COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 16 au 21 JUIN 75

UN ALGORITHME DE CLASSEMENT TRES GENERAL : LES DICHOTOMIES HIERARCHISEES. APPLICATION A L'ETUDE DE SIGNAUX ELECTROPHYSIOLOGIQUES

J. DEMARTINI, H. FOREST, A. VINCENT-CARREFOUR

LABORATOIRE DE SIGNAUX ET SYSTEMES - I.U.T. 95, avenue de Fabron - 06041 NICE CEDEX

RESUME

Lé problème de la reconnaissance de formes présente deux aspects : un aspect classification et un aspect classement.

Le problème du classement d'un ensemble de formes peut être modélisé sous la forme d'un ensemble-quotient défini sur l'ensemble des formes.

L'algorithme des dichotomies hiérarchisées réalise la synthèse d'une relation d'équivalence définissant un ensemble-quotient donné sur un ensemble de formes à partir des éléments de chaque classe de l'ensemble quotient.

Cet algorithme permet la synthèse de la plupart des classeurs utilisés en reconnaissance de formes ; il permet de plus la synthèse de classeurs, non classiques, bien adaptés au problème de la reconnaissance de formes issues de signaux électro-physiologiques.

SUMMARY

The pattern recognition problem involves two aspects:

- taxonomy problem
- classification problem.

Classification of a set of patterns can be modeled by a quotient-set defined on the pattern set.

Algorithm of "Dichotomies hiérarchisées" synthesizes an equivalence-relation approximating a given quotient-set defined on a pattern set, from elements of all the classes of the quotient set.

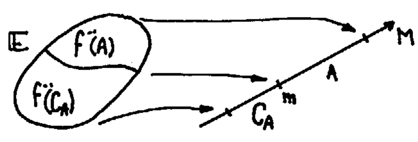
This algorithm permits the synthesis of most of the classifiers used in pattern recognition techniques. It permits also the synthesis of non classic classifiers well adapted to recognition of electrophysiological patterns.

UN ALGORITHME DE CLASSEMENT TRES GENERAL : LES DICHOTOMIES HIERARCHISEES. APPLICATION A L'ETUDE DE SIGNAUX ELECTROPHYSIOLOGIQUES

J. Demartini, H. Forest, A. Vincent-Carrefour

I - POSITION DU PROBLEME DE CLASSEMENT : Le problème de classement est un aspect inhabituel du problème universel de la description d'un phénomène physique à travers un modèle. Il se pose lorsqu'on se propose de trouver un modèle dans le contexte suivant :-le processus physique est appréhendé par un ensemble fini de "mesures" : $E = \{e_1, e_2, e_3, \ldots e_m\}$
- le comportement du processus physique est provoqué par un ensemble fini de "causes" : $C = \{c_1, c_2, c_3, \ldots c_p\}$
Le modèle qui va nous permettre de décrire ce processus sera l'application g qui permet de passer de C à E : $C \xrightarrow{g} E$
g permet d'attribuer à chaque cause possible le ou les effets qui en découleront. En général, la nature est ainsi faite qu'à une cause présumée peut correspondre plusieurs effets ou même qu'un effet peut être engendré par plusieurs causes ; ce paradoxe n'est qu'apparent, il provient essentiellement de deux faits expérimentaux : - décider que telle ou telle mesure est représentative est une présomption de modèle qui n'a a priori aucune raison d'être justifiée, - toute mesure expérimentale est entachée de bruit de fond. Avant de pouvoir dire quoique ce soit, il nous faut commencer par énoncer une première hypothèse simplificatrice : H.S.1 : nous supposerons que, si une même cause peut conduire à plusieurs effets, l'inverse n'est pas vrai. Nous verrons rapidement que cette hypothèse est suffisamment peu restrictive pour être gênante. Rappelons les deux autres hypothèses sur lesquelles va s'appuyer tout ce qui suit : H.S.2 : nous supposerons que l'ensemble des mesures est fini. H.S.3 : nous supposerons que l'ensemble des causes est fini. Considérons alors l'application g^{-1} (les flèches prises à l'envers) ; sous les hypothèses H.S.1, H.S.2 et H.S.3, elle définit une relation d'équivalence sur E : cela signifie que dans E les e_i se regroupent, le point commun de chaque groupe étant la cause qui provoque chaque élément du groupe ; ces groupes sont des classes d'équivalence sur E et l'ensemble de ces classes d'équivalence forment l'ensemble-quotient (E/g^{-1}).Etudier un phénomène physique dans un tel contexte peut se faire sous deux optiques différentes : a) étant donné E, étant donné une forme d'application g, trouver l'ensemble quotient (E/g^{-1}) qui en résulte. La signification de ce point de vue est la suivante : étant donné une population représentée par E et étant donné une mesure de ressemblance représentée par g, on recherche les regroupements cohérents au sens de g qui se produisent dans la population considérée. Un algorithme classique existe pour résoudre ce problème : l'algorithme des "nuées dynamiques" de Diday. On a affaire, dans ce cas, à un problème dit de "classification" ou de "taxonomie". b) étant donné E, étant donné l'ensemble quotient C, trouver une forme d'application g^{-1} permettant de passer de E à C. La signification de cette optique d'étude est la suivante : étant donné une population représentée par E et dans laquelle on sait devoir distinguer des classes, on cherche une application "g^{-1}" permettant de déterminer à quelle classe appartient un élément quelconque de la population considérée. On réalise alors un "classeur".

II - ALGORITHME DES DICHOTOMIES HIERARCHISEES APPLIQUE AU PROBLEME DU CLASSEMENT. Nous exposerons, sans rien perdre de la généralité de l'algorithme, le cas où on est en présence de deux classes. II-1 Contexte de l'algorithme : le problème de classement peut être posé de la façon suivante : a) on considère un ensemble quelconque $\mathbb{E}$ (ensemble universel du problème), b) on considère deux sous-ensembles E_1 et E_2 de $\mathbb{E}$, non nécessairement disjoints et dont la réunion est l'ensemble des mesures E ; E_1 et E_2 sont finis. Pour synthétiser le classeur, c'est-à-dire l'application g^{-1}, on se donne les "outils" suivants : a) un ensemble totalement ordonné M dont la relation d'ordre est notée $\geqslant$ b) un ensemble F d'applications f de $\mathbb{E}$ dans M. L'image de $x \in \mathbb{E}$ par $f \in F$ est notée f(x). II-2 Propriétés des outils de travail introduits: a) doublet séparateur : on appelle doublet séparateur tout élément(f,m) de l'ensemble F x M qui, à tout élément f de F et à tout élément m de M fait correspondre la partition de $\mathbb{E}$ en deux classes d'équivalence définies par :
$f^{-1}(A)$ et $f^{-1}(C_A)$ avec $A = \{ a/a \in M,\ a \geqslant m\}$



b) Séparatrice dans $\mathbb{E}$: étant donné le doublet séparateur (f,m), on appellera séparatrice dans $\mathbb{E}$ associée au doublet séparateur (f,m) le sous-ensemble $f^{-1}(m)$ de $\mathbb{E}$. Nous avons dès lors le moyen de "fabriquer" dans $\mathbb{E}$, donc éventuellement dans E deux classes d'équivalence.

L'algorithme des dichotomies hiérarchisées va permettre, en agençant des doublets séparateurs, de reconstituer les classes d'équivalence cherchées. REMARQUE : Etant donné les outils que nous nous sommes fabriqués (générateurs de classe d'équivalence), il est important de noter, dès à présent, que si E_1 et E_2 ne sont pas disjoints, ils ne peuvent pas constituer deux classes d'équivalence. Il faut alors reposer le pro-

UN ALGORITHME DE CLASSEMENT TRES GENERAL : LES DICHOTOMIES HIERARCHISEES
APPLICATION A L'ETUDE DE SIGNAUX ELECTROPHYSIOLOGIQUES

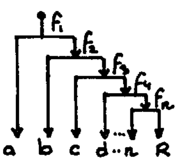
J. Demartini, H. Forest, A. Vincent-Carrefour

blème dans les termes suivants : étant donné E et trois classes d'équivalence E'_1, E'_2 et $E'_3 = E_1 \cap E_2$ trouver "g^{-1}". On peut ainsi, sans perte de généralité, supposer que $E_1 \cap E_2 = \emptyset$. Nous allons maintenant montrer qu'un tel agencement est possible :

PROPOSITION : Sous certaines conditions, il existe une suite finie (f_1, m_1) ; (f_2, m_2)...; (f_n, m_n) de doublets séparateurs permettant de reconstituer les classes d'équivalence données E_1 et E_2.

II-3 Reconstitution des classes d'équivalence E_1 et E_2 : Supposons que l'on sache trouver un doublet séparateur engendrant deux classes d'équivalence dont l'une est tout ou partie de E_1 ou E_2 : par exemple de E_1. Ce doublet séparateur a alors permis d'obtenir un morceau de E_1. Si on enlève ce morceau de E_1 de E, sur l'ensemble résultant E', on saura trouver un deuxième doublet-séparateur ayant les mêmes propriétés que le premier ; il engendrera, par exemple, un morceau de la classe E_2. Ainsi, au fur et à mesure qu'on ajoute des doublets-séparateurs, on reconstitue par morceaux les classes d'équivalence E_1 et E_2. Ces doublets sont hiérarchisés car chacun ne peut opérer que dans l'ordre où il a été obtenu. Cette reconstitution s'arrête dans deux cas :

- ce qui reste de l'ensemble E initial, après avoir enlevé successivement tous les morceaux de classe, est vide : la reconstitution est complète,
- on n'a plus d'applications f pour fabriquer un nouveau doublet (on n'en avait pas prévu assez dans F) et ce qui subsiste de E est le résidu de classement. Le fonctionnement du classeur ainsi constitué peut être représenté par le schéma suivant :



avec, par exemple :

si $R = \emptyset$ $\begin{cases} E_1 = (a \cup c \cup \ldots) \\ E_2 = (b \cup \ldots \cup n) \end{cases}$

si $R \neq \emptyset$ (résidu) $\begin{cases} (a \cup c \cup \ldots) \subset E_1 \\ (b \cup \ldots \cup n) \subset E_2 \end{cases}$

Voyons maintenant comment trouver un doublet-séparateur permettant cet agencement.

II-4 Obtention d'un doublet-séparateur "convenable"

La seule contrainte à laquelle doit satisfaire ce doublet est la suivante : étant donné f, une des deux classes engendrée par ce doublet doit être tout ou partie de E_1 ou E_2. Cette classe d'équivalence va être définie par un représentant x^* et par m qui, avec f, définit la séparatrice.

a) Choix de x^* : prenons $x^* = \underset{x \in E}{\mathrm{Inf}}\ f(x) = m_f$

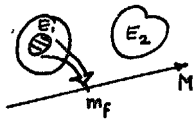
m_f existe toujours puisque M est ordonné.

b) Choix de m : x^* étant choisi, deux cas de figure peuvent se produire :

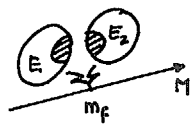
1 - $f^{-1}(m_f) \subset E_1$ ou E_2 (par exemple E_1)

2 - $f^{-1}(m_f) \subset E_1 \cup E_2$ et ($\not\subset E_1$ ou $\not\subset E_2$)

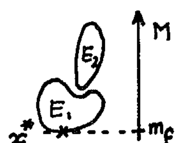
1er cas : on peut trouver un sous-ensemble non vide $E'_1 \subset E_1$ tel que : $\forall x \in E'_1$ et $\forall y \in E_2$ $f(x)<f(y)$ (E'_1 contiendra au moins x^*)



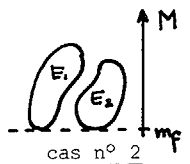
2ème cas : on ne peut pas trouver un sous-ensemble E'_1 et dans ce cas, le doublet séparateur n'existe pas.



Exemple : E = {points de $\mathbb{R}^2$}
f : projection sur un axe



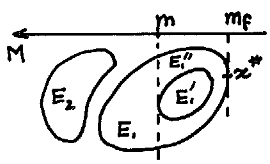
cas n° 1



cas n° 2

Ayant trouvé E'_1, on lui associe m tel que $m = \underset{x \in E'_1}{\mathrm{Sup}}\ f(x)$

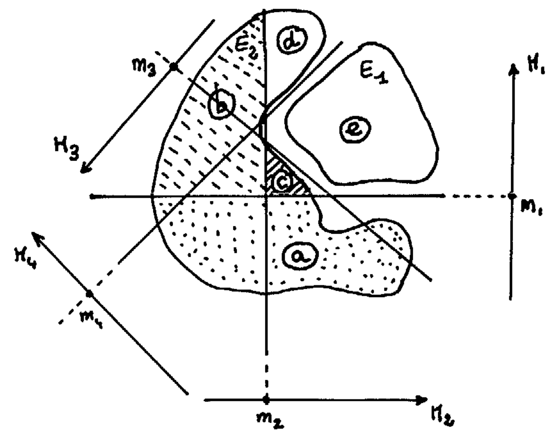
Le doublet-séparateur (f,m) engendre deux classes d'équivalence : $E''_1 = \{x/f(x) \leq m\}$ $R_1 = \{x/f(x) > m\}$ avec $E'_1 \subseteq E''_1 \subseteq E_1$. En reprenant l'exemple ci-dessus, on obtient la disposition suivante :



finalement, on obtiendra un doublet-séparateur "convenable" correspondant à une application f, pourvu que $f^{-1}(m_f)$ soit entièrement contenu dans E_1 ou dans E_2.

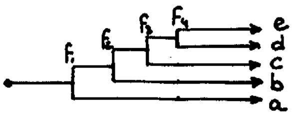
II-5 Quelques exemples d'application de l'algorithme des dichotomies hiérarchisées au problème du classement. Exemple n° 1 : $\mathbb{E} = \mathbb{R}^2$;

E = {points de R^2} et on dispose des quatre applications suivantes :



f_1 : projection sur un axe ↑
f_2 : projection sur un axe →
f_3 : projection sur un axe ↙
f_4 : projection sur un axe ↖

On obtient ainsi la suite $(f_1,m_1)(f_2,m_2)(f_3,m_3)(f_4,m_4)$ et le classeur ainsi réalisé peut être représenté par le schéma :



$E_1 = e$

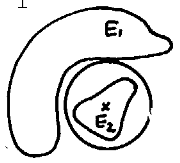
$E_2 = (a \cup b \cup c \cup d)$

Si un seul doublet-séparateur avait suffit, on aurait obtenu le classeur linéaire traditionnel ; nous avons ici obtenu un classeur linéaire par morceaux.

Exemple n° 2 : $\mathbb{E} = R^2$; $E = \{\text{points de } \mathbb{R}^2\}$ on dispose des applications suivantes :

$f_w(x) = \sqrt{(x-w)^2}$ w étant un élément de E et x étant un autre élément de E. Prenons w, arbitrairement dans E. $x^*_1 = w_1$ car $f_{w_1}(w_1) = 0$ et prenons m, tel que $\forall\, x \in E'_2$

$f_{w_1}(x) \leq m_1$ et $\forall\, x \in E_1$ $f_{w_1}(x) > m_1$. Dans ce cas, cela revient à tracer un cercle centré sur w_1 et englobant le plus possible de E_2 (ici la totalité) sans toucher E_1.



On a obtenu un classeur à séparatrice circulaire ; c'est une des formes du classeur quadratique traditionnel.

Exemple n° 3 = $\mathbb{E} = \mathbb{B}^n$ (espace booléen de dimension n) $E = \{\text{points de } \mathbb{B}^n\}$; on utilise par l'exemple les applications suivantes :

$f_w(x) = \Sigma\, x \oplus w$ (distance de Harmming)

w et x étant des éléments de E. On obtiendrait un classeur particulièrement adapté au traitement d'images du type "bélino".

REMARQUE IMPORTANTE : Si les applications choisies à l'exemple n° 1 peuvent conduire à des cas où le doublet-séparateur n'existe pas, celles choisies à l'exemple n° 2 et n° 3 conduisent toujours à un doublet-séparateur.

II-6 Traitement des fonctions continues : signal ECG

Considérons à titre d'exemple le phénomène simulé suivant (cf. figures en annexe) : l'état cardiaque d'un patient peut être de deux types : a) normal, b) avec un défaut de conduction. Ce phénomène est mis en évidence sur l'électrocardiogramme (dérivation D II par exemple). On constate expérimentalement que l'ECG d'un patient dans l'état a) conduit aux enregistrements regroupés sous le nom de "classe 1" et que l'ECG d'un patient dans l'état b) conduit aux enregistrements regroupés sous le nom de "classe 2".

Dans ce contexte : - l'ensemble $\mathbb{E}$ est l'ensemble des fonctions continues, définies sur un intervalle fini (ici, une période cardiaque) ; - l'ensemble E est l'ensemble des enregistrements obtenus expérimentalement ; les deux classes d'équivalence sont la classe 1 et la classe 2 : c'est la séquence-test du problème.

- les applications f choisies (arbitrairement) pour construire le classeur sont les suivantes :

$f_w(x) = \Sigma\, |x - w|$ (x et w étant deux éléments de E) évaluée aux points d'échantillonnage.

- l'ensemble M est la droite réelle positive.

L'algorithme des dichotomies hiérarchisées conduit à un classeur à trois doublets séparateurs donnant un classement complet sur la séquence-test.

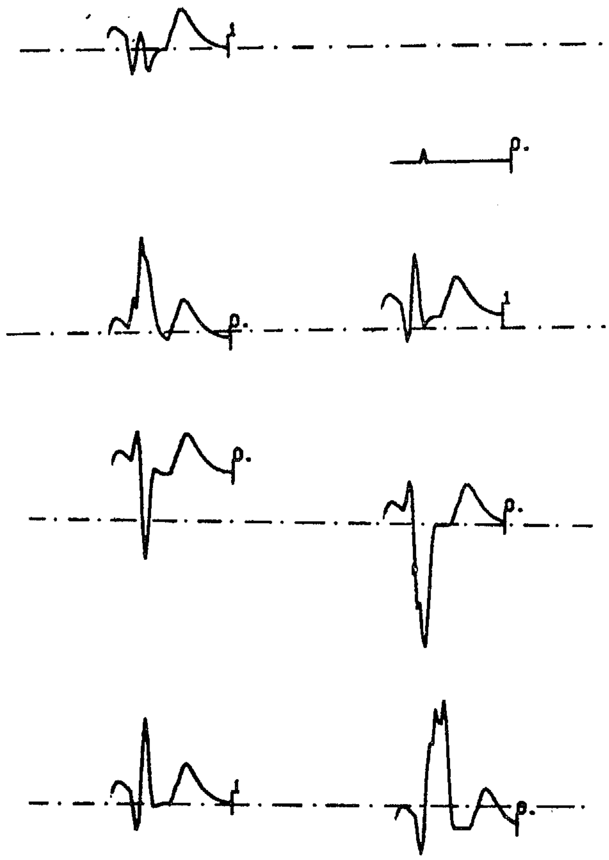
REMARQUE : On constate, en considérant la séquence d'essai, que, hors de la séquence-test, ce classeur peut conduire à des erreurs de classement. C'est tout à fait normal si on se souvient qu'un modèle de processus n'est jamais extrapolable sans risques au-delà du domaine dans lequel il a été défini.

III - CONCLUSION : L'exemple final de classement, bien que simpliste, montre que l'analyse d'un processus par classement peut conduire à des résultats très intéressants. Le classement permet d'obtenir des modèles de processus bien adaptés à la prise de décision en contexte empirique. L'algorithme des dichotomies hiérarchisées permet, dans ce contexte, d'obtenir pratiquement à tout coup un classeur satisfaisant dans la mesure où la connaissance expérimentale qu'on a du processus est suffisante (séquence test convenable).

ANNEXE : SEQUENCE-TEST

Classe 1

Classe 2



SEQUENCE D'ESSAI

Classeur à 3 unités de base

1 : classe 1

2 : classe 2

0 : forme inconnue