaine) n'exigeant pas la mise en oeuvre d'un matériel ou d'une expérimentation compliquée, et des "tests chers" exigeant cette mise en oeuvre.

Etant donné la suite des "réponses" aux tests à vue, une simple itération de la formule de Bayes permet d'obtenir la nouvelle liste des probabilités $(\pi'_1,\ldots,\pi'_n)$ des toxiques. C'est donc en fait sur le détecteur D constitué des "tests chers" que sera construit le pseudoquestionnaire K^*, et K^* opèrera sur $(\pi'_1,\ldots,\pi'_n)$.

La collecte des données [15], et le programmation ont été mises en oeuvre dans le cadre du contrat DGRST N° 68-01-446. Etant donné l'importance des données nécessaires, ce programme n'a pu être testé que sur un sous-ensemble de 30 toxiques, le nombre de tests chers pris en compte étant de 20 (ces tests sont dichotomiques).

Ce programme calcule tout d'abord la liste des probabilités $(\pi'_1,\ldots,\pi'_n)$ à partir des réponses aux tests à vue, il construit ensuite le pseudoquestionnaire K^* opérant sur π' et fournit la fréquence d'apparition des différents tests chers figurant dans K^*. Ces résultats permettent au médecin de mettre en oeuvre simultanément les analyses susceptibles d'affiner son diagnostic.
Disposant ensuite des réponses aux tests chers ainsi mis en oeuvre, le programme fournit une dernière liste des probabilités des toxiques.

Dans le but de préfigurer d'une manière plus réaliste une utilisation future, la technique de programmation a été rendue conversationnelle grâce au système CP.CMS.
Les résultats obtenus ayant été très encourageants, une équipe médicale est en train de rassembler les données concernant l'ensemble des 200 toxiques existants.

L'exemple suivant reproduit les sorties du programme concernant l'intoxication d'un malade par le Parathion.

PROJET D'AIDE AU DIAGNOSTIC EN TOXICOLOGIE CLINIQUE

VEUILLEZ INDIQUER LES NUMEROS DES SYMPTOMES OBSERVES CHEZ LE MALADE

-35

VOMISSEMENTS

-15

HYPERSUDATION

-20

MYOSIS

-33

TR.PSYCHIQUES DELIRE AGITATION

-7

CRAMPES MEMBRES INFERIEURS

-2

CEPHALEES

-4

CLONIES FIBRIL. MUSCULAIRES

-fin

AU VU DES SEULES INFORMATIONS FOURNIES PAR L'EXAMEN CLINIQUE DU MALADE

LES PROBABILITES EN MILLIEMES DES DIFFERENTS TOXIQUES SONT

DERIVES AMPHETAMINES	419
PARATHION	374
DIGITALINE	49
PLOMB	33
I. N. H.	26
OPIACES	22
LARGACTIL	14
TOFRANIL	11

LA SOMME DES PROBABILITES DES TOXIQUES NON CITES EST INFERIEURE A 5 POUR CENT

AFIN DE PRECISER OU DE CONFIRMER CES PREMIERES INDICATIONS

VOICI LES PROBABILITES D'AVOIR A CONSIDERER LE OU LES ELEMENTS PARACLINIQUES SUIVAN

TROUBLES DE REPOLARISATION	1000
CHUTE DES CHOLINESTERASES	1000
TROUBLES DE CONDUCTION	647
CHUTE DE LA RESERVE ALCALINE	639
CHUTE DE POTASSIUM (<3MEQ)	391
POTASSIUM AUGMENTE (>6MEQ)	194
GLOBULES BLANCS (>15000)	102

VEUILLEZ INDIQUER LES RESULTATS OBTENUS AUX EXAMENS SUIVANTS

CHUTE DES CHOLINESTERASES

-oui

LA LISTE DES PROBABILITES EST AINSI MODIFIEE

PARATHION	827
DERIVES AMPHETAMINES	115
DIGITALINE	13
PLOMB	9
I.N.H.	7
OPIACES	6
LARGACTIL	4

TROUBLES DE REPOLARISATION

-non

LA LISTE DES PROBABILITES EST AINSI MODIFIEE

PARATHION	834
DERIVES AMPHETAMINES	116
PLOMB	9
I.N.H.	7
OPIACES	6
DIGITALINE	4
LARGACTIL	4

CHUTE DU POTASSIUM (<3 MEQ)

-non

LA LISTE DES PROBABILITES EST AINSI MODIFIEE

PARATHION	850
DERIVES AMPHETAMINES	105
PLOMB	9
OPIACES	6
DIGITALINE	4
LARGACTIL	4

CETTE DERNIERE LISTE CONSTITUE LES MEILLEURES INDICATIONS QUE NOUS PUISSIONS VOUS FOURNIR

Y A T'IL UNE AUTRE ETUDE A TRAITER ?

-non

NOUS VOUS REMERCIONS DE VOTRE ATTENTION

2) Problème P2 : un problème de segmentation

a) Exposé du problème [4], [16], [25].

Nous considérons une population Ω, et par ailleurs :

- un processus aléatoire T sur Ω, dont les différents états possibles sont $(T_1,\ldots,T_n)$
- une famille D de processus aléatoires q sur Ω ayant chacun au plus a états possibles (a entier ≥ 2).

Nous cherchons à prévoir les comportements des individus de Ω pour la variable T dite "à expliquer", d'après leurs comportements pour une suite de variables de D, dites "explicatives".

Nous cherchons ainsi à réaliser à l'aide des variables explicatives une partition de la population Ω en "segments", dans lesquels les individus aient un comportement le plus homogène possible vis à vis de la variable à expliquer T.

Dans ce cas, on est moins attaché à la précision du résultat que dans le problème P1, mais pour que celui-ci soit utilisable, il faut que le nombre de variables explicatives qui participent à la description d'un segment ne soit pas trop grand.

b) Etude du problème P2

De la même façon qu'à l'occasion du problème P2, une étude statistique permet d'appréhender les probabilités conditionnelles

$$p(q_i|T_j)\ ,\ j \in \{1,\ldots,n\},\ i \in \{1,\ldots,a(q)\},\ q \in D$$

dans la population Ω, ainsi que les fréquences

$$\pi_1,\ldots,\pi_n$$

des états $T_1,\dots,T_n$ dans Ω.

Ainsi, nous sommes en mesure de construire sur D un pseudoquestionnaire K opérant sur π ; ayant fixé ε (généralement bien plus grand que dans P1), l'interprétation d'un sommet x se fera à ε près comme dans P1.

La qualité de l'explication peut se mesurer par ρ_2, probabilité de l'ensemble $E(\pi,\varepsilon)$ des sommets terminaux de K non interprétables à ε près pour π .

Pour que le résultat donné par K soit utilisable, il faut que K ait une hauteur petite. Ainsi, R_o étant un entier fixé par l'utilisateur, nous considérons l'ensemble

$$\mathcal{K}(D,\pi,\varepsilon,R_o)$$

des pseudoquestionnaires construits sur D, opérant sur π, avec interprétation à ε près, et de hauteur au plus égale à R_o.

Le problème est de trouver un élément $\bar{K} \in \mathcal{K}(D,\pi,\varepsilon,R_o)$ qui minimise ρ_2.

Le problème d'optimisation ainsi soulevé est différent de celui qui a été dégagé lors de l'étude de P1.

Ce problème reste ouvert, et nous faisons encore la conjecture que les pseudoquestionnaires K^* décrits lors de l'étude de P1, donnent des résultats approchant l'optimum.

Remarquons qu'il n'y a pas lieu ici, de soulever la question d'existence d'un élément de $\mathcal{K}(D,\pi,\varepsilon,R_o)$ dont le risque ρ_2 soit au plus égal à un seuil η_o fixé à l'avance.

En effet, en raison de la contrainte de rang, la mise en oeuvre de K^* de hauteur R_o (R_o dépasse rarement 5 dans les applications pratiques) est relativement de faible coût. Il suffira donc de vérifier à postériori l'intérêt du résultat.

BIBLIOGRAPHIE

[1] R.B. ASH — *"Information theory"*
John Wiley and Sons, Inc 1965.

[2] J.R. BARRA — *"Polyèdres convexes"*
Polycopié. Faculté des Sciences de Grenoble - 1964

[3] C. BERGE — *"Théorie des graphes et ses applications"*
Dunod - Paris - 1963.

[4] H. BERGONIER, L. BOUCHARENC — *"Une méthode de segmentation basée sur la théorie de l'information".*
Prix "Marcel Dassault" pour la recherche sur les supports de publicité - 1966.

[5] D. CHENAIS, M. TERRENOIRE — *"Aide au diagnostic en toxicologie"*
Journées de l'Informatique au service de l'Homme - Grenoble - Mars 1969.

[6] D. CHENAIS, M. TERRENOIRE — *"Détermination de la réalisation d'un processus aléatoire à l'aide de pseudoquestionnaires".*
Colloquium on Combinatorial Mathematics - Balatonfüred (24-29 Août 1969), organisé par l'Académie des Sciences de Hongrie.

[7] D. CHENAIS, M. TERRENOIRE, M. MATTEI, J. FAURE, M. YACOUB — *"L'Aide au diagnostic en toxicologie".*
Journées internationales d'Informatique Médicale.
Toulouse, 4-5-6 mars 1970.

[8] D.F. FADDEEV — *"Zum Begriff der Entropie eines endlichen Warhrschermlichkeits Schema ; Arbeiten zur Informationstheorie".*
V.E.B. Berlin 1957.

[9] R.M. FANO — *"Transmission of information"*
M.I.T. Press Cambridge, Mass.

[10] S. GUIASU
R. THEODORESCU
"La théorie mathématique de l'information".
Dunod - Paris - 1968.

[11] HALMOS
"Measure Theory".
Van Nostrand - 1965.

[12] D.A. HUFFMAN
"A Method for the construction of minimum redondancy codes".
P.I.R.E. V9 - p. 1098- 1952.

[13] A.I. KHINCHIN
"Mathematical foundations of information Theory".
Dover publications - New-York - 1957.

[14] S. KULLBACK
"Information and Statistics".
J. Wiley.

[15] M. MATTEI (-Mathieu)
"Les ordinateurs à l'aide du diagnostic médical, application en toxicologie".
Thèse de Médecine - Grenoble - 16 décembre 1969.

[16] Ph. T. NGHIEM
Kh. VO KHAC
"Etude sur les aspects théoriques et pratiques de la segmentation aux moindres carrés".
R.I.R.O. n° 8 - V1 - Dunod - 1968.

[17] C.F. PICARD
"Théorie des questionnaires".
Gauthier-Villars - Paris 1966.

[18] C.F. PICARD
"Théorèmes de l'Information traitée".
Colloquium on information Theory - Debrecen - 1967.

[19] C.F. PICARD
"Valeur maximale de l'information traitée".
C.R.A.S. Série A - 265 (1967) p. 624.

[20] C.F. PICARD
"Graphes et questionnaires".
Gauthier-Villars (sous presse).

[21] A. RENYI
"Calcul des probabilités avec un appendice sur la théorie de l'information".
Dunod - Paris.

[22] C.E. SHANNON, W. WEAVER — *"The mathematical theory of communication" (1947). Université of Illinois Press - 1964.*

[23] M. SOBEL — *"Group Testing to classify all defectures in a binomial sample, in information and decision process". Mac Graw Hill - 1960.*

[24] M. SOBEL — *"Optimal group testing". Tech. Report n° 72, Standford University - 1964.*

[25] M. TERRENOIRE — *"Utilisation d'un critère informationnel pour le dépouillement d'une enquête". Congrès d'Informatique Appliquée à l'Economie - Bucarest - Octobre 1969.*

[26] M. TERRENOIRE — *"Une généralisation des questionnaires : les pseudoquestionnaires". C.R.A.S. Série A - 270 (1970) p. 263.*

[27] M. TERRENOIRE — *"Pseudoquestionnaires et Information" C.R.A.S. Série A - (1970).*

[28] P.M. WEICHSEL — *"The Kronecker product of graphs". Pro. Amer. Math. Soc. 13 (1962) p. 47.*

[29] WOLFOWITZ — *"Coding theorems of Information Theory". Springer-Verlag - 2ème Edition - New-York - 1964.*

TABLE DES MATIERES

VU

Grenoble, le

Le Président de la Thèse

VU

Grenoble, le

Le Doyen de la Faculté des Sciences

VU, et permis d'imprimer

Le Recteur de l'Académie de GRENOBLE