

# Temporal granular logic for temporal data mining

Paul Cotofrei

University of Neuchâtel  
2000 Neuchâtel, Switzerland  
Email: paul.cotofrei@unine.ch

Kilian Stoffel

University of Neuchâtel  
2000 Neuchâtel, Switzerland  
Email: kilian.stoffel@unine.ch

**Abstract**—In this article, a formalism for a specific temporal data mining task (the discovery of rules, inferred from databases of events having a temporal dimension), is defined. The proposed theoretical framework, based on first-order temporal logic, allows the definition of the main notions (*event*, *temporal rule*, *constraint*) in a formal way. This formalism is then extended to include the notion of temporal granularity and a detailed study is made to investigate the formal relationships between semantics for the same event in linear time structures with different granularities.

## I. INTRODUCTION

The domain of temporal data mining focuses on the discovery of causal relationships among events that are ordered in time and may be causally related. The contributions in this domain encompass the discovery of temporal rule, of sequences and of patterns. However, in many respects this is just a terminological heterogeneity among researchers that are, nevertheless, addressing the same problem, albeit from different starting points and domains.

Although there is a rich bibliography concerning formalism for temporal databases, there are very few articles on this topic for temporal data mining. In [1]–[3] general frameworks for temporal mining are proposed, but usually the researches on causal and temporal rules are more concentrated on the methodological/algorithmic aspect, and less on the theoretical aspect. Based on a methodology for temporal rule extraction, described in [4], we proposed in [5] an innovative formalism based on first-order temporal logic, which permits an abstract view on temporal rules. The formalism is developed around a time model for which the events are those that describe system evolution (event-based temporal logics). Each formula expresses what the system does at each event, events are referring to other events, and so on: this results in specifying relationships of precedence and cause-effect among events. But the real systems are systems whose components (events) have dynamic behavior regulated by very different (even by orders of magnitude) time granularities. Analyzing such systems (hereinafter granular systems) means to approach theories, methodologies, techniques and tools that make use of granules (or groups, classes, clusters of a universe) in the process of problem solving. Granular computing (the label which covers these approaches) is essential to human problem solving, and hence has a very significant impact on the design and implementation of intelligent systems. By focusing on different levels of granularities, one can obtain various levels

of knowledge, as well as inherent knowledge structure [6]–[9].

Despite the widespread recognition of its relevance in the fields of formal specifications, knowledge representation and temporal databases, there is a lack of a systematic framework for time granularity. Hobbs [10] proposed a formal characterization of the general notion of granularity, but gives no special attention to time granularity. Clifford et al. [11] provide a set-theoretic formalization of time granularity, but they do not attempt to relate the truth value of assertions to time granularity. Extensions to existing languages for formal specifications, knowledge representation and temporal databases that do support a limited concept of time granularity are proposed in [12]–[14]. Finally, Bettini et al. [15], [16] provide a formal framework for expressing data mining tasks involving time granularities, investigate the formal relationships among event structures that have temporal constraints, define the pattern-discovery problem with these structures and study effective algorithms to solve it.

The purpose of this paper is to extend our formalism to include the concept of time granularity. We define the process for which a given structure of time granules  $\mu$  (called temporal type) induces a first-order linear time structure  $M_\mu$  on the basic (or absolute) linear time structure  $M$ . The major change for the temporal logic based on  $M_\mu$  is at the semantic level: for a formula  $p$ , the interpretation do not assign a meaning of truth (one of the values  $\{true, false\}$ ), but a degree of truth (a real value from  $[0, 1]$ ). Consequently, we can give an answer to the following question: if the temporal type  $\mu$  is *finer than* temporal type  $\nu$ , what is the relationship between the interpretations of the same formula  $p$  in the linear time structures  $M_\mu$  and  $M_\nu$ . We also study the variation process for the set of satisfiable events (degree of truth equal one) during the transition between two time structures with different granularity. By an extension at the syntactic and semantic level we define a mechanism of aggregation for events, that reflects the following intuitive phenomenon: in a coarser world, not all events inherited from a finer world are satisfied, but in exchange there are new events which become satisfiable.

The rest of the paper is structured as follows. In the next section, the first-order temporal logic formalism is extensively described (the main terms and concepts). The definitions and theorems concerning the extension of the formalism towards a temporal granular logic are presented in Section 3. Finally, the last section summarizes our work and the appendix contains the proofs of the theorems in the paper.

## II. THE FORMALISM OF TEMPORAL RULES

Time is ubiquitous in information systems, but the mode of representation/perception varies in function of the purpose of the analysis [17], [18]. Firstly, there is a choice of a *temporal ontology*, which can be based either on *time points* (instants) or on *intervals* (periods). Secondly, time may have a *discrete* or a *continuous* structure. Finally, there is a choice of *linear* vs. *nonlinear* time (e.g. acyclic graph). Our selection, imposed by the discrete representation of all databases, is a temporal domain represented by linearly ordered discrete instants.

Databases being first-order structures, the first-order logic represents a natural formalism for their description. Consequently, the first-order temporal logic expresses the formalism of temporal databases. For the purpose of our approach we consider a restricted first-order temporal language  $L$  which contains only constant symbols  $\{c, d, \dots\}$ ,  $n$ -ary ( $n \geq 1$ ) function symbols  $\{f, g, \dots\}$ , variable symbols  $\{y_1, y_2, \dots\}$ ,  $n$ -ary predicate symbols ( $n \geq 1$ , so no proposition symbols), the set of relational symbols  $\{=, <, \leq, >, \geq\}$ , the logical connective  $\{\wedge\}$  and a temporal connective of the form  $X_k$ ,  $k \in \mathbb{Z}$ , where  $k$  strictly positive means *after  $k$  time instants*,  $k$  strictly negative means *before  $k$  time instant* and  $k = 0$  means *now*.

The syntax of  $L$  defines terms, atomic formulae and compound formulae. A Horn clause is a formula of the form  $B_1 \wedge \dots \wedge B_m \rightarrow B_{m+1}$  where each  $B_i$  is a positive (non-negated) atom. Syntactically, we cannot express Horn clauses in our language  $L$  because the logical connective  $\rightarrow$  is not defined. However, to allow the description of rules, which formally look like a Horn clause, we introduce a new logical connective,  $\mapsto$ , which practically will represent a rewrite of the connective  $\wedge$ . Therefore, a formula in  $L$  of the form  $p \mapsto q$  is syntactically equivalent to the formula  $p \wedge q$ . When and under what conditions we may use the new connective, is explained in the following definitions.

**DEFINITION 1** *An event (or temporal atom) is an atom formed by the predicate symbol  $E$  followed by a bracketed  $n$ -tuple of terms ( $n \geq 1$ )  $E(t_1, t_2, \dots, t_n)$ . The first term of the tuple,  $t_1$ , is a constant symbol representing the name of the event and all others terms are expressed according to the rule  $t_i = f(t_{i1}, \dots, t_{ik_i})$ . A short temporal atom (or the event's head) is the atom  $E(t_1)$ .*

**DEFINITION 2** *A constraint formula for the event  $E(t_1, t_2, \dots, t_n)$  is a conjunctive compound formula,  $E(t_1, t_2, \dots, t_n) \wedge C_1 \wedge C_2 \wedge \dots \wedge C_k$ . Each  $C_j$  is a relational atom  $tpc$ , where the first term  $t$  is one of the terms  $t_i$ ,  $i = 1 \dots n$ ,  $p$  is a relational symbol and the second term is a constant symbol.*

For a short temporal atom  $E(t_1)$ , the only constraint formula that is permitted is  $E(t_1) \wedge (t_1 = c)$ . We denote such constraint formula as *short constraint formula*.

**DEFINITION 3** *A temporal rule is a formula of the form  $H_1 \wedge \dots \wedge H_m \mapsto H_{m+1}$ , where  $H_{m+1}$  is a short constraint formula*

and  $H_i, i = 1..m$  are constraint formulae, prefixed by the temporal connectives  $X_{-k}$ ,  $k \geq 0$ . The maximum value of the index  $k$  is called the *time window of the temporal rule*.

*Remark.* The reason for which we did not permit the expression of the implication connective in our language is related to the truth table for a formula  $p \rightarrow q$ : even if  $p$  is false, the formula is still true, which is unacceptable for a temporal rationing of the form *cause*  $\rightarrow$  *effect*.

If we change in Definition 1 the conditions imposed on the terms  $t_i, i = 1..n$  to "each term  $t_i$  is a variable symbol", we obtain the definition of a temporal atom template. We denote such a template as  $E(y_1, \dots, y_n)$ . Following the same rationing, a constraint formula template for  $E(y_1, \dots, y_n)$  is a conjunctive compound formula,  $C_1 \wedge C_2 \wedge \dots \wedge C_k$ , where the first term of each relational atom  $C_j$  is one of the variables  $y_i$ ,  $i = 1 \dots n$ . Consequently, a short constraint formula template is the relational atom  $y_1 = c$ . Finally, by replacing in Definition 3 the notion "constraint formula" with "constraint formula template" we obtain the definition of a temporal rule template. Practically, the only formulae constructed in  $L$  are temporal atoms, constraint formulae, temporal rules and the corresponding templates.

The semantics of  $L$  is provided by an interpretation  $I$  over a domain  $D$  (in our formalism,  $D$  is always a linearly ordered domain). The interpretation assigns an appropriate meaning over  $D$  to the (non-logical) symbols of  $L$ . Usually, the domain  $D$  is imposed during the discretisation phase, which is a pre-processing phase used in almost all knowledge extraction methodologies. Based on Definition 2, an event can be seen as a labelled (constant symbol  $t_1$ ) sequence of points extracted from raw data and characterized by a finite set of features (terms  $t_2, \dots, t_n$ ). Consequently, the domain  $D$  is the union  $D_e \cup D_f$ , where the set  $D_e$  contains all the strings used as event names and the set  $D_f$  represents the union of all domains corresponding to chosen features.

**EXAMPLE 1** Consider a database containing daily price variations of a given stock. After the application of the discretisation phase we obtain an ordered sequence of events. Each event has the form  $(name, v_1, v_2)$ , where the name is one of the strings from  $D_e = \{peak, flat, valley\}$  and the features  $v_1, v_2$  represents the mean, respectively, the standard error. The statistics are calculated using daily prices, supposed to be subsequences of length  $w = 12$ . Consequently, a temporal atom in  $L$  is defined as  $E(d_i, f(c_{j1}, \dots, c_{j12}), g(c_{k1}, \dots, c_{k12}))$ , where  $d_i \in D_e$ ,  $c_i \in D_f = \mathbb{R}$  and the meaning of symbol function  $f$  (respectively  $g$ ) is given by the statistical function  $\bar{x}$  (respectively  $se(\bar{x})$ ).

To define a first-order linear temporal logic based on  $L$ , we need a structure having a temporal dimension and capable to capture the relationship between a time moment and the interpretation  $I$  at this moment.

**DEFINITION 4** *Given  $L$  and a domain  $D$ , a (first order) linear time structure is a triple  $M = (S, x, I)$ , where  $S$  is a set*

of states,  $x : \mathbb{N} \rightarrow S$  is an infinite sequence of states  $(s_1, s_2, \dots, s_n, \dots)$  and  $\mathbf{I}$  is a function that associates with each state  $s$  an interpretation  $\mathbf{I}_s$  of all symbols from  $L$ .

In the framework of linear temporal logic, the set of symbols is divided into two classes, the class of global symbols and the class of local symbols. Intuitively, a global symbol  $w$  has the same interpretation in each state, i.e.  $\mathbf{I}_s(w) = \mathbf{I}_{s'}(w) = \mathbf{I}(w)$ , for all  $s, s' \in S$ ; the interpretation of a local symbol may vary, depending on the state at which is evaluated. The formalism of temporal rules assumes that all function symbols (including constants) and all relational symbols are global, whereas the predicate symbols and variable symbols are local. Consequently, as the temporal atoms, constraint formulae, temporal rules and the corresponding templates are expressed using the predicate symbol  $E$  or the variable symbols  $y_i$ , the meaning of truth for these formulae depend on the state at which they are evaluated. Given a first order time structure  $M$  and a formula  $p$ , we denote the instant  $i$  (or equivalently, the state  $s_i$ ) for which  $\mathbf{I}_{s_i}(p) = \text{true}$  by  $i \models p$ , i.e. at time instant  $i$  the formula  $p$  is true. Therefore,  $i \models E(t_1, \dots, t_n)$  means that at time  $i$  an event with the name  $\mathbf{I}(t_1)$  and characterized by the global features  $\mathbf{I}(t_2), \dots, \mathbf{I}(t_n)$  occurs. Concerning the event template  $E(y_1, \dots, y_n)$ , the interpretation of the variable symbols  $y_j$  at the state  $s_i$ ,  $\mathbf{I}_{s_i}(y_j)$ , is chosen such that  $i \models E(y_1, \dots, y_n)$  for all time moment  $i$ . Because

- $i \models p \wedge q$  if and only if  $i \models p$  and  $i \models q$ , and
- $i \models X_k p$  if and only if  $i + k \models p$ ,

a constraint formula (template) is true at time  $i$  if and only if all relational atoms are true at time  $i$  and  $i \models E(t_1, \dots, t_n)$ , whereas a temporal rule (template) is true at time  $i$  if and only if  $i \models H_{m+1}$  and  $i \models (H_1 \wedge \dots \wedge H_m)$ .

Now suppose that the following assumptions are true:

- For each formula  $p$  in  $L$ , there is an algorithm that calculates the value of the interpretation  $\mathbf{I}_s(p)$ , for each state  $s$ , in a finite number of steps.
- There are states (called incomplete states) that do not contain enough information to calculate the interpretation for all formulae defined at these states.
- It is possible to establish a measure, (called *general interpretation*) about the degree of truth of a compound formula along the entire sequence of states  $(s_0, s_1, \dots, s_n, \dots)$ .

The first assumption express the calculability of the interpretation  $\mathbf{I}$ . The second assumption express the situation when only the body of a temporal rule can be evaluated at a time moment  $i$ , but not the head of the rule. Therefore, for the state  $s_i$ , we cannot calculate the interpretation of the temporal rule and the only solution is to estimate it using a general interpretation. This solution is expressed by the third assumption. (*Remark:* The second assumption violates the condition about the existence of an interpretation in each state  $s_i$ , as defined in Definition 4. But it is well known that in data mining sometimes data is incomplete or is missing. Therefore, we must modify this condition as " $\mathbf{I}$  is a function

that associates with almost each state  $s$  an interpretation  $\mathbf{I}_s$  of all symbols from  $L$ ").

However, to ensure that this general interpretation is well defined, the linear time structure must present some property of consistency. Practically, this means that if we take any sufficiently large subset of time instants, the conclusions we may infer from this subset are sufficiently close from those inferred from the entire set of time instants. Therefore,

**DEFINITION 5** Given  $L$  and a linear time structure  $M$ , we say that  $M$  is a consistent time structure for  $L$  if, for every formula  $p$ , the limit  $\text{supp}(p) = \lim_{n \rightarrow \infty} n^{-1} \#A$  exists, where  $\#$  means "cardinality" and  $A = \{i = 1..n \mid i \models p\}$ . The notation  $\text{supp}(p)$  denotes the support (of truth) of  $p$ .

Now we define the general interpretation for an  $n$ -ary predicate symbol  $P$  as:

**DEFINITION 6** Given  $L$  and a consistent linear time structure  $M$  for  $L$ , the general interpretation  $I_G$  for an  $n$ -ary predicate  $P$  is a function  $D^n \rightarrow [0, 1]$ , such that, for each  $n$ -tuple of terms  $\{t_1, \dots, t_n\}$ ,  $I_G(P(t_1, \dots, t_n)) = \text{supp}(P(t_1, \dots, t_n))$ .

The general interpretation is naturally extended to constraint formulae, temporal rules and the corresponding templates. There is another useful measure, called *confidence*, but available only for temporal rules (templates). This measure is calculated as a limit ratio between the number of certain applications (time instants where both the body and the head of the rule are true) and the number of potential applications (time instants where only the body of the rule is true). The reason for this choice is related to the presence of incomplete states, where the interpretation for the implicated clause cannot be calculated.

**DEFINITION 7** The confidence of a temporal rule (template)  $H_1 \wedge \dots \wedge H_m \mapsto H_{m+1}$  is the limit  $\lim_{n \rightarrow \infty} (\#B)^{-1} \#A$ , where  $A = \{i = 1 \dots n \mid i \models H_1 \wedge \dots \wedge H_m \wedge H_{m+1}\}$  and  $B = \{i = 1 \dots n \mid i \models H_1 \wedge \dots \wedge H_m\}$ .

For different reasons, (the user has not access to the entire sequence of states, or the states he has access to are incomplete), the general interpretation cannot be calculated. A solution is to estimate  $I_G$  using a finite linear time structure, i.e. a model.

**DEFINITION 8** Given  $L$  and a consistent time structure  $M = (S, x, \mathbf{I})$ , a model for  $M$  is a structure  $\tilde{M} = (\tilde{T}, \tilde{x})$  where  $\tilde{T}$  is a finite temporal domain  $\{i_1, \dots, i_n\}$ ,  $\tilde{x}$  is the subsequence of states  $\{x_{i_1}, \dots, x_{i_n}\}$  (the restriction of  $x$  to the temporal domain  $\tilde{T}$ ) and for each  $i_j, j = 1, \dots, n$ , the state  $x_{i_j}$  is a complete state.

For a given model  $\tilde{M}$ , the local support for a formula  $p$  (denoted  $\text{supp}(p, \tilde{M})$ ) is defined as the ratio  $\frac{\#A}{\#\tilde{T}}$ , where  $A = \{i \in \tilde{T} \mid i \models p\}$ . Now we may define the estimator for a general interpretation:

DEFINITION 9 Given  $L$  and a model  $\tilde{M}$  for  $M$ , an estimator of the general interpretation for an  $n$ -ary predicate  $P$ ,  $I_{G(\tilde{M})}(P)$ , is a function  $D^n \rightarrow [0, 1]$ , assigning to each formula  $p = P(t_1, \dots, t_n)$  the value  $\text{supp}(p, \tilde{M})$ .

Similarly, the estimation of the confidence for a temporal rule (template) is defined as:

DEFINITION 10 Given a model  $\tilde{M} = (\tilde{T}, \tilde{x})$  for  $M$ , the estimation of the confidence for the temporal rule (template)  $H_1 \wedge \dots \wedge H_m \mapsto H_{m+1}$  is the ratio  $(\#B)^{-1} \#A$ , where  $A = \{i \in \tilde{T} \mid i \models H_1 \wedge \dots \wedge H_m \wedge H_{m+1}\}$  and  $B = \{i \in \tilde{T} \mid i \models H_1 \wedge \dots \wedge H_m\}$ .

EXAMPLE 2 If  $M$  is a consistent linear time structure for the temporal language  $L$  defined in Example 1, consider the model  $\tilde{M}$  given by the sequence of states  $\{s_1, \dots, s_{100}\}$ . Let be  $X_{-3}(y_1 = \text{peak}) \wedge X_{-3}(y_2 < 11) \wedge X_{-1}(y_1 = \text{peak}) \mapsto X_0(y_1 = \text{valley})$  a temporal rule template. If the local support for the rule is 0.6 and the local support for the body of the rule is 0.8 then the estimated confidence for the rule, based on model  $\tilde{M}$ , is  $0.6/0.8 = 0.75$ .

### III. THE GRANULARITY MODEL

We start with the concept of a temporal type to formalize the notion of time granularities, as described in [19].

DEFINITION 11 Let  $(\mathcal{T}, <)$  (index) be a linearly ordered temporal domain isomorphic to a subset of the integers with the usual order relation, and let  $(\mathcal{A}, <)$  (absolute time) be a linearly ordered set. Then, a temporal type on  $(\mathcal{T}, \mathcal{A})$  is a mapping  $\mu$  from  $\mathcal{T}$  to  $2^{\mathcal{A}}$  such that

- 1)  $\mu(i) \neq \emptyset$  and  $\mu(j) \neq \emptyset$ , where  $i < j$ , imply that each element in  $\mu(i)$  is less than all the elements in  $\mu(j)$ , and
- 2) for all  $i < j$ , if  $\mu(i) \neq \emptyset$  and  $\mu(j) \neq \emptyset$ , then  $\forall k, i < k < j$  implies  $\mu(k) \neq \emptyset$ .

Each set  $\mu(i)$ , if non-empty, is called a granule of  $\mu$ . Property (1) says that granules do not overlap and that the order on indices follows the order on the corresponding granules. Property (2) disallows an empty set to be the value of a mapping for a certain index value if a lower index and a higher index are mapped to non-empty sets. Two temporal types  $\mu_1$  and  $\mu_2$  are said to be *shifting equivalent* (denoted  $\mu_1 \rightleftharpoons \mu_2$ ) if there is a bijection function  $h : \mathcal{T} \rightarrow \mathcal{T}$  such that  $\mu_1(i) = \mu_2(h(i))$ , for all  $i \in \mathcal{T}$ . In the following we disallow multiple types that are equivalent with respect to shifting of their indices, i.e.

$$\mu_1 \rightleftharpoons \mu_2 \Rightarrow \mu_1 = \mu_2. \quad (1)$$

DEFINITION 12 Let  $\mu$  and  $\nu$  be temporal types on  $(\mathcal{T}, \mathcal{A})$ .

- Finer-than:  $\mu$  is said to be finer than  $\nu$ , denoted  $\mu \preceq \nu$ , if for each  $i \in \mathcal{T}$ , there exists  $j \in \mathcal{T}$  such that  $\mu(i) \subseteq \nu(j)$ .
- Groups-into:  $\mu$  is said to group into  $\nu$ , denoted  $\mu \trianglelefteq \nu$ , if for each non-empty granule  $\nu(j)$ , there is a subset  $S$  of  $\mathcal{T}$  such that  $\nu(j) = \bigcup_{i \in S} \mu(i)$ .

When a temporal type  $\mu$  is finer than a temporal type  $\nu$ , we also say that  $\nu$  is coarser than  $\mu$ . The *finer-than* relationship formalizes the notion of finer partitions of the absolute time. By definition, this relation is reflexive, i.e.  $\mu \preceq \mu$  for each temporal type  $\mu$ . Furthermore, the finer-than relation is obviously transitive and antisymmetric (a consequence of (1)) and, hence, it is a partial order. Therefore, there exists a unique least upper bound of the set of all temporal types, denoted by  $\mu_{\top}$ , and a unique greatest lower bound, denoted by  $\mu_{\perp}$ . These top and bottom elements are defined as follows:  $\mu_{\top}(i) = \mathcal{A}$  for some  $i \in \mathcal{T}$  and  $\mu_{\top}(j) = \emptyset$  for each  $j \neq i$ , and  $\mu_{\perp}(i) = \emptyset$  for each  $i \in \mathcal{T}$ . We formalize this result in the following theorem [19]:

THEOREM 1 Any temporal type system having an infinite index, and satisfying (1), is a lattice with respect to the finer-than relationship.

Concerning the *groups-into* relationship, it also satisfies the properties of a partial order (reflexivity, transitivity and antisymmetry), but generally  $\mu \preceq \nu$  not imply  $\mu \trianglelefteq \nu$  or viceversa.

Consider now  $\mathcal{A} = \mathcal{T} = \mathbb{N}$ . If we impose to any temporal type  $\mu$  the condition

$$0 < \#\mu(i) < \infty, \text{ for all } i \in \mathbb{N} \quad (2)$$

then it can be proven (see Appendix) that the condition (1) is a consequence of the condition (2). Therefore, for the set of temporal types satisfying (2) and the following restriction

$$\forall i \in \mathbb{N}, \mu^{-1}(i) \neq \emptyset \quad (3)$$

it can be shown that there exists a unique greatest lower bound,  $\mu_{\perp}(i) = i, \forall i \in \mathbb{N}$ , but no least upper bound  $\mu_{\top}$ . Furthermore, the condition (3) is a sufficient condition for the equivalence of the relationships *finer-than* and *groups-into*, according to the following lemma:

LEMMA 1 If  $\mu$  and  $\nu$  are temporal types on  $(\mathbb{N}, \mathbb{N})$  which satisfy the conditions (2) and (3) (shortly, are of type  $\mathcal{G}_1$ ), then  $\mu \preceq \nu \iff \mu \trianglelefteq \nu$ .

If  $M = (S, x, \mathbf{I})$  is a first-order linear time structure, then let the absolute time  $\mathcal{A}$  be given by the sequence  $x$ , by identifying the time moment  $i$  with the state  $s_i$  (on the  $i^{\text{th}}$  position in the sequence). If  $\mu$  is a temporal type from  $\mathcal{G}_1$ , then the temporal granule  $\mu(i)$  may be identified with the set  $\{s_j \in S \mid j \in \mu(i)\}$ . Therefore, the temporal type  $\mu$  induces a new sequence,  $x_{\mu}$ , defined as  $x_{\mu} : \mathbb{N} \rightarrow 2^S, x_{\mu}(i) = \mu(i)$ . (Remark: In the following the set  $\mu(i)$  will be considered, depending of the context, either as a set of natural numbers, or as a set of states).

Consider now the linear time structure derived from  $M$ ,  $M_{\mu} = (2^S, x_{\mu}, \mathbf{I}^{\mu})$ . To be well defined, we must precise the interpretation  $\mathbf{I}_{\mu(i)}^{\mu}$ , for each  $i \in \mathbb{N}$ . Because for a fixed  $i$  the set  $\mu(i)$  is a finite sequence of states, it defines (if all the states are complete states) a model  $\tilde{M}_{\mu(i)}$  for  $M$ . Therefore the

estimated general interpretation  $I_{G(\tilde{M}_{\mu(i)})}$  is well defined and we consider, by definition, that for all temporal free<sup>1</sup> formula  $p$  in  $L$ ,

$$\mathbf{I}_{\mu(i)}^\mu(p) = I_{G(\tilde{M}_{\mu(i)})}(p) = \text{supp}(p, \tilde{M}_{\mu(i)}) \quad (4)$$

This interpretation is extended to any temporal formula in  $L$  according to the rule:

$$\mathbf{I}_{\mu(i)}^\mu(X_{k_1}p_1 \wedge \dots \wedge X_{k_n}p_n) = \frac{1}{n} \sum_{j=1}^n \mathbf{I}_{\mu(i+k_j)}^\mu(p_j) \quad (5)$$

where  $p_i$  are temporal free formulae and  $k_i \in \mathbb{Z}, i = 1 \dots n$ .

**DEFINITION 13** *If  $M = (S, x, \mathbf{I})$  is a first-order linear time structure and  $\mu$  is a temporal type from  $\mathcal{G}_1$ , then the temporal time structure induced by  $\mu$  on  $M$  is the triple  $M_\mu = (2^S, x_\mu, \mathbf{I}^\mu)$ , where  $x_\mu : \mathbb{N} \rightarrow 2^S$ ,  $x_\mu(i) = \mu(i)$  and  $\mathbf{I}^\mu$  is a function that associates with almost each set of states  $\mu(i)$  an interpretation  $\mathbf{I}_{\mu(i)}^\mu$  according to the rules (4) and (5).*

Of a particular interest is the linear time structure induces by the greatest lower bound temporal type  $\mu_\perp(i) = i$ . In this case,  $M_{\mu_\perp} = (S, x, \mathbf{I}^{\mu_\perp})$ , where the only difference from the initial time structure  $M$  is at the interpretation level: for  $p = P(t_1, \dots, t_n)$  a formula in  $L$ , if the interpretation  $\mathbf{I}_s(p)$  is a function defined on  $D^n$  with values in  $\{\text{true}, \text{false}\}$  – giving so the meaning of truth – the interpretation  $\mathbf{I}_s^{\mu_\perp}(p)$  is a function defined on  $D^n$  with values in  $[0, 1]$  – giving so the degree of truth. The relation linking the two interpretations is given by  $\mathbf{I}_s(p) = \text{true}$  iff  $\mathbf{I}_s^{\mu_\perp}(p) = 1$ . Indeed, supposing the state  $s_i$  is a complete state, it defines the model  $M_i = (i, s_i)$  and we have, for  $p$  a temporal free formula,

$$\mathbf{I}_{\mu_\perp(i)}^{\mu_\perp}(p) = \text{supp}(p, \tilde{M}_i) = \begin{cases} 1, & \text{if } \mathbf{I}_s(p) = \text{true}, \\ 0, & \text{if } \mathbf{I}_s(p) = \text{false} \end{cases} \quad (6)$$

For a formula  $\pi = X_{k_1}p_1 \wedge \dots \wedge X_{k_n}p_n$ , we have  $\mathbf{I}_{s_i}(\pi) = \text{true}$  iff  $\forall j \in \{1 \dots n\}, i + k_j \models p_j$ , which is equivalent with

$$\begin{aligned} \text{supp}(p_1, \tilde{M}_{i+k_1}) = \dots = \text{supp}(p_n, \tilde{M}_{i+k_n}) = 1 &\Leftrightarrow \\ \Leftrightarrow \frac{1}{n} \sum_{j=1}^n \mathbf{I}_{\mu_\perp(i+k_j)}^{\mu_\perp}(p_j) = \mathbf{I}_{\mu_\perp(i)}^{\mu_\perp}(\pi) = 1. \end{aligned}$$

Consequently, the linear time structure  $M_{\mu_\perp}$  can be seen as an extension, at the interpretation level, of the classical linear time structure  $M$ . Furthermore, all the time structures induced by a temporal type have in common interpretations which take values in  $[0, 1]$  if are applied on predicate symbols in  $L$ . This observation allows us to establish the relation linking the interpretations  $\mathbf{I}^\mu$  and  $\mathbf{I}^\nu$ , from two linear time structures induced by  $\mu$  and  $\nu$ , when exists a relationship *finer-than* between these two temporal types. According to the lemma 1, for each  $i \in \mathbb{N}$  there is a subset  $N_i \subset \mathbb{N}$  such that  $\nu(i) = \bigcup_{j \in N_i} \mu(j)$ . If  $p$  is a temporal free formula in  $L$ ,

then the interpretation  $\mathbf{I}^\nu$  for  $p$  at  $\nu(i)$  is the weighted sum of the interpretations  $\mathbf{I}_{\mu(j)}^\mu(p)$ , where  $j \in N_i$ . We formalize this result in the following theorem:

**THEOREM 2** *If  $\mu, \nu$  are temporal types from  $\mathcal{G}_1$ , such that  $\mu \preceq \nu$ , and  $\mathbf{I}^\mu, \mathbf{I}^\nu$  are the interpretations from the induced linear time structures  $M_\mu$  and  $M_\nu$  on  $M$ , then for each  $i \in \mathbb{N}$ ,*

$$\mathbf{I}_{\nu(i)}^\nu(p) = \frac{1}{\#\nu(i)} \sum_{j \in N_i} \#\mu(j) \mathbf{I}_{\mu(j)}^\mu(p), \quad (7)$$

where  $N_i$  is the subset of  $\mathbb{N}$  which satisfies  $\nu(i) = \bigcup_{j \in N_i} \mu(j)$  and  $p$  is a temporal free formula in  $L$ .

If we consider  $\mu = \mu_\perp$  then  $\#\mu(j) = 1$ , for all  $j \in \mathbb{N}$  and  $\mathbf{I}_{\mu(j)}^\mu(p) = \text{supp}(p, \tilde{M}_j)$  (see (6)). Therefore,

$$\begin{aligned} \mathbf{I}_{\nu(i)}^\nu(p) &= \frac{1}{\#\nu(i)} \sum_{j \in \nu(i)} \text{supp}(p, \tilde{M}_j) \\ &= \frac{1}{\#\nu(i)} \#\{j \in \nu(i) \mid j \models p\} \\ &= \text{supp}(p, \tilde{M}_{\nu(i)}) = \mathbf{I}_{G(\tilde{M}_{\nu(i)})}(p) \end{aligned}$$

result which is consistent with the definition 13. But the significance of the theorem 2 is revealed in a particular context. Firstly, let  $\mathcal{G}_2$  be the subset of  $\mathcal{G}_1$ , obtained when the condition (2) is replaced by the stronger condition (2'),  $\#\mu(i) = c_\mu$ , where  $c_\mu \in \mathbb{N}$ . If  $\mu, \nu \in \mathcal{G}_2$  and  $\mu \preceq \nu$ , it can be shown that  $\#N_i = \frac{c_\nu}{c_\mu}, \forall i \in \mathbb{N}$  and so the relation (7) becomes

$$\mathbf{I}_{\nu(i)}^\nu(p) = \frac{1}{\#N_i} \sum_{j \in N_i} \mathbf{I}_{\mu(j)}^\mu(p). \quad (8)$$

Generally speaking, consider three worlds,  $W_1, W_2$  and  $W_3$  – defined as sets of granules of information – where  $W_1$  is finer than  $W_2$  which is finer than  $W_3$ . Suppose also that the conversion between granules from two different worlds is given by a constant factor. If the independent part of information in each granule is transferred from  $W_1$  to  $W_2$  and then the world  $W_1$  is "lost", the theorem 2 in the form (8) affirm that is possible to transfer the independent information from  $W_2$  to  $W_3$  and to obtain the same result as for the transfer from  $W_1$  to  $W_3$ .

**EXAMPLE 3 :** Consider a linear time structure  $M$  (here, the world  $W_1$ ) and a temporal free formula  $p$  such that, for the first six time moments, we have  $i \models p$  for  $i \in \{1, 3, 5, 6\}$  – the concept of independence, in this example, means that the interpretation of  $p$  depends only on the state  $i$ . Let be  $\mu, \nu \in \mathcal{G}_2$ ,  $\mu \preceq \nu$ , with  $\mu(i) = \{2i - 1, 2i\}$  and  $\nu(i) = \{6i - 5, \dots, 6i\}$ . Therefore,  $\nu(1) = \mu(1) \cup \mu(2) \cup \mu(3)$ . According to the definition 13,  $\mathbf{I}_{\mu(1)}^\mu(p) = 0.5$ ,  $\mathbf{I}_{\mu(2)}^\mu(p) = 0.5$ ,  $\mathbf{I}_{\mu(3)}^\mu(p) = 1$ , whereas  $\mathbf{I}_{\nu(1)}^\nu(p) = 0.66$ . If the linear time structure  $M$  is "lost", the temporal types  $\mu$  and  $\nu$  are "lost" too (we don't know the absolute time  $\mathcal{A}$  given by  $M$ ). But if we know the induced time structure  $M_\mu$  (world  $W_2$ ) and the relation between  $\mu$  and  $\nu$ , we can deduce the time structure

<sup>1</sup>A formula  $p$  in  $L$  is called *temporal free formula* if it don't contain the temporal connective  $X_k$ .

$M_\nu$  (world  $W_3$ ) completely. As example, according to (8),  $\mathbf{I}_{\nu(1)}^\nu(p) = \frac{1}{3} \sum_{i=1}^3 \mathbf{I}_{\mu(i)}^\mu(p) = 0.66$ .

The theorem 2 is not effective for temporal formulae (which can be seen as the dependent part of the information from a temporal granule). In this case we will show that the interpretation, in the coarser world, of a temporal formula with a given time window is linked with the interpretation, in the finer world, of a similar formula but having a larger time window.

**THEOREM 3** *If  $\mu$  and  $\nu$  are temporal types from  $\mathcal{G}_2$  such that  $\mu \preceq \nu$  and  $\mathbf{I}^\mu, \mathbf{I}^\nu$  are the interpretations from the induced linear time structures  $M_\mu$  and  $M_\nu$  on  $M$ , then for each  $i \in \mathbb{N}$ ,*

$$\mathbf{I}_{\nu(i)}^\nu(p \wedge Xq) = \frac{1}{k} \sum_{j \in N_i} \mathbf{I}_{\mu(j)}^\mu(p \wedge X_k q) \quad (9)$$

where  $k = c_\nu/c_\mu$ ,  $\nu(i) = \bigcup_{j \in N_i} \mu(j)$  and  $p, q$  are temporal free formulae in  $L$ .

According to this theorem, if we know the confidence of a temporal rule (template) in the world  $W_1$ , we can say nothing about the confidence of the same rule in the world  $W_2$ , coarser than  $W_1$ . If we define the operator

$$\text{zoom}_k(X_{k_1} q_1 \wedge \dots \wedge X_{k_n} q_n) = X_{k \cdot k_1} q_1 \wedge \dots \wedge X_{k \cdot k_n} q_n$$

then the confidence of a temporal rule  $H$  in  $W_2$  depends only on the confidence of the temporal rule  $\text{zoom}_k(H)$  in  $W_1$ , where  $k$  is the coefficient of conversion between the two worlds.

All the deduction process made until now was conducted to obtain an answer to the following question: how the degree of truth of a formula  $p$  is changing if we pass from a linear time structure with a given granularity to a coarser one. And we proved that we can give a well-defined method if we impose some restrictions on the temporal types which induce these time structures. But there is another phenomenon which follows the process of transition between two real worlds with different time granularities: new kinds of events appear, some kinds of events disappear.

**DEFINITION 14** *An event type (denoted  $E[t]$ ) is the set of all temporal atoms from  $L$  having the same name (or head).*

All the temporal atoms of a given type  $E[t]$  are constructed using the same symbol predicate and we denote by  $N[t]$  the arity of this symbol. Consider  $E(t, t_2, \dots, t_n) \in E[t]$  (where  $n = N[t]$ ). According to the definition 1, a term  $t_i, i \in \{2, \dots, n\}$  has the form  $t_i = f(t_{i1}, \dots, t_{ik_i})$ . Suppose now that for each index  $i$  the function symbol  $f$  from the expression of  $t_i$  belongs to a family of function symbols with different arities, denoted  $\mathcal{F}_i[t]$  (so different sets for different event types  $E[t]$  and different index  $i$ ). This family has the property that the interpretation for each of its member is given by a real functions which

- is applied on a variable number of arguments, and

- is invariant in the order of the arguments.

A good example of a such real function is a statistical function, e.g.  $\text{mean}(x_1, \dots, x_n)$ . Based on the set  $\mathcal{F}_i[t]$  we construct the set of terms expressed as  $f_k(c_1, \dots, c_k)$ , where  $f_k$  is a  $k$ -ary function symbol from  $\mathcal{F}_i[t]$  and  $c_i, i = 1..k$  are constant symbols. We denote such a set as  $T_i[t]$ . Consider now the operator  $\oplus$  defined on  $T_i[t] \times T_i[t] \rightarrow T_i[t]$  such that

$$f_n(c_1, \dots, c_n) \oplus f_m(d_1, \dots, d_m) = f_{n+m}(c_1, \dots, c_n, d_1, \dots, d_m)$$

Of course, because the interpretation of any function symbol from  $\mathcal{F}_i[t]$  is invariant in the order of arguments, we have

$$\begin{aligned} f_n(c_1, \dots, c_n) \oplus f_m(d_1, \dots, d_m) &= \\ &= f_n(c_{\sigma(1)}, \dots, c_{\sigma(n)}) \oplus f_m(d_{\varphi(1)}, \dots, d_{\varphi(m)}) \end{aligned}$$

where  $\sigma$  (respectively  $\varphi$ ) is a permutation of the set  $\{1, \dots, n\}$  (respectively  $\{1, \dots, m\}$ ). Furthermore, it is evidently that the operator  $\oplus$  is commutative and associative.

We introduce now a new operator (denoted  $\boxplus$ ) defined on  $E[t] \times E[t] \rightarrow E[t]$ , such that, for  $E(t, t_2, \dots, t_i, \dots, t_n) \in E[t]$ ,  $E(t, t'_2, \dots, t'_i, \dots, t'_n) \in E[t]$  we have:

$$\begin{aligned} E(t, t_2, \dots, t_i, \dots, t_n) \boxplus E(t, t'_2, \dots, t'_i, \dots, t'_n) &= \\ &= E(t, t_1 \oplus t'_1, \dots, t_i \oplus t'_i, \dots, t_n \oplus t'_n) \quad (10) \end{aligned}$$

Once again, it is evidently that the operator  $\boxplus$  is commutative and associative. Therefore, we can apply this operator on a subset  $\mathcal{E}$  of temporal atoms from  $E[t]$  and we denote the result as  $\boxplus_{e_i \in \mathcal{E}} e_i$ .

If  $M = (S, x, \mathbf{I})$  is a linear time structure, for each event type  $E[t]$  we define the subset  $E[t]_M$  satisfying the condition:

$$E[t]_M = \{e \in E[t] \mid \exists s_i \in x \text{ such that } \mathbf{I}_{s_i}(e) = \text{true}\} \quad (11)$$

In a similar manner we define  $E[t]_{\tilde{M}}$ , where  $\tilde{M} = (\tilde{T}, \tilde{x})$  is a model for  $M$ :

$$E[t]_{\tilde{M}} = \{e \in E[t] \mid \exists s_i \in \tilde{x} \text{ such that } \mathbf{I}_{s_i}(e) = \text{true}\} \quad (12)$$

In other words,  $E[t]_M$  is the set of all temporal events of type  $E[t]$  which are satisfied by  $M$ , whereas  $E[t]_{\tilde{M}} \subseteq E[t]_M$  is the set of events from  $E[t]$  satisfied by  $\tilde{M}$ . If we consider now  $M_\mu$ , the linear time structure induced by temporal type  $\mu$  on  $M$ , the definition of  $E[t]_{M_\mu}$  is derived from (11) only by changing the condition  $\mathbf{I}_{s_i}(e) = \text{true}$  in  $\mathbf{I}_{\mu(i)}^\mu(e) = 1$ . Of course, only for  $\mu = \mu_\perp$  we have  $E[t]_M = E[t]_{M_\mu}$  (we proved that  $\mathbf{I}_{s_i}(p) = \text{true} \Leftrightarrow \mathbf{I}_{\mu_\perp(i)}^{\mu_\perp}(p) = 1$ ). Generally  $E[t]_M \supset E[t]_{M_\mu}$ , which is a consequence of the fact that a coarser world satisfies less temporal events than a finer one.

**EXAMPLE 4** : If  $M$  is a linear time structure such that for the event  $e \in E[t]$  we have  $i \models e$  iff  $i$  odd, and  $\mu$  is a temporal type given by  $\mu(i) = \{2i - 1, 2i\}$ , then it is obviously that  $e \in E[t]_M$  but  $e \notin E[t]_{M_\mu}$ .

In the same time a coarser world may satisfies new events, representing a king of aggregation of local, "finer" events.

DEFINITION 15 If  $\mu$  is a temporal type from  $\mathcal{G}_1$ , we call the aggregate event of type  $t$  induced by the granule  $\mu(i)$  (denoted  $e[t]_{\mu(i)}$ ) the event obtained by applying the operator  $\boxplus$  on the set of events of type  $t$  which are satisfy by the model  $\tilde{M}_{\mu(i)}$ , i.e.

$$e[t]_{\mu(i)} = \boxplus_{e_i \in E[t]_{\tilde{M}_{\mu(i)}}} e_i \quad (13)$$

If an event  $e \in E[t]$  is not satisfied by  $M$  (or  $e \notin E[t]_M$ ), it is obviously that according to (4)  $\mathbf{I}_{\mu(i)}^\mu(e) = 0$ , for all  $\mu$  and all  $i \in \mathbb{N}$ . Therefore, the relation (4) is not appropriate to give the degree of truth for  $e[t]_{\mu(i)}$ . Before to give the rule expressing the interpretation for an aggregate temporal atom, we impose to  $M$  the following restriction: there is a one-to-one relationship between the set of events satisfied by  $M$  and the set of states, or

$$\exists h : \bigcup_t E[t]_M \rightarrow S, h(e) = s \text{ where } \mathbf{I}_s(e) = \text{true} \quad (14)$$

We define the interpretation of the aggregate event induced by  $\mu(i_0)$ , evaluated at  $\mu(i)$ , as:

$$\mathbf{I}_{\mu(i)}^\mu(e[t]_{\mu(i_0)}) = \sum_{e_j \in \mathcal{E}} \mathbf{I}_{\mu(i)}^\mu(e_j) \quad (15)$$

where  $\mathcal{E} = E[t]_{\tilde{M}_{\mu(i_0)}}$ . The restriction (14) is sufficient to assure that  $\mathbf{I}_{\mu(i)}^\mu(e[t]_{\mu(i_0)}) \leq 1$ , for all  $i, i_0 \in \mathbb{N}$ . Furthermore,  $e[t]_{\mu(i_0)} \in E[t]_{M_\mu}$  iff there is  $i \in \mathbb{N}$  such that  $h^{-1}(\mu(i)) = \mathcal{E}$ , where  $h^{-1}$  is the inverse of the function defined in (14). Practically, this means that an aggregate event of type  $t$  is satisfiable if there is a granule  $\mu(i)$  such that  $j \in \mu(i), j \models e$  implies  $e \in E[t]$ .

EXAMPLE 5 Let be  $M$  a linear time structure,  $e_1, e_2 \in E[t]$  such that  $i \models e_1$  for  $i \in \{1, 4\}$  and  $i \models e_2$  for  $i \in \{3, 5, 6\}$ . (Remark: restriction (14) means that  $s_1 = s_4$  and  $s_3 = s_5 = s_6$ ). If  $\mu(1) = \{1, 2, 3\}$  and  $\mu(2) = \{4, 5, 6\}$  then

$$\begin{aligned} e[t]_{\mu(1)} &= e[t]_{\mu(2)} = e_1 \boxplus e_2 \\ \mathbf{I}_{\mu(1)}^\mu(e[t]_{\mu(1)}) &= \mathbf{I}_{\mu(1)}^\mu(e[t]_{\mu(2)}) = 2/3 \\ \mathbf{I}_{\mu(2)}^\mu(e[t]_{\mu(1)}) &= \mathbf{I}_{\mu(2)}^\mu(e[t]_{\mu(2)}) = 1 \end{aligned}$$

#### IV. CONCLUSIONS

In this article we developed a formalism for a specific temporal data mining task: the discovery of knowledge, represented in the form of general Horn clauses, inferred from databases with a temporal dimension. The theoretical framework we proposed, based on first-order temporal logic, permits to define the main notions (*event*, *temporal rule*, *constraint*) in a formal way. The concept of a consistent linear time structure allows us to introduce the notions of *general interpretation*, of *support* and of *confidence*, the lasts two measure being the expression of the two similar concepts used in data mining.

Starting from the inherent behavior of temporal systems – the events and their interactions are determined, mainly, by the temporal scale – we extended the capability of our formalism to “capture” the concept of time granularity. To

keep an unitary viewpoint of the meaning of the same formula at different scales of time, we changed the usual definition of the interpretation  $\mathbf{I}^\mu$  for a predicate symbol in the frame of a temporal granular logic: it return the degree of truth (a real value between zero and one) and not only the meaning of truth (*true* or *false*).

The consequence of the definition for  $\mathbf{I}^\mu$  is formalized in theorem 2: only the independent information (here, the temporal free formulae) may be transferred without lost between worlds with different granularities. Concerning the temporal rules (scale dependent information), we proved that the confidence of a rule in a coarser world is linked with the confidence of a similar rule in a finer world, obtained by applying the operator  $zoom_k$  on the initial temporal rule.

We defined also a mechanism to aggregate events of the same type, that reflects the following intuitive phenomenon: in a coarser world, not all events inherited from a finer world are satisfied, but in exchange there are new events which become satisfiable. To achieve this we extended the syntax and the semantics of L by allowing “family” of function symbols and by adding two new operators.

In our opinion, the logical next step in our work consists in adding a probabilistic dimension to the formalism. Preliminary results confirm that this approach allows a unified framework for the initial formalism and its granular extension, framework in which many of the concept definitions become consequences of a fundamental stochastic structure.

#### V. APPENDIX. PROOFS

In the following, if  $S_1, S_2$  are two non-empty sets, the notation  $S_1 \ll S_2$  is equivalent with  $(\forall x \in S_1 \forall y \in S_2) x < y$ . The set of natural numbers,  $\mathbb{N}$ , do not contains zero.

PROPOSITION If  $\mu_1, \mu_2$  are temporal types on  $(\mathbb{N}, \mathbb{N})$  satisfying  $0 < \#\mu_k(i) < \infty$ ,  $k = 1, 2$  and  $i \in \mathbb{N}$ , then  $\mu_1 \rightleftharpoons \mu_2 \Rightarrow \mu_1 = \mu_2$ .

PROOF:  $\mu_1 \rightleftharpoons \mu_2$  is equivalent with the existence of a bijection function  $h : \mathbb{N} \rightarrow \mathbb{N}$  such that  $\mu_1(i) = \mu_2(h(i))$ , for all  $i$ . We will prove by induction that  $h(i) = i$ , which is equivalent with  $\mu_1 = \mu_2$ .

- $i = 1$ : suppose that  $h(1) > 1$ . If  $a = \min(\mu_1(1))$ , then  $\mu_1(1) = \mu_2(h(1)) \Rightarrow a \in \mu_2(h(1))$ . Because  $1 < h(1)$  we have  $\mu_2(1) \ll \mu_2(h(1))$  (see def. 11) and so there is  $b \in \mu_2(1)$  such that  $b < a$ . The inequality  $1 < h(1)$  implies  $h^{-1}(1) > 1$ , and so  $\mu_2(1) = \mu_1(h^{-1}(1)) \gg \mu_1(1)$ . But the last inequality ( $\gg$ ) is contradicted by the existence of  $b \in \mu_1(h^{-1}(1))$  which is smaller than  $a \in \mu_1(1)$ . In conclusion,  $h(1) = 1$ .
- $i = n + 1$ : suppose that  $h(n + 1) > n + 1$ . The rationing is similar to the previous case (it's sufficient to replace 1 with  $n + 1$ ) because  $h(i) = i, \forall i \leq n$  and  $h(n + 1) > n + 1$  implies also  $h^{-1}(n + 1) > n + 1$ . The contradiction of the hypothesis, in this case, means that  $h(n + 1) = n + 1$  and, by induction, that  $h(i) = i, \forall i \in \mathbb{N}$ . ■

Proofs of lemma 1

PROOF: Let be  $\mu \in \mathcal{G}_1, \nu \in \mathcal{G}_1$ .

- $\mu \preccurlyeq \nu$  : let  $j \in \mathbb{N}$ . The condition (3) means that for all  $k \in \nu(j)$ ,  $\mu^{-1}(k) \neq \emptyset$  and so  $S = \{\mu^{-1}(k) \in \mathbb{N} \mid k \in \nu(j)\} \neq \emptyset$ . Because  $\mu \preccurlyeq \nu$  we have  $\mu(i) \subset \nu(j)$ ,  $\forall i \in S$ , which implies  $\bigcup_{i \in S} \mu(i) \subset \nu(j)$  (a). In the same time,  $\forall k \in \nu(j)$  we have  $k \in \mu(\mu^{-1}(k))$  which implies  $\nu(j) \subseteq \bigcup_{i \in S} \mu(i)$  (b). From (a) and (b) we have  $\nu(j) = \bigcup_{i \in S} \mu(i)$ , which implies  $\mu \leq \nu$ .
- $\mu \leq \nu$ : let  $i_0 \in \mathbb{N}$  and let  $k \in \mu(i_0)$ . According to (3), there exists  $j = \nu^{-1}(k)$ . Because  $\mu \leq \nu$  there is a set  $S$  such that  $\nu(j) = \bigcup_{i \in S} \mu(i)$ . Because the sets  $\mu(i), i \in S$  are disjoint and  $k \in \nu(j) \cap \mu(i_0)$  we have  $i_0 \in S$ . Therefore, for each  $i_0$  there is  $j \in \mathbb{N}$  such that  $\mu(i_0) \subseteq \nu(j)$ , which implies  $\mu \preccurlyeq \nu$ . ■

### Proofs of theorem 2

PROOF: According to the definition 13, we have

$$\mathbf{I}_{\nu(i)}^{\nu}(p) = \text{supp}(p, \tilde{M}_{\nu(i)}) = \frac{\#\{j \in \nu(i) \mid j \models p\}}{\#\nu(i)} \quad (a)$$

On the other hand, because  $\nu(i) = \bigcup_{j \in N_i} \mu(j)$ , we have also

$$\begin{aligned} \frac{1}{\#\nu(i)} \sum_{j \in N_i} \#\mu(j) \mathbf{I}_{\mu(j)}^{\mu}(p) &= \frac{1}{\#\nu(i)} \sum_{j \in N_i} \#\mu(j) \text{supp}(p, \tilde{M}_{\mu(j)}) \\ &= \frac{1}{\#\nu(i)} \sum_{j \in N_i} \#\{k \in \mu(j) \mid k \models p\} \\ &= \frac{\#\{j \in \nu(i) \mid j \models p\}}{\#\nu(i)} \quad (b) \end{aligned}$$

From (a) and (b) we obtain (7). ■

### Proofs of theorem 3

PROOF: If  $\mu, \nu \in \mathcal{G}_2$  such that  $\#\mu(i) = c_\mu$  and  $\#\nu(i) = c_\nu$ , for all  $i \in \mathbb{N}$ , it is easy to show that the sets  $N_i$  satisfying  $\nu(i) = \bigcup_{j \in N_i} \mu(j)$  have all the same cardinality,  $\#N_i = c_\nu/c_\mu = k$  and contain successive natural numbers,  $N_i = \{j_i, j_i + 1, \dots, j_i + k - 1\}$ . From the relations (5) and (8) we have:

$$\begin{aligned} \mathbf{I}_{\nu(i)}^{\nu}(p \wedge Xq) &= \frac{1}{2} \left( \mathbf{I}_{\nu(i)}^{\nu}(p) + \mathbf{I}_{\nu(i+1)}^{\nu}(q) \right) = \\ &= \frac{1}{2} \left( \frac{1}{\#N_i} \sum_{j \in N_i} \mathbf{I}_{\mu(j)}^{\mu}(p) + \frac{1}{\#N_{i+1}} \sum_{j \in N_{i+1}} \mathbf{I}_{\mu(j)}^{\mu}(q) \right) = \\ &= \frac{1}{2} \left( \frac{1}{k} \sum_{j=j_i}^{j_i+k-1} \mathbf{I}_{\mu(j)}^{\mu}(p) + \frac{1}{k} \sum_{j=j_i+k}^{j_i+2k-1} \mathbf{I}_{\mu(j)}^{\mu}(q) \right) = \\ &= \frac{1}{2k} \left( \sum_{j=j_i}^{j_i+k-1} \left( \mathbf{I}_{\mu(j)}^{\mu}(p) + \mathbf{I}_{\mu(j+k)}^{\mu}(q) \right) \right) = \\ &= \frac{1}{2k} \left( \sum_{j=j_i}^{j_i+k-1} 2\mathbf{I}_{\mu(j)}^{\mu}(p \wedge X_k q) \right) = \\ &= \frac{1}{k} \sum_{j \in N_i} \mathbf{I}_{\mu(j)}^{\mu}(p \wedge X_k q) \quad \blacksquare. \end{aligned}$$

### REFERENCES

- [1] S. Al-Naemi, "A theoretical framework for temporal knowledge discovery," in *Proceedings of International Workshop on Spatio-Temporal Databases*, Spain, 1994, pp. 23–33.
- [2] X. Chen and I. Petrounias, "A Framework for Temporal Data Mining," *Lecture Notes in Computer Science*, vol. 1460, pp. 796–805, 1998.
- [3] D. Malerba, F. Esposito, and F. Lisi, "A logical framework for frequent pattern discovery in spatial data," in *Proceedings of 5th Conference Knowledge Discovery in Data*, 2001.
- [4] P. Cotofrei and K. Stoffel, "Classification Rules + Time = Temporal Rules," in *Lecture Notes in Computer Science*, vol. 2329. Springer Verlag, 2002, pp. 572–581.
- [5] —, "From temporal rules to temporal meta-rules," in *Proceedings of 6th International Conference Data Warehousing and Knowledge Discovery, DaWaK 2004*, ser. Lecture Notes in Computer Science, vol. 3181, Zaragoza, Spain, 2004, pp. 169–178.
- [6] Y. Yao and N. Zhong, "Potential applications of granular computing in knowledge discovery and data mining," in *Proceedings of World Multiconference on Systemics, Cybernetics and Informatics*, M. Torres, B. Sanchez, and J. Aguilar, Eds. Orlando, Florida, USA: International Institute of Informatics and Systematics, 1999, pp. 573–580.
- [7] Y. Yao, "Granular computing: basic issues and possible solutions," in *Proceedings of the 5th Joint Conference on Information Sciences*, P. Wang, Ed. Atlantic City, New Jersey, USA: Association for Intelligent Machinery, 2000, pp. 186–189.
- [8] L. A. Zadeh, "Information granulation and its centrality in human and machine intelligence," in *Rough Sets and Current Trends in Computing*, 1998, pp. 35–36.
- [9] T. Y. Lin and E. Louie, "Data mining using granular computing: fast algorithms for finding association rules," *Data mining, rough sets and granular computing*, pp. 23–45, 2002.
- [10] J. Hobbs, "Granularity," in *Proceedings of the IJCAI-85*, 1985, pp. 432 – 435.
- [11] J. Clifford and A. Rao, "A simple general structure for temporal domains," in *Temporal Aspects of Information Systems*. Elsevier Science, 1988.
- [12] G.-C. Roman, "Formal specification of geographic data processing requirements," *IEEE Trans. Knowl. Data Eng.*, vol. 2, no. 4, pp. 370–380, 1990.
- [13] C. Evans, "The macro-event calculus: representing temporal granularity," in *Proceedings of PRICAI*, Japan, 1990.
- [14] E. Ciapessoni, E. Corsetti, A. Montanari, and P. S. Pietro, "Embedding time granularity in a logical specification language for synchronous real-time systems," *Sci. Comput. Program.*, vol. 20, no. 1-2, pp. 141–171, 1993.
- [15] C. Bettini, X. S. Wang, and S. Jajodia, "Mining temporal relationships with multiple granularities in time sequences," *Data Engineering Bulletin*, vol. 21, no. 1, pp. 32–38, 1998.
- [16] C. Bettini, X. S. Wang, S. Jajodia, and J.-L. Lin, "Discovering frequent event patterns with multiple granularities in time sequences," *IEEE Trans. Knowl. Data Eng.*, vol. 10, no. 2, pp. 222–237, 1998.
- [17] J. Chomicki and D. Toman, "Temporal Logic in Information Systems," *BRICS Lecture Series*, vol. LS-97-1, pp. 1–42, 1997.
- [18] E. A. Emerson, "Temporal and Modal Logic," *Handbook of Theoretical Computer Science*, pp. 995–1072, 1990.
- [19] C. Bettini, X. S. Wang, and S. Jajodia, "A general framework for time granularity and its application to temporal reasoning," *Ann. Math. Artif. Intell.*, vol. 22, no. 1-2, pp. 29–58, 1998.