On Evidential Markov Chains

Hélène Soubaras
Thales R&T - RD128 - F91767 Palaiseau
helene.soubaras@thalesgroup.com

Abstract

Evidential Markov chains (EMCs) are a generalization of classical Markov chains to the Dempster-Shafer theory, replacing the involved states by sets of states. They have been proposed recently in the particular field of an image segmentation application, as hidden models. With the aim to propose them as a more general tool, this paper explores new theoretical aspects about the conditioning of belief functions and the comparison to classical Markov chains and HMMs will be discussed. New computation tools based on matrices are proposed. The potential application domains seem promising in the information-based decision-support systems and an example is given.
Keywords: Markov chains, belief functions, Dempster-Shafer theory, Hidden Markov Models, evidential networks

1 Introduction

Markov chains [4] are well-known statistical models for memoryless systems. They are applied to a wide range of application domains, and they are a mathematically powerful tool [16] [8].

But the parameters they involve are precise probabilities, which will not be available in a family of decision-making problems where the data are imprecise or incomplete, or in systems whose behavior can be described only roughly. This is why the generalization of Markov chains to belief functions has recently been proposed in works around W. Pieczynsky [3] [7]. This new model, called *evidential Markov chain* (EMC), was used as hidden model in a particular application of image segmentation. These works proposed an algorithm to solve the hidden model based on HMM approaches, and examined the computational complexity.

The objective of this paper is to explore some theoretical aspects about EMCs, and to show their relevance to a wide panel of possible applications. Basics of the Dempster-Shafer theory [11] [15] will first be reminded, then the Markov chains and the EMC will be defined. Aspects about conditioning will be discussed [13], and some possible applications will be proposed.

2 Basics of the Dempster-Shafer theory

This section will remind the basics of the Dempster-Shafer theory and provide tools to understand them (probabilistic point of view and matrix) that will be useful in the sequel.

2.1 Basic belief assignment

One calls *frame of discernment* a set Ω of all possible hypotheses; Ω can be discrete or continuous.

A *mass function*, also called BBA (Basic Belief Assignment) [11], is a mapping *m* on the power set 2^Ω, which is the set of all subsets of Ω, to $[0; 1]$ such that

$$\sum_{A \subseteq \Omega} m(A) = 1$$

A subset $A \subset \Omega$ is called a *focal set* as soon as its mass is nonzero.

If $m(\oslash) \neq 0$, some belief on an hypothesis that would be outside Ω. This is the *Open World Assumption (OWA)* [6]. Otherwise, the mass function is said normalized. m becomes a classical probability when the focal sets are disjoint atoms. $\mathcal{F} \subseteq 2^{\Omega}$ will denote the set of all focal sets.

2.2 Induced probability space

In this paragraph shows that the belief functions can be manipulated through a probability μ, as did Shafer [12].
The set $2^{\mathcal{F}}$ is then the set of all collections of focal sets. Note that $2^{\mathcal{F}} \subseteq 2^{2^{\Omega}}$, which is the set of all collections of subsets of Ω. The elements of $2^{\mathcal{F}}$ are then of the form:

$$A \in 2^{\mathcal{F}} \iff A = \{B_1, B_2 ... B_n\}$$

with $B_i \in \mathcal{F}\, \forall i$. So, $2^{\mathcal{F}}$ is a σ-algebra on $\mathcal{F}$. Let's define on $2^{\mathcal{F}}$ the following function:

$$\mu : 2^{\mathcal{F}} \to [0; 1]$$

such that $\forall A \in 2^{\mathcal{F}}$,

$$\mu\left(A = \{B_1, B_2 ... B_n\}\right) = \sum_{i=1}^{n} m(B_i) \qquad (1)$$

It is easy to see that μ is a measure, since $\mu(\oslash) = 0$ and μ is additive, i.e. $\mu(\bigcup_i A_i) = \sum_i \mu(A_i)$ as soon as the A_i are pairwise disjoint. Furthermore, $\mu(\mathcal{F}) = 1$. Thus, $(\mathcal{F}, 2^{\mathcal{F}}, \mu)$ is a probability space. In other words, focal sets can be seen as set-valued random variables. The probability $\mu(A = \{B_1, B_2 ... B_n\})$ corresponds to the fact that one of the focal sets B_i, $1 \leq i \leq n$ occurs (thus the truth is in one of these sets). At this stage one doesn't take into account the fact that the B_i are disjoint or not.

One can define two functions that provide collections in $2^{\mathcal{F}}$ for any given $A \subseteq \Omega$ (even if A is not in $\mathcal{F}$):

$$\overline{\mathcal{F}}(A) = \{B \in \mathcal{F} / B \cap A \neq \oslash\}$$

(these are the elements of $\mathcal{F}$ *hitting* the given subset A), and the dual collection, which is:

$$\underline{\mathcal{F}}(A) = \{B \in \mathcal{F} / B \subseteq A\}$$

Note that $\overline{\mathcal{F}}(A) = \underline{\mathcal{F}}^c(A^c)$.

$\underline{\mathcal{F}}(A)$ and $\overline{\mathcal{F}}(A)$ are called respectively the *inner* and the *outer restriction* of $\mathcal{F}$ with respect to A [18].

2.3 Belief function, plausibility and commonality

For a given mass function m a *belief* function Bel, a *plausibility* function Pl and a *commonality* function q have been defined as follows [11] for all $A \subseteq \Omega$:

$$Bel(A) = \sum_{B \subseteq A, B \neq \oslash} m(B) \qquad (2)$$

$$Pl(A) = \sum_{B \cap A \neq \oslash} m(B) \qquad (3)$$

$$q(A) = \sum_{B \supseteq A} m(B) \qquad (4)$$

They can also be written as: $Bel(A) = \mu(S \subseteq A, S \neq \oslash)$, $Pl(A) = \mu(S \cap A \neq \oslash)$, and $q(A) = \mu(A \subseteq S)$.

One can remark that Bel and Pl can be written using the inner and the outer extension of A in $\mathcal{F}* = \mathcal{F} \setminus \{\oslash\}$, which denotes the set of nonempty focal sets of the frame Ω. The belief function can be expressed as:

$$Bel(A) = \mu\left(\underline{\mathcal{F}*}(A)\right)$$

and the plausibility as:

$$Pl(A) = \mu(\overline{\mathcal{F}*}(A))$$

The two functions Bel and Pl are dual, related by:

$$Pl(A) = 1 - m(\oslash) - Bel(A^c) \qquad (5)$$

where A^c denotes the complementary set of A in Ω.

Smets [15] introduced the *pignistic probability* Bet associated to m. When Ω is discrete, it is defined $\forall x \in \Omega$ by:

$$Bet(x) = \frac{1}{1 - m(\oslash)} \sum_{A/x \in A} \frac{m(A)}{|A|} \qquad (6)$$

where $|A|$ is the cardinality of A, i.e. the number of elements of A. Bel, Pl and Bet are all equal to classical probabilities when the focal sets are disjoint atoms.

It is important to notice that the three functions (belief, plausibility and commonality) are not measures because they are not additive, but sub-additive since:

$$Bel(A \cup B) \geq Bel(A) + Bel(B)$$

3 Matrix tools

We consider a BBA on a finite discrete frame Ω. N_f is the number of focal sets. One will define the mass vector M by its coordinates:

$$M(j) = m(A_j) = m_j \qquad (7)$$

for all focal set A_j, $1 \leq j \leq N_f$.

3.1 Matrix tools for belief functions

It is known [5] that the relation between the BBA m and the belief function Bel is a bijection. For a given discrete space Ω containing N elements and any function Bel defined on a set of subsets $\mathcal{F} \subseteq 2^\Omega$, if Bel satisfies the two assumptions:

(i). $Bel(\Omega) \leq 1$

(ii). Bel is completely monotone,, i.e. if $A \subset B$ then $Bel(A) \leq Bel(B)$

(iii). Bel is subadditive

a mass function m can be deduced thanks to the so-called *Möbius transform*:

$$m(A) = \sum_{B \subseteq A} (-1)^{|A \setminus B|} Bel(B) \qquad (8)$$

If one denotes as M the column vector of the masses of all the subsets of Ω, its size will be 2^N. The column vector Bel containing all the values of the belief function on the nonempty subsets will also be of size 2^N, and it can be calculated from M thanks to a matrix product:

$$Bel = BfrM.M$$

and the Möbius transform is then performed by the inverse matrix $BfrM^{-1}$. $BfrM$ is a generalization matrix G in the particular case where its nonzero elements are equal to 1. G is the $2^N \times 2^N$ matrix defined by Smets [14] as:

Definition 3.1. A *generalization matrix* of a collection of subsets A_i is a stochastic matrix G satisfying $G(i,j) = 0$ if $A_j \not\subseteq A_i$.

Smets [14] also defined, similarly:

Definition 3.2. A *specialization matrix* is a stochastic matrix S satisfying $S(i,j) = 0$ if $A_i \not\subseteq A_j$.

In this paper we propose a new matrix in order to compute the plausibility function. It will be called the gauge matrix:

Definition 3.3. The *gauge* (pattern) matrix of a collection of subets A_i is defined by:

$$G_a(i,j) = \begin{cases} 1 & \text{if} \qquad A_i \cap A_j \neq \oslash \\ 0 & \text{otherwise} \end{cases}$$

The 2^N-size column vector Pl of the plausibility function is then defined by:

$$Pl = G_a M$$

The commonality can also be computed through such a matrix product.

3.2 Markov kernel matrix

Let Ω be a frame of discernment. One suppose there is a finite partition $\mathcal{H} = \{X_i \ / \ 1 \leq i \leq N_c\} \subseteq 2^\Omega$ on the frame

Ω. The couple $(\Omega, \mathcal{H})$ is called a *propositional space*. Each subset X_i can be called a *class*. Let m be a BBA on Ω, with a finite set of focal sets $\mathcal{F} = \{A_k / 1 \leq k \leq N_f\}$. One would like to estimate in which class X is the truth for a given BBA.
Classes and focal sets can be viewed as random variables X and A., taking values in $\mathcal{H}$ and $\mathcal{F}$ respectively. Each focal set A_k can occur with a probability m_k.
The assumption that will be made now is that there exists a fixed *Markov kernel* K, which is the matrix of conditional probabilities of the occurence of one of the two random variables given the other. (As it will be developed below, in the particular case of Markov chains, the Markov kernel is the state transition matrix). The kernel K is defined by $K(i,k) = p(i|k)$ such that

$$p_i = \sum_{k=1}^{N_f} p(i|k) m_k \tag{9}$$

where $p_i = Pr(X_i)$ and $m_k = m(A_k)$. This can also be written with P, the vector of the probabilities of the classes:

$$P = KM \tag{10}$$

As $0 \leq p(i|k) \leq 1$ for all (i,k), one can notice from (2) and (3) that for all compatible kernel K,we get the following relation, for all set $X = X_i$:

$$Bel(X) \leq p(X) \leq Pl(X) \tag{11}$$

Thus the probabilities $p_i(X_i)$ of each class X_i are *imprecise probabilities* since they belong to an interval.

3.3 Matrix representations for classes

Let N_f be the number of focal sets and N_c the number of classes. One supposes that N_f and N_c are finite. One can still define the gauge matrix G_a of size $N_c \times N_f$ by $G_a(i,j) = 1$ if $X_i \cap A_j \neq \oslash$, and 0 otherwise, for all classes X_i and for all focal sets A_j. Any Markov kernel K compatible with the BBA is zero where G_a is zero. The lines of the transposed matrix G_a^T can be seen as base-2 representations of the focal sets. One can describe entirely a belief mass by its gauge matrix G_a and its mass vector M.
When the classes are not atoms, the cardinality of a focal set can be defined as the number of classes it meets. This number is obtained by

$$(11...1)\, G_a = \begin{pmatrix} |A_1| \\ |A_2| \\ \vdots \\ |A_{N_f}| \end{pmatrix}$$

The computation of the belief function, the plausibility function, the commonality and the pignistic probability with matrix products is still possible, as it was shown by Smets [14] and at paragraph 3, for the 2^{N_c} subsets of Ω that are unions of subsets X_i:

$$Bel = G.M \tag{12}$$

$$Pl = G_a.M \tag{13}$$

$$q = S.M \tag{14}$$

$$Bet = PetPfrM.M \tag{15}$$

where $BfrM$, $QfrM$ and $BetPfrM$ where defined by Smets [14] ($BfrM$ and $QfrM$ are respectively a generalization and a specialisation matrix whose non null elements are 1); G_a is the gauge matrix (3.3). They are all $2^{N_c} \times N_f$-sized matrices.

4 Evidential Markov chains

4.1 Definition

Let $\Omega = \{a_1, a_2 ... a_N\}$ be the set of the possible random *states* x_t of a system for each time t.

Definition 4.1. The probability Pr for each state of the system satisfies the *Markov property* if and only if:

$$Pr(x_t | x_0, x_1, x_2, x_3, ... x_{t-1}) = Pr(x_t | x_{t-1})$$

Definition 4.2. A *transition matrix* for a system is the $N \times N$ matrix defined by:

$$Q = (q_{ij})_{1 \leq i,j \leq N}$$

where

$$q_{ij} = Pr(x_{t+1} = a_i | x_t = a_j)$$

If a transition matrix exists, the Markov property is satisfied. If one denotes as P_t the vector of the probabilities of each state:

$$P_t = Pr(x_t) = \begin{pmatrix} Pr(x_t = a_1) \\ Pr(x_t = a_2) \\ \vdots \\ Pr(x_t = a_N) \end{pmatrix}$$

one has the following relation:

$$P_t = QP_{t-1} = Q^t P_0$$

Definition 4.3. A *Markov chain* is a triple (Ω, Q, P_0) where $P_0 = Pr(x_0)$ is the *initial probability vector.*

An evidential Markov chain is a classical Markov chain where the random variable representing the possible states of the system is replaced by random (focal) sets [3] [9]:

Definition 4.4. Let Ω be a frame of discernment. An *evidential Markov chain* (EMC) is a Markov chain $(\mathcal{F}, Q, M_{o)})$ where $\mathcal{F}$ is a set of focal sets and M_0 is the vector of the initial masses of all the focal sets.

If the vector of masses at time *t* is denoted as M_t, one can write the following relation:

$$M_t = QM_{t-1} \qquad (16)$$

In the particular case where $\mathcal{F}$ is the set of the N atoms of Ω, the EMC becomes a classical probabilistic Markov chain.
One can verify through equations 12, 13, 14 and 15 that the belief function, the plausibility, the commonality and the pignistic probability of an EMC are Markov chains whose transition matrices are of the form

$$Q' = HQ(H^T H)^{-1} H^T$$

where the matrix H represents $BfrM$, G_a, $QfrM$ and $BetPfrM$ respectively. Of course, if the topology of the focal sets does not satisfy some conditions, the matrix $H^T H$ will not be invertible.
The fact that theses functions are Markov chains is still true when the matrices H are restricted to a subcollection of subsets, for example to atoms. If H is square, it may be invertible, the resulting transition matrix will be:

$$Q' = HQH^{-1}$$

4.2 Conditioning and the Generalized Bayes Theorem

Evidential Markov chains are particular cases of evidential networks [19] since they rely on conditional masses. There have been some works about conditioning in the Dempster-Shafer theory [1] [17].
In probability theory the conditional probability of a subset A given a subset B is the new probability defined on B as probability space, and it is given by the Bayes formula:

$$Pr(A|B) = \frac{Pr(A \cap B)}{Pr(B)} \qquad (17)$$

Such relations of conditioning have been proposed for belief masses [11]. Thanks to the probability μ introduced at 2.2, ere are four ways to express the conditional certainty that the truth is *always* ($\mu(S \subseteq A) = Bel(A)$) / *possibly* ($\mu(S \cap A \neq \oslash) = Pl(A)$) in A given that it is *always / possibly* in B: $\frac{Bel(A \cap B)}{Bel(B)}$, $\frac{Bel(B) - Bel(B \setminus A)}{Bel(B)}$; $\frac{Pl(B) - Pl(B \setminus A)}{Pl(B)}$ (Dempster's $Bel(A|B)$ [15]) and $\frac{Pl(A \cap B)}{Pl(B)}$ (Dempster's $Pl(A|B)$ [15]).
Dempster also defined the underlying conditional mass [15], which is equal to, in the unnormalized case:

$$m(A \wr B) = \begin{cases} \mu(S \cap B = A) & \text{for } A \subseteq B \\ 0 & \text{otherwise} \end{cases}$$

and, in its normalized version: $m(A|B) = m(A \wr B)/\mu(B^c)$. Smets [13] proposes expressions for the generalization of the Bayes theorem to belief functions (GBT). In this article, the expression proposed for the masses themselves is the Conjunctive Rule of Combination [2], which is exactly the one expressed using $m(A \wr B)$:

$$m_t(A) = \sum_B m(A \wr B) m_{t-1}(B)$$

This is exactly the operation performed in the EMC transition matrix product 16.

Nevertheless, Smets' expressions for the GBT for the belief and the plausibility functions suppose that one of the two frames of discernment is a partition. To be applied to a EMC, the GBT implies that F is a partition of Ω. This is a particular case which is not very interesting since it corresponds to a classical probabilistic case where the belief and the plausibility are all equal to a simple probability.

In conclusion, through its transition matrix, the EMC performs one form of conjunctive rule of combination of Demspter's unnormalized conditional masses.

4.3 Associated Hidden Markov Model

Definition 4.5. A *Hidden Markov Model* (HMM) is a 5-uple $(\Omega_x, \Omega_y, Q, K, P_0)$ where (Ω_x, Q, P_0) is a Markov chain, and the observation is a random variable y taking values in Ω_y and such that the Markov kernel of y given x is K.

The internal states x of the Markov chain are not known, except through the knowledge of y. To estimate x_t from an observed sequence y_t (when $dim(y) < dim(x)$), algorithms have been proposed such as the Baum-Welch algorithm and the Viterbi algorithm [10].

To compare a HMM and a EMC (see figures 1 and 2) let's consider a given HMM with transition matrix Q. If there exists a compatible EMC with transition matrix Q', the following condition must be satisfied:

$$(KQ - Q'K)P = 0 \text{ for all probability vector} P$$

If $dim(M) \leq dim(P)$, one solution is $Q' = K.Q.K^T.(K.K^T)^{-1}$ (if K is not degenerated). Thus it is possible to find one EMC which is compatible with a given HMM, but it is not so easy to find one HMM compatible with a given EMC since $K^T K$ is not invertible.

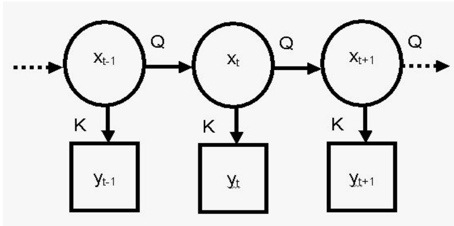

Figure 1: Scheme of a HMM

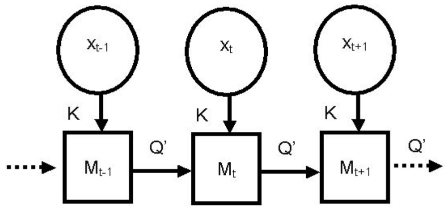

Figure 2: Scheme of a EMC

Now suppose we have a EMC, and let's consider again the partition of Ω into classes $\mathcal{H} = \{X_i / i \leq i \leq N_c\} \subseteq 2^\Omega$; one assumes that the conditional belief functions $Bel(A_i|X_k)$ are known, and that the BBA on $\mathcal{H}$ is normalized. The GBT can then be applied [13]:

$$\alpha Bel(X_i|A_j) = \prod_{k \neq i} Bel(A_j^c|X_k) - \prod_k Bel(A_j^c|X_k)$$

where α is the normalizing factor:

$$\alpha = 1 - \prod_k Bel(A_j^c|X_k)$$

Thus, thanks to the Möbius transform, a BBA can be calculated for the classes. If this BBA were a classical probability, the classes X_i could be seen as the hidden internal state of the Markov chain whose observations are the (random) focal sets. Thus, an EMC can be viewed as a generalization of a classical HMM.

In conclusion, we showed that it can be easy to find one EMC compatible with a given HMM. For a given EMC, a family of compatible HMMs exists. When the EMC observations are random focal sets, the EMC internal state can be solved as in a HMM.

5 Applications

The EMC model have been first proposed to achieve image segmentation [7] [3]. It was supposed to be hidden in those cases; this means the BBA could not be observed directly, but through a measurement y such that the conditional probabilities $Pr(y|A)$ are known to be fixed for each focal set A. An algorithm derived from the classical HMM identification was proposed [7].
EMCs can be also interesting models for other uncertain systems, particularly if they involve phenomena that are difficult to quantify, like human behaviors. Such modelling can be applied to the forecasting of the future evolution of a system; it can also be useful for simulations in order to measure the performances of other algorithms. Techniques used in the classical statistics, such as Monte-Carlo or Importance Sampling, could be generalized to EMCs.
As an example, an EMC can be used as a simulation model for the tenseness between two countries that could lead to a conflict or a war. The tenseness is an underlying value that can be estimated only through open sources of information (journal articles, television news...) and indirectly (e.g. through symptomatic events such as demonstrations, declarations, political decisions...). A decision-support system should be able to extract a BBA from these events; to validate such systems on can use an EMC whose focal sets are overlapping rough estimations of the tenseness, and the tenseness itself is quantized on several values (4 in the example shown figure 3. These intervals of tenseness are the classes X_i.

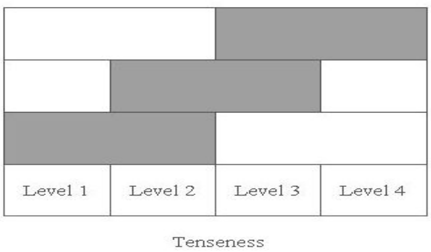

Figure 3: Example for an EMC with three focal sets ”low”, ”medium” and ”high” for 4 levels of tenseness between two countries

The 3×3 transition matrix of the EMC could reflect the fact that when the tenseness begins to rise, it can easily rise more. For example, it can be:

$$Q = \begin{pmatrix} 0.9 & 0.1 & 0 \\ 0.1 & 0.6 & 0.2 \\ 0 & 0.3 & 0.8 \end{pmatrix}$$

This matrix shows for example that the tenseness can increase easily from medium to high, but resolution of the crisis from high to medium is less likely to happen. The EMC allows to translate such approximately described phenomena into a model that can be implemented.
An example obtained by running such an EMC is shown at figure 4. There, a mass function is generated by the EMC model at each time. Previously, an expert had assigned a mass function to each one of the predefined possible classes of events. The simulator computes then the Euclidian distance between the mass vector generated by the EMC and the mass vectors of each class, and it chooses randomly one class amongst the nearest ones.

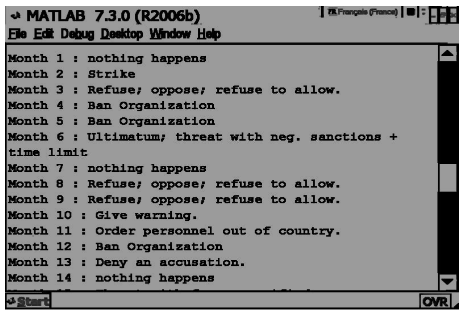

Figure 4: Random sequence of classes of geopolitical events following an evidential Markov-modelled increase of tenseness

6 Conclusion

Evidential Markov chains (EMCs) are a generalization of the classical Markov chains. They are Markov chains involving masses on focal sets instead of probabilities on elementary states. They have been proposed only in the image segmentation model [3] [7]. This paper examines some theoretical aspects of EMCs: it relates them to the Dempster’s rules of conditioning and the Smets’ Generalized Bayes Theorem; it points out that an EMC is a generalization of a HMM. Some computation tools based on matrix are also proposed.

EMC models have potentially interesting applications in the field of uncertain systems, particularly those involving human behaviors or imprecise data such as text. An example is given for the simulation of the tenseness between two conflicting countries.

References

[1] Didier Dubois, Serafín Moral, and Henri Prade. Belief change rules in ordinal and numerical uncertainty theories, 311-392.

[2] Didier Dubois and Henri Prade. A set theoretical view of belief functions. *Int. J. Gen. Systems*, 12:193–226, 1986.

[3] L. Fouque, A. Appriou, and W. Pieczynski. An evidential markovian model for data fusion and unsupervised image classification. In *Proceedings of 3rd International Confernce on Information Fusion, FUSION 2000*, pages YuB4–25 – TuB4–31, Paris, France, July 10-13th 2000.

[4] David Freedman. *Markov chains*. Holden-Day, 1971.

[5] Michel Grabisch, Toshiaki Murofushi, and Michio Sugeno. *Fuzzy Measures and Integrals*. Physica-Verlag, 2000.

[6] Ralf Haenni. Ignoring ignorance is ignorant, 2003.

[7] Pierre Lanchantin and Wojciech Pieczynski. Chaînes et arbres de markov évidentiels avec applications à la segmentation des processus non stationnaires. *Revue Traitement du Signal*, 22, 2005.

[8] Gregory Nuel and Bernard Prum. *Analyse statistique des squences biologiques*. Editions Herms, Labvoisier, Paris, 2007.

[9] Wojciech Pieczynski. Multisensor triplet markov chain and theory of evidence. *Int. J. on Approximate Reasoning*, 45:1–16, May 2007.

[10] Lawrence R. Rabiner. A tutorial on hidden markov models and selected applications in speech recognition. *Proceedings of the IEEE*, 77(2):257–286, Feb. 1989.

[11] Glenn Shafer. *A Mathematical Theory of Evidence*. Princeton University Press, Princeton, NJ, 1976.

[12] Glenn Shafer. Propagating belief functions in qualitative markov trees. *International Journal on Approximate Reasoning*, 1:349–400, 1987.

[13] Philippe Smets. Belief functions: the disjunctive rule of combination and the generalized Bayesian theorem. *Int. J. Approximate Reasoning*, 9:1–35, 1993.

[14] Philippe Smets. The application of the matrix calculus to belief functions. *Int. J. Approximate Reasoning*, 31:1–30, 2002.

[15] Philippe Smets and Robert Kennes. The transferable belief model. *Artificial Intelligence*, 66:191–234, 1994.

[16] Hélène Soubaras and Juliette Mattioli. Une approche markovienne pour la prvision du risque. In *Proceedings of 7th Congrs International Pluridisciplinaire Qualit et Sret de Fonctionnement QUALITA 2007*, pages 64–71, Tanger, Maroc, 20-22 March 2007.

[17] Hong Xu and Philippe Smets. Reasoning in evidential networks with conditional belief functions. *Int. J. on Approximate Reasoning*, 14:155–185, 1996.

[18] Amel Ben Yaghlane, Thierry Denœux, and Khaled Mellouli. Coarsening approximations of belief functions. In *Lecture Notes in Computer Science*, 2001.

[19] Boutheina Ben Yaghlane and Khaled Mellouli. Inference in directed evidential networks based on the transferable belief model. *Int. J. on Approximate Reasoning*, in press, available on-line 26 January 2008.