Belief Functions for Dynamical Systems

1^{st} spring school on belief functions theory and applications

Emmanuel Ramasso

Associate Professor at FEMTO-ST institute - Besançon
UMR CNRS 6174 - UFC / ENSMM / UTBM
Dep. Control Systems and Micro-Mechatronic Systems (AS2M)
http://www.femto-st.fr/~emmanuel.ramasso/

April 7th 2011

Dynamical systems

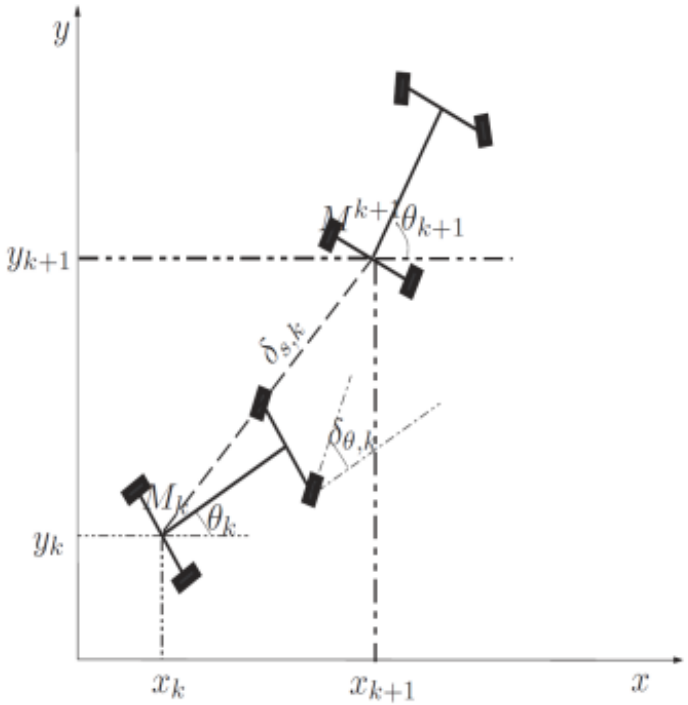
- Consider a dynamical system (note: generally with constant parameters), Ex: vehicle, human, machine. Provides a time-series (given sensors), generally multi-dimensional.
- Sensors are chosen to collect data (observations) concerning the system state.
- States are considered imprecise and uncertain
- Continuous states: Dynamics generally known, one wants to perform state filtering/smoothing and sometimes classification. Use of filters like Kalman (linear systems) or particle (non linear systems). Ex: positionning of targets.
- Discrete states (with continuous or discrete observations): Dynamics are generally estimated, one wants a segmentation of time-series and perform sequence classification. Ex: human motion analysis, fault detection in machines.

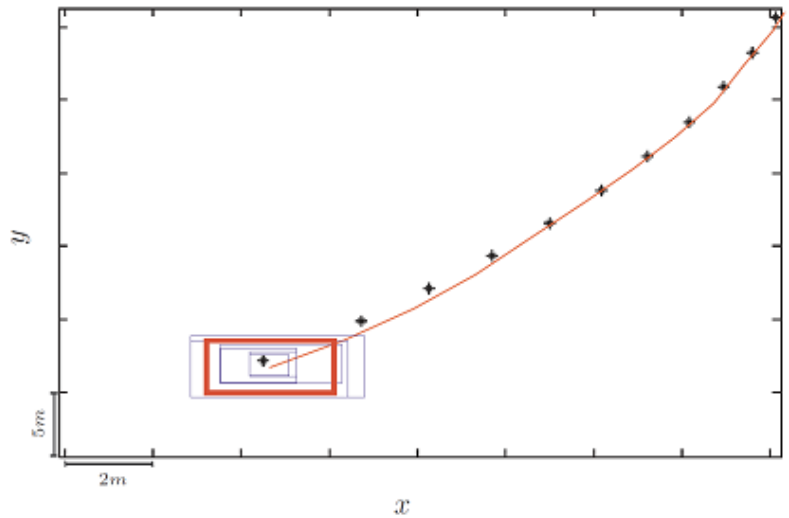
Temporal Belief Functions

- In this talk, system modelling is assumed to be made with belief functions
  - Continuous states: belief functions on reals (or intervals)
  - Discrete states: belief functions on discrete frame of discernement (FoD)
- What are Temporal Belief Functions
  - Belief on states at t **depends on** some previous belief (e.g $t-1$) and on observations
  - Goals: makes belief or/and observations smooth, classify and/or segment time-series

Illustration of continuous states

Imprecise state filtering with temporal beliefs on intervals: *[G. Nassreddine, F. Abdallah and T. Denoeux. State estimation using interval analysis and belief function theory: Application to dynamic vehicle localization. IEEE Trans. on Systems, Man and Cybernetics B, vol. 40, Issue 5, pages 1205-1218, 2010]*

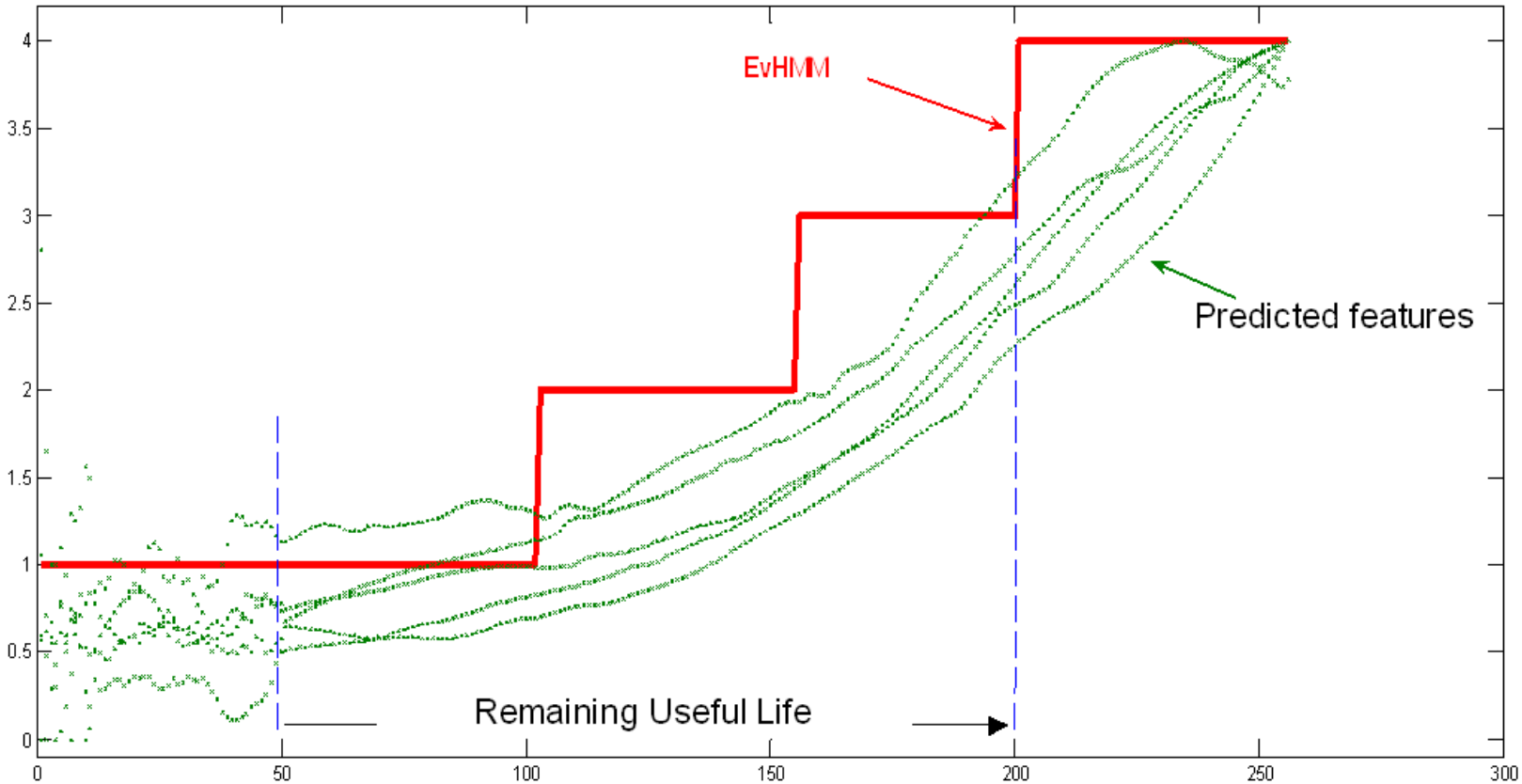




Intervals, interval expectations (bold), estimated position (star)

Illustration of discrete states

Prediction and classification of data using belief functions and neuro-fuzzy systems: *[Emmanuel Ramasso and Rafael Gouriveau, Prognostics in Switching Systems: Evidential Markovian Classification of Real-Time Neuro-Fuzzy Predictions, IEEE Int. Conf. on Prognostics and System Health Management, Hong-Kong, 2010]*



Objectives of this talk

The management of temporal belief functions is quite recent (end of 90's). This talk will:

- Focus mainly on discrete states
- Provide current solutions for dynamical systems analysis (*try* to be exhaustive in primary and relevant sources)
- Describe some tools, give some examples

- Note: assume you were serious students ☺

1. Belief functions
2. Implicit Markov Models and Belief Functions
3. Explicit Markov Models and Belief Functions
4. EvHMM: Evidential HMM
5. Time-Sliced Temporal Evidential Networks
6. Continuous state filtering
7. Conclusions and perspectives

1 Belief functions

2 Implicit Markov Models and Belief Functions

3 Explicit Markov Models and Belief Functions

4 EvHMM: Evidential HMM

5 Time-Sliced Temporal Evidential Networks

6 Continuous state filtering

7 Conclusions and perspectives

Origin

- Dempster in 1967 then Shafer in 1976: "Evidence Theory" (DST)
- Theory elaborated by Smets from 80's: "Transferable Belief Model" (TBM)
- The "belief function" formalism offers a general framework for the representation and combination of uncertain and imprecise data
- Generalizes probabilistic and set-membership approaches

Advantages

- Allow to represent the available knowledge, without unjustified prior, by relaxing the additivity constraint: doubt is explicitly represented
- Smets work allowed to
  - Generalize Bayesian inference to belief functions without prior in the GBT *Smets 93 - [Belief functions: The disjunctive rule of combination and the generalized Bayesian theorem, International Journal of Approximate Reasoning, Vol. 9(1), 1993, pp 1-35]*
  - Show that information brought by the emptyset can be useful, in particular for the open world assumption *[Analyzing the combination of conflicting belief functions, Information Fusion, Vol. 8(4), 2007]*
- A lot of tools have been proposed (in particular from 2000) and a lot of communications have appeared

Basic Belief Assignment (BBA)

- A distribution of masses (BBA) is $m^{\Omega_t} : 2^{\Omega_t} \rightarrow [0,1], A \rightarrow m^{\Omega_t}(A), \sum_{A \subseteq \Omega_t} m^{\Omega_t}(A) = 1$
- With N states in $\Omega_t = \{s_1, s_2, \ldots s_N\}$
- The conditional mass is denoted here $m^{\Omega_t|\Omega_{t-1}}[B](C),\ C \subseteq \Omega_t, B \subseteq \Omega_{t-1}$
- A BBA is in one-to-one correspondance with communalities q, plausibilities pl, and credibilities bel... improve interpretation and simplify some computations. For example the conjunctive rule: $q_{1,2}^{\Omega_t}(A) = q_1^{\Omega_t}(A) \cdot q_2^{\Omega_t}(A)$
- $m^{\Omega_t}(\emptyset) = 1 - pl^{\Omega_t}(\Omega_t)$ is called the mass of conflict
- Example:

A	$\emptyset$	s_1	s_2	$\{s_1, s_2\}$	s_3	$\{s_1, s_3\}$	$\{s_2, s_3\}$	$\{s_1, s_2, s_3\}$
$m_1(A)$	0.1	0.2	0.05	0.15	0.08	0.12	0.17	0.13
$m_2(A)$	0	0	0	0	0.2	0.5	0	0.3
$m_3(A)$	0	0.6	0.1	0	0.3	0	0	0
$m_4(A)$	0	1	0	0	0	0	0	0
$m_5(A)$	1	0	0	0	0	0	0	0
$m_6(A)$	0	0	0	0	0	0	0	1

Propagations

- The "f"-total law [Smets 93] - Specialization:

$$f^{\Omega_t}(C) = \sum_{B \subseteq \Omega_{t-1}} f^{\Omega_t|\Omega_{t-1}}[B](C) \cdot m^{\Omega_{t-1}}(B)$$

  with $f \in \{m, q, pl, bel, w\}$

- The Generalized Bayesian Theorem (GBT):

$$q^{\Omega_{t-1}|\Omega_t}[C](B) = \prod_{\theta \in B} pl^{\Omega_t|\Omega_{t-1}}[\theta](C)$$

- Decision making
  - using the pignistic transform *[P. Smets, Decision making in the TBM: the necessity of the pignistic transformation, Int. J. of Approximate Reasoning, Vol. 38(2), pp. 133-147, 2005]*:
  - or the plausibility transform PlP *[Barry R. Cobb, Prakash P. Shenoy: On the plausibility transformation method for translating belief function models to probability models. Int. J. Approx. Reasoning 41(3): 314-330 (2006)]*: takes plausibilities on singletons, normalise

1 Belief functions

2 Implicit Markov Models and Belief Functions

3 Explicit Markov Models and Belief Functions

4 EvHMM: Evidential HMM

5 Time-Sliced Temporal Evidential Networks

6 Continuous state filtering

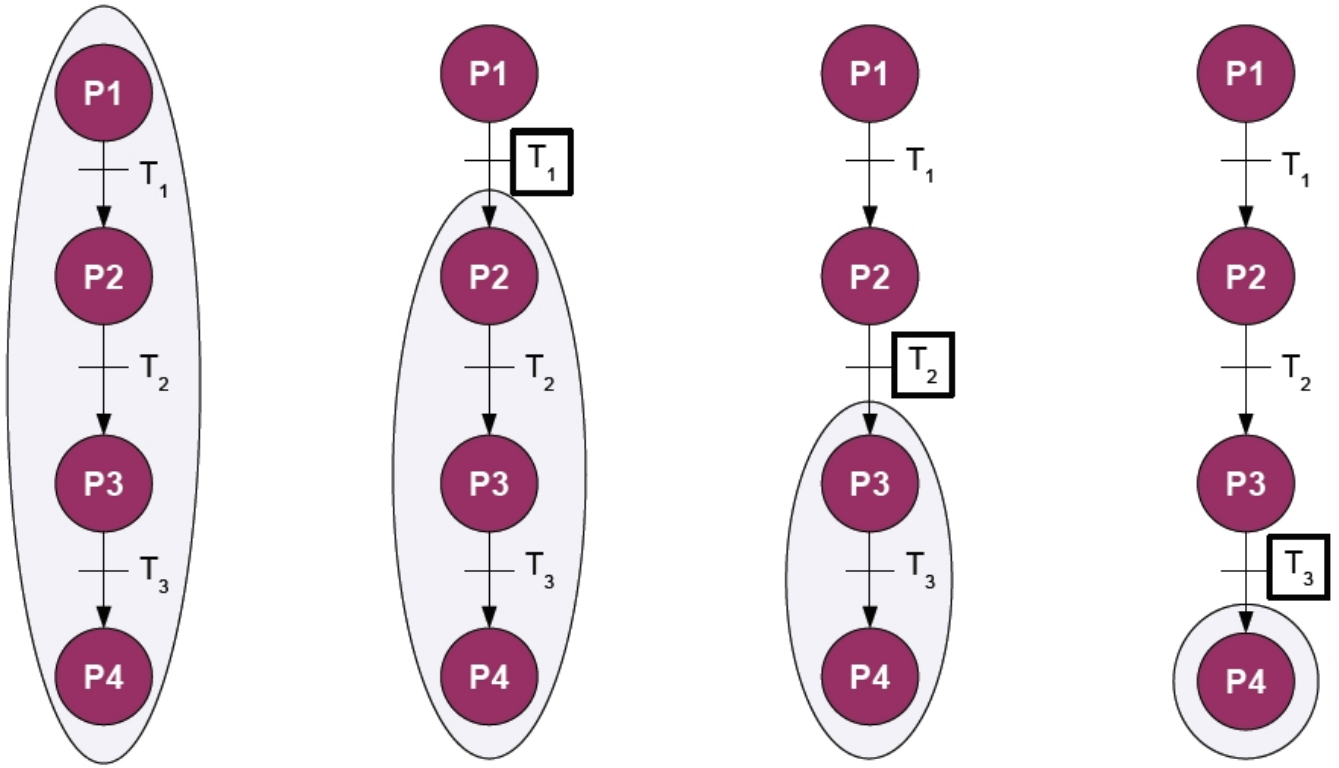
7 Conclusions and perspectives

The "Belief Petri Net" (BPN, 1999)

- Reference: *[M. Rombaut, I. Jarkass, and T. Denoeux, State recognition in discrete dynamical systems using Petri nets and Evidence theory. In Europ. Conf. on Symbolic and Quantitative Approaches to Reasoning with Uncertainty (ECSQARU), pages 352-361, 1999]*
- First application of the DRC + total plausibility law on temporal belief functions
- The sequence of states is modelled by a simple tree
- The tree is initialized by the vacuous BBA
- The transitions evolve with observations (need to find rules)
- The mass on states flows down in the tree according to the transitions

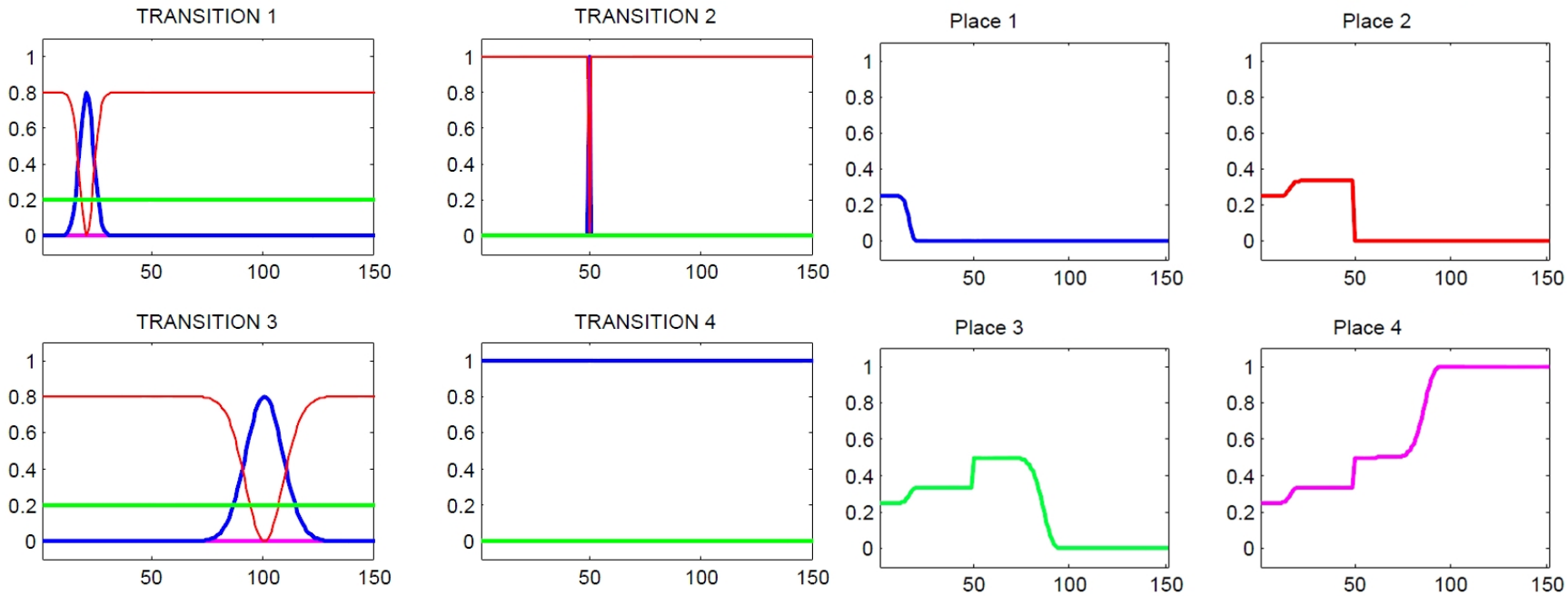
The "Belief Petri Net" (BPN, 1999)

- Apply the "f"-total law to make belief flow-down
- Ex: transitions are precise and certain (i.e. conditioning)



The "Belief Petri Net" (BPN, 1999)

- An example with uncertain and imprecise transitions



- Classification: see the mass on the last state

Details

- Transitions between state k and $k+1$ can be true or false: binary BBA on $\Omega_R^k = \{t, f\}$
- BBA on transitions are translated into a transition matrix using a refinement:

$$m^{\Omega_t|\Omega_{t-1}}[s_k](s_k) \leftarrow m^{\Omega_R^k}(f)$$

$$m^{\Omega_t|\Omega_{t-1}}[s_k](s_{k+1}) \leftarrow m^{\Omega_R^k}(t)$$

$$m^{\Omega_t|\Omega_{t-1}}[s_k](s_{k+1} \cup s_k) \leftarrow m^{\Omega_R^k}(t \cup f)$$

- Note: the transition matrix is sparse
- How to get $m^{\Omega_t|\Omega_{t-1}}[A](\cdot), A \subseteq \Omega_{t-1}$? Apply the DRC [Smets 93] between conditionals to singletons in A to obtain conditionals on A
- Then apply the m-total law for transfering the preceeding mass defined on Ω_{t-1} onto Ω_t:

$$m^{\Omega_t}(C) = \sum_{B \subseteq \Omega_{t-1}} m^{\Omega_t|\Omega_{t-1}}[B](C) \cdot m^{\Omega_{t-1}}(B)$$

- Matlab demo

An Evidence Filtering approach (TEF)

- Reference: *D.A. Dewasurendra, Peter H. Bauer and Kamal Premaratne, Evidence Filtering, IEEE Trans. Signal Processing, vol. 55(12), pp. 5796-5805, Dec. 2007*
- Problem: decision in distributed sensors (network), data imprecision, sensor reliability, several sources...
- Dempster-Shafer approach, combined with discrete time filtering techniques, filter called TEF (Temporal Evidential Filter)
- Provides inferences in the frequency domain based on sensors fusion

Principle

- They consider the uncertainty on a situation A estimated by $[bel(A)\ pl(A)]$
- They update these bounds by:

$$bel(B)_{1,k+1} = \alpha_k bel(B)_{1,k} + \beta_k bel(B|A)_{2,k}$$

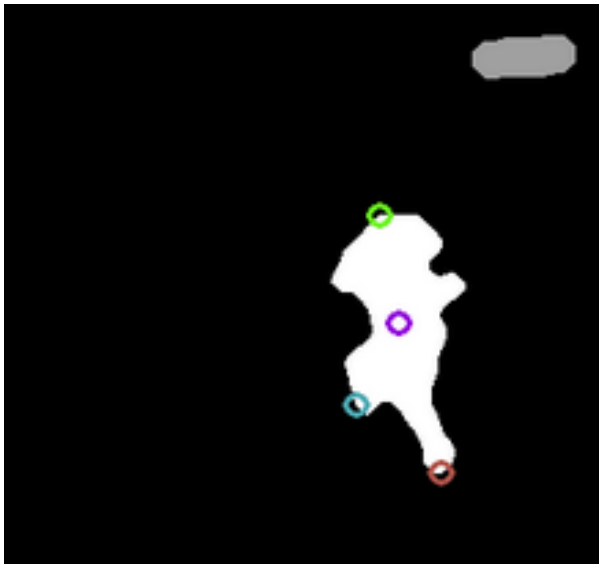
$$pl(B)_{1,k+1} = \alpha_k pl(B)_{1,k} + \beta_k pl(B|A)_{2,k}$$

- The first part concerns past knowledge, the second part integrates the current measurements
- Equations are rewritten recursively, and then interpreted as a 1-D first-order linear filter (weights supposed constant) - exploiting concepts in automatic control
- This common type of filter can be characterized in the time domain by its impulse response and in the spectral domain by its z-transform
- How they generated masses? (not described)

An Evidential Filtering approach (TCF)

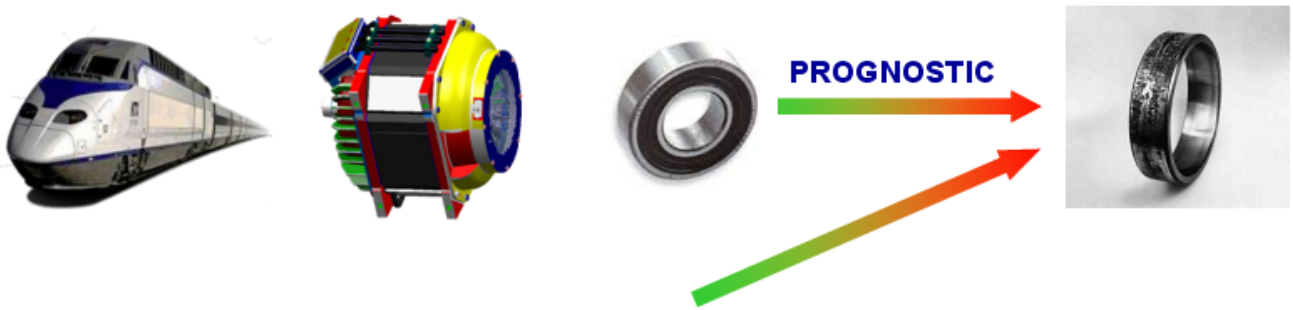
- Reference: *[E. Ramasso and M. Rombaut and D. Pellerin, State filtering and change detection using TBM conflict - Application to human action recognition in athletics videos, IEEE Trans. on Circuits and Systems for Video Technology, 17(7):944-949, 2007]*
- Context [PhD]: State sequence recognition applied to human motion analysis - Online detection of actions (states) and recognition of activities (sequences) of athletics in videos, collaboration with University of Crete

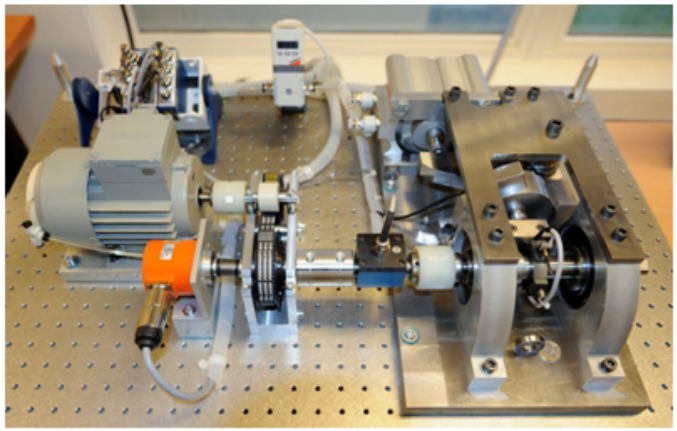




An Evidential Filtering approach (TCF)

- Context [current]: Machine fault prognosis and diagnosis - Lowest Replaceable Unit, Prognostics of health state





TCF principle

- Assumes a set of N binary states, each can be more or less true or false ($\Omega_t = \{T_t, F_t\}$, or Ω_t^k if required)
- Transitions between hypotheses are doubtful
- The TCF makes beliefs on states smooth
- Based on a prediction-fusion scheme, coupled with switching models
- BBA generated here by the likelihood model [Ref. Appriou and Smets] and the distance model [Ref. Denoeux]

Switching models

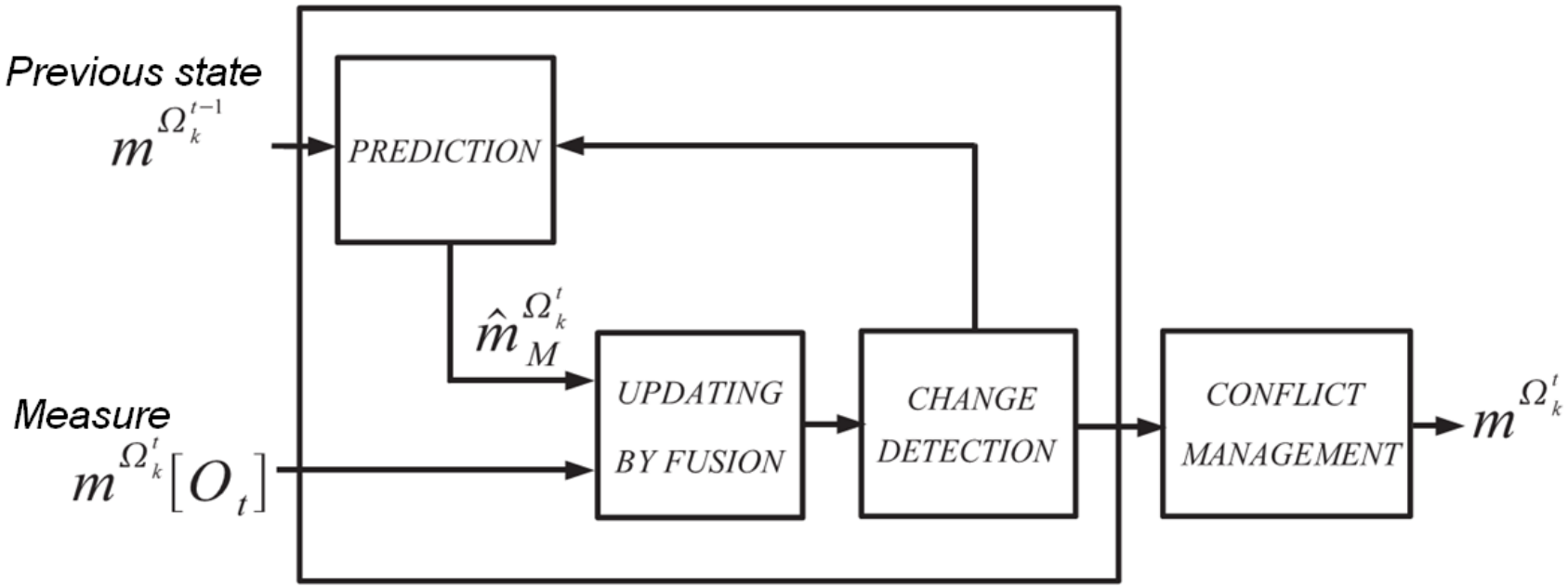
- Evolution rule: if an action is true (resp. false) at $t-1$ then it remains true (resp. false) at $t \Rightarrow$ *[Branko Ristic and Philippe Smets, Target Identification Using Belief Functions and Implication Rules, IEEE transactions on aerospace and electronic systems, vol. 41(3), pp. 1097-1103, 2005]* (implication and deconditioning)
- Two modes $\mathcal{M} = \{\mathcal{T}, \mathcal{F}\} \Rightarrow$ two rules, but only one applied at each instant
- Translation of evolution rules into distributions of masses (BBA) conditional to the mode, example for model "state is true"

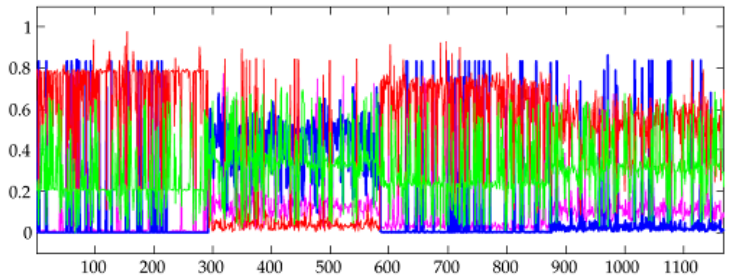
$$m_{\mathcal{T}}^{\Omega_t}[T_{t-1}](T_t) = \gamma_{\mathcal{T}}$$

$$m_{\mathcal{T}}^{\Omega_t}[T_{t-1}](\Omega_t) = 1 - \gamma_{\mathcal{T}}$$

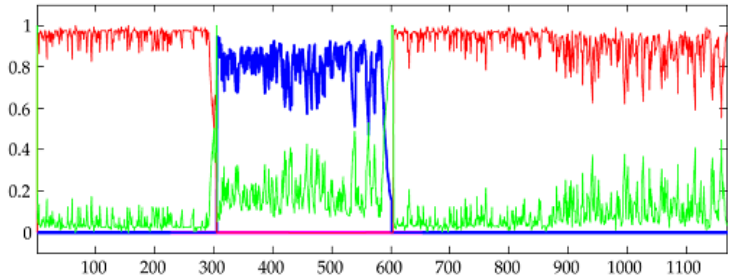
- A similar model for "state is false"

TCF principle





Input BBA



Output BBA

Prediction-Fusion

- Combination of the evolution model output with the BBA obtained at from observations:

$$m^{\Omega_t} = \hat{m}^{\Omega_t}_{\mathcal{M}} \textcircled{\cap} \tilde{m}^{\Omega_t}$$

- Models in the form: "*if true then true with belief* $\gamma_{\mathcal{M}}$*, remainder on ignorance...*", so applying the models is equivalent to a discounting:

$$\hat{m}^{\Omega_t}_{\mathcal{T}}(T_t) = \gamma_{\mathcal{T}} \cdot m^{\Omega_{t-1}}(T_{t-1})$$

$$\hat{m}^{\Omega_t}_{\mathcal{T}}(\Omega_t) = \gamma_{\mathcal{T}} \cdot m^{\Omega_{t-1}}(\Omega_{t-1}) + 1 - \gamma_{\mathcal{T}}$$

- Similarly for $\hat{m}^{\Omega_t}_{\mathcal{F}}(F_t)$

Exploiting conflict for model change

- Conflict value:

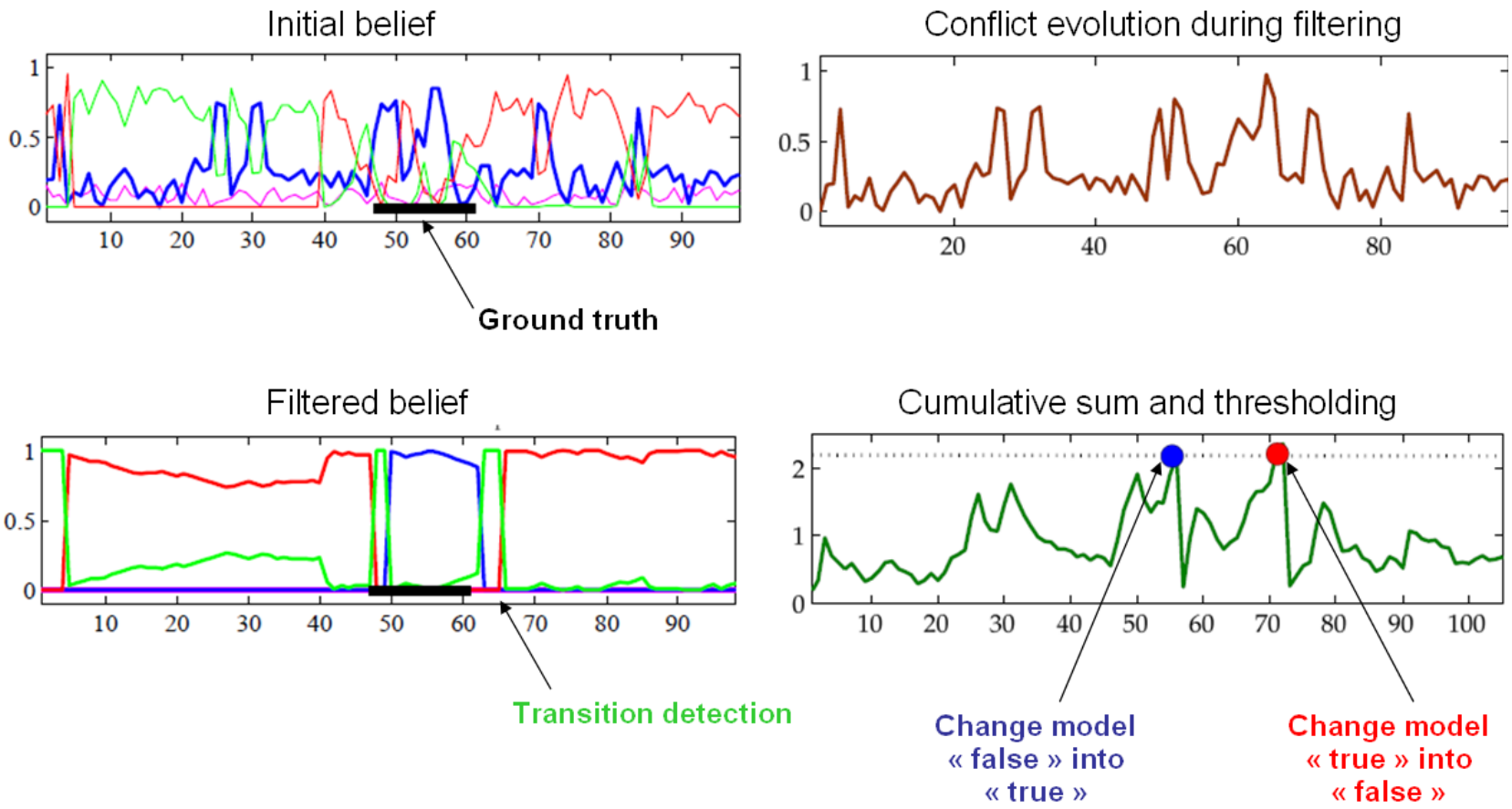
$$\epsilon_t = \hat{m}^{\Omega_t}_{\mathcal{M}} \textcircled{\cap} \tilde{m}^{\Omega_t}(\emptyset)$$

- Reflects the degree of contradiction between the model (evolution rule) and observations
- If conflict is too high then change the model
- Decision on the cumulative sum of conflict (with fading $\lambda \in [0, 1]$)

$$\mathbf{CS}(t) \leftarrow \mathbf{CS}(t-1) \cdot \lambda + \epsilon_t$$

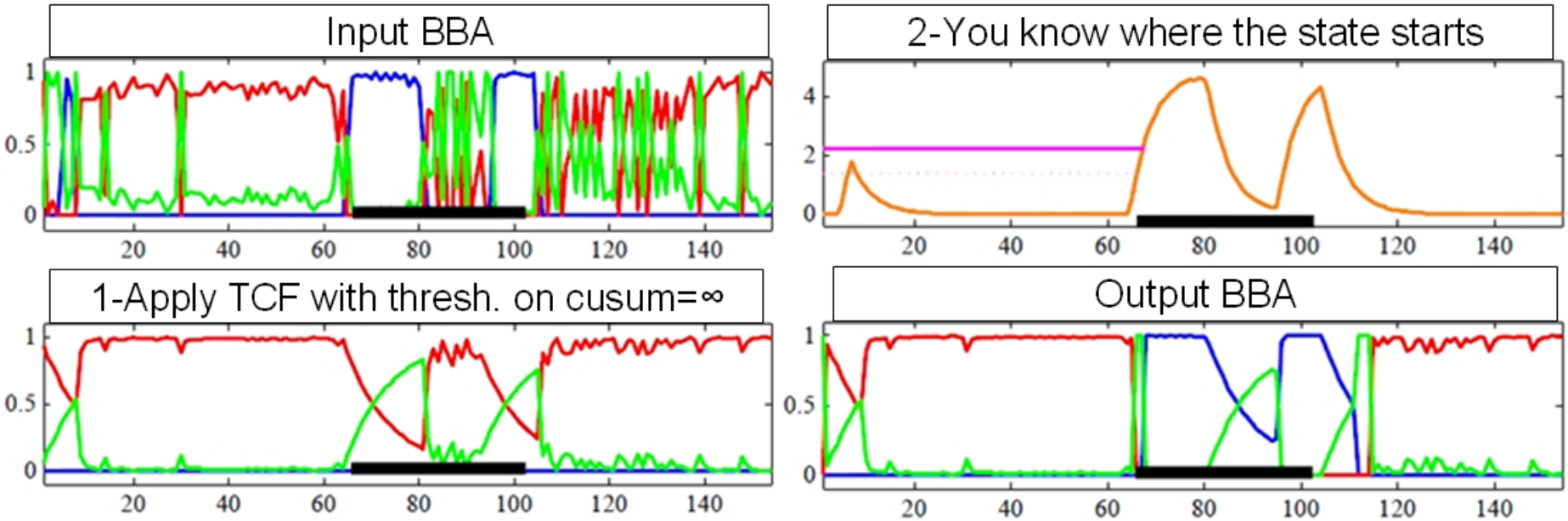
- Using a threshold to detect the change: if $\mathbf{CS}(t) > \mathbb{T}_s$ then change the model

Illustrate model change



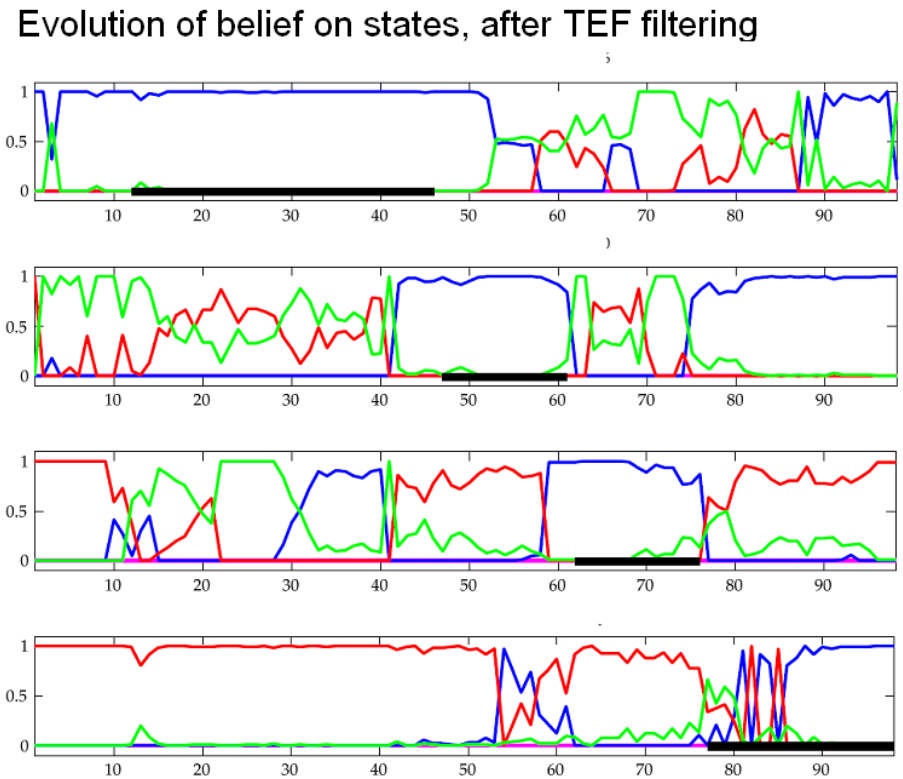
Color : state is true in blue, false in red and doubt in green

Find the stop threshold



An example

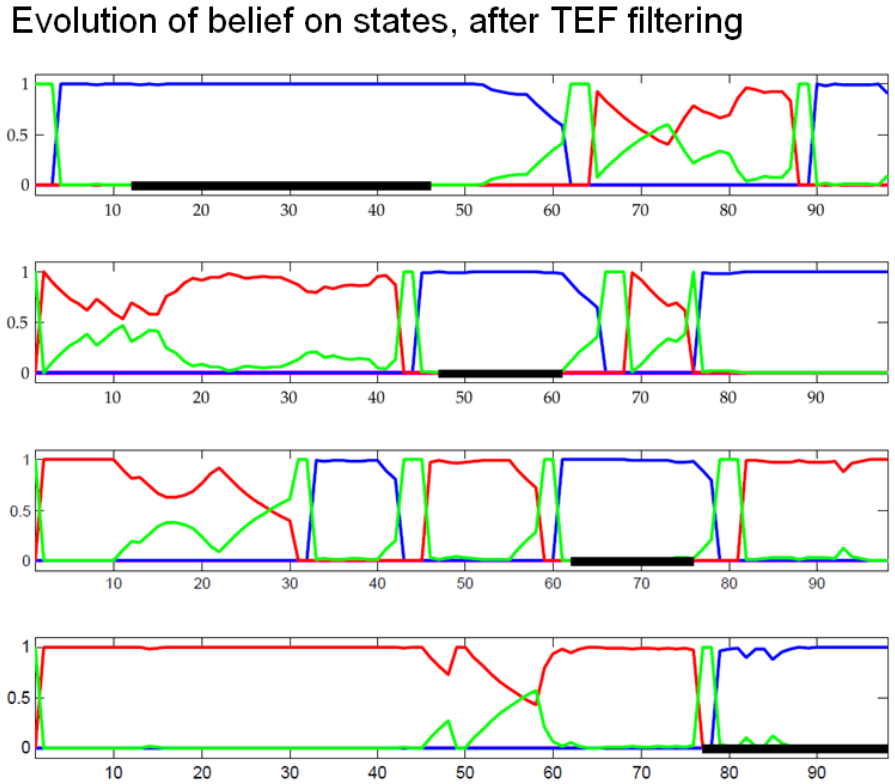
- Segmenting a data stream (before)



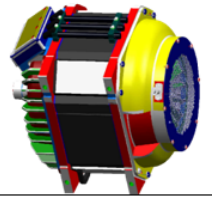


An example

- Segmenting a data stream (after)



Indexing a video database from motion





Engine OK



Anomaly ?



Degrading

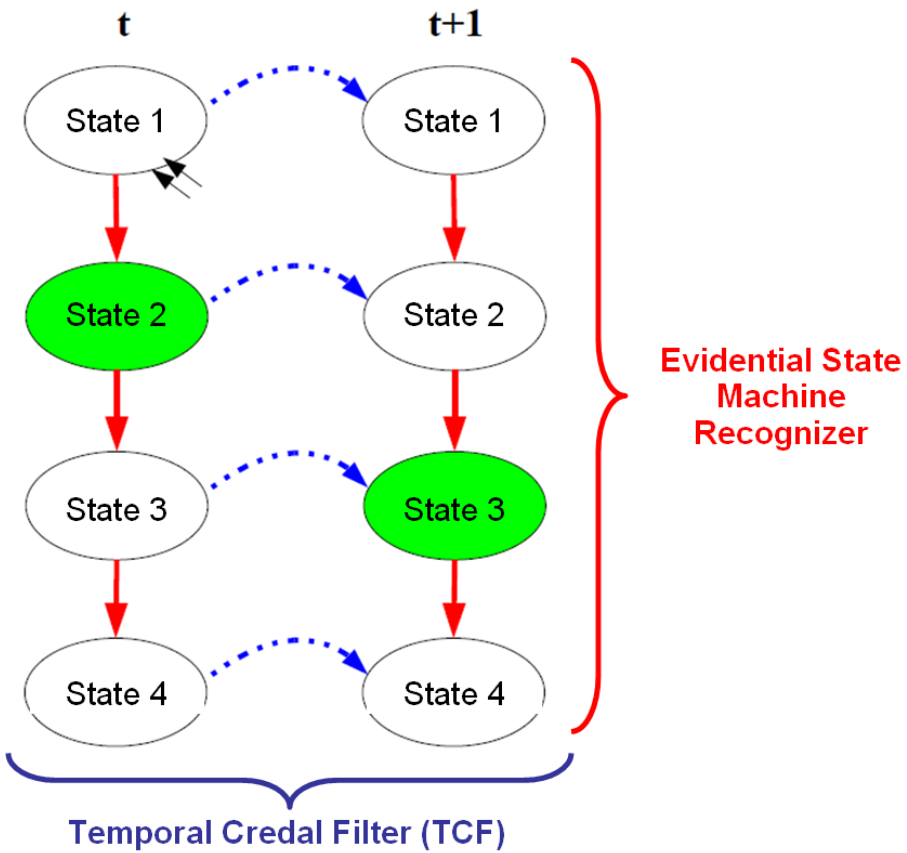


Fault mode

- Ensuring state synchronisation ?

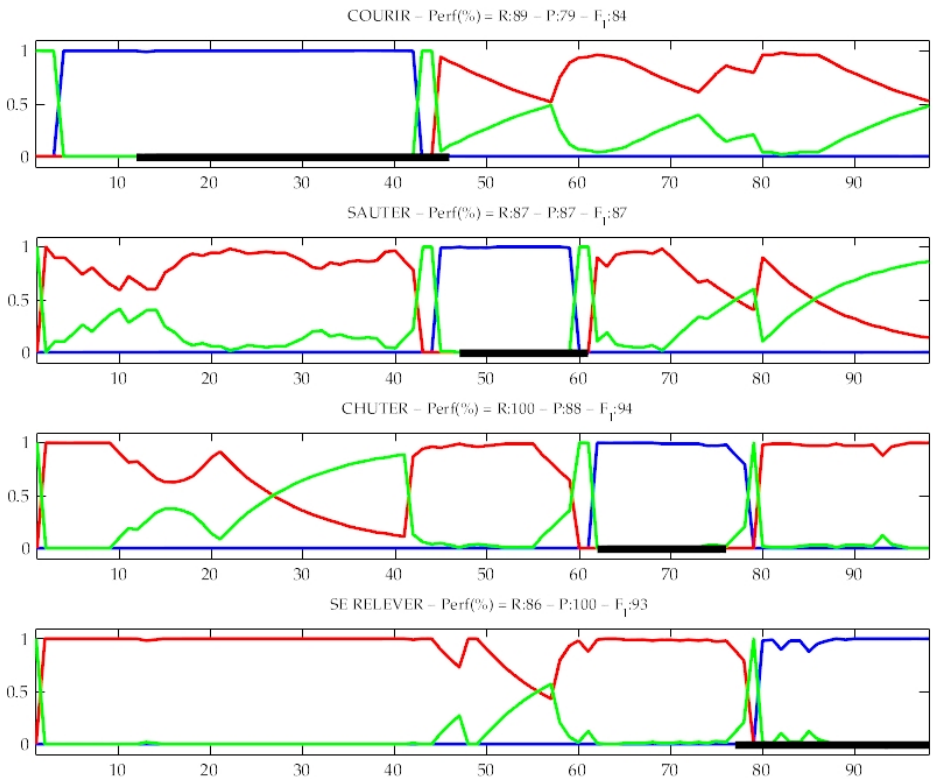
State machine for sequence recognition (BSS, 2006, 2010)

- Reference: *[E. Ramasso and C. Panagiotakis and M. Rombaut and D. Pellerin, Belief Scheduler based on model failure detection in the TBM framework. Application to human activity recognition, International Journal of Approximate Reasoning, 51(7), 2010]*



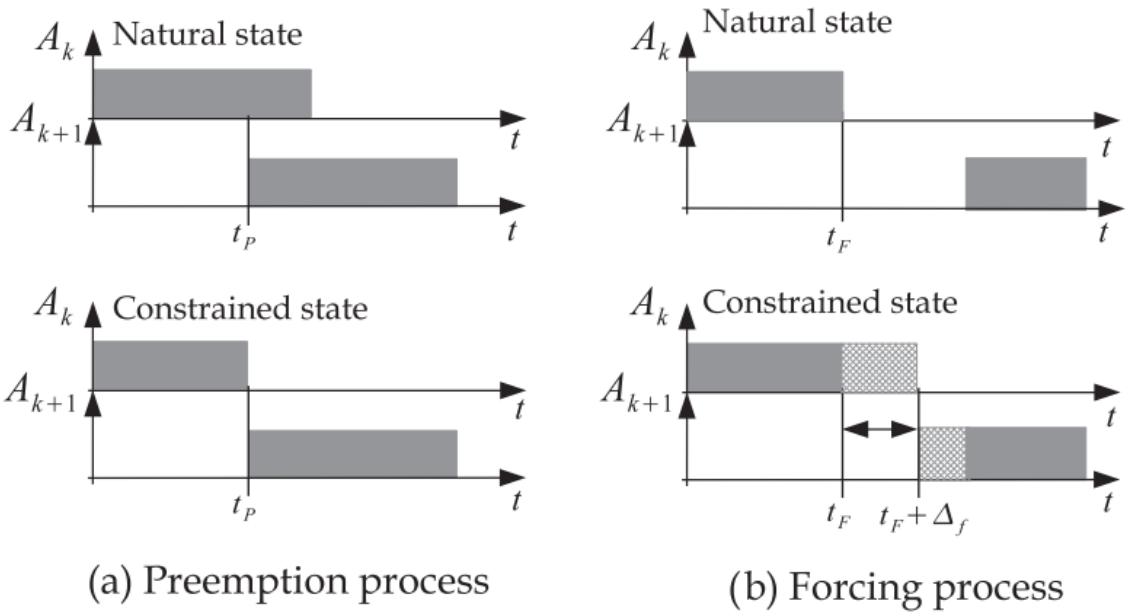
Goal

- BSS manages disjoint (but consecutive) states and superposed states
- Hypothesis: one state true at a given time
- Thus one TCF designed for each state, the BSS manages the commutations of models according to the sequence of states



Principle

- Two rules: one for "preemption", one for "forcing"

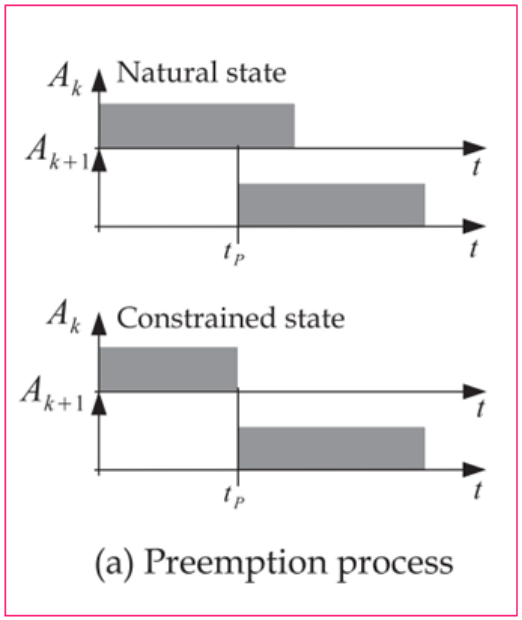


(a) Preemption process

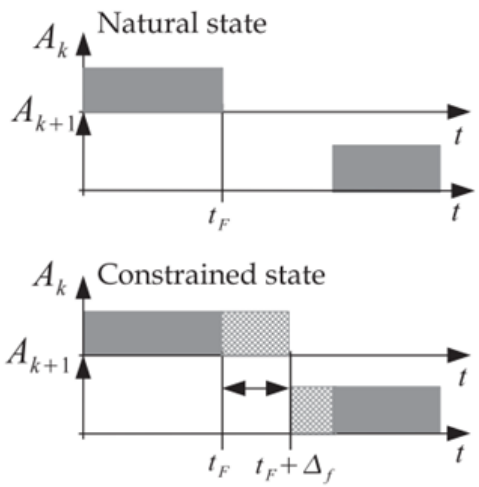
(b) Forcing process

Preemption process

- Manages overlapped states. At $t = t_P$, A_k is still true while A_{k+1} becomes true ($CS_{k+1}(t)$ is greater than its stop threshold), thus two states are true at the same time (A_{k+1} wants to preempt A_k)
- The natural state of A_{k+1} is true from time t_P and the constrained state of A_k is false (unless validation is not fulfilled, cf. paper)



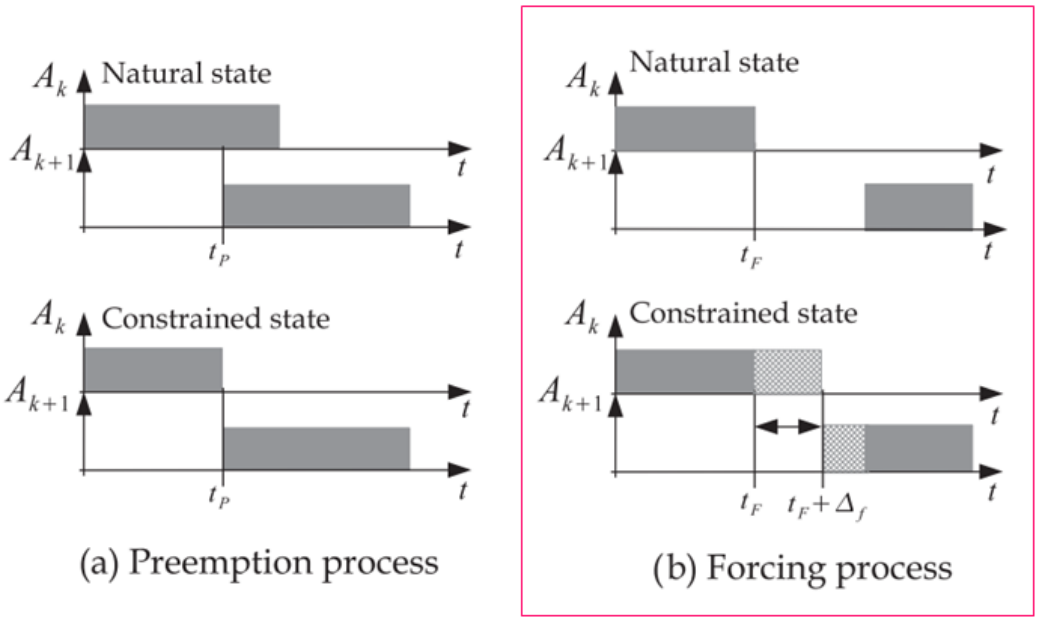
(a) Preemption process



(b) Forcing process

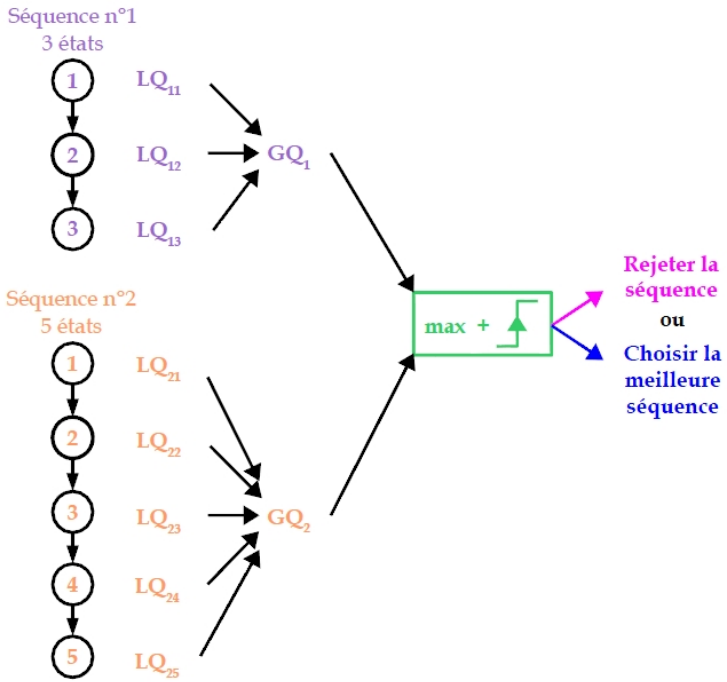
Forcing process

- Manages disconnected states. At $t = t_F$, the current state A_k is false as well as the next state A_{k+1}
- The constrained state of A_k is forced to be true until A_{k+1} becomes true (or after $t_F + \Delta_F$)



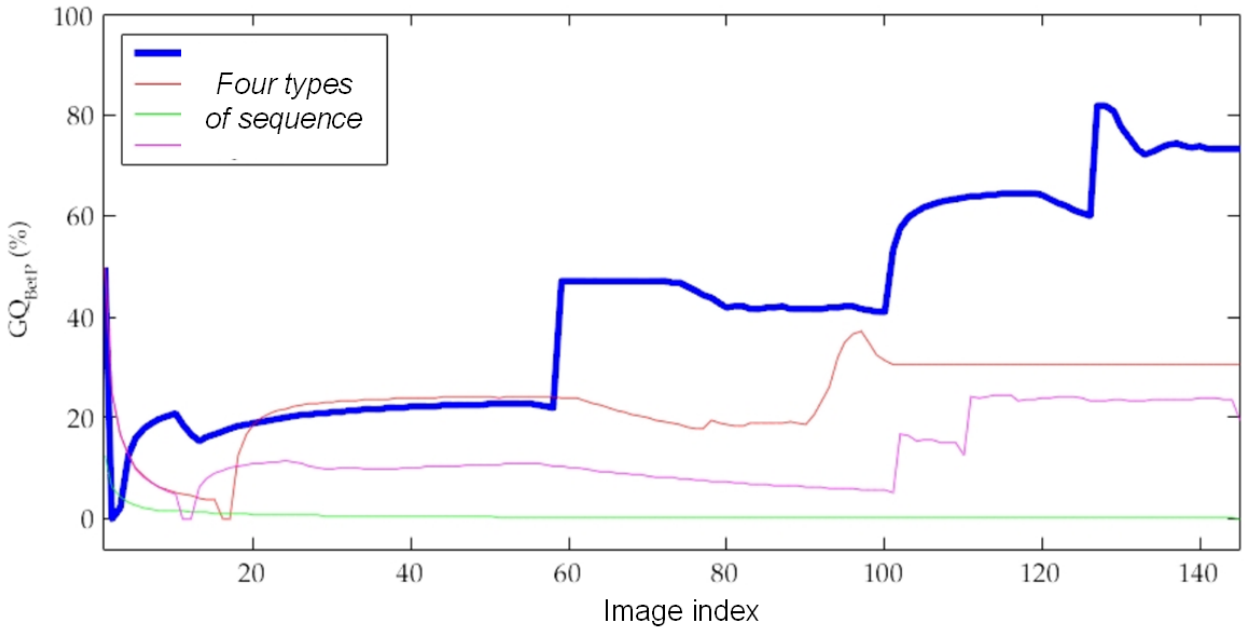
Classification of sequences with BSS

- Local criterion: For each state, given the evolution model is $\mathcal{T}$, compute the average value of the belief on hypothesis "state is true" (recur.)
- Global criterion: For each sequence, compute the mean value of local criteria over states where the TCF uses the model $\mathcal{T}$



Classification of sequences with BSS

- Online calculation of the classification criterion, example with four possible sequences: is the athlete performing pole-vault, high-jump, long-jump, triple-jump?



- Note (fuzzy approach) *[L. Rodriguez-Benitez, C. Solana-Cipres, J. Moreno-Garcia, L. Jimenez-Linares, Approximate reasoning and finite state machines to the detection of actions in video sequences, Int. Journal of Approximate Reasoning, Volume 52, Issue 4, June 2011, Pages 526-540]*

1 Belief functions

2 Implicit Markov Models and Belief Functions

3 Explicit Markov Models and Belief Functions

4 EvHMM: Evidential HMM

5 Time-Sliced Temporal Evidential Networks

6 Continuous state filtering

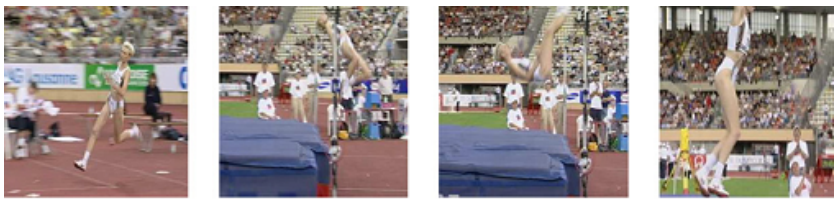
7 Conclusions and perspectives

Time series modelling

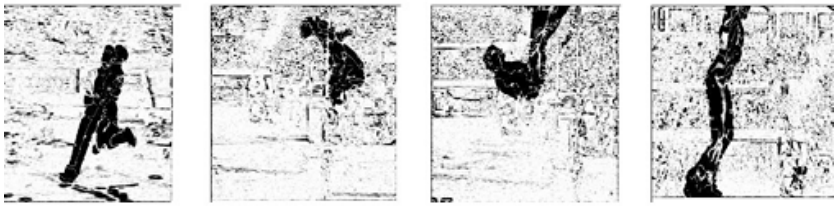
- Reference: *[L. Rabiner, A tutorial on hidden markov models and selected applications in speech recognition, Proc. IEEE, vol. 77, pp. 257-285, 1989]*
- Model a time series as being generated by two processes: one generating observations and one generating hidden discrete states
- Ex: in human motion analysis from videos, you only observe some characteristics of relevant points on the silhouette not the motion itself
- Ex: in speech recognition, you only observe some characteristics extracted from an audio stream (by a microphone), not the words themselves
- Note: "Hidden" considered in previous examples! But in HMM, all parameters are optimized jointly to maximize likelihood
- Use of HMM to discriminate (ex: classification) or to generate (ex: DNA applications)
- Notations of the FoD: $\Omega_t = \{s_1, s_2 \dots s_N\}$

Example

- Five time series (multi-dim. observations) extracted

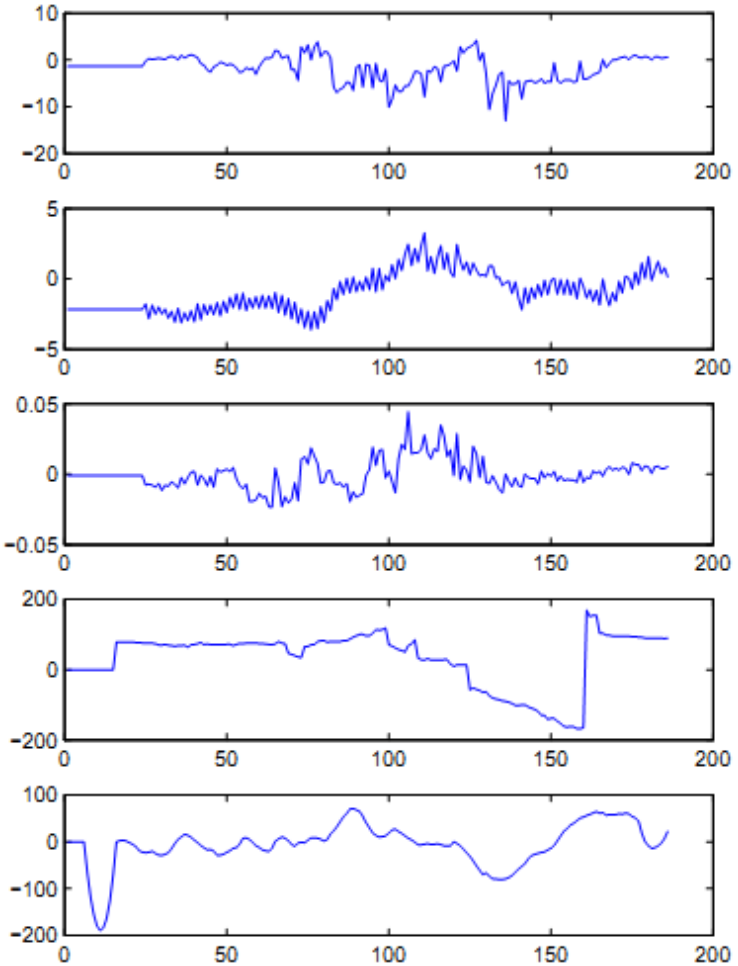


(a) Original video sequence.



(b) Dominant motion images.





Can you see the 4 states?

Goals

Given a (multi-dim.) time-series (sequence of observations)

- P1 - Classification: Classify the sequence into possible HMM models (Ex: identify the motion, the speaker or the words given some observations) $\Rightarrow$ Forward algorithm
- P2 - Segmentation: Determine the underlying sequence of (hidden) discrete states from the noisy observations (Ex: segment the motion into elementary parts, segment a word into letters) $\Rightarrow$ Viterbi algorithm
- P3 - Training: Estimate parameters of HMM $\Rightarrow$ Baum-Welch algorithm
  - governing the generation of likelihoods of observations (given states)
  - links (conditional probabilities) between states at consecutive instants
  - prior on first states
  - all jointly!

Solving P1 and P2

- P1 is solved by applying the forward algorithm relying on:

$$\alpha_t(s_j) = \sum_{\forall s_i \in \Omega_{t-1}} \alpha_{t-1}(s_i) \cdot a_{ij} \cdot b_j(\mathbf{O}_t)$$

$$L_p = \sum_{j \in \Omega_T} \alpha_T(s_j)$$

- P2 is solved by applying the Viterbi algorithm relying on:

$$\delta_t(s_j) = \max_{\forall s_i \in \Omega_{t-1}} \delta_{t-1}(s_i) \cdot a_{ij} \cdot b_j(\mathbf{O}_t)$$

and store arguments for backtracking

Propositions of generalization of HMM to belief functions - Beginning in 2000's

- Why?
  - Take imprecision into account (of sensors or prior)
  - Improve fusion
  - Avoid assuming independance of observations

In 2000's

- With fuzzy measures (Choquet capacity): *[M.A. Mohamed and P. Gader, Generalized HMM - part I: Theoretical frameworks, IEEE Tr. on Fuzzy Systems, vol. 8, pp. 67-81, 2000]*

$\Rightarrow$ Initially, general model but assumes several hypotheses at several stages of their algorithms to tend to simple product likewise to probabilistic HMM: expressions similar at the end

- Work of W. Pieczynski et al. [2000-2007]

$\Rightarrow$ Approach rather probabilistic (explications to come ☺)

- Work at GIPSA-lab, then FEMTO-ST [since 2007]

$\Rightarrow$ General approach without any probabilistic assumption (to be described in the sequel)

Fouque, Appriou et Pieczynski (2000)

- Reference: *[L. Fouque, A. Appriou, and W. Pieczynski, An evidential Markovian model for data fusion and unsupervised image classification, Int. Conf. on Information Fusion, Vol. 1, 2000, Paris, France, pp. TuB4-25 - TuB4-31]*
- Sensors provide belief masses (likelihood model): their combination by Dempster rule is a BBA m_1
- The prior is a Bayesian BBA m_2
- Their combination is thus Bayesian
- Propagation to the vicinity made using probabilistic transitions

- Note: For sensors, it is sufficient to keep only plausibilities on singletons

- Probabilistic transitions are sometimes sufficient if the evolution model is reliable while sensors report are imprecise

Lanchantin et Pieczynski (2005), Pieczynski (2007)

- Reference: *[P. Lanchantin and W. Pieczynski, Unsupervised restoration of hidden non stationary Markov chains using evidential priors, IEEE Trans. on Signal Proc., Vol. 53(8), pp. 3091-3098, 2005]*
- Reference: *[W. Pieczynski, Multisensor triplet Markov chains and theory of evidence, Int. Journal of Approx. Reasoning, Vol. 45, No. 1, pp. 1-16, 2007]*
- Only prior are belief
- The same conclusion as previously

Pieczynski et Benboudjema (2006)

- Reference: *[W. Pieczynski and D. Benboudjema, Multisensor triplet Markov fields and theory of evidence, Image and Vision Comp., Vol. 24, No. 1, pp.61-69, 2006]*
- Section 3: Only priors are belief
- Section 4: Only sensors provides belief
- The same conclusion as previously

HMM and partially supervised training [MLSP 09]

- Reference: [MLSP 09] - *[E. Ramasso, Contribution of belief functions to HMM with an application to fault diagnosis, In IEEE Int. Workshop on Machine Learning and Signal Proc., 2009]*
- Problem: we know *partially* the states for some training data, i.e. with a belief function expressed on the set of states. How to train prob. HMM?
- Solutions exist for semi-supervised training in prob. HMM $\Rightarrow$ called *Partially Hidden Markov Models* in the litterature, but the state is either known of unknown
- A solution proposed in [MLSP 09], called Partially-Hidden Markov Models. Example of forward propagation:

$$\alpha_t(s_j) = \sum_{\forall s_i \in \Omega_{t-1}} \alpha_{t-1}(s_i) \cdot a_{ij} \cdot b_j(\mathbf{O}_t) \cdot \mathbf{pl_o}^{\Omega_t}(s_j)$$

- Based on an idea developed for static data in *[E. Come, L. Oukhellou, T. Denoeux and P. Aknin, Learning from partially supervised data using mixture models and belief functions. Pattern Recognition, vol. 42, 334-348, 2009]*

1 Belief functions

2 Implicit Markov Models and Belief Functions

3 Explicit Markov Models and Belief Functions

4 EvHMM: Evidential HMM

5 Time-Sliced Temporal Evidential Networks

6 Continuous state filtering

7 Conclusions and perspectives

Everything is evidence

- Prior $\Rightarrow$ BBA at $t = 1$ (on subsets of states)
- Observations $\Rightarrow$ BBA at t (related to subsets of states)
- Transitions $\Rightarrow$ BBA on subsets at t given subsets at $t - 1$

- Propagations (based on [Smets 93] + management of the temporal case)

  PhD: *E. Ramasso, Reconnaissance de séquences d'états par le modèle des croyances transférables, application à l'analyse de vidéos d'athlétisme, thèse de l'UJF (au GIPSA-lab), 2007*

  ESCQARU 07: *E. Ramasso and M. Rombaut and D. Pellerin, Forward, Backward and Viterbi procedures in the Transferable Belief Model for state sequence analysis using belief functions, In ECSQARU 2007*

  MLSP 09 *E. Ramasso, Contribution of belief functions to HMM with an application to fault diagnosis, In IEEE International Workshop on Machine Learning and Signal Processing, 2009*

Advantages of EvHMM

- First states unknown
- "Emission" models generate belief on states given (imprecise) observations
- Representation of doubt between states (explicit transitions)
- Data labelled by belief functions
- Observations not always independent
- Reliability, fusion, conflict management ...

Forward-Backward Propagations

- Forward - filtering variable:

$$q_{\alpha}^{\Omega_t}(S_j) = q_{b}^{\Omega_t}(S_j|\mathbf{O_t}) \cdot \left(\sum_{S_i \subseteq \Omega_{t-1}} m_{\alpha}^{\Omega_{t-1}}(S_i) \cdot q_{a}^{\Omega_t|\Omega_{t-1}}(S_j|S_i) \right)$$

  Expression is correct with communalities

  Be careful to the BBA $m_{\alpha}^{\Omega_{t-1}}$

- Backward variable:

$$q_{\beta}^{\Omega_t}(S_i) = \sum_{S_j \subseteq \Omega_{t+1}} m_{\beta \cap b}^{\Omega_{t+1}}(S_j) \cdot q_{a}^{\Omega_t|\Omega_{t+1}}(S_i|S_j)$$

  Be careful to the GBT

- Smoothing variable:

$$q_{\gamma}^{\Omega_t}(S_i) = q_{\alpha}^{\Omega_t}(S_i) \cdot q_{\beta}^{\Omega_t}(S_i)$$

- Local decision with $\mathrm{BetP}\{\gamma\}$ or $\mathrm{PlP}\{\gamma\}$

Case of a Evidential Markov Chain

- In a chain, the state is completely known at t. In an EMC, the state can be imprecise $\Rightarrow m_b^{\Omega_t}(A_t) = 1$ for some subset $A_t \subseteq \Omega_t$.
- Proposition in *[W. Pieczynski, Multisensor triplet Markov chains and theory of evidence, Int. J. of Approx. Reasoning, 45(1), pp. 1-16, 2007]*: "*Déf. 1.4: An EMC can be written* $m(A_1)m(A_2|A_1)\ldots m(A_t|A_{t-1})$ *where m is a BBA*"

- Alternative - Since we have $m_b^{\Omega_{t-1}}(B_{t-1}) = 1$ for some subset $B_{t-1} \subseteq \Omega_{t-1}$:
  - $\Rightarrow$ Condition transitions: $m_a^{\Omega_t|\Omega_{t-1}}[B_{t-1}, A_t]$
  - $\Rightarrow$ The "impact" of these conditionings is quantified by $pl_a^{\Omega_t|\Omega_{t-1}}[B_{t-1}](A_t)$ (recall that "conflict = 1 - plausibility of everything")
  - $\Rightarrow$ Proposition: "Given a sequence of observations (states), an EMC can be written $pl(A_1)pl(A_2|A_1)\ldots pl(A_t|A_{t-1})$ where pl is a plausibility" [Interpretation of MLSP 09 results, from "evidential Viterbi proposition"]

Classification by EvHMM (P1)

(from $t = 1$ to T)

1. Prediction: Apply the q-total law given transitions (in q) and previous BBA on states (in m)
2. Fusion: Combine prediction and observation (in q)
3. Normalisation: Store the mass of conflict and normalise by one rule
4. Criterion of classification: Integrate conflict $\Rightarrow$ generalise the criterion in prob. HMM if conjunctions and normalisation are made by Dempster rule

$\Rightarrow$ The less the conflict the better is the correspondance between model and observations, recall that "conflict = 1 - plausibility of everything":

$$\lambda^* = \underset{n}{\text{argmax}} \ \frac{1}{T} \sum_{t=1}^{T} \log(pl_{\alpha}^{\Omega_t}(\Omega_t|\lambda_n))$$

Decision "global" - Viterbi (P2)

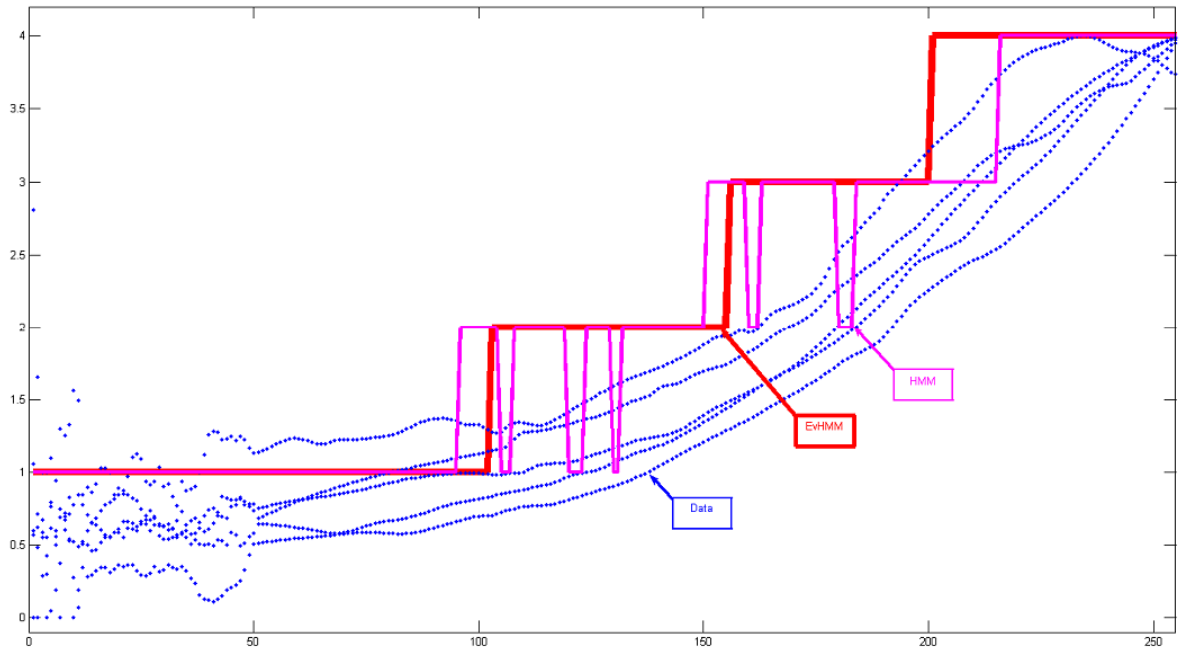
- Goal: determine the best sequence of states within noisy observations. Prob. version: maximum likelihood.
- A suboptimal evidential version: drawing benefits from belief functions [PhD]
- An optimal version [MLSP 09], for sequences of singletons
  - Given a state sequence $\{s_1, s_2, \ldots, s_T\}$, we can assess its plausibility
  - Condition the propagations by the singletons $\Rightarrow$ 2 focal sets for each conditioning
  - Propagate this information with the usual Viterbi algorithm with plausibility on singletons [MLSP 09]
  - Require the transition matrix, prior and observations as plausibilities

Parameter training (P3)

- Prior: not required
- Transitions [PhD, MLSP 09]:
  - Average the joint belief on states at t and $t-1$
  - Condition on states in Ω_{t-1} and marginalize on Ω_t to obtain the mass on subsets at t given singletons at $t-1$.
  - Store $|2^{\Omega_t}| \cdot |\Omega_{t-1}|$ elements, deduce conditionals on subsets by the disjunctive rule of combination (DRC) during propagations
  - The same procedure used to estimate coefficients in evolution models in the TCF [PhD]
- Models (emission): proposition in [PHM 11] (to appear)

Example of application

- Reference: *E. Ramasso and R. Gouriveau, Prognostics in Switching Systems: Evidential Markovian Classification of Real-Time Neuro-Fuzzy Predictions, In IEEE Prognostics and System Health Management Conference 2009*
- State recognition using EvHMM



1 Belief functions

2 Implicit Markov Models and Belief Functions

3 Explicit Markov Models and Belief Functions

4 EvHMM: Evidential HMM

5 Time-Sliced Temporal Evidential Networks

6 Continuous state filtering

7 Conclusions and perspectives

Evidential Graphical Models

- Graphical models: Powerful tool to represent variables of interest and their causal links
- VBS [see Shenoy papers]: Propagation tools to assess belief on some variables
- We choose the inference mechanisms in Evidential Graphical Models proposed in "Smets bible" [Smets 93] to propagate belief based on conditionals

Evidential Graphical Models + Time

- Proposition of a Time-Sliced Temporal Evidential Network (TSEN) [IEEE PHM 2011] for diagnosis and prognosis: "extends" the propagation in Evidential Graphical Models [Smets 93] with the temporal dimension
- EvHMM is a TSEN where some observations Y are measured and some hidden states X are present, Y is conditioned on X and we want to infer belief on X_t for some t.

Notes

- Reference: *[E. Pollard, M. Rombaut, and B. Pannetier, Bayesian networks vs. evidential networks. An application to convoy detection, in IPMU conference, 2010]*
  - Evidential network + time
  - Use the TCF to propagate belief between states at consecutive instants (estimation of parameters presented in TCSVT), discounting
  - No hidden variables
  - EvHMM inference can be applied easily, using general belief functions
- Reference: *[System prediction combining state estimation with an evidential influence diagram J. Schubert, P. Svenson, C. Martenson, 12th Int. Conf. on Information Fusion, Seattle, USA, 2009]*
  - Influence diagram is a general framework for decision-aiding problems (prob. method): observed+hidden+decision variables
  - Propose to work on the joint space between nodes in the diagram
  - Here decision does not play any role, only fusion+prediction is considered
  - How masses are generated?
  - Prediction-Fusion made as in TCF and EvHMM, no method to estimate the links between variables

1 Belief functions

2 Implicit Markov Models and Belief Functions

3 Explicit Markov Models and Belief Functions

4 EvHMM: Evidential HMM

5 Time-Sliced Temporal Evidential Networks

6 Continuous state filtering

7 Conclusions and perspectives

Underlying continuous states, decision on discrete states

- Reference: *[M. Rombaut. Decision in multi-obstacle matching process using the theory of belief. In AVCS'98, pages 63-68, 1998]*, then some improvements by D. Gruyer
- Reference: *[Y. Lemeret, E. Lefevre et D. Jolly, Improvement of an association algorithm for obstacle tracking, Information Fusion, vol. 9(2), pp. 234-245, 2008]*
- An association algorithm for object tracking based on belief functions
- At t some objects are known. At $t+1$, given some measurements, makes a decision: a new object appears, an object disappears, or an object had already been detected before
- Takes into account the imperfections due to each item of information and each sensor
- There is no prediction phase (requires an evolution model on data or on hypotheses)

Kalman filter - Linear systems

- Reference: *[B. Ristic and Ph. Smets. Kalman filters for tracking and classification and the Transferable Belief Model. In Int. Conf. on Information Fusion, 2004]*
- Joint Tracking and Classification (JTC) based on belief functions: using Kalman filter and continuous belief functions (`http://bfas.iutlan.univ-rennes1.fr`). Equations for filtering are similar, except:
  - Vacuous prior (first target positions unknown): the Kalman-based method is good at every iteration even the first ones
  - Use the q-LC (Least Committed) isopignistic for the evolution model noises in order to better model the target behavior
  - Imprecise observations: take into account lack of integrity in sensor reports or in evolution model predictions
  - When information are redundant, or when the model is not correct, or when data are corrupted: conflict may occur. Individual information may be discounted in order to improve the integrity of the filtering.
  - Efficient TBM-based classification phase: better description of expert rules *[Branko Ristic and Philippe Smets, Target Identification Using Belief Functions and Implication Rules, IEEE transactions on aerospace and electronic systems, vol. 41(3), pp. 1097-1103, 2005]*

Particle filter - Non linear systems

- Particle filter + GBT for classification: *[F. Caron, M. Davy, E. Duflos, P. Vanheeghe. Particle Filtering for Multisensor Data Fusion with Switching Observation Models. Application to Land Vehicle Positioning. IEEE Trans. on Signal Processing, vol. 55(6), 2007]*
- Note - Useful relation - expression of posterior plausibilities obtained from a Gaussian mixtures: *[F. Caron, B. Ristic, E. Duflos and P. Vanheeghe, Least committed basic belief density induced by a multivariate Gaussian: Formulation with applications, International Journal of Approximate Reasoning, 48(2), 2008]*
- Note - Interval-valued observations + data fusion + Kalman: *[Caron, F., Smets, P., Duflos, E. and Vanheeghe, P., Multisensor data fusion in the frame of the TBM on reals. Application to land vehicle positioning. 7th Int. Conf. on Information Fusion, Philadelphia, USA, 2005]*

Belief functions + Time + Intervals

- Imprecise state filtering with temporal beliefs on intervals: *[G. Nassreddine, F. Abdallah and T. Denoeux. State estimation using interval analysis and belief function theory: Application to dynamic vehicle localization. IEEE Trans. on Systems, Man and Cybernetics B, vol. 40, Issue 5, pages 1205-1218, 2010]*

Algorithm 2 Belief State Estimation (BSE) algorithm.

Input: $m^{\mathbf{x}}_k$, $m^{\mathbf{u}}_k$, $m^{\mathbf{z}}_{k+1}$, $p^{\mathbf{u}}$, $p^{\mathbf{z}}$, q
Output: $m^{\mathbf{x}}_{k+1}$, $[\widehat{\mathbf{x}}_{k+1}]$, $\widehat{\mathbf{x}}_{k+1}$

Read the input $\mathbf{u}_k$ and its error. Deduce $m^{\mathbf{u}}_k$ with $p^{\mathbf{u}}$ focal elements
Read the output $\mathbf{z}_{k+1}$ and its error. Deduce $m^{\mathbf{z}}_{k+1}$ with $p^{\mathbf{z}}$ focal elements
$r \leftarrow 0$
for $i = 1$ to $p^{\mathbf{x}}_k$ **do**
  for $j = 1$ to $p^{\mathbf{u}}$ **do**
    for $\ell = 1$ to $p^{\mathbf{z}}$ **do**
      $r \leftarrow r + 1$
      $[\mathbf{x}^r_{k+1}] \leftarrow \mathcal{C}\left(\mathbf{f}, \mathbf{g}, [\mathbf{x}^i_k], [\mathbf{u}^j_k], [\mathbf{z}^\ell_{k+1}]\right)$.
      $m^{\mathbf{x}}_{k+1}([\mathbf{x}^r_{k+1}]) \leftarrow m^{\mathbf{x}}_k([\mathbf{x}^i_k]) \cdot m^{\mathbf{u}}_k([\mathbf{u}^j_k]) \cdot m^{\mathbf{z}}_{k+1}([\mathbf{z}^\ell_{k+1}])$
    end for
  end for
end for
Normalize $m^{\mathbf{x}}_{k+1}$
Summarize $m^{\mathbf{x}}_{k+1}$ to keep at most q focal elements.
% Computation of interval and point estimates %
$[\widehat{\mathbf{x}}_{k+1}] \leftarrow [\mathbb{E}](m^{\mathbf{x}}_{k+1})$
$\widehat{\mathbf{x}}_{k+1} \leftarrow \mathbb{E}(m^{\mathbf{x}}_{k+1})$

Particle filter - An evidential filter (IJAR 2009)

- Reference: *R. Muñoz-Salinas, R. Medina-Carnicer, F.J. Madrid-Cuevas, A. Carmona-Poyato, Multi-camera people tracking using evidential filters, International Journal of Approximate Reasoning, Volume 50, Issue 5, May 2009, Pages 732-749*
- Multiple Evidential Particle Filtering (MEPF) algorithm
- The dynamical system might be comprised of a set of n subsystems, each with its own dynamics
- Underlying idea: the true target state is estimated from a set of possible states (called particles), using a binary frame for each particle $\Rightarrow$ MEPF does not evaluate the likelihood of particles (Bayesian sense) but their degree of evidence
- The algorithm is specifically conceived to simultaneously deal with multiple sensors, the evidence of particles is evaluated using all the available sensors and finally fused

Conclusions

- Presentation of tools to manage temporal belief functions: continuous vs discrete states
- Some recent work on temporal belief functions + Markov for the classification of time-series
- Presentation of EvHMM
- Applications to image and signal analysis
- Belief functions for dynamical systems enable (in addition to many other things presented in the school ☺)
  - Partial labelling of imprecise and uncertain data
  - Transition modelling (between states)
  - Conflict management
  - Finer description of states evolution
- Complexity: bigger than prob., reduced using efficient programming and simplifications

Perspectives

- Theoretical: mainly on parameters training, already hard in static case, imagine in the temporal case ☹: Proposition of a method to IEEE PHM 2011 [PhD of Lisa Serir at FEMTO-ST]
  - On exploiting *[D. Dubois and T. Denoeux, Statistical Inference with Belief Functions and Possibility Measures: a discussion of basic assumptions, SMPS 2010]*
  - On exploiting *[Thierry Denoeux, Maximum Likelihood from Evidential Data: An Extension of the EM Algorithm, SMPS 2010]* (ref?)
  - Modelling using Valuation-Based Systems [cf. Shenoy]
- Applicative: e.g. EvHMM or TCF (model change) for image and audio processing, and for prediction (current work at FEMTO-ST)

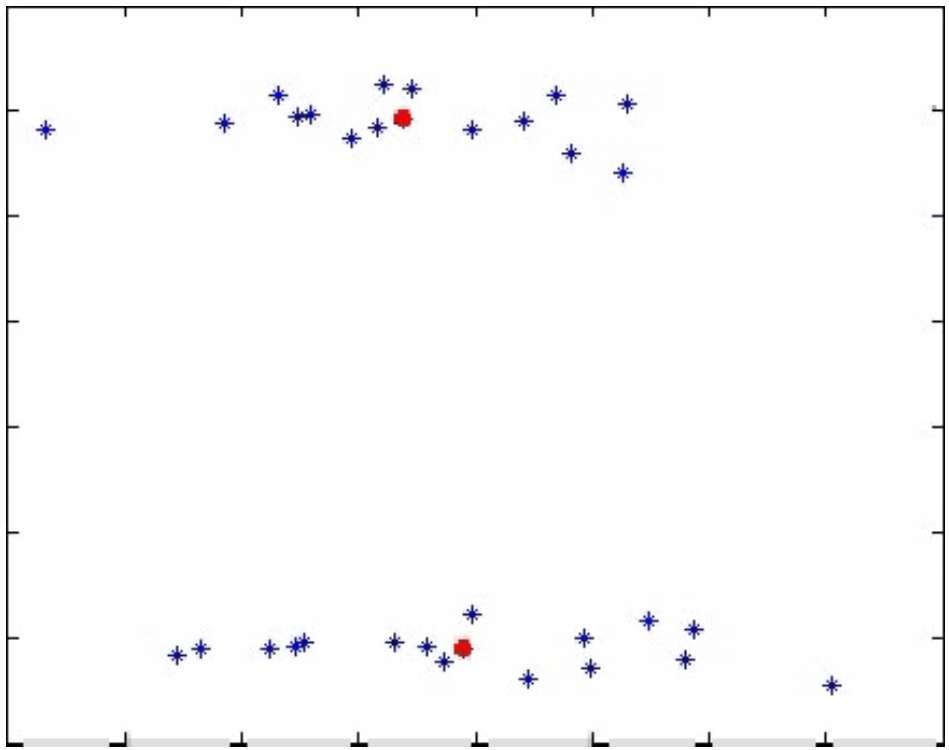
Perspectives

(To Do list ☺)

- Rewrite in english the first chapter of my PhD: *E. Ramasso, M. Rombaut, D. Pellerin, Modèle des Croyances Transférables : Représentation des connaissances, Fusion d'informations, Décision – Transferable Belief Model: Knowledge representation, Information fusion, Decision, Technical Report hal-00283144, 29 May 2008*
- Complete the references of this presentation, write a report (under work)
- Update the wiki on belief functions (part on Temporal Belief Functions on my webpage with pres.+codes+ref. . .)

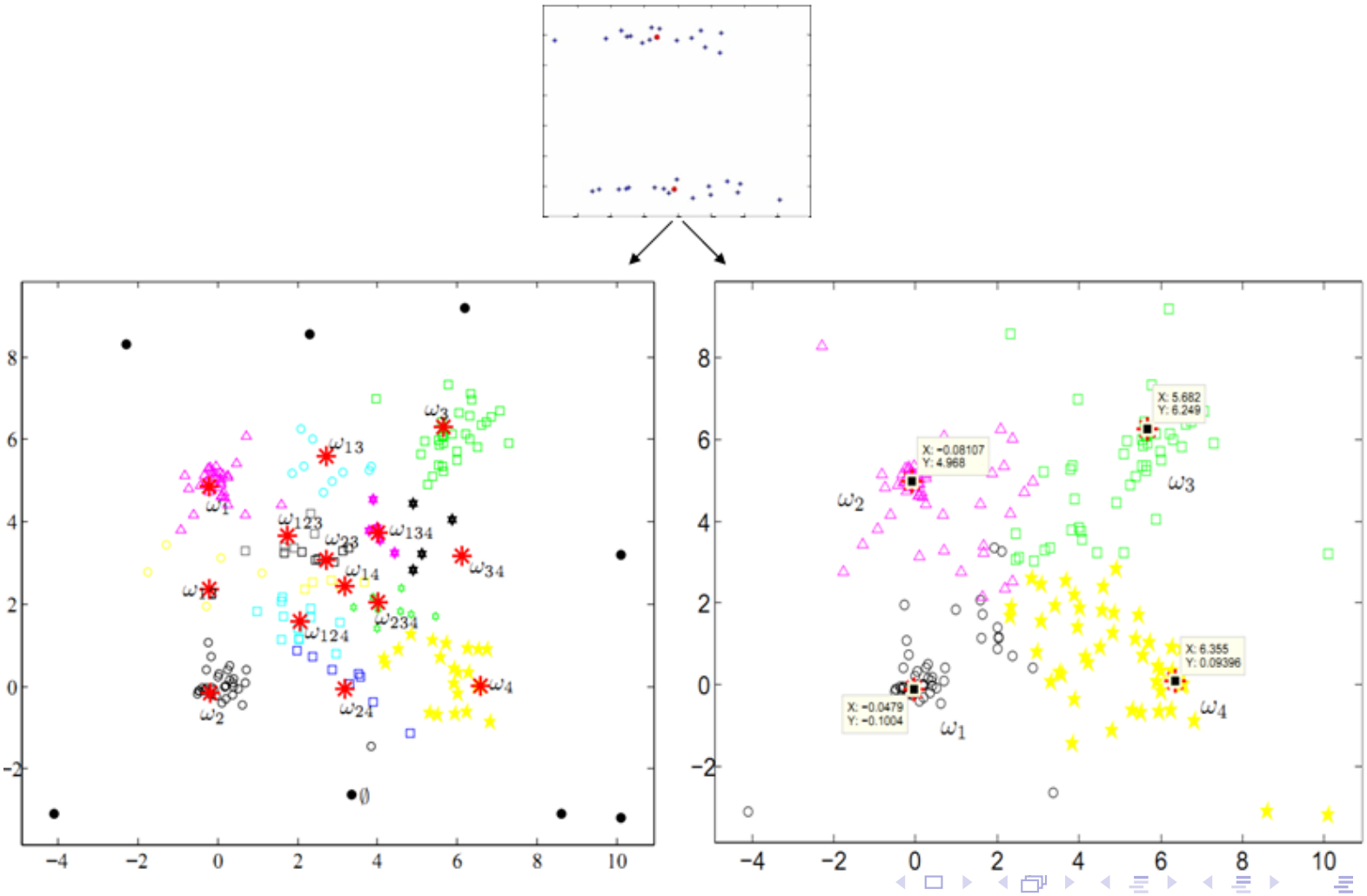
Perspectives

- Problem: online building of a credal partition $\Rightarrow$ online estimation of models generating belief masses
- A new algorithm for building a credal partition [ECSQARU 2011]: first algorithm based on belief functions for *online data clustering*
- Based on ideas developed in fuzzy clustering and in ECM



Perspectives

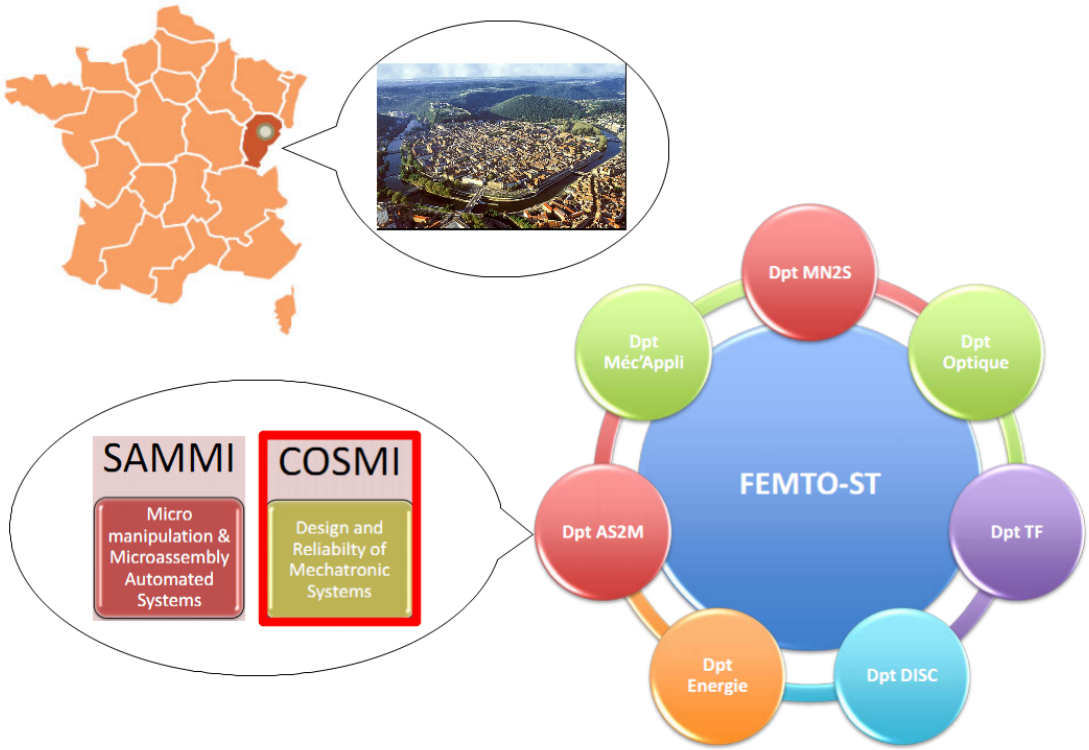
- A new algorithm for building a credal partition [ECSQARU 2011]: first algorithm based on belief functions for *online data clustering*



Thanks for your attention

Acknowledgement to the organization committee for this school

Where are Besançon and FEMTO-ST?



Papers, references, codes (to come), web page on temporal belief functions

`http://www.femto-st.fr/~emmanuel.ramasso/`