

50376
1983
305
N° d'ordre : 318

50376
1983
305

THÈSE

présentée à

L'UNIVERSITE DES SCIENCES ET TECHNIQUES DE LILLE 1

pour l'obtention du titre de

DOCTEUR-INGENIEUR

par

Brigitte CHEBEL

Ingénieur E.N.P.A.

APPLICATION DE L'ANALYSE DE STRUCTURE ET DE LA FONCTION ENTROPIE AU TRAITEMENT DU SIGNAL



030 07821 5 6

Soutenue le 22 mars 1983 devant la Commission d'Examen

Membres du Jury :

MM.

P.

VIDAL

Président

C.

VASSEUR

Rapporteur

B.

DUBUISSON

Examineur

Y.

MOSCHETTO

Examineur

J.M.

TOULOTTE

Examineur

Travail financé par le Conseil Régional Nord/Pas-de-Calais
et par l'INSERM (CRL N° 805014).

AVANT - PROPOS

Le travail présenté dans ce mémoire a été effectué au laboratoire d'automatique de l'Université des Sciences et Techniques de Lille I et au centre de réanimation médicale de l'Hôpital Albert Calmette de Lille.

J'exprime toute ma gratitude à Monsieur le Professeur P. VIDAL pour sa sollicitude et la gentillesse avec laquelle il m'a accueillie dans son laboratoire. Je le remercie d'avoir bien voulu accepter la présidence de mon jury de thèse.

Que Monsieur J.M. TOULOTTE, professeur à l'Université des Sciences et Techniques de Lille I, trouve ici l'expression de ma gratitude pour avoir accepté de juger ce travail.

Je tiens à remercier Monsieur C. CHOPIN, Professeur du C.H.R. de Lille, pour les moyens et le temps qu'il a mis à ma disposition pour la réalisation de ce travail.

J'adresse mes vifs remerciements à Monsieur B. DUBUISSON, Professeur à l'Université de Compiègne, pour l'honneur qu'il me fait en voulant bien juger ce travail.

Je remercie également Monsieur Y. MOSCHETTO, Directeur à l'INSERM à Lille, de l'intérêt qu'il a bien voulu porter à ces travaux en acceptant de participer au jury.

Je suis profondément reconnaissante à Monsieur C. VASSEUR, Maître Assistant à l'Université des Sciences et Techniques de Lille I, pour m'avoir guidé avec tant de compétence et de clairvoyance, pour ses critiques constructives, et le temps qu'il n'hésitait pas à consacrer à la réalisation de ce travail.

Un grand merci également à Madame M. LELONG, à Madame A. PIGNON et à Monsieur J. HOUZE, pour leur contribution amicale et généreuse à l'élaboration matérielle de ce mémoire.

INTRODUCTION

De toutes les sciences de l'ingénieur, l'automatique et l'informatique sont probablement celles qui ont le plus contribué ces dernières années au prodigieux développement des méthodes d'analyse et de traitement des données utilisées dans le cadre du génie biologique et médical. /1/ /2/ /3/ /4/ /7/

Ainsi les appareils informatisés ou microinformatisés les plus sophistiqués permettent l'aide au diagnostic et sont en mesure de proposer différents types de solution avec telle ou telle probabilité selon tel ou tel symptôme / 6/. Par contre l'interprétation finale des résultats reste bien évidemment de la compétence du médecin.

De la même façon l'analyse de signaux comme par exemple l'EEG (électroencéphalogramme) qui a pour but de caractériser l'évolution de l'activité du cerveau dans le temps / 5/ conduit à des traitements très élaborés. Toutefois ces traitements sont effectués en temps différé comme dans le cas d'appareils d'aide au diagnostic et nécessitent un ordinateur de capacité mémoire assez importante pour le traitement des bases de données que constitue l'évolution du signal dans le temps.

L'objet du travail que nous présentons est de décrire les méthodes à appliquer lors de la réalisation d'un appareil destiné à surveiller en temps réel l'évolution de l'état d'un malade placé dans un service de réanimation.

Cette surveillance qui doit être intensive, nécessite un traitement complexe. Dans ce cadre une méthodologie générale sur l'analyse de signaux prélevés simultanément sur un même système a été mise au point.

Le développement s'appuie sur l'exemple de signaux délivrés par le corps humain.

L'étude se décompose alors en quatre chapitres :

.../...

chapitre 1 - Analyse et reconnaissance syntaxique du signal

chapitre 2 - Analyse informationnelle des relations intersignaux

chapitre 3 - Analyse prédictive de l'évolution temporelle

chapitre 4 - Application en réanimation respiratoire

A - CHAPITRE 1

ANALYSE ET RECONNAISSANCE SYNTAXIQUE DU SIGNAL

Etudier le comportement d'un système revient à analyser les variables ou les signaux caractérisant son activité.

Pour cela une première sélection de signaux parmi tous ceux disponibles s'avère indispensable afin d'optimiser le temps de traitement et l'interprétation. Cette première sélection s'appuie d'une part sur des impératifs technologiques et d'autre part sur l'expérience d'un spécialiste du système étudié (ici le médecin).

L'analyse des signaux présélectionnés pris un par un fait alors l'objet du 1er chapitre.

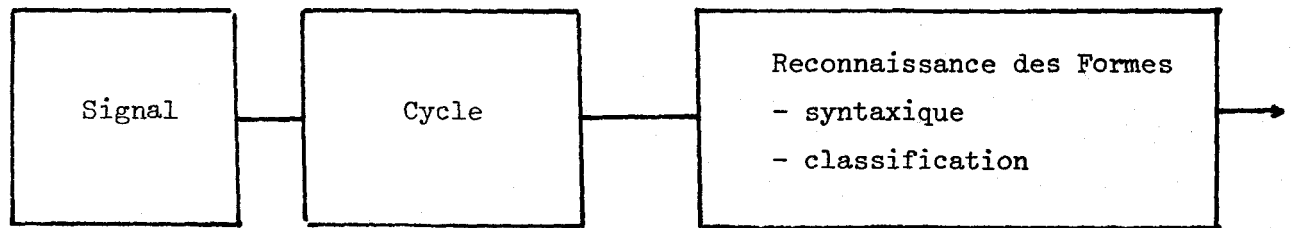
Toutefois cette analyse ne s'appuie pas sur les méthodes fréquentielles conventionnelles (FFT, transformée de Walsh...), mais sur une description de la structure des signaux.

Plus précisément il est apparu dans notre application que la majorité des signaux pouvait se décomposer en cycles successifs. Nous avons donc étudié ces cycles et les différentes formes possibles qu'ils sont susceptibles de prendre.

Dans ce cadre l'utilisation d'outils syntaxiques associés à des critères de classification est particulièrement bien adaptée à une description permettant de différencier les diverses formes prises par les cycles.

Cette première démarche se schématise comme suit.

.../...



Ainsi l'évolution dans le temps, de chaque signal se traduit par une suite de classes caractérisant la forme des cycles extraits.

B - CHAPITRE 2 ANALYSE INFORMATIONNELLE DES RELATIONS INTERSIGNAUX

Le but de ce chapitre est d'étudier les relations éventuelles liant les suites de classes définies au chapitre 1.

Pour cela il est impossible d'utiliser les méthodes conventionnelles telles que les fonctions de corrélation. Par contre le calcul d'entropie proposé par la théorie de l'information apparait comme un outil intéressant de mise en évidence de ces relations. /34/

L'approche adoptée dans le cadre de cette étude utilise la fonction d'entropie en l'adaptant à ces suites temporelles.

C - CHAPITRE 3 ANALYSE PREDICTIVE DE L'EVOLUTION TEMPORELLE

L'objet de ce 3ème chapitre est d'effectuer une observation des signaux sur une période longue afin de prédire les tendances d'évolution.

.../...

Pour cela on définit des paramètres statistiques permettant :

- de réaliser l'étude macroscopique du signal.
- d'envisager la prédiction d'une tendance sur une série de cycles donnée

Plus précisément une représentation géométrique de ces outils statistiques permet de ramener l'étude prédictive à l'étude de trajectoires de tendance.

D - CHAPITRE 4 APPLICATION EN REANIMATION RESPIRATOIRE

Finalement la méthodologie décrite dans les trois premiers chapitres est appliquée à des signaux électrophysiologiques prélevés sur des malades transitant dans un service de réanimation respiratoire. En particulier on s'y attache à définir les traitements permettant d'exploiter au mieux les informations ne nécessitant pas de prélèvements invasifs.

- C H A P I T R E I -

ANALYSE ET RECONNAISSANCE SYNTAXIQUE DU SIGNAL

CHAPITRE I

I.1 - INTRODUCTION	page 1
I.2 - SUIVI D'UNE COURBE	page 2
I.2.1. - Introduction	page 2
I.2.2. - Structure d'une forme	page 3
I.2.3. - Les différentes primitives	page 3
I.2.4. - L'extraction du cycle	page 5
I.3 - OUTILS DE LA MISE EN OEUVRE	page 5
I.3.1. - Extraction de primitives	
I.3.3.1. - Cas de pentes positives	page 6
I.3.3.2. - Cas de pentes négatives	page 7
I.3.2. - Les paramètres descriptifs	page 9
I.4 - DESCRIPTION D'UN CYCLE	page 10
I.4.1. - L'approche grammaticale	page 10
I.4.1.1. - Rappel sur les grammaires	page 10
I.4.1.2. - Application	page 11
I.4.2. - Approche paramétrique	page 13
I.5 - CLASSIFICATION	page 13
I.5.1. - Cas d'un cahier des charges existant	page 13
I.5.2. - Cas de cahier des charges inexistant	page 15
I.6 - TRAITEMENT TEMPS REEL	page 17
I.6.1. - But	page 17
I.6.2. - Rappel sur le temps réel	page 18
I.6.3. - Application	page 18
I.7 - CONCLUSION	page 24

I.1 - INTRODUCTION

L'analyse du comportement d'un système peut être réalisée en suivant l'évolution d'un ensemble de signaux caractérisant l'état de celui-ci.

En effet dans de nombreux cas toute modification du comportement du système se traduit par une affectation au niveau de la forme des signaux émis. Dans ces conditions il apparaît que l'étude de la structure du signal constitue un outil intéressant pour caractériser l'état d'un système. /12/

Dès lors l'extraction des différentes formes prises par un signal, puis la caractérisation de ces formes sont à la base du suivi dans le temps de l'évolution du système. L'objet de ce chapitre est alors d'étudier le signal sous son aspect structurel et syntaxique. De plus nous envisageons un suivi en temps réel afin de pouvoir corriger suffisamment rapidement une évolution indésirable du système.

La méthodologie présentée comporte quatre étapes :

- la décomposition du signal en formes successives
- l'extraction des primitives de syntaxe et des paramètres caractérisant chaque forme
- la reconnaissance du mot syntaxique caractérisant la forme
- l'affinement de cette reconnaissance par une classification utilisant les paramètres extraits.

Il importe de noter que les deux dernières étapes décrites réalisent une procédure de reconnaissance de formes à deux niveaux.

- . le premier niveau qui est syntaxique, permet de faire une préreconnaissance de classes
- . le second niveau qui est métrique, permet de réaliser une partition des classes afin d'aboutir à une reconnaissance définitive plus fine.

.../...

I.2 - SUIVI D'UNE COURBE

I.2.1 - INTRODUCTION

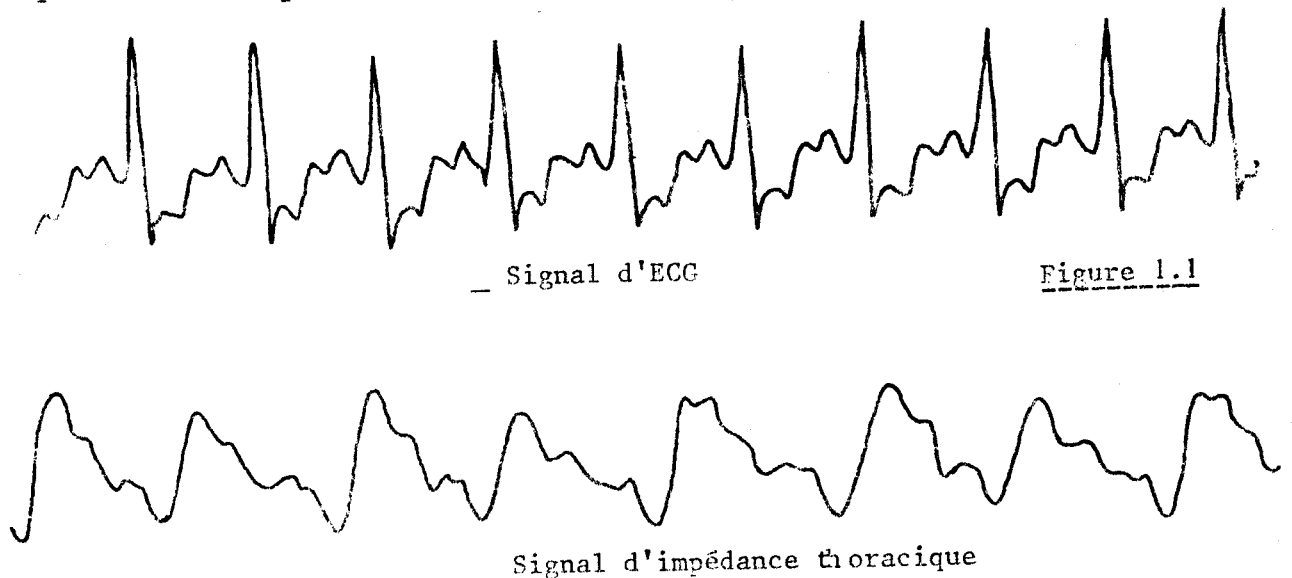
Afin d'assurer le suivi du signal, on segmente celui-ci en intervalles de temps successifs, définis selon un critère spécifique de sa nature.

Par exemple, le découpage d'un signal cyclique est immédiat. Chaque segment correspond en effet à un cycle qui caractérise le système.

L'extraction du cycle dépend des critères de début et de fin de cycle qui sont soit imposés par le signal, soit déterminés par l'utilisateur lui-même.

La fin d'un cycle correspond le plus souvent au début du cycle suivant de telle sorte que l'extraction du cycle se ramène au test d'un seul critère.

Par contre la segmentation d'un signal non cyclique dépend de la forme de celui-ci et du cahier des charges imposé par l'utilisateur. Nous envisageons dans ce mémoire essentiellement l'étude des signaux cycliques dont la figure 1.1 illustre deux représentations tirées du domaine biomédical : l'impédance thoracique* et l'ECG (électrocardiogramme). /6/ /26/.



* L'impédance thoracique est un signal électrophysiologique, prélevé à l'aide d'électrodes sur le thorax humain et reflétant l'activité mécanique pulmonaire du patient surveillé /28/.

I.2.2 - STRUCTURE D'UNE FORME

La détermination des paramètres caractéristiques pour chaque cycle est obtenue en associant au signal une enveloppe formée de segments de droite séparés par des points de cassures ainsi que l'illustre la figure 1.2 .

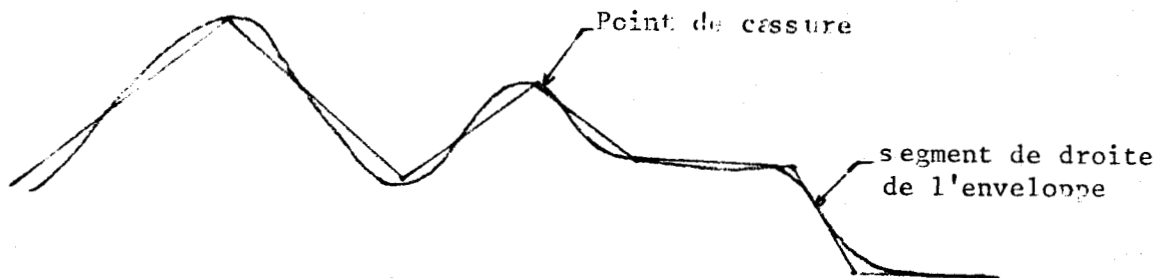


Figure 1.2

Les particularités du signal se traduisent par des changements de direction qui se font progressivement ou brutalement. Chacun des segments de l'enveloppe est obtenu en évaluant les pentes moyennes successives prises par le signal au cours de son évolution dans le temps.

Ces segments de droite sont entièrement définis par leurs pentes et les positions des points de cassures. Ceci conduit à une mise en oeuvre particulièrement simple et rapidement exploitable pour une implantation sur microprocesseur.

Ce type de description dans la mesure où il réalise une schématisation du signal, engendre une certaine perte d'information. Toutefois nous verrons qu'il conduit à la mise en oeuvre d'une grammaire simple et non spécifique permettant de reconnaître facilement l'allure générale d'un cycle.

Par ailleurs, l'affinement de cette reconnaissance par une classification utilisant les paramètres du signal permettra de compenser les pertes d'informations consenties au niveau de l'approche syntaxique.

I.2.3 - LES DIFFERENTES PRIMITIVES

Ainsi que l'illustre la figure 1.3, on peut recenser six types de points de cassures sur un signal quelconque.

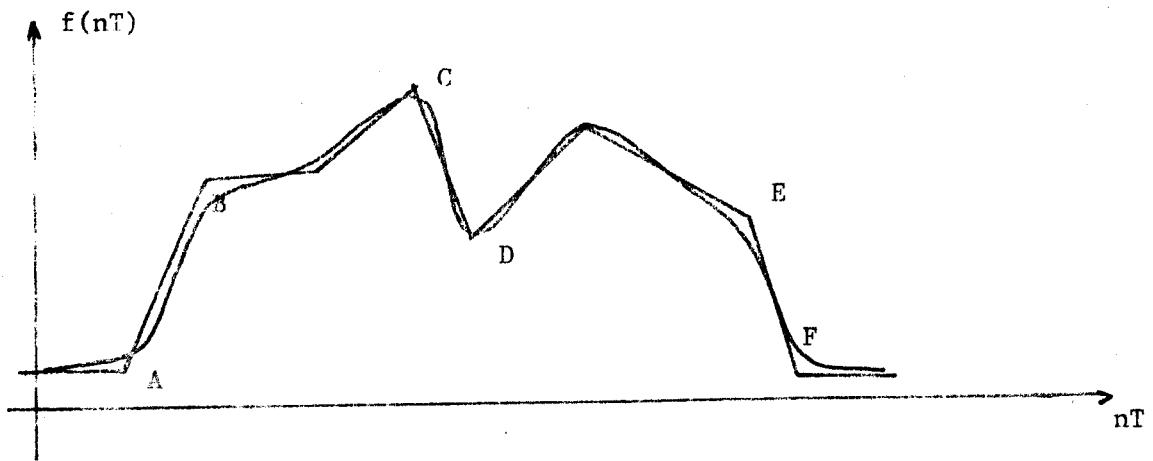


Figure 1.3

- A passage d'une pente faiblement positive à une pente fortement positive
- B passage d'une pente fortement positive à une pente faiblement positive
- C passage d'une pente positive à une pente négative
- D passage d'une pente négative à une pente positive
- E passage d'une pente faiblement négative à une pente fortement négative
- F passage d'une pente fortement négative à une pente faiblement négative

Figure 1.3

En ce qui concerne la forme, elle est donc essentiellement décrite par la suite des points de cassures ainsi définis.

Par contre si l'on veut différencier deux formes ayant une même allure, on peut affiner l'analyse en utilisant les points de cassures et les valeurs des pentes entre les points de cassures.

En conclusion, l'utilisation d'une grammaire permet dans un premier temps de déterminer l'allure général du cycle, donc de le classer dans une catégorie de forme prédéterminée. Dans un second temps l'affinement de cette classification est obtenue en tenant compte des éléments quantitatifs de la description, la grammaire permettant l'approche qualitative du problème.

I.2.4 - L'EXTRACTION DU CYCLE

L'extraction du cycle se déduit du signal traité et des différentes allures qu'il peut prendre.

Elle dépend des critères de début et de fin de cycle, le début d'un cycle pouvant coïncider dans de nombreux cas, avec la fin de son prédécesseur.

Considérons l'exemple de la figure 1.4.

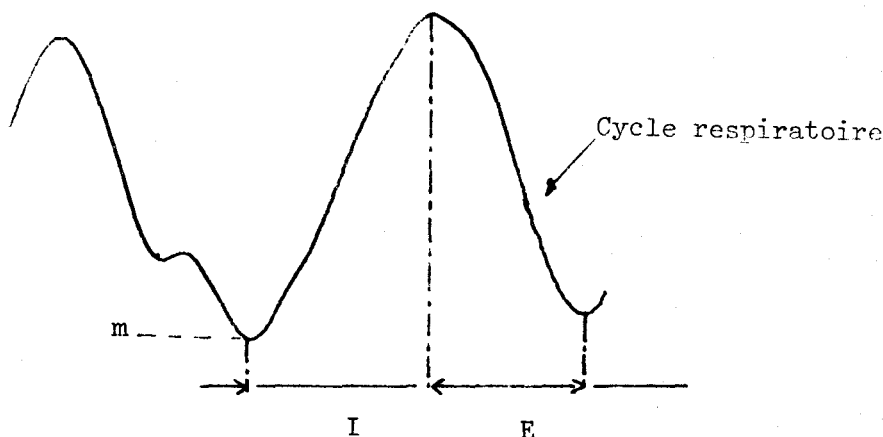


Figure 1.4

Il représente un signal d'impédance thoracique, pour lequel chaque cycle est l'image d'un cycle respiratoire déterminé par une phase inspiratoire I suivie d'une phase expiratoire E. Le début de chaque cycle coïncide avec un minimum négatif m. /26/ /29/

Le problème de l'isolation du cycle est de rechercher le critère de début. Toutefois ce critère peut changer dans certains cas pathologiques, de telle sorte qu'il est nécessaire de recenser les différents "cas de figure" et de les tester lors d'une extraction de cycle.

I.3 - OUTILS DE LA MISE EN ŒUVRE

Soit un signal échantillonné à la période T. Afin de caractériser les différents types de points de cassures A, B, C, D, E et F, nous en donnons

ci-après les définitions formelles où nT ($n \in \mathbb{N}$) désigne un instant d'échantillonnage quelconque et $\frac{d}{dt} f(nT)$ la dérivée du signal échantillonné à l'instant nT .

3.1 - EXTRACTION DE PRIMITIVES

3.1.1. - Cas de pente positive

. Primitive A (figure 1.5a)

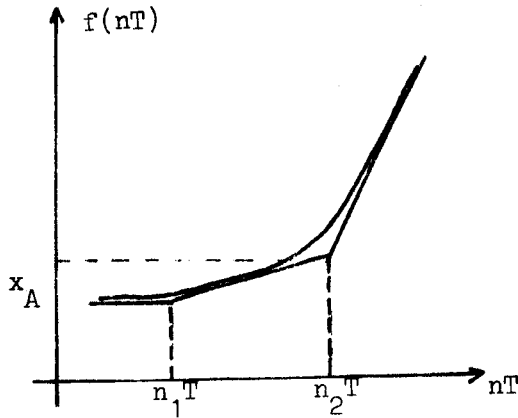


Figure 1.5a

Soit a_1 et b_1 2 valeurs données telles que

$$0 < a_1 < 1$$

$$0 < b_1 < 1$$

Soit par ailleurs une suite d'instantanés d'échantillonnage successifs nT ou

$$\frac{d}{dt} f(nT) \geq 0.$$

Alors n_1T est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} f[(n_1 - 1)T] < a_1 \frac{d}{dt} f(n_1T)$$

Soit enfin la suite des instantanés d'échantillonnage nT successifs suivant n_1T ($n > n_1$) alors n_2T est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} f[(n_2 - 1)T] < b_1 \frac{d}{dt} f(n_2T)$$

Dans ces conditions on choisit comme point de cassure relatif à la primitive A : $x_A = f(n_2T)$.

Primitive B (figure 1.5b)

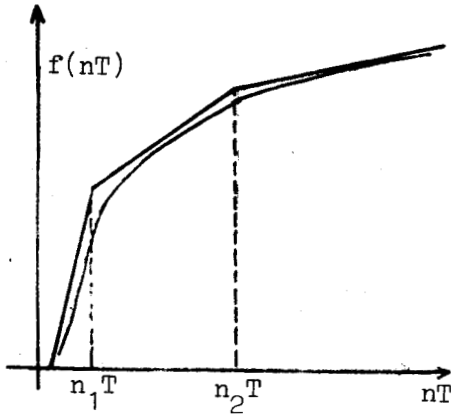


Figure 1.5b

Soit a_2 et b_2 2 valeurs données telles que

$$a_2 \text{ et } b_2 > 1$$

Soit par ailleurs une suite d'instantanés d'échantillonnage nT où $\frac{d}{dt} f(nT) > 0$

alors

n_1T est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} f[(n_1 - 1)T] > a_2 \frac{d}{dt} f(n_1T)$$

Soit enfin la suite des instantanés d'échantillonnage nT successifs

$$n_1T \quad (n > n_1)$$

alors

n_2T est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} f(n_2 - 1)T > b_2 \frac{d}{dt} f(n_2T)$$

Dans ces conditions on choisit comme point de cassure relatif à la primitive B

$$x_B = f(n_2T).$$

3.1.2. - Cas de pentes négatives

. Primitive E (figure 1.5c)

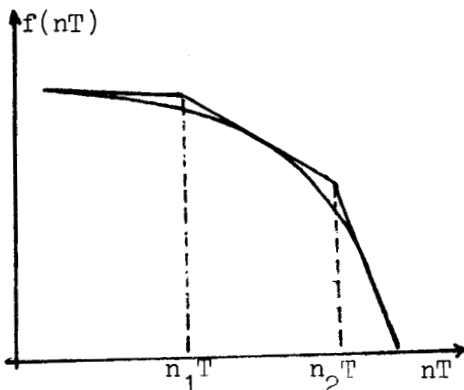


Figure 1.5c

Soit a_3 et b_3 2 valeurs données telles que

$$0 < a_3 < 1$$

$$0 < b_3 < 1$$

Soit par ailleurs une suite d'instantanés d'échantillonnage nT où $\frac{d}{dt} f(nT) \leq 0$

alors

n_1T est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} [f(n_1 - 1) T] > a_3 \frac{d}{dt} f(nT)$$

Soit enfin la suite des instants d'échantillonnage nT successifs suivant $n_1 T$ ($n > n_1$)

alors

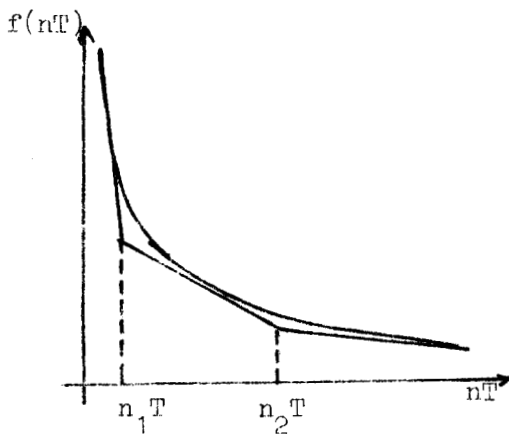
$n_2 T$ est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} f[(n_2 - 1) T] > b_3 \frac{d}{dt} f(n_2 T)$$

Dans ces conditions, on choisit comme point de cassure relatif à la primitive E

$$x_E = f(n_2 T)$$

. Primitive F (figure 1.5d)



Soit a_4 et b_4 2 valeurs données telles que

$$a_4 \text{ et } b_4 > 1$$

Soit par ailleurs une suite d'instants d'échantillonnage nT où $\frac{d}{dt} f(nT) \leq 0$ alors

$n_1 T$ est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} [f(n_1 - 1) T] < a_4 \frac{d}{dt} (f(nT)).$$

Soit enfin la suite des instants d'échantillonnages nT successifs suivant $n_1 T$ ($n > n_1$)

alors

$n_2 T$ est le premier instant d'échantillonnage de cette suite où

$$\frac{d}{dt} (f(n_2 - 1) T) < b_4 \frac{d}{dt} (f(n_2 T))$$

Dans ces conditions, on choisit comme point de cassure relatif à la primitive F

$$x_F = f(n_2 T)$$

3.1.3. - Les extrémums

1.3.1. - Minimum D

$$x_D / x_D = f(nT) \quad \frac{d}{dt} f(nT) \leq 0 \text{ et } \frac{d}{dt} f(n+1)T > 0$$

1.3.2. - Maximum C

$$x_C / x_C = f(nT) \quad \frac{d}{dt} f(nT) \geq 0 \text{ et } \frac{d}{dt} f(n+1)T < 0$$

I.3.2 - LES PARAMETRES DESCRIPTIFS

L'extraction des primitives décrites au paragraphe précédent, nécessite des outils mathématiques très simples à mettre en oeuvre tels que les seuils, le comptage du temps, le calcul de pente, les mesures d'amplitudes du signal.

. amplitude

Elle permet de différencier 2 points caractéristiques de même nature.

. dérivée

Elle permet la reconnaissance proprement dite des points caractéristiques par calcul des pentes du signal.

. seuils

Ils sont déterminés expérimentalement (apprentissage) et permettent la détection des points de cassure.

. compteurs

Ils permettent de mesurer les intervalles de temps par comptage des périodes d'échantillonnage.

La simplicité de mise en oeuvre permet une implantation facile sur microprocesseur.

I.4 - DESCRIPTION D'UN CYCLE

Après avoir extrait un cycle, nous envisageons d'en décrire la structure selon deux approches complémentaires.

. une approche qualitative

qui décrit la forme par un mot syntaxique construit à l'aide d'une grammaire dont le langage de base est l'ensemble des primitives décrites précédemment. Une analyse syntaxique élémentaire permet ainsi de donner l'allure générale du cycle.

. une approche quantitative

qui permet de définir plus précisément l'allure de l'enveloppe. Cette approche vise à caractériser plus finement la courbe étudiée en donnant les valeurs des pentes et les amplitudes des points de cassures de l'enveloppe du cycle. Ce sont ces paramètres associés à l'analyse syntaxique qui permettent d'achever la classification des formes recueillies.

I.4.1 - L'APPROCHE GRAMMATICALE

4.1.1. - Rappel sur les grammaires

En reconnaissance des formes, l'utilisation des grammaires est de plus en plus répandue. Une reconnaissance par petits morceaux de la forme, décomposée en primitives est préférée à une analyse globale./9/ /7/

La concaténation de ces primitives définit la forme à reconnaître suivant la nature de celles-ci; on crée ou l'on utilise des grammaires construites sur un même modèle issu du langage naturel.

Une grammaire est dite formelle quand elle est définie par le quadruplet $G = (VT, VN, S, P)$ dans lequel : /23/ /24/ /25/

- . a . - V_T désigne un ensemble de primitives symboles constituant le vocabulaire terminal.
- . b . - V_N désigne un ensemble de symboles non terminaux que l'on nomme vocabulaire auxiliaire avec
 $V_T \cup V_N$ constituant le vocabulaire total V de G et
 $V_T \cap V_N = \emptyset$
- . c . - P désigne un ensemble fini de règles dites de réécriture ou de production qui sont notées :
 $\alpha \rightarrow \beta$ ou α et β sont des ensembles concaténés d'éléments du vocabulaire V .
- . d . - S désigne un symbole de début de phrase $\in V_N$ appelé axiome.

Cette grammaire peut s'utiliser de deux manières :

- en mode génératif

qui permet la construction de chaînes de symboles terminaux par application des règles de production.

. en mode analytique

qui décompose le mot obtenu à l'aide de règles de grammaires connues / 24/ jusqu'à obtenir des primitives.

4.1.2. - Application

Le vocabulaire terminal V_T de la grammaire envisagée est composée des différentes primitives déterminées au paragraphe précédent :

$$V_T = (S, A, B, C, D, E, F)$$

$$V_N = (\alpha, \beta, \gamma, \delta, \epsilon, \lambda)$$

primitives	A	B	C	D	E	F
formes						

Nous donnons, ci-après, les règles de production déterminées dans le cas général. Toutefois ces règles prennent en compte certaines restrictions telles que par exemple l'impossibilité d'avoir deux minimums successifs, c'est-à-dire deux primitives D successives sans transiter par un maximum.

$S \rightarrow A\alpha$	$\alpha \rightarrow B\beta$ $\alpha \rightarrow B$	$\alpha \rightarrow C\gamma$ $\alpha \rightarrow C$	
$S \rightarrow B\beta$	$\beta \rightarrow C\gamma$ $\beta \rightarrow C$	$\beta \rightarrow A\alpha$ $\beta \rightarrow A$	
$S \rightarrow C\gamma$	$\gamma \rightarrow E\epsilon$ $\gamma \rightarrow E$	$\gamma \rightarrow F\lambda$ $\gamma \rightarrow F$	$\gamma \rightarrow D\delta$ $\gamma \rightarrow D$
$S \rightarrow D\delta$	$\delta \rightarrow A\alpha$ $\delta \rightarrow A$	$\delta \rightarrow B\beta$ $\delta \rightarrow B$	$\delta \rightarrow C\gamma$ $\delta \rightarrow C$
$S \rightarrow E\epsilon$	$\epsilon \rightarrow F\lambda$ $\epsilon \rightarrow F$	$\epsilon \rightarrow D\delta$ $\epsilon \rightarrow D$	
$S \rightarrow F\lambda$	$\lambda \rightarrow E\epsilon$ $\lambda \rightarrow E$	$\lambda \rightarrow D\delta$ $\lambda \rightarrow D$	

La figure 1.5 illustre un exemple de courbes correspondant à une même phrase.

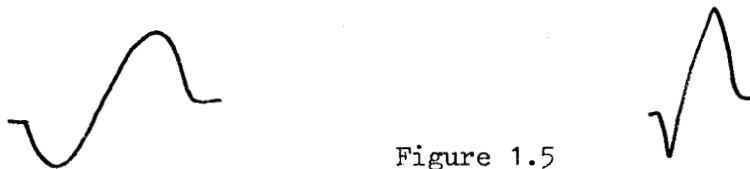


Figure 1.5

On peut aisément donner " l'équivalent mot " de cette forme grâce aux règles de production.

$$S \rightarrow E\epsilon \rightarrow ED\delta \rightarrow EDC\gamma \rightarrow EDCF$$

I.4.2 - APPROCHE PARAMETRIQUE

Elle consiste simplement à évaluer sur l'enveloppe du signal l'ensemble des paramètres caractéristiques définis précédemment :

- *pentés*
- *amplitudes d'extremums*
- *amplitudes des points de cassures*
- *durée séparant deux points de cassures*

Cette approche vient en complément de l'approche grammaticale.

I.5 - CLASSIFICATION

Outre que la classification des cycles utilise, comme la description, les approches grammaticale et paramétrique, on peut envisager deux schémas de travail selon la connaissance que l'on a, a priori, du système :

- dans le premier schéma, on considère que l'utilisateur fournit un cahier des charges permettant à l'analyste de disposer d'information a priori sur le signal.
- dans le second schéma, l'utilisateur ne fournit aucune information a priori et l'analyse doit passer par une phase d'apprentissage permettant de déterminer les formes à reconnaître /22/.

I.5.1 - CAS D'UN CAHIER DES CHARGES EXISTANT

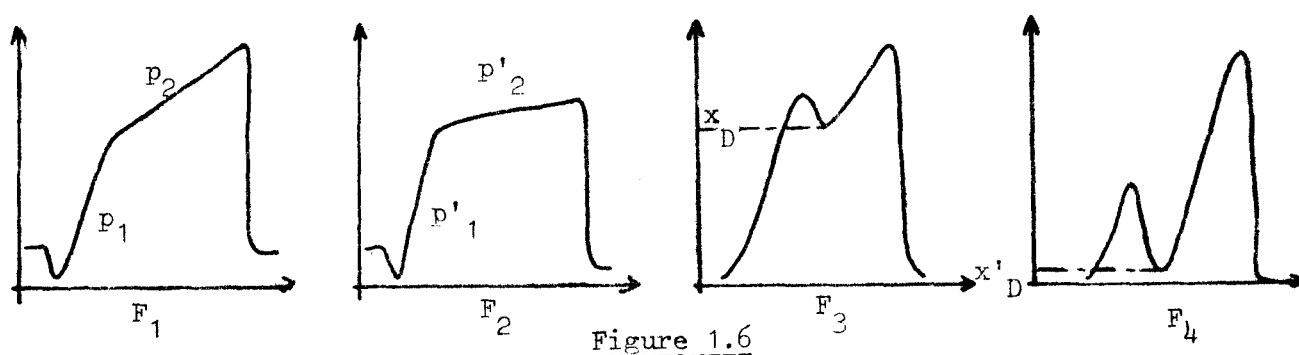
dans le cas où un cahier des charges est en mesure de spécifier pour un signal donné les formes caractéristiques à extraire, nous pouvons associer à ces formes des mots syntaxiques.

Ces mots syntaxiques sont construits à partir du vocabulaire terminal déterminé précédemment.

La reconnaissance d'une forme revient alors à comparer le mot construit aux mots décrivant les formes prototypes du cahier des charges.

Cette première étape permet de déterminer l'allure générale de la courbe. En conséquence, l'étude des paramètres caractérisant l'enveloppe de celle-ci doit venir en complément de la classification syntaxique.

Considérons à titre d'exemple la famille des formes illustrées figure 1.6 .



Cet ensemble fait apparaître quatre formes à reconnaître F_1 F_2 F_3 F_4 . F_1 et F_2 ont la même allure générale. Elles sont donc décrites par un même mot. Il en va de même pour F_3 et F_4 .

F_1 et F_2 se contruisent comme :

$S \rightarrow E\epsilon \rightarrow ED\delta \rightarrow EDB\beta \rightarrow EDBC\gamma \rightarrow EDBC\phi$

F_3 et F_4 se contruisent comme :

$S \rightarrow A\alpha \rightarrow AC\gamma \rightarrow ACD\delta \rightarrow ACDC\gamma \rightarrow ACDC\phi$

Il apparait clairement sur cet exemple simple que le second niveau de classification (classification paramétrique) est conditionné par les résultats donnés par le premier niveau (niveau syntaxique). Cet aspect est particulièrement intéressant car il permet d'utiliser les paramètres caractéristiques de manière parfaitement adaptée.

Après avoir reconnu le mot, on doit étudier certains paramètres spécifiques des formes :

F_1 et F_2 se différencient par le rapport des pentes p_2/p_1 et p'_2/p'_1 . Ce rapport est donc un élément de séparation pour F_1 et F_2 . F_3 et F_4 sont séparées quant à elles par les amplitudes de leurs minima x_D et x'_D .

I.5.2 - CAS DE CAHIER DES CHARGES INEXISTANT

Lorsque l'on n'a aucune information sur la nature des différentes formes susceptibles d'être rencontrées, on doit passer par une phase d'apprentissage qui a pour but de :

- 1°) recenser les différentes phrases générées par l'allure générale du signal étudié.
- 2°) affiner cette première classification en étudiant les paramètres décrivant l'enveloppe de la courbe.

2.1. - Phase syntaxique

La phase d'apprentissage permet d'associer à toute forme possible du signal un mot syntaxique généré à partir du langage terminal décrit aux paragraphes précédents.

Cette phase constitue l'inférence grammaticale qui a pour but de générer une grammaire à partir d'un langage connu. La reconnaissance devient ensuite immédiate. Elle consiste à vérifier si le mot descriptif de la forme appartient au langage.

Les grammaires les plus faciles à inférer sont les grammaires dites régulières. /23/ /24/.

Il existe un grand nombre d'algorithmes d'inférence différents plus ou moins performants /23/ /24/ /25/ .

Les trois grandes familles d'algorithmes sont :

- . les méthodes d'heuristique $u v^k w$
- . des k finales
- . d'agrégation des finales.

2.2. - Analyse des données

Pour affiner la classification entre deux courbes de même allure syntaxique, on doit faire appel à l'analyse des données.

Celle-ci se compose d'une très grande quantité de méthodes dont certaines s'adaptent mieux à la nature des données à traiter que d'autres /20/ /21/

Le but de ce paragraphe est de donner un aperçu succinct des méthodes les plus utilisées.

On peut diviser l'analyse des données en deux grosses branches /18/ /19/ 21/.

1° l'analyse factorielle

2° la classification

A . L'analyse factorielle

Elle a pour but de réduire le nombre des paramètres non pas par sélection de certains d'entre eux mais par construction de nouveaux paramètres qui ne sont que des combinaisons des paramètres initiaux /14/ /16/ /17/.

Dans ces conditions si p est le nombre de paramètres caractérisant une forme, chaque forme est représentée par un point dans un espace à p dimensions .

pour étudier la proximité de ces points entre eux, on doit alors munir les espaces d'une métrique. Le problème se pose ensuite en terme de géométrie et fait appel à des notions de combinaison linéaire, de distance, de projection, de centre de gravité, d'inertie etc ... /14/ /15/

Les différents types d'analyse factorielle se différencient alors par :

- la construction des nuages de points considérés
- le choix de la métrique qui sert à calculer les distances et les inerties.

B . Classification automatique

On peut distinguer deux grands types de méthodes

- non hiérarchiques
- hiérarchiques

. Les méthodes non hiérarchiques

Ce sont des méthodes grâce auxquelles on obtient un nombre fixé de classes. /11/

. Les méthodes hiérarchiques

Dans ce cas, chaque classe d'une partition est incluse dans une classe de la partition suivante ou précédente.

Cette classification peut se faire de deux manières :

- . *ascendante*
- . *descendante*

Les méthodes ascendantes utilisent une démarche d'agrégation binaire qui partant de l'ensemble des éléments les réunit deux par deux.

Les méthodes descendantes sont fondées sur le principe de désagrégation binaire.

Elles ont l'avantage d'être moins coûteuses que les précédentes mais elles sont moins précises.

I.6 - TRAITEMENT TEMPS RÉEL

I.6.1 - BUT

Le suivi en continu d'un signal demande une détermination suffisamment rapide de toute anomalie ou de toute forme particulière afin de réagir

éventuellement sur le système. Nous envisageons donc une gestion temps réel de la classification des formes.

I.6.2 - RAPPEL SUR LE TEMPS REEL

Soit un signal échantillonné à une période d'échantillonnage T . Le suivi temps réel de ce signal consiste à traiter ce signal entre deux prises d'échantillons successives. Cette procédure implique une limitation du temps de calcul à la période d'échantillonnage.

Dans le cas où T est très courte, la limitation est évidemment considérable. Afin de pallier cette difficulté, le traitement à effectuer est scindé en plusieurs traitements indépendants et exclusifs les uns des autres, mais dépendant de l'instant dans lequel se trouve le signal. En conséquence un choix entre ces différents traitements doit être opéré.

Chacun de ces traitements est composé de différentes tâches qui se complètent et s'enchainent. Celles-ci peuvent être également exclusives les unes des autres.

La gestion de ces tâches et de ces traitements est réalisée par un superviseur. Le rôle du superviseur est alors de mémoriser l'état de calcul à chaque période afin d'activer pour la période suivante, les traitements à effectuer compte tenu du nouvel échantillon acquis.

Cette mémorisation se fait à l'aide de drapeaux que le superviseur positionne suivant l'état d'avancement du traitement.

I.6.3 - APPLICATION

Les différents traitements à effectuer sont au nombre de quatre .

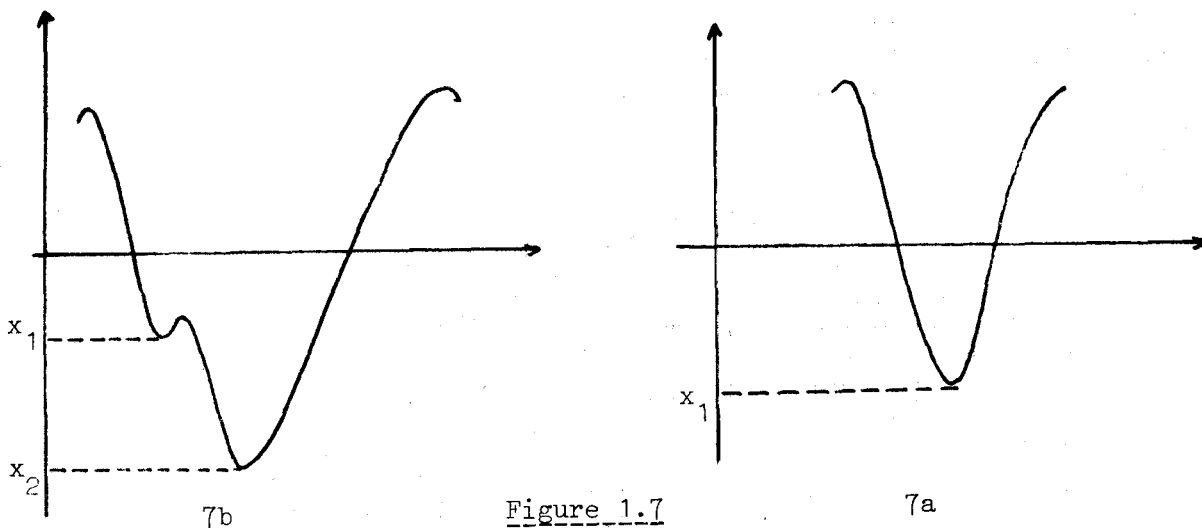
- T1 : recherche de début de cycle
- T2 : test de validité du point de début
- T3 : recherche des primitives A, B, C, D, E, F
détermination des paramètres caractéristiques
- T4 : classification des formes

Les trois premiers traitements T1, T2, T3 sont activés à chaque instant d'échantillonnage et se terminent par l'extraction d'un cycle. La classification de formes n'intervient qu'à l'issue de l'extraction d'un cycle, après la validation du point de début de cycle (un début de cycle correspond à la fin d'un cycle précédent). Cette classification se fera au rythme d'un cycle.

La réalisation du superviseur est entièrement conditionnée par la recherche des six primitives A, B, C, D, E, F à laquelle s'ajoute la recherche du début de cycle.

Il est à noter que compte tenu des restrictions dans les règles de production de la grammaire, la détection d'une primitive, restreint la recherche de la primitive suivante. Ainsi et grâce au superviseur, on évite de scruter systématiquement tous les " cas de figure " .

En outre le traitement temps réel nécessite dans certains cas, l'utilisation de la technique de confirmation ou d'infirmité d'hypothèses. Considérons à titre d'exemple les cycles de la figure 1.7 .



On considère que le début de cycle coïncide avec un minimum absolu négatif suivi d'un maximum positif.

Dans les deux cas, à la détection du minimum x_1 , on considère que ce point est un début possible aussi bien qu'un minimum local. Il faut attendre quelques échantillons plus tard, c'est-à-dire la détection d'un maximum positif figure 1.7a ou négatif 1.7b pour pouvoir conclure.

L'arbre programmatique décrivant les différents traitements de notre application est décrit figure 1.8 .

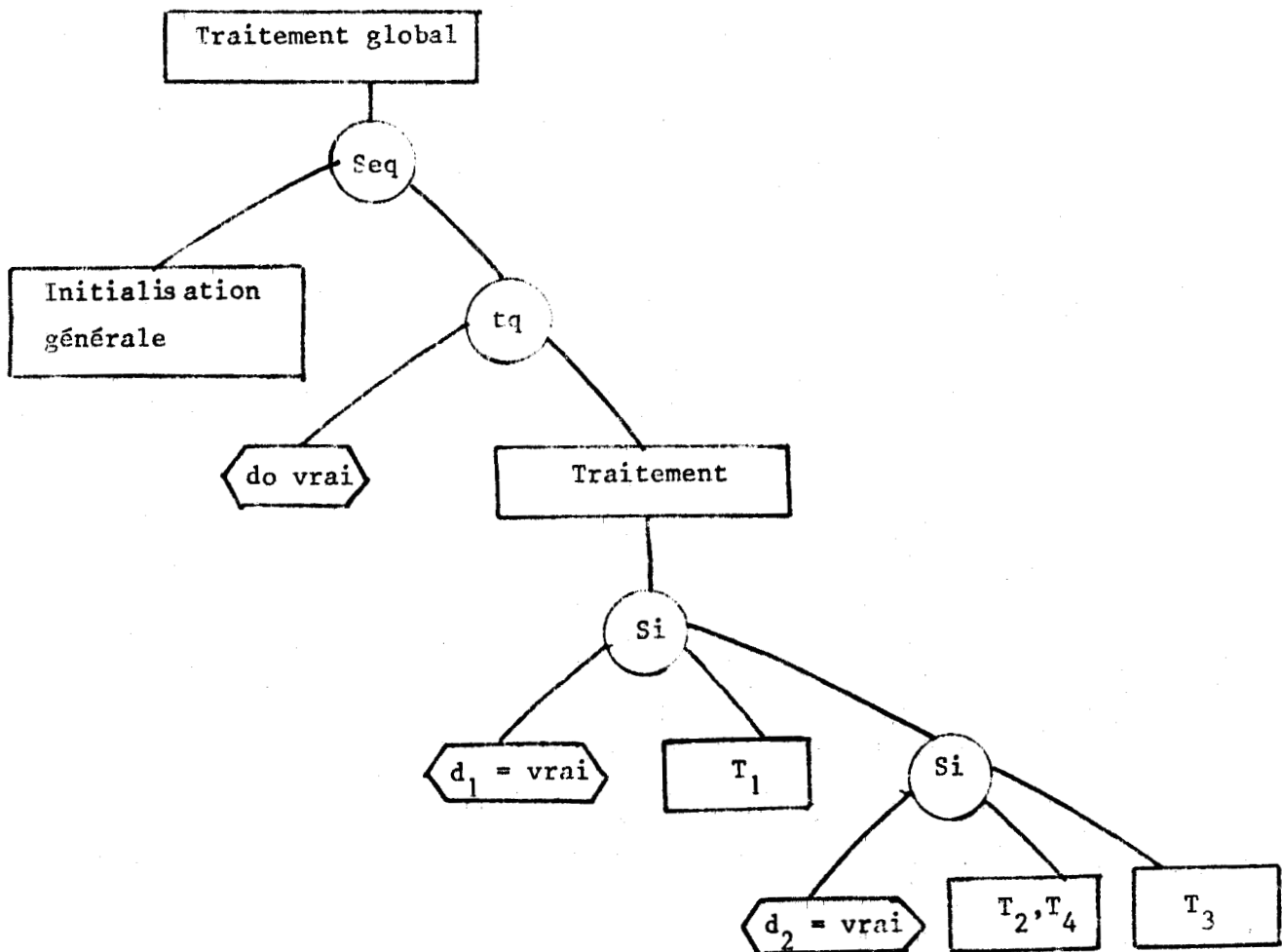


Figure 1.8

Afin d'assurer un traitement continu des échantillons prélevés sur le signal d'entrée, sans qu'une interruption intervienne un drapeau d_0 est positionné à une valeur vraie, et testé à cette valeur.

L'arrêt du traitement se fait en positionnant d_0 à faux (équivalent à une rupture).

L'étape d'initialisation générale ne devant être faite qu'une seule fois, cette tâche est à l'extérieur du test de d_0 .

Pour effectuer le traitement T_1 , on teste le drapeau d_1 qui est mis à la valeur vraie, lorsque l'on doit rechercher un début de cycle, ceci lors de l'initialisation ou après le traitement T_3 . On donne à d_1 la valeur " faux " quand la tâche T_1 est effectuée et à d_2 la valeur vraie.

Cela permet de tester dans un premier temps la validité du point de début du cycle, c'est-à-dire la fiabilité du point déterminé par T_1 . S'il y a possibilité, on procède alors à la classification de la forme.

La figure 1.9 illustre l'algorithme de la tâche T_2 , T_4 .

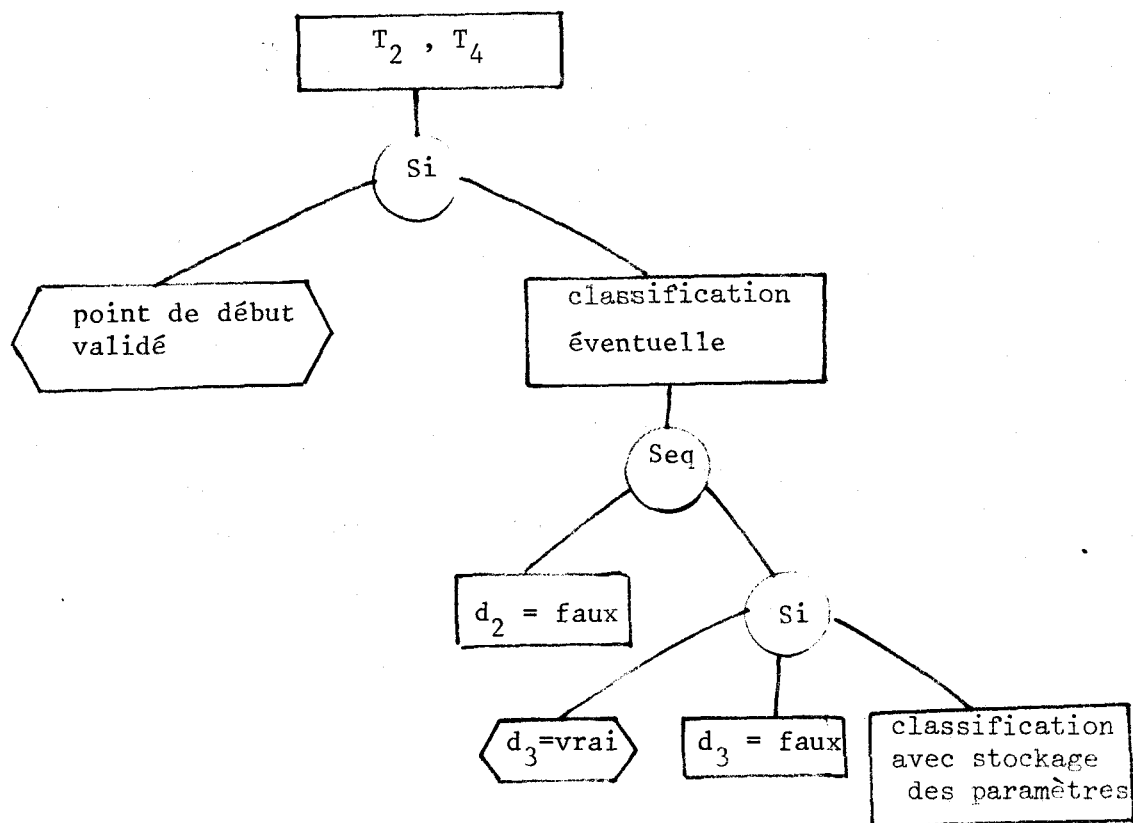


Figure 1.9

d_3 est un drapeau qui décide s'il y a matière à faire une classification ou non.

Dans le cas d'un début de traitement, la validation d'un point de début contrairement au cas suivant ne correspond pas à l'isolation d'un cycle. On doit donc éviter tout traitement sur le cycle ce qui est fait en initialisant d_3 à la valeur vraie.

Par contre après un premier passage qui rendra d_3 faux, un cycle sera extrait à chaque validation du point de début de cycle, on pourra donc activer un programme de reconnaissance de formes (classification).

La classification des formes se fait donc en deux temps :

- 1°) recherche d'un mot syntaxique donnant la structure de la forme
- 2°) test des valeurs des points spécifiques.

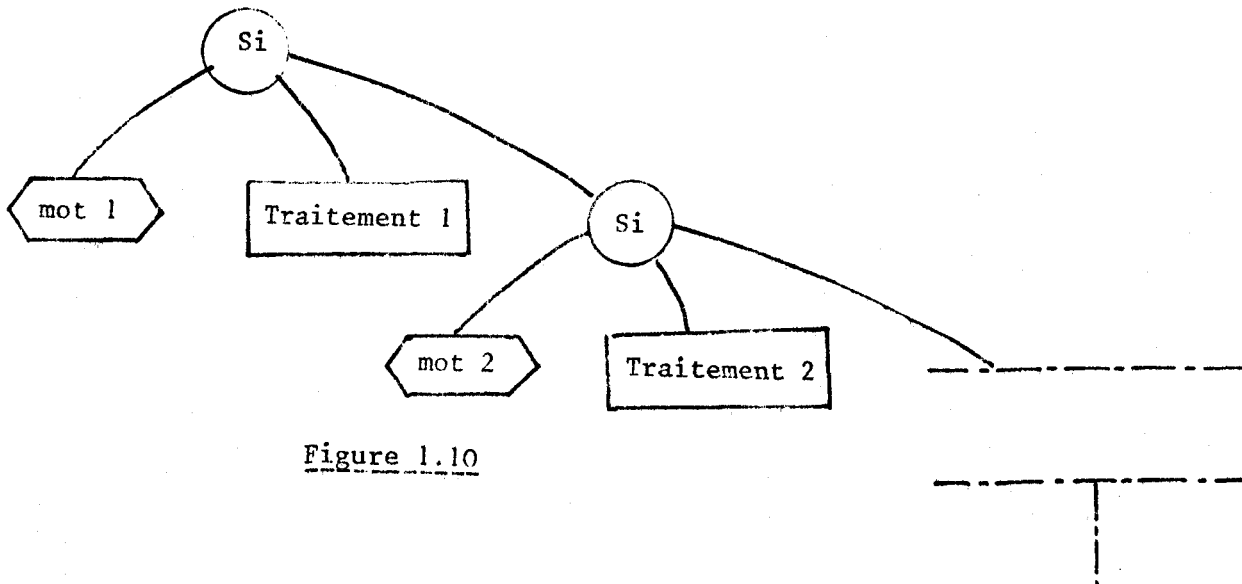


Figure 1.10

Et l'on positionne d_2 à faux pour qu'à l'échantillon suivant le traitement T_3 s'effectue. Il se compose d'une série de tests qui positionne le superviseur dans les différentes tâches de recherche de primitives illustrées figure 1.11.

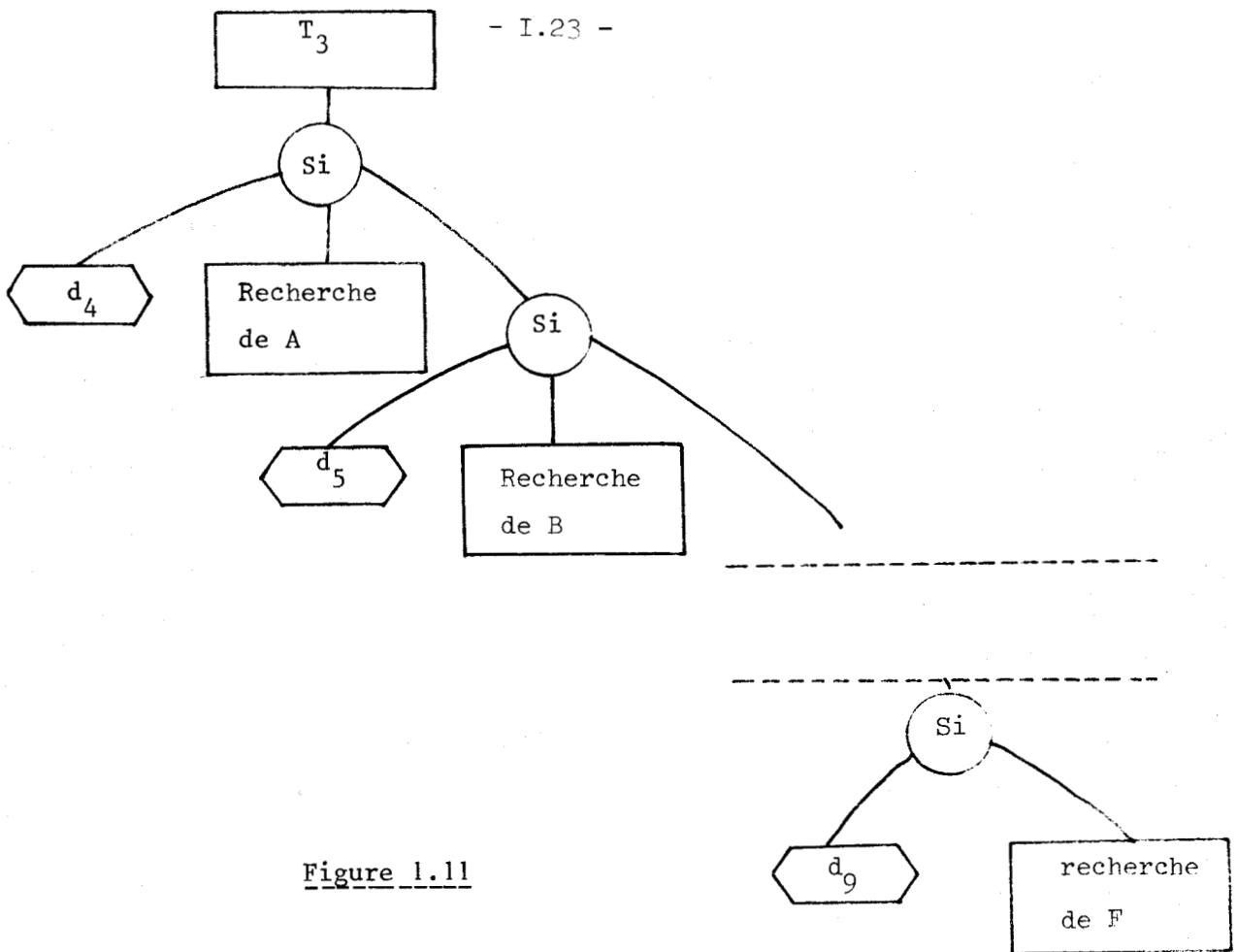


Figure 1.11

Les drapeaux $d_4, d_5, d_6 \dots d_9$ servent à mémoriser l'état du calcul et grâce à un jeu de tests appropriés positionne le traitement sur l'une des différentes tâches. (recherche de A, recherche de F).

Chacune de celles-ci peut se décomposer en éléments eux-mêmes gérés par un mini superviseur. A titre d'exemple, la figure 1.12 illustre la recherche de B.

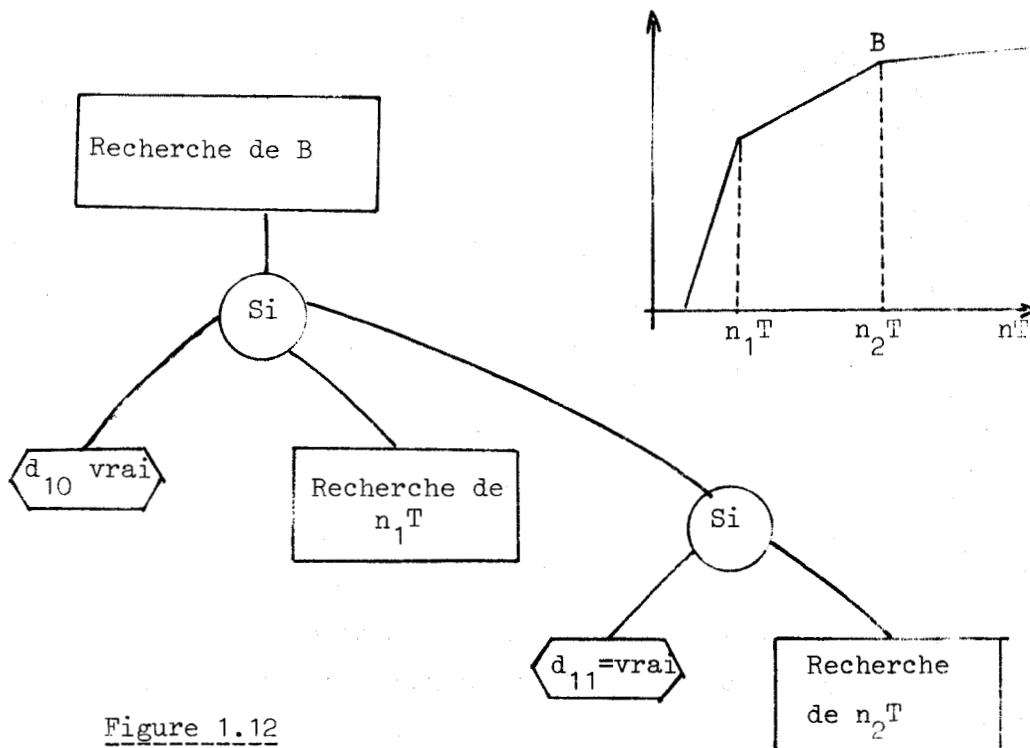


Figure 1.12

A l'issue de la détection de $n_2 T$, il y aura un positionnement de d_5 à faux.

A chaque fin de tâche ou de traitement, il y a un changement de valeur des différents drapeaux qui permet la bonne gestion du traitement global.

I.7 - CONCLUSION

Les méthodes de reconnaissance de formes exposées dans ce chapitre ont pour but de faire une description structurelle d'un signal évoluant dans le temps.

Elles ont permis de définir pour chaque cycle de signal, un prototype reflétant un état spécifique du processus étudié.

Il est à noter que dans certaines études la forme ne suffit pas à elle seule à caractériser entièrement le processus.

Dans ce cas, nous sommes amenés à prendre en compte d'autres paramètres qui permettent d'affiner la classification afin d'obtenir des prototypes suffisamment précis.

L'utilisation des méthodes de reconnaissance de forme a permis de créer à partir du signal évoluant dans le temps, une suite de formes caractéristiques.

Par ailleurs, la nécessité d'effectuer les traitements en temps réel nous contraint à diviser le traitement général en tâches spécifiques qui sont effectuées à des moments précis relatifs au signal.

L'utilisation de variables Booléennes d_0 , d_1 nommées drapeaux, permet par simple scrutation de celles-ci de se positionner à telle ou telle tâche ceci dans le but de minimiser le temps de traitement.

L'objet du chapitre suivant est alors d'utiliser les propriétés de ces suites de formes afin d'analyser les relations pouvant lier plusieurs signaux différents.

- C H A P I T R E I I -

ANALYSE INFORMATIONNELLE DES RELATIONS INTERSIGNALS

CHAPITRE II

II.1 - INTRODUCTION	page 1
II.2 - NOTION SUR LES SYSTEMES	page 2
II.3 - RAPPEL SUR LA THEORIE DE L'INFORMATION	page 6
II.3.1. - Introduction	page 6
II.3.2. - Entropie	page 7
II.3.3. - Propriétés de l'entropie	page 8
II.3.4. - Transinformation	page 9
II.3.5. - Cas de deux variables	page 9
3.5.1. - Notations	page 9
3.5.2. - Cas où X_1 et X_2 sont indépendants	page 10
3.5.3. - Cas où X_1 et X_2 sont dépendants	page 11
II.4 - APPLICATION	page 13
II.4.1. - Le tableau des données	page 13
4.1.1. - Acquisition des données	page 13
4.1.2. - Importance du changement d'état	page 14
II.4.2. - Cas de signaux synchronisés	page 14
II.4.3. - Prise en compte du retard	page 18
II.4.4. - Calcul d'entropie	page 19
II.4.5. - Comparaison des entropies H_{12} et H_{21}	page 24
II.4.6. - Exemple d'illustration	page 25
II.4.7. - Algorithme	
II.5 - CONCLUSION	page 38

II.1 - INTRODUCTION

Le chapitre précédent a permis d'envisager la caractérisation d'un signal donné en le décomposant en primitives appropriées et en opérant une reconnaissance puis une classification automatique sur ces primitives.

L'objet du deuxième chapitre est de sélectionner parmi un ensemble de signaux donnés, les signaux les plus représentatifs du fonctionnement d'un système. Pour cela nous envisageons d'étudier les relations existant entre les signaux.

La mise en évidence de ces relations permet d'éviter tout calcul redondant qui alourdit l'interprétation des résultats lors d'une surveillance automatique temps réel.

Pour atteindre cet objectif, deux solutions sont possibles. La première issue du traitement du signal utilise les intercorrélations/13/ La deuxième est donnée par la théorie de l'information et met en jeu le calcul d'entropie.

C'est ce dernier point qui a été retenu pour cette étude et qui fait l'objet de ce chapitre. /8 /

En premier lieu, nous nous attacherons donc à situer le système duquel sont extraits les signaux dans la hiérarchie épistémologique* des systèmes, ceci dans le but de formaliser le problème et de déterminer la méthode adéquate pour sa résolution.

Après un bref rappel sur la théorie de l'information, qui permet de mesurer les échanges d'informations entre les différentes parties d'un système, nous appliquerons ces notions aux signaux.

Le problème sera ensuite formalisé par construction d'un tableau de données reflétant la connaissance que l'on a du système.

Enfin le chapitre se termine par la présentation d'arbres programmatiques donnant les tâches à effectuer pour la résolution du problème dans le cas particulier de deux signaux.

* George J. KLIR a défini une hiérarchie épistémologique des systèmes qui établit les différents niveaux de connaissance qu'un observateur peut avoir d'un système
/31 / / 32/.

II.2 - NOTION SUR LES SYSTÈMES

Si a priori aucune étude préalable n'a été faite sur le système, il importe d'appréhender un grand nombre de variables ou de signaux pour en cerner avec précision les comportements.

Par contre si un spécialiste du processus étudié peut guider le choix des signaux les plus significatifs, il importe d'exploiter cette expérience avant toute considération mathématique. Ceci permet de simplifier et surtout de rendre abordable l'analyse qui devient très délicate lorsque la quantité de données est très importante.

La connaissance que l'on a du système peut se situer à différents niveaux (hiérarchie épistémologique), allant d'une simple connaissance des données jusqu'à une connaissance des relations internes au système /33/. Dans cette hiérarchie la présente étude se situe au premier niveau , celui des données.

Dans ces conditions la démarche utilisée pour l'observation d'un système peut se schématiser selon l'illustration de la figure 2.1 sur laquelle apparaissent les éléments suivants :

$\Omega = \{\omega_1, \omega_2, \dots, \omega_L\}$ ensemble fini de L observations

$\Sigma = \{X_1, X_2, \dots, X_n\}$ ensemble de n paramètres attachés à ces observations

$M_{X_1} M_{X_2} \dots M_{X_n}$: n ensembles finis d'états pouvant être pris respectivement par $X_1, X_2, \dots, X_n$.

avec $M_{X_i} = \{e_1^i, e_2^i, e_j^i, \dots, e_{\omega_i}^i\}$

et e_j^i modalité correspondant au $j^{\text{ème}}$ état éventuel que peut prendre la variable X_i

M_{X_i} possède ω_i modalités

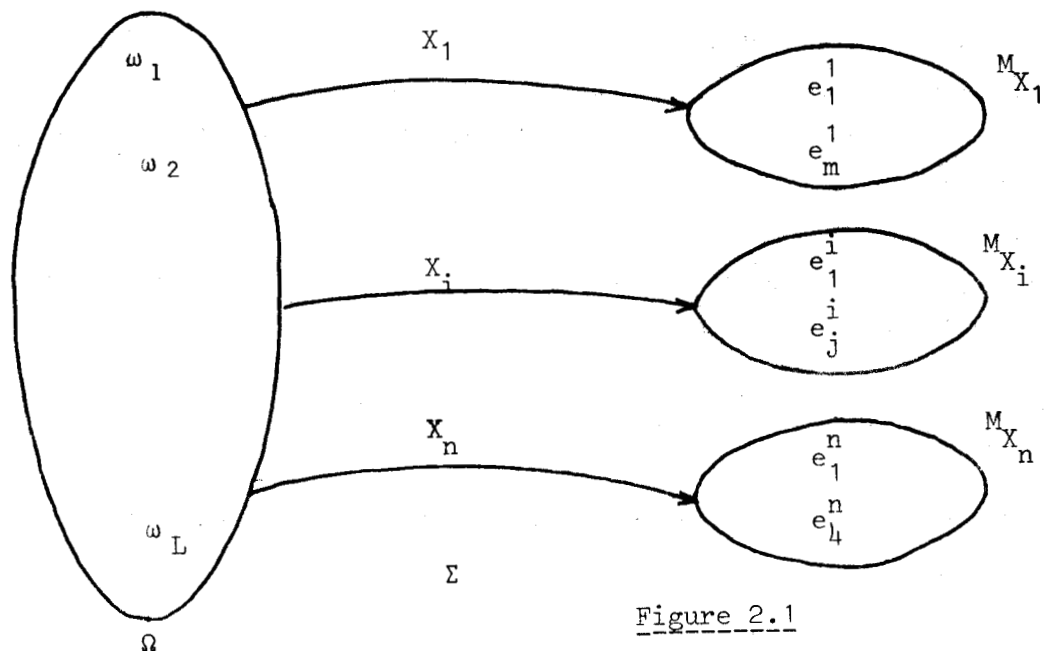


Figure 2.1

Sur l'ensemble Ω formé des observations qui constituent la connaissance initiale que l'on a du système, on prélève des paramètres qui prennent leurs valeurs dans des ensembles de modalités M_{X_i} /34/ /36/.

Le tableau de la figure 2.2, appelé tableau des observations, recense les différents états pris par les X_i au cours des L observations.

$\Omega \quad \Sigma$	X_1	X_2 X_i X_n
ω_1	e_α^1	e_β^2 e_j^i e_δ^n
ω_2	e_δ^1	e_β^2 e_ξ^2 e_γ^2
ω_j	e_ω^1	e_ξ^2 e_γ^2 e_γ^2
ω_L	e_α^1	e_γ^2 e_γ^2 e_γ^2

Figure 2.2

Exemple d'illustration.

Cet exemple illustre l'étude des relations existant entre la consommation en essence, la vitesse et la marque de fabrication, pour un ensemble de véhicules donné.

Pour ce faire, l'espace d'observation s'identifie à un ensemble de voitures. Ainsi, pour notre exemple nous étudions une série de 6 voitures.

$$\Omega = \{ v_1, v_2, \dots, v_6 \} \quad (v_i = \text{voiture n}^\circ i)$$

Par ailleurs les paramètres étudiés sont, bien évidemment au nombre de 3 :

X_1 = consommation en essence aux 100 km

X_2 = vitesse en km/h

X_3 = marque du véhicule | Renault, Peugeot ..., |

L'espace d'observation doit être constitué avec soin afin de tenir compte de toutes sortes de véhicules à notre disposition sans pour autant être redondant. Par exemple si 3 R5 figurent dans l'espace d'observation, une seule de ces R5 sera prise en compte. Les observations sont donc sélectives.

Enfin, il faut définir des modalités pour chacun des paramètres X_i . Ainsi la consommation d'essence X_1 , s'échelonne de 51 à 101 et l'ensemble M_{X_1} correspondant comporte les différents états que peut prendre le paramètre X_1 dans cette plage de valeurs.

e_1^1 représente une consommation de] 51 à 61]

e_2^1 " " " "] 61 à 71]

e_3^1 " " " "] 71 à 81]

e_4^1 " " " "] 81 à 91]

e_5^1 " " " "] 91 à 101]

Les états e_i^2 issus du paramètre X_2 sont créés à partir d'une vitesse de 40 km/h.

e_1^2					représente une vitesse entre [40 km/h et 60 km/h]
e_2^2	"	"	"	"	] 60 km/h et 80 km/h]
e_3^2	"	"	"	"	] 80 km/h et 100 km/h]
e_4^2	"	"	"	"	] 100 km/h et 120 km/h]
e_5^2	"	"	"	"	supérieure à 120 km/h

Les états e_j^3 se rapportant aux paramètres X_3 représentent respectivement :

e_1^3 : marque 1

e_2^3 : marque 2

e_3^3 : marque 3

e_4^3 : marque 4

Nous avons ainsi formé les ensembles de modalités :

$$M_{X_1} = \{ e_1^1 \quad e_2^1 \quad e_3^1 \quad e_4^1 \quad e_5^1 \}$$

$$M_{X_2} = \{ e_1^2 \quad e_2^2 \quad e_3^2 \quad e_4^2 \quad e_5^2 \}$$

$$M_{X_3} = \{ e_1^3 \quad e_2^3 \quad e_3^3 \quad e_4^3 \}$$

Les six observations permettent alors de parvenir aux conclusions suivantes :

La voiture A de marque 1 consomme 6l50 à la vitesse de 80 km/h

La voiture B de marque 1 consomme 81 à la vitesse de 100km/h

La voiture C de marque 4 consomme 7l60 à la vitesse de 45 km/h

La voiture E de marque 2 consomme 5l60 à la vitesse de 50 km/h

La voiture F de marque 4 consomme 91 à la vitesse de 120km/h

La voiture G de marque 3 consomme 6l30 à la vitesse de 90 km/h

Le tableau des observations, en fonction des notations prises dans cet exemple est alors donné figure 2.3 .

Ω	x_1	x_2	x_3
ω_1	e_2^1	e_2^2	e_1^3
ω_2	e_4^1	e_3^2	e_1^3
ω_3	e_3^1	e_1^2	e_4^3
ω_4	e_1^1	e_1^2	e_2^3
ω_5	e_5^1	e_4^2	e_4^3
ω_6	e_2^1	e_3^2	e_3^3

Figure 2.3

II.3 - RAPPEL SUR LA THÉORIE DE L'INFORMATION

II.3.1 - INTRODUCTION

L'application que nous envisageons vise à rechercher les relations pouvant exister entre les formes extraites des signaux prélevés sur un processus. Pour cela, la théorie classique du signal qui utilise des mesures pour le calcul des fonctions de corrélation ne peut pas être appliquée.

Cette impossibilité provient du fait que les formes sont des entités classifiables , mais non mesurables.

Par contre, la théorie de l'information, qui s'appuie sur des éléments statistiques offre des outils intéressants pour mesurer les échanges d'informations existant entre différentes classifications. /37/ /38/.

C'est cet aspect qui est développé dans ce paragraphe.

II.3.2 - ENTROPIE

Selon les notations de la figure 2.1 un premier objectif est de caractériser l'information qu'il est possible d'acquérir pour chacun des paramètres X_i .

L'information que l'on a d'un paramètre X_i , ou son degré d'indétermination n'est autre que l'entropie que l'on note $H(X_i)$ /34/.

Soit l'ensemble M_{X_i} que l'on a partitionné en modalités deux à deux disjointes, comme le montre l'exemple précédent. (consommation d'essence, vitesse,...).

Ces différentes modalités $e_1^i, e_2^i, \dots, e_{\omega_i}^i$ se produisent avec des probabilités $\text{pr}(X_i = e_1^i), \text{pr}(X_i = e_2^i), \dots, \text{pr}(X_i = e_{\omega_i}^i)$

Pour chaque X_i donné, donc pour $i = 1, n$, les propriétés suivantes sont vérifiées :

$$\forall j \in [1, \omega_i].$$

$$\text{pr}(X_i = e_j^i) \geq 0 \text{ et}$$

$$\sum_{j=1}^{\omega_i} \text{pr}(X_i = e_j^i) = 1.$$

L'entropie est alors l'information fournie par la variable aléatoire X_i et sa loi de probabilité correspondante.

Elle est définie par l'espérance mathématique de l'information donnée par chaque cas particulier dans lequel la variable prend l'état e_j^i $\left[j=1, \omega_i \right]$

$$H(X_i) = - \sum_{j=1}^{\omega_i} \text{pr}(X_i = e_j^i) \log_2 \text{pr}(X_i = e_j^i)$$

avec card $M_{X_i} = \omega_i$

II.3.3 - PROPRIETES DE L'ENTROPIE

1. L'information ne dépend que de la distribution de probabilité $P = \{ \text{pr}(X_i = e_j^i) \}$ et est donc indépendante du support.

2. La base logarithmique de l'entropie est 2

3. La fonction $\varphi(X) = -X \log X$ est définie positive sur l'intervalle $] 0, 1]$ d'où

$$\forall X_i \quad H(X_i) \geq 0 \text{ puisque } \text{pr}(X_i = e_j^i) \in [0, 1]$$

$H(X_i) = 0$ X_i est constante. Pas de variation donc pas d'indétermination.

4. L'entropie $H(X_i)$ prend sa valeur maximale lorsque les événements élémentaires sont équiprobables.

$$\text{Si } \forall j \text{ pr}(X_i = e_j^i) = \frac{1}{\omega_i} \quad \Rightarrow$$

$$H(X_i) = - \sum_{j=1}^{\omega_i} \text{pr}(X_i = e_j^i) \log \text{pr}(X_i = e_j^i) = \log \omega_i$$

5. Soient X_1 et X_2 deux variables.

$$H(X_1 \cup X_2) \leq H(X_1) + H(X_2)$$

II.3.4 - TRANSINFORMATION

La 5ième propriété suggère qu'il y a une quantité d'information commune aux deux variables X_1 et X_2 et qui doit être comptée deux fois dans $H(X_1) + H(X_2)$.

On considère que cette information est transmise d'une variable à une autre et est appelée transinformation ou information mutuelle.

De manière plus générale la quantité d'information transmise entre plusieurs variables d'un même ensemble $\{X_1, X_2, \dots, X_{n_n}\} = \Sigma$ se mesure par la transinformation $I(\Sigma)$ qui a pour expression $I(\Sigma) = \sum_{i=1}^n H(X_i) - H(\bigcup_{i=1}^n X_i)$

II.3.5 - CAS DE DEUX VARIABLES

Nous donnons ci-après les éléments permettant d'évaluer la transinformation (ou information mutuelle) dans le cas de deux variables X_1 et X_2 .

3.5.1. - Notations

Soient les deux ensembles.

$$M_{X_1} = \{e_1^1, e_2^1, \dots, e_q^1\} \text{ avec card } M_{X_1} = q$$

$$M_{X_2} = \{e_1^2, e_2^2, \dots, e_r^2\} \text{ avec card } M_{X_2} = r$$

afin de simplifier les notations, nous écrirons :

$$p_i(X_1) = \text{pr}(X_1 = e_i^1) \quad i = 1, q$$

$$p_j(X_2) = \text{pr}(X_2 = e_j^2) \quad j = 1, r$$

$$p_{ij} = \text{pr}(X_1 = e_i^1, X_2 = e_j^2)$$

$$\text{avec } p_i(X_1) = \sum_{j=1}^r p_{ij}$$

$$p_j(X_2) = \sum_{i=1}^q p_{ij}$$

Dans ces conditions, on peut définir la fonction d'entropie suivante :

$$H(X_1, X_2) = - \sum_{j=1}^r \sum_{i=1}^q p_{ij} \log_2 p_{ij}$$

relative à l'ensemble produit $M_{X_1} \times M_{X_2}$

3.5.2. - Cas où X_1 et X_2 sont indépendants

Dans ce cas nous avons :

$$p_{ij} = p_i(X_1) \cdot p_j(X_2)$$

$$H(X_1, X_2) = - \sum_{i=1}^q \sum_{j=1}^r p_{ij} \cdot \log_2 p_{ij}$$

$$= - \sum_{i=1}^q \sum_{j=1}^r p_i(X_1) \cdot p_j(X_2) \cdot \log [p_i(X_1) \cdot p_j(X_2)]$$

$$H(X_1, X_2) = - \sum_{i=1}^q \sum_{j=1}^r p_i(X_1) p_j(X_2) \log p_i(X_1) - \sum_{i=1}^q \sum_{j=1}^r p_i(X_1) p_j(X_2) \log p_j(X_2)$$

$$= - \sum_{i=1}^q p_i(X_1) \log p_i(X_1) \sum_{j=1}^r p_j(X_2) - \sum_{i=1}^q p_i(X_1) \sum_{j=1}^r p_j(X_2) \log p_j(X_2)$$

$$\text{sachant que } \sum_{j=1}^r p_j(X_2) = 1 \text{ et } \sum_{i=1}^q p_i(X_1) = 1$$

L'expression devient

$$H(X_1, X_2) = - \sum_{i=1}^q p_i(X_1) \log p_i(X_1) - \sum_{j=1}^r p_j(X_2) \log p_j(X_2)$$

$$H(X_1, X_2) = H(X_1) + H(X_2)$$

Dans ces conditions, la transinformation est nulle (variables indépendantes).

3.5.3. - Cas où X_1 et X_2 sont dépendants

La loi de BAYES nous permet dans ce cas de dépendance entre les variables de définir une formule à base de probabilités conditionnelles $P(X_2/X_1)$: probabilité d'obtenir X_2 sachant X_1 qu'on notera $p_{j/i}(X_1, X_2)$.

Suivant les notations précédentes, nous avons :

$$p_{ij}(X_1, X_2) = p_i(X_1) \times p_{j/i}(X_1, X_2)$$

d'où

$$\begin{aligned} H(X_1, X_2) &= - \sum_{i=1}^q \sum_{j=1}^r p_{ij}(X_1, X_2) \cdot \log p_{ij}(X_1, X_2) \\ &= - \sum_{i=1}^q \sum_{j=1}^r p_i(X_1) p_{j/i}(X_1, X_2) \cdot \log [p_i(X_1) \cdot p_{j/i}(X_1, X_2)] \\ &= - \sum_{i=1}^q \sum_{j=1}^r p_i(X_1) \cdot p_{j/i}(X_1, X_2) \log p_i(X_1) - \sum_{i=1}^q \sum_{j=1}^r p_i(X_1) p_{j/i}(X_1, X_2) \log p_{j/i}(X_1, X_2) \\ &= - \sum_{i=1}^q p_i(X_1) \log p_i(X_1) \cdot \sum_{j=1}^r p_{j/i}(X_1, X_2) - \sum_{i=1}^q p_i(X_1) \sum_{j=1}^r p_{j/i}(X_1, X_2) \log p_{j/i}(X_1, X_2) \end{aligned}$$

$$\text{sachant que } \sum_{j=1}^r p_{j/i}(X_1, X_2) = 1 \text{ et } \sum_{i=1}^q p_i(X_1) = 1$$

nous obtenons

$$H(X_1, X_2) = - \sum_{i=1}^q p_i(X_1) \log p_i(X_1) - \sum_{j=1}^r p_{j/i}(X_1, X_2) \log p_{j/i}(X_1, X_2)$$

en notant

$$H(X_2/X_1) = - \sum_{j=1}^r p_{j/i}(X_1, X_2) \log p_{j/i}(X_1, X_2)$$

la formule devient

$$H(X_1, X_2) = H(X_1) + H(X_2/X_1)$$

où

$H(X_2/X_1)$ est appelée entropie conditionnelle.

La transinformation $\overset{\leftrightarrow}{I}$ peut alors être redéfinie à partir de cette nouvelle donnée.

$H(X_2/X_1) = H(X_1, X_2) - H(X_1)$. Nous savons que

$$\overset{\leftrightarrow}{I} = H(X_2) + H(X_1) - H(X_1, X_2)$$

$$\overset{\leftrightarrow}{I} = H(X_2) - H(X_2/X_1)$$

La transinformation normalisée paraît plus adéquate à manipuler

$$\frac{\overset{\leftrightarrow}{I}}{H(X_2)} = 1 - \frac{H(X_2/X_1)}{H(X_2)}$$

. Cas de variables X_1 et X_2 indépendantes

$$H(X_2/X_1) = H(X_2) \Rightarrow \frac{H(X_2/X_1)}{H(X_2)} = 1 \Rightarrow \frac{\overset{\leftrightarrow}{I}}{H(X_2)} = 0$$

. Cas de variables X_1 et X_2 dépendantes

$$0 \leq H(X_2/X_1) \leq H(X_2) \Rightarrow 0 \leq \frac{H(X_2/X_1)}{H(X_2)} \leq 1 \Rightarrow$$

$$0 \leq \frac{\overset{\leftrightarrow}{I}}{H(X_2)} < 1, \text{ cette valeur est nulle s'il y a indépendance entre les}$$

variables traitées.

Par contre on peut chiffrer un degré de dépendance compris entre 0 et 1.

II.4 - APPLICATION

L'application des notions décrites précédemment pose le problème fondamental de la collecte des données et de la formation des tableaux permettant d'estimer les paramètres statistiques mis en jeu.

Par ailleurs l'entropie et la transinformation entre deux signaux X_1 et X_2 étant des opérations commutatives, elles ne permettent pas de définir une direction dans le passage de l'information. (soit de X_1 vers X_2 soit de X_2 vers X_1).

De plus elles n'intègrent pas parfaitement la variable temps et ne permettent pas, par exemple, de tenir compte des retards pouvant intervenir dans le passage de l'information d'un signal vers l'autre.

Le but de ce paragraphe est d'essayer de pallier ces difficultés.

II.4.1 - LE TABLEAU DES DONNEES

C'est un tableau qui recense toutes les informations que l'on a du système. Il dépend directement des observations et se trouve à la base de tous les calculs ce qui explique le soin particulier que l'on doit mettre à l'établir.

4.1.1. - Acquisition des données

Le problème est de recenser un nombre important de formes pour que l'estimation des paramètres statistiques ne soit pas biaisée.

Toutefois dans le cadre d'un suivi, plus que la prise en compte systématique de toutes les formes ou états, c'est la prise en compte des changements de formes qui importe.

4.1.2. - Importance du changement d'état

Il est intéressant d'examiner les transitions des divers signaux et d'analyser les relations liant les changements d'états. L'idée est d'évaluer l'influence d'un signal sur un autre.

Cette influence se traduit par le fait qu'un changement de forme au niveau de l'un des deux signaux engendre un changement de forme sur l'autre signal, soit simultanément soit avec un certain retard.

Ne sachant pas, a priori, quel est le signal qui influe sur l'autre, nous rechercherons, dans un premier temps, l'existence d'un transfert d'information, aussi bien vers un signal que vers l'autre, ceci en tenant compte de tous les changements sur les signaux.

Ces changements d'états pouvant être indépendants ou non les uns des autres, le premier problème à résoudre est de mettre en évidence l'existence ou l'absence de liens entre les signaux.

L'objectif est ensuite de déterminer quel est le signal le plus représentatif du processus étudié. Un tel signal sera appelé "signal maître", tandis que l'autre sera appelé "signal esclave". Pour cette étude les calculs d'entropies sont utilisés.

Toutefois et afin de privilégier un signal par rapport à un autre, les tableaux de données sont construits de manière adaptée.

II.4.2 - CAS DE SIGNAUX SYNCHRONISES

Deux signaux sont synchronisés lorsque l'occurrence d'un état sur un signal s'accompagne de l'occurrence simultanée d'un état sur l'autre signal.

L'évolution dans le temps de deux signaux S_1 et S_2 synchronisés se schématise alors par une suite d'états synchrones 2 à 2 (figure 2.4).

EXEMPLE

S_1	e_α^1	e_β^1	e_β^1	e_β^1	e_α^1	e_α^1	e_α^1	e_γ^1	e_α^1	e_α^1
S_2	e_γ^2	e_γ^2	e_γ^2	e_γ^2	e_λ^2	e_λ^2	e_γ^2	e_γ^2	e_γ^2	e_γ^2

Etats synchrones

Figure 2.4

Un changement d'état peut intervenir aussi bien au niveau de la première chaîne S_1 que de la deuxième chaîne S_2 .

EXEMPLE

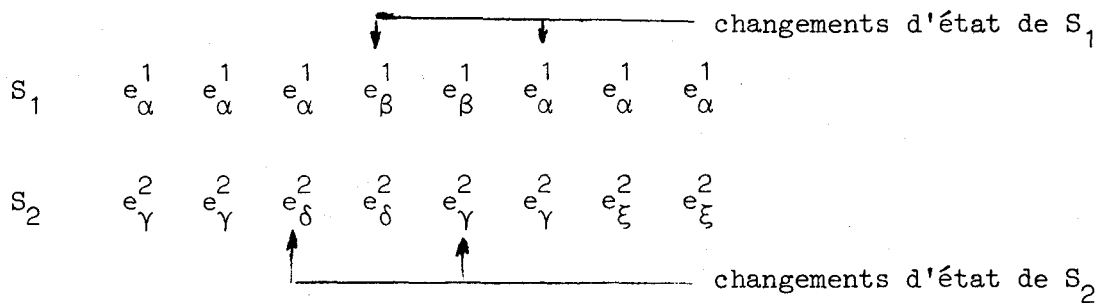


Figure 2.5

Afin d'évaluer un éventuel lien entre ces changements d'états, une observation en est effectuée de deux manières différentes.

En effet on privilégie chaque signal à tour de rôle en supposant qu'il est le signal "maître". Le principe consiste à vérifier cette hypothèse en observant l'évolution du signal maître et les répercussions de ses changements d'états sur l'autre signal.

Ceci conduit à ignorer un changement d'état du signal "esclave" si celui-ci n'intervient pas à la suite d'un changement d'état du signal "maître".

Par exemple à la figure 2.5 si nous considérons que S_1 est le signal "maître" on ne tient pas compte de e_δ^2 mais seulement de e_β^1 et de sa répercussion qui est e_γ^2 .

Les observations sont les suivantes :

$$\begin{array}{cc} e_{\beta}^1 & e_{\gamma}^2 \\ e_{\alpha}^1 & e_{\xi}^2 \end{array}$$

Cet exemple permet de constater que la répercussion d'un changement d'état de S_1 sur S_2 peut se faire soit immédiatement soit avec un retard de n_k cycles.

La valeur de ce retard peut être variable pour une même combinaison de changement d'états donné $e_i^1 e_j^2$.

En conséquence chacune des observations est composée du couple $e_i^1 e_j^2$ ainsi que du nombre de cycles de retard n_k

Il est possible qu'un changement d'état du signal maître ne se répercute pas sur le signal esclave. Afin de tenir compte de ce changement d'état, nous prenons la nouvelle valeur du signal maître combiné à la valeur correspondante du signal esclave.

Cette combinaison se différencie des autres par la valeur de son retard qui n'est pas recensée.

Technique de remplissage du tableau des données

Dans un souci de clarté, quelques exemples sont donnés et recensés dans des tableaux de données qui reflètent les observations effectuées sur les signaux.

Dans ces exemples, on suppose que le signal S_1 est le signal maître.

EXEMPLES

$$\begin{array}{l} S_1 : e_{\alpha}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad \left| \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\alpha}^1 \quad e_{\alpha}^1 \right. \\ S_2 : e_{\beta}^2 \quad e_{\beta}^2 \quad e_{\gamma}^2 \quad e_{\delta}^2 \quad e_{\beta}^2 \quad e_{\beta}^2 \quad e_{\beta}^2 \quad e_{\beta}^2 \quad \left| \quad e_{\beta}^2 \quad e_{\beta}^2 \right| \quad e_{\delta}^2 \quad e_{\delta}^2 \end{array}$$

$n = 2$

Figure 2.6a

$$\begin{array}{l}
 S_1 : e_{\beta}^1 \quad e_{\beta}^1 \quad \left| \quad e_{\alpha}^1 \quad e_{\alpha}^1 \quad \left| \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\beta}^1 \right. \right. \\
 S_2 : e_{\beta}^2 \quad e_{\beta}^2 \quad \left| \quad e_{\gamma}^2 \quad e_{\gamma}^2 \quad \left| \quad e_{\alpha}^2 \quad e_{\alpha}^2 \quad e_{\alpha}^2 \quad e_{\alpha}^2 \quad \right| \quad e_{\delta}^2 \quad e_{\delta}^2
 \end{array}$$

$n = 4$

Figure 2.6b

$$\begin{array}{l}
 S_1 : e_{\beta}^1 \quad \left| \quad e_{\delta}^1 \quad e_{\delta}^1 \quad \left| \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\beta}^1 \quad e_{\beta}^1 \quad \left| \quad e_{\alpha}^1 \quad e_{\alpha}^1 \right. \right. \\
 S_2 : e_{\alpha}^2 \quad \left| \quad e_{\alpha}^2 \quad e_{\alpha}^2 \quad \left| \quad e_{\alpha}^2 \quad e_{\xi}^2 \quad e_{\alpha}^2 \quad e_{\alpha}^2 \quad \left| \quad e_{\alpha}^2 \quad e_{\gamma}^2 \right. \right.
 \end{array}$$

$n=1 \qquad n=1$

Figure 2.6c

Pour chaque changement d'état du signal maître, on attend un changement d'état du signal esclave et on note le nombre de cycles séparant ces deux états.

<u>observations</u>	<u>signal 1</u>	<u>signal 2</u>	<u>retard</u>	
ω_1	e_{β}^1	e_{δ}^2	2	fig.2.6a
ω_2	e_{α}^1	e_{γ}^2	0	} fig.2.6b
ω_3	e_{β}^1	e_{δ}^2	4	
ω_4	e_{δ}^1	x		} fig.2.6c
ω_5	e_{β}^1	e_{ξ}^2	1	
ω_6	e_{α}^1	e_{γ}^2	1	

figure 2.7

Remarque

=====

Il faut noter que dans certains cas, le changement d'état sur S_1 ne s'accompagne pas de changement sur S_2 . Ce cas est illustré sur l'observation ω_4 de la figure 2.6c.

II.4.3 - PRISE EN COMPTE DU RETARD

A chaque combinaison peut correspondre divers retards. Le tableau de la figure 2.7 présente une combinaison possédant 2 retards différents : $e_\beta^1 e_\delta^2$. (retards 2 et 4 cycles sur les observations ω_1 et ω_3).

Le retard est donc une donnée importante au même titre que les combinaisons de changement d'états.

Par exemple une combinaison qui se fait à peu près constamment avec le même retard et qui se retrouve fréquemment donne une excellente indication sur une relation $S_1 \rightarrow S_2$. Par contre, si elle s'échelonne sur une plage importante de valeurs, le lien est moins évident que précédemment.

Afin de tenir compte de ce paramètre important, pour chaque combinaison de changement d'états, un histogramme sur le retard correspondant est élaboré.

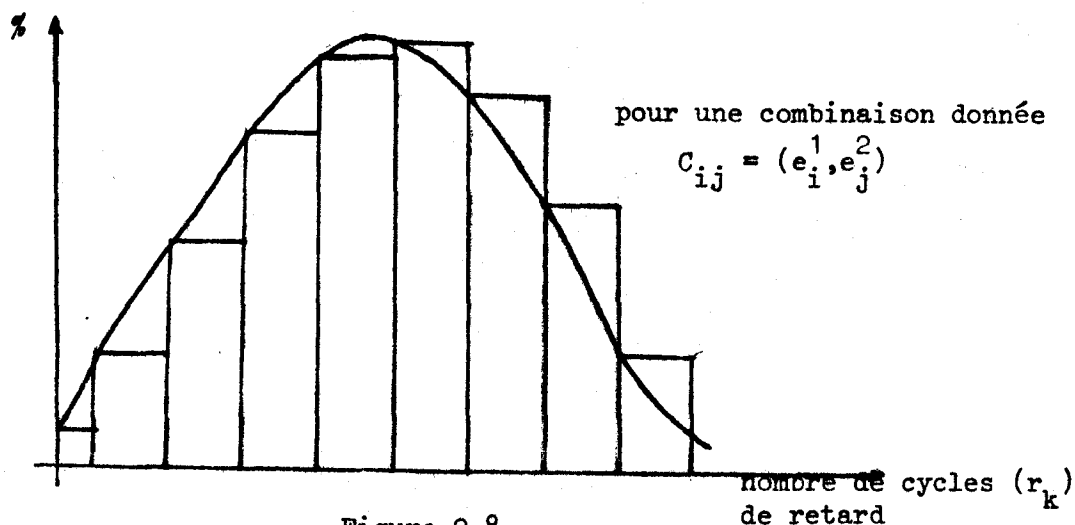


Figure 2.8

A partir du tableau de la figure 2.7 la construction de l'histogramme de distribution des retards pour une combinaison de changements d'états donnée, est immédiate (figure 2.8)

Cette fréquence n'est autre que la probabilité conditionnelle $p(r_k/C_{ij})$, d'obtention du retard r_k sachant que nous avons la combinaison C_{ij}

/35/

$$p(r_k/C_{ij}) = \frac{n(r_k/C_{ij})}{n(C_{ij})} = \frac{\text{nombre de retard } r_k \text{ pour la combinaison } C_{ij}}{\text{nombre de combinaisons } C_{ij}}$$

Par exemple pour le tableau de la figure 2.7
avec $C_{\beta,\delta} = (e_{\beta}^1, e_{\delta}^2)$

$$r_1 = 2$$

$$r_2 = 4$$

$$p(r_1/C_{3,\delta}) = \frac{1}{2}$$

$$p(r_2/C_{\beta,\delta}) = \frac{1}{2}$$

II.4.4 - CALCUL D'ENTROPIE

Pour mesurer le degré de dépendance entre les signaux, le calcul d'entropie relatif aux combinaisons C_{ij} des deux signaux, associés à leurs retards r_k est réalisé comme suit :

$H_{12}(C,R)$: entropie globale relative à toutes les combinaisons et tous les retards, S_1 étant supposé maître et S_2 esclave

$$H_{12}(C,R) = - \sum_{ij} \sum_k p(C_{ij}, r_k) \cdot \log p(C_{ij}, r_k)$$

$p(C_{ij}, r_k)$: probabilité d'obtenir le couple (C_{ij}, r_k)

L'histogramme à deux dimensions de la figure 2.9 donne une illustration de la distribution $p(C_{ij}, r_k)$.

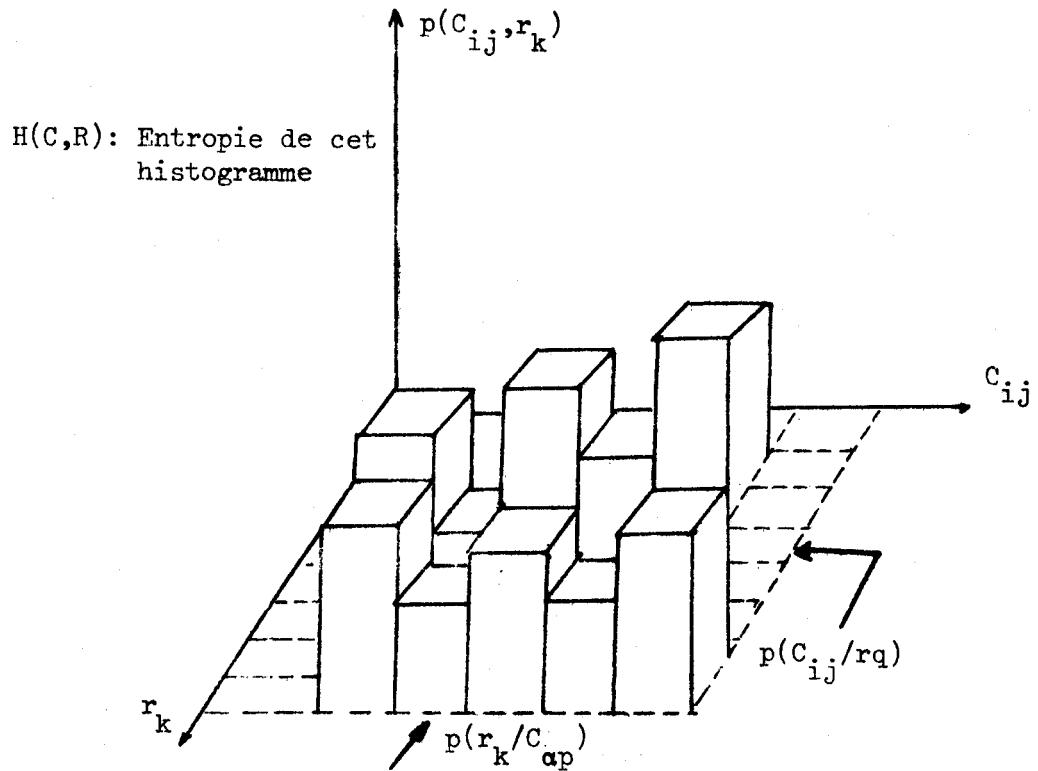


Figure 2.9

Sachant que

$$P(C_{ij}, r_k) = p(C_{ij}) \times p(r_k/C_{ij})$$

$$H_{12}(C,R) = - \sum_{ij} \sum_k p(C_{ij}) \times p(r_k/C_{ij}) \cdot \log (p(C_{ij}) \cdot p(r_k/C_{ij}))$$

$$= - \sum_{ij} \sum_k p(C_{ij}) \times p(r_k/C_{ij}) \log p(C_{ij})$$

$$- \sum_{ij} \sum_k p(r_k/C_{ij}) \times p(C_{ij}) \cdot \log p(r_k/C_{ij})$$

$$H_{12}(C,R) = - \sum_{ij} p(C_{ij}) \cdot \log p(C_{ij}) \sum_k p(r_k/C_{ij})$$

$$- \sum_{ij} p(C_{ij}) \sum_k p(r_k/C_{ij}) \log p(r_k/C_{ij})$$

sachant que $\sum_k p(r_k/C_{ij}) = 1$

en notant $H(r/C_{ij}) = - \sum_k p(r_k/C_{ij}) \cdot \log p(r_k/C_{ij})$

L'expression devient

$$H_{12}(C,R) = - \underbrace{\sum_{ij} p(C_{ij}) \log p(C_{ij})}_{H_{12}(C)} - \underbrace{\sum_y p(C_{ij}) \times H(r/C_{ij})}_{H_{12}(R)}$$

que l'on note

$$H_{12}(C,R) = H_{12}(C) + H_{12}(R)$$

$H(C)$ correspond à l'entropie de la projection de l'histogramme de la figure 2.9 sur le plan $[C_{ij}, p(C_{ij}, r_k)]$

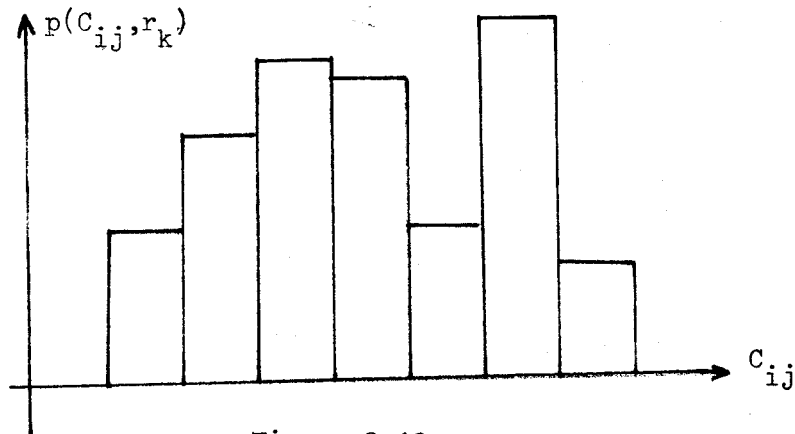
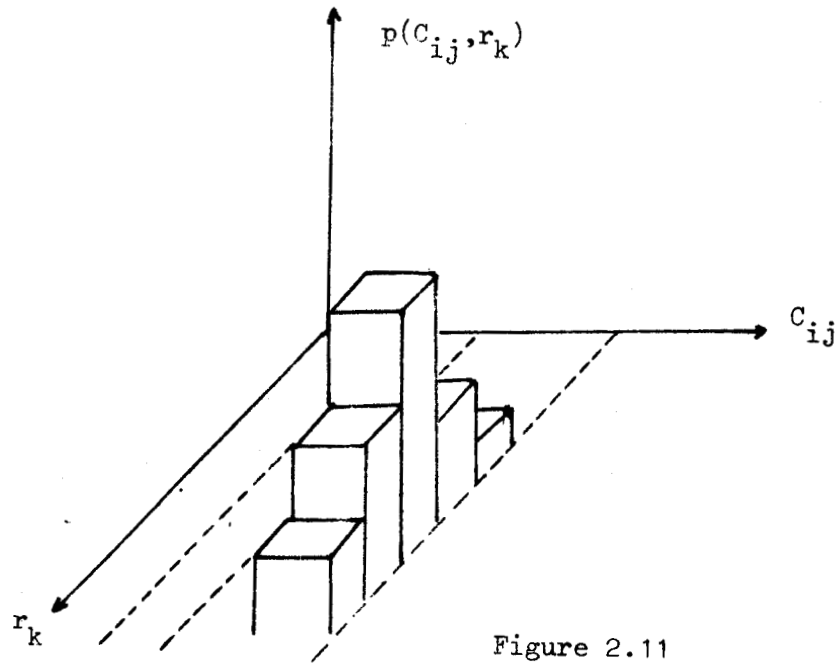


Figure 2.10

Les $H(r/C_{ij})$ sont les entropies des coupes de l'histogramme de la figure 2.9 dans le plan $(r_k, p(C_{ij}, r_k))$. La figure 2.11 illustre une coupe d'histogramme donc l'entropie $H(r/C_{ij})$.



et $H_{12}(R)$ est formé de la somme des entropies de ces histogrammes (coupes d'histogrammes), pondérées par les probabilités d'apparition des combinaisons.

Le calcul de $H_{12}(C, R)$ nécessite la construction dans un premier temps d'un tableau où l'on recense les différentes entropies $H(r/C_{ij})$ et les probabilités d'occurrence des combinaisons.

	e_1^1	e_2^1	...	e_n^1
e_1^2	$H(r/C_{11})$	$H(r/C_{11})$	---	$H(r/C_{n1})$
e_2^2	$H(r/C_{12})$	$H(r/C_{22})$	---	---
...	---	---	---	---
e_1^2	$H(r/C_{12})$	$H(r/C_{21})$		$H(r/C)$

Figure 2.12

L'obtention des probabilités $p(C_{ij})$ est immédiate

$$p(C_{ij}) = \frac{\text{nombre d'apparition de } C_{ij}}{\text{nombre total de combinaisons}}$$

	e_1^1	e_2^1		e_n^1
e_1^2	$p(1,1)$	$p(2,1)$	-----	-----
e_2^2	$p(1,2)$	-----	-----	-
---	-----	-----	-----	-
e_1^2	-	-		

Figure 2.13

La valeur de $H(C,R)$ donne une mesure de l'influence des changements d'état du signal maître sur le signal esclave (causalité).

Si cette entropie est faible, elle indique une bonne dépendance du signal esclave par rapport au signal maître.

La même analyse doit être faite en privilégiant le signal S_2 afin d'évaluer son influence sur S_1 .

La démarche dans ces conditions sera la même que précédemment où l'on devra permutter les rôles de S_1 et de S_2 .

L'interprétation des résultats se fait alors par comparaison des entropies obtenues dans les deux hypothèses.

soit $H_{12}(C,R)$: entropie relative à la première hypothèse
(S_1 signal maître)

et $H_{21}(C,R)$ entropie relative à la deuxième hypothèse
(S_2 signal maître)

Les valeurs de H_{12} (C,R) et H_{21} (C,R) nous renseignent sur l'existence d'un lien entre les signaux S_1 et S_2 .

II.4.5 - COMPARAISON DES ENTROPIES H_{12} ET H_{21}

Le résultat de cette comparaison est illustré dans le tableau ci-dessous.

$H_{21} \backslash H_{12}$	0	a	∞
0	C tester $ H_{12} - H_{21} $ par rapport à k	B_2 S_2 maître	
a	B_1 S_1 maître	A pas de lien	
∞			

a est un seuil en deça duquel les entropies sont considérées comme faible.

- A : les deux entropies H_{12} et H_{21} sont supérieures à un seuil prédéterminé $a > 0$. Dans ce cas il n'y a pas de lien entre les deux suites étudiées.
- B_1, B_2 : l'une des deux entropies est supérieure à a, l'autre inférieure à a. Dans ce cas c'est le signal d'entropie inférieure à a qui est maître.

B_2 : S_2 maître

B_1 : S_1 maître

- C Les deux entropies sont inférieures à a. Dans ce cas il faut poursuivre le test de la manière suivante.

* $| H_{12} - H_{21} | \leq k$ k seuil prédéterminé. Alors

l'interdépendance entre S_1 et S_2 s'effectue dans les deux sens

$S_1 \rightarrow S_2$ et $S_2 \rightarrow S_1$

* $| H_{12} - H_{21} | \geq k.$

Dans ce cas c'est l'entropie la plus petite qui détermine le signal maître.

II.4.6 - EXEMPLE D'ILLUSTRATION

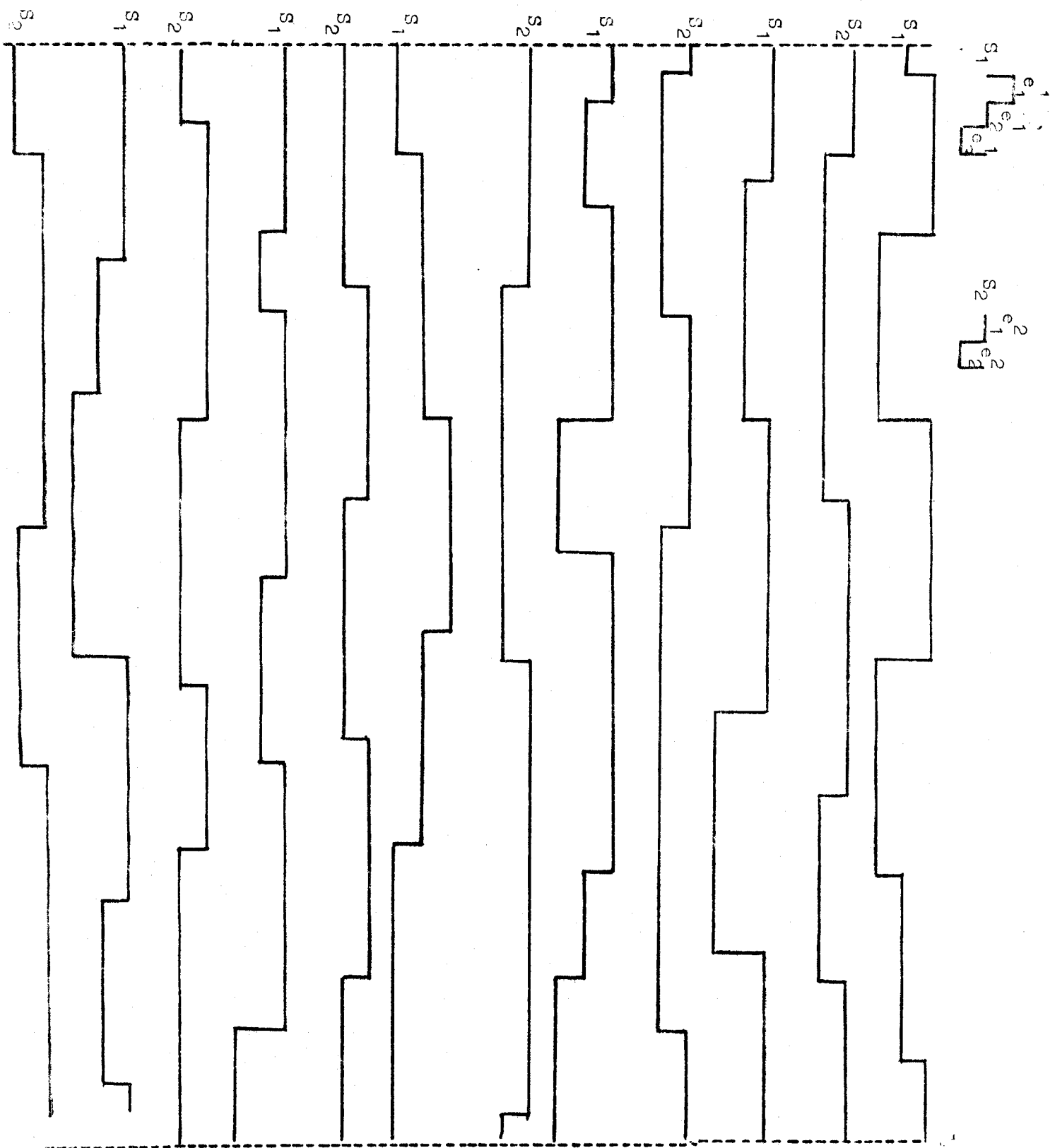
Soient 2 signaux S_1 et S_2 représentant l'évolution dans le temps de deux variables issues du même processus.

Chaque signal peut prendre différents états

Le signal S_1 peut avoir trois états e_1^1, e_2^1, e_3^1

Le signal S_2 peut avoir deux états e_1^2, e_2^2

L'évolution de ces deux signaux dans le temps peut être représentée par deux suites synchrones illustrées figure 2.14



BUS
LINE

Figure 2.14

La recherche de l'existence ou non d'un lien entre ces deux signaux conduit à deux hypothèses exclusives l'une de l'autre.

1°) S_1 est le signal maître et S_2 le signal esclave

2°) S_2 est le signal maître et S_1 le signal esclave.

1°) Hypothèse où S_1 est le signal maître

Le remplissage du tableau de données est fait en privilégiant le signal S_1 , ainsi qu'il a été indiqué aux paragraphes précédents.

Combinaison	observation	signal 1	signal 2	retard	nombre de combinaisons
C_{11}	ω_1	e_1^1	e_1^2	2	3
C_{11}	ω_2	e_1^1	e_1^2	3	3
C_{12}	ω_3	e_1^1	e_2^2	2	4
C_{12}	ω_4	e_1^1	e_2^2	3	3
C_{21}	ω_5	e_2^1	e_1^2	3	3
C_{21}	ω_6	e_2^1	e_1^2	4	2
C_{21}	ω_7	e_2^1	e_2^2	--	4
C_{32}	ω_8	e_3^1	e_2^2	4	4
C_{32}	ω_9	e_3^1	e_2^2	--	4

Figure 2.1

Remarque

=====

Toutes les combinaisons possibles entre les différents états de S_1 et de S_2 n'apparaissent pas au complet dans le tableau des observations. Il n'y a que la partie observée qui est recensée.

CALCUL DES PROBABILITES

1°) des combinaisons de changements d'états.

$$P(C_{11}) = \frac{6}{30} = 0.20$$

$$P(C_{12}) = \frac{7}{30} = 0.23$$

$$P(C_{21}) = \frac{9}{30} = 0.30$$

$$P(C_{32}) = \frac{8}{30} = 0.27$$

e_j^2	e_i^1	e_1^1	e_2^1	e_3^1	
e_1^2		0.20	0.30	-	
e_2^2		0.23	-	0.27	

$P(C_{ij})$

2°) des retards de ces combinaisons.

pour C_{11}

$$P(r=2/C_{11}) = \frac{3}{6} = 0.5$$

$$P(r=3/C_{11}) = \frac{3}{6} = 0.5$$

pour C_{21}

$$P(r=3/C_{21}) = \frac{3}{9} = 0.33$$

$$P(r=4/C_{21}) = \frac{2}{9} = 0.22$$

pour C_{12}

$$P(r=2/C_{12}) = \frac{4}{7} = 0.57$$

$$P(r=3/C_{12}) = \frac{3}{7} = 0.42$$

pour C_{32}

$$P(r=4/C_{32}) = \frac{4}{8} = 0.5$$

CALCUL DES ENTROPIES

1°) des combinaisons de chargement d'états

$$H_{12}(C) = - \sum_{ij} P(C_{ij}) \log_2 P(C_{ij}) = - \sum_{i=1}^3 \sum_{j=1}^2 p(C_{ij}) \log_2 p(C_{ij})$$

$$H_{12}(C) = - p(C_{11}) \log_2 p(C_{11}) - p(C_{12}) \log_2 p(C_{12}) - p(C_{21}) \log_2 p(C_{21}) \\ - p(C_{32}) \log_2 p(C_{32})$$

Les calculs conduisent à la valeur suivante

$$H_{12}(C) = H(C_{11}) + H(C_{21}) + H(C_{32}) + H(C_{12}) \\ = 0.46 \quad + 0.52 \quad + 0.50 \quad + 0.48$$

$$H_{12}(C) = 1.01$$

2°) des retards de ces combinaisons

on recense dans le tableau suivant les $\sum_k H(r_k/C_{ij}) = H(R/C_{ij})$

$\begin{matrix} e_i^1 \\ e_j^2 \end{matrix}$	e_1^1	e_2^1	e_3^1
e_1^2	1	1.01052	-
e_2^2	0.985269	-	0.499967

$$(C_{ij}) = e_i^1 e_j^2$$

$$H(R/C_{ij})$$

$$H_{12}(R) = \sum_y H_{12}(R_k / C_{ij}) \cdot p(C_{ij}) =$$

tout calcul fait nous obtenons :

$$H_{12}(R) = 0,86$$

comme

$$H_{12}(C,R) = H_{12}(C) + H_{12}(R) \quad \text{nous avons } H_{12}(C,R) = 0,86 + 1,01 = 1,87$$

2°) Hypothèse où S_2 est le signal maître

Le remplissage du tableau de données est fait en privilégiant le signal S_2 . (voir figure 2.16)

Le CALCUL DE PROBABILITE

1°) des combinaisons de changements d'états

$$P(C_{11}) = 4/23$$

$$P(C_{21}) = 5/23$$

$$P(C_{31}) = 2/23$$

$$P(C_{12}) = 3/23$$

$$P(C_{22}) = 5/23$$

$$P(C_{32}) = 4/23$$

	e_i^1	e_1^1	e_2^1	e_3^1
e_j^2				
$P(C_{ij})$	e_1^2	4/23	5/23	2/23
	e_2^2	3/23	5/23	4/23

2°) des retards de ces combinaisons

pour C_{11}

$$P(r=2/C_{11}) = \frac{2}{4}$$

$$P(r=3/C_{11}) = \frac{1}{4}$$

$$P(r=1/C_{11}) = \frac{1}{4}$$

pour C_{31}

$$P(r=3/C_{31}) = \frac{1}{2}$$

$$P(r=5/C_{31}) = \frac{1}{2}$$

Combinaison	Observation	Signal 1	Signal 2	Retard	Nombre de combinaisons
C_{11}	ω_1	e_1^1	e_1^2	2	2
C_{11}	ω_2	e_1^1	e_1^2	3	1
C_{11}	ω_3	e_1^1	e_1^2	4	1
C_{21}	ω_4	e_2^1	e_1^2	7	1
C_{21}	ω_5	e_2^1	e_1^2	5	1
C_{21}	ω_6	e_2^1	e_1^2	4	1
C_{21}	ω_7	e_2^1	e_1^2	3	2
C_{31}	ω_8	e_3^1	e_1^2	3	1
C_{31}	ω_9	e_3^1	e_1^2	5	1
C_{12}	ω_{10}	e_1^1	e_2^2	2	1
C_{12}	ω_{11}	e_1^1	e_2^2	3	1
C_{12}	ω_{12}	e_1^1	e_2^2	5	1
C_{22}	ω_{13}	e_2^1	e_2^2	2	1
C_{22}	ω_{14}	e_2^1	e_2^2	3	1
C_{22}	ω_{15}	e_2^1	e_2^2	4	2
C_{22}	ω_{16}	e_2^1	e_2^2	5	1
C_{32}	ω_{17}	e_3^1	e_2^2	2	1
C_{32}	ω_{18}	e_3^1	e_2^2	4	1
C_{32}	ω_{19}	e_3^1	e_2^2	6	2

Figure 2.16

pour C_{21}

$$P(r=7/C_{21}) = \frac{1}{5}$$

$$P(r=5/C_{21}) = \frac{1}{5}$$

$$P(r=4/C_{21}) = \frac{1}{5}$$

$$P(r=3/C_{21}) = \frac{2}{5}$$

pour C_{12}

$$P(r=2/C_{12}) = \frac{1}{3}$$

$$P(r=3/C_{12}) = \frac{1}{3}$$

$$P(r=5/C_{12}) = \frac{1}{3}$$

pour C_{22}

$$P(r=2/C_{22}) = \frac{1}{5}$$

$$P(r=3/C_{22}) = \frac{1}{5}$$

$$P(r=4/C_{22}) = \frac{2}{5}$$

$$P(r=1/C_{22}) = \frac{1}{5}$$

pour C_{32}

$$P(r=2/C_{32}) = \frac{1}{4}$$

$$P(r=4/C_{32}) = \frac{1}{4}$$

$$P(r=6/C_{32}) = \frac{2}{4}$$

CALCUL DES ENTROPIES

1°) des combinaisons de changements d'états

$$H_{21}(C) = - \sum_{ij} P(C_{ij}) \log_2 P(C_{ij}) = \sum_{ij} H(C_{ij})$$

$$H_{21}(C) = H(C_{11}) + H(C_{12}) + H(C_{21}) + H(C_{22}) + H(C_{31}) + H(C_{32})$$

$$= 0,44 + 0,38 + 0,48 + 0,48 + 0,31 + 0,44$$

$$H_{21}(C) = 2,53$$

2°) des retards de ces combinaisons

on recense dans le tableau suivant les $\sum_k H(r_k/C_{ij}) = H(R/C_{ij})$

e_j^2	e_i^1	e_1^1	e_2^1	e_3^1	(C_{ij})
e_1^2	:	1.5	1.92193	1	$H(R/C_{ij})$
e_2^2	:	1.58496	1.92193	1.5	

$$H_{21}(R) = \sum_y H(R/C_{ij}) \cdot P(C_{ij})$$

tout calcul fait, nous obtenons

$$H_{21}(R) = 1,65$$

comme

$$\begin{aligned} H_{21}(C,R) &= H_{21}(C) + H_{21}(R) \\ &= 2,53 + 1,65 = 4,18 \end{aligned}$$

Commentaires des résultats

Niveau des retards

Dans le cas n°1, recherche de $H_{12}(C,R)$, l'histogramme des retards présente un pic. Cela se répercute sur la valeur de son entropie qui est faible. Elle est inférieure à 1

$$H_{12}(R) = 0,86$$

Par contre dans les cas n°2, l'histogramme des retards est diffus. Son entropie est donc plus conséquente que dans le cas précédent. Elle est supérieure à 1

$$H_{21}(R) = 1,65$$

Niveau des combinaisons

De même qu'en ce qui concerne les retards, l'histogramme des combinaisons est centré sur des valeurs bien définies. Cette propriété se traduit par une faible entropie.

Au contraire l'histogramme associé à la seconde hypothèse contient de nombreuses raies correspondant à de faibles probabilités d'apparition des combinaisons. Ainsi dans ce cas, l'entropie devient nettement supérieure à un.

CONCLUSION

Compte-tenu des comparaisons effectuées respectivement entre $H_{12}(C)$ et $H_{21}(C)$ d'une part, puis $H_{12}(R)$ et $H_{21}(R)$ d'autre part, il apparaît ainsi clairement que S_1 est le signal maître. Autrement dit S_1 apparaît comme étant la cause de S_2 .

L'intérêt de la méthode présentée réside dans la possibilité de mesurer ce lien de causalité.

II.4.7. - ALGORITHME

Ce paragraphe a pour objet de donner un aperçu des algorithmes mis en jeu pour la résolution du problème traité dans ce chapitre.

7.1 - Tableau de données

Cet algorithme de création de tableau de données a pour but d'enregistrer les combinaisons relatives aux changements d'états et leurs retards associés, et ceci dans notre champ d'observation.

Ce dernier n'est autre qu'une série de fichiers sur lesquelles sont stockés les différents états synchronisés des deux signaux S_1 et S_2 , extraits par l'algorithme décrit dans le chapitre 1.

Tant qu'il y a des valeurs à traiter, le traitement est effectué.

La variable booléenne "fini" indique la fin d'apparition des données. cette variable est donc testée pour arrêter le traitement.

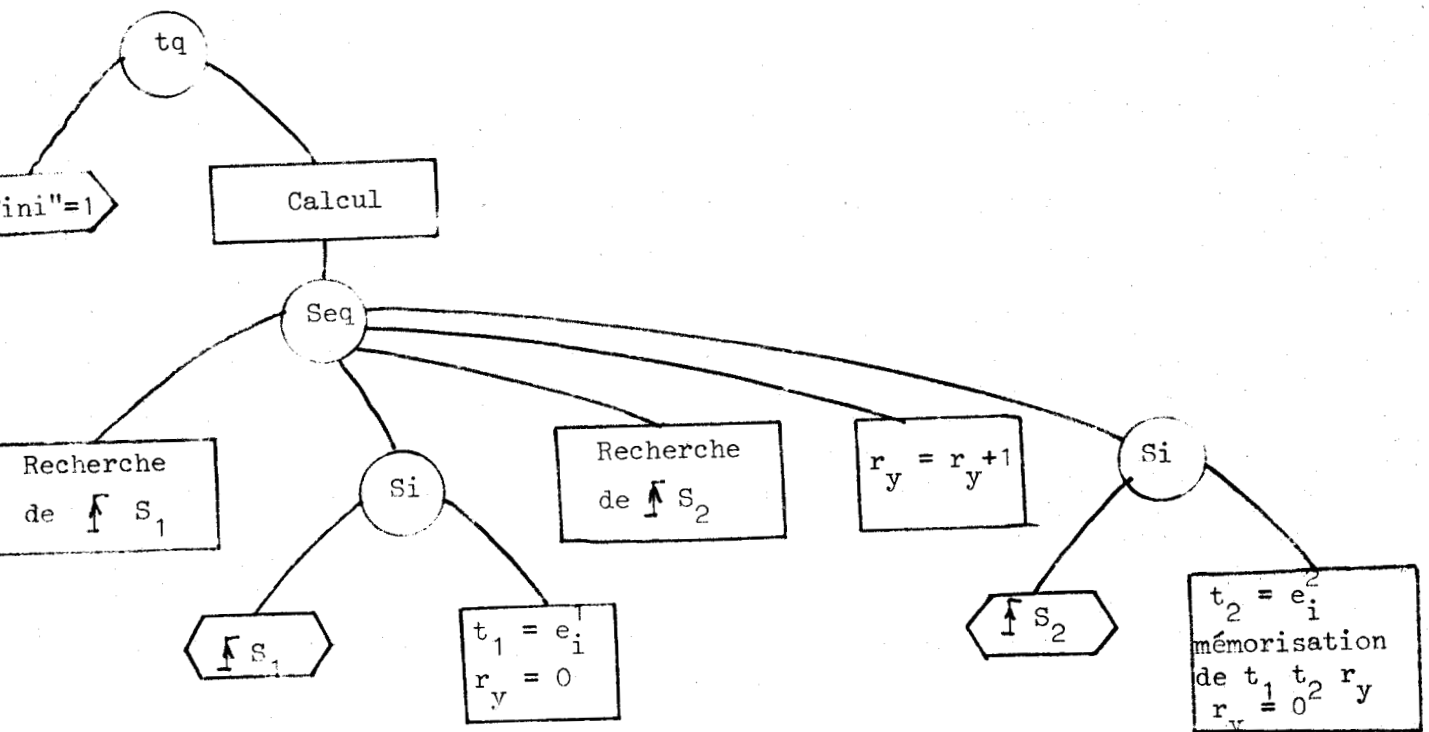


Figure 2.19

$\mathcal{F}(S_1)$ ($\mathcal{F}(S_2)$) symbolise le changement d'état intervenant au niveau du signal S_1 (S_2)

. La recherche de $\mathcal{F}(S_1)$ ($\mathcal{F}(S_2)$) est un traitement évident, permettant de détecter que l'état à l'instant $j+1$ est différent de l'état à l'instant j

. t_1 (t_2) est la variable qui mémorise l'état changé.

r_y : est le retard pris par $\mathcal{F}(S_2)$ vis à vis de $\mathcal{F}(S_1)$

Il est à remarquer que cet algorithme ne tient compte que des couples $\{S_1\}$ et $\{S_2\}$. S'il n'y a pas de $\{S_2\}$ à un $\{S_1\}$, la valeur présente de $\{S_1\}$ sera effacée au profit de la nouvelle valeur et le retard r_y par conséquent est remis à zéro.

Afin d'avoir le retard exact, on incrémente la variable r_y , l'algorithme décrit à la figure 2 est celui correspondant à l'hypothèse S_1 "signal maître".

Dans l'autre cas, l'algorithme est similaire.

7.2 - Tableau des probabilités et des entropies

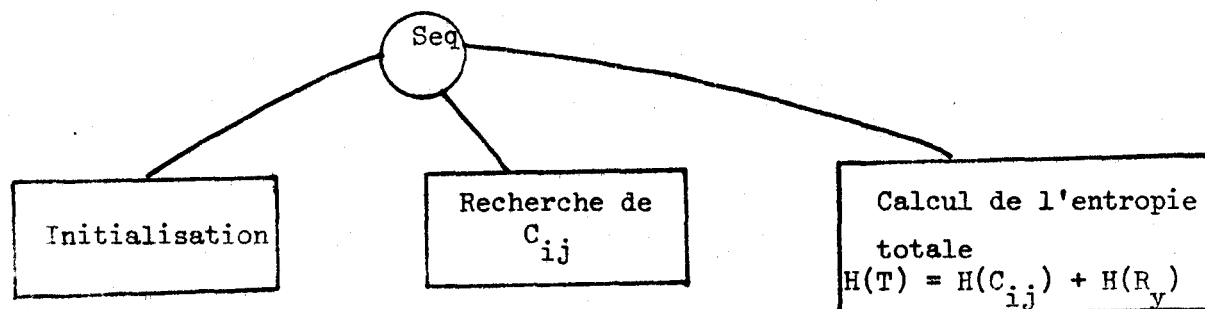
L'algorithme qui permet d'obtenir ces probabilités, se sert de différentes variables toutes utiles à ce calcul d'entropie.

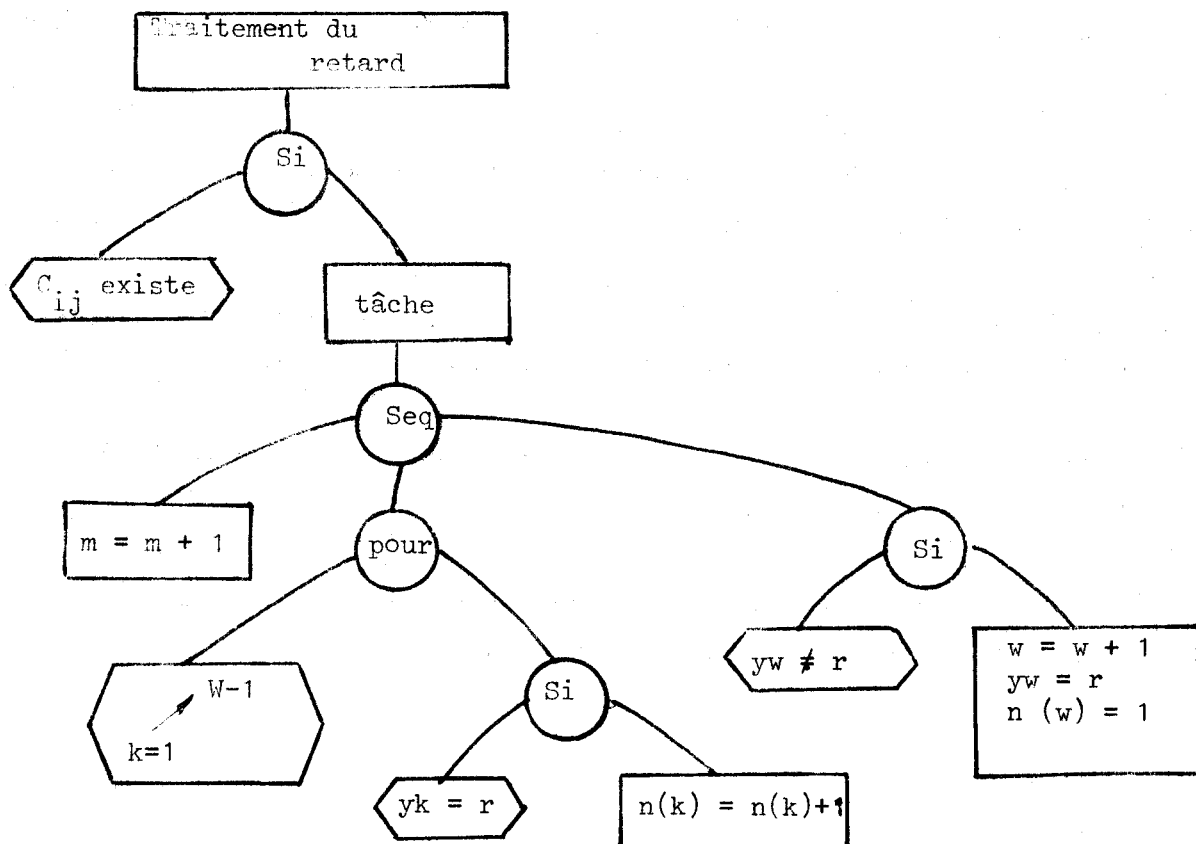
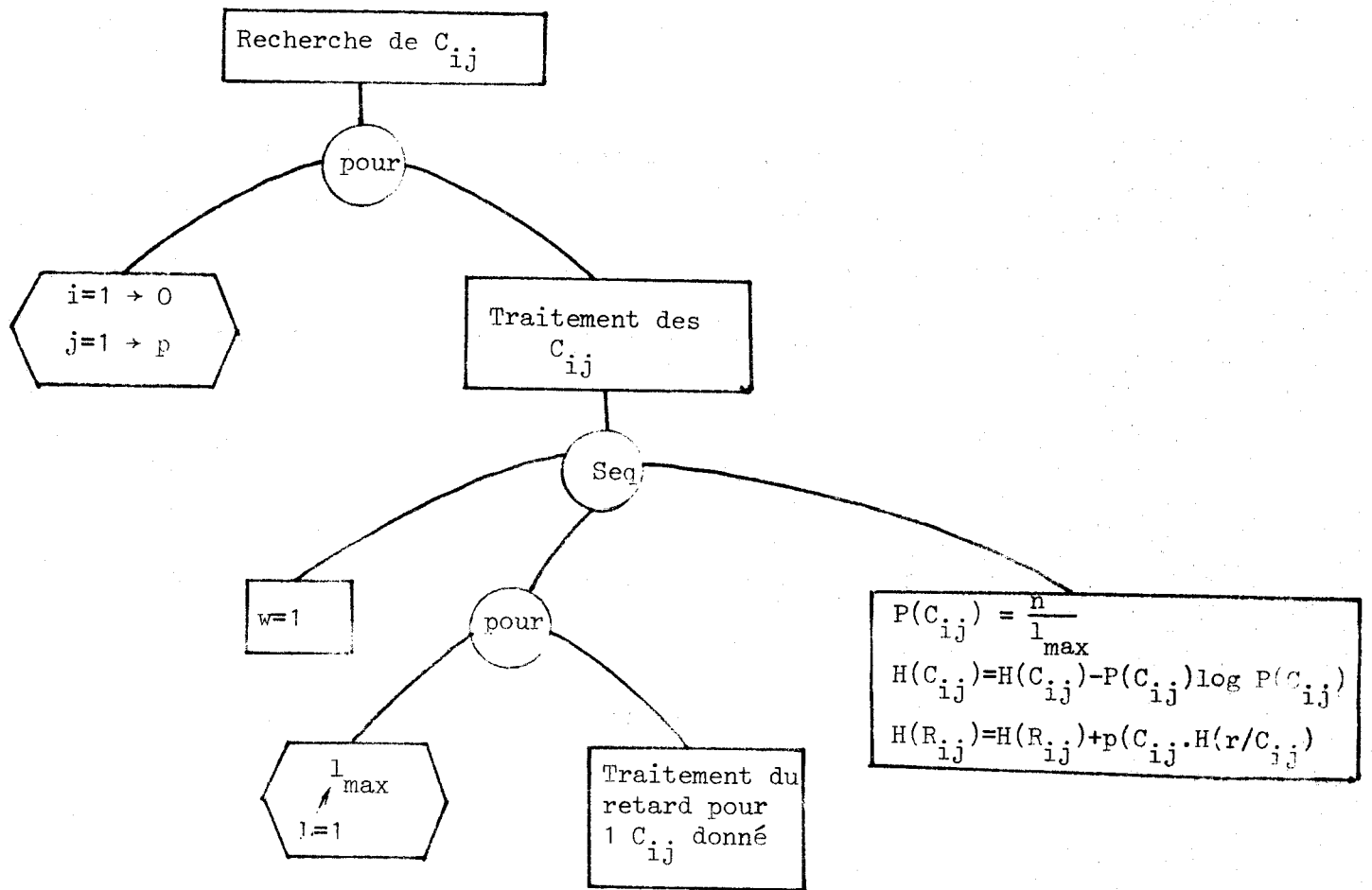
Cet algorithme traite un fichier contenant les combinaisons de changement d'états $C_{ij} = e_i^1 e_j^2$ avec leur retard associé.

La première tâche effectuée est la reconnaissance dans le fichier des diverses combinaisons.

Pour chaque combinaison, le fichier est passé en revue. Donc de la valeur 1 à 1 max qui est le nombre total d'éléments du fichier, on recherche une combinaison C_{ij} pour laquelle on sélectionne les divers retards y_1 à y_r afin de calculer les probabilités du retard.

Puis cette opération de reconnaissance de C_{ij} est faite jusqu'à épuisement des combinaisons, c'est-à-dire pour les valeurs de i et de j variant suivant leurs domaines respectifs.





k : variable muette qui permet de passer en revue tous les retards
 m : nombre de C_{ij} dans le fichier
 r : retard associé aux combinaisons dans le fichier
 l : variable qui indique le nombre d'éléments du fichier pris en compte
sa valeur maximale est $l_{\max}$.
 $y_1, y_2 \dots y_w$: les différents retards existants pour une combinaison C_{ij}
donnée.

le calcul des $H(R_{ij})$ et des $H(C_{ij})$ se cumule au fur et à mesure de
la recherche des C_{ij} .

Cet algorithme permet donc d'avoir l'entropie totale liant le
signal S_1 et S_2 .

II.5 - CONCLUSION

Le traitement décrit dans ce chapitre a pour objectif de sélectionner
un nombre minimum de signaux à prélever sur un système afin d'étudier totalement
son comportement.

L'outil de base employé, l'entropie, a permis de mesurer les liens de
causalité établis entre divers signaux.

Généralement la théorie de l'information est appliquée à des tableaux
de contingence et non pas à des valeurs évoluant dans le temps.

Malgré cette évolution, nous envisageons la possibilité de créer un
tableau de données recensant tous les cas.

La formation de ce tableau de données nécessite un soin particulier
et demande un nombre considérable d'observations.

De plus les observations doivent être faites d'une manière choisie et non pas anarchique de façon à balayer tous les états possibles que peut prendre un signal. Ceci demande un travail de recensement assez long mais nécessaire.

Le programme utilisé dans ce chapitre contrairement aux programmes présentés dans ce mémoire, n'est pas un programme temps réel. Il permet la sélection des signaux qui doivent assurer la surveillance d'un processus dans le temps.

- C H A P I T R E I I I -

ANALYSE PREDICTIVE DE L'EVOLUTION TEMPORELLE

CHAPITRE III

3.1 - INTRODUCTION	page 1
3.2 - NOTION DE SEQUENCES	page 1
3.2.1. - Introduction	page 1
3.2.2. - Notion de séquences	page 3
3.2.3. - Propriétés des séquences	page 5
3.3 - EVOLUTION DANS LE TEMPS DES SEQUENCES	page 8
3.3.1. - Introduction	page 8
3.3.2. - Interprétation géométrique	page 8
3.3.3. - Etude de la trajectoire	page 10
3.3.4. - Paramètres caractéristiques de la trajectoire	page 11
3.3.5. - Quelques cas types	page 12
3.3.6. - Modélisation markovienne	page 13
3.3.7. - Algorithme	page 18
CONCLUSION	page 23

3.1 - INTRODUCTION

Après avoir, dans le chapitre précédent mis en évidence les signaux les plus significatifs du fonctionnement d'un processus, nous vous proposons, dans ce chapitre, d'étudier l'évolution dans le temps de ces signaux.

Pour cela on effectue une surveillance sur une période longue afin de prendre en compte toutes les fluctuations. /42/

Cette surveillance a pour but de fournir à l'expérimentateur une information sur l'état général vers lequel le processus a tendance à évoluer.

Il est clair, que dans un tel contexte, l'analyse ne peut pas rester au niveau du cycle proprement dit, mais elle doit traiter une information beaucoup plus synthétique afin de mettre en évidence les évolutions temporelles.

Dans ce but, nous nous appliquons à introduire des notions permettant une étude macroscopique du signal. En particulier, nous cherchons à définir des quantités permettant de réaliser une prédiction et donc de préciser les tendances d'évolution du processus.

3.2 - NOTION DE SÉQUENCES

3.2.1 - INTRODUCTION

La méthode décrite au chapitre 1 a permis d'extraire, des signaux traités, des informations au rythme du processus qui les a engendrées (cycles successifs).

Ces cycles sont ensuite reconnus comme appartenant à telle ou telle classe de fonctionnement du processus.

La sélection des signaux les plus significatifs proposée au deuxième chapitre permet alors de fournir une suite temporelle de cycles caractérisant les différents états que traverse le processus lors de son évolution.

Un cycle isolé pris hors de son contexte ne peut fournir aucune information sur la dynamique de fonctionnement du processus. Pour obtenir une telle information, il faut lui associer les cycles le précédant et le suivant.

Considérons à titre d'exemple le signal de la figure 3.1

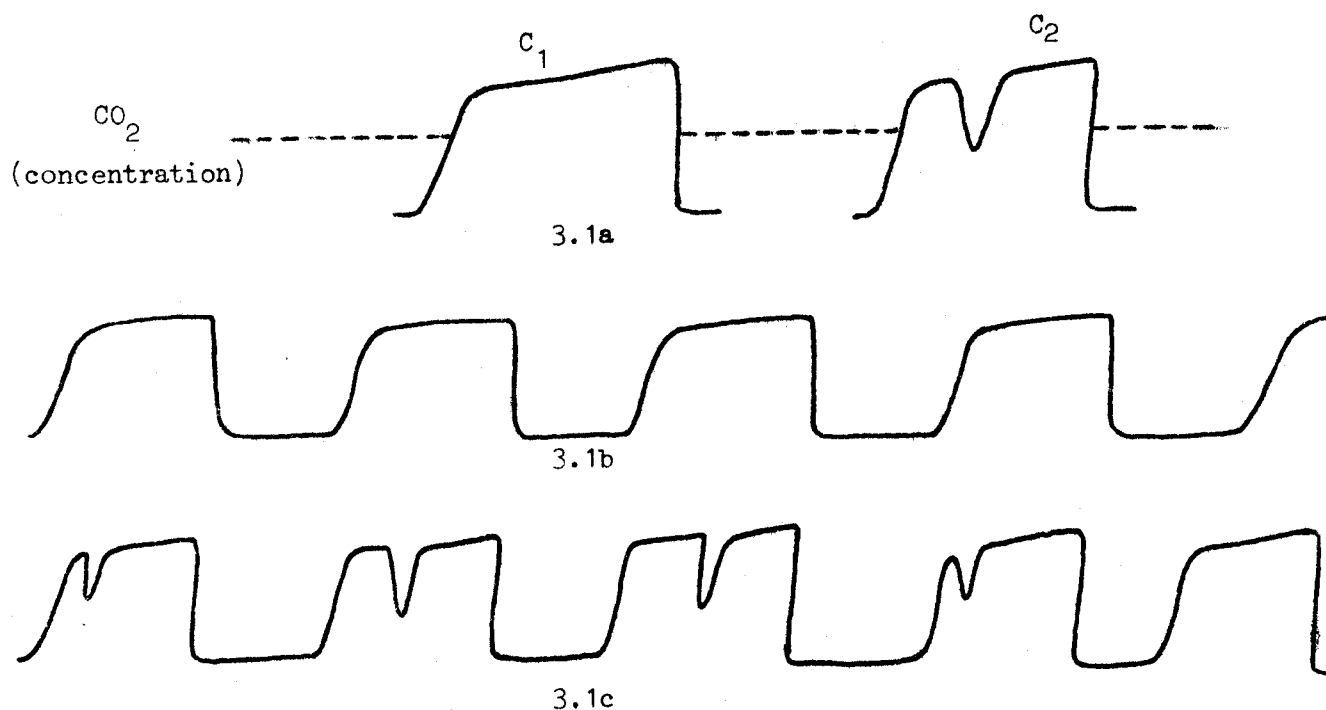


Figure 3.1

Il représente l'évolution de la concentration en CO_2 des gaz expirés par un malade placé en réanimation respiratoire. Le cycle C_1 reflète un fonctionnement normal de la respiration du patient, alors que le cycle C_2 marque une anomalie au niveau du fonctionnement des bronches. Aux figures 3.1b et 3.1c les cycles C_1 et C_2 sont placés dans des contextes différents permettant des interprétations différentes.

Le cas de la figure 3.1b reflète une période calme sans danger d'évolution alors que dans, le cas de la figure 3.1c le cycle C_1 indique une amélioration suivant une période de mauvais fonctionnement.

Pour mesurer cette différence d'interprétation selon le contexte, nous proposons de donner une information condensée donnant l'état global du processus pendant un certain intervalle de temps, appelé temps d'observation.

De plus on applique sur l'ensemble des états émis par le processus une relation d'ordre définie en fonction des besoins de l'analyse.

3.2.2. - Notion de séquences

Définition 1

Dans l'ensemble des états, on distingue 2 groupes :

- *états statiques*
- *états dynamiques*

Un état statique reflète une situation stable correspondant au fonctionnement normal du processus.

Un état dynamique au contraire montre une évolution du fonctionnement normal vers un accident.

Exemple

En ce qui concerne le signal de concentration de CO_2 d'un malade en réanimation respiratoire, le cas normal d'un être sain diffère de celui d'un asthmatique.

Les formes de signaux engendrés par leurs respirations, bien que différentes peuvent être considérées dans les deux cas comme normales, car elles traduisent l'état d'un malade qui peut être stationnaire sans nécessiter une surveillance intensive.

Par contre certaines formes descriptives d'accidents relatifs à la mécanique ventilatoire peuvent intervenir dans l'évolution du signal et nécessiter l'intervention d'un médecin (agent extérieur au processus).

PROPRIETE

Une suite d'états dynamiques est toujours suivie d'une suite d'états statiques.

En effet une suite d'états dynamiques contient les éléments permettant de prédire un accident éventuel pouvant survenir au niveau du processus.

Le retour à un fonctionnement stable intervient ensuite et spécifie qu'on obtient une suite d'états statiques qu'elle soit similaire à la précédente ou dégradée.

Exemples

=====

Les anomalies apparaissant sur des signaux physiologiques sont suivies par des périodes normales.

1°) dans le cas de l'impédance thoracique chez le nouveau né qui, prise sous certaines conditions, reflète l'activité mécanique pulmonaire du patient, une apnée intervient lors d'un stress respiratoire. Elle est toujours suivie par des périodes calmes : le retour à la normale.

2°) dans le cas de l'ECG, une tachycardie est suivie d'un retour à un rythme normal.

REMARQUE

Lorsque le processus à surveiller tend à se dégrader cela se traduit par l'occurrence d'accidents [états dynamiques], se succédant les uns aux autres parmi des états statiques avec une fréquence plus ou moins grande.

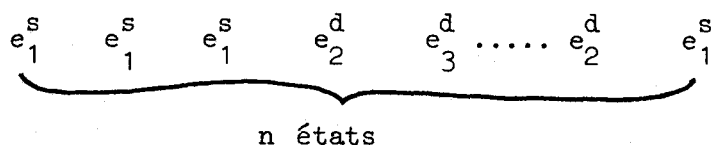
Définition 2

Soit Σ_s une suite d'états avec e_k^s un état appartenant à l'ensemble des états statiques, e_k^d un état appartenant à l'ensemble des états dynamiques. Une séquence est une suite de $n + p$ états quelconques encadrés d'états statiques.

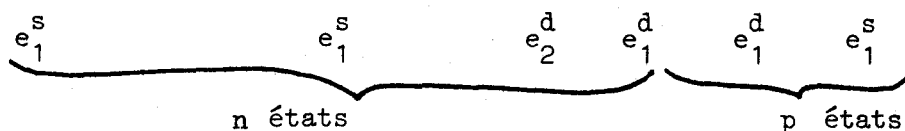
- n est un entier positif fixé permettant d'étudier une fenêtre minimale de Σ_s (effet de filtrage)
- p est un entier positif variable qui permet de tenir compte des états apparaissant après le n ème état, si ces états ne sont pas statiques.

Si le n ème état est un état dynamique, il sera pris en compte jusqu'à l'obtention d'un état statique au $(n + p)$ ème état ainsi que le montrent les exemples suivants :

séquence 1



séquence 2



3.2.3. - Propriétés des séquences

La reconnaissance des séquences peut s'effectuer par analyse syntaxique. Toutefois cette approche oblige à recenser toutes les combinaisons pouvant se produire pour la formation d'un mot de dimensions minimum n prenant ses primitives dans un ensemble d'états fini.

Compte-tenu de la lourdeur de cette démarche, on caractérise une séquence non pas par la suite des cycles la composant, mais par un paramètre quantifiable.

De même il apparaît qu'en général dans l'évolution d'un signal, il peut y avoir beaucoup d'éléments normaux n'apportant pas d'information nouvelle quant à l'évolution du processus.

Par contre un paramètre quantifiable reflétant globalement le fonctionnement peut donner des informations très importantes sur l'évolution (dégradation, amélioration ...).

Pour évaluer ce paramètre quantifiable, une relation d'ordre est définie sur les états.

Définition 3

A chaque état aussi bien statique que dynamique, est affectée une mesure de gravité.

Pour cela, on pondère les états par des poids de plus en plus important, lorsqu'on passe des états statiques aux états dynamiques.

Ce poids affecté aux états est appelé "degré de gravité".

Définition 4

On appelle gravité, d'une séquence la somme des degrés de gravité des états composant la séquence. Soit $X(e_i)$: le degré de gravité d'un état e_i et $\varphi(\Sigma_j)$: gravité de la séquence Σ_j .

$$\text{Alors } \varphi(\Sigma_j) = \sum_{i=1}^{n+p} X(e_i)$$

Ainsi une séquence se différencie d'une autre par sa gravité qui évalue les périodes calmes ou agitées du processus.

REMARQUE 1

La bonne différenciation état statique/état dynamique est obtenue en affectant aux états statiques un degré de gravité négligeable.

REMARQUE 2

Afin que la gravité d'une séquence soit représentative du nombre variable d'états composant une séquence, il importe d'affecter un degré de gravité non nul aux états statiques.

La gravité tient ainsi compte des p éléments de la séquence qui traduisent un accident intervenu au $n^{\text{ième}}$ état de la séquence.

Définition 5

La classification des séquences relativement à leur gravité permet de définir un ensemble de situations S_j selon les règles suivantes :

$$S_1 / \quad \varphi(\Sigma_i) \leq a_1$$

$$S_2 / \quad a_1 < \varphi(\Sigma_i) \leq a_2$$

$$S_3 / \quad a_2 < \varphi(\Sigma_i) \leq a_3$$

$$S_4 / \quad a_3 < \varphi(\Sigma_i) \leq a_4$$

Les seuils a_k définissent l'appartenance d'une séquence à telle ou telle situation selon sa gravité.

3.3 - ÉVOLUTION DANS LE TEMPS DES SÉQUENCES

3.3.1 - INTRODUCTION

La surveillance du processus se ramène à l'observation d'une suite de séquences pouvant traverser un nombre k de situations données, notées $S_1, S_2, \dots, S_k$.

L'étude de l'évolution du processus est maintenant envisagée par le biais des lois de probabilités calculées sur une fenêtre glissante dans le temps comprenant un nombre fixe de séquences ainsi que l'illustre la figure ci-dessous. /40/ /43/

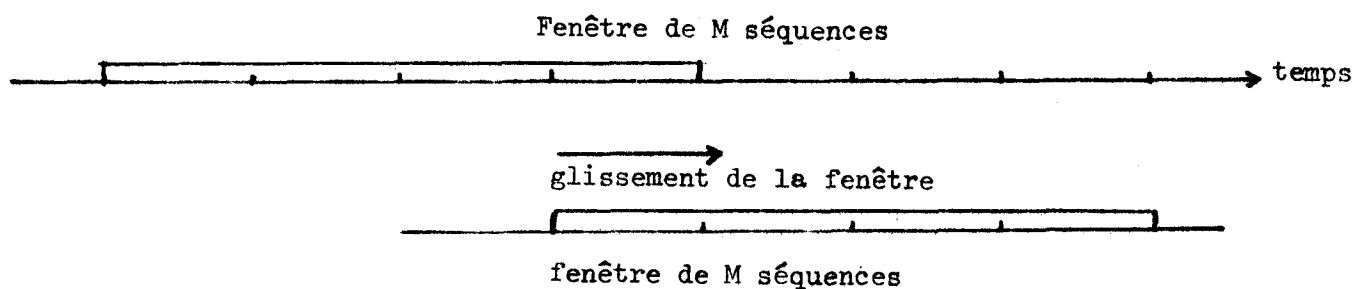


Figure 3.2

le glissement effectué peut être d'une ou plusieurs séquences.

3.3.2 - INTERPRÉTATION GEOMETRIQUE

La distribution des situations $S_1, S_2, \dots, S_k$ à l'intérieur d'une fenêtre donnée peut se traduire par un histogramme dont les composantes sont les $P(S_i)$.
(probabilités d'avoir la situation S_i à l'intérieur de la fenêtre considérée)
(figure 3.3).

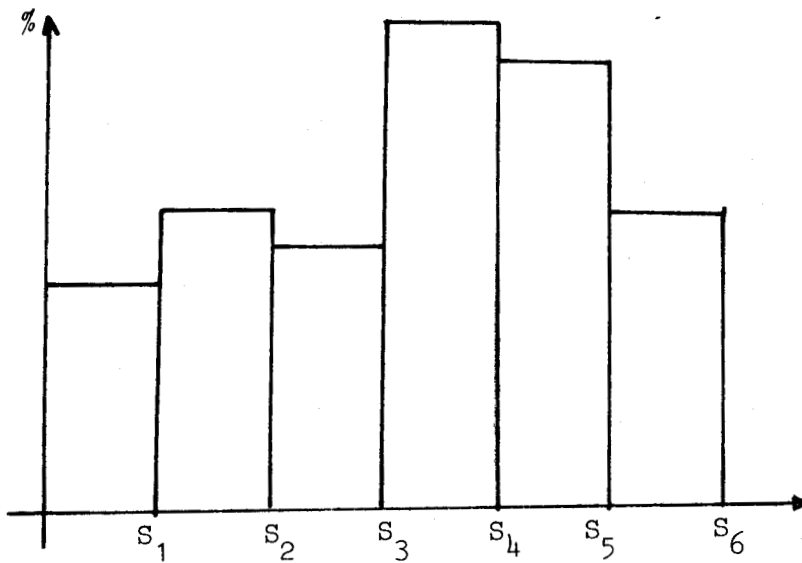


Figure 3.3

Une autre représentation géométrique possible de cette distribution à l'intérieur d'une fenêtre est donnée figure 3.4., pour 3 situations données.

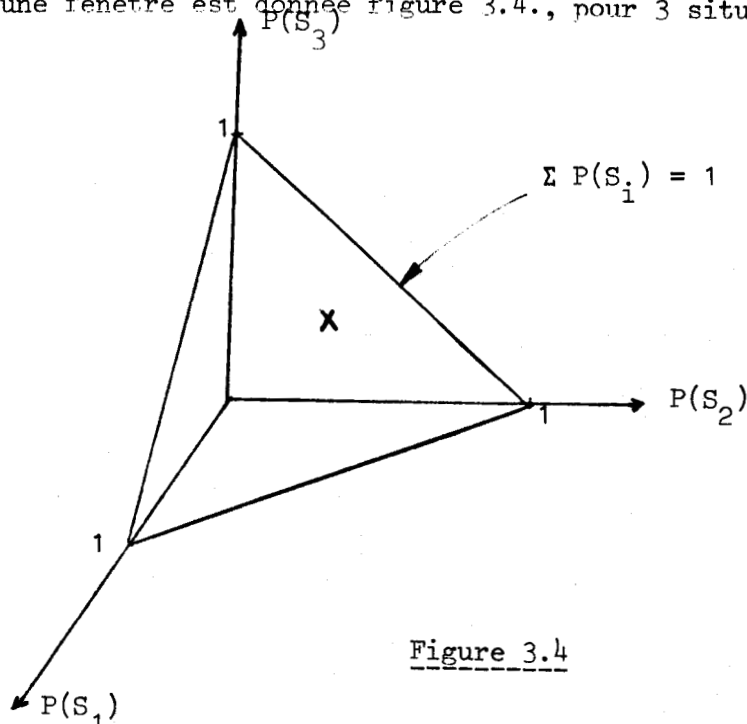


Figure 3.4

Dans cette représentation, on définit un espace Euclidien à k dimensions représentant chacune un $P(S_i)$ $i = 1, k$. Dans ces conditions la distribution des $P(S_i)$ dans la fenêtre considérée a pour image un point de coordonnées les $P(S_i)$ dans l'espace ainsi défini.

Lorsque l'on passe d'une fenêtre d'observation à une autre, le point image des distributions décrit une trajectoire appartenant au plan d'équation $\Sigma P(S_i) = 1$.

3.3.3 - ETUDE DE LA TRAJECTOIRE

L'étude cinématique de la trajectoire du point image permet d'obtenir deux types d'informations.

1°) une information statique qui est la position du point et qui donne la tendance actuelle du processus.

2°) une information dynamique qui est la vitesse du point le long de sa trajectoire qui permet d'évaluer l'évolution du processus vers telle ou telle situation.

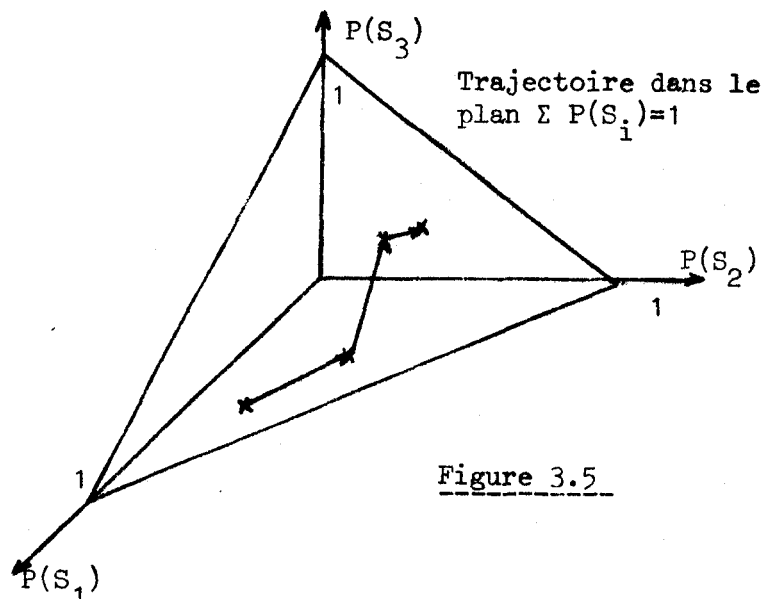
Par ailleurs soit $v^t = \begin{pmatrix} P^t(S_i) \\ P^t(S_k) \end{pmatrix}$ le vecteur correspondant au point image d'une fenêtre donnée et

$v^{t+1} = \begin{pmatrix} P^{t+1}(S_i) \\ P^{t+1}(S_k) \end{pmatrix}$ le vecteur correspondant au point image de la fenêtre immédiatement suivante.

Dans ces conditions $\Delta v = v^{t+1} - v^t$, qui n'est autre que la dérivée discrète du vecteur v , mesure la vitesse du point image.

En d'autre terme, Δv évalue la différence entre deux histogrammes correspondant aux deux fenêtres successives observées.

De plus la succession dans l'espace des vecteurs Δv donne la trajectoire du point image ainsi que l'illustre la figure 3.5 dans le cas de trois situations.



3.3.4 - PARAMETRES CARACTERISTIQUES DE LA TRAJECTOIRE

L'analyse cinématique de la trajectoire ainsi définie permet de prédire les tendances d'évolution du processus vers une des situations S_i ($i = 1, k$) données.

Pour cela, ainsi que l'illustre la figure 3.6, on se propose d'évaluer pour le point P, image de la fenêtre t, les écarts angulaires θ_i entre le vecteur Δv et les directions PM_i telles que M_i désigne le point du plan $\Sigma P(S_i) = 1$ où $P(S_i) = 1$ ($\theta_i \in]-\pi, +\pi[$).

Dans ces conditions, la situation S_d telle que :

$$|\theta_d| = \min_{i=1, k} |\theta_i|$$

désigne la situation privilégiée vers laquelle le processus a tendance à évoluer.

Par ailleurs, la projection de Δv sur PM_d donne la vitesse de cette évolution.

évaluation de θ_d

$$\text{On a : } \cos \theta_i = \frac{\vec{\Delta v} \cdot \vec{PM}_i}{||\vec{\Delta v}|| \cdot ||\vec{PM}_i||}$$

Par ailleurs on peut définir θ_d par la relation

$$\cos \theta_d = \max_{i=1, k} \cos \theta_i$$

évaluation de la projection de Δv sur PM_d

La mesure de cette projection est définie par :

$$W_d = ||\Delta v|| \cdot \cos \theta_d$$

$$W_d = \frac{\vec{\Delta v} \cdot \vec{PM}_d}{||\vec{PM}_d||}$$

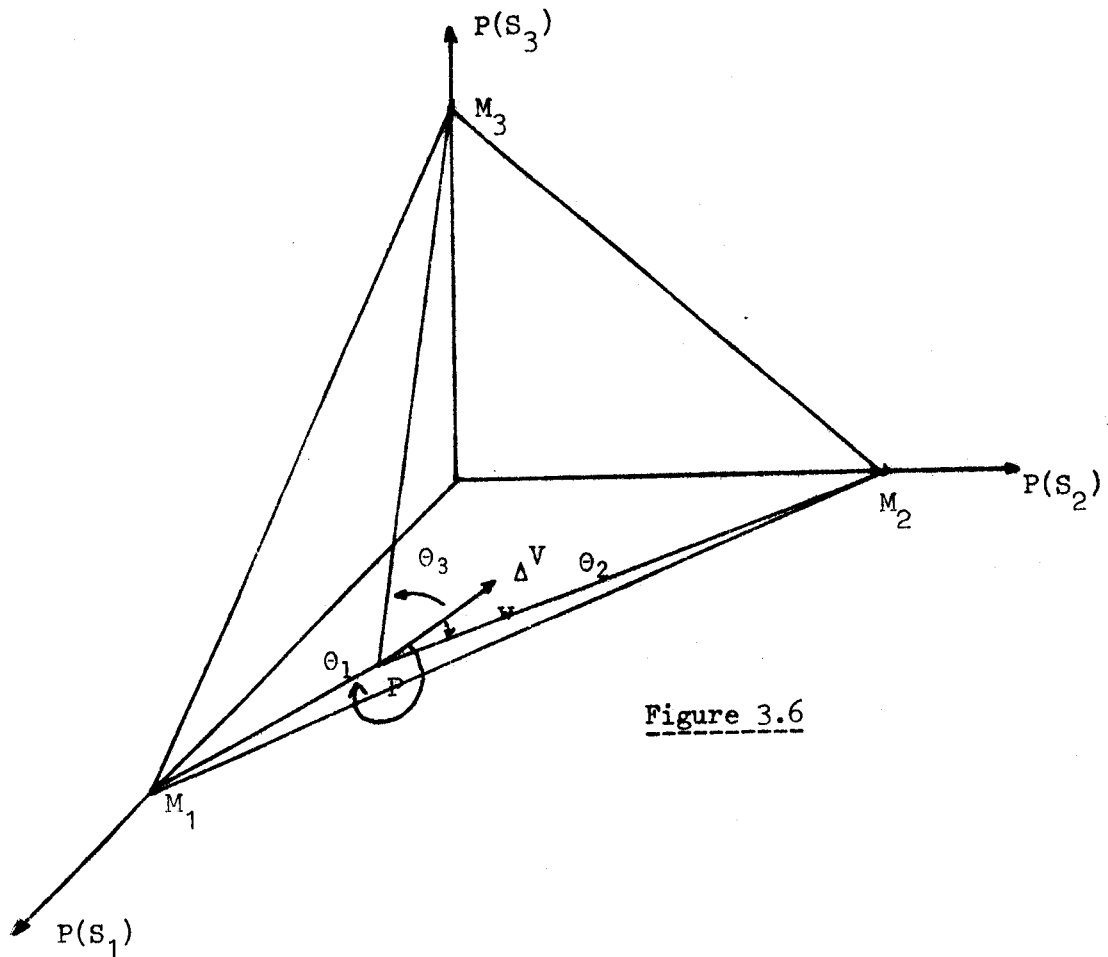


Figure 3.6

3.3.5 - QUELQUES CAS TYPES

Soit à titre d'exemple un processus à trois situations S_1 , S_2 , S_3 classés de la façon suivante :

- S_1 : situation normale
- S_2 : situation intermédiaire
- S_3 : situation grave

1°) dégradation lente

Elle se traduit par une trajectoire telle que celle présentée figure 3.7.

L'amélioration lente se traduit par une trajectoire du même type mais décrite dans le sens contraire.

2°) dégradation brutale

Elle se traduit par la trajectoire illustrée à la figure 3.7.

3°) prédominance d'une situation

Elle se traduit par une trajectoire localisée dans des zones réduites telles que celles illustrées figure 3.8 .

4°) aucune tendance définie

Dans ce cas la trajectoire évolue dans une zone intermédiaire et aucune conclusion quant à la prédiction de l'évolution du processus ne peut être donnée. (figure 3.8) .

Cas particuliers

=====

Ce sont notamment les cas où la trajectoire du point image est cyclique, on se ramène à un point.

Dans ces cas très particuliers, l'évolution du processus est dite stationnaire et peut être modélisée par une chaîne de Markov homogène dont nous donnons ci-après un bref rappel.

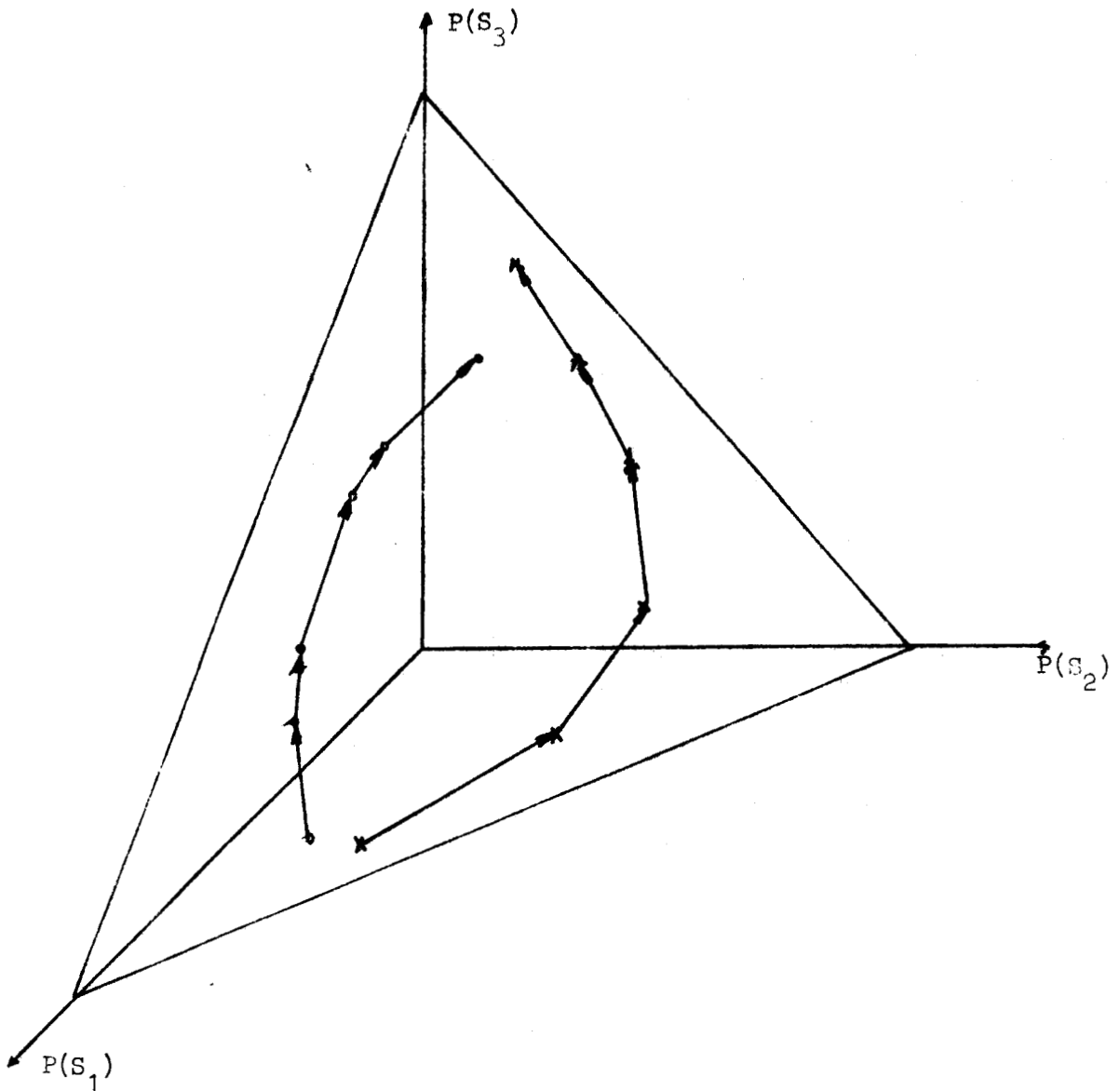
3.3.6 - MODELISATION MARKOVIENNE

Cette modélisation permet d'estimer à partir d'une séquence à l'instant t la séquence vers laquelle évolue le processus à l'instant $t+1$. Pour évaluer l'évolution du processus, on examine la transition d'une séquence vers une autre /38/.

Pour cela on calcule les probabilités d'avoir une séquence S_j à l'instant $t+1$ sachant qu'à l'instant t la séquence était S_i et on note

$P(S_j^{t+1} / S_i^t)$ cette probabilité. /23/

Dans ces conditions la probabilité d'obtenir S_j^{t+1} se note :



○—○—○ dégradation rapide

×—×—× dégradation lente

si les états sont classés par ordre de gravité dans le sens contraire des aiguilles d'une montre sur la figure, en l'occurrence de S_1 S_2 à S_3 .

Figure 3.7

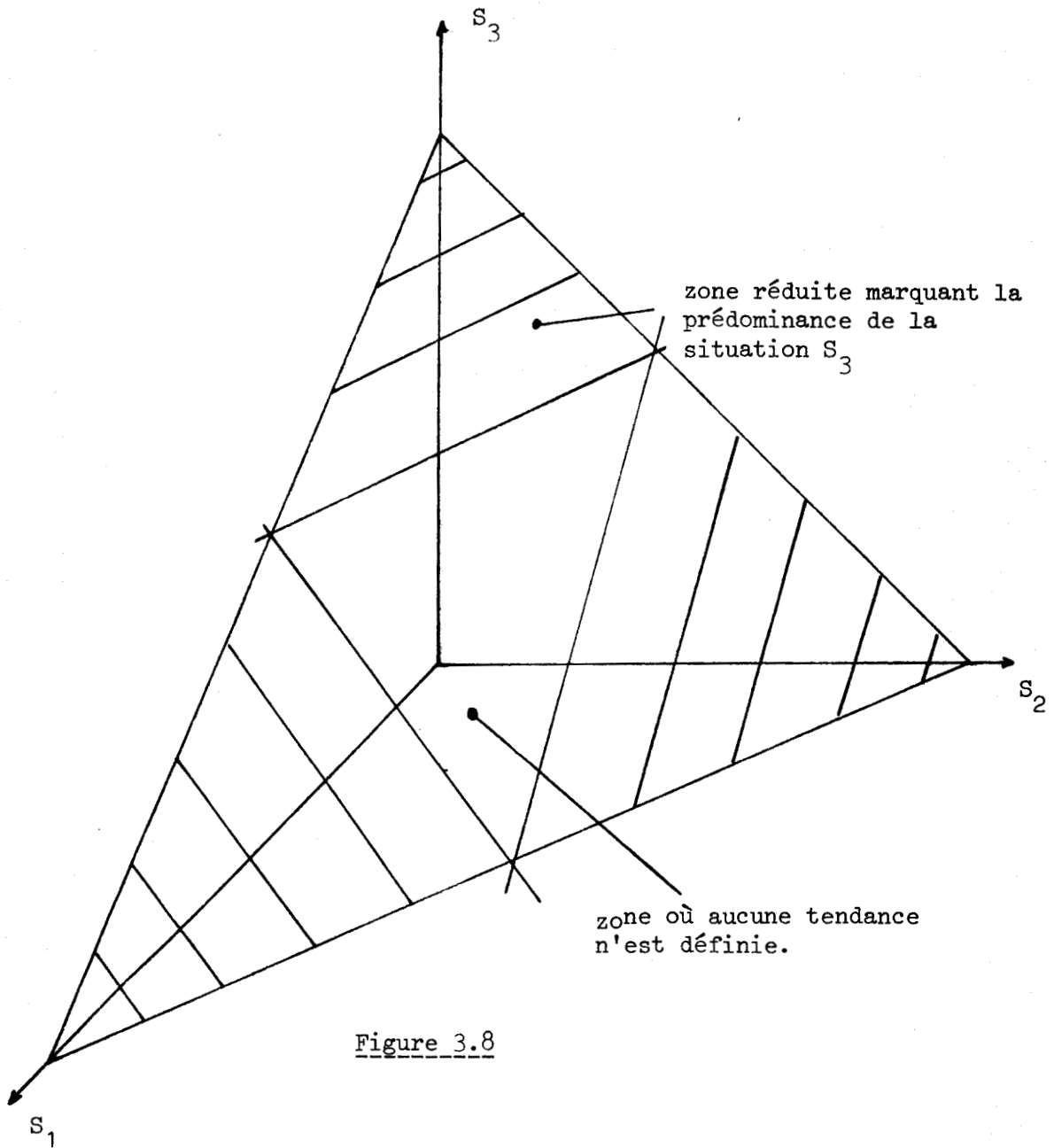


Figure 3.8



$$P(S_j^{t+1}) = \sum_{i=1}^k a_i P(S_j^{t+1} / S_i^t) \quad j = 1, k.$$

avec a_i : coefficient de présence tel que

$$a_i = \begin{cases} 1 & \text{si la situation } S_i^t \text{ est présente au temps } t \\ 0 & \text{si " } S_i^t \text{ est absente "} \end{cases}$$

L'expression précédente conduit à la forme matricielle suivante :

$$\begin{pmatrix} P(S_1^{t+1}) \\ P(S_2^{t+1}) \\ \vdots \\ P(S_k^{t+1}) \end{pmatrix} = \begin{pmatrix} P(S_1^{t+1}/S_1^t) & P(S_1^{t+1}/S_2^t) & \dots & P(S_1^{t+1}/S_k^t) \\ & P(S_2^{t+1}/S_1^t) & & " \\ & & & \\ & & & P(S_k^{t+1}/S_k^t) \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_k \end{pmatrix}$$

Expression dans laquelle la matrice $\{P(S_i^{t+1}/S_j^t)\}$ est notée $\mu(t)$.
 Cette matrice $\mu(t)$ est formée d'éléments dépendant du temps $P(S_i^{t+1}/S_j^t)$ i et j pouvant prendre k valeurs.

On peut alors noter :

$$P(S_i^{t+1}/S_j^t) = \frac{P(S_i^{t+1}, S_j^t)}{P(S_j^t)}$$

avec

$$P(S_i^{t+1}, S_j^t) = \frac{\text{nombre de transitions } S_j^t \rightarrow S_i^{t+1} \text{ observées}}{\text{nombre total de transitions observées}}$$

$$P(S_j^t) = \frac{\text{nombre de situations } S_j^t \text{ observées}}{\text{nombre total de situations observées}}$$

Par construction $\mu(t)$ est une matrice stochastique.

En effet c'est une matrice carrée d'ordre k , et d'éléments

$$\mu_{ij} = P(S_i^{t+1}/S_j^t) \text{ tels que :}$$

$$\text{Pour tout couple d'indice } 0 < \mu_{ij} < 1 \text{ et pour tout indice } j \quad \sum_{i=1}^k \mu_{ij} = 1$$

i : indice de ligne

j : indice de colonne

Pour notre application $\mu(t)$ ($t \in \mathbb{N}$) est la suite des matrices stochastiques associée à la suite des situations S_i ($i = 1, k$).

On peut alors définir un processus Markovien par une suite de matrices stochastiques et par un ensemble de valeurs initiales de probabilités.

De plus nous dirons qu'une chaîne de Markov S_i est homogène si sa suite de matrices stochastiques est constante. /41/

Une telle chaîne est donc décrite par une seule matrice stochastique μ et les probabilités de transition $P(S_j^{t+1}/S_j^t)$ sont indépendantes du temps.

REMARQUE

La modélisation du processus par une chaîne de Markov nécessite l'obtention des matrices $\mu(t)$ ainsi que des probabilités initiales /39/ /41/ .

Dans le cas d'une surveillance de signaux, le début de l'analyse ne correspond pas à l'état initial du signal.

exemple : surveillance de l'ECG.

Nous ne pouvons pas donner les valeurs initiales. La seule possibilité est donc de déterminer une matrice qui soit indépendante du temps. Cela oblige à émettre l'hypothèse de stationnarité du processus, qui peut s'avérer fausse et donner des résultats erronés.

Dans cette optique et ne connaissant à priori rien sur le processus étudié, nous n'avons émis aucune hypothèse de manière à rester dans le cadre général des observations avec histogrammes et analyse de l'évolution de ces histogrammes dans le temps.

3.3.7 - ALGORITHME

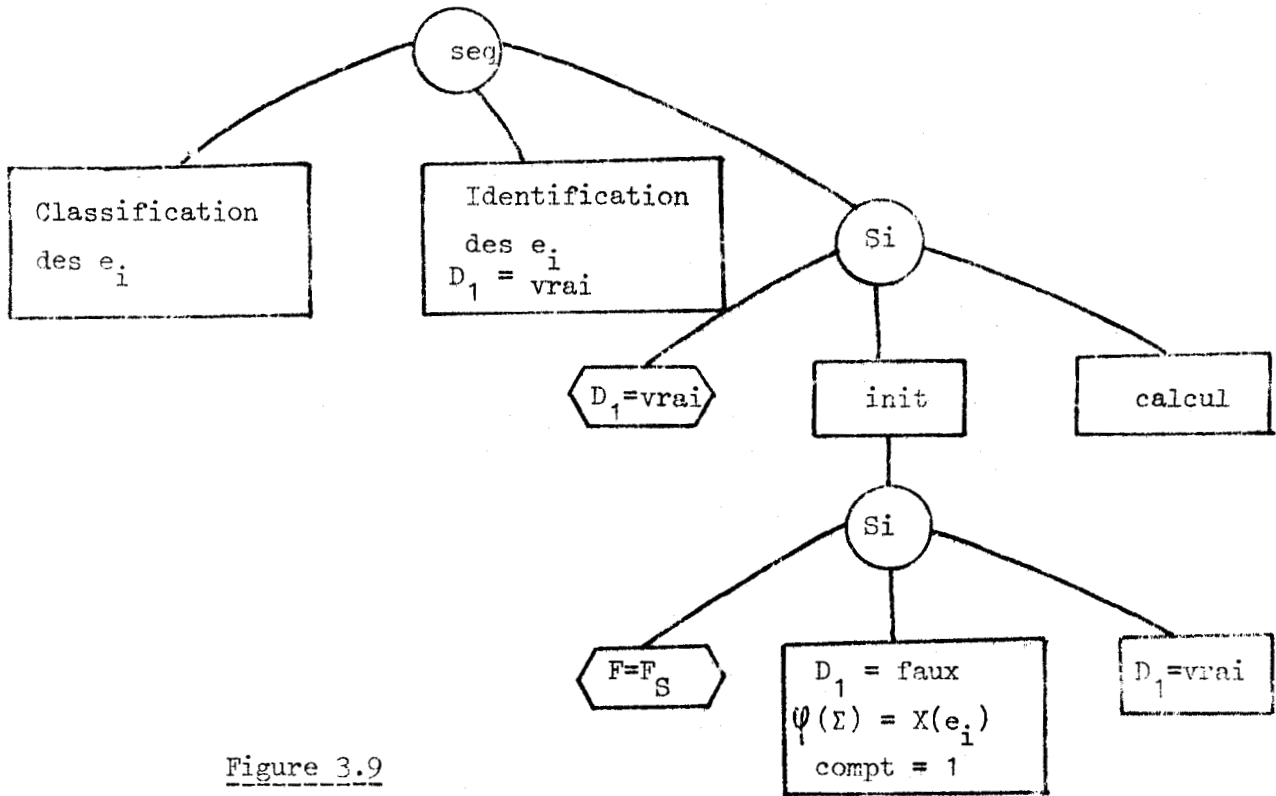


Figure 3.9

init : sert à initialiser le calcul. C'est-à-dire à rechercher une forme statique, pour déterminer le début de la première séquence.

F_D : forme dynamique

F_S : forme statique

F : forme existante

$\psi(\Sigma)$: gravité de la séquence Σ

$X(e_i)$: degré de gravité d'un état e_i

D_1 : drapeau permettant l'initialisation quand il est à vrai

"classification des e_i " : programme décrit au chapitre 1.



" Identification des e_i ", est une partie de programme s'insérant à la suite des programmes décrits au chapitre 1 en l'occurrence derrière le programme de " classification des e_i ". Afin de permettre un calcul de gravité assez simple, le taux de gravité associé à chaque e_i ainsi que la nature de leur forme (statique ou dynamique) ont été recensés dans un tableau (figure 3.10).

e_1	F_S	$X(e_1)$
e_2	F_S	$X(e_2)$
"		
e_i	F_D	$X(e_i)$
"		
e_n	F_D	$X(e_n)$

Figure 3.10

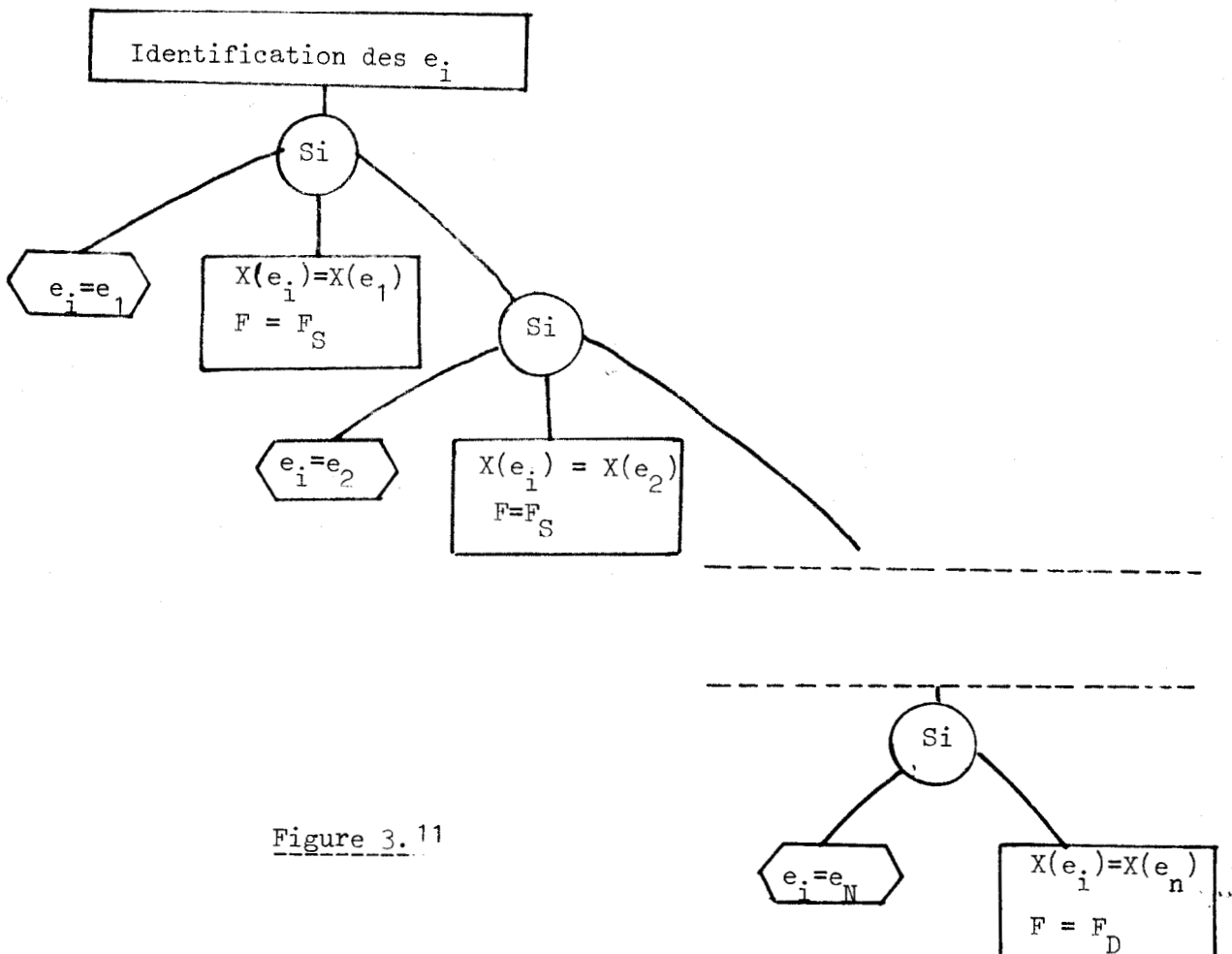
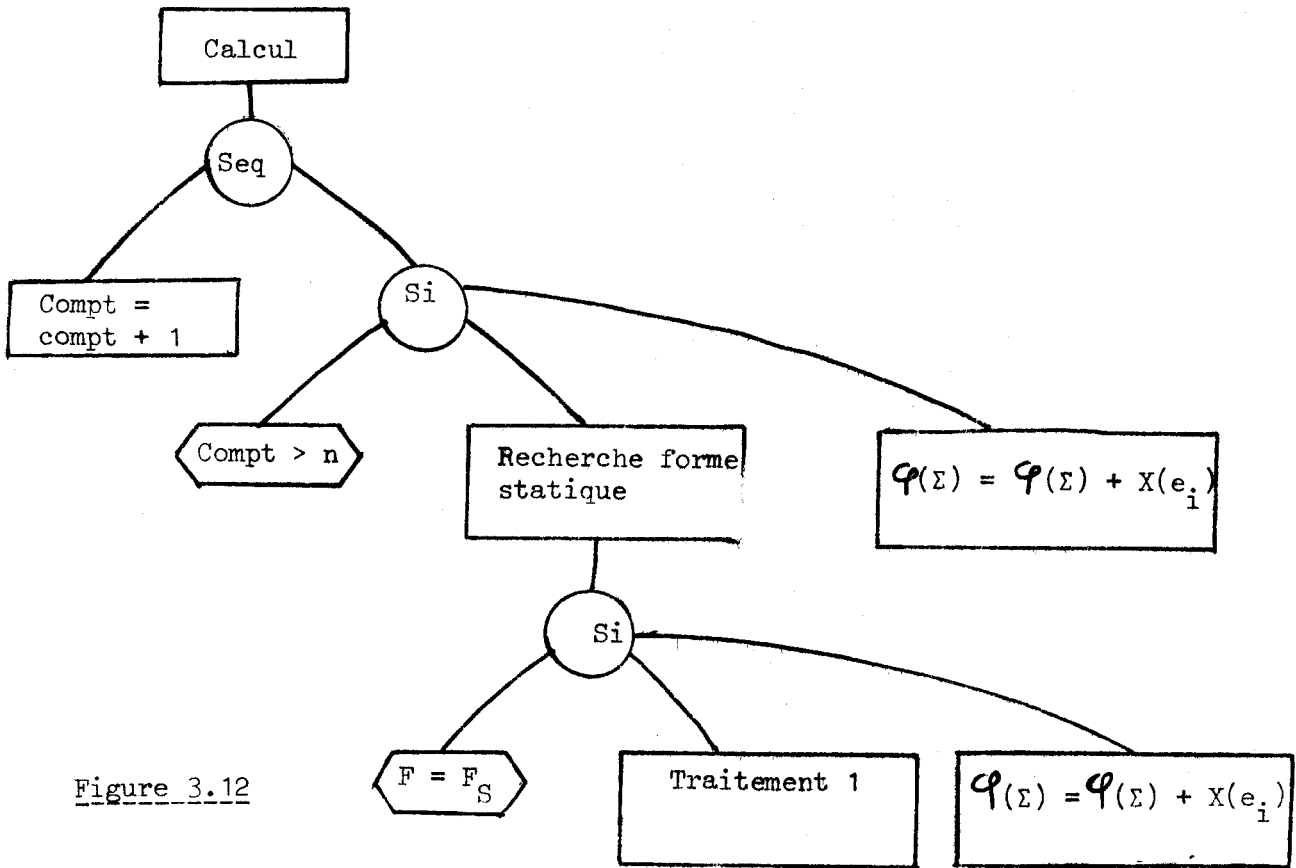
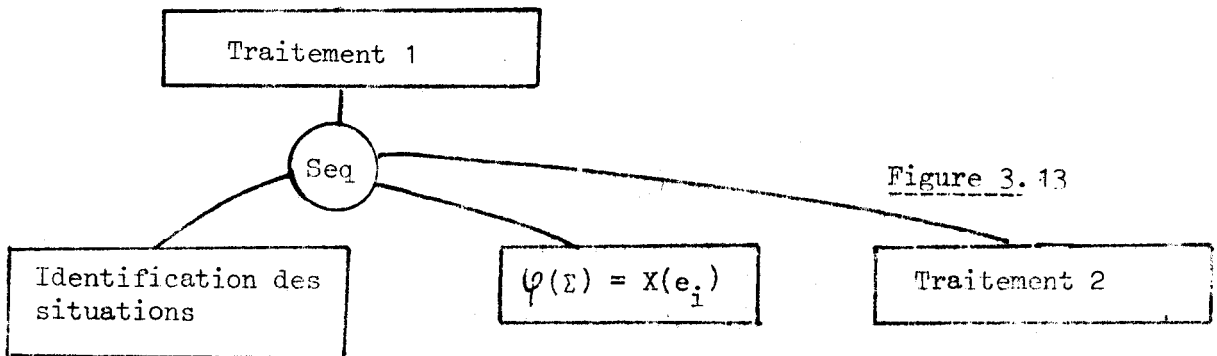


Figure 3.11



"traitement" a pour but d'identifier les séquences aux situations existantes et d'initialiser le $\varphi(\Sigma)$ pour la prochaine séquence, car le fait d'avoir $F = F_S$ lorsque le compteur "compt" est $> n$, nous indique qu'une séquence a été isolée.



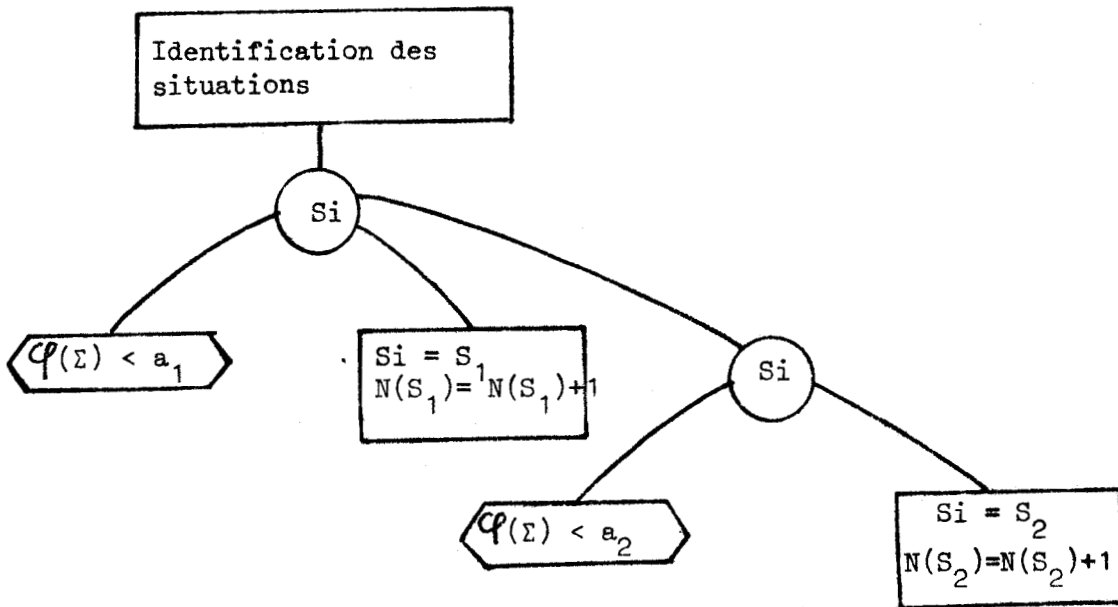


Figure 3.14

"identification des situations" n'est qu'un jeu de test sur la gravité et le recensement du nombre de fois qu'une situation S_i apparait.

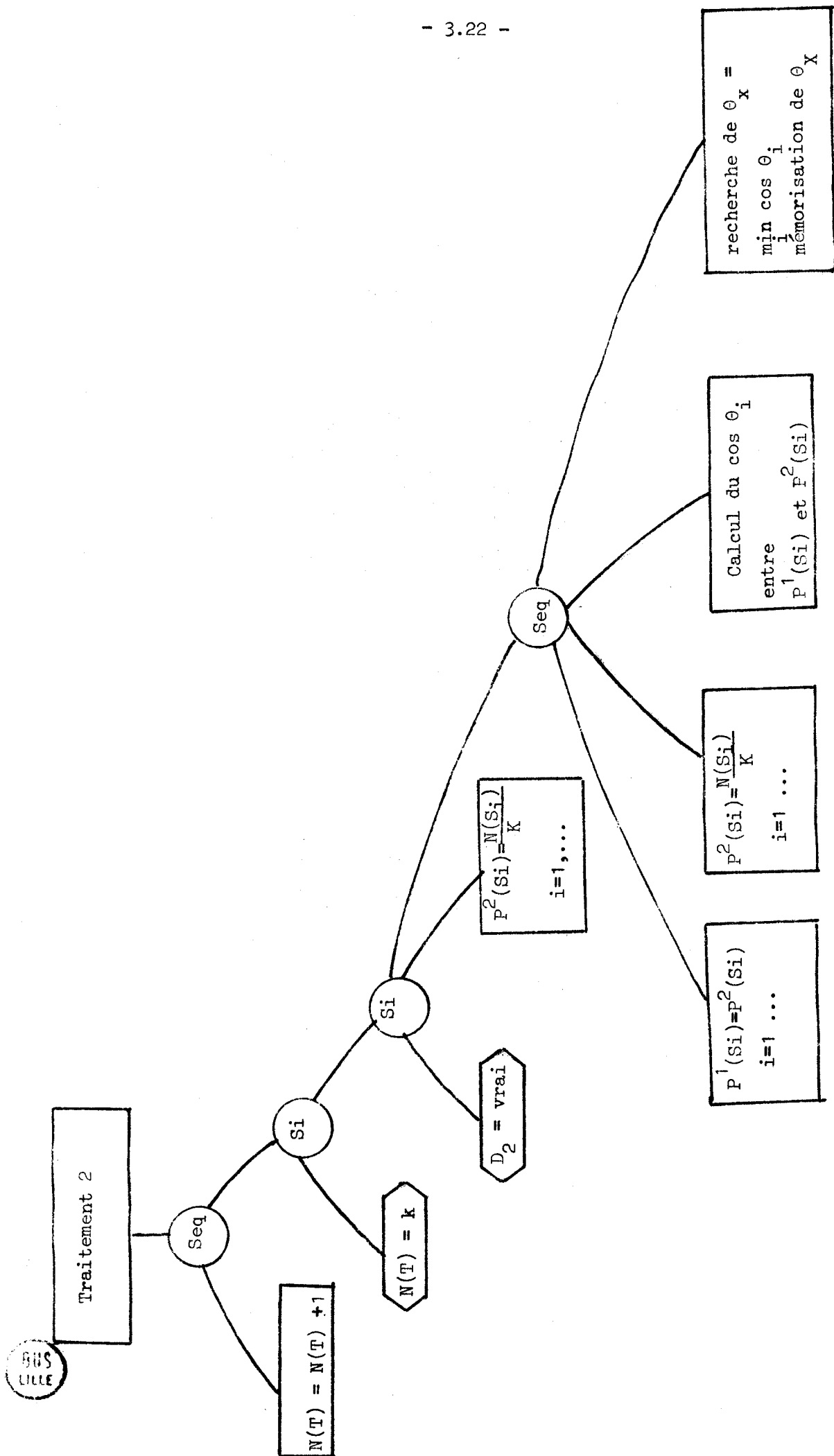


Figure 3.15

CONCLUSION

La surveillance d'un processus par observation des signaux permet de vérifier la bonne marche du système et d'alerter l'observateur dans les cas critiques afin qu'une intervention extérieure puisse être effectuée.

L'objectif visé dans ce chapitre est de prévenir toute détérioration du processus. Partant du signal présent, nous avons tenté de prédire vers quel état on tendait. Ceci en définissant un nombre restreint de situations donnant une vue globale des états du système sur une fenêtre donnée du temps. Cette évaluation globale minimise l'erreur susceptible d'être commise lors d'une prédiction.

Les prédictions sont ensuite effectuées en faisant appel à des outils statistiques simples tels que les histogrammes.

L'évolution dans le temps de ces éléments statistiques a pu ensuite être évaluée en adoptant une représentation géométrique élémentaire permettant de mettre en évidence des trajectoires de tendance.

L'évolution de ces trajectoires, dans l'espace des situations, est analogue à l'évolution des trajectoires d'état dans l'espace d'état.

Il est alors apparu que certains cas particuliers d'évolution des trajectoires pouvaient se ramener à l'étude de chaînes de Markov homogènes.

Nous donnons au chapitre suivant quelques éléments d'application de ces méthodes au traitement du signal biologique.

- C H A P I T R E I V -

APPLICATION EN REANIMATION RESPIRATOIRE

CHAPITRE IV

4.1 - INTRODUCTION	page 1
4.2 - DEFINITION DU CAHIER DES CHARGES	page 2
4.2.1. - Sélection des signaux	page 2
4.2.2. - Description des signaux	page 4
a - mesure du volume	page 4
b - mesure de la concentration en CO ₂	
4.2.3. - Choix de la méthode	page 7
4.3 - TRAITEMENT SUR CALCULATEUR	page 7
4.3.1. - Observateur numérique	page 7
4.3.2. - Traitement global	page 8
4.3.3. - Choix des paramètres	page 9
4.3.3.1. - Le volume	page 9
4.3.3.2. - La concentration en CO ₂	page 10
4.3.3.3. - La pression	page 14
4.3.4. - Synchronisation	page 16
4.4 - EXTRACTION DE CYCLE	page 17
4.4.1. - Traitement du volume	page 17
4.4.2. - Traitement de la concentration CO ₂	page 19
4.4.3. - Traitement de la pression	page 22
4.5 - ETUDE DANS LE TEMPS DES FORMES	page 24
4.5.1. - Création d'état	page 24
4.6 - PROTOCOLE DE TRAVAIL	page 24
4.6 1. - Acquisition des signaux	page 26
4.6 2. - Gestion des fichiers	page 27
4.6 3. - Résultats et discussions	page 27
4.7 - CONCLUSION	page 28

4.1 - INTRODUCTION

Dans les chapitres précédents, nous avons décrit une méthodologie permettant l'étude dans le temps d'un système quelconque.

Dans ce chapitre, nous appliquons cette méthodologie à des signaux physiologiques prélevés sur des patients soumis à une ventilation assistée et nécessitant une surveillance continue de leur état.

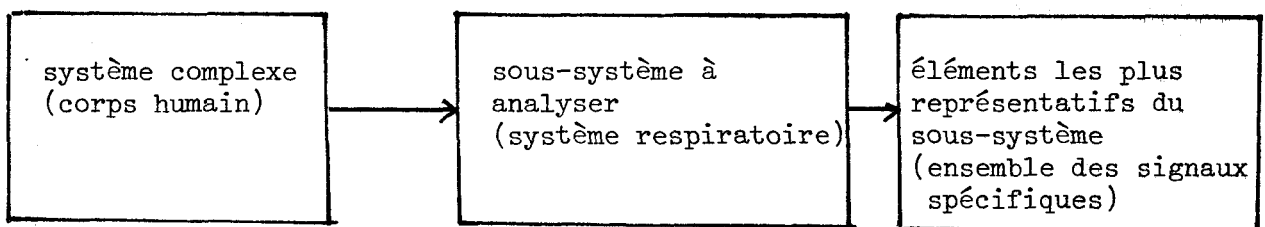
Ces malades, en transit dans un service de réanimation respiratoire ont des problèmes de natures diverses ce qui se répercute sur la structure des signaux prélevés en clinique.

Le but de cette application est de réaliser un système de surveillance permettant l'analyse temps réel de plusieurs signaux pris simultanément sur le patient, afin de prévenir l'évolution de son état.

Notre intérêt s'est porté sur l'activité du système respiratoire du patient. Cela nécessite une sélection des signaux les plus représentatifs de cette activité, afin d'éviter une redondance de traitements.

Les signaux ne permettant pas une étude prolongée dans le temps, tels que ceux qui sont prélevés d'une manière invasive ont donc été mis à l'écart.

La sélection entreprise s'illustre de la manière suivante :



Après captation des signaux au niveau du malade, le traitement de ceux-ci consiste à étudier :

- 1°) la structure de leur forme qui n'est autre que le reflet du mécanisme ventilatoire du poumon étudié.
- 2°) son évolution dans le temps.

4.2 - DEFINITION DU CAHIER DES CHARGES

4.2.1. - Sélection des signaux

En réanimation respiratoire, les malades sous ventilation nécessitent une surveillance intensive de leur système pulmonaire.

En particulier la pression artérielle à l'inspiration et à l'expiration les échanges gazeux, le PH, la saturation de l'hémoglobine doivent être prélevés. /44/

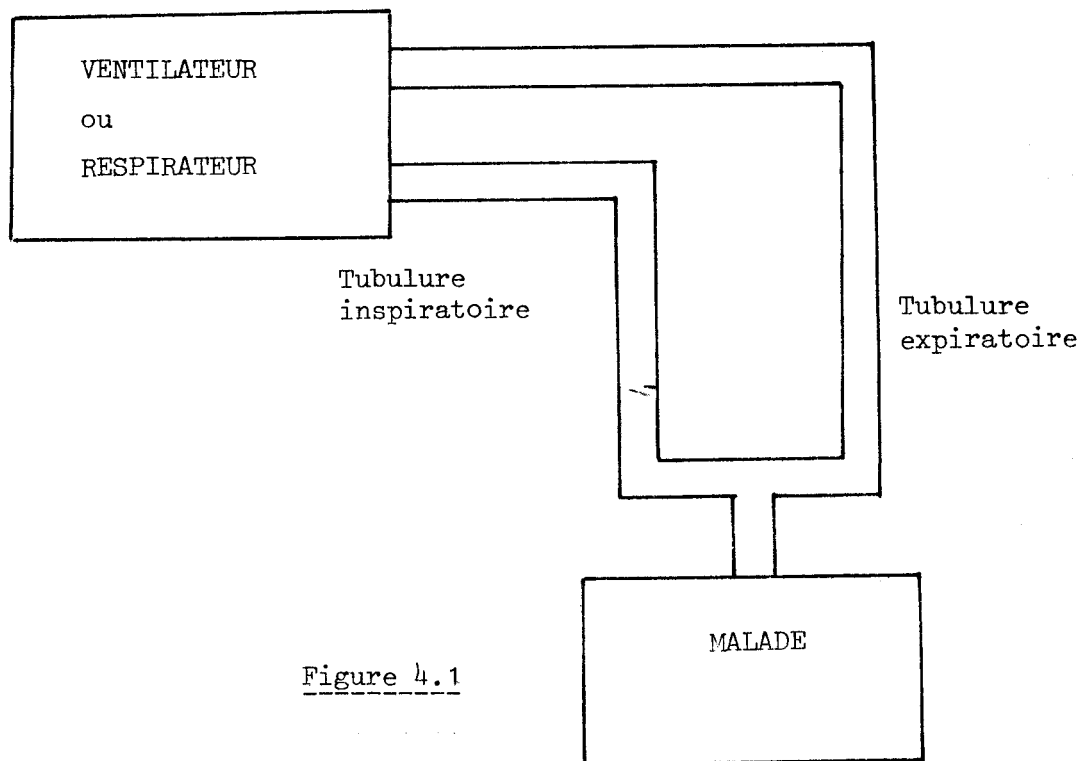
Les échanges respiratoires demande une étude de la consommation d'oxygène, de la production de CO_2 et de la mécanique ventilatoire des échanges gazeux /45/.

Dans notre travail, nous nous sommes intéressés à la mécanique ventilatoire qui excluait tout prélèvement invasif comme les prises de sang et permettait l'étude de signaux prélevés simultanément à l'aide de capteurs spécifiques.

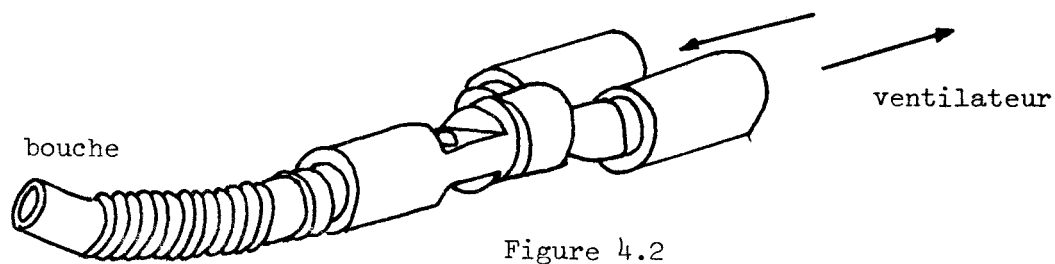
Chaque malade ventilé reçoit de l'air ambiant ou de l'air enrichi grâce à un respirateur auquel il est relié et qui contrôle sa respiration à une fréquence réglée par le médecin.

Ce respirateur est pourvu d'un débitmètre mesurant le débit instantané d'air qui peut être intégré électriquement et donner le volume transitant entre le malade et le ventilateur.

Il utilise des transducteurs aussi bien électroniques que mécaniques permettant une mesure de pression, ou un contrôle des valves dans les tubulures inspiratoire et expiratoire (figure 4.1).



La connexion ventilateur malade se fait par une pièce en Y et permet un branchement au niveau de la bouche, aisé et non invasif / 46/.



Le relevé des signaux est donc effectué à la bouche du patient comme l'indique la figure 4.3 .



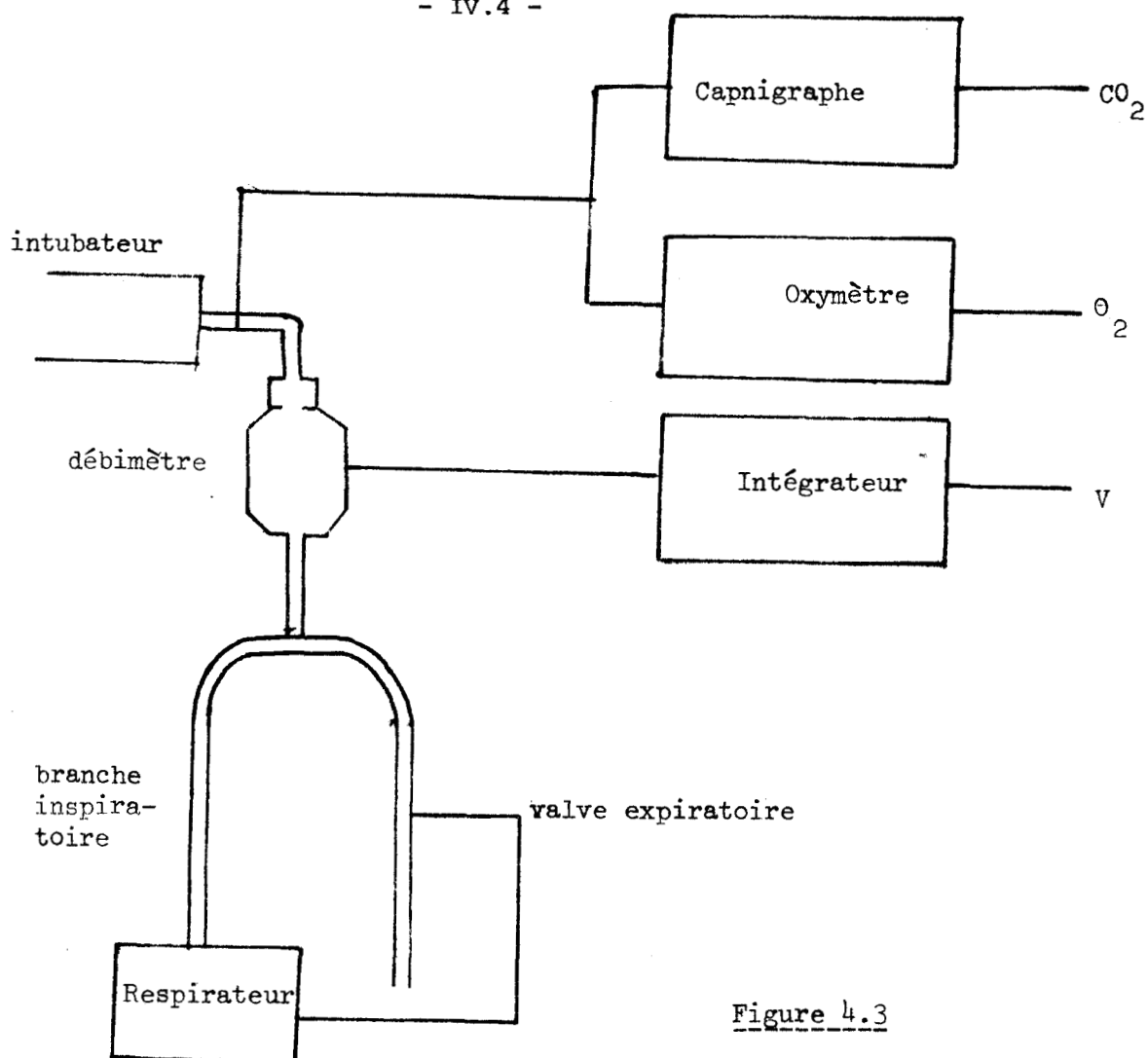


Figure 4.3

Dans le cadre de cette étude, les signaux suivants ont été analysés.

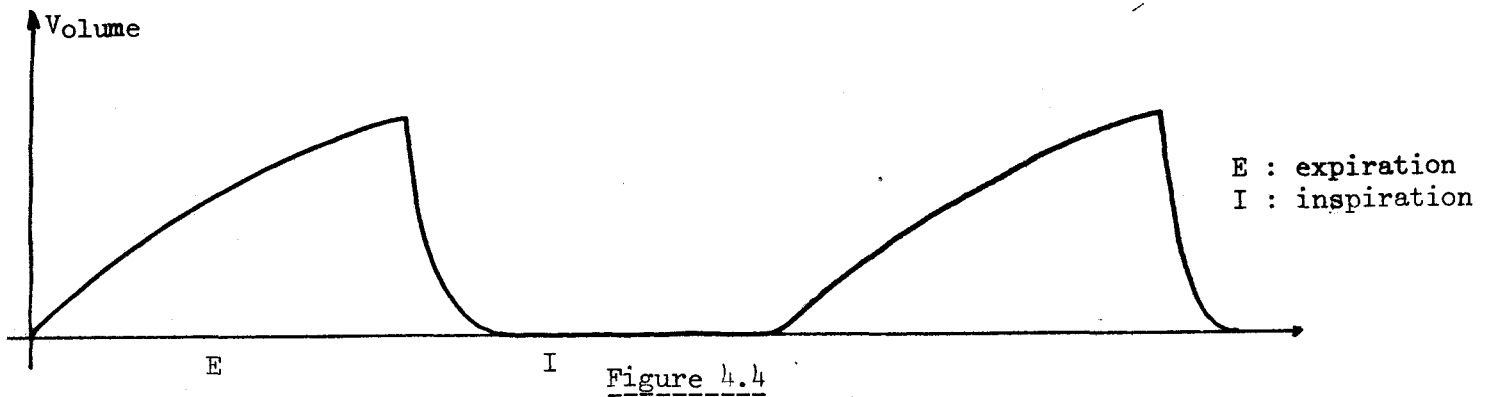
- 1°) variation de volume pour chaque cycle ventilatoire
- 2°) variation de la concentration de gaz carbonique
- 3°) évolution de la pression

4.2.2. - Description des signaux

a - Mesure du volume

La mesure du volume est faite par un débitmètre dont le signal est intégré électriquement. Il permet la mesure des débits instantanés sans inertie appréciable. /47/

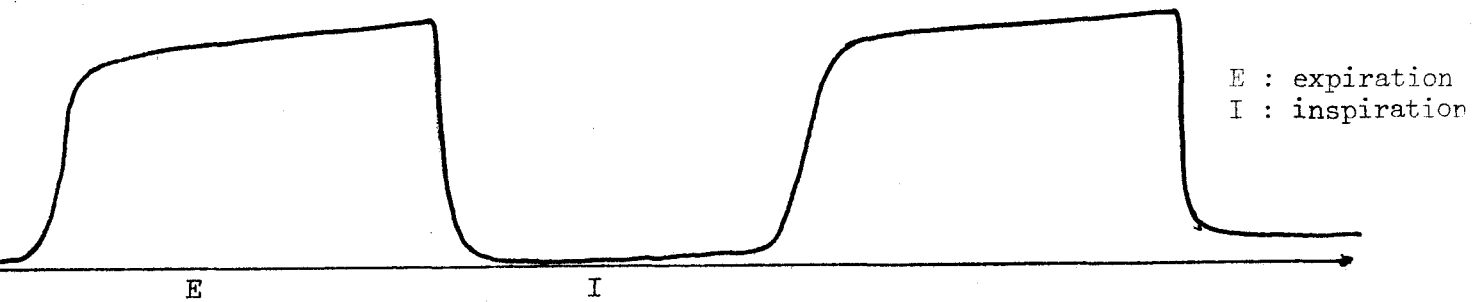
L'allure de la courbe du volume est de la forme



b) mesure de la concentration en CO_2

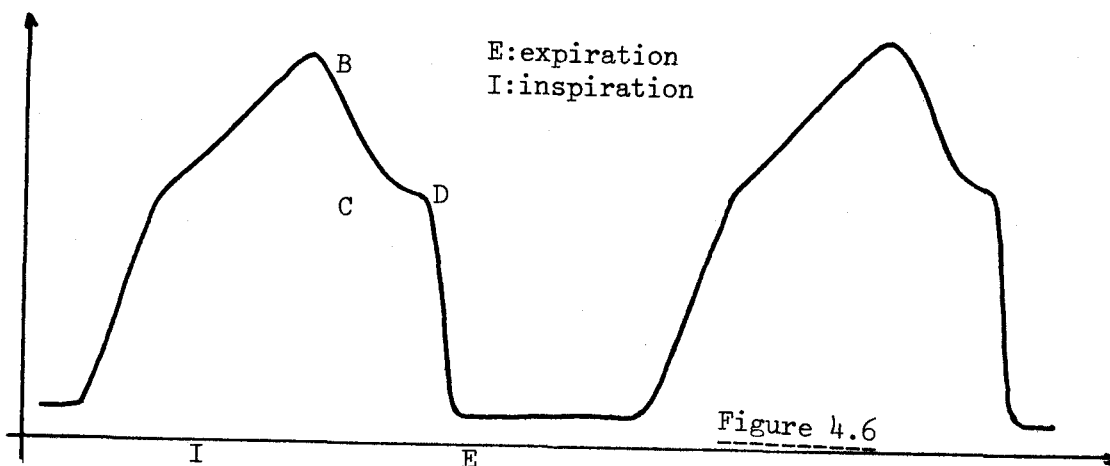
La mesure de la concentration en gaz carbonique est faite par un analyseur de CO_2 appelé capnigraphe.

La courbe du CO_2 prend l'allure suivante dans les cas normaux.



La mesure de la pression a été faite grâce à un capteur intégré au respirateur.

L'allure du signal de pression est donnée par le schéma suivant.



4.2.3. - Choix de la méthode

Le cahier des charges spécifie toutes les formes susceptibles d'être rencontrées ainsi que leurs interprétations.

En ce qui concerne le signal de concentration en CO_2 expiré l'interprétation est mécanique.

En ce qui concerne le volume, la structure du signal conduit à un nombre de formes limité imposé par le signal ventilé lui-même.

En ce qui concerne le signal de pression, l'interprétation est beaucoup moins évidente et conduit à distinguer un nombre important de formes.

Par ailleurs, compte tenu de l'activité répétitive du corps humain, chaque signal est périodique. Ainsi pour chacun des trois signaux, nous avons extrait en temps réel chaque cycle respiratoire et procédé à leurs classifications parmi les formes prototypes.

4.3 - TRAITEMENT SUR CALCULATEUR

Au chapitre 1, nous avons proposé une certaine méthodologie que nous allons appliquer pour les signaux de CO_2 , de volume et de pression.

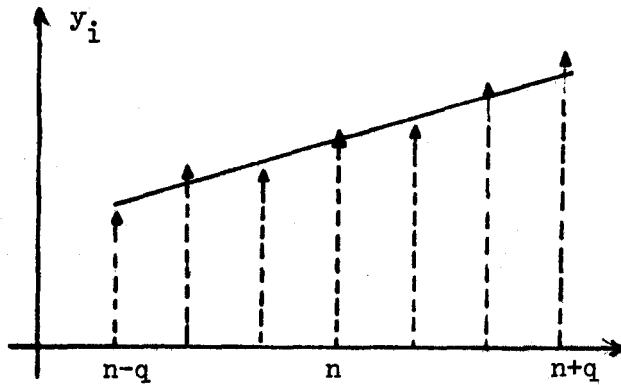
Pour appliquer le traitement proposé, le signal est associé à sa dérivée selon une procédure de segmentation récursive /26 /

4.3.1. - Observateur numérique (segmentation récursive)

L'observateur numérique de segmentation génère 2 suites b_n et a_n représentant respectivement

- l'estimation du signal filtré et
- l'estimation de sa dérivée.

Ces suites sont générées par un algorithme dont le mécanisme est illustré ci-dessous.



y_i : signal échantillonné

Figure 4.7

Soit T la période d'échantillonnage au $n^{\text{ième}}$ instant d'échantillonnage (nT) nous associons les q échantillons précédant et les q échantillons suivant d'instant nT .

La séquence ainsi constituée comporte $2q + 1$ échantillons notés : $\{ y_{n+i} \}$, $i = -q \dots +q$.

La méthode consiste alors à

- 1°) Tracer le segment de régression linéaire du signal d'entrée à partir des $\{ y_{n+i} \}$.
- 2°) Faire glisser ce segment le long du signal. Il vient alors :

$$b_n = \frac{1}{2q+1} \sum_{i=-q}^{+q} y_{n+i}$$

$$a_n = \frac{3}{q(q+1)(2q+1)T} \cdot \sum_{i=-q}^{+q} i y_{n+i}$$

Nous appliquerons cet algorithme à chaque signal.

La valeur de q se détermine expérimentalement. Elle vaut dans notre cas 3 pour les 3 signaux. Par ailleurs $T = 20$ ms.

4.3.2. - Traitement_global

Le traitement envisagé doit être effectué simultanément sur les 3 signaux en temps réel.

La structure temps réel est obtenue en réalisant un superviseur chargé de gérer des drapeaux validant l'activation de telle ou telle tâche. /28/

Ainsi que l'illustre l'arbre de la figure 4.8, le traitement correspond à une structure "tant que" permettant le retour continu à une même tâche. "Trait 3S" (traitement des 3 signaux)

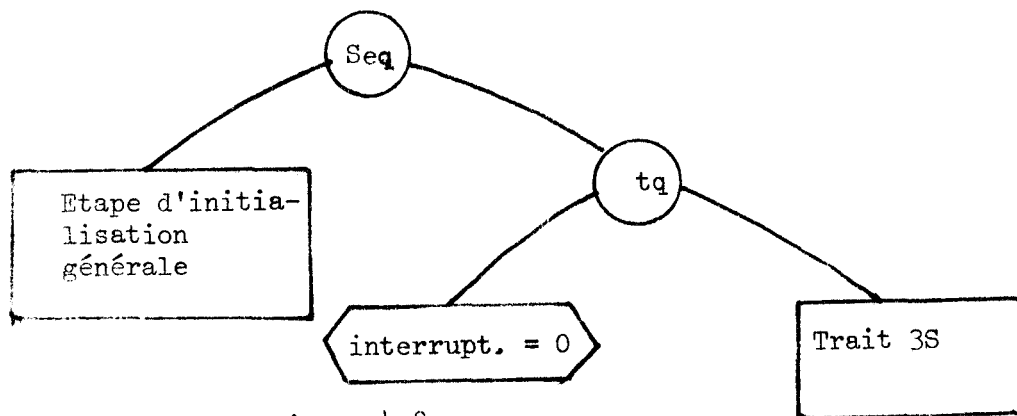


Figure 4.8

interrupt est une variable booléenne marquant si une interruption du traitement doit être faite ou non.

La tâche "Trait 3S" (figure 4.9) comporte elle-même

- l'acquisition des échantillons
- la génération des différentes suites a_n et b_n
- le traitement des différents cycles extraits.

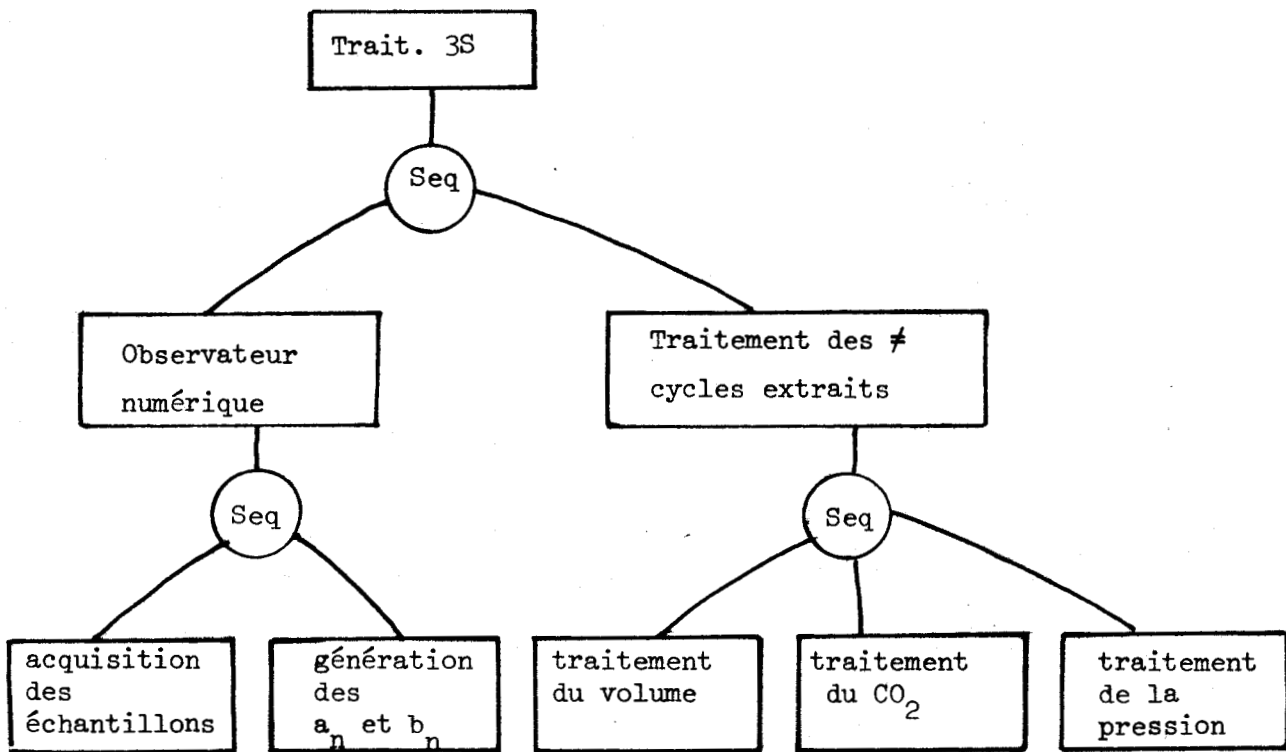


Figure 4.9

Les prochains paragraphes décrivent en détail ces différents traitements.

4.3.3. - Choix des paramètres

Les formes prototypes déterminées par le cahier des charges sont caractérisées par des paramètres extraits des enveloppes des signaux selon la procédure décrite au chapitre 1.

4.3.3.1. - Le volume

Sur le schéma suivant sont représentées les 2 classes rencontrées

- cycle normal
- cycle désadapté B

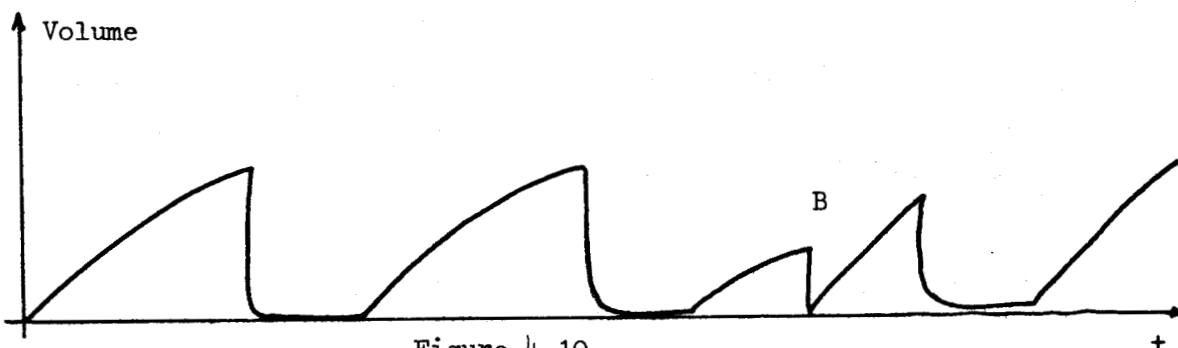
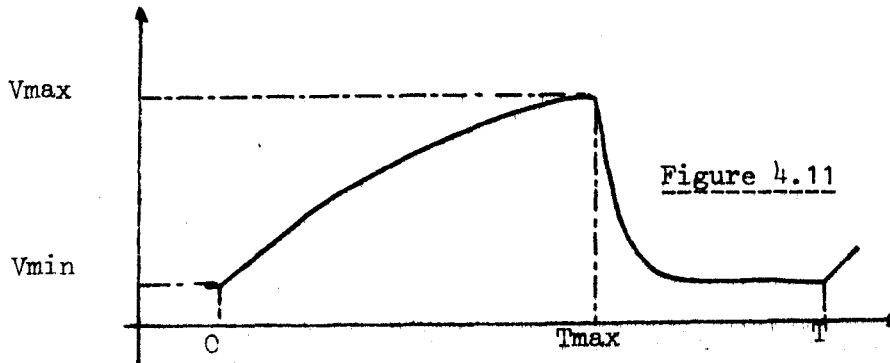


Figure 4.10

Ces 2 classes ne se différencient pas par leurs formes mais par les valeurs des maxima et des longueurs de cycle.

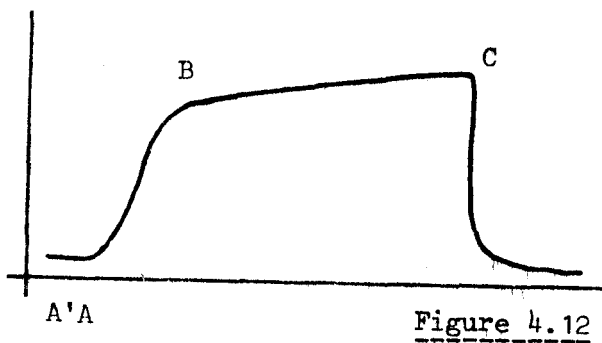
Paramètres utilisés (figure 4.11)



Vmax : valeur du volume maximal
Vmin : valeur du volume minimal
Tmax : temps correspondant à Vmax
T : période correspondant à Vmin.

4.3.3.2. - La concentration en CO_2

Le signal de la concentration en gaz carbonique expiré a l'allure suivante (cas normal)



A'A est le temps de réponse dû au trajet de l'air dans les bronches (espace mort).

Le segment BC représente le plateau alvéolaire qui a une pente positive mais minimale. Quand cette pente devient importante, elle traduit une forme spécifique en S relative aux asthmatiques.

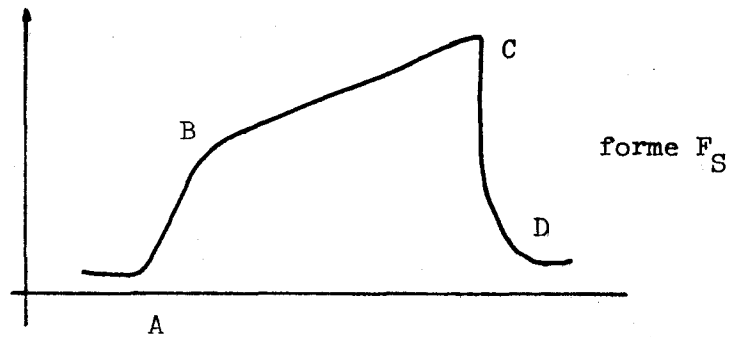


Figure 4.13

Ce premier type de forme anormale est noté F_S (forme en S). Il est donc important de déterminer le point B qui fait la jonction entre les pentes de AB et de BC.

Les formes F_N et F_S ont la même allure et se décrivent par un même mot syntaxique selon la grammaire définie au chapitre 1.

Il existe d'autres formes pour le signal de CO_2 qui traduisent chacune un mécanisme ventilatoire. Ces formes ne peuvent pas décrire à elles seules une pathologie. Pour cela, il faut leurs associer d'autres paramètres.

Par exemple une obstruction de la canule se traduit par le signal suivant de la figure 4.14a

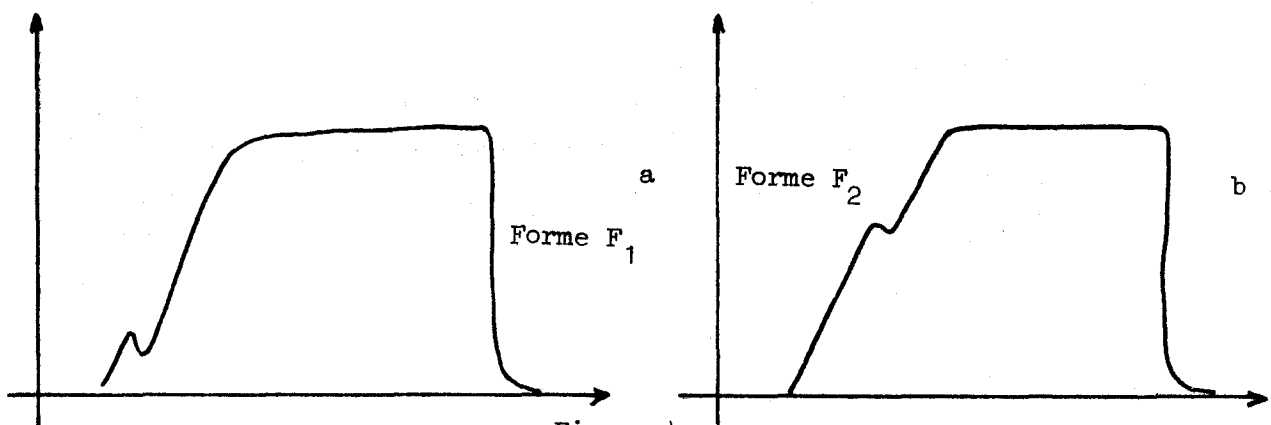


Figure 4.14

Si l'obstruction se situe plus en aval dans les poumons, elle occasionne une gêne et se traduit par la figure 4.14b

Par contre lorsque l'un des poumons a une activité moindre que l'autre, cela se traduit par un asynchronisme et l'on obtient l'allure suivante.

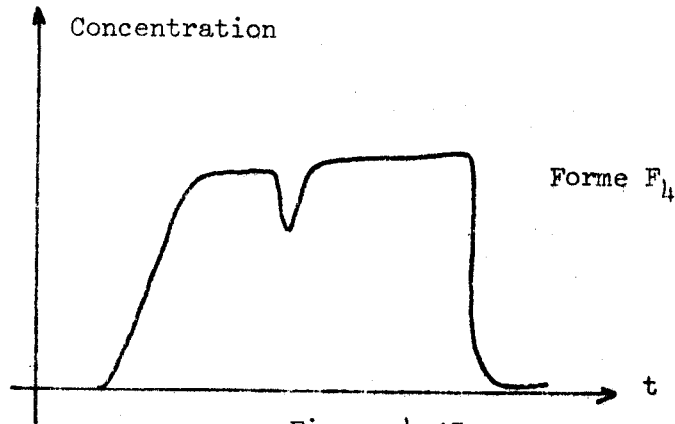


Figure 4.15

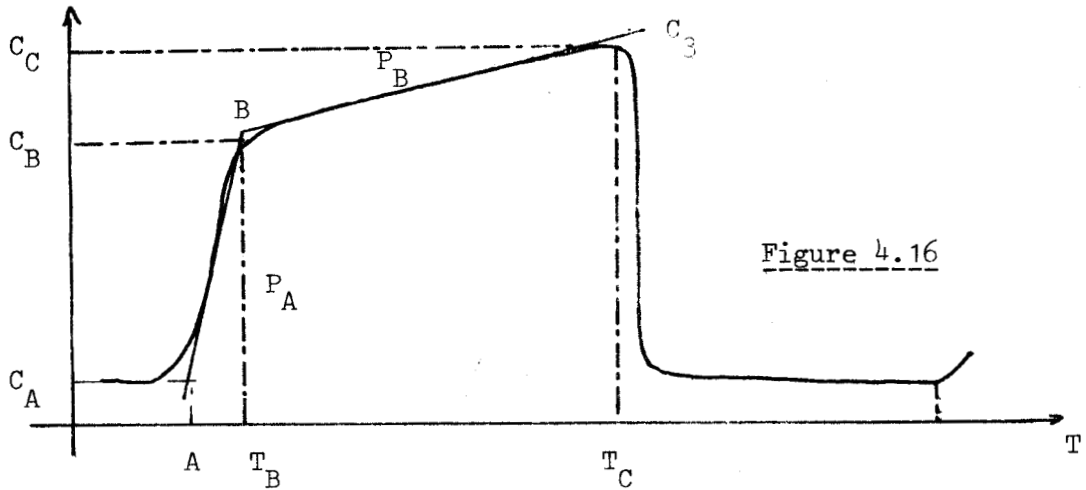
Lorsqu'un malade est dans un état très relâché, cela se traduit par des oscillations cardiaques qui apparaissent au niveau du plateau alvéolaire.

Enfin lorsque la respiration du malade se désadapte par rapport à la machine, nous avons des oscillations anarchiques sur le plateau qui fournissent des formes méconnaissables.

En résumé, il apparaît que toutes les anomalies se localisent dans le tronçon AB ou le tronçon BC.

-- PARAMETRES UTILISES

Chaque signal est décrit par une enveloppe où sont recensés les paramètres permettant de décrire toutes les formes données par le cahier des charges. La figure 4.16 illustre quelle est la méthode de détermination des paramètres.



- P_A, P_B pentes moyennes des segments AB et BC
- C_A, C_B, C_C valeurs de la concentration aux points A,B,C
- T_B, T_C valeurs des temps associé aux points B et C
- T période

Il faut bien entendu tenir compte des pathologies qui se traduisent par des minima et des maxima sur la courbe. Pour cela, on définit des indicateurs qui comptabilisent les anomalies intervenues par rapport à la forme normale.

Pour le tronçon A B, deux indicateurs ont été définis :

T(1) pour la comptabilité des anomalies caractérisant les formes F_1
 T(2) " " " " les formes F_2

Pour le plateau, l'indicateur T(3) recense le nombre d'anomalies, les amplitudes et les temps à mémoriser pour déterminer si les oscillations sont cardiaques où s'il y a désadaptation.

4.3.3.3. - La pression

- formes normales

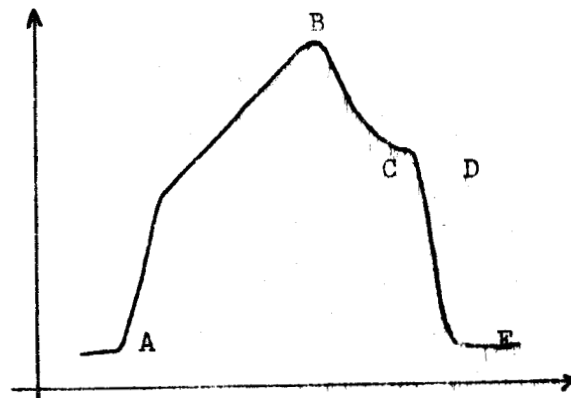


Figure 4.17

La fréquence ou l'absence du plateau CD permet de définir plusieurs types de formes normales.

Par ailleurs pour une même allure de courbe, l'amplitude du plateau et les valeurs des pentes permettent également de différencier plusieurs classes.

- formes anormales

Elles sont caractérisées :

- soit par des accidents se produisant sur l'un des segments AB ou CD (voir figure 4.18a et 4.18b)
- soit par une désadaptation complète où l'on obtient des formes diverses (voir figure 4.19)

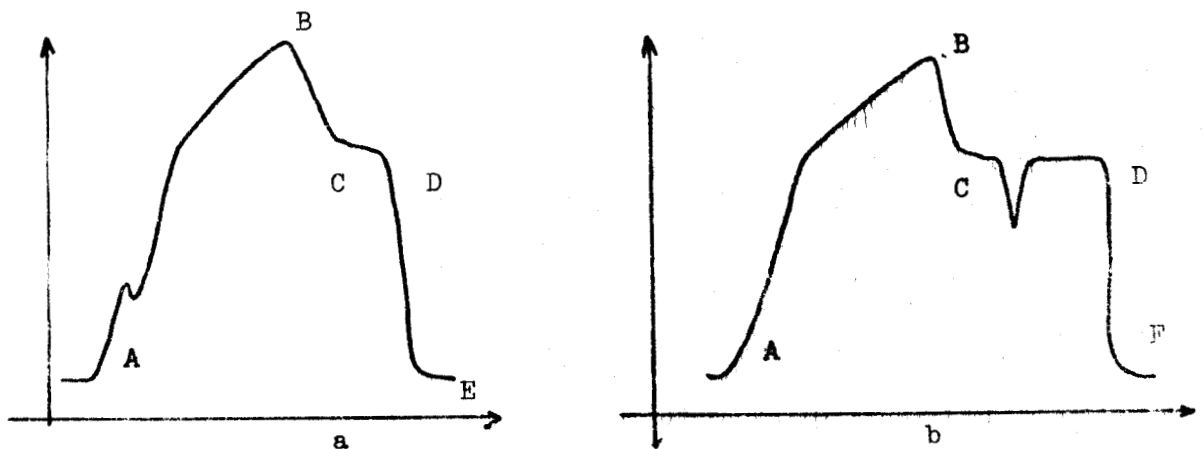


figure 4.18

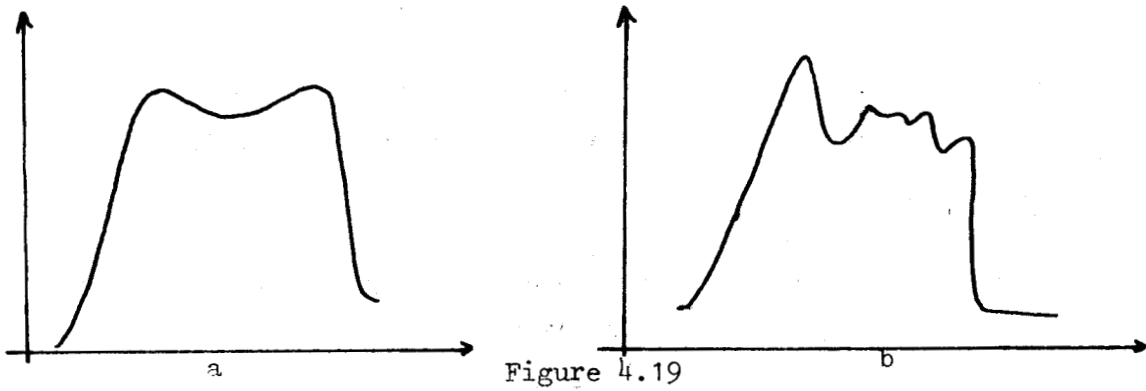


Figure 4.19

Dans ce cas nous associons à ces deux formes une même classe notée " désadaptée " alors que sur le plan géométrique, elles diffèrent totalement.

Il existe également des formes de désadaptation du malade envers la machine qui se traduisent par une amplitude très basse de P_{max} , ou par une période complètement différente de la période normale (figure 4.20).

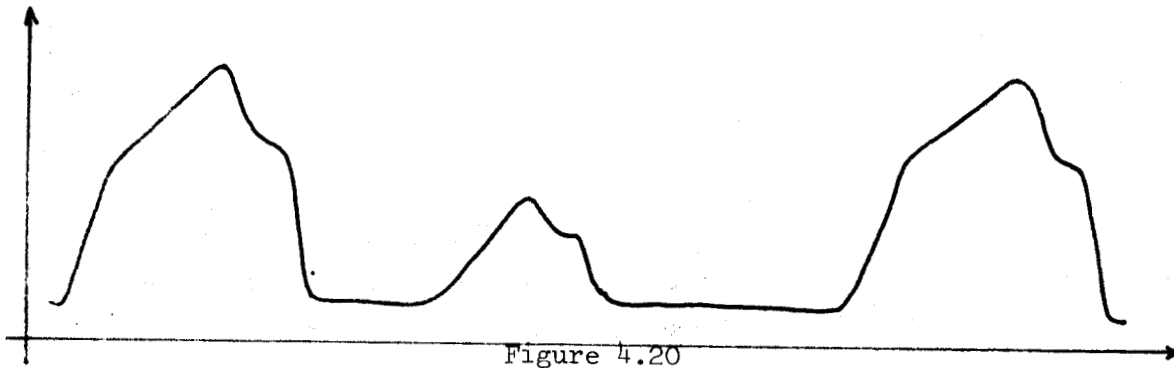


Figure 4.20

Paramètres utilisés

La forme est caractérisée à partir de son enveloppe ainsi que l'illustre la figure 4.21.

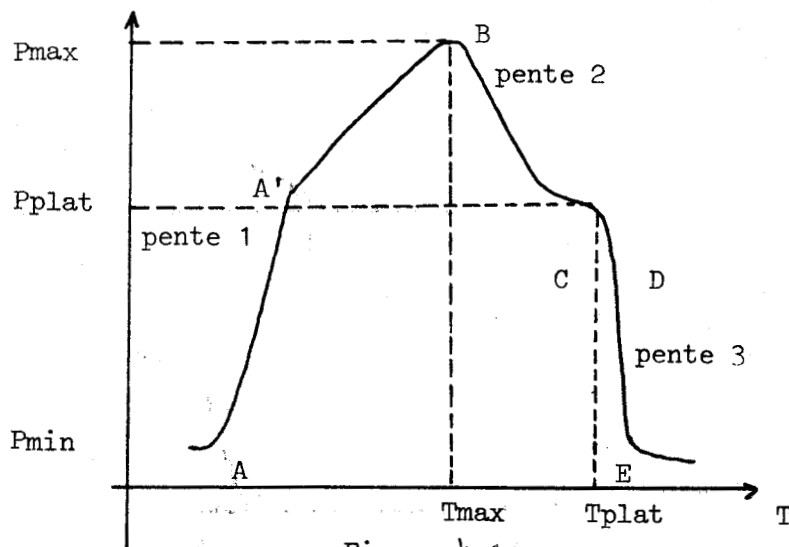


Figure 4.21

pente 1 : c'est la pente moyenne du segment AB qui peut se décomposer en 2 pentes AA' et A'B suivant l'application que l'on veut réaliser.

pente 2 et pente 3 : sont les pentes négatives moyennes des segments BC et DE marquant l'expiration.

Pmax, Pplat, Pmin : sont les amplitudes des pressions respectivement du maximum, du plateau, et du minimum.

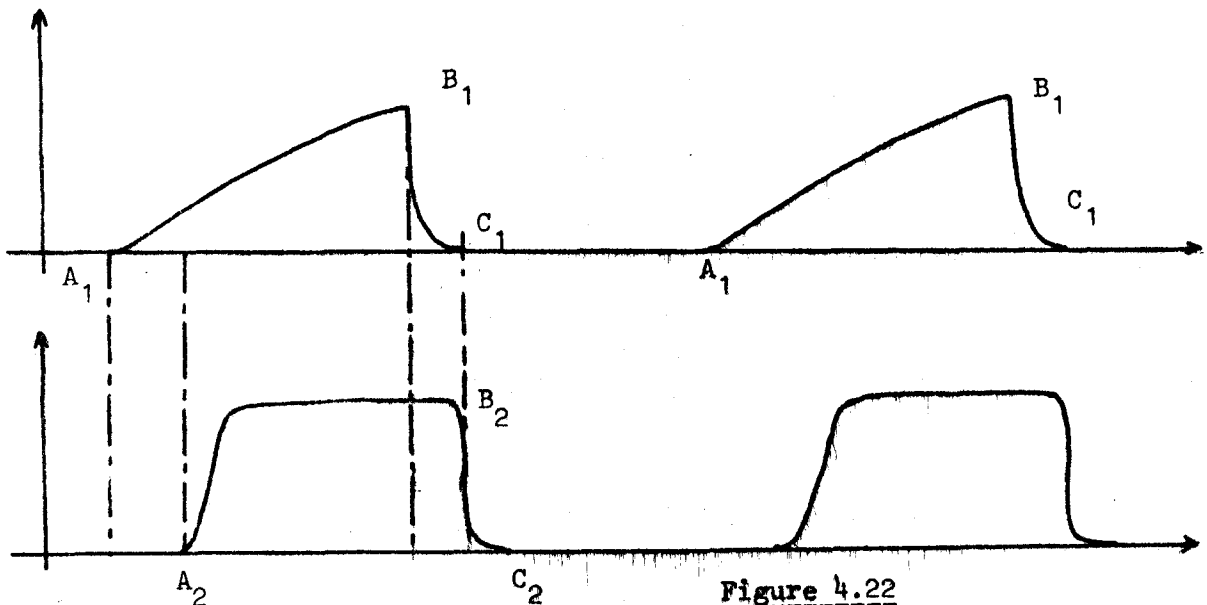
Tmax, Tplat, T sont les instants correspondant à ces différentes amplitudes.

Par ailleurs on associe à chaque segment un indicateur permettant de comptabiliser les accidents survenant sur chacun de ces segments.

4.3.4. - Synchronisation

Les 3 signaux à traiter sont générés au même instant par les poumons. Par contre en ce qui concerne le signal de CO_2 un léger retard apparaît au prélèvement. Ceci est dû au temps de transit des gaz dans la tubulure.

Il importe donc de resynchroniser le signal de CO_2 étant entendu qu'il peut avoir plus d'un cycle de retard par rapport aux signaux de volume et de pression.



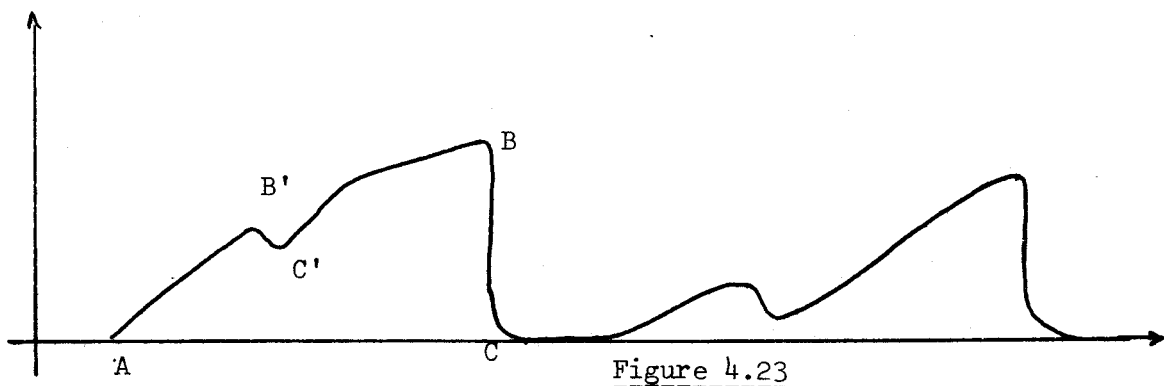
4.4 - EXTRACTION DE CYCLE

Pour chaque signal on effectue un traitement spécifique d'extraction de paramètres en vue d'une classification. Le plus simple concerne celui du volume.

4.4.1. - Traitement du volume

L'algorithme mis au point a pour objet d'extraire en temps réel un cycle et les paramètres caractérisant V_{max} V_{min} T_{max} T_{min} .

Nous n'avons que 2 classes à reconnaître dans ce signal.



Le traitement en temps réel conduit à prendre des décisions a priori qui sont infirmées ou confirmées a posteriori par les échantillons suivants.

Une telle procédure oblige à prendre en compte toute perturbation apparaissant sur le signal, quitte à la négliger au moment de la décision de classification.

Afin d'illustrer ceci, nous donnons ci-après la succession des décisions prises par l'algorithme lors du traitement du signal de la figure 4.23.

Pour la recherche du maximum B nous devons lors du traitement temps réel émettre l'hypothèse que B' est le maximum jusqu'à la détection de B qui permettra d'infirmar cette hypothèse.

Il en va de même pour la détermination de fin de cycle qui coïncide avec le début du cycle suivant.

Pour éviter une mauvaise détection, le début de cycle n'est validé que si les pentes ont dépassées un certain seuil spécifique au volume.

Lorsque le début de cycle est validé, le cycle précédent est extrait et les paramètres le caractérisant peuvent être envoyés en sortie, c'est-à-dire aiguillés vers un programme de classification.

L'algorithme de traitement du volume est présenté figure 4.24.

La gestion temps réel des programmes s'effectue à l'aide d'un jeu de drapeaux que l'on teste en cascade pour pouvoir traiter telle ou telle tâche.

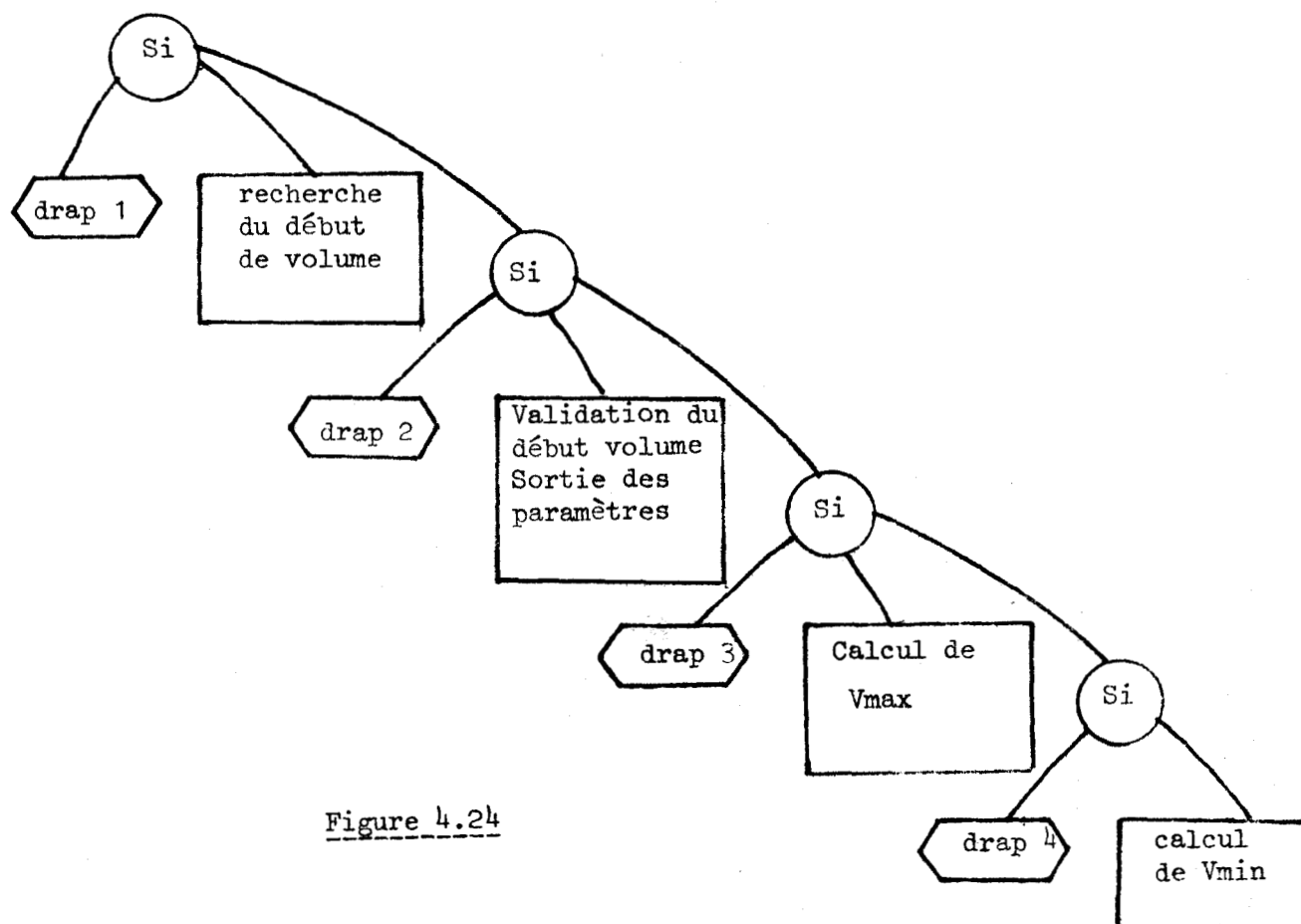


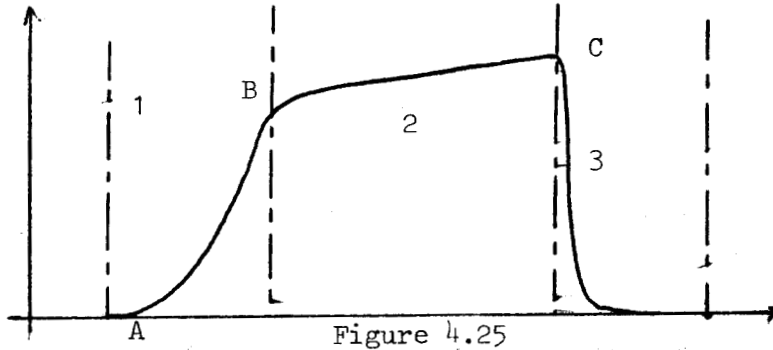
Figure 4.24

Les différents traitements se font par tests des dérivées et des amplitudes afin de déterminer un minimum, un maximum ou une fin de cycle.

A chaque fin de traitement, les drapeaux drap 1 drap 2 ... sont repositionnés pour aiguiller le calcul vers une autre tâche.

4.4.2. - Traitement de la concentration de CO_2

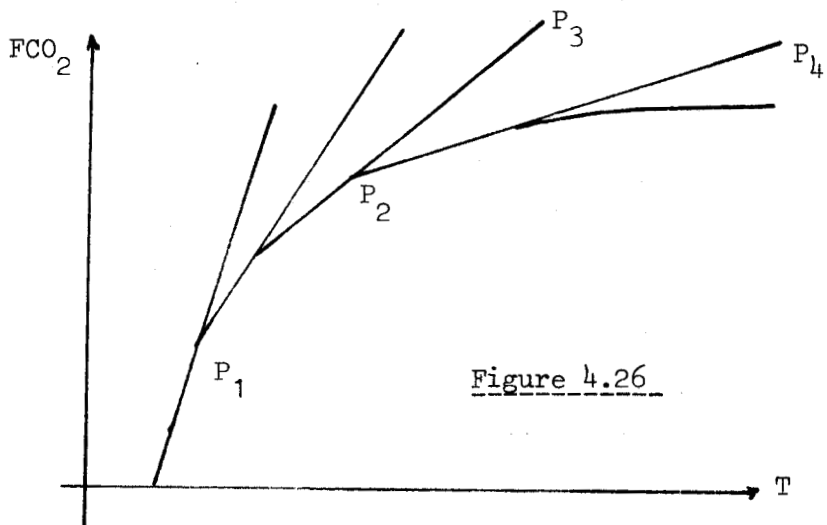
Pour assurer le suivi de la courbe de concentration, nous allons la partager en tronçons et rechercher les points caractéristiques du signal tout en envisageant pour chaque tronçon les anomalies potentielles pouvant perturber le cas normal.



Tronçon 1

L'enveloppe de la concentration est déterminée par plusieurs points dont le plus délicat à trouver, surtout en cas d'anomalie est le point B. Nous allons passer en revue les cas typiques survenant sur le tronçon 1.

Pour bien cerner l'arrondi de la courbe au niveau du point B, on teste 3 pentes que l'on considère tangentes à la courbe (p_1 , p_2 et p_3) et une 4^{ème} pente p_4 qui n'apparaît pas pour une forme F_S .

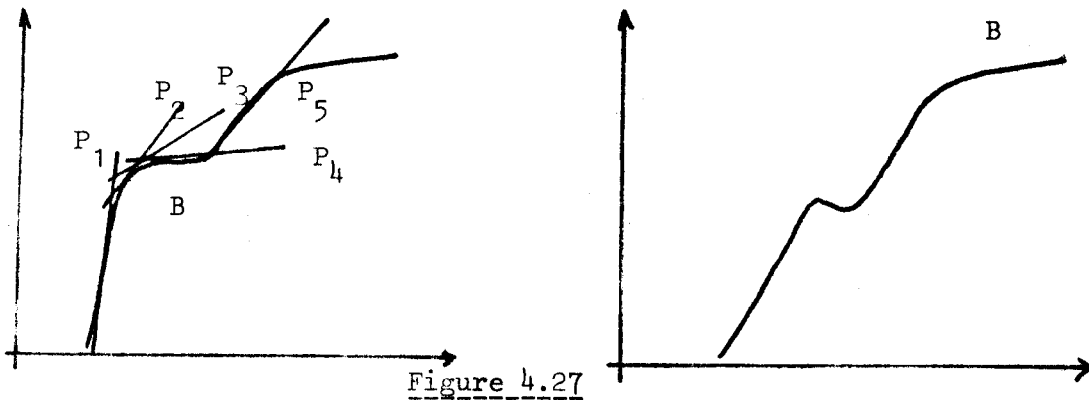


$$p_2 = \alpha_1 p_1$$

$$p_3 = \alpha_2 p_2 \quad (0 < \alpha_1, \alpha_2, \alpha_3 < 1)$$

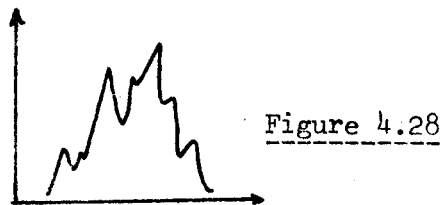
$$p_4 = \alpha_3 p_2$$

Dans ces conditions les anomalies intervenant au niveau du tronçon 1 ont les configurations suivantes :

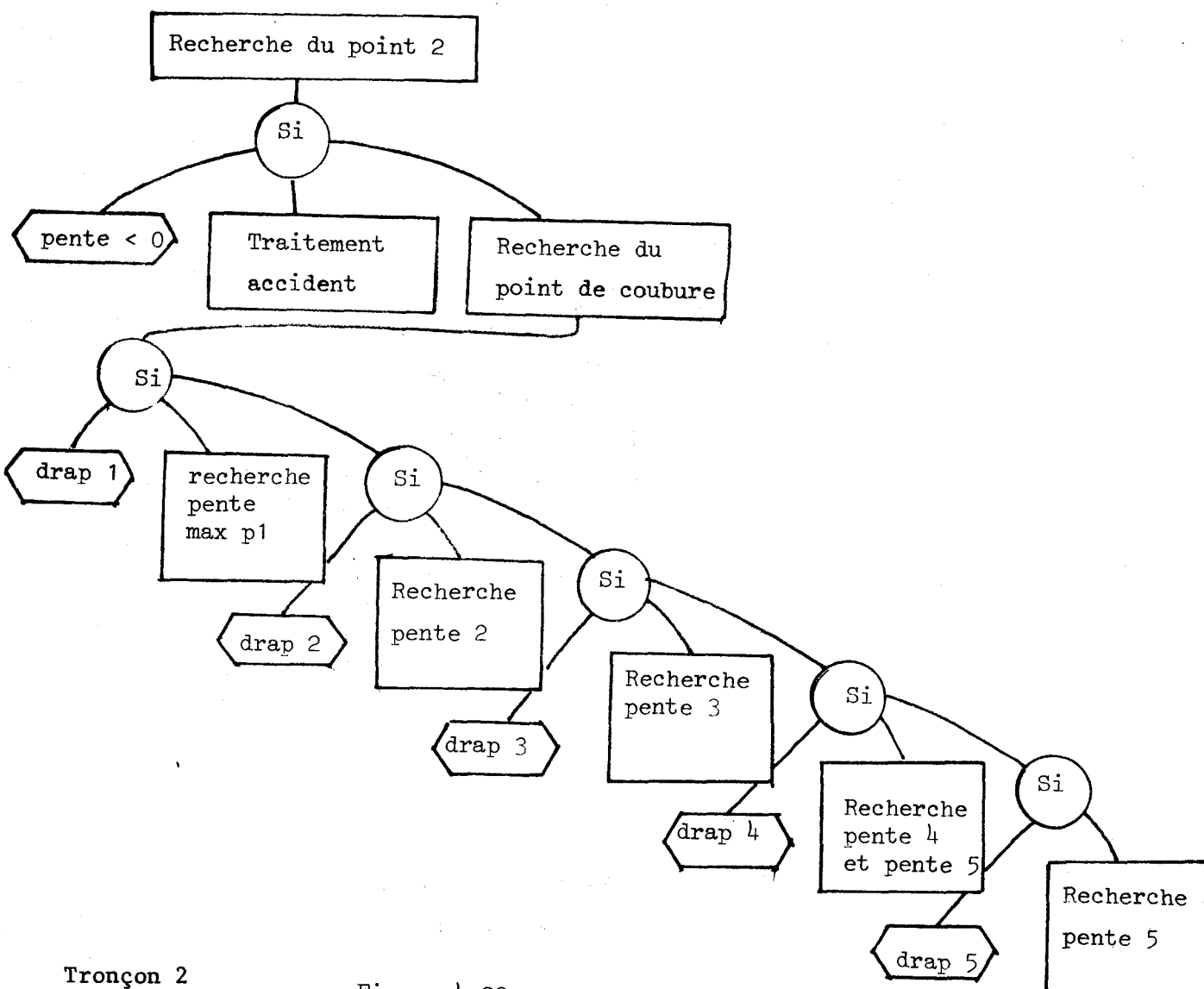


Dans le 1er cas, on doit détecter une remontée de la pente p_3 en p_5 avec $p_5 = \alpha_5 p_4$ ($\alpha_5 > 1$) conduisant au point B qui infirme B'.

Dans le 2ème cas, on doit tester un changement de signe de la dérivée avant toute recherche de courbure. Tant que la dérivée est négative, on attend l'arrivée d'un minimum que l'on teste pour savoir si il y a désadaptation totale ou seulement anomalie au niveau du tronçon 1. En cas de désadaptation totale, le tronçon 1 n'existe pas (figure 4.28).



L'arbre de programmation correspondant est présenté ci-dessous.



Tronçon 2

Figure 4.29

En ce qui concerne le tronçon 2, on effectue simplement un recensement des anomalies en surveillant les pentes et leurs changements de signe.

En particulier on stocke les amplitudes et la fréquence de ces anomalies afin de pouvoir différencier les oscillations cardiaques d'une désadaptation lors de la classification.

L'arbre programmatique de l'algorithme global de traitement du signal de CO₂ est donné par la figure suivante.

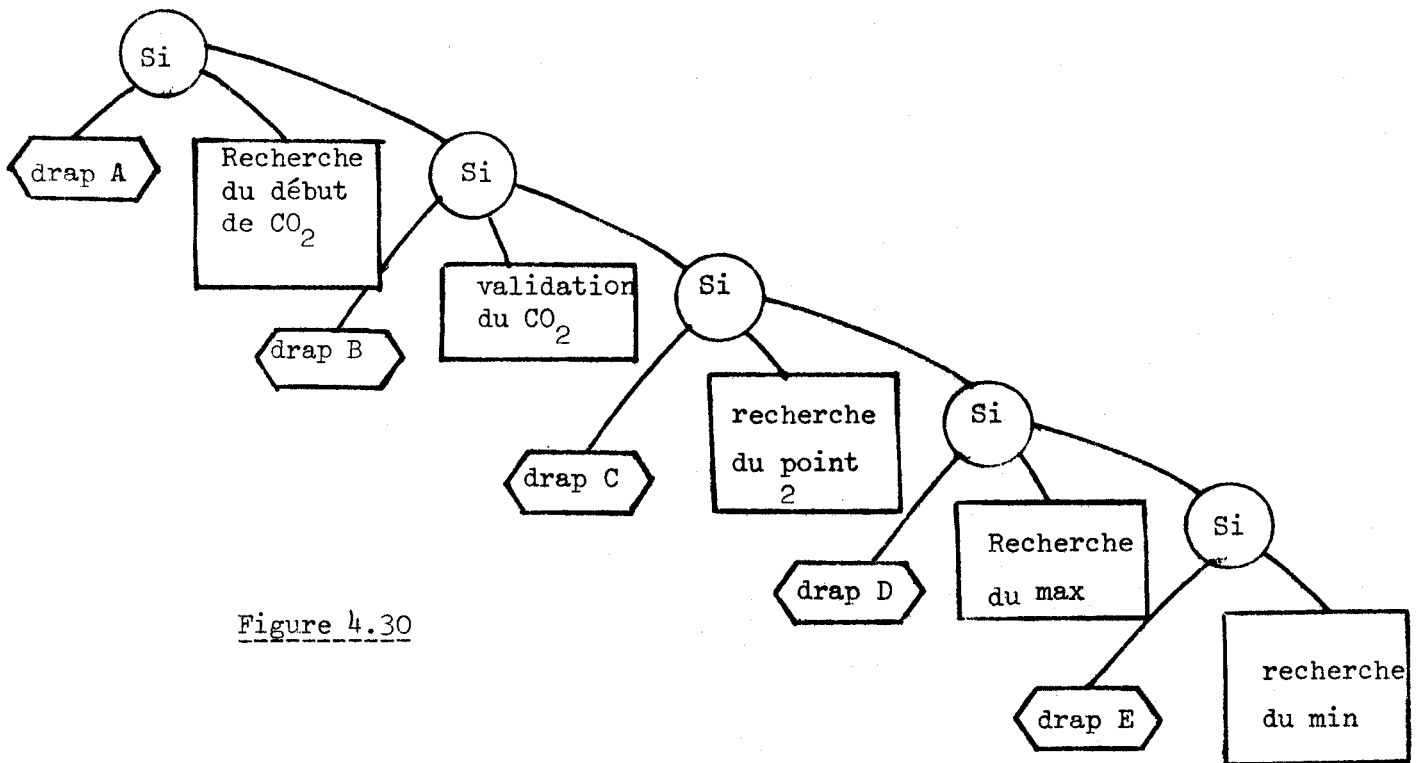


Figure 4.30

Chaque tâche est activée selon le positionnement des drapeaux. La gestion des drapeaux est réalisée à la fin de chaque tâche.

4.4.3. - Traitement de la pression

Le signal est divisé en tronçons et conduit à une approche du problème similaire à celle faite pour la concentration en CO_2

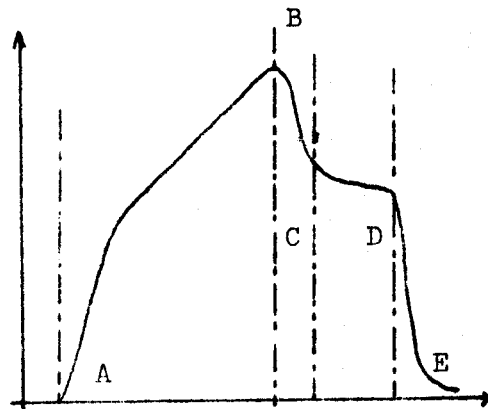
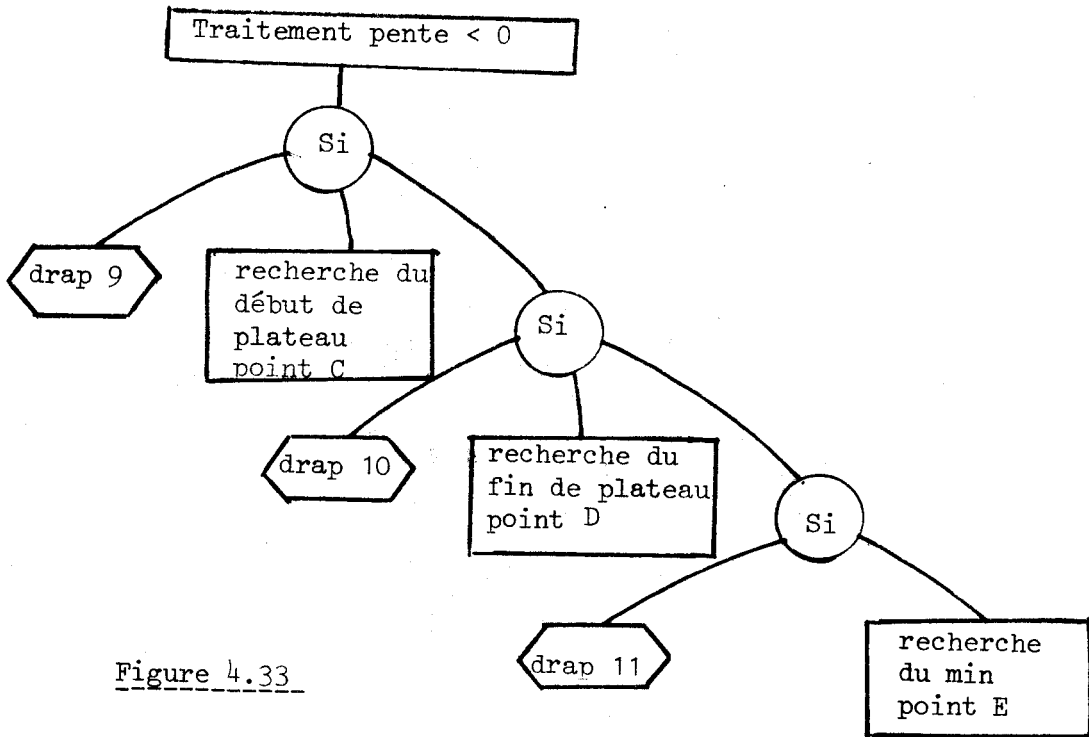
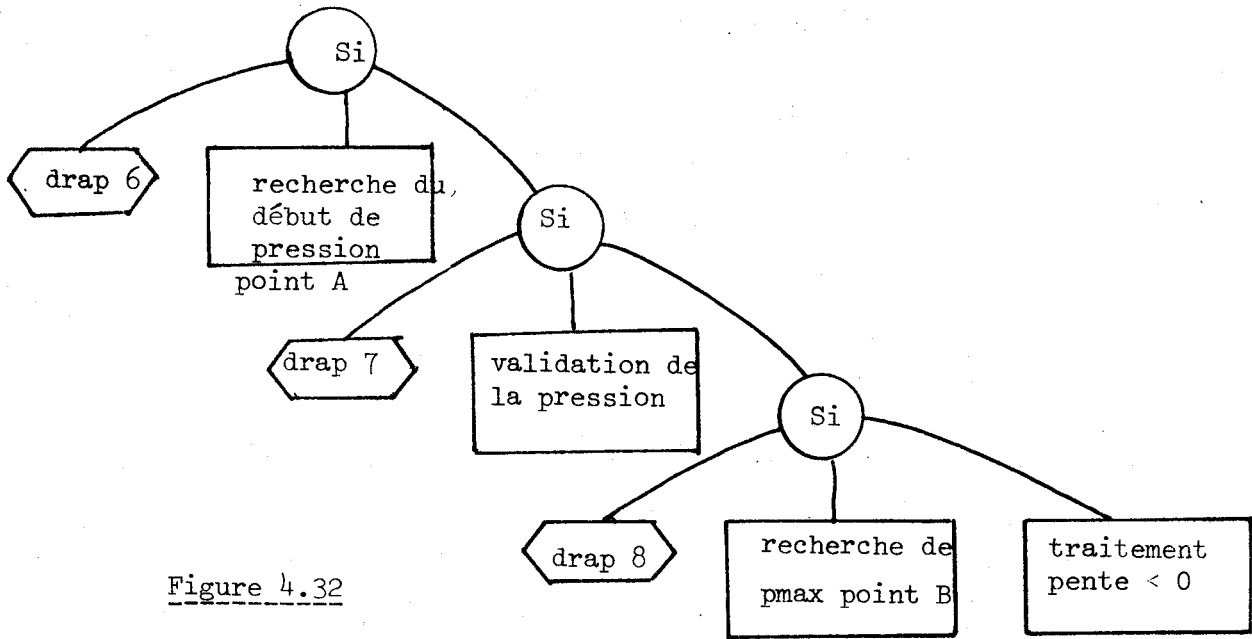


Figure 4.31

Les points caractéristiques à détecter sont ici B C D. L'algorithme correspondant est illustré ci-dessous :



4.5 - ETUDE DANS LE TEMPS DES FORMES

L'isolation et la classification des cycles ayant été effectuées à partir des paramètres extraits précédemment, l'étape suivante de l'étude consiste à définir une méthode de suivi des signaux.

4.5.1. - Création d'Etat

Pour étudier l'évolution d'un signal dans le temps, il convient de générer des séquences telles que celles définies au chapitre précédent. Ces séquences sont caractérisées par leur taux de gravité qui est égal à la somme des degrés de gravité des événements appartenant à la séquence.

A titre d'exemple, nous allons examiner l'évolution dans le temps de la concentration en CO_2 pour laquelle on définit les 2 types d'évènements.

- . statiques
- . dynamiques

Les événements statiques sont associés aux formes F_N et F_S c'est-à-dire aux formes normales.

En effet F_N est une forme normale pour un individu sain
et F_S " " pour un individu asthmatique.

Les événements dynamiques sont quant à eux répertoriés ci-dessous.

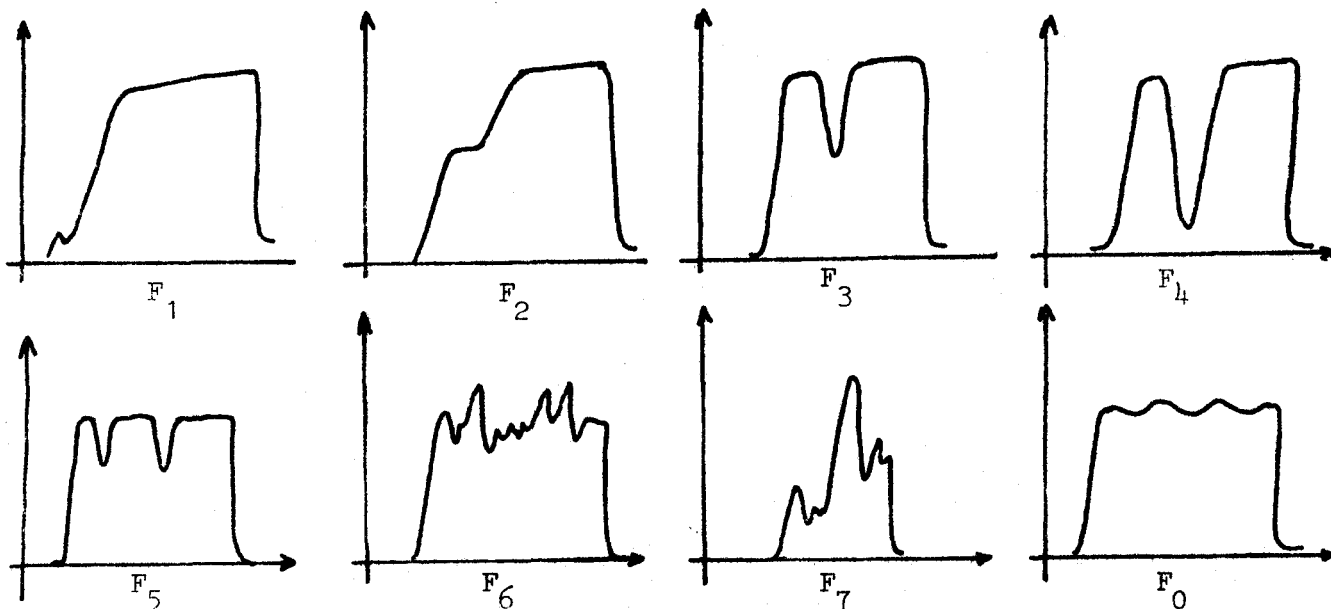


Figure 4.34

L'échelle de gravité correspondante est alors la suivante :

très grave	grave	moyen	bon
$F_6 F_7$	$F_5 F_4$	$F_3 F_2 F_0$	$F_1 F_N F_S$

Les valeurs choisies pour ces degrés de gravité ont été définies par apprentissage et ont conduit à la détermination suivante.

$$X(F_N) = X(F_S) = 0.0001$$

$$X(F_1) = 0.005$$

$$X(F_2) = X(F_3) = 0.05$$

$$X(F_4) = 0.5$$

$$X(F_5) = 1$$

$$X(F_6) = 10$$

$$X(F_7) = 50$$

- séquences

Les séquences sont définies sur un ensemble minimum de 20 cycles $n = 20$

- classification

4 états E_1, E_2, E_3 et E_4 ont été définis selon les règles suivantes.

$$\begin{aligned} \varphi(\Sigma) &\leq 0.01 \rightarrow E_1 \text{ état bon} \\ 0.01 < \varphi(\Sigma) &\leq 1 \rightarrow E_2 \text{ état moyen} \\ 1 < \varphi(\Sigma) &\leq 100 \rightarrow E_3 \text{ état grave} \\ 100 < \varphi(\Sigma) &\rightarrow E_4 \text{ état très grave.} \end{aligned}$$

4.6 - PROTOCOLE DE TRAVAIL

Les algorithmes de traitement ont été mis au point sur un ordinateur DEC MINC de 64 K de mémoires fonctionnant sous système RT 11.

4.6.1. - Acquisition des signaux

Pour la mise au point des algorithmes, nous avons enregistré les signaux de CO₂, volume, pression sur un enregistreur 4 voix Hewlett Packard. Chaque enregistrement utilise des bandes magnétiques d'1/4 de pouce et enregistre en modulation de fréquence sur 4 canaux. Il y a 6 vitesses d'enregistrement dépendant de la fréquence utilisée.

Nous avons structuré les programmes temps réel, de manière à faciliter la transposition de l'algorithme de traitement des signaux sur microprocesseur.

La mise en forme des signaux comporte 2 opérations (voir figure 4.35)

- un filtrage analogique par un filtre passe bas de fréquence de coupure 15 Hz et de pente - 12 dB/octave.
- une amplification pour l'acquisition des signaux sur ordinateur.

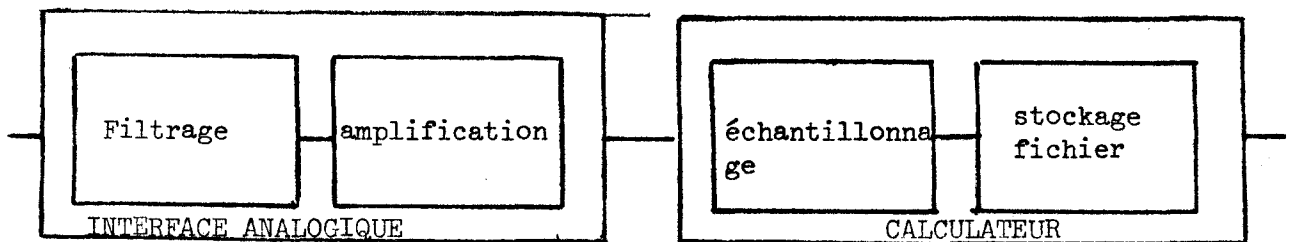


Figure 4.35

L'acquisition des données échantillonnées se fait par l'intermédiaire des Convertisseurs Analogiques/Numériques du MINC à une fréquence de 50 Hz.

Par ailleurs le tampon d'entrée recevant les échantillons est limité à 5500 cases mémoires. Pour créer un fichier conséquent, bien que limité par la capacité mémoire du MINC, nous devons donc faire un transfert du tampon sur disquette après chaque série d'acquisitions. Ceci demande un arrêt du magnétophone à chaque transfert puis une remise en route lors de la série d'échantillons suivante.

Le programme d'acquisition des données est illustré figure 4.36.

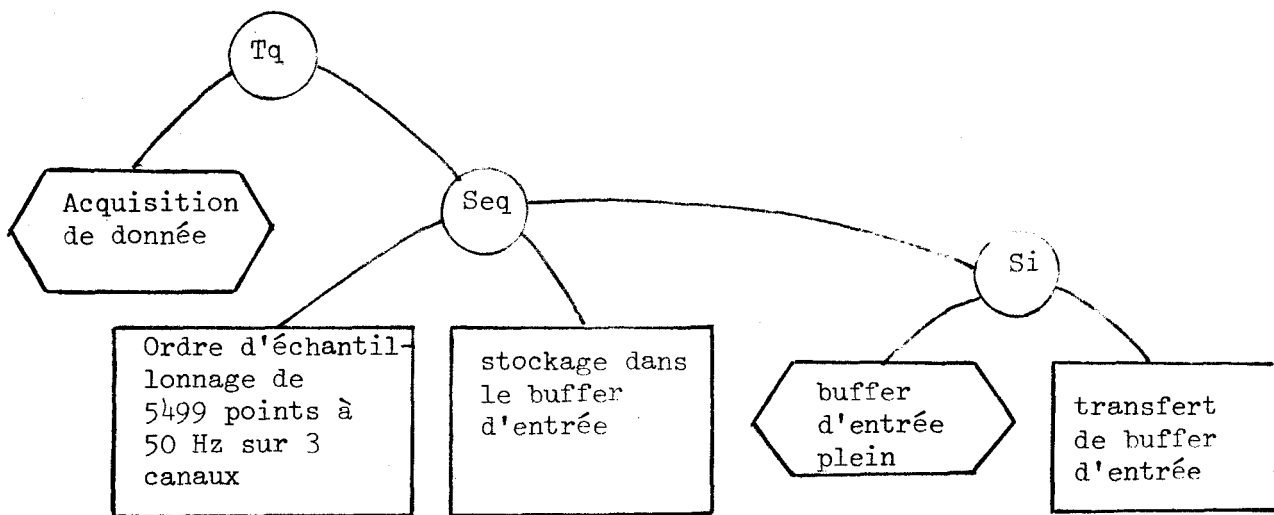


Figure 4.36

Tous les programmes ont été écrits en PASCAL OMSI du MINC.

4.6.2. - Gestion des fichiers

Pour la mise au point du programme une simulation temps réel a été effectuée; le minicalculateur ne permettant pas une évolution temps réel effective.

La simulation a consisté en une gestion de fichier . Plus précisément les données, au lieu d'être prélevées directement sur les convertisseurs A/D afin d'être traitées, ont été lues sur des fichiers.

4.5.3. - Résultats et discussions

Les fichiers utilisés nécessitent un emplacement mémoire important, ce qui limite assez rapidement les possibilités d'expérimentation et de validation

en "vraie grandeur" de la méthode.

Les essais effectués ont toutefois permis de valider la méthode jusqu'au stade de la classification temps réel. La validation de la méthode de prédiction nécessite quant à elle la réalisation d'une machine de traitement temps réel permettant de tester l'algorithme en clinique.

Cette machine construite autour de cartes industrielles à base de microprocesseur fait l'objet d'une étude spécifique en cours au Laboratoire d'Automatique.

4.7 - CONCLUSION

Ce travail a pris appui sur un cahier des charges bien défini. Il a toutefois nécessité une longue observation des signaux afin de permettre d'envisager toutes les possibilités de formes susceptibles d'apparaître. En particulier les perturbations bien que négligeables pour le cahier des charges, doivent être prises en compte pour lever toutes ambiguïtés au niveau de la détection des formes.

L'étude réalisée a résulté en la constitution d'un outil d'analyse visant non seulement à effectuer une classification temps réel des formes apparaissant mais également à fournir une prédiction résultant de l'observation de ces formes.

La validation en clinique ne pourra alors s'effectuer, qu'en implantant les méthodes sur des systèmes de traitement numérique temps réel dédiées à cette application spécifique.

- C O N C L U S I O N -

L'étude effectuée a permis de montrer la faisabilité d'un outil de surveillance basé sur l'analyse temps réel de signaux prélevés simultanément sur un même processus.

Les méthodes utilisées tiennent particulièrement compte des observations expérimentales effectuées sur les signaux eux-mêmes. Cette démarche subjective permet d'opérer une présélection simple des paramètres utilisés, alors que des outils mathématiques purement objectifs ne le permettent pas toujours.

Ceci conduit à envisager l'étude des signaux sous un aspect structurel. Pour cela et en s'appuyant sur le caractère cyclique de la plupart des phénomènes étudiés, on a défini une méthode de paramétrisation des cycles observés.

Cette description qui est de nature syntaxique nous a ensuite conduit à utiliser une méthode informationnelle pour étudier les relations entre signaux . Dans ce cadre nous avons utilisé la notion d'entropie que nous avons adaptée afin notamment de prendre en compte les aspects temporels (décalage entre signaux ...).

L'étape suivante de l'étude a consisté à définir une méthode de prédiction de l'évolution des signaux. Dans ce but un nombre restreint de situations caractérisant l'activité du système a été défini. Dès lors l'adoption d'une représentation géométrique élémentaire des paramètres statistiques décrivant ces situations permet de mettre en évidence des trajectoires de tendance. Ces trajectoires constituent un outil simple de prédiction.

Les signaux d'origine électrophysiologiques constituent un support particulièrement intéressant à ce type de démarche. Toutefois la contrainte temps réel ne peut être pleinement satisfaite qu'en envisageant l'implantation du système de traitement sur une machine dédiée. Une telle machine dont le développement est actuellement en cours doit permettre de constituer une base de données de signaux utilisable pour valider les méthodes.

"BIBLIOGRAPHIE"

- . 1 . A.S GINSBERG, F.L. OFFENSEND
" An application of decision theory to a medical diagnosis "
I.E.E.E. Trans. Systems Sciences
SSC n°3 Sept.68 - pp.355-362
- . 2 . D.B DRANE - W.B MARTIN
Pattern Recognition applied to sleep state classification
Electroenceph. Clin. Neurophysiol., 1969, 26 p.238
- . 3 . P.I. KUZNELZOV - L.A PCHEHNTZEV
" The application of some mathematical methods in medical diagnostics "
Mathematical Biosciences 5, 1969 - p.365-377
- . 4 . MOISEEVA and USOV
" Some medical and Mathematical Aspects of the Computer diagnostics "
Proceeding of the I.E.E.E. vol.57 n°11 - November 69
- . 5 . P.S BOWLING and J.R BOURNE
" Discriminant analysis of Electroencephalograms recorded from renal patients "
Biomedical Engineering January 78 n°1
- . 6 . D. GUSTAFSON - A. WILLSKY - JYH - YUN WANG - M.C LANCASTER - J.H TRIEBWASSER
ECG/VCG Rhythm diagnosis using statistical signal analysis
I.E.E.E. Transactions of biomedical engineering vol BME-25 n°4
July 1978
- . 7 . M. BEN BASSAT - R.W. CARLSON and ALL
Pattern - Based Interactive diagnosis of multiple disorders :
the Medas System.
I.E.E.E. Transactions on pattern analysis and machine intelligence,
vol. PAMI-2, n°2 - March 1980

- . 8 . JAN KRYSPIN, ALAN M. NORWICH
Application of information calculus to Medical data analysis and reduction
 Mathematical Biosciences A, 165-172 - 1973.

- . 9 . OKAYIMA, STARK, WHIPPLE, YASUI
Computer pattern recognition technique : some results with real Electrocardiographic data "
 I.E.E.E. Transactions of biomedical Electronics BME-10
 pp. 106-114 - July 1963.

- . 11 . RIICHIRO MIZOGUCHI - MASAMICHI SCHIMURA
À non parametric algorithms for detecting clusters using Hierarchical structure
 I.E.E.E. Transactions on pattern analysis and machine Intelligence
 vol. PAMI-2, n°4 - July 1980.

- . 12 . J.E. ALBUS
Electrocardiogram Interpretation Using stochastic finite state Model. in syntactic Pattern Recognition application
 Editor K.S. FU, Springer-Verlag, Berlin 1976.

- . 13 . B. PICINBONO
Introduction à l'étude des signaux et phénomènes aléatoires
 Dunod Paris 1977

- . 14 . J.P BENZECRI & collaborateurs
" La taxinomie "
 Dunod 1980 - Paris 3° Edition

- . 15 . J.P BENZECRI & collaborateurs
" L'analyse des correspondances "
 Dunod 1980 - Paris 3° Edition

.../...

- . 16 . L. LEBART, A. MORINEAU, N. TABARD
Techniques de la description statistique
Dunod 1977

- . 17 . P. BERTIER - BOUROCHE J.M
Analyse des données multidimensionnelles
PUF 1975.

- . 18 . Michel VOLLE
Analyse des données cours de l'ENSAE 1977

- . 19 . EDWIN DIDAY
Problems of clustering and recent advances
Invited paper at the 11th European Congress of statistics
OSLO August 1978.

- . 20 . M. LOSFELD
Documents pour le cours de D.E.A. - Analyse des Données
Enseignement de 3° Cycle, Université des Sciences & Techniques de Lille 1

- . 21 . M. LERMAN
Les bases de la classification automatique
Paris 1970 - Gauthier Villars

- . 22 . HARRY C. ANDREWS
" Introduction to mathematical techniques in pattern recognition "
WILEY INTERSCIENCE 1972

- . 23 . Michel HUGHES
Initiation mathématique aux grammaires formelles
Librairie Larousse 1972

- . 24 . K.S FU
sequential methods in pattern recognition and machine learning
Academic Press 1968

- . 25 . MICLET - KALIZEWSKI - GRENIER
Reconnaissance des formes - ENST 1978

- . 26 . C.V. RAJAGOPALAN
*segmentation des suites temporelles application à l'extraction
heuristique des paramètres fondamentaux du cardiorespirogramme
en pathologie néonatale*
Thèse de Docteur-Ingénieur 1979, Lille

- . 27 . C. VASSEUR
*La notion d'évènement dans les systèmes dynamiques : détection ,
classification temps réel et application à la conception d'une instrumenta-
tion distribuée - Thèse de Docteur ès Sciences Physiques - 2 Juin 1982*

- . 28 . B. CHEBEL
*Implantation sur microprocesseur d'un système de surveillance
respiratoire*
Mémoire de D.E.A. 1980

- . 29 . C. VASSEUR
*A microprocessor oriented segmentation technique and efficient tool
for electrophysiological signal analysis*
EEMTIC 79, OTTAWA - May 1979

- . 30 . KEINOSUKE FUKUNAGA
Introduction to statistical pattern recognition
Academic press 1972 - New York San Fransisco London

- . 31 . G.J KLIR
" On the representation of activity arrays "
International journal of general systems - vol.2, n°43
pp 149-168 - 1975

- . 32 . G.J KLIR
" Identification of generative structures in Empirical data "
international journal of general systems vol.3 - pp.89-104 - 1976.

.../...

- . 33 . G.J KLIR
 *" On the problems of computer-aided structure identification some
 experimental observations and resulting guidelines "*
 International journal of Man Machine studies
 vol.9 - pp. 593-623, 1977

- . 34 . V.M TORO CORDOBA
 *" Contribution à l'analyse structurale de systèmes complexes
 à l'aide de l'entropie et ses généralisations "*
 Thèse 3° Cycle soutenue le 8 Mars 1982 - Lille

- . 35 . J.N CRONHOLM
 *" a general method of obtaining exact sampling probabilities
 of the Shannon - Wiener measure of information H "*
 Psychometrika, vol.28 n° 4 pp 405-413 - 1963.

- . 36 . J. SALLATIN
 *" représentations d'observations dans le contexte de la théorie
 de l'information "*
 Thèse d'Etat - Université Pierre et Marie Curie - Paris 14 Juin 1979

- . 37 . ROGER C. CONANT
 " Laws of information which govern systems "
 I.E.E.E. Transactions on systems, man, and cybernetics
 vol. SMC-6 n°4 April 1976

- . 38 . ARTHUR C. SANDERSON, JAKUB SEGEN, E. RICHEY
 Hierarchical Modeling of EEG Signals
 I.E.E.E. Transactions of pattern analysis and Machine Intelligence
 vol. PAMI-2 n°5, Sept.1980

- . 39 . Y. ROZANOV
 Processus Aléatoires
 Traduction française - Edition Mir 1975 - Moscou

.../...

- . 40 . FERGUSON T.S
 Mathematical statistics
 1967 Academic Press N.Y San Francisco London

- . 41 . M. DEKKER
 Markov chains
 New-York 1970 Academic Press

- . 42 . MANABU ICHINO
 *Multiclass Pattern Recognition systems based on independant
 subrecognition systems*
 I.E.E.E. Transactions on systems man and Cybernetics
 vol. SMC-6 n°4, April 1976

- . 43 . FOURGEAU D.C FUCHS A.
 statistique
 Année 1972 - Academic Press

- . 44 . STUART F. SULLIVAN - RICHARD N. PATTERSON - EMANUEL M.PAPPER
 *Evaluation of pulmonary function data using
 Computer time - sharing*
 Proceeding of the I.E.E.E. Vol.57 n°11 November 1969

- . 45 . S. TURNEY - T. Mc ASLAN - R. COWLEY
 The continous measurement of pulmonary gas exchanges and mechanics
 Ann Thorac Surg 13 229-242 - 1972

- . 46 . S. TURNEY - T. Mc CLUGGAGE - W. BLUMENFELD - T. Mc ASLAN - R. COWLEY
 Automatic respiratory gas monitoring
 Ann Thorac Surg 14 - 159-171 - 1972

- . 47 . M.C CHAMBRIN
 *" Réalisation à l'aide de microprocesseurs d'un système de surveillance
 de l'efficacité ventilatoire en réanimation respiratoire "*
 Thèse de Docteur-Ingénieur, Lille 1979.



R É S U M É

Ce travail vise à définir une méthodologie de description d'analyse et de traitement du signal prenant en compte la structure de celui-ci ainsi que son contenu informationnel.

Le premier chapitre du mémoire décrit, dans un premier temps, les méthodes permettant de donner aux signaux une structure de messages décomposables en primitives successives. Dans un second temps, les messages ainsi constitués font l'objet d'une procédure de reconnaissance hiérarchisée à deux niveaux : reconnaissance syntaxique au premier niveau, puis affinement par une reconnaissance métrique au second niveau.

Le deuxième chapitre propose une adaptation du calcul de transformation afin d'intégrer les spécificités des suites temporelles dans l'évaluation des relations pouvant lier plusieurs signaux à structure de messages. C'est ainsi que les retards (déphasages) d'un signal sur un autre ont pu être pris en compte permettant de dégager un indice de causalité d'un signal sur un autre.

Dans le troisième chapitre, on définit un histogramme décrivant la statistique des différents états pris par un signal sur un horizon temporel donné. Les variations de cet histogramme induites par un glissement dans le temps constituent alors un indicateur de prédiction de l'évolution du signal. Une interprétation géométrique élémentaire de cette opération de différentiation permet de définir des trajectoires de tendance facilement interprétables.

Finalement, le dernier chapitre propose une application de la méthodologie à la définition d'un dispositif de surveillance des signaux respiratoires d'un patient placé en réanimation.