We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

4,800
Open access books available

122,000
International authors and editors

135M
Downloads

Our authors are among the

154
Countries delivered to

TOP 1%
most cited scientists

12.2%
Contributors from top 500 universities

CLARIVATE ANALYTICS BOOK CITATION INDEX INDEXED

WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com

24

Temporal Rules Over Time Structures with Different Granularities - a Stochastic Approach

Paul Cotofrei and Kilian Stoffel
Information Management Institute, University of Neuchâtel
Switzerland

1. 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.
It is obvious that there is an implicit relationship between the characteristics of the knowledge extracted from data with temporal dimension and the time scale of the same data. Therefore, a natural question is how (or when) an information discovered based on a given level of time granularity may be reused if this granularity changes. For example, a temporal rule ($\mathbb{T}$) expressed as "If the sales **increase** *today* by 10% and *tomorrow* by 5% then the *day after tomorrow* they will **increase** only by 1%" may be used to make predictions if its degree of confidence remains stable along a time scale using days as basic time granules. But what can we say about the confidence of the rule ($\mathbb{T}_{new}$) obtained by replacing "*today*", "*tomorrow*" and "*day after tomorrow*" with "*this week*", "*next week*" and "*the week after next*", when applied on a time scale using weeks as granules? Is this confidence still well-defined? And if the answer is yes, can its degree be deduced from the confidence of $\mathbb{T}$?
There are at least two important issues related to the conversion of $\mathbb{T}$ into $\mathbb{T}_{new}$ that must be emphasized. First, if the event "**increase**" is in fact the elementary event "**daily increase**" (denoted e), it is quite simple (if daily sales are available) to check the truth of the proposition "e by p% during day n" and, consequently, the truth of the implication/implicated clause of $\mathbb{T}$ for a given day (or time granule). But if we consider the week as the new time granule, we cannot interpret as true or false the proposition "e by p% during week n" and, consequently, not the truth of the implication/implicated clause of $\mathbb{T}_{new}$ either. So the classical definition of confidence for a rule is no more effective for this new time scale and a new, more appropriate definition of confidence must be introduced. Second, if we consider, in the context of week as time granule, the new event "**weekly increase**" (denoted e^*), then we may retrieve the classical definition of confidence for the rule ($\mathbb{T}_{new}$) if it were possible to check the truth of the proposition "e^* by p% during week n". As a remark, the value of truth for this proposition cannot be checked in the initial context (day as time granule) simply because e^* does not exist here. The new event may be seen as an *aggregation* of basic events of type e (considering, for example, the rate of the weekly increase as the mean of all daily increase rates, during the week). Therefore, a formula linking the truth values of the basic events with the truth value of the aggregated event must be introduced. And it is also clear that proposing either a new

definition for the confidence measure or formulae linking the truth value of the same event in worlds with different granularities cannot be made outside of a coherent, logical framework.

1.1 Previous work

Although there is a rich literature concerning the formalism for temporal databases, there are few articles on this topic for temporal data mining. In Al-Naemi (1994); Chen & Petrounias (1998); Malerba et al. (2001), general frameworks for temporal mining are proposed, but usually the research on causal and temporal rules is more concentrated on the methodological/algorithmic aspect and less on the formal aspect. An innovative formalism based on first-order temporal logic, which permits an abstract view on temporal rules, was proposed in Cotofrei & Stoffel (2005). The central concept defined in this formalism is the property of *consistency* for a linear time structure M, which guarantees the preservation over time of the confidence/support of a temporal rule (defined as the limit of a given sequence). The formalism was developed around a time model in which the events are those that describe system evolution.

But the real world systems are systems whose components (events) have dynamic behavior regulated by very different – even by magnitude – time granularities. Analyzing such systems (hereinafter, granular systems) means approaching theories and methodologies that make use of granules (or groups, clusters of a universe) in the process of problem solving. Granular computing (the label which covers this approach) is a way of thinking that relies on our ability to perceive the real world under various grain sizes, to abstract and to consider only those things that serve our present interest, and to switch among different granularities. By focusing on different levels of granularities, we can obtain various levels of knowledge, as well as inherent knowledge structure. Granular computing is essential to human problem solving, and hence has a very significant impact on the design and implementation of intelligent systems, as in Yao & Zhong (1999); Zadeh (1998); Lin & Louie (2002).

The notions of granularity and abstraction are used in many subfields of artificial intelligence. The granulation of time and space leads naturally to temporal and spatial granularities. They play an important role in temporal and spatial reasoning (Euzenat, 1995; Hornsby, 2001; Combi et al., 2004). Based on granularity and abstraction, many authors studied fundamental topics in artificial intelligence, such as knowledge representation (Zhang & Zhang, 1992), search (Zhang & Zhang, 2003), natural language understanding (Mani, 1998) or machine learning (Saitta & Zucker, 1998). Concerning data mining tasks, Bettini et al. (Bettini, Wang, Jajodia & Lin, 1998; Bettini, Wang & Jajodia, 1998a;b) investigated the formal relationships among event structures having temporal constraints, defined the pattern-discovery problem with these structures and studied effective algorithms to solve it.

To include the concept of time granularity in the initial formalism, we defined (Cotofrei & Stoffel, 2009) a process by which a given structure of time granules μ (called temporal type) induces a first-order linear time structure M_μ (called granular world) on the basic (or absolute) linear time structure M. The major change for the temporal logic based on M_μ is at the semantic level: for a formula p, the interpretation does no more assign a meaning of truth (one of the values $\{true, false\}$), but a degree of truth (a real value from $[0,1]$). By an extension at the syntactic and semantic level, we were able to define an aggregation mechanism for events reflecting 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.

A deeper analysis of how these results may be applied to make predictions (i.e., the issues

related to temporal rules evolution due to changes in data time scale) shows the limitation of the pure logical approach to solve some open questions. One of these questions is to what extent is it possible to guarantee the existence of the support for an aggregated event, and consequently, the consistency of a given granular world? A second question is linked to the way in which a temporal rule implying aggregated events can be used to make predictions in a granular world.

Unfortunately, the answers to these questions cannot be given inside the granular temporal logic formalism we developed, because the existence of the limit cannot be proven. The solution we propose in this paper consists of adding a probabilistic dimension to the granular formalism using the stochastic approach introduced in Cotofrei & Stoffel (2007). This extension was put forward in response to the difficulty of checking the consistency property for a linear time structure M (which involves verifying the existence of the support for *each* well-defined formula). By providing a probability system to the set of states S, we could define a stochastic linear time structure such that for each realization of the stochastic sequence $\psi(\omega)$ obtained by randomly drawing a point ω in $S^{\mathbb{N}}$, there is a corresponding (ordinary) linear time structure M_ω. The key answer to the consistency question is the equivalence, as we proved, between the existence of the support for a given formula p and the property of a particular stochastic sequence to obey the strong law of large numbers (SLLN). The sequence corresponding to p (*characteristic sequence*) is constructed, using appropriate transformations, from the stochastic sequence ψ.

This stochastic layer added to the temporal granular logic allows us to define a unified framework (the *stochastic granular time formalism*) in which many of the initially defined concepts become consequences of the properties of a fundamental stochastic structure. Among the results we will prove in the paper based on this formalism we may cite the theorems concerning the existence of the consistency property for any granular time structure induced from a stochastic structure $\mathbb{M}$ under the hypothesis that the random process ψ contains a certain amount of dependence (i.i.d, α-mixing or L_2-NED). These results are stronger than those obtained in the framework of the classical temporal granular logic, due to the fact that in a probabilistic framework we may apply fundamental results which go beyond a simple algebraic manipulation. Furthermore, we could prove the consistency property based only on the requirement that the function giving the interpretation of a temporal formula is a Borel transformation. Concerning the support of an aggregate event, we could establish that the characteristic sequence for this type of event is obtained by applying a particular type of transformation, which asks certain restrictions for the temporal type. Beside these theoretical results, the rationale we followed to prove them emphasizes the main advantage of the stochastic approach: the possibility of inferring, from specific properties of a stochastic process, the existence of different types of consistency, defined based on user necessity.

The structure of the chapter is as follows. In the next section, the main terms (*temporal event, temporal rule*) and concepts (*support, consistency, confidence*) of the first-order temporal logic formalism are described. The definitions and theorems concerning the extension of the formalism towards a temporal granular logic are presented in Sec. 3, whereas the limits of these results from a practical viewpoint and the proposed solution to overcome these issues (the stochastic extension) are described in the following section. Finally, the last section summarizes the work and proposes some promising future developments.

2. Logical 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 (Chomicki & Toman, 1997; Emerson, 1990). The temporal ontology, on which the logical formalism introduced in Cotofrei & Stoffel (2004; 2005) was constructed, is represented by linearly ordered discrete instants. Syntactically, all the terms and formulae are defined over a restricted first-order temporal language $\mathbb{L}$ containing constant symbols, n-ary function symbols, variable symbols $\{y_1, y_2, ...\}$, n-ary predicate symbols ($n \geq 1$), the set of relational symbols $\{=, <, \leq, >, \geq\}$, the logical connective $\wedge$ and a temporal connective of the form ∇_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*.
A Horn clause cannot be expressed in $\mathbb{L}$ because the logical connective $\rightarrow$ is not included. However, to allow the description of rules, which formally look like Horn clauses, a new logical connective, $\mapsto$, was introduced (in practical terms, it is a rewrite of the connective $\wedge$). The next definitions introduce the main types of formulae (based on known concepts from temporal data mining) and the conditions allowing the use of the new connective.

Definition 1 *An event (or temporal atom) is an atom formed by the predicate symbol E followed by a bracketed n-tuple of terms* $(n \geq 1)$ $E(t_1, t_2, \ldots, t_n)$. *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* $t_i = f(t_{i1}, \ldots, t_{ik_i})$.

Definition 2 *A constraint formula for the event* $E(t_1, \ldots t_n)$ *is a conjunctive compound formula,* $E(t_1, t_2, \ldots t_n) \wedge C_1 \wedge \cdots \wedge C_k$. *Each* C_j *is a relational atom* $t \rho c$, *where the term t is one of the terms* $t_i, i = 1 \ldots n$, *the term c is a constant symbol and* ρ *a relational symbol.*

Definition 3 *A temporal rule in standard form is a formula of the form* $H_1 \wedge \cdots \wedge H_m \mapsto H_{m+1}$, *where* H_{m+1} *is a constraint formula prefixed by* ∇_0 *and* $H_i, i = 1..m$ *are constraint formulae, prefixed by the temporal connectives* ∇_{-k}, $k > 0$. *The maximum value of the index k is called the time window of the temporal rule.*

Remark. The reason for which the implication connective was not included in $\mathbb{L}$ 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 all the terms $t_i, i = 1 \ldots n$, from the expression of a temporal atom, constraint formula or temporal rule are represented by variable symbols (and not by constant symbols), then the new formula is denoted a temporal atom template $E(y_1, \ldots, y_n)$ (respectively, a constraint formula template or temporal rule template). These templates are considered as general patterns for events or temporal rules. Practically, the only formulae constructed in $\mathbb{L}$ are temporal atoms, constraint formulae, temporal rules and the corresponding templates.
The semantics of $\mathbb{L}$ is provided by an interpretation $\mathbf{I}$ over a domain $\mathbf{D}$. The interpretation assigns an appropriate meaning over $\mathbf{D}$ to the (non-logical) symbols of $\mathbb{L}$. Based on Definition 1, 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 $t_2, \cdots, t_n$). Consequently, the domain $\mathbf{D}$ is the union $\mathbf{D}_e \cup \mathbf{D}_f$, where the set $\mathbf{D}_e$ contains all the strings used as event names and the set $\mathbf{D}_f$ represents the union of all domains corresponding to chosen features. But a temporal logic cannot be defined without a structure having a temporal dimension and capable of capturing the relationship between a time moment and the interpretation $\mathbf{I}$ at this moment.

Definition 4 *Given* $\mathbb{L}$ *and a domain* $\boldsymbol{D}$*, a (first order) linear time structure is a triple* $M = (S, x, \boldsymbol{I})$*, where S is a set of states,* $x : \mathbb{N} \to S$ *is an infinite sequence of states* $(s_1, s_2, \ldots, s_n, \ldots)$ *and* $\boldsymbol{I}$ *is a function that associates with each state s an interpretation* $\mathbf{I}_s$ *of all symbols from* $\mathbb{L}$*.*

In the framework of a linear temporal logic, the set of symbols is divided into two classes: the class of global symbols (having the same interpretation in each state) and the class of local symbols (the interpretation depends on the state in which they are 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, the meaning of truth for any temporal atom, constraint formula, temporal rule or corresponding template depends on the state at which they are evaluated. Given a first order time structure M and a formula p, the instant i (or equivalently, the state s_i) for which $\mathbf{I}_{s_i}(p) = true$ is denoted by $i \models p$. Therefore, $i \models E(t_1, \ldots, t_n)$ means that at time i an event with the name $\mathbf{I}(t_1)$ and characterized by the global features $\mathbf{I}(t_2), \ldots, \mathbf{I}(t_n)$ occurs. Concerning the event template $E(y_1, \ldots, y_n)$, the interpretation of the variable symbols y_j at the state s_i, $\mathbf{I}_{s_i}(y_j)$, is chosen such that $i \models E(y_1, \ldots, y_n)$ for each time moment i. Finally, $i \models \nabla_k p$ if and only if $i + k \models p$, whereas a temporal rule (template) is true at time i if and only if $i \models H_{m+1}$ and $i \models (H_1 \wedge \cdots \wedge H_m)$.
The connection between the restricted first-order temporal logic and the temporal data mining task this logic tries to formalize (temporal rules extraction) is given by the following assumptions:

A. For each formula p in $\mathbb{L}$, 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.

B. There are states (called incomplete states) that do not contain enough information to calculate the interpretation for all formulae defined at these states.

C. 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 $(s_0, s_1, .., s_n, .)$.

The first assumption expresses the calculability of the interpretation **I**. The second assumption expresses the situation (e.g., due to missing data) 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, the interpretation of the temporal rule cannot be calculated, and the only solution (expressed by the third assumption) is to estimate it using a general interpretation. However, to ensure that this general interpretation is well-defined, the linear time structure must present some property of consistency. In practical terms, this means that the conclusions inferred from a sufficiently large subset of time instants are sufficiently close to those inferred from the entire set of time instants. Therefore,

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

Consequently, if M is a consistent time structure, the general interpretation - seen as a function taking values in $[0,1]$ - is well-defined for any formula p defined in $\mathbb{L}$, by the relation $I_G(p) = \text{supp}(p)$. 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 evaluated

as true) and the number of potential applications (time instants where only the body of the rule is evaluated as true). Furthermore, it can be proved that, for a consistent structure, the confidence is the ratio between the support of the entire rule and the support, not null, of its body.
If, for different reasons (e.g., the states the user has access to are incomplete or are missing), the support measure cannot be calculated, then a possible solution is to estimate $\mathrm{supp}(p)$ using a finite linear time structure, i.e. a model.

Definition 6 *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 $\{i_1, \ldots, i_n\}$, $\tilde{x}$ is the subsequence of states $\{x_{i_1}, \ldots, x_{i_n}\}$ (the restriction of x to the temporal domain $\tilde{T}$) and for each $i_j, j = 1, \ldots, n$, the state x_{i_j} is a complete state.*

This particular structure can be used to obtain an estimator for the support measure ($\mathrm{supp}(p, \tilde{M}) = \#\tilde{T}^{-1}\#\{i \in \tilde{T} \mid i \models p\}$) or for the confidence measure ($\mathrm{conf}(H, \tilde{M}) = \#B^{-1}\#A$, where $A = \{i \in \tilde{T} \mid i \models H_1 \wedge \cdots \wedge H_m \wedge H_{m+1}\}$ and $B = \{i \in \tilde{T} \mid i \models H_1 \wedge \cdots \wedge H_m\}$).
EXAMPLE 1. Consider raw data representing the price variations (six daily records) of a given stock and suppose that a particular methodology for event detection reveals two types of potentially useful events. The output of the methodology is a database of events, where each tuple (v_1, v_2) with record index i expresses the event occurring at time moment i, labeled (v_1 value) with one of the strings $\{increse, decrease\}$ and characterized (v_2 value) by the feature given by the mean of the daily records. In the frame of the temporal logic formalism, the language $\mathbb{L}$ will contain a 2-ary predicate symbol E, two variable symbols y_1, y_2, a 6-ary function symbol f, two sets of constant symbols – $\{d_1, d_2\}$ and $\{c_1, \ldots, c_n\}$ – and the usual set of relational symbols and logical (temporal) connectives. According to the syntactic rules of $\mathbb{L}$, a temporal atom is defined as $E(d_i, f(c_{j_1}, .., c_{j_6}))$, an event template as $E(y_1, y_2)$, whereas a possible temporal rule ($\mathcal{H}$) is

$$E(t_1, t_2) \wedge (t_1 = increase) \wedge (t_2 \leq 5) \mapsto \nabla_1 (E(t_1, t_2) \wedge (t_1 = decrease) \wedge (t_2 \geq 3))$$

("translated" in a natural language as "IF at time t the price increases in average with at most five units THEN at time $t + 1$ the price decreases in average with at least 3 units"). A linear time structure $M = (S, x, \mathbf{I})$ may be defined by considering the set S as the set of all distinct tuples from the event database and the sequence x as the ordered sequence of tuples in the database (see Table 1). At this stage the interpretation of all symbols (global and local) can be defined. For the global symbols, the interpretation is quite intuitive: the meaning $\mathbf{I}(d_1)$ is *increase*, $\mathbf{I}(d_2)$ is *decrease* and $\mathbf{I}(f)$ is the function $f : R^6 \mapsto R, f(x_1, \ldots x_6) = 1/6 \sum_{i=1}^{6} x_i$. For the predicate symbol E, the function $\mathbf{I}_{s_i}(E(t_1, t_2)) : D \to \{true, false\}$ is provided by a finite algorithm, receiving as input the state $s_i = (v_1, v_2)$ and providing as output the value *true* if $\mathbf{I}_{s_i}(t_j) = v_j$ for all $j = 1..2$ and *false* otherwise. If M is a consistent linear time structure having a model $\tilde{M}$ given by the first n states from the sequence x, then the finite structure $\tilde{M}$ can be used to estimate the confidence of the temporal rule $\mathcal{H}$ (based on the first ten states, this estimation is 2/5). And, due to the consistency property, this degree of confidence is reliable information about the prediction power of this rule when applied to future data.

v_1	increase	increase	decrease	increase	increase	increase	decrease	increase	decrease	decrease
v_2	3	5	5	1	8	4	2	4	3	2

Table 1. The first ten states of the linear time structure M

3. The granularity model

The inherent granularity of time implies necessary granular-dependent structure of knowledge extracted by any general temporal data mining methodology. Once the framework in which temporal data is analyzed accepts the dynamic of time scale change, natural questions arise, such as how temporal rule interpretations related to different levels of granularities are connected, or when the consistency property is preserved under time scale changes. In order to find answers to these issues, we extended the initial "static" formalism to include the concept of time granularity (Cotofrei & Stoffel, 2009) by defining a process from which a given structure of time granules μ (called temporal type) induces a first-order linear time structure M_μ (called granular world) on the basic (or absolute) linear time structure M. The concept of a temporal type, formalizing the notion of time granularity, was introduced by Bettini, Wang & Jajodia (1998a).

Definition 7 *Let $(\mathcal{T},<)$ (index) be a linearly ordered temporal domain isomorphic to a subset of 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 μ 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 $\mu(j)$,*
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 μ. Property (1) says that granules do not overlap and that the order on indexes 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.
When considering a particular application or formal context, we can specialize this very general model along in different directions, as the choice of the sets $\mathcal{T}$ or $\mathcal{A}$, or the restrictions on the structure of granules. We call the resulting formalization a *temporal type system*. Also following Bettini, Wang & Jajodia (1998a), a number of interesting relationships between two temporal types, μ and ν, on $(\mathcal{T},\mathcal{A})$, are defined.

A. *Finer-than*: μ is said to be finer than ν, denoted $\mu \preccurlyeq \nu$, if $\forall i \in \mathcal{T}$, $\exists j \in \mathcal{T}$ such that $\mu(i) \subseteq \nu(j)$.
B. *Groups-into*: μ is said to group into ν, denoted $\mu \trianglelefteq \nu$, if $\forall \nu(j) \neq \emptyset$, $\exists S \subset \mathcal{T}$ such that $\nu(j) = \bigcup_{i \in S} \mu(i)$.
C. *Shifting*: μ and ν are said to be shifting equivalent, denoted $\mu_1 \rightleftharpoons \mu_2$, if there is a bijective function $h : \mathcal{T} \to \mathcal{T}$ such that $\mu(i) = \nu(h(i))$, for all $i \in \mathcal{T}$.

When a temporal type μ is finer than a temporal type ν, we also say that ν is *coarser* than μ. The *finer-than* relationship formalizes the notion of finer partitions of the absolute time. This relation is reflexive, transitive, but if no restrictions are given, it is not antisymmetric, and hence it is not a partial order. Considering the *groups-into* relation, $\mu \trianglelefteq \nu$ ensures that for each granule of μ there exists a set of granules of ν covering exactly the same span of time. The *groups-into* relation has the same properties as the *finer-than* relation, but generally $\mu \preccurlyeq \nu$ does not imply $\mu \trianglelefteq \nu$ or vice-versa. Finally, *shifting* is clearly an equivalence relation. But by considering only temporal type systems satisfying the restriction that no pair of different types can be shifting equivalent, we obtain a class of systems for which the relationships $\preccurlyeq$ and $\trianglelefteq$ are partial order, i.e, are reflexive, transitive and antisymmetric.

Let $\mathcal{G}_0$ denote the set of temporal types for which the index set and the absolute time set are isomorphic with the set of positive natural numbers, i.e. $\mathcal{A} = \mathcal{T} = \mathbb{N}$. Consider now the following particular subsets of $\mathcal{G}_0$, represented by temporal types with a) non-empty granules, b) with granules covering all the absolute time and c) with constant size granules:

$$\mathcal{G}_1 = \{\mu \in \mathcal{G}_0 \,|\, \forall i \in \mathbb{N}, 0 < \#\mu(i)\} \tag{1}$$

$$\mathcal{G}_2 = \{\mu \in \mathcal{G}_1 \,|\, \forall i \in \mathbb{N}, \mu(i)^{-1} \neq 0\} \tag{2}$$

$$\mathcal{G}_3 = \{\mu \in \mathcal{G}_2 \,|\, \forall i \in \mathbb{N}, \mu(i) = c_\mu\} \tag{3}$$

The membership of a temporal type defined by one of these subsets implies very useful properties, the most important being:

$$\mu, \nu \in \mathcal{G}_2 \Longrightarrow \mu \preccurlyeq \nu \Leftrightarrow \mu \trianglelefteq \nu. \tag{4}$$

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^{th} position in the sequence). If μ is a temporal type from $\mathcal{G}_2$, then the temporal granule $\mu(i)$ may be identified with the set $\{s_j \in S \,|\, j \in \mu(i)\}$. Therefore, the temporal type μ induces a new sequence, x_μ, defined as $x_\mu : \mathbb{N} \to 2^S$, $x_\mu(i) = \mu(i)$. Consider now the linear time structure derived from M, $M_\mu = (2^S, x_\mu, \mathbf{I}^\mu)$. To be well-defined, we must give the interpretation $\mathbf{I}^\mu_{\mu(i)}$ 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 support measure exists, and we consider, by definition, that for a temporal free formula (which does not contains any ∇ operator, e.g. a temporal atom) p in $\mathbb{L}$

$$\mathbf{I}^\mu_{\mu(i)}(p) = \text{supp}(p, \tilde{M}_{\mu(i)}) \tag{5}$$

This interpretation is extended to any temporal formula in $\mathbb{L}$ according to the rule:

$$\mathbf{I}^\mu_{\mu(i)}(\nabla_{k_1} p_1 \wedge \ldots \wedge \nabla_{k_n} p_n) = \frac{1}{n} \sum_{j=1}^{n} \mathbf{I}^\mu_{\mu(i+k_j)}(p_j) \tag{6}$$

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

Definition 8 *If $M = (S, x, \boldsymbol{I})$ is a first-order linear time structure and μ is a temporal type from $\mathcal{G}_2$, then the linear granular time structure induced by μ on M is the triple $M_\mu = (2^S, x_\mu, \boldsymbol{I}^\mu)$, where $x_\mu : \mathbb{N} \to 2^S$, $x_\mu(i) = \mu(i)$ and $\boldsymbol{I}^\mu$ is a function that associates with almost each set of states $\mu(i)$ an interpretation $\boldsymbol{I}^\mu_{\mu(i)}$ according to rules (5)-(6).*

3.1 Linking two granular structures

All the granular time structures induced by a temporal type have in common interpretations which take values in $[0,1]$ if applied to formulae in $\mathbb{L}$. This observation allows us to establish relationships linking the interpretations $\mathbf{I}^\mu$ and $\mathbf{I}^\nu$, from two linear granular time structures induced by μ and ν, when there exists a relationship *finer-than* ($\mu \preccurlyeq \nu$) between these two temporal types. The key for establishing such a relation is given by property 4, which guarantees that for each $i \in \mathbb{N}$ there is a subset $N_i \subset \mathbb{N}$ such that $\nu(i) = \bigcup_{j \in N_i} \mu(j)$. As we proved in Cotofrei & Stoffel (2009), the capacity to "transfer" information (here, formula interpretation) from a finer world to a coarser one depends strictly on the nature of information.

– The **time independent** part of information may be "transferred" between two granular worlds, i.e. knowing the interpretation of an event in a finer structure allows the calculation of its interpretation (degree of truth) in each coarser structure (see Theorem 1).
– The **time dependent** part of information can't be "transferred" without loss between two granular worlds, and concerns especially the interpretation of temporal rules (see Theorem 2). One consequence of this theorem is that all the information related to temporal formulae having a time window less than k (where k is the coefficient of conversion between the two worlds) is lost during the transition to the coarser world.

Theorem 1 *If μ, ν are temporal types from $\mathcal{G}_2$ such that $\mu \preccurlyeq \nu$, and $\boldsymbol{I}^{\mu}, \boldsymbol{I}^{\nu}$ are the interpretations from the induced linear time structures M_μ and M_ν on M, then $\forall i \in \mathbb{N}$,*

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

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

Theorem 2 *If M_μ, M_ν are granular time structures induced by $\mu, \nu \in \mathcal{G}_3$ (constant size granular worlds), $\mu \preccurlyeq \nu$, then $\forall i \in \mathbb{N}$,*

$$\boldsymbol{I}^{\nu}_{\nu(i)}(\nabla_{k_1} p_1 \wedge ... \nabla_{k_n} p_n) = \frac{1}{k} \sum_{j \in N_i} \boldsymbol{I}^{\mu}_{\mu(j)} \mathbb{Z}_k(\nabla_{k_1} p_1 \wedge ... \nabla_{k_n} p_n)) \tag{8}$$

where $k = c_\nu / c_\mu$, $\nu(i) = \bigcup_{j \in N_i} \mu(j)$, p_i, $i = 1..n$, are temporal free formulae in $\mathbb{L}$ and $\mathbb{Z}_k$ is the operator defined over the set of formulae in $\mathbb{L}$, as $\mathbb{Z}_k(\nabla_{k_1} p_1 \wedge \ldots \wedge \nabla_{k_n} p_n) = \nabla_{k \cdot k_1} p_1 \wedge \ldots \wedge \nabla_{k \cdot k_n} p_n$.

3.2 The consistency problem

The importance of the concepts of consistency, support and confidence (see Sec. 2) for the process of information transfer between worlds with different granularity may be highlighted by analyzing the analogous expressions for a linear granular time structure M_μ.

Definition 9 *Given $\mathbb{L}$ and a linear granular time structure M_μ on M, we say that M_μ is a consistent granular time structure if, for every formula p, the limit $supp(p, M_\mu) = \lim_{n \to \infty} \frac{1}{n} \sum_{i=1}^{n} \boldsymbol{I}^{\mu}_{\mu(i)}(p)$ exists. The notation $supp(p, M_\mu)$ denotes the support (degree of truth) of p under M_μ.*

A natural question concerns the inheritance of the consistency property from the basic linear time structure M by the induced time structure M_μ. The answer is formalized in the following theorem (see (Cotofrei & Stoffel, 2009) for proof).

Theorem 3 *If M is a consistent time structure and $\mu \in \mathcal{G}_3$ then the granular time structure M_μ is also consistent.*

The implications of Theorem 3 are extremely important, because by defining the confidence of a temporal rule $\mathcal{H}$, $H_1 \wedge \ldots \wedge H_m \mapsto H_{m+1}$ over a consistent granular time structure M_μ as:

$$\text{conf}(\mathcal{H}, M_\mu) = \frac{\text{supp}(H_1 \wedge \ldots \wedge H_m \wedge H_{m+1}, M_\mu)}{\text{supp}(H_1 \wedge \ldots \wedge H_m, M_\mu)} \tag{9}$$

we could prove the corollary that the confidence of $\mathcal{H}$, over any granular time structure M_μ induced on a consistent time structure M by a temporal type $\mu \in \mathcal{G}_3$, **exists and is independent of** μ. In other words, the property of consistency is a sufficient condition for the independence of the measure of support/confidence, during the process of information transfer between worlds with different granularities, all derived from an absolute world using constant conversion factors. In practice, this means that even if we are not able to establish, for a given granule $\mu(i)$ in a given world M_μ, the degree of truth for the temporal rule $\mathcal{H}$, we are sure that the confidence of $\mathcal{H}$, given by (9), is the same in each world $M_\mu, \forall \mu \in \mathcal{G}_3$.

3.3 Events aggregation

Another inherent phenomenon accompanying the process of transition between two real worlds with different time granularities is related to the creation of new kinds of significant events. Intuitively, a new event is obtained by applying a kind of "aggregation" mechanism on a set of "similar" events. Formally we need to define the syntax and semantics of these concepts.

We introduce the notion of event type (denoted $E[t]$) as the set of all temporal atoms from $\mathbb{L}$ having the same name (or head). Consider $E(t,t_2,\ldots,t_n) \in E[t]$. According to Definition 1, a term $t_i, i \in \{2,..,n\}$ has the form $t_i = f(t_{i1},\ldots,t_{ik_i})$. 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). $\mathcal{F}_i[t]$ has the property that the interpretation for each of its members is given by a real function which is applied to a variable number of arguments, and is invariant in the order of arguments. A good example of a such real function is a statistical function, e.g. $\text{mean}(x_1,..,x_n)$. Let $T_i[t]$ be the set of terms expressed as $f_k(c_1,\ldots,c_k)$, where f_k is a function symbol from $\mathcal{F}_i[t]$ and c_j are constant symbols. Consider now the following two operators, $\oplus : T_i[t] \times T_i[t] \to T_i[t]$ and $\boxplus : E[t] \times E[t] \to E[t]$ such that:

$$f_n(c_1,..,c_n) \oplus f_m(d_1,..,d_m) = f_{n+m}(c_1,..,c_n,d_1,..,d_m)$$
$$E(t,t_2,..,t_n) \boxplus E(t,t'_2,..,t'_n) = E(t,t_2 \oplus t'_2,..,t_n \oplus t'_n)$$

Obviously the operators $\oplus$ and $\boxplus$ are commutative and associative. Therefore, we can apply the operator $\boxplus$ on a subset $\mathcal{E}$ of temporal atoms from $E[t]$ and denote the result as $\boxplus_{e_i \in \mathcal{E}} e_i$.

By definition, a formula p is satisfied by a linear time structure $M = (S,x,\mathbf{I})$ (or by a model $\tilde{M}$ of M) if there is at least a state $s_i \in x$ (respectively in $\tilde{x}$) such that $\mathbf{I}_{s_i}(p) = true$. Therefore, the set of events of type t satisfied by M is given by $E[t]_M = \{e \in E[t] \,|\, \exists s_i \in x, \mathbf{I}_{s_i}(e) = true\}$. Similarly, the set of events of type t satisfied by M_μ (the structure induced by μ on M) is defined as $E[t]_{M_\mu} = \{e \in E[t] \,|\, \exists \mu_i \in x_\mu, \mathbf{I}^\mu_{\mu(i)}(e) = 1\}$. Generally $E[t]_M \supset E[t]_{M_\mu} \supset E[t]_{M_\nu}$, for $\mu \preccurlyeq \nu$, which is a consequence of the fact that a coarser world satisfies less temporal events than a finer one. At the same time a coarser world may satisfy new events, representing a kind of aggregation of local, "finer" events.

Definition 10 *If $\mu \in \mathcal{G}_2$ then the aggregate event of type t induced by the subset of satisfied events $\mathcal{A} \subset E[t]_M$ (denoted $e[t]_\mathcal{A}$) is the event obtained by applying the operator $\boxplus$ on the set of events from $\mathcal{A}$, i.e.*

$$e[t]_\mathcal{A} = \boxplus_{e_i \in \mathcal{A}} e_i \tag{10}$$

Of a practical interest is the aggregate event induced by the subset $\mathcal{A}$ containing all the events of type t satisfied by a model $\tilde{M}_{\mu(i)}$ (for a given i) denoted $e[t]_{\mu(i)}$. According to (5), the interpretation of an event e in any world M_μ depends on the interpretation of the same event in M. Therefore, if e is not satisfied by M it is obvious that $\mathbf{I}^{\mu}_{\mu(i)}(e) = 0$, for all μ and all $i \in \mathbb{N}$. Because an aggregate event (conceived as a new, "federative" event) is not usually satisfied by M, the relation (5) is not appropriate to give the degree of truth for $e[t]_{\mu(i)}$. By restricting to linear time structures M satisfying the condition that two different events of type t cannot be evaluated as *true* at the same state $s \in S$, the formula expressing the interpretation for an aggregate temporal atom is given by the following definition:

Definition 11 *If M_μ is a linear granular time structure ($\mu \in \mathcal{G}_2$) and $e[t]_{\mathcal{A}}$ is an aggregate event, then the interpretation of $e[t]_{\mathcal{A}}$ in the state $\mu(i)$ is defined as:*

$$\boldsymbol{I}^{\mu}_{\mu(i)}(e[t]_{\mathcal{A}}) = \frac{\#(\mathcal{E}_i \cap \mathcal{A})}{\#\mathcal{A}} \sum_{e_j \in \mathcal{A}} \boldsymbol{I}^{\mu}_{\mu(i)}(e_j) \tag{11}$$

where $\mathcal{E}_i = E[t]_{\tilde{M}_{\mu(i)}}$.

The restriction is necessary to assure that the interpretation of an aggregate event is well-defined, i.e. $\mathbf{I}^{\mu}_{\mu(i)}(e[t]_{\mathcal{A}}) \leq 1$. Furthermore, the interpretation is equal one if and only if all (and only these) satisfied events of type t from $\mathcal{A}$ are also satisfied by $\tilde{M}_{\mu_i}$.

4. Temporal rules in a granular world: toward a stochastic approach

The theoretical framework used to define a linear granular time structure allowed us to prove some nice mathematical results, as the heritage of the consistency property by the worlds with constant granule size, or the independence regarding μ of the confidence measure for a temporal rule over a world M_μ. But if we return to the "real world" and reflect on how these results could be practically applied, we found a number of issues which cannot be avoided. To start our reasoning, let's consider another more simple temporal rule $\mathcal{H}$, defined in the context of Example 1 (according to the described context, the absolute unit time of the linear time structure M is "day"). Suppose that the confidence of this rule (in a natural language, it says that "*IF today there is a daily increase in average of five percent THEN tomorrow will be a daily decrease in average of three percent*") is 0.6. Consider now the granular time structure M_μ, induced on M by $\mu(i) =$ week i (constant granule size equals five).

– **First issue: the temporal rule meaning.** The rule in the new context M_μ (translated as "*IF this week there is a **daily** increase in average of five percent THEN the next week will be a **daily** decrease in average of three percents*") is lacking utility, due to the fact that the events $E(t_1, t_2)$, even if formally theorem 1 permits them to calculate their interpretation for each week granule, reflects well the system behavior only for a time scale using days as basic granules. A "***daily** increase in average of five percent during a week*" has no meaning, except the particular case where in *each* day of the week there is an increase in average of five percent. On the other hand, if during a given week there are three daily *increase* events in average of 5, 8 and 11 percent (and two daily *decrease* events in average of 3 and 1 percents), a natural way to define a meaningful event of type *increase* for a week granule is to consider a "weekly increase" with an average of $(5 + 8 + 11)/3 = 8$ percent (the feature is the mean function), with a degree of truth for this week equal to $3/5$ (respectively a

"weekly decrease" event with an average of $(3+1)/2 = 4$ percent). This is exactly the approach "events aggregation", introduced in Subsec. 3.3, which allows us to consider the new (now meaningful) rule $\mathcal{H}_{new}$: *"IF this week there is a* ***weekly*** *increase in average of eight percent THEN the next week will be a* ***weekly*** *decrease in average of two percent"*.

– **Second issue: predictive power.** If events aggregation (when applicable) represents a solution to obtain meaningful events for a given granular time structure, let's analyse how a temporal rule implying such events can be used to make predictions. In the world M, if the implication clause of $\mathcal{H}$ is true for a particular day, then we expect, with a confidence (probability) of 0.6, that the implicated clause will also be true the next day. The corollary of Theorem 3 assures us that the confidence – defined in (9) as a ratio of the degree of truth for all rule clauses/all implication clauses – of the rule $\mathcal{H}$ over the world M_μ exists. But the fact that in a particular week the implication clauses of $\mathcal{H}_{new}$ have a degree of truth equal to (lets say) 0.4 and that the confidence is (lets say) 0.8 does not allow us to infer any information about the degree of truth for the implicated clause. A useful way to employ the rule would be a rationing of type *"If the degree of truth for the implication clauses of $\mathcal{H}_{new}$ is at least (lets say) 0.7 for a particular week then we expect, with a confidence of 0.8, that the degree of truth of the implicated clause to be at least (lets say) 0.6 the next week"*. In order to be able to apply this inference schema we need a new definition for the confidence measure over a granular world, seen as the limit ratio between the number of granules where both the implication clauses and the implicated clause have a degree of truth greater than some chosen constants and the number of granules where only the implication clauses have a degree of truth greater than the chosen constant.

Definition 12 *Given M_μ a linear time structure on M and $\mathcal{H}$ a temporal rule $H_1 \wedge .. \wedge H_m \mapsto H_{m+1}$, the confidence (α,β) of $\mathcal{H}$, denoted $conf_{\alpha,\beta}(\mathcal{H},M_\mu)$, is the limit (if exists) $\lim_{n\to\infty}(\#B)^{-1}\#A$, where $A = \{i \leq n | \boldsymbol{I}^{\mu}_{\mu(i)}(H_1 \wedge .. \wedge H_m) \geq \alpha, \boldsymbol{I}^{\mu}_{\mu(i)} H_{m+1} \geq \beta\}$ and $B = \{i \leq n | \boldsymbol{I}^{\mu}_{\mu(i)}(H_1 \wedge .. \wedge H_m) \geq \alpha\}$, where $\alpha,\beta \in (0,1]$.*

At this moment two questions naturally arise: (1) If M_μ is consistent, does the support of an aggregate event exist? (2) If M_μ is consistent, does the confidence (α,β) for a temporal rule exist? Unfortunately, the answers to these questions cannot be given inside the granular temporal logic formalism we developed because the existence of the limits can't be proved. A possible solution is to extend this pure logical formalism using the stochastic approach introduced in Cotofrei & Stoffel (2007). This extension was proposed in response to the difficulty of checking the consistency property for a linear time structure M (which involves verifying the existence of the support for *each* well-defined formula), by deriving the consistency as an objective consequence of a specific property of a stochastic process.

4.1 The stochastic model

The key of the stochastic extension for a first order time structure $M = (S,x,\mathbf{I})$ is given by the observation that the sequence x may be considered as a particular realization of a stochastic process. Technically, this can be done by providing a probability system $(S,\sigma(S),P)$ for the set of states S. Indeed, if $S = \{s_0,s_1,\ldots\}$ is a countable set of states, consider $\sigma(S)$ the $\sigma-$algebra generated by S. The probability measure P on $\sigma(S)$ is defined such that $P(s_i) = p_i > 0, \forall i \in \mathbb{N}$. Consider now a random variable $\mathbb{X} : S \to \mathbb{R}$ such that the probability $P(\mathbb{X} = s_i) = p_i$ for all $i \in \mathbb{N}$. If $S^{\mathbb{N}} = \{\omega \,|\, \omega = (\omega_1,\omega_2,\ldots,\omega_t,\ldots), \omega_t \in S, t \in \mathbb{N}\}$, then the variable $\mathbb{X}$ induces the stochastic sequence $\psi : S^{\mathbb{N}} \to \mathbb{R}^{\mathbb{N}}$, where $\psi(\omega) = \{\mathbb{X}_t(\omega), t \in \mathbb{N}\}$ and $\mathbb{X}_t(\omega) = \mathbb{X}(\omega_t)$ for all

$t \in \mathbb{N}$. The fact that each $\omega \in S^{\mathbb{N}}$ may be uniquely identified with a function $x : \mathbb{N} \to S$ and that $\mathbb{X}$ is a bijection between S and $\mathbb{X}(S)$ allows us to uniquely identify the function x with a single realization of the stochastic sequence. In other words, the sequence $x = (s_{(1)}, s_{(2)}, \ldots, s_{(i)}, \ldots)$ from the structure M can be seen as one of the outcomes of an infinite sequence of experiments, each experiment being modelled by the probabilistic system $(S, \sigma(S), P)$.

Definition 13 *Given $\mathbb{L}$ and a domain D, a stochastic (first order) linear time structure is a quintuple $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$, where*

- *$S = \{s_1, s_2, \ldots\}$ is a (countable) set of states,*
- *P is a probability measure on the $\sigma-$algebra $\sigma(S)$ such that $P(s_i) = p_i > 0, i \in \mathbb{N}$*
- *$\mathbb{X}$ is a random variable such that $P(\mathbb{X} = s_i) = p_i$,*
- *ψ is a random sequence, $\psi(\omega) = \{\mathbb{X}(\omega_i)\}_1^{\infty}$ where $\omega \in S^{\mathbb{N}}$,*
- *$\boldsymbol{I}$ is a function that associates with each state s an interpretation $\boldsymbol{I}_s$ for all symbols from $\mathbb{L}$.*

To each realization of the stochastic sequence ψ, obtained by random drawing of a point in $\mathbb{R}^{\infty}$ (or equivalently, of a point ω in $S^{\mathbb{N}}$), corresponds a realization of the stochastic structure $\mathbb{M}$. This realization is given by the (ordinary) linear time structure $M_\omega = (S, \omega, \mathbf{I})$, which implies that the semantics attached to the symbols of $\mathbb{L}$, described in Section 2, is totally effective. Moreover, if p is a formula defined in language $\mathbb{L}$ and A_p the event[1] "the interpretation of the formula p is true", then

$$\overline{\mathbf{1}_{A_p}}_n(\omega) = \frac{\sum_{t=1}^{n} \mathbf{1}_{A_p}(\omega_i)}{n} = \frac{\#\{i \leq n \,|\, \mathbf{1}_{A_p}(\omega_i) = 1\}}{n} = \frac{\#\{i \leq n \,|\, \mathbf{I}_{s_{(i)}}(p) = true\}}{n} \tag{12}$$

where the last term is exactly the expression which gives, at the limit, the support of p. Consequently, ***supp(p) exists (almost sure) if the stochastic sequence $\{\mathbf{1}_{A_p}\}_1^{\infty}$ satisfies the strong law of large numbers***.

To obey the law of large numbers, a sequence must satisfy regularity conditions relating to two distinct factors: the probability of extreme values (limited by bounding absolute moments) and the degree of dependence between coordinates. The necessity of a set of regularity conditions is usually hard to prove (except if the sequences are independent), but various configurations of dependency and boundedness conditions can be shown to be sufficient. The characteristic sequences, which are derived from the stochastic process ψ using appropriate Borel transformation (depending on p), have all absolute moments bounded by 0 and 1. Therefore, the only regularity condition which may vary (and which is inherited from ψ) is the degree of dependence. After a deeper analysis of the various types of dependence restrictions a stochastic process may contain, we proved (Cotofrei & Stoffel, 2007) the following results:

- **Independence and Consistency.** *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$ is i.i.d., then almost all linear time structures $M_\omega = (S, \omega, \boldsymbol{I}_s)$ are consistent.* But the independence condition represents a serious drawback for any temporal rule extraction methodology, because it implies a null correlation between the body and the head of a rule, i.e. not at all meaningful temporal rules.

[1] In this context, an event is a set of possible outcomes of a random experiment

– **Dependence and Consistency.** *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$ is α-mixing[2] or is $L_2 - NED$[3], then almost all linear time structures $M_\omega = (S, \omega, \boldsymbol{I}_s)$ are consistent.* Therefore, with up to a certain amount of dependence between events (that makes the rules meaningful), it is possible to guarantee the "correctness" of the temporal rules (expressed by the confidence measure) when applied to future data.

4.2 The granularity stochastic model

If we apply the stochastic approach, as developed in the previous subsection, to the granularity model, we obtain what we call a stochastic granular time formalism. Let $\psi = \{\mathbb{X}_i\}_1^\infty$ be a stochastic process and μ a temporal type. If we denote $\mathbf{X}_{\mu(i)}$ the random vector $(\mathbb{X}_{j_1}, \ldots, \mathbb{X}_{j_k})$, where $j_i, i = 1..k$ are all the indices from $\mu(i)$, then the random sequence induced by μ on ψ is simply $\mu[\psi] = \{\mathbf{X}_{\mu(i)}\}_{i=1}^\infty$. Similarly, if $\omega \in S^{\mathbb{N}}$ then $\omega_{\mu(i)} = (\omega_{j_1}, \ldots, \omega_{j_k})$ and $\mu[\omega] = \{\omega_{\mu(i)}\}_1^\infty$. Therefore, we define a stochastic granular time structure as:

Definition 14 *If $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$ is a stochastic (first-order) linear time structure and μ is a temporal type from $\mathcal{G}_1$, then the stochastic granular time structure induced by μ on $\mathbb{M}$ is the quintuple $\mathbb{M}_\mu = (2^S, P, \mathbb{X}, \mu[\psi], \boldsymbol{I}^\mu)$, where $\boldsymbol{I}^\mu$ is given by (5)-(6).*

Practically, the random process $\mu[\psi]$ from the stochastic granular time structure M_μ is a sequence of random vectors obtained by grouping the coordinates of the process ψ according to the mapping μ. To each realization of the stochastic sequence ψ, obtained by a random drawing of a point ω in $S^{\mathbb{N}}$, corresponds to a realization of the stochastic structure $\mathbb{M}$ (i.e., the time structure $M_\omega = (S, \omega, \mathbf{I}_\omega)$) and a corresponding realization of the stochastic structure $\mathbb{M}_\mu$ (i.e., the granular time structure $M_{\mu[\omega]} = (2^S, \mu[\omega], \mathbf{I}^\mu_\omega)$).
In the following we establish the expression linking the interpretation $\mathbf{I}^\mu$ of a given formula in $\mathbb{L}$ with the random process $\mu[\psi]$. For this we introduce the function $\overline{\mathcal{S}}$ defined by $\overline{\mathcal{S}}(\mathbf{X}_{\mu(i)}) = (\#\mu(i))^{-1} \sum_{j \in \mu(i)} \mathbb{X}_j$. If $\{\mathbb{X}_i\}$ are identical distributed, with $E(\mathbb{X}_i) = \gamma$, then it is evident that $E(\overline{\mathcal{S}}(\mathbf{X}_{\mu(i)})) = \gamma$, for all $i \in \mathbb{N}$. Consider the following two situations:

– *Temporal free formula*: According to (5) and to (12),

$$\mathbf{I}^\mu_{\mu[\omega](i)}(p) = \mathrm{supp}(p, \tilde{M}_{\mu[\omega](i)}) = \overline{\mathcal{S}}\left((\mathbf{1}_{A_p})_{\mu[\omega](i)}\right). \tag{13}$$

– *Temporal formula*: According to (6) and (13), for a temporal formula $\nabla_{k_1} p_1 \wedge \ldots \wedge \nabla_{k_n} p_n$

$$\mathbf{I}^\mu_{\mu[\omega](i)}(\nabla_{k_1} p_1 \wedge \ldots \wedge \nabla_{k_n} p_n) = \frac{1}{n} \sum_{j=1}^n \mathbf{I}^\mu_{\mu[\omega](i+k_j)}(p_j) = \frac{1}{n} \sum_{j=1}^n \overline{\mathcal{S}}\left((\mathbf{1}_{A_{p_j}})_{\mu[\omega](i+k_j)}\right). \tag{14}$$

The consequence of these relations is that ***the support of any formula p under $M_{\mu[\omega]}$ exists if and only if the characteristic sequence $\{\overline{\mathcal{S}}(\mathbf{1}_{A_p})_{\mu[\omega]}\}$ satisfies the strong law of large numbers***. The sequence corresponding to p is constructed by applying a mapping μ and a particular Borel transformation on the stochastic sequence ψ. By analyzing the sufficient conditions (the dependence degree) for ψ which assure, through the transformation function, the applicability of SLLN for any characteristic sequence, we arrived at the following results:

[2] the degree of dependence converges to zero if the distance between coordinates converges to ∞
[3] a function of a mixing sequence with an infinite number of parameters

Theorem 4 ***(Independence and Consistency)*** *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$ is i.i.d., then almost all granular time structures induced by a temporal type $\mu \in \mathcal{G}_1$, $M_{\mu[\omega]} = (2^S, \mu[\omega], \boldsymbol{I}^\mu)$, are consistent.*

(For proof see Appendix). This result is stronger than those obtained in Theorem 3, where the temporal type has to satisfy a more restrictive condition, i.e. $\mu \in \mathcal{G}_3$. This is explained by the fact that in a probabilistic framework we can apply fundamental results which go beyond a simple algebraic manipulation. Furthermore, we can prove that the consistency is preserved even if we replace the function giving the interpretation of a temporal formula (the arithmetic mean, see (6)) with any t-norm transformation.

Theorem 5 ***($\alpha-$Mixing and Consistency)*** *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$ is $\alpha-$mixing, then almost all granular time structures induced by a temporal type $\mu \in \mathcal{G}_2$, $M_{\mu[\omega]} = (2^S, \mu[\omega], \boldsymbol{I}^\mu)$, are consistent.*

(For proof see Appendix). This result is, once again, stronger than those obtained in the pure granular logical formalism, but we must remark on the supplementary condition imposed on μ (now in $\mathcal{G}_2$) compared with the independence case.

Theorem 6 ***(Near-Epoch Dependence and Consistency)*** *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M} = (S, P, \mathbb{X}, \psi, \boldsymbol{I})$ is L_2-NED on an $\alpha-$mixing sequence, then almost all granular time structures induced by a temporal type $\mu \in \mathcal{G}_3$, $M_{\mu[\omega]} = (2^S, \mu[\omega], \boldsymbol{I}^\mu)$, are consistent.*

(For proof see Appendix). For the near-epoch dependence case we were forced to impose the stronger restriction to the temporal type μ (constant size and total coverage) to compensate the higher degree of dependence of the stochastic process ψ. This type of dependence is, according to the stochastic limit theory (Davidson, 1994; Davidson & de Jong, 1997), the highest degree of dependence for which theorems concerning SLLN still hold.

4.2.1 Aggregated event support and (α, β) confidence

All these results were obtained by analyzing the characteristic sequences for the formulae constructed in $\mathbb{L}$, for which the interpretation is given by the expressions (5)-(6). The transformations applied to the process ψ to generate characteristic sequences belong to a family of Borel functions $\mathcal{G}$,

$$g_p(\mathbb{X}_{i+k}(\omega)) = (g_p \circ \mathbb{X}_{i+k})(\omega) = \begin{cases} 1 & \text{if } \omega_{i+k} \in A_p, \\ 0 & \text{if not} \end{cases} \quad \text{for all } k \geq 0 \tag{15}$$

(for independence and $\alpha-$mixing dependence) and to a slightly different family $\tilde{\mathcal{G}}$, satisfying *i)* $\tilde{g}_p(\mathbb{X}_i(\omega)) = g_p(\mathbb{X}_i(\omega))$, *ii)* $\tilde{g}_p$ continuous and iii) $|\tilde{g}(\boldsymbol{X}^1) - \tilde{g}(\boldsymbol{X}^2)| \leq M \sum_{i=1}^n |x_i^1 - x_i^2|$ a.s., where $\boldsymbol{X}^1$, $\boldsymbol{X}^2$ are random vectors from $\mathbb{R}^n$ (for near-epoch dependence).
The degree of truth of an aggregate event is given by a different rule (11), implying that a particular type of transformation ($\mathcal{T}$) must be applied to ψ to obtain the corresponding characteristic sequence. To start the rationale, let ω be a sequence of states generated by the process ψ and $e[t]_{\mathcal{A}}$ an aggregate event induced by the set $\mathcal{A}$. The expression given at the limit

the support of $e[t]_{\mathcal{A}}$ in the granular time structure M_μ is

$$\frac{1}{n}\sum_{i=1}^{n}\mathbf{I}^{\mu}_{\mu[\omega](i)}(e[t]_{\mathcal{A}})=\frac{1}{n}\sum_{i=1}^{n}\left(\frac{\#(\mathcal{E}_i\cap\mathcal{A})}{\#\mathcal{A}}\sum_{e_j\in\mathcal{A}}\mathbf{I}^{\mu}_{\mu[\omega](i)}(e_j)\right)=\frac{1}{\#\mathcal{A}}\sum_{e_j\in\mathcal{A}}\left(\frac{1}{n}\sum_{i=1}^{n}\#(\mathcal{E}_i\cap\mathcal{A})\mathbf{I}^{\mu}_{\mu[\omega](i)}(e_j)\right)$$

For a fixed j and according to (13), $\mathbf{I}^{\mu}_{\mu[\omega](i)}(e_j)=\overline{\mathcal{S}}\left((\mathbf{1}_{A_{e_j}})_{\mu[\omega](i)}\right)$, which represents the i^{th} coordinate of a random sequence $(X_i^j)_1^\infty$ of variables obtained from ψ by applying a Borel transformation (an application of the mean function on functions from family $\mathcal{G}$). Consider now the sequence $(\mathbf{1}_{A_{e_j}})_1^\infty$, $N_i^j=\sum_{k\in\mu[\omega](i)}\mathbf{1}_{A_{e_j}}(\omega_k)$ (the variable counting the number of times e_j is satisfied in $\mu[\omega](i)$) and $gt(\cdot)$ a function defined as

$$gt(N_i^j)=\begin{cases}1 & \text{if } N_i^j\geq 1,\\ 0 & \text{if not}\end{cases}$$

Consequently we have $\#(\mathcal{E}_i\cap\mathcal{A})=\sum_{k\in\mu[\omega](i)}gt(N_i^k)$, which represents the i^{th} coordinate of a second random sequence $(Y_i)_1^\infty$, obtained again from ψ by applying a Borel transformation. The variables Y_i^j are identical distributed if and only if $\#\mu[\omega](i)$ is a constant, which implies $\mu\in\mathcal{G}_3$. Under this restriction, the characteristic sequence for the aggregate event $e[t]_{\mathcal{A}}$ (which is $\left(\frac{1}{\#\mathcal{A}}\sum_{e_j\in\mathcal{A}}Y_i\cdot X_i^j\right)_1^\infty$ and is obtained from ψ by applying a particular Borel transformation) inherits the dependence degree of ψ. Therefore, the answer to our first question is given by the following theorem:

Theorem 7 *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M}=(S,P,\mathbb{X},\psi,\mathbf{I})$ is i.i.d or α-mixing or L_2-NED on an $\alpha-$mixing sequence, then in almost all granular time structures induced by a temporal type $\mu\in\mathcal{G}_3$, an aggregated event has a support.*

Concerning the confidence (α,β) for a temporal rule, let introduces a particular type of measure, the (α)support for a formula p under M_μ, denoted $\text{supp}_\alpha(p,M_\mu)$ and defined as $\lim_{n\to\infty}n^{-1}\#\{i\leq n\,|\,\mathbf{I}^{\mu}_{\mu(i)}(p)\geq\alpha\}$. Furthermore, following a similar rationing used in the proofs of theorems 4-6 and treating the two cases (p temporal free formula and p temporal rule), we can prove the possibility of constructing a specific characteristic sequence - denoted $\left((X_p^\alpha)_i\right)_1^\infty$ - derived from ψ by applying a specific Borel transformation, such that the existence of $\text{supp}_\alpha(p)$ is guaranteed by the capacity of the specific sequence to obey SLLN. Because this capacity is assured only if $\mu\in\mathcal{G}_3$, we can assert that

Theorem 8 *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M}=(S,P,\mathbb{X},\psi,\mathbf{I})$ is i.i.d or α-mixing or L_2-NED on an $\alpha-$mixing sequence, then in almost all granular time structures induced by a temporal type $\mu\in\mathcal{G}_3$, an aggregated event has an (α)support, for any $\alpha\in(0,1]$*

The expression given $\text{conf}_{\alpha,\beta}(\mathcal{H},M_\mu)$ may be rewritten as:

$$\lim_{n\to\infty}\frac{\#\{i\leq n|\mathbf{I}^{\mu}_{\mu(i)}(H_1\wedge..\wedge H_m)\geq\alpha,\mathbf{I}^{\mu}_{\mu(i)}H_{m+1}\geq\beta\}}{\#\{i\leq n|\mathbf{I}^{\mu}_{\mu(i)}(H_1\wedge..\wedge H_m)\geq\alpha\}}=\lim_{n\to\infty}\frac{\left(\overline{X^{\alpha}_{H_1\wedge..\wedge H_m}\cdot X^{\beta}_{H_{m+1}}}\right)_n}{\left(\overline{X^{\alpha}_{H_1\wedge..\wedge H_m}}\right)_n}$$

so the existence of the confidence(α,β) is directly related to the existence of a non null $\text{supp}_\alpha(H_1 \wedge .. \wedge H_m)$ and of the property of the sequence $\left((X^\alpha_{H_1\wedge..\wedge H_m} \cdot X^\beta_{H_{m+1}})_i\right)_1^\infty$ to obey SSLN (which is assured by Theorem 8). In conclusion, the answer to our second question is:

Theorem 9 *If the random process ψ from the stochastic first-order linear time structure $\mathbb{M} = (S,P,\mathbb{X},\psi,\boldsymbol{I})$ is i.i.d or α-mixing or L_2-NED on an $\alpha-$mixing sequence then in almost all granular time structures induced by a temporal type $\mu \in \mathcal{G}_3$, a temporal rule $(\mathcal{H})$, $H_1 \wedge .. \wedge H_m \mapsto H_{m+1}$ for which $supp_\alpha(H_1 \wedge .. \wedge H_m) \neq 0$ has a well-defined confidence $conf_{(\alpha,\beta)}(\mathcal{H},M_\mu)$, $\forall \alpha,\beta \in (0,1]$.*

The rationale we followed to prove the existence of the confidence (α,β) for a temporal rule emphasizes the main advantage of the stochastic approach: the possibility to infer, from specific properties of a stochastic process, the existence of *different types of consistency*, defined based on user necessity. For a temporal granular structure M_μ, as an example, the consistency may be defined either in the classical way (Definition 9) or as the existence, for any formula p, of $\text{supp}_\alpha(p,M_\mu)$. And depending on the user's needs for other types of confidence measures, other concepts of consistency may be defined (of course, under the hypothesis that SLLN still implies these new concepts). Therefore, by considering all the consistency concepts introduced in this chapter we could prove that:

Theorem 10 *If the random process ψ from the stochastic linear time structure $\mathbb{M} = (S,P,\mathbb{X},\psi,\boldsymbol{I})$ is i.i.d, α-mixing or L_2-NED, then almost all granular time structures induced by a temporal type $\mu \in \mathcal{G}_3$, are consistent.*

From a practical point of view, after testing (see Cotofrei & Stoffel (2007) for a discussion about possible statistical tests) that the sequence ω of states, derived from raw data, contains a certain amount of dependence, this theorem assures us that any temporal rule $H_1 \wedge .. \wedge H_m \mapsto H_{m+1}$ (implying *any* type of defined temporal events - aggregated events included), for which the support of $H_1 \wedge .. \wedge H_m$ is not null, has a well-defined (but not μ independent) confidence over any M_μ with $\mu \in \mathcal{G}_3$.

5. Conclusions

Starting from the inherent behavior of temporal systems - the perception of events and of their interactions is determined, in a large measure, by the temporal scale - the question about the mechanisms of transferring (transforming) discovered knowledge from a finer time scale to a coarser one is naturally imposed. We approached this question using a theoretical framework based on first-order temporal logic and extended to "capture" the concept of time granularity. The introduced formalism allows us to define main notions such as *event, temporal rule, support and confidence* in a formal way, based on the fundamental concept of consistency for a linear time structure M.

To keep a unitary viewpoint on the semantics of the same formula at different scales of time, the usual definition of the interpretation $\mathbf{I}^\mu$ for a formula was changed: now it returns the degree of truth (a real value between zero and one) and not only the meaning of truth (*true* or *false*). Based on the concept of consistency extended to granular time structures, we could prove that this property is inherited from the basic time structure M if the temporal type μ is of type $\mathcal{G}_3$ (granules with constant size). The major consequence of this theorem is that a given form of confidence, expressed by (9), is preserved in all granular time structures derived from the same consistent time structure.

By reflecting on how the changes of the time scale affect the meaning (in the "real world") of the temporal rules, we could emphasize an intrinsic connection between the significance of an event for a user and the granularity of the time. Indeed, any methodology which extracts events from raw temporal data acts based on an implicit granularity (usually given by the time scale of raw data). Therefore, all the generated temporal events have a specific meaning *only* for this initial time scale, and any change in the time granularity implies the "loss" of this meaning. Our solution to this problem was the introduction of the concept of "event aggregation", a mechanism generating new events with an appropriate significance and satisfied in a coarser world. To achieve this we extended the syntax and the semantics of the language $\mathbb{L}$ by allowing "families" of function symbols and by adding two new operators. Due to the limitations in proving the existence of the support for aggregate events and of the new introduced confidence (α, β) (allowing the use of temporal rules for prediction purposes under a granular world), we extended our formalism by a stochastic dimension. In this framework, using the relation between the capacity of a stochastic process to obey the strong law of large numbers and the consistency property, we proved that under a given amount of dependency (which implies meaningful rules), the existence of the confidence for any temporal rule is guaranteed (which implies preserving the predictive power of the rule on any future data sets).
In our opinion, the conclusion of our analysis may be summarized as follows: *only the fundamental properties (knowledge) concerning the time structures conceived as a whole may be transferred (preserved) during a granularity time change process. On the other hand, the information linked to a granule, seen as "local knowledge" cannot be transferred during the same process (or if it can, it's a meaningless transfer).*

6. Appendix

6.1 Proof of theorem 4

If ψ is an i.i.d. process, then for p a temporal free formula the sequence $\{\mathbf{1}_{A_p}\}_1^\infty$ is also i.i.d. By applying Pfeiffer (1989, page 255) Theorem , the vectors $(\mathbf{1}_{A_p})_{\mu(i)}$ are independent, and consequently, according to the Theorem (Pfeiffer, 1989, page 254) and to the fact that the function $\overline{\mathcal{S}}$ is a Borel transformation, the sequence $\left\{\overline{\mathcal{S}}\left((\mathbf{1}_{A_p})_{\mu[\omega](i)}\right)\right\}_{i=1}^{\infty}$ is independent. Therefore, the classical Kolmogorov theorem may be applied, and so the support of the formula p, under the granular time structure $M_{\mu[\omega]}$, exists almost sure. For the temporal formula $\nabla_{k_1} p_1 \wedge \ldots \wedge \nabla_{k_n} p_n$, similar considerations assure that, for a fixed i, the random variables $\overline{\mathcal{S}}\left((\mathbf{1}_{A_{p_1}})_{\mu[\omega](i+k_1)}\right), \ldots, \overline{\mathcal{S}}\left((\mathbf{1}_{A_{p_n}})_{\mu[\omega](i+k_n)}\right)$ are independent. The sequence corresponding to the temporal formula (see 14) is not independent, but k_n-dependent, and so the conditions of the Theorem (Hall & Heyde, 1980, page 40) are satisfied. As a consequence, this sequence obeys the law of large numbers, i.e. the support of the temporal formula exists.

6.2 Proof of theorem 5

If ψ is α-mixing then it is evident that any subsequence of ψ is also α-mixing. The following result, necessary for our rationale, is a consequence of the fact that mixing is a property of σ-fields generated by $\{\mathbb{X}_i\}$.

Lemma 1 *Consider $\mathbb{X}_i$ an α-mixing sequence of size $-\varphi$ and let be k sequences ${}_j\mathbb{Y}_i$ obtained by applying on $\{\mathbb{X}_i\}$ the measurable functions $g_j(\mathbb{X}_t, \ldots, \mathbb{X}_{t-\tau_j})$, $j = 1 \ldots k$. Then the sequence*

${}_1\mathbb{Y}_{i_1}, {}_2\mathbb{Y}_{i_2}, \ldots, {}_k\mathbb{Y}_{i_k}, {}_1\mathbb{Y}_{i_{k+1}}, \ldots$, obtained by tacking successively from each sequence ${}_j\mathbb{Y}_i$ coordinates with indices in an increasing order, is also α-mixing of size $-\varphi$.

The utility of this lemma is due to the fact that the granules of a temporal type from $\mathcal{G}_2$ have a variable size, and so we cannot apply a single measurable function $g(\cdot)$, with a fixed number of parameters, on $\{\mathbf{1}_{A_p}\}$. By considering for each effective size $k \in \mathbb{N}$ the function $mean_k(x_1, \ldots, x_k) = k^{-1}\sum x_i$ and applying Lemma 1 on $\{\mathbf{1}_{A_p}\}$ we obtain that $\overline{\mathcal{S}}\left((\mathbf{1}_{A_p})_{\mu[\omega](i)}\right)$, p a temporal free formula, is $\alpha-$mixing. Concerning a temporal formula $\nabla_{k_1} p_1 \wedge \ldots \wedge \nabla_{k_n} p_n$, by applying n times Lemma 1 for the $\alpha-$mixing sequences $\{\mathbf{1}_{A_{p_j}}\}$, $j = 1 \ldots n$, we obtain the $\alpha-$mixing sequences $\overline{\mathcal{S}}\left((\mathbf{1}_{A_{p_j}})_{\mu[\omega](i)}\right)$, $j = 1 \ldots n$. From these sequences we extract the subsequence $\overline{\mathcal{S}}\left((\mathbf{1}_{A_{p_1}})_{\mu[\omega](i+k_1)}\right), \ldots, \overline{\mathcal{S}}\left((\mathbf{1}_{A_{p_n}})_{\mu[\omega](i+k_n)}\right)$, $i \in \mathbb{N}$ (which is $\alpha-$mixing, according to the same Lemma), on which we apply the function $g_n(\cdot)$. The resulting sequence is again $\alpha-$mixing, according to the Theorem (Davidson, 1994, page 210). Finally, the corresponding sequence for any formula in L is $\alpha-$mixing, bounded by the interval $[0,1]$, thus fulfilling the conditions of Theorem (Hall & Heyde, 1980, page 40).

6.3 Proof of theorem 6

According to a corollary proved in Cotofrei & Stoffel (2007), any sequence $\{\mathbf{1}_{A_p}\}$ is also L_2-NED on the same sequence $\{\mathbb{V}_i\}$. If $\#\mu(i) = k$ then it is easy to show that the function $mean_k(\cdot)$ is continuous and satisfies the uniform Lipschitz condition. Therefore, according to the Theorem (Davidson, 1994, page 269), the sequence corresponding to the temporal free formula p, $\mathcal{S}\left((\mathbf{1}_{A_p})_{\mu[\omega](i)}\right)$, is also L_2-NED on $\{\mathbb{V}_i\}$. The same theorem, applied to the sequence of vectors $\left(\mathcal{S}\left((\mathbf{1}_{A_{p_1}})_{\mu[\omega](i+k_1),}\right) \ldots, \mathcal{S}\left((\mathbf{1}_{A_{p_n}})_{\mu[\omega](i+k_n)}\right)\right)$, all L_2-NED on $\{\mathbb{V}_i\}$, and for the Lipschitz function $mean_n(\cdot)$, assures that the sequence $\frac{1}{n}\sum_{j=1}^{n} \mathcal{S}\left((\mathbf{1}_{A_{p_j}})_{\mu[\omega](i+k_j)}\right)$ is L_2-NED on $\{\mathbb{V}_i\}$. Therefore, for any formula in L the corresponding sequence is L_2-NED on the α-mixing sequence $\{\mathbb{V}_i\}$. Furthermore, these sequences fulfil the conditions of the Theorem (Davidson & de Jong, 1997, page 258) for $q = 2$ and so obey the strong law of large numbers.

7. References

Al-Naemi, S. (1994). A theoretical framework for temporal knowledge discovery, *Proc. of Int. Workshop on Spatio-Temporal Databases*, Spain, pp. 23–33.

Bettini, C., Wang, X. S. & Jajodia, S. (1998a). A general framework for time granularity and its application to temporal reasoning., *Ann. Math. Artif. Intell.* 22(1-2): 29–58.

Bettini, C., Wang, X. S. & Jajodia, S. (1998b). Mining temporal relationships with multiple granularities in time sequences, *Data Engineering Bulletin* 21(1): 32–38.

Bettini, C., Wang, X. S., Jajodia, S. & Lin, J.-L. (1998). Discovering frequent event patterns with multiple granularities in time sequences, *IEEE Trans. Knowl. Data Eng.* 10(2): 222–237.

Chen, X. & Petrounias, I. (1998). A Framework for Temporal Data Mining, *Lecture Notes in Computer Science* 1460: 796–805.

Chomicki, J. & Toman, D. (1997). Temporal Logic in Information Systems, *BRICS Lecture Series* LS-97-1: 1–42.

Combi, C., Franceschet, M. & Peron, A. (2004). Representing and reasoning about temporal granularities, *J. Log. Comput.* 14(1): 51–77.

Cotofrei, P. & Stoffel, K. (2004). From temporal rules to temporal meta-rules, *LNCS, vol 3181*, pp. 169–178.

Cotofrei, P. & Stoffel, K. (2005). First-order logic based formalism for temporal data mining, *Fundation of Data Mining and Knowledge Extraction*, Vol. 6 of *Studies in Computational Intelligence*, pp. 185–210.

Cotofrei, P. & Stoffel, K. (2007). Stochastic processes and temporal data mining, *Proceedings of the 13th KDD*, San Jose, USA, pp. 183–190.

Cotofrei, P. & Stoffel, K. (2009). *Foundations in Computational Intelligence*, Springer Verlang, chapter Time Granularity in Temporal Data Mining, pp. 67–96.

Davidson, J. (1994). *Stochastic Limit Theory*, Oxford University Press.

Davidson, J. & de Jong, R. (1997). Strong law of large numbers for dependent and heterogenous processes: a synthesis of new and recent results, *Econometric Reviews* 16: 251–279.

Emerson, E. A. (1990). Temporal and Modal Logic, *Handbook of Theoretical Computer Science* pp. 995–1072.

Euzenat, J. (1995). An algebraic approach to granularity in qualitative time and space representation., *IJCAI (1)*, pp. 894–900.

Hall, P. & Heyde, C. (1980). *Martingale Limit Theory and Its Application*, Academic Press.

Hornsby, K. (2001). Temporal zooming, *Transactions in GIS* 5: 255–272.

Lin, T. Y. & Louie, E. (2002). Data mining using granular computing: fast algorithms for finding association rules, *Data mining, rough sets and granular computing*, Physica-Verlag GmbH, pp. 23–45.

Malerba, D., Esposito, F. & Lisi, F. (2001). A logical framework for frequent pattern discovery in spatial data, *Proceedings of FLAIRS*, pp. 557 – 561.

Mani, I. (1998). A theory of granularity and its application to problems of polysemy and underspecification of meaning, *Proceedings of the Sixth International Conference Principles of Knowledge Representation and Reasoning*, pp. 245–255.

Pfeiffer, P. (1989). *Probability for Applications*, Springer-verlag.

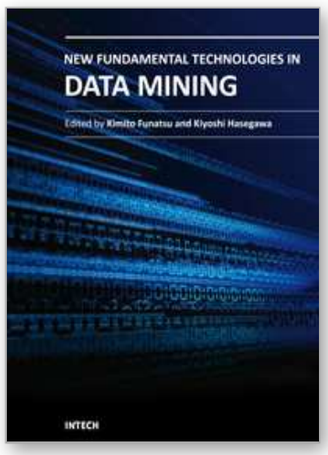
Saitta, L. & Zucker, J.-D. (1998). Semantic abstraction for concept representation and learning., *Proc. of the Symp. on Abstraction, Reformulation and Approximation*, pp. 103–120.

Yao, Y. & Zhong, N. (1999). Potential applications of granular computing in knowledge discovery and data mining, *Proceedings of WMSCI*, Orlando, pp. 573–580.

Zadeh, L. A. (1998). Information granulation and its centrality in human and machine intelligence., *Rough Sets and Current Trends in Computing*, pp. 35–36.

Zhang, B. & Zhang, L. (1992). *Theory and Applications of Problem Solving*,, North-Holland, Amsterdam.

Zhang, L. & Zhang, B. (2003). The quotient space theory of problem solving, *Proceedings of International Conference on Rough Sets, Fuzzy Set, Data Mining and Granular Computing*, pp. 11–15.



New Fundamental Technologies in Data Mining

Edited by Prof. Kimito Funatsu

ISBN 978-953-307-547-1
Hard cover, 584 pages
Publisher InTech
Published online 21, January, 2011
Published in print edition January, 2011

The progress of data mining technology and large public popularity establish a need for a comprehensive text on the subject. The series of books entitled by "Data Mining" address the need by presenting in-depth description of novel mining algorithms and many useful applications. In addition to understanding each section deeply, the two books present useful hints and strategies to solving problems in the following chapters. The contributing authors have highlighted many future research directions that will foster multi-disciplinary collaborations and hence will lead to significant development in the field of data mining.

How to reference

In order to correctly reference this scholarly work, feel free to copy and paste the following:

Paul Cotofrei and Kilian Stoffel (2011). Temporal Rules over Time Structures with Different Granularities - a Stochastic Approach, New Fundamental Technologies in Data Mining, Prof. Kimito Funatsu (Ed.), ISBN: 978-953-307-547-1, InTech, Available from: http://www.intechopen.com/books/new-fundamental-technologies-in-data-mining/temporal-rules-over-time-structures-with-different-granularities-a-stochastic-approach

INTECH
open science | open minds

InTech Europe
University Campus STeP Ri
Slavka Krautzeka 83/A
51000 Rijeka, Croatia
Phone: +385 (51) 770 447
Fax: +385 (51) 686 166
www.intechopen.com

InTech China
Unit 405, Office Block, Hotel Equatorial Shanghai
No.65, Yan An Road (West), Shanghai, 200040, China
中国上海市延安西路65号上海国际贵都大饭店办公楼405单元
Phone: +86-21-62489820
Fax: +86-21-62489821

IntechOpen