



**HAL**  
open science

## Modélisation des dysfonctionnements d'un système dans le cadre d'activités de maintenance

Pascal Vrignat, Manuel Avila, Florent Duculty, Frédéric Kratz

### ► To cite this version:

Pascal Vrignat, Manuel Avila, Florent Duculty, Frédéric Kratz. Modélisation des dysfonctionnements d'un système dans le cadre d'activités de maintenance. 16ème Congrès de Maîtrise des Risques et de Sécurité de Fonctionnement, Oct 2008, AVIGNON, France. pp.Communication 4A-1. hal-00594893

**HAL Id: hal-00594893**

**<https://hal.science/hal-00594893>**

Submitted on 23 Jun 2011

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# MODELISATION DES DYSFONCTIONNEMENTS D'UN SYSTEME DANS LE CADRE D'ACTIVITES DE MAINTENANCE

## Modelling of the dysfunctions of a system (in maintenance activities)

Pascal VRIGNAT , Manuel AVILA et Florent DUCULTY, Frédéric KRATZ.  
Institut PRISME  
Equipe MCDS  
IUT de l'Indre  
2, Av. François MITTERRAND  
36000 Châteauroux

### Résumé

Les événements qui précèdent une panne sont souvent récurrents. Une suite d'événements particuliers peut, éventuellement, informer d'une panne prochaine. Notre étude tente d'appréhender « cette signature » à l'aide d'un Modèle de Markov Caché. Les premiers résultats montrent que notre approche permet d'apporter une aide à la décision. Ce système expert peut permettre de reconsidérer les priorités d'activités de maintenance journalières.

### Summary

The events preceding a failure are often recurring. A series of special events can eventually inform a blackout coming soon. In this way, the result of our study is the implementation of an Hidden Markov Model. Early results show that our approach will provide assistance to the decision. This expert system can reconsider the priorities of activities of daily maintenance.

## 1. Introduction

Hypothèse : les événements qui précèdent une panne sont souvent récurrents. Une suite d'événements particuliers peut, éventuellement, informer d'une panne prochaine ! Quelques exemples connus peuvent illustrer notre hypothèse.

- En mécanique : un bruit, une vibration précèdent la rupture. Une baisse de performances traduit une ou des anomalies.
- En informatique : un déplacement inopiné du pointeur, un ralentissement, un dysfonctionnement d'une application comme un navigateur Internet peuvent traduire la présence d'un virus sur l'ordinateur concerné.

Notre étude tente d'appréhender « cette signature » à l'aide d'un Modèle de Markov Caché ou HMM<sup>1</sup>. Le processus caché correspondra à l'état du système (ou sous-système) et les observations seront les informations observables sur le système.

Nos travaux de recherche présentés dans cet article, montrent qu'il est possible de modéliser le niveau de dégradation d'un processus (une peseuse volumétrique insérée dans un processus dit « continu » pour l'agro-alimentaire).

Les résultats montrent également, que notre démarche de travail peut apporter une aide à la décision dans l'organisation quotidienne de la maintenance (Figure 1).

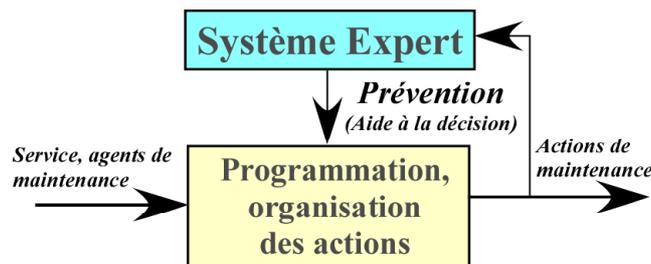


Figure 1 : Objectif de nos travaux

## 2. Les modèles stochastiques

Les modèles stochastiques sont des représentations de systèmes dynamiques basées sur les probabilités. Bien que cet aspect aléatoire semble assez fruste de prime abord, c'est en fait une représentation très naturelle. Souvent, lorsque l'on observe un système, on est capable de trouver intuitivement des règles que le système suit, au moins globalement. L'approche la plus employée consiste à prendre le cas moyen, celui qui ressemble le plus à toutes les observations, et à considérer qu'il modélise le système de façon suffisante. La variance du processus, autour du cas moyen, est tout simplement ignorée. L'approche statistique consiste à tenir compte de chacune des différentes possibilités, tout en lui attribuant une probabilité qui exprime la fréquence de ses occurrences. Ainsi, plus un cas se produira souvent, plus sa probabilité sera importante. L'étude des processus stochastiques a commencé au début du 20<sup>ème</sup> siècle grâce à un mathématicien Russe, Markov Andreï Andreïevitch. Son étude statistique du langage l'a conduit à formuler l'hypothèse Markovienne, qui peut se résumer ainsi : « L'évolution future d'un système ne dépend que de son état présent ». Autrement dit, cette hypothèse implique que l'état courant du système contient toute l'information apportée par le passé. C'est donc une hypothèse très forte, mais qui semble relativement logique. En pratique, on constate que de nombreux systèmes enfreignent cette condition. Cependant, en affinant le modèle, on peut souvent le rendre Markovien. Des travaux originaux [10] montrent qu'il est possible de modéliser, à l'aide d'un processus markovien de saut, l'usure des pièces caténaïres pour la SNCF. La Figure 2 montre la hiérarchie de ces modèles. Dans ce diagramme, les modèles les plus simples sont au sommet, et chaque modèle dérivé ajoute de nouveaux paramètres (entre parenthèses), ainsi que de nouvelles fonctions.

<sup>1</sup> Hidden Markov Model

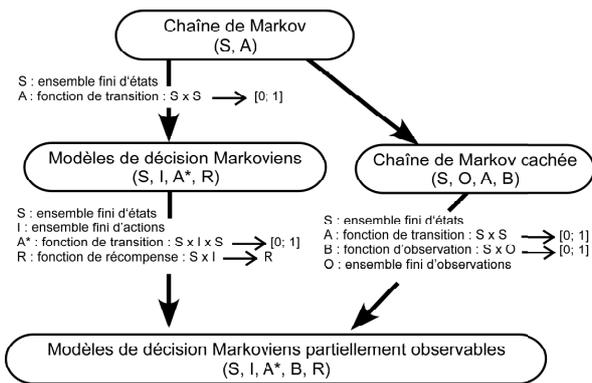


Figure 2 : Diagramme des modèles Markoviens [9]

## 2.1 Approches Markoviennes cachées

La théorie des Modèles de Markov Cachés (MMCs) a pour origine les travaux du mathématicien russe Markov en analyse du langage par chaînes à états interconnectés, dites chaînes de Markov. C'est au début des années 70 que les modèles Markoviens cachés ont véritablement connu leur essor en traitement du signal, essentiellement dans le cadre de la reconnaissance automatique de la parole. Ceci s'explique en partie par la découverte par Baum [3] d'une procédure générale d'ajustement des paramètres des modèles aux caractéristiques du signal à reconnaître et par Viterbi [17] d'un algorithme de décodage utilisé en théorie de l'information puis adapté par la suite au décodage des états cachés de la chaîne. Depuis lors, moyennant l'introduction quasi continue de nouveaux éléments de modélisation, le cadre théorique des modèles markoviens cachés s'est considérablement élargi tandis que les approches par MMC se sont généralisées à d'autres domaines d'application [4] :

- les travaux en reconnaissance de la parole [9], [6],
- biosciences [7], [15],
- climatologie [8],
- reconnaissance de l'écriture [1],
- vérification de l'identité (basée sur l'écriture) [14], [13],
- médecine (aide au diagnostic) [16],
- ...

Aujourd'hui, les réseaux bayésiens dynamiques peuvent être perçus comme les héritiers des MMCs. La richesse de modélisation est beaucoup plus importante avec des réseaux bayésiens dynamiques dans la mesure où le contexte utilisé pour définir les observations peut incorporer plusieurs états cachés. Ils permettent également de modéliser des processus en interaction et donc de fusionner différents types d'information. Le philosophe Dennett [5] parle d'un point de vue d'interprétation qui permet de comprendre les phénomènes qui sont régis par un mécanisme : « le point de vue du plan » (design stance), dont l'idée est qu'un objet se comporte en général « de la manière dont il est construit pour se comporter », c'est-à-dire selon son plan. D'après Dennett, de ce point de vue « presque tout le monde peut prédire le moment où un réveil-matin va sonner simplement en observant son état extérieur. Si l'on veut faire des prédictions plus précises et plus détaillées au niveau du plan, il faut descendre à un degré de description de son plan qui soit moins abstrait ; par exemple, au niveau où l'on décrit les engrenages, mais où leur composition matérielle n'est pas spécifiée » : autrement dit parce qu'on sait comment fonctionne théoriquement le réveil, on a besoin uniquement d'accéder à son état extérieur pour prédire son comportement.

Les chaînes de Markov Cachées modélisent des phénomènes aléatoires dont on suppose qu'ils sont composés à un premier niveau d'un processus aléatoire de transition entre des états inobservables (les états cachés) et, à un second niveau, d'un autre processus aléatoire qui, dans chaque état, engendre des valeurs observables (appelées observations). On retrouve donc dans ce formalisme une application de l'idée énoncée par Dennett de deux niveaux : l'état extérieur (les observations) et un degré de description plus détaillé que l'on n'a pas besoin d'observer parce qu'on connaît son fonctionnement théorique (on

possède un modèle théorique de ses états cachés). L'articulation de ces deux niveaux confère aux modèles de type MMC une grande flexibilité et ces modèles basés sur des transitions entre états sont bien adaptés pour rendre compte de processus organisés dans le temps. De plus, les modèles de Markov cachés permettent un alignement élastique entre un nombre quelconque d'observations temporelles et un ensemble d'états définis a priori.

### 2.1.1 Définitions formelles d'une chaîne de Markov cachée

De manière générale, les CMCs permettent de représenter des problèmes où les états ne sont pas directement observables, ce qui était supposé dans le cas précédent. Ces états sont donc cachés, mais ils peuvent être estimés à travers des processus stochastiques qui produisent les séquences d'observations. L.R. Rabiner [11] propose une présentation très générale des modèles de Markov cachés ou non. Ces modèles cachés permettent de représenter des processus plus complexes que ceux représentés par des MMs. La composante cachée est la réalisation  $q = \{q_1, q_2, \dots, q_T\}$  d'une suite de variables aléatoires  $Q = \{Q_1, Q_2, \dots, Q_T\}$  où les  $q_t$  prennent leurs valeurs dans l'ensemble  $S = \{s_1, s_2, \dots, s_N\}$  de  $N$  classes (les états). La suite observable  $o = \{o_1, o_2, \dots, o_T\}$  est la réalisation du processus aléatoire  $O = \{O_1, O_2, \dots, O_T\}$  où les  $o_t$  prennent leurs valeurs dans l'ensemble  $V$  de  $M$  classes (les symboles). La propriété de Markov est supposée vraie pour la composante cachée :  $\forall t \in [2 \dots T]$ ,

$$P(Q_t = q_t | Q_1 = q_1 \dots Q_{t-1} = q_{t-1}) = P(Q_t = q_t | Q_{t-1} = q_{t-1})$$

Nous avons pour cette composante les définitions suivantes :

\* loi de probabilité initiale :  $P(Q_1 = s_j) = \pi_j, 1 \leq j \leq N$ ,

\* matrice des probabilités de transition :

$$P(Q_t = s_j | Q_{t-1} = s_i) = a_{ij}, 1 \leq i, j \leq N, t \in [2 \dots T]$$

Les éléments d'une CMC sont donc :

\*  $N$  : le nombre d'états du modèle.

\*  $M$  : le nombre de symboles différents générés par la CMC. Ces symboles forment l'ensemble  $V = \{v_1, v_2, \dots, v_m\}$ .

\*  $A = [\dots a_{ij} \dots]$  : matrice de distribution des probabilités de transition. Cette matrice est de dimension  $N \times N$ . Le terme générique  $a_{ij}$  désigne la probabilité de transiter d'un état  $s_i$  vers un état  $s_j$ . La matrice est stochastique, i.e. la relation

$$\sum_{j=1}^N a_{ij} = 1, \forall i \in [1 \dots N]$$

est vérifiée. La définition même de cette matrice suppose que la connectivité est a priori totale entre tous les états. Certaines liaisons pouvant toutefois être affectées de probabilités de transition nulles (Figure 3).

\*  $B = [\dots b_j(k) \dots]$  : matrice de distribution des probabilités de génération des symboles. Cette matrice est de dimension  $N \times M$ .

Le terme générique  $b_j(k) = P(O_t = v_k | Q_t = s_j)$  désigne la probabilité pour la CMC se trouvant dans l'état  $s_j$  génère le symbole de rang  $k$ ,  $v_k$ . La matrice  $B$  est stochastique, i.e.

$$\sum_{k=1}^M b_j(k) = 1, \forall j \in [1 \dots N].$$

\*  $\Pi = [\dots \pi_i \dots]$  matrice de distribution des probabilités initiales.

Cette matrice est de dimension  $1 \times N$ . Le terme générique  $\pi_i$  désigne la probabilité que la CMC se situe dans l'état  $s_i$  à

l'instant initial. La matrice  $\Pi$  est stochastique, i.e.  $\sum_{i=1}^N \pi_i = 1$ .

En résumé, une CMC est définie par le triplet de matrices noté

$\lambda = (A, B, \Pi)$  qui suppose implicitement la connaissance de  $N$  (nombre d'états) et de  $M$  (nombre de symboles).

### 2.1.2 Graphe d'états

La topologie d'une CMC dépend directement des éléments non nuls de la matrice de transition. En particulier, une CMC dont la matrice  $A$  est pleine (aucun terme nul) est dite ergodique : toutes les possibilités d'évolution sont permises (Figure 3 (a)). Les CMCs dont la matrice de transition est triangulaire supérieure sont appelées modèles gauches-droits (Figure 3 (b)). Ils sont particulièrement adaptés à la modélisation de processus temporels.

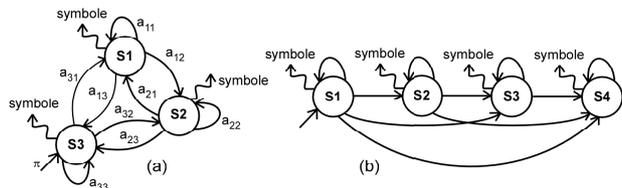


Figure 3 : Graphe d'états d'une chaîne de Markov cachée complète. (a) chaîne à topologie ergodique ( $N = 3$ ) et (b) chaîne à topologie gauche-droite ( $N = 4$ )

### 2.1.3 Approche des CMCs par l'exemple

Considérons un lancer de pièces de monnaie. Nous disposons de trois pièces (biaisées). Le lanceur ne nous informe pas de la pièce qu'il a choisie, et pour une observation donnée, les symboles (pile ou face) peuvent être générés tour-à-tour par la pièce une, deux ou trois. Une chaîne de Markov cachée modélise l'expérience de la manière suivante (Figure 4).

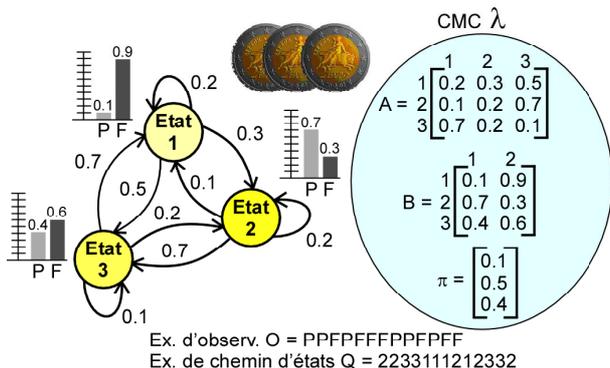


Figure 4 : Exemple d'une chaîne de Markov cachée (modélisation d'un lancer avec 3 pièces de monnaie)

On considère qu'un état modélise une pièce, et que cet état peut produire deux symboles : pile et face. On admet alors que l'on peut passer d'un état à un autre (c'est-à-dire changer de pièce) suivant certaines lois de probabilités, et que chaque pièce possède sa propre distribution de probabilités de générer les symboles pile et face. La séquence de (pile-face) est toujours observable, mais la séquence d'états qui a engendré cette séquence n'est pas observable. On dit alors qu'elle est cachée. Les probabilités de transitions entre les états (non observables) sont stockées dans une matrice notée  $A$ , alors que les probabilités d'apparitions des symboles (observables) sont mises dans une matrice notée  $B$ .

## 3. Aspects algorithmiques

Les CMCs trouvent leur intérêt dans la résolution de trois problèmes fondamentaux que sont l'évaluation, le décodage et l'apprentissage [11]. Nous utilisons l'évaluation et l'apprentissage.

### 3.1 Premier problème : l'évaluation

Etant données une CMC  $\lambda$  et une observation  $o$ , on cherche à calculer la vraisemblance de l'observation  $o$  avec la CMC  $\lambda$ , c'est-à-dire avec quelle probabilité la CMC  $\lambda$  engendre l'observation  $o$ . Cette valeur est notée  $P(O = o | \lambda)$ . Un calcul direct de cette valeur est bien sûr possible, mais très « coûteux » en nombre d'opérations car il énumérerait tous les chemins  $q$  possibles tel que  $P(O = o | \lambda) = \sum_q P(Q = q, O = o | \lambda)$ .

On préfère utiliser un algorithme de complexité moindre : l'algorithme Forward [11] qui réalise le calcul en  $N^2T$  opérations contre  $2TN^T$  opérations pour le calcul direct (avec  $N$  le nombre d'états et  $T$  la longueur de l'observation). Pour cela, on définit la variable Forward comme la probabilité d'observer la suite  $o = (o_1, o_2, \dots, o_t)$  à l'instant  $t$  dans l'état  $s_i$  en connaissant la CMC  $\lambda$ .  $\alpha_t(i) = P(o_1 \dots o_t; q_t = s_i | \lambda)$

A l'instant initial ( $t = 1$ ), la variable Forward  $\alpha_1(i)$  est la probabilité jointe d'observer le premier symbole de l'observation ( $o_1$ ) et de se trouver dans l'état  $s_i$ . On a donc  $\alpha_1(i) = \pi_i \times b_i(o_1)$  avec  $o_1 = v_{k_1}$ . Puis, une phase d'induction calcule  $\alpha_t(j)$  pour  $t = 2 \dots T$ . A l'instant  $t$ , ce calcul s'effectue comme le produit de la probabilité d'observer la suite partielle  $o = (o_1, o_2, \dots, o_{t-1})$  menant à l'état  $s_i$  et de la probabilité que cet état  $s_i$  génère l'observation de rang  $t$  ( $o_t$ ).

$$\alpha_t(j) = \left[ \sum_{i=1}^N \alpha_{t-1}(i) \times a_{ij} \right] \times b_j(o_t) \quad \begin{cases} 2 \leq t \leq T \\ 1 \leq j \leq N \\ o_t = v_k \end{cases}$$

La variable  $\alpha_t(j)$  est par définition la probabilité d'avoir toute l'observation en terminant dans l'état  $s_j$ . La sommation de cette variable sur l'ensemble des états finaux possibles donne  $P(O = o | \lambda)$ . On peut considérer le problème d'une façon analogue. On obtient alors l'algorithme Backward, où le terme générique est :  $\beta_t(j) = P(o_{t+1} \dots o_T | Q_t = s_j)$ , la probabilité de générer l'observation ( $o_{t+1}, \dots, o_T$ ) et d'être dans l'état  $s_j$  à l'instant  $t$ .

### 3.2 Second problème : l'apprentissage (estimation des paramètres non structuraux de la CMC)

L'apprentissage d'un modèle de Markov caché consiste à estimer son vecteur paramètre  $\lambda$  sur la base d'un ensemble de séquences d'observations appelé corpus d'apprentissage.

L'algorithme d'apprentissage le plus communément utilisé est l'algorithme de Baum-Welch [2], lequel est une instance de l'algorithme EM<sup>2</sup> appliquée aux CMCs. Il se fonde sur le critère du Maximum de Vraisemblance (MV). Etant données une CMC  $\lambda$  et une observation  $o$ , on cherche la CMC  $\lambda^*$  qui a la plus forte probabilité d'engendrer l'observation  $o$ .

C'est une procédure de ré-estimation itérative des matrices  $A$ ,  $B$  et  $\Pi$  d'une CMC. Etant donné un modèle  $\lambda = (A, B, \Pi)$  quelconque et une séquence d'observations  $o = (o_1, o_2, \dots, o_T)$  l'algorithme de Baum-Welch ré-estime les valeurs des matrices  $A$ ,  $B$  et  $\Pi$  de façon à maximiser la vraisemblance de l'observation  $o : P(O = o | \lambda)$ . On obtient après exécution de

l'algorithme, la CMC  $\lambda^*$  telle que :

$$\lambda^* = \operatorname{argmax}_{\lambda} P(O = o | \lambda) \quad [1]$$

<sup>2</sup> Expectation-Maximisation

Pour ré-estimer les valeurs des matrices  $A$ ,  $B$  et  $\Pi$ , on définit deux matrices :  $\Xi$  et  $\Gamma$ . On note :  $\Xi = [\dots \xi_t(i, j) \dots]$  où  $\xi_t(i, j)$  est la probabilité d'être dans l'état  $S_i$  à l'instant  $t$  et de passer dans l'état  $S_j$  à l'instant  $t + 1$ , connaissant le modèle  $\lambda$  et la séquence d'observations  $O$ , ce qui peut s'écrire :

$$\xi_t(i, j) = P(q_t = S_i, q_{t+1} = S_j | O = o, \lambda) \quad [2]$$

Cette matrice est de dimension  $T \times N \times N$ . En utilisant les variables Forward ( $\alpha$ ) et Backward ( $\beta$ ) précédemment vues, on établit la relation suivante :

$$\xi_t(i, j) = \frac{\alpha_t(i) \times a_{ij} \times b_j(o_{t+1}) \times \beta_{t+1}(j)}{P(O = o | \lambda)} \quad [3]$$

Ce qui peut s'interpréter par : c'est la probabilité d'observer  $(o_1 \dots o_t)$ , d'être dans l'état  $S_i$  à l'instant  $t$ , de transiter dans l'état  $S_j$  à l'instant  $t + 1$ , d'observer le symbole  $o_{t+1}$  étant dans l'état  $S_j$  et d'observer la séquence  $(o_{t+2} \dots o_T)$ . Le dénominateur  $P(O = o | \lambda)$  sert de « normalisateur ».

Rappel :

$$P(O | \lambda) = \sum_{i=1}^N \sum_{j=1}^N \alpha_t(i) \times a_{ij} \times b_j(o_{t+1}) \times \beta_{t+1}(j) \quad [4]$$

La seconde matrice ( $\Gamma$ ) est définie par :  $\Gamma = [\dots \gamma_t(i) \dots]$  où  $\gamma_t(i)$  est la probabilité d'être dans l'état  $S_i$  à l'instant  $t$  sachant le modèle  $\lambda$  et la séquence  $O$ . Cette matrice est de dimension  $T \times N$ . Si l'on reprend la définition de  $\xi_t(i, j)$ , on remarque que si l'on somme  $\xi_t(i, j)$  sur  $j$  (considérer tous les chemins d'états permettant d'observer  $o_1 \dots o_{t+1}$ , avec  $q_{t+1} = S_j$ ) pour tous les états d'arrivée, on retrouve  $\gamma_t(i)$ . D'où la relation :

$$\gamma_t(i) = \sum_{j=1}^N \xi_t(i, j) \quad [5]$$

Si l'on somme  $\gamma_t(i)$  suivant  $t$ , on obtient une quantité qui peut être interprétée comme le nombre de fois que l'état  $S_i$  est « visité », ou de manière équivalente, le nombre de transitions depuis l'état  $S_i$  (si l'on ne différencie pas l'instant  $t$  pour la sommation). De la même manière, la sommation de  $\xi_t(i, j)$  suivant le temps  $t$  (depuis  $t = 1$  à  $t = T-1$ ) peut être interprétée comme le nombre de transitions depuis l'état  $S_i$  à l'état  $S_j$ .

Ainsi :

$$\sum_{t=1}^{T-1} \gamma_t(i) = \text{nombre de transitions depuis } S_i \quad [6]$$

$$\sum_{t=1}^{T-1} \xi_t(i, j) = \text{nombre de transitions depuis } S_i \text{ vers } S_j \quad [7]$$

En utilisant les formules ci-dessus, on peut donner une méthode de ré-estimation des paramètres d'une CMC. Les formules de ré-estimations pour  $A, B, \Pi$  sont :

$\bar{a}_{ij} = \frac{\text{nombre de transitions depuis } S_i \text{ vers } S_j}{\text{nombre de transitions depuis } S_i}$  soit :

$$\bar{a}_{ij} = \frac{\sum_{t=1}^{T-1} \xi_t(i, j)}{\sum_{t=1}^{T-1} \gamma_t(i)} \quad [8]$$

$\bar{b}_i(k) = \frac{\text{nombre de fois où le symbole } v_k \text{ est observé pour l'état } S_i}{\text{nombre de fois où l'on a l'état } S_i}$

soit :

$$\bar{b}_i(k) = \frac{\sum_{t=1}^T \sum_{O_t = v_k} \gamma_t(i)}{\sum_{t=1}^T \gamma_t(i)} \quad [9]$$

$\bar{\pi}_i = \text{fréquence (nombre d'apparitions) de l'état } S_i \text{ à l'instant } t = 1$

### 3.3 A la recherche du meilleur chemin : procédure de Viterbi

La procédure de Viterbi (Andrew Viterbi) (segmental k-means algorithm en anglais [12]) maximise la probabilité des séquences d'observations de la base d'apprentissage uniquement le long des chemins optimaux obtenus par décodage de Viterbi [12]. La variable  $\delta(t, i)$  est définie comme la probabilité maximale que les observations jusqu'à l'instant  $t$  aient été émises par le modèle de Markov caché  $\lambda$  en suivant un chemin dont l'état à l'instant  $t$  est l'état d'indice  $i$  :

$$\delta(t, i) = \max_{q_1 \dots q_{t-1}} P(o_1 \dots o_t, q_1 \dots q_{t-1} = i | \lambda) \quad [10]$$

avec  $O = (o_1 \dots o_T)$  la suite d'observations et  $i$  le numéro de l'état atteint par le processus caché  $q$  à l'instant  $t$ . Une récurrence sur le temps, en parallèle pour tous les états, s'applique, à laquelle s'ajoute la mémorisation du meilleur chemin à travers la fonction  $\delta$ . L'algorithme suivant est donné pour un modèle « gauche-droite » où seul l'état 1 est considéré comme l'état initial et l'état  $N$  comme l'état final. On obtient :

-initialement (état d'indice 1) :

$$\delta(1, i) = \begin{cases} b_1(o_1), i = 1 & \dots \\ 0, 1 < i \leq N \end{cases} \quad [11]$$

-par récurrence sur les valeurs de  $\delta(t, i)$  :

-pour  $t$  variant de 2 à  $T$ ,

$$\delta(t, j) = \max_{1 \leq i \leq N} \delta(t-1, i) a_{ij} b_j(o_t) \quad [12]$$

$$\psi(t, j) = \arg \max_{1 \leq i \leq N} \delta(t-1, i) a_{ij} \quad [13]$$

-la probabilité du meilleur chemin est donnée par :

$$P(O, \tilde{Q} | \lambda) = \delta(T, N) \quad [14]$$

et le meilleur chemin est :

$$\tilde{Q} = (\tilde{q}_1 \dots \tilde{q}_T) \quad [15]$$

$$\tilde{q}_T = N$$

et :

$$\tilde{q}_{t-1} = \psi(t, \tilde{q}_t) \quad [16]$$

## 4. Etude de cas : modélisation du taux de dégradation d'une peseuse volumétrique dans le cadre d'activités de maintenance

### 4.1 Environnement de travail et activités de maintenance dans le processus de fabrication

Harry's est une société née en France, d'une double histoire : un boulanger de Châteauroux, qui a su importer et valoriser le savoir-faire américain en matière de fabrication de pain de mie, et un boulanger de Talmont, qui a développé une gamme de spécialités de viennoiseries inspirées de recettes de brioches vendéennes. De ce lignage, Harry's France a hérité des deux principes fondateurs de son activité : qualité et innovation. Au lendemain de la Libération, un certain Paul Picard, fils de boulanger à Châteauroux, fait la connaissance des Américains de la base de l'OTAN, ce qui incite le boulanger français à effectuer un premier voyage aux États-Unis. Quelques années plus tard, Paul Picard décide d'importer le savoir-faire des boulangers d'Outre-Atlantique pour fabriquer du pain de mie. Harry's détient une place unique sur le marché de la boulangerie de marque à courte durée de vie. A partir de sa position en France, le Groupe s'est développé en Europe et est aujourd'hui présent en Belgique, Espagne, Russie et en Turquie, à travers

des créations, des reprises et joints ventures<sup>3</sup>. L'association avec Barilla, qui a fait l'acquisition du Groupe Harry's en 2003, a permis de constituer le premier groupe européen de boulangerie pré-emballée.

L'usine située à Montierchaume (36000-France) fonctionne toute l'année sans interruption en équipes organisées en 3x8. Pour chaque équipe, rattachée à la production autour d'un processus dit « continu », il y a 20 agents de maintenance encadrés par un responsable global. A l'issue d'une intervention de maintenance (préventive ou non), chaque agent consigne le détail de ses actes sur une base de données centralisée et gérée par Access<sup>4</sup> (pour exemple, voir Tableau 1). Dans le cas de la peseuse volumétrique étudiée et repérée M2, il n'y a pas de variation du temps d'ouverture sur l'année (144 heures par semaine sur l'année) et la quantité de produit fabriqué est de 1,6T/H. La Figure 5 décrit l'organisation du processus continu.

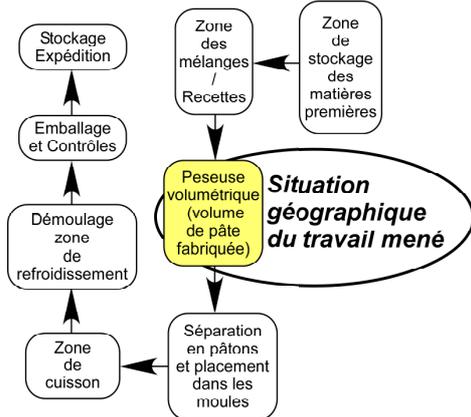


Figure 5 : Principe d'organisation du processus continu

L'arrêt d'un sous-système engendre l'arrêt total de production en amont de celui-ci. Une maintenance préventive est donc indispensable ! Le principe de fonctionnement de la peseuse volumétrique peut s'assimiler au fonctionnement d'une seringue que l'on remplirait par un effet de dépression.

A partir des données dont nous disposons, nous avons travaillé dans un premier temps à synthétiser le bilan des observations codifiées sur l'année 2005 entre le 3 janvier 2005 et le 8 novembre 2005 soit au total 475 relevés relatés chronologiquement dans le Tableau 1. La base de temps choisie étant la journée (cf Tableau 1), nous avons inséré le code Ras dans l'empilage des interventions lorsque le service de maintenance n'était pas sollicité (Figure 6).

RM	RM	SEC	Ras	Ras	Ras	SEC	TEP	...
2	2	6	10	10	10	6	5	...

↑ ↑ ↑ Insertions journalières

Figure 6 : Principe de codage de la séquence d'observations conformément au Tableau 2 avec ses insertions de codage (10) n'étant pas présent dans le tableau initial (Tableau 1)

Numero dans la liste	NOM	FOIHPF	DATE	LIGNE	NOM DE MACHINE	TRAVAUX EFFECTUES	ORS	DUREES INTERNEES A LA MAINTENANCE	TI (min)	PIECES CHANGEES	REPEREES
1	M Dupond	AM	03/01/2005	M2	Peseuse	regle cadence a la demande du RL		RM	20		
2	M Dupond	AM	03/01/2005	M2	Peseuse	regle cadence a la demande du RL		RM	20		
3	M Dupond	M	04/01/2005	M2	Peseuse	consigné et deconsigne amorce		SEC	20		
7	M Dupond	M	08/01/2005	M2	Peseuse	consigné et deconsigne		SEC	10		
8	M Dupond	VSD	08/01/2005	M2	Peseuse	pas de bruit a vide		TEP	30		
9	M Dupond	VSD	08/01/2005	M2	Peseuse	pas de bruit a vide		TEP	30		

Tableau 1 : Exemple de consignation des évènements sur la peseuse M2

Etat du processus	
MARCHE	
ARRET	
Codes symboliques : Observations	
Nature des interventions (Observations pour modèle)	
1	DEP (Dépannage/arrêt prod) sur la période
2	RM (Réglage Machine) sur la période
3	AU (Autre) sur la période
4	OBS (Observation) sur la période
5	TEP (Travaux Entretien Préventif pas d'arrêt prod) sur la période
6	SEC (Sécurité) sur la période
7	RAN (Remise A Niveau / planifié) sur la période
8	NET (Nettoyage Machine) sur la période
9	VEP (Visite Entretien Préventif) sur la période
10	Ras pas d'intervention

Tableau 2 : présentation du choix de codification symbolique (effectué à la main) des différentes occurrences d'observations

#### 4.2 Premier essai

Avant de décrire les résultats sur nos premiers essais montrant qu'ils peuvent apporter une aide à la décision dans l'organisation quotidienne de la maintenance, nous nous posons la question suivante :

*Combien de niveaux de vigilance significatifs et interprétables peut-on choisir ?*

Force est de constater que dans bon nombre de domaines, les usages et pratiques en utilisent quatre. Pour exemples :

- le renforcement de la politique de prévention des risques liés aux inondations (Ministère de l'écologie et du développement durable – 2006),
- plan national canicule, fortes chaleurs, sécurité des élèves et des personnels (Ministère de la santé – 2004),
- présentation du dispositif piloté par la préfecture de police, préfecture de la zone de défense de Paris, pour la mise en œuvre du plan vigipirate en Ile-de-France (rapport présenté à Jacques Chirac Président de la République le lundi 17 octobre 2005),
- ...

Nous sommes restés au constat des pratiques reconnues et efficaces sur quatre niveaux.

L'apprentissage de notre modèle (Figure 7) CMC consiste à estimer son vecteur paramètre  $\lambda$  sur la base d'un ensemble de séquences d'observations. Dans notre cas, la totalité des actions de maintenance enregistrées sur l'année 2005 (475 relevés relatés chronologiquement) respectant la codification symbolique du Tableau 2.

#### 4.2.1 Choix du modèle CMC

Tenant compte des éléments évoqués du chapitre 4.2, nous avons souhaité faire un premier essai sans contrainte sur un modèle à topologie à 4 états (Figure 7).

<sup>3</sup> Ce terme est habituellement associé aux fonctionnements des grandes affaires rattachés aux multinationales

<sup>4</sup> Microsoft Access ou MS Access (officiellement Microsoft Office Access) est un système de gestion de base de données relationnelles édité par Microsoft

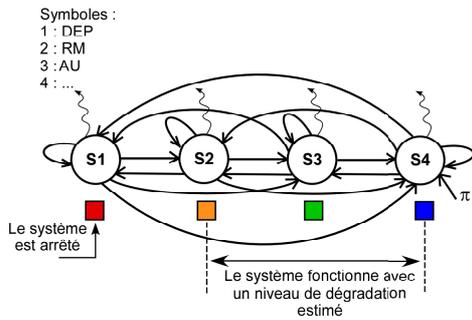


Figure 7 : Modèle ergodique à 4 états

#### 4.2.2 Estimation des paramètres non structuraux du modèle CMC à topologie ergodique

Comme le montre la Figure 8, l'algorithme Baum-Welch a besoin pour effectuer une ré-estimation d'un modèle initial  $\lambda = (A, B, \Pi)$  et d'une séquence d'observations choisie. Le modèle initial  $\lambda$  peut faire l'objet d'un tirage aléatoire des différentes probabilités ou d'une pré-disposition définie en fonction de critères choisis et judicieux. Dans notre cas, nous avons effectué un « calage préalable » des matrices A et B en tenant compte des points suivants :

- l'activité DEP découle réellement et uniquement d'un arrêt de la peseuse (voir matrice B, Figure 8),
- compte-tenu que l'état S1 est dépendant du symbole DEP, un prorata a été effectué tenant compte du volume d'activité DEP par rapport aux autres activités sur la totalité de l'année 2005 (voir matrice A, Figure 8).

**Matrices de départ pour ré-estimations du modèle CMC**

A	1	2	3	4
1	0,2516	0,2495	0,2495	0,2495
2	0,2500	0,2500	0,2500	0,2500
3	0,2500	0,2500	0,2500	0,2500
4	0,2500	0,2500	0,2500	0,2500

B	1	2	3	4	5	6	7	8	9	10
1	1,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0000	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111
3	0,0000	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111
4	0,0000	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111

$\Pi$  0,0000 0,0000 0,0000 1,0000

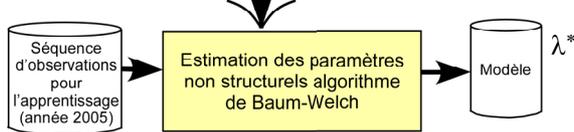


Figure 8 : Ré-estimation du modèle à topologie ergodique

#### 4.2.3 A la recherche du meilleur chemin

Le résultat de l'estimation du « meilleur chemin (Figure 9) » et donc l'arrêt dans un état donné auraient pu se traduire concrètement par l'affichage d'un pictogramme comme l'utilise Météo-France dans ses prévisions. Nous avons choisi une autre stratégie dans la modalité d'affichage du résultat : « un pseudo-VU-mètre progressif (Figure 10) ».

Ce « pseudo-VU-mètre progressif » permet au responsable de la maintenance de disposer visuellement d'un outil capable d'afficher des niveaux d'estimations relatifs au taux de dégradation du système en fonction des activités concrètes en maintenance sur une journée.

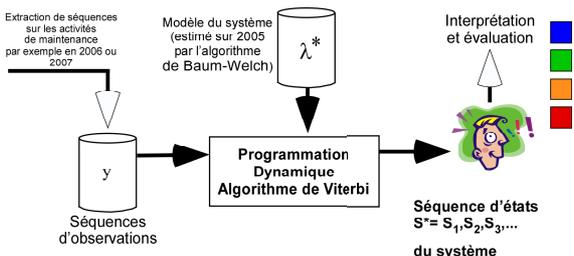


Figure 9 : Estimation du meilleur chemin

#### 4.2.4 Résultats sur le modèle à topologie ergodique

La Figure 10 montre un exemple d'estimations en fonction de l'empilage des interventions de maintenance (empilage des symboles relatifs au Tableau 2). Excepté que notre modèle estime bien l'état « ARRET » en fonction de l'activité de maintenance rattachée, « notre pseudo-VU-mètre » a tendance à fonctionner sur une loi presque tout-ou-rien ! Ce 1<sup>er</sup> résultat est mauvais et n'apporte aucune aide possible. Dans la pratique, nous conserverons le raisonnement suivant : l'estimation commence à la suite d'un état « ARRET » et s'arrêtera au prochain état « ARRET » et ainsi de suite.

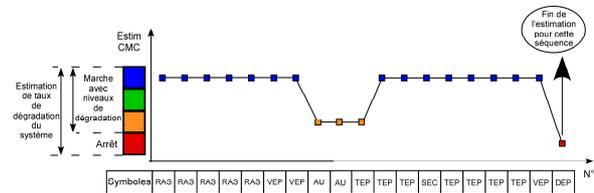


Figure 10 : Résultat des estimations obtenues à partir du modèle à topologie ergodique sur une séquence complète en 2006

#### 4.3 Deuxième essai

Compte-tenu, des premiers résultats ci-dessus, nous avons choisi d'orienter le modèle CMC en choisissant une topologie forcée (Figure 11). **Ce 2<sup>ème</sup> choix, doit nous permettre de disposer d'un outil d'estimations progressif.** En effet, nous voyons (Figure 11) que le passage de l'état S4 à l'état S1 ne pourra être que progressif.

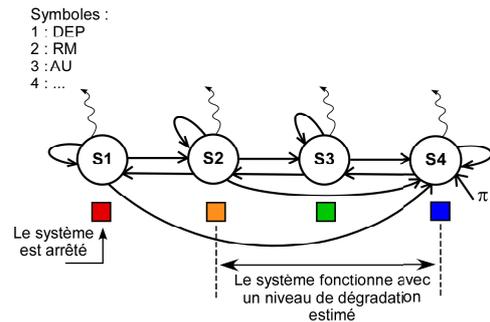


Figure 11 : Modèle à topologie forcée à 4 états

#### 4.3.1 Estimation des paramètres non structuraux du modèle CMC forcé

La procédure reste identique qu'au paragraphe 4.2.2. Les matrices A et B sont re-calibrées en fonction (Figure 12).

**Matrices de départ pour ré estimations du modèle CMC**

A	1	2	3	4
1	0,3347	0,3326	0,0000	0,3326
2	0,2500	0,2500	0,2500	0,2500
3	0,0000	0,3333	0,3333	0,3333
4	0,0000	0,0000	0,5000	0,5000

B	1	2	3	4	5	6	7	8	9	10
1	1,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000	0,0000
2	0,0000	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111
3	0,0000	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111
4	0,0000	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111	0,1111

$\Pi$  0,0000 0,0000 0,0000 1,0000

Figure 12 : Calibrage des matrices A et B en fonction du modèle CMC à topologie forcée

#### 4.3.2 Résultats sur le modèle à topologie forcée

Les premiers résultats pour ce choix topologique montrent que le modèle propose une estimation progressive du taux de dégradation du système (Figure 13, Figure 14). Pour exemples, deux séquences estimées sur 2006 et 2007 (Figure 13, Figure 14) montrent que l'arrêt est bien détecté. La Figure 16 propose une comparaison entre les quantités réelles et estimées par le modèle à topologie forcée (4 états) sur les arrêts du système lors des campagnes de maintenance du 1<sup>er</sup> trimestre 2006 et 2007.

La Figure 15 montre un exemple de résultat d'estimations obtenues à partir du modèle à topologie forcée sur une séquence en cours en 2007. L'estimation du processus se dégradant (passage de l'état S4 à l'état S2), les interventions de maintenance stabilisent dans un premier temps la dégradation avant de revenir dans l'état S4 (situation la plus favorable). Nous constatons Figure 13 qu'une activité de travaux d'entretien préventif permet de faire remonter la qualité du système. Ne rien faire sur le système peut engendrer une dégradation de celui-ci (Figure 14) ! Une activité opérationnelle de maintenance comme des travaux d'entretien préventifs cumulés permettent de faire remonter le niveau de dégradation du système. De simples activités de visites d'entretien préventifs montrent que le système continue de se dégrader (Figure 14).

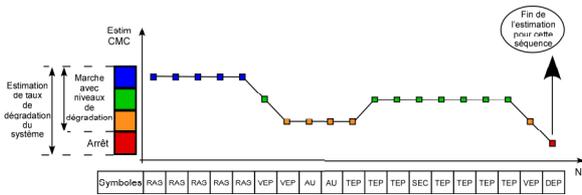


Figure 13 : Résultat des estimations obtenues à partir du modèle à topologie forcée sur la même séquence complète en 2006

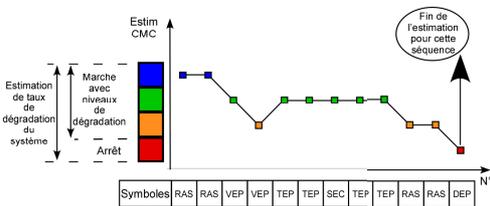


Figure 14 : Résultat des estimations obtenues à partir du modèle à topologie forcée sur une séquence complète en 2007

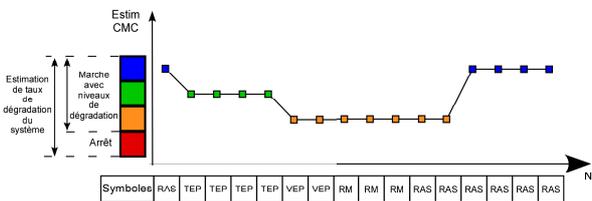


Figure 15 : Résultat des estimations obtenues à partir du modèle à topologie forcée sur une séquence en cours en 2007

Très souvent, effectuer des réglages sur un système, entraîne ce système vers une situation opérationnelle « périlleuse ». Il faudra attendre, suite aux réglages effectués, que ces activités aient redonné un bon niveau dans l'estimation du taux de dégradation (Figure 15).

La Figure 16 confirme nos choix concernant la quantité des niveaux de vigilance significatifs et interprétables. L'essai présenté montre qu'un modèle CMC à 5 états à topologie forcée n'apporte pas d'éléments supplémentaires par rapport au modèle CMC à topologie forcée.

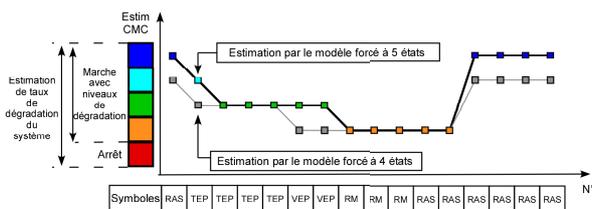


Figure 16 : Comparaison de résultats issus d'une estimation : topologie forcée à 4 états, topologie forcée à 5 états sur une séquence en cours en 2007

## 5. Conclusion

Au-delà des commentaires que nous n'apporterons pas dans cet article sur le niveau de qualité du service de maintenance, nous montrons qu'il est possible d'estimer un niveau de dégradation sur notre système à partir de séquences d'observations (activité de maintenance) et conformément au postulat fixé initialement. Cependant, il nous reste à optimiser le modèle selon les pistes suivantes (perspectives) :

- il faut sans doute rendre l'état de panne absorbant.
- notre approche peut-elle apporter une prédiction de panne ? Par exemple, existe-t-il une corrélation entre le nombre d'états n°2 estimés et le temps qui passe avant d'arriver à l'arrêt ?

Dernier point relatif à la Figure 17 : L'estimation montre que le système est fort dégradé alors qu'un entretien préventif est prévu très prochainement.

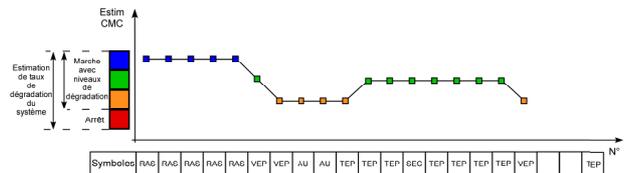


Figure 17 : Estimation en cours

Le responsable de la maintenance doit être amené à se poser deux questions :

- doit-on avancer l'action TEP ? Si oui, il rentre dans une démarche d'optimisation de la maintenance en « reprogrammant des priorités d'activités ».
- doit-on prendre le risque de continuer dans l'état ?

Enfin, dans une approche plus globale, le responsable de la maintenance pourrait orienter ses priorités au sein de son système de production, s'il disposait de plusieurs VU-mètres rattachés à des sous-systèmes ou systèmes complets.

## 6. Références

- [1] M. Avila, *Optimisation de modèles Markoviens pour la reconnaissance de l'écrit*, Thèse de doctorat, Université, Rouen, 1996.
- [2] L. E. Baum, *An inequality and associated maximisation technique in statistical estimation for probabilistic functions*, *Inequalities*, 3 (1972), pp. 1 - 8.
- [3] L. E. Baum and T. Petrie, *Statistical inference for probabilistic functions of Markov chains*, *Annals Math. Stat.*, 37, 1966, pp. 1554 -1563.
- [4] T. Brouard, *Hybridation de Chaînes de Markov Cachées : conception d'algorithmes d'apprentissage et applications*, Thèse de doctorat, Université François Rabelais, Tours, 1999.
- [5] D. C. Dennett, *la stratégie de l'interprète*, essais Gallimard, 1990, pp. 29.
- [6] M. M. Doss, *Using Auxiliary Sources of Knowledge for Automatic Speech Recognition*, Thèse de doctorat n°3263, Ecole Polytechnique Fédérale, Lausanne, 2005.
- [7] W. N. Grundy, T. L. Bailey and M. E. Baker, *Meta-MEME : Motif-based Hidden Markov Models of protein families*, *Computer applications in the biosciences*, 13(4) (1997), pp. 397.
- [8] J. P. Hugues and P. Guttorp, *A hidden Markov model for downscaling synoptic atmospheric patterns to precipitation amounts*, *Climate research*, 15(1), p. 1, (2000).
- [9] L. Jeanpierre, *Apprentissage et adaptation pour la modélisation stochastique de systèmes dynamiques réels*, Thèse de doctorat, Université Henri Poincaré, Nancy-I, 2002.
- [10] C. Meier-Hirmer, F. Sourget and M. Roussignol, *Estimation de la durée de vie des pièces caténaies*, in

- R. s. JESA, ed., *Journal européen des systèmes automatisés*, Vol.40, n°7, 2006.
- [11] L. R. Rabiner, *A tutorial on hidden Markov models and selected applications in speech recognition*, Proceeding of the IEEE, 77(2) (1989), pp. 257 - 286.
- [12] L. R. Rabiner, B. H. Juang, S. E. Levinson, & and M. M. Sondhi, *Recognition of isolated digits using hidden Markov models with continuous mixture densities*, AT&T Technical Journal, 64 (1986), pp. 1211 - 1222.
- [13] A. Schalapbach and H. Bunke, *Off-line handwriting identification using HMM-based recognizers Proc. 17th Int. Conference on Pattern Recognition*, vol. II, 2004, pp. 654 - 658.
- [14] A. Schalapbach and H. Bunke, *Using HMM-based recognizers for writer identification and verification Proc. 9th Int. Workshop on frontiers in Handwriting Recognition (2004)*, pp. 167 - 172.
- [15] S. Schbath, *Les chaînes de Markov cachées : présentation et usage en analyse de séquences biologiques*, Unité Mathématique, Informatique & Génome, INRA, 2007.
- [16] F. B. Vialatte, *Aide au diagnostic d'anomalies cardiaques mémoire de DEA de Sciences Cognitives*, Paris VI, Paris, 2002.
- [17] A. J. Viterbi, *Error bounds for conventional codes and asymptotically optimum decoding algorithm*, IEEE Trans. on Information Theory, 13 (1967), pp. 260 - 269.