A Hierarchical Pattern Matching Procedure for Signal Abstraction

A. Otero[1*], P. Félix[1], S. Fraga[1], S. Barro[1] and F. Palacios[2]

[1] Dpto. de Electrónica e Computación, Universidade de Santiago de Compostela. Santiago de Compostela. 15782, Spain. (abraham@dec.usc.es)
[2] Hospital Universitario de Getafe, Madrid. Spain

Abstract. The Multivariable Fuzzy Temporal Profile model enables human experts to project their knowledge of a signal pattern, described over a set of parameters, into a computable description. Here it is endowed with a hierarchical capability through the definition of algorithms that recognize a set of findings over a signal, and aggregate them into a set of abstraction levels. We also present a heuristics that takes advantage of the continuity properties of real signals to allow the matching process to meet real time requirements, even over high frequency signals.

1 Introduction

This work has its origin in the thesis that processes of abstraction over perception are organized into a hierarchical structure. Thus, the recognition and interpretation of observable events is carried out by successively aggregating pieces of information -findings- that are extracted from the data acquired. This operation results in a decrease in the volume of data handled, and an increase in their semantics. With this idea in mind, we have developed a hierarchical matching procedure for the Multivariable Fuzzy Temporal Profile (MFTP) model.

This model, presented in [5], makes use of the fuzzy set theory [6] to represent and handle the imprecision and uncertainty that are characteristic of human knowledge, and to allow the direct acquisition of knowledge on patterns from human experts. Constraint Satisfaction Problems (CSP) [1] are used to supply a structural description of signal patterns that enables the hierarchy of abstraction that is present in pattern recognition to be captured, making it possible to give detailed explanations of the matching results to the expert, as well as to describe sets of patterns that have common findings.

This paper is organized as follows: we now go on to explain certain fuzzy concepts on which the MFTP model is based, in order to then present in Section 3 a real example from the clinical domain that we use to introduce the theoretical concepts of the model, which are dealt with in Section 4. Section 5 explains how the hierarchical matching is carried out, and a heuristics that exploits the hierarchical nature of the model to significantly improve matching efficiency is presented in Section 6. Finally, we present the conclusions of this work.

* The authors wish to acknowledge the support from the Spanish CICyT and the Xunta de Galicia under the projects TIC2003-09400-C04-03 and PGIDIT04SIN206003PR, respectively.

2 Fuzzy concepts

We will consider time as being projected onto a one-dimensional discrete axis $\tau = \{t_0, t_1, ..., t_i, ...\}$, where t_i represents a *precise* instant and t_0 represents the temporal origin. We consider that for every $i \in \mathbb{N}$, $t_{i+1} - t_i = \Delta t$, where the constant Δt is the minimum step of the temporal axis.

Given as discourse universe the set of real numbers $\mathbb{R}$, a **fuzzy number** A is a normal ($\exists\ v \in \mathbb{R},\ \mu^A(v) = 1$) and convex ($\forall v, v', v'' \in \mathbb{R},\ v' \in [v, v'']$, $\mu^A(v') \geq \min\ \{\mu^A(v), \mu^A(v'')\}$) fuzzy subset of $\mathbb{R}$. We obtain a fuzzy number A from a flexible constraint given by a possibility distribution π^A, which defines a mapping from $\mathbb{R}$ to the real interval $[0, 1]$. Given a precise number $v \in \mathbb{R}$, $\pi^A(v) \in [0, 1]$ represents the possibility of A being precisely v. By means of π^A we define a fuzzy subset A of $\mathbb{R}$, which contains the possible values of A.

We introduce the concept of **fuzzy increment** with the aim of representing quantities, such as the difference between two numbers, which may be fuzzy or not. Following Zadeh's extension principle [6], the fuzzy increment between a pair of fuzzy numbers A and B is given by D such $\pi^D(i) = \max_{i=t-s} \min\{\pi^A(t), \pi^B(s)\}$.

3 Clinical example

We use an example from the clinical domain to introduce the MFTP model: the recognition of the sinusal P wave in a multichannel electrocardiogram (ECG) recording. The P wave is studied as a sign of auricular electrical activity, and is of great use in the diagnosis of arrhythmias and auriculo-ventricular conduction abnormalities. Nevertheless, its detection is still problematic in automatic ECG monitoring and currently there is no commercial bedside monitor that provides this information. The difficulty in their detection lies in the low relative voltage of these waves, the frequency with which small artifacts simulate them, as well as in the morphological variations produced even by breathing movements.

The events of the auricle-ventricular complex that enable or help in the diagnosis of arrhythmias and conduction abnormalities are the onset of the P wave and the onset of the QRS [3]. Among the multiple ECG leads of the monitoring, II and aV_R are chosen for their suitability (see Fig. 1). Over Lead II the pattern is defined as a gradual increase from a basal ECG value, corresponding with the onset of the P wave, and a very sharp decrease, corresponding with the onset of the QRS. Both events must be separated by approximately 120-200 ms. Over aV_R the P wave must be negative, thus there must be a gentle decrease in aV_R, practically simultaneously with the gentle increase over II. Non-compliance with these criteria precludes a sinusal origin for the auricular activity.

4 The MFTP model

The MFTP model enables the identification of a pattern $\mathcal{M}$ of special significance over the temporal evolution of a set of parameters $\mathcal{P} = \{P^1, ..., P^s\}$, where each parameter P^j is obtained by means of an acquisition and sampling process: $P^j =$

$\{(v^j_{[1]}, t^j_{[1]}), ..., (v^j_{[k]}, t^j_{[k]}), ...\}$. This pattern is described by a human expert and consists of a set of findings and relations between them.

The MFTP model is an extension of the Fuzzy Temporal Profile model [2], which allows a finding to be represented as a morphology described over a single physical parameter. The fact of being able to relate the occurrence of different findings among parameters is of great importance, as often the appearance of a finding over a single parameter, which on its own may not be a major determinant, may well be of interest if it appears to be related with other findings on other parameters which also do not seem to be definitive when considered in isolation. In our clinical example, the appearance of the pattern described over Lead II does not imply a sinusal beat if in aV_R P wave is positive.

Furthermore, the MFTP model is able to explicitly represent the hierarchy of abstraction levels that the expert employs to reason about the system. This enables detailed explanations of the reasons behind the occurrence or non-occurrence of a pattern to be given. Going back to the example, the independent representation of the onset of the P wave and of the QRS complex makes it possible to reason over whether the non-occurrence of the pattern was caused by an abnormal origin of the beat (P wave negative in II), or whether there has been a first-degree blockage (excessive distance between the P wave and the QRS complex) or a second-degree one (the P wave is not followed by a QRS complex).

The MFTP model is based on the CSP formalism and on the fuzzy set theory. An MFTP allows a pattern to be represented by means of a network of fuzzy constraints between a set of significant points.

Definition 1. *We define* ***significant point*** *on a physical parameter* P^j, X^j_i, *as the pair formed by a variable from the domain* V^j_i *and a temporal variable* T^j_i. *A significant point* $X^j_i = < V^j_i, T^j_i >$ *represents an unknown value* V^j_i *for* P^j *at an unknown temporal instant* T^j_i. *In the absence of constraints,* V^j_i *and* T^j_i *may take any precise value* $v^j_{[k]}$ *and* $t^j_{[k]}$, *respectively, where* $(v^j_{[k]}, t^j_{[k]}) \in \mathcal{P}^j$.

By $\mathcal{A}^j_i$ we denote the assignment of precise values from the evolution P^j to the variables of X^j_i; i.e., $\mathcal{A}^j_i = (v^j_{[k]}, t^j_{[k]})$. We define a general fuzzy constraint between a set of significant points, providing a computable support for soft descriptions of the form of a signal.

Definition 2. *A* ***fuzzy constraint*** R *between a set of significant points* $\{X^{j_1}_{i_1}, X^{j_2}_{i_2}, ..., X^{j_g}_{i_g}\}$ *is defined by means of a fuzzy relation* $C = C(X^{j_1}_{i_1}, X^{j_2}_{i_2}, ..., X^{j_g}_{i_g})$. C *is represented by means of a membership function* μ^C, *which associates a degree of satisfaction of* R *to each assignment of precise values to* $X^{j_1}_{i_1}, X^{j_2}_{i_2}, ..., X^{j_g}_{i_g}$.

In principle, nothing restricts the form of the constraints that make up a MFTP. However, experience has shown that a set of constraints limiting the fuzzy increment, fuzzy temporal extension and fuzzy slope between a pair of significant points is able to capture a good number of nuances. Thus we define a constraint $L^{j_1 j_2}_{i_1 i_2}$ between two significant points $X^{j_1}_{i_1}$ y $X^{j_2}_{i_2}$ by means of a normal, convex

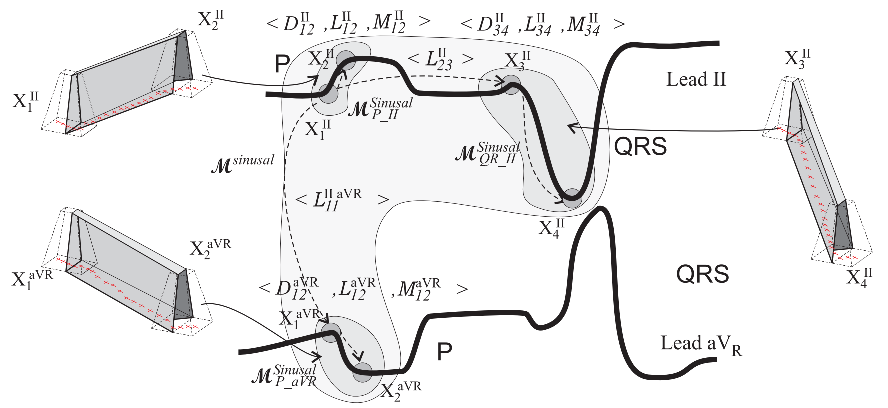

Fig. 1. Graph of the sinusal rhythm MFTP along with the graphical representation of the fuzzy trajectories that the MFTP defines for each finding

possibility distribution $\mu^{L_{i_1 i_2}^{j_1 j_2}}(X_{i_1}^{j_1}, X_{i_2}^{j_2}) = \pi^{L_{i_1 i_2}^{j_1 j_2}}(h)$, $h \in \tau$, which represents the possibility of the fuzzy temporal extension between $X_{i_1}^{j_1}$ and $X_{i_2}^{j_2}$ being h. The assignations $T_{i_1}^{j_1} = t_{i_1}^{j_1}$ and $T_{i_2}^{j_2} = t_{i_2}^{j_2}$ are possible if $\pi^{L_{i_1 i_2}^{j_1 j_2}}(t_{i_2}^{j_2} - t_{i_1}^{j_1}) > 0$. In Fig. 1 L_{23}^{II} models the linguistic description "*approximately of 120-200 ms*".

A constraint $D_{i_1 i_2}^{j_1 j_2}$ between a pair of significant points $X_{i_1}^{j_1}$ y $X_{i_2}^{j_2}$ is defined by means of a normal and convex possibility distribution $\mu^{D_{i_1 i_2}^{j_1 j_2}}(X_{i_1}^{j_1}, X_{i_2}^{j_2}) = \pi^{D_{i_1 i_2}^{j_1 j_2}}(d)$, $d \in \mathbb{R}$, which represents the possibility of the fuzzy increase between $X_{i_1}^{j_1}$ and $X_{i_2}^{j_2}$ being d. The assignations $V_{i_1}^{j_1} = v_{i_1}^{j_1}$ and $V_{i_2}^{j_2} = v_{i_2}^{j_2}$ are possible if $\pi^{D_{i_1 i_2}^{j_1 j_2}}(v_{i_2}^{j_2} - v_{i_1}^{j_1}) > 0$. In Fig. 1 D_{12}^{II} models the description "*gentle increase*".

A constraint $M_{i_1 i_2}^{j}$ between a pair of significant points $X_{i_1}^{j}$ and $X_{i_2}^{j}$, defined over the same parameter P^j, is defined by means of a normal and convex possibility distribution $\mu^{M_{i_1 i_2}^{j}}(X_{i_1}^{j}, X_{i_2}^{j}) = \pi^{M_{i_1 i_2}^{j}}(m)$, $m \in \mathbb{R}$, which represents the possibility of the fuzzy slope between $X_{i_1}^{j}$ and $X_{i_2}^{j}$ being m. The assignations $V_{i_1}^{j} = v_{i_1}^{j}$, $V_{i_2}^{j} = v_{i_2}^{j}$, $T_{i_1}^{j} = t_{i_1}^{j}$ and $T_{i_2}^{j} = t_{i_2}^{j}$ are possible if $\pi^{M_{i_1 i_2}^{j}}((v_{i_2}^{j} - v_{i_1}^{j})/(t_{i_2}^{j} - t_{i_1}^{j})) > 0$. In Fig. 1 M_{34}^{II} models the description "*...sharp* decrease", where "*sharp*" is modelled by means of a high slope value.

Definition 3. *We define a* ***Multivariable Fuzzy Temporal Profile*** *(MFTP)* $\mathcal{M} =< \mathcal{W}^{\mathcal{M}}, \mathcal{X}^{\mathcal{M}}, \mathcal{R}^{\mathcal{M}} >$ *as a finite set of MFTPs* $\mathcal{W}^{\mathcal{M}} = \{\mathcal{M}_1^{\mathcal{M}}, ..., \mathcal{M}_s^{\mathcal{M}}\}$, *a finite set of significant points* $\mathcal{X}^{\mathcal{M}} = \{X_{i_1}^{j_1}, X_{i_2}^{j_2}, ..., X_{i_g}^{j_g}\}$ *and a finite set of constraints* $\mathcal{R}^{\mathcal{M}} = \{R_1, ..., R_f\}$ *amongst the points of* $\mathcal{W}^{\mathcal{M}}$ *and* $\mathcal{X}^{\mathcal{M}}$.

The constraints $R_i \in \mathcal{R}^{\mathcal{M}}$ can be defined between significant points belonging to $\mathcal{X}^{\mathcal{M}}$, between significant points belonging to the set of subMFTPs $\mathcal{W}^{\mathcal{M}}$ or between both types of significant points. The recursive structure of the MFTP model is based in the way that humans define patterns; a complex pattern is often made up of a set of findings and a set of relations between them. Each of the findings of the pattern may also be a pattern, and may comprise a set of findings and relations between them, and so on, successively.

The pattern of the sinusal P wave is made up of three findings for which certain temporal relations between them must be satisfied: $\mathcal{M}^{Sinusal} =< \{\mathcal{M}^{Sinusal}_{P\_II}, \mathcal{M}^{Sinusal}_{QR\_II}, \mathcal{M}^{Sinusal}_{P\_aV_R}\}, \varnothing, \{L^{II}_{23}, L^{II\ aV_R}_{11}\} >$. Each of these findings is in turn an MFTP that is defined over its corresponding lead; thus $\mathcal{M}^{Sinusal}_{P\_II} =< \varnothing, \{X^{II}_1, X^{II}_2\}, \{L^{II}_{12}, D^{II}_{12}, M^{II}_{12}\} >$, $\mathcal{M}^{Sinusal}_{QR\_II} =< \varnothing, \{X^{II}_3, X^{II}_4\}, \{L^{II}_{34}, D^{II}_{34}, M^{II}_{34}\} >$ y $\mathcal{M}^{Sinusal}_{P\_aV_R} =< \varnothing, \{X^{aV_R}_1, X^{aV_R}_2\}, \{L^{aV_R}_{12}, D^{aV_R}_{12}, M^{aV_R}_{12}\} >$.

An MFTP can be represented by a graph in which nodes correspond to significant points, and arcs correspond to constraints (see Fig. 1). The MFTP model also enables us to restrict the evolution of a parameter P^j between each pair of significant points $X^j_{i_1}$ and $X^j_{i_2}$ (see Fig. 1) by means of a membership function $\mu^{S^j_{i_1 i_2}}(\mathcal{A}^j_{i_1}, \mathcal{A}^j_{i_2})$ which defines a fuzzy course within which the temporal evolution of the parameter must remain in order to satisfy the constraint [2].

5 Matching

The ultimate aim of the MFTP model is to identify a pattern $\mathcal{M}$ over a set of parameters $\mathcal{P}$, which represent the temporal evolution of a physical system S, automatically generating information organized in a hierarchy of levels of abstraction. Given that the MFTP model is based on the formalism of constraint networks, comparing an MFTP with $\mathcal{P}$ is formally equivalent to solving a CSP [1], where the domains of the variables are determined by $\mathcal{P}$.

Definition 4. *We define a* ***solution*** *of $\mathcal{M}$ as a set of assignations $\mathcal{A} = \{A^1_1, ..., A^1_{n^1}, ..., A^s_1, ..., A^s_{n^s}\}$ to all the significant points of the pattern that satisfies the set of constraints that make up $\mathcal{M}$, with a degree greater than zero. The degree of satisfaction of a solution $\mathcal{A}$ is given by:*

$$\pi^{\mathcal{M}}(\mathcal{A}) = min\{\min_{\mathcal{M}^{\mathcal{M}}_h \in \mathcal{W}^{\mathcal{M}}} \{\pi^{\mathcal{M}^{\mathcal{M}}_h}(\mathcal{A}^{\mathcal{M}^{\mathcal{M}}_h})\}, \min_{R_k \in \mathcal{R}^{\mathcal{M}}} \{\pi^{R_k}(\mathcal{A}^{R_k})\}\} \tag{1}$$

Where $\mathcal{A}^{\mathcal{M}^{\mathcal{M}}_h}$ and $\mathcal{A}^{R_k}$ are the projection of $\mathcal{A}$ over the set of significant points involved in $\mathcal{M}^{\mathcal{M}}_h$ and in R_k, respectively. π^{R_k} is the degree of satisfaction of $R_k \in \mathcal{R}^{\mathcal{M}}$ and $\pi^{\mathcal{M}^{\mathcal{M}}_h}$ is the degree of satisfaction of $\mathcal{M}^{\mathcal{M}}_h \in \mathcal{W}^{\mathcal{M}}$. $\pi^{\mathcal{M}}(\mathcal{A})$ represents the degree of similarity between a fragment of the evolution $\mathcal{P}$ with the description represented by $\mathcal{M}$.

Imitating human experts, we divide the matching into as many stages as the number of levels of abstraction given by the composition of findings. In the lowest level is the temporal evolution of the system $\mathcal{P}$, and in the highest the global pattern $\mathcal{M}$. Figure 2a shows a recursive procedure which runs through the hierarchy of the pattern until it reaches the lowest level MFTPs; it resolves them and then goes up the next level of abstraction and resolves the MFTPs there, by assembling solutions found for the MFTPs in lower levels. The procedure comes to an end when solutions for the global pattern are obtained. For example, in order to match the sinusal beat pattern, firstly we search for solutions for the three subMFTPs: $\mathcal{M}^{Sinusal}_{P\_II}, \mathcal{M}^{Sinusal}_{QR\_II}$, $\mathcal{M}^{Sinusal}_{P\_aV_R}$ and we then attempt to assemble the solutions encountered in order to obtain solutions for $\mathcal{M}^{Sinusal}$.

The solutions to the MFTPs on each level are constructed incrementally, by exploring a search tree. We suppose that the global pattern $\mathcal{M}$ has f nesting levels, where $\mathcal{M}^l_h$ is the subMFTP h in level l. The matching starts with the lowest-level MFTPs, those which are defined directly over the system's signal: $\mathcal{M}^o_h = \{\varnothing, X^{\mathcal{M}^o_h}, \mathcal{R}^{\mathcal{M}^o_h}\}$. We start from an empty tuple of assignations, and we extend it with an assignation to a new significant point, such that the degree of satisfaction of the extended tuple is greater than or equal to a limit $c_{\inf}$, below which the degree of satisfaction of a solution is considered to be unacceptable. The degree of satisfaction of the tuple $\mathcal{A}_{[r+1]}$ of assignations to $r+1$ significant points of $X^{\mathcal{M}^o_h}$, $\mathcal{A}_{[r+1]} = \{A^{j_1}_{i_1}, ..., A^{j_r}_{i_r}, A^{j_{r+1}}_{i_{r+1}}\}$, is calculated on basis of the degree of satisfaction of $\mathcal{A}_{[r]}$ by means of:

$$\pi^{\mathcal{M}^o_h}(\mathcal{A}_{[r+1]}) = min\{\pi^{\mathcal{M}^o_h}(\mathcal{A}_{[r]}), \min_{R\in\mathcal{R}^{[\mathcal{A}_{[r]}, A^{j_{r+1}}_{i_{r+1}}]}}\{\pi^R(\mathcal{A}_{[r+1]})\}\} \qquad (2)$$

where $\mathcal{R}^{[\mathcal{A}_{[r]}, A^{j_{r+1}}_{i_{r+1}}]}$ is the set of those constraints that are defined over points to which $\mathcal{A}_{[r+1]}$ assigns value, and which involve $X^{j_{r+1}}_{i_{r+1}}$. For example, for matching $\mathcal{M}^{Sinusal}_{P\_II}$ we start by carrying out an assignation $\mathcal{A}^{II}_1 = (v^{II}_1, t^{II}_1)$ to X^{II}_1; we then go on to search for an assignation $\mathcal{A}^{II}_2 = (v^{II}_2, t^{II}_2)$ that satisfies the set of constraints $\{L^{II}_{12}, D^{II}_{12}, M^{II}_{12}\}$. If no such assignation is found, we backtrack to the last significant point to which a value has been assigned, and we attempt to carry out another assignation. The same procedure is followed for the recognition of $\mathcal{M}^{Sinusal}_{QR\_II}$ and $\mathcal{M}^{Sinusal}_{P\_aV_R}$. We then search for solutions to those MFTPs for which all subMFTPs have been resolved; i.e. those that are in nesting level 1. For these we start from an empty tuple of assignations, and we initially extend it with solutions that have previously been found for their subMFTPs. On the basis of the degree of satisfaction of the initial tuple, $\mathcal{A}_{[r]}$, and of the solution of the lower level subMFTP that has been assembled, $\mathcal{A}^{\mathcal{M}^o_h}$, we obtain the degree of satisfaction for the extended tuple $\mathcal{A}_{[s]}$, $\mathcal{A}_{[s]} = \mathcal{A}_{[r]} \cup \mathcal{A}^{\mathcal{M}^o_h}$, by means of:

$$\pi^{\mathcal{M}^1_k}(\mathcal{A}_{[s]}) = min\{\pi^{\mathcal{M}^1_k}(\mathcal{A}_{[r]}), \pi^{\mathcal{M}^o_h}(\mathcal{A}^{\mathcal{M}^o_h}), \min_{R\in\mathcal{R}^{[\mathcal{A}_{[r]}, \mathcal{A}^{\mathcal{M}^o_h}]}}\{\pi^R(\mathcal{A}^R_{[s]})\}\} \qquad (3)$$

where $\mathcal{M}^1_k$ is the MFTP from nesting level 1 that we are resolving and $\mathcal{R}^{[\mathcal{A}_{[r]}, \mathcal{A}^{\mathcal{M}^o_h}]}$ is the set of constraints that only involve the points to which $\mathcal{A}_{[s]}$ assigns value, and which involve at least one point to which $\mathcal{A}_{[r]}$ assigns value, and at least one other to which $\mathcal{A}^{\mathcal{M}^o_h}$ assigns value; i.e. the set of constraints that are defined between points of the assembled subMFTP and those points that already have a value assigned to them. Thus, if $\mathcal{M}^1_k$ has points that are not linked to one of the

```
procedure MATCH(M);
 if(M contains MFTPs) then
  for each (M^M ∈ M)
   MATCH(M^M);
 end if;
 BUILD_SOL(M,0);
end procedure;
```

(a)

```
procedure BUILD_SOL(M,r);
 if(r = n) then H^M ←< A_[r], π^M(A_[r]) >;
 if(r < m) then
  L_r= A^{M_h^M} ∈H^{M_h^M} /π^M(A_[r]∪A^{M_h^M}) > c_inf;
  while(L_l ≠ ∅ and c_inf < 1) do
   take and delete A^{M_h^M} from L_l;
   A^[d]= A_[r]∪A^{M_h^M};
   BUILD_SOL(M,r + n^{M_h^M})
  end while;
 end if;
 else
  L_r = A_i^j ∈ P/π^M(A_[r] ∪ A_i^j) > c_inf;
  while(L_r ≠ ∅ and c_inf < 1) do
   take and delete A_i^j from L_r;
   A_[r+1]= A_[r]∪A_i^j;
   BUILD_SOL(M,r + 1);
  end while;
 end else;
end procedure;
```

(b)

Fig. 2. (a) Procedure that makes it possible to solve a MFTP hierarchically; (b) procedure that assembles the solutions of the MFTPs.

subMFTPs, we continue to extend the tuple of assignations, in a similar manner to 2. Once all level 1 MFTPs have been completed, we go up to the next level and resolve its MFTPs by assembling solutions for its subMFTPs and assignations to points by means of expressions equivalent to 3 and 2, respectively, but which are defined between the nesting levels l and $l-1$. The process finishes when the global pattern is resolved. Throughout the entire search process we prune those branches in the search space that give rise to a solution with a degree of satisfaction lower than c_{inf}.

Figure 2b shows a procedure that searches for solutions for an MFTP by extending a locally valid tuple of assignations. The list L_r stores the solutions of the next subMFTP, or assignations to the next significant point, whose degree of local consistency with respect to the tuple that has been assembled up to this point is higher than c_{inf}; the list $\mathcal{H}^{\mathcal{M}}$ stores solutions found for the MFTP $\mathcal{M}$ and its degree of satisfaction; n is the number of significant points of $\mathcal{M}$; $n^{M_h^M}$ is the number of significant points of $\mathcal{M}_h^{\mathcal{M}}$ and $m = \Sigma_{h=1}^{m} n^{M_h^M}$. After the execution of the MATCH procedure, the solutions found for $\mathcal{M}$ and for each $\mathcal{M}_h^{\mathcal{M}}$ remain stored in $\mathcal{H}^{\mathcal{M}}$ and $\mathcal{H}^{\mathcal{M}_h^{\mathcal{M}}}$, respectively.

In the example of the P wave, in order to assemble solutions for $\mathcal{M}^{Sinusal}$, we start from an empty tuple of solutions, and we add to it a solution $\mathcal{A}_{P\_II}^{Sinusal}$. We then search for a solution $\mathcal{A}_{QR\_II}^{Sinusal}$ that satisfies the constraint L_{23}^{II}, and with this we construct a tuple of assignations, whose degree of consistency is calculated on the basis of 3. Finally, we search for a solution $\mathcal{A}_{P\_aV_R}^{Sinusal}$ that satisfies the constraint $L_{11}^{II\ aV_R}$, having then obtained a solution for $\mathcal{M}^{Sinusal}$. If, during the search, no solution is found for the following subMFTP that is compatible with the solutions that have been assembled up to that point, we backtrack to the previous subMFTP, and we attempt to assemble a different solution.

5.1 A heuristics for increasing matching efficiency

In order to guarantee the completeness of the search, we have to find all the possible solutions for the global pattern, regardless of their degree of satisfaction. This is due to the fact that in CSPs local solutions, even the optimal ones, do not necessarily form part of a global solution. This may significantly diminish the efficiency of the hierarchical matching.

In order to avoid searching for an excessive amount of solutions we have developed a heuristics that, by making use of the continuity properties of real signals, aims to obtain a set of solutions that is as representative as possible of each occurrence of a finding. The intuitive idea behind the heuristics is to attempt to obtain a "sampling" of the occurrence. In order to do so, we search for solutions within a temporal window whose length L_W must be greater than or the same as the maximum temporal extension of the MFTP, to guarantee that no solutions are lost. Thus, for example, in order to match $\mathcal{M}_{P\_II}^{Sinusal}$ the width of the temporal window must be the same or greater than the maximum value allowed by L_{12}^{II}. Each new solution must be better than the best solution up to that point, and from the set of solutions that are found, we only store those that differ more than d_{inf} from the ones that have already been stored. We define the distance between solutions as the Euclidian distance of the vectors formed by the temporal assignations to the significant points: $d(\mathcal{A}^1, \mathcal{A}^2) = \sum_{A_i^{1j} \in \mathcal{A}^1, A_i^{2j} \in \mathcal{A}^2} (t_i^{1j} - t_i^{2j})^2$, where $\mathcal{A}^1$ and $\mathcal{A}^2$ are two solutions of $\mathcal{M}$. When the data from this window have been processed, the window is shifted by a constant interval ΔL_W, and the process is repeated. In this way for each occurrence of a finding we obtain a set of solutions that are scattered around its temporal evolution. Searching in every window for better solutions than those already found considerably shortens the search space, as it avoids spreading out all the branches that would lead to a poorer solution than the best of those already found. Storing only those solutions that are significantly different limits the search space in the following stages of the matching, at the same time as it supplies reasonable guarantees of the stored solutions being representative of the temporal evolution of the system.

ΔL_W and d_{inf} act as algorithm control parameters: if they have high values less time will be needed to perform the matching, but less solutions will be generated and stored, and there will be a higher probability of non-optimal solutions or of losing occurrences of the pattern. Using lower values results in more time being required for the search, but is less probable that solutions will be lost or that non-optimal ones will be found. The adjustment of these parameters depends on the system's dynamics: for systems that evolve slowly we use higher values, while in rapidly-evolving systems lower values are used.

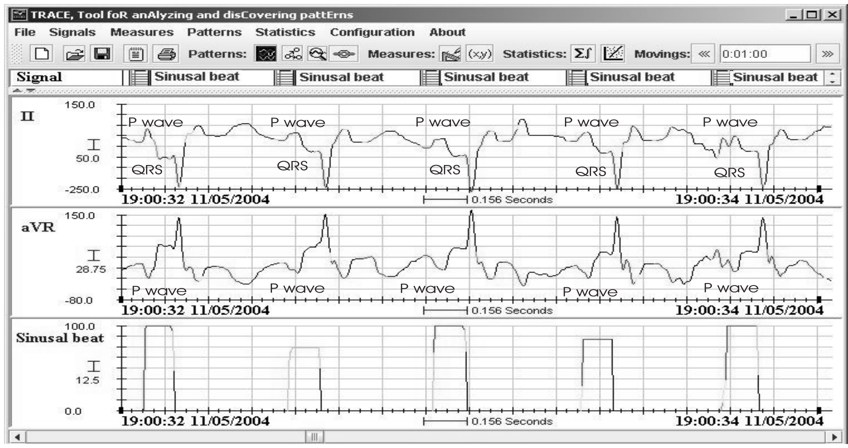

Fig. 3. Detection of the sinusal rhythm on the reading used in the validation

The most precise matching is obtained when $\Delta L_W = \Delta t$, Δt being the sampling period and $d_{\text{inf}}=0$, i.e. all the solutions found are stored.

Beyond this heuristics, common constraint satisfaction problems heuristics can be exploited to speed up the matching, as searching first for uncommon events of the pattern (a high value, a sharp peak, etc.).

6 Validation

The validation that is offered here should be considered as a proof of concept and of the efficiency of the algorithms, and not as an exhaustive validation of a medical pattern. This validation consists of detecting the sinusal rhythm pattern with a reading lasting a little over 10 minutes that has a total of 1,263 beats, and to do so we use the Tool foR anAlyzing and disCovering pattErns [4] (TRACE, see Fig. 3), with which clinical research on medical signals is currently being carried out. The results of the detection are contrasted with the visual detection made by a medical specialist; there were 1,129 true positives (TP), 102 true negatives (TN), 4 false positives (FP) and 32 false negatives (FN). In our evaluation, the TPs are detected P waves that belong to a beat from the sinusal basal rhythm, while their non-detection represents a FN. Electrical activity confusion that does not correspond to any beat results in a FP. Lack of confusion in the electrical activity by auricular extrasystole followed by an early QRS complex constitutes a TN. Thus we obtain a sensitivity of 0.917 and a specificity of 0.889, with a confidence interval of (95%,0.9–0.931) and (95%,0.747–0.956), respectively. In spite of dealing with a pattern of great physiological variability and having carried out the detection over a complex recording, without filtering, and being subjected to all kinds of artifacts, the results are highly satisfactory.

The detection, carried out on a Pentium III at 800 MHz, required 75 seconds. This proves that, in spite of the theoretically high computational complexity of the recognition algorithms, they can be used for real-time detection, even in difficult cases: the ECG was sampled at 250 Hz. This, along with the quality of the results, shows the validity of the recognition algorithms.

7 Discussion and conclusions

We have presented a hierarchical matching procedure for signal abstraction. The use of fuzzy logic enables us to represent and handle the imprecision and uncertainty that are characteristic of human knowledge. Explicitly capturing the abstraction hierarchy that is present in the pattern simplifies the acquisition and revision of expert knowledge and allows the matching to be performed in as many levels of abstraction as the human expert employs to reason about the system, enabling the complete histories of findings of interest at each level of abstraction to be constructed. Thus, the human expert can be given detailed explanations as how the lower-level information has been combined in order to generate higher-level information. It is also more suitable for agent-based implementation, where each agent can take charge of matching each finding, using the results from the previous agents.

The heuristics that we have presented sacrifices the completeness of the matching algorithms. Nevertheless, it is highly probable that the solutions found will enable the global pattern, if it exists, to be found, since they act as a sampling of each occurrence of a finding. Using values for ΔL_W and $d_{\inf}$ that are adapted to the dynamics of the system the heuristics has never led to occurrences of the pattern being lost, and very rarely has it generated non-optimal solutions. On the other hand, the heuristics has made it possible to fully comply with real time requirements in the detection of patterns studied in the domains of patient supervision and mobile robotics.

With regard to future work, we aim to study the problem of network consistency: the description of the MFTP may be inconsistent, which means that there is no completely possible solution; hence the MFTP should be revised. Even when is consistent, it maybe possible to refine the knowledge projected onto a MFTP, which would result in a more efficient matching process.

References

1. R. Dechter. *Constraint Processing*. Morgan Kaufmann Publishers, 2003.
2. P. Félix, S. Barro, and R. Marín. Fuzzy constraint networks for signal pattern recognition. *Artificial Intelligence*, 148:103–140, 2003.
3. HJ. Marriott and MB. Conover. *Advanced Concepts in Arrhythmias*. Mosby, 1998.
4. A. Otero, S. Barro P. Félix, and F. Palacios. A tool for the analysis and synthesis of alarms in patient monitoring. In *FUSION2004*, pages 951–958, 2004.
5. A. Otero, CV. Rodríguez, J. Correa, M. Rodríguez, and P. Félix. A fuzzy constraint satisfaction approach for landmark recognition in mobile robotics. In *IPMU2004*, pages 183–191, 2004.
6. LA. Zadeh. The concept of a linguistic variable and its application to approximate reasoning. *Information Science*, 8:199–249, 1975.