# A String-based Model for Infinite Granularities (Extended Abstract) 

Jef Wijsen<br>Université de Mons-Hainaut<br>Jef.Wijsen@umh.ac.be


#### Abstract

In the last few years, the concept of time granularity has been defined by several researchers, and a glossary of time granularity concepts has been published. These definitions often view a time granularity as a (mostly infinite) sequence of time granules. Although this view is conceptually clean, it is extremely inefficient or even practically impossible to represent a time granularity in this manner. In this paper, we present a practical formalism for the finite representation of infinite granularities. The formalism is string-based, allows symbolic reasoning, and can be extended to multiple dimensions to accommodate, for example, space.


## Introduction

In the last few years, formalisms to represent and to reason about temporal and spatial granularity have been developed in several areas of computer science. Although several researchers have used different definitions of time granularity, they commonly agree upon the overall view of a time granularity as a possibly infinite sequence of time granules, where a time granule is a subset of some fixed time domain. In (Bettini et al. 1998; Bettini, Wang, \& Jajodia 1998; Wang et al. 1997), for example, a time granularity is defined as a mapping from integer numbers to subsets of the time domain. In (Wijsen 1999), we defined a time granularity as a possibly infinite but computable partition of the natural numbers. Although all these definitions are conceptually clean, they mostly do not address one important and practical question: How do we represent these infinite structures in a finite way that is amenable to manipulation by a computer system? Clearly, the complexity of many problems involving granularities depends on such effective representations (Wijsen 1998).

In the time granularity glossary (Bettini et al. 1998), this problem is partially addressed by the concept of "groups periodically into." The following example is slightly adapted from that glossary. The following granularities $G$ and $H$ are such that $H$ can be finitely represented in terms of $G$.

Copyright (C) 2000, American Association for Artificial Intelligence (www.aaai.org). All rights reserved.


It suffices to specify that the first granule of $H$ consists of the first two granules of $G$, that the second granule of $H$ coincides with the third granule of $G$, and that this pattern repeats every three granules of $G$. (Bettini $\&$ De Sibi 1999) further investigates the expressiveness of extensions to this basic formalism.

A challenging and important issue that deserves more attention is how to reason efficiently about such periodical granularities. For example, given some granularity lattice and a second time granularity $H^{\prime}$ defined in terms of $G$, can we easily compute the supremum and the infimum of $H$ and $H^{\prime}$ ? In this paper, we study in more depth such periodic patterns and related reasoning problems.

We propose a new, simple syntax for effectively describing infinite time granularities. The syntax uses three symbols: called filler, $U$ called gap, and $/$ called separator. Granules are constructed from fillers and gaps, and are delimited by separators. The repeating pattern of $H$ introduced above is denoted 1 【III. The gap symbol $\sqcup$ is used to denote gaps. It has been recognized that gaps are needed to represent some common real-life time granularities. A typical example is the granularity "business week," which contains a repetition of five-day periods separated by weekend gaps. Moreover, our formalism accounts for a finite "offset" preceding the infinite repetition of the finite pattern. In this way, one can model, for example, that business weeks used to contain 6 working days before the Saturday became a free day.

Figure 1 shows a comprehensive example. Time instant 1 represents a Thursday, time instant 2 the Friday after, and so on. For the time granularity $J$, from time instant 3 on, one observes a repeating pattern consisting of a weekend followed by a five-day working week; every working week has a gap on Tuesday and Wednes-
 The infinite repetition is preceded by an offset ©il? We represent a granularity as an ordered pair where the first component is the offset, and the second the


Figure 1: Representation of granularity.
repeating pattern. For the current example, this yields:

Such a representation is called a granspec. Although it is common practice to let the repeating pattern coincide with granule boundaries, our formalism allows repeating patterns that start or end in the middle of a granule. For example, the time granularity represented in Figure 1 can also be described as a repetition


In the latter granspec, the end of the repeating pattern does not coincide with the end of a granule. So the same granularity can be presented by different granspecs. This raises some interesting questions that will be addressed in this paper: What is the shortest granspec? How can we verify whether two distinct granspecs actually represent the same granularity? Other questions concern the reasoning about time granularities in the proposed formalism. In particular, we are interested in performing certain reasoning tasks by the symbolic manipulation of granspecs.

Time granularity constructs found in the literature are typically one-dimensional. We indicate how our formalism can be extended to multiple dimensions. Such extra dimensions could be temporal or spatial. For example, temporal databases often use two time dimensions, called transaction time and valid time. Spatiotemporal databases use spatial and temporal dimensions. Still other dimensions are typical in OLAP (Wijsen \& Ng 1999). Another extension concerns moving from infinite strings to infinite trees, which may be more natural for representing different granularity levels. Such multi-layered temporal structures have been investigated in the area of temporal logic (Montanari, Peron, \& Policriti 1999).
One may argue that in many practical situations, the problem of infiniteness can be circumvented by fixing some end date far ahead, say December 31, 2100 . However, this line of reasoning misses two important points. First, the representation with fixed end date may not be more concise, and second, the end date complicates certain reasoning tasks because it results in a special, incomplete last granule for certain granularities.

This paper is organized as follows. We first propose our notion of granularity. Regular granularities are then defined as a class of granularities that can be described by granspecs. In general, the same granularity can be
represented by different granspecs. We therefore introduce a unique canonical granspec, which turns out to be the most concise description. We indicate how certain computations on regular granularities can be performed by symbolic manipulation of granspecs. Finally, we indicate two possible extensions. In the appendix, we give two theorem proofs. Interestingly, in terms of automata theory, a granspec induces an ultimately periodic word on the alphabet $\{\mathbf{\bullet}, \sqcup, \imath\}$. This link with the theory of combinatorics on words is interesting, as it allows the application of well-established results (Theorem 13, for example).

## Granularity

We view time as isomorphic to the naturals. The set of natural numbers $\{1,2,3, \ldots\}$ is denoted $\mathbb{N}$. Our starting point is a classical, set-theoretic approach to time granularity (Clifford \& Rao 1987), where a given granularity constitutes a partition of the set of granules of a finer granularity. But unlike other approaches, we take as definition for granularity not the partition itself, but the corresponding equivalence relation that can be constructed from it (Wijsen 1998; 1999). More precisely, we define a granularity as a possibly infinite equivalence relation on (some subset of) the naturals, in which the equivalence classes are not interleaved w.r.t. the natural order. For example, the equivalence classes for the time granularity of Figure 1 are $\{1,2\},\{3,4\},\{5,8,9\},\{10,11\},\{12,15,16\}, \ldots$ These are the equivalence classes; the time granularity itself is the corresponding equivalence relation, which contains, for example, the pair $(5,8)$ but not $(5,10)$, meaning that 5 and 8 belong to the same granule, but 5 and 10 do not.
Definition 1 A granularity is an equivalence relation $R$ on some subset $N \subseteq \mathbb{N}$ such that whenever $E_{1}, E_{2} \subseteq$ $N$ are two distinct equivalence classes, then either

1. for every $i \in E_{1}$, for every $j \in E_{2}, i<j$, or
2. for every $i \in E_{1}$, for every $j \in E_{2}, j<i$.

There is no explicit "index set" in this formalism, but such index set can be easily derived. For the granularity of Figure 1, for example, we obtain $1 \mapsto\{1,2\}, 2 \mapsto$ $\{3,4\}, 3 \mapsto\{5,8,9\}, \ldots$ As we argued in previous work, the absence of an explicit index set simplifies certain reasoning issues. In particular, the common relation "finer than" between time granularities coincides with
set inclusion. The set of all granularities ordered by set inclusion is a lattice.
Theorem 1 The set of all granularities, ordered by $\subseteq$, is a lattice. If $R_{1}, R_{2}$ are two granularities, then $\inf \left\{R_{1}, R_{2}\right\}=R_{1} \cap R_{2}$ and $\sup \left\{R_{1}, R_{2}\right\}=\bigcap\{R \subseteq$ $\mathbb{N} \times \mathbb{N} \mid R$ is a granularity and $\left.R_{1} \cup R_{2} \subseteq R\right\}$.
Note: the set of all granularities is closed for intersection, and the intersection of two granularities coincides with the infimum. The supremum does not coincide with union, however; the set of all granularities is not closed for union. The lattice of granularities is infinite. Other studies often assume a finite partial order of granularities (Dyreson et al. 1998), corresponding to real-life calendars.

## Granspecs

A granspec is an ordered pair of finite strings over the alphabet $\{\mathbf{\bullet}, \sqcup, \ell\}$, the first string representing the offset, and the second the repeating pattern.
Every granspec defines an infinite string. Formally, the relation between granspecs and the infinite strings they represent is established by a function ${ }^{\circ}$, such that if $\alpha$ is any granspec, then $\alpha^{\infty}$ is the infinite string represented by $\alpha$. We say that $\alpha^{\infty}$ is the trace produced

 produced by the granspec $\beta=$ (III; பi!』). In this case, we say that $\alpha$ and $\beta$ are trace-equal, denoted $\alpha \equiv_{\mathrm{T}} \beta$.
Definition 2 We fix the alphabet $\{\mathbf{I}, \omega, \imath\}$. The symbols 【, ப, and \ are called filler, gap, and separator respectively. The symbols and $L$ are both nonseparators. A string is a (possibly infinite) sequence of symbols from this alphabet. If $w$ is a string, we denote by $w[i, j]$ the string $w(i) w(i+1) \ldots w(j)$. The empty string is denoted $\varepsilon$. String concatenation is defined as usual. If $w$ is a finite string and $k \in \mathbb{N}$, then $w^{k}=\overbrace{w w \ldots w}^{k \text { times }}$ and $w^{0}=\varepsilon$. The length of a finite string $w$ is denoted $|w|$. If $v$ is a finite string, then $\lceil v \rrbracket$ denotes the number of occurrences of $!$ and $\sqcup$ in $v$; that is $\llbracket v \rrbracket$ denotes the cardinality of the set $\{i \in \mathbb{N}|i \leq|v|$ and $v(i) \neq\{ \}$. The last symbol of a non-empty, finite string $w$ is denoted $\operatorname{tail}(w)$.

A granspec is a pair $(v ; w)$ where $v$ and $w$ are finite strings and $\llbracket w \rrbracket \geq 1$. We call $v$ the offset and $w$ the repeating part of the granspec ( $v ; w$ ).
The trace produced by the granspec ( $v ; w$ ), denoted $(v ; w)^{\infty}$, is equal to the infinite string $v w w w \ldots$ Two granspecs $\alpha$ and $\beta$ are said to be trace-equal, denoted $\alpha \equiv \mathrm{T} \beta$, iff $\alpha^{\infty}=\beta^{\infty}$.
Every string with infinitely many non-separator symbols represents a granularity. Formally, the relation between infinite strings and the granularities they represent is established by a function $\operatorname{gran}(\cdot)$, such that if $t$ is any string with infinitely many non-separator symbols, then $\operatorname{gran}(t)$ is the granularity represented by $t$.

Intuitively, $\operatorname{gran}(t)$ contains $(i, j)$ if and only if the $i^{\text {th }}$ non-separator symbol in $t$ is a filler, the $j^{\text {th }}$ nonseparator symbol is a filler as well, and there is no separator between these two fillers. For example, the string

## 

produced by the granspec $\alpha=(\mathbf{1} ; \boldsymbol{\square} \mathrm{l})$ introduced above, represents the granularity with as equivalence classes $\{1\},\{2,3\},\{5,6\},\{8,9\}, \ldots$ The same granularity is represented by the trace

## 

produced by the granspec $\gamma=$ (1) IU case, we say that $\alpha$ and $\gamma$ are gran-equal, denoted $\alpha \equiv_{\mathrm{G}}$ $\gamma$. Note incidentally $\alpha \not \equiv_{\mathrm{T}} \gamma$. In general, substituting ¿ $\mathrm{\cup}$ for U , or vice-versa, in a string does not change the granularity represented.
Definition 3 If $t$ is an infinite string and $k \in \mathbb{N}_{0}$, then $\square(t, k)$ denotes the number of occurrences of 1 or $u$ among the first $k$ symbols of $t$. That is, $\square(t, k)$ is equal to $\rrbracket t[1, k] \rrbracket$. This notation is also used for finite strings.

Every string $t$ with infinitely many non-separator symbols induces a granularity, denoted $\operatorname{gran}(t)$, as follows. For all $i, j \in \mathbb{N},(i, j) \in \operatorname{gran}(t)$ iff there exists $k, l \in \mathbb{N}$ such that

1. $t(k)=t(l)=\mathbf{~}$,
2. $k-\square(t, k)=l-\square(t, l)$, i.e., no separator occurs between positions $k$ and $l$, and
3. $\square(t, k)=i$ and $\square(t, l)=j$.

Every granularity that can be expressed as $\operatorname{gran}\left(\alpha^{\infty}\right)$ for some granspec $\alpha$, is called a regular granularity. Two granspecs $\alpha$ and $\beta$ are said to be gran-equal, denoted $\alpha \equiv_{\mathrm{G}} \beta$, iff $\operatorname{gran}\left(\alpha^{\infty}\right)=\operatorname{gran}\left(\beta^{\infty}\right)$.
For example, the following table shows the string (call it $t$ ) introduced in Figure 1:

| $i-\square(t, i):$ | 0 | 0 | 1 | 1 | 1 | 2 | 2 | 2 | 2 | 2 |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: | :---: |
| $\square(t, i)$ : | 1 | 2 | 2 | 3 | 4 | 4 | 5 | 6 | 7 | 8 |  |
| $t(i):$ | - | 1 | 1 | - | - | 1 | - | ப | $\pm$ | - |  |
| $i$ : | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 |  |

We have $t(7)=t(10)=$ and $7-\square(t, 7)=10-$ $\square(t, 10)=2$, i.e., no separator occurs between positions 7 and $10 . \square(t, 7)=5$ and $\square(t, 10)=8$, hence $(5,8) \in$ $\operatorname{gran}(t)$.
Theorem 2 If $t$ is a string with infinitely many nonseparator symbols, then gran( $(t)$ is indeed a granularity.
For a granularity to be regular, it must show a linearly repeating pattern from some time instant on. This seems to be the case for all common real-life granularities, like days, business weeks, weeks, months, years, and so on. The granularity with equivalence classes $\left\{1, \ldots, 2^{2}\right\},\left\{3^{2}, \ldots, 4^{2}\right\},\left\{5^{2}, \ldots, 6^{2}\right\}, \ldots$, for example, is not regular.

In general, the same regular granularity can be represented by several granspecs, as we showed above. If


Figure 2: Overview of equivalences.
$\alpha$ and $\beta$ are two granspecs, then $\alpha=\beta$ implies $\alpha \equiv_{\mathrm{T}} \beta$, and $\alpha \equiv \equiv_{\mathrm{T}} \beta$ in turn implies $\alpha \equiv_{\mathrm{G}} \beta$. The inverse, however, is not true. That is, $\alpha \equiv_{\mathrm{G}} \beta$ does not imply $\alpha \equiv_{\mathrm{T}} \beta, \alpha \equiv_{\mathrm{T}} \beta$ does not imply $\alpha=\beta$, and $\alpha \equiv_{\mathrm{G}} \beta$ does not imply $\alpha=\beta$. The following theorem gives some transformations that preserve $\equiv_{\mathrm{G}}$ or $\equiv_{\mathrm{T}}$.
Theorem 3 Let $v, v^{\prime}, w, w^{\prime}$ be placeholders for a (possibly empty) string, and $i \in \mathbb{N}_{0} .^{1}$

$$
\begin{array}{rll}
\left(v \imath \imath v^{\prime} ; w\right) & \equiv_{\mathrm{G}} & \left(v \imath v^{\prime} ; w\right) \\
\left(v ; w l \imath w^{\prime}\right) & \equiv_{\mathrm{G}} & \left(v ; w \imath w^{\prime}\right) \\
\left(v \sqcup^{i} \imath v^{\prime} ; w\right) & \equiv_{\mathrm{G}} & \left(v \imath \sqcup^{i} v^{\prime} ; w\right) \\
\left(v ; w \sqcup^{i} \imath w^{\prime}\right) & \equiv_{\mathrm{G}} & \left(v ; w \imath \sqcup^{i} w^{\prime}\right) \\
\left(v ; \imath \sqcup^{i}\right) & \equiv_{\mathrm{G}} & (v ; \sqcup) \\
(\imath v ; w) & \equiv_{\mathrm{G}} & (v ; w) \\
\left(v ; \imath w \imath \sqcup^{i}\right) & \equiv_{\mathrm{G}} & \left(v ; \imath w \sqcup^{i}\right) \\
\left(v \imath \sqcup^{i} ; \imath w\right) & \equiv_{\mathrm{G}} & \left(v \sqcup^{i} ; \imath w\right) \\
\left(v \imath \sqcup^{i} ; \sqcup^{j}\right) & \equiv_{\mathrm{G}} & (v ; \sqcup) \\
\left(v w^{\prime} ; w w^{\prime}\right) & \equiv_{\mathrm{T}} & \left(v ; w^{\prime} w\right) \\
\left(v ; w^{k}\right) & \equiv_{\mathrm{T}} & (v ; w) \tag{11}
\end{array}
$$

We now address the following problem: Decide whether two given granspecs $\alpha$ and $\beta$ are gran-equal. To this extent, we will define a unique canonical form, such that $\alpha$ and $\beta$ are gran-equal if and only if they share the same canonical form. The canonical form is introduced in two steps. First, we define aligned granspecs as granspecs for which every separator is directly preceded by a filler in the produced trace. We show that if two granspecs $\alpha$ and $\beta$ are aligned, then $\alpha \equiv_{\mathrm{G}} \beta$ implies $\alpha \equiv_{\mathrm{T}} \beta$. In a second step, canonical granspecs are defined as a restricted class of aligned granspecs. The steps are summarized in Figure 2.

## Alignment

We say that a granspec $\alpha$ is aligned if every occurrence of $\}$ in $\alpha^{\infty}$ is immediately preceded by an occurrence of $\llbracket$, and eventually followed by an occurrence of $\boldsymbol{\square}$.
Definition 4 A granspec $\alpha$ is aligned iff it satisfies the following conditions:

[^0]1. for every $i \in \mathbb{N}$, if $\alpha^{\infty}(i)=$ l then for some $j>i$, $\alpha^{\infty}(j)=\llbracket ;$
2. $\alpha^{\infty}(1) \neq$; and
3. for every $i \in \mathbb{N}$ with $i>1$, if $\alpha^{\infty}(i)=1$ then $\alpha^{\infty}(i-1)=$.

Theorem 4 A granspec $(v ; w)$ is aligned iff

1. $v(1) \neq l$;
2. for every $i$ with $1<i \leq|v|$, if $v(i)=$ ! then $v(i-1)=$ ■;
3. if $w(1)=\ell$ then $v \neq \varepsilon$ and $\operatorname{tail}(v)=\operatorname{tail}(w)=\mathbf{~ ; ~}$
4. for every $i$ with $1<i \leq|w|$, if $w(i)=$ 2 then $w(i-1)=\mathbf{I}$; and
5. if $w=\sqcup^{i}$ for some $i>0$ then $\left.v \neq v^{\prime}\right\} \sqcup^{j}$ for all strings $v^{\prime}$ and $j \geq 0$.
The following theorem states that alignment can always be achieved.
Theorem 5 For every granspec $\alpha$, there exists an aligned granspec $\beta$ such that $\alpha \equiv_{\mathrm{G}} \beta$.
Finally, we obtain the desired result.
Theorem 6 Let $\alpha$ and $\beta$ be two aligned granspecs. If $\alpha \equiv_{\mathrm{G}} \beta$ then $\alpha \equiv_{\mathrm{T}} \beta$.
The proof of the latter theorem is given in the appendix.

## Canonical

Two aligned granspecs can be distinct, and still be gran-equal (and hence trace-equal by Theorem 6). We now introduce a canonical form such that two distinct canonical granspecs represent distinct granularities. In particular, a canonical granspec is an aligned granspec whose offset and repeating part do not end with the same symbol, and whose repeating part is not itself a repetition of some smaller pattern.
Definition 5 A granspec ( $v ; w$ ) is canonical iff

1. it is aligned,
2. $w=u^{k}$ for some $u$ implies $u=w$ and $k=1$, and
3. if $v \neq \varepsilon$ then $\operatorname{tail}(v) \neq \operatorname{tail}(w)$.

A canonical form can always be achieved.
Theorem 7 (Existence.) For every granspec $\alpha$, there exists a canonical granspec $\beta$ such that $\alpha \equiv_{\mathrm{G}} \beta$.
Proof. Let $\alpha$ be a granspec. By Theorem 5, there exists an aligned granspec $\alpha^{\prime}$ such that $\alpha \equiv_{\mathrm{G}} \alpha^{\prime}$. By repeated application of (10) and (11), we can obtain from $\alpha^{\prime}$ a canonical granspec $\alpha^{\prime \prime}$ such that $\alpha^{\prime \prime} \equiv \mathrm{T} \alpha^{\prime}$.

For example,

| granspec | canonical form |
| :---: | :---: |
|  |  |
| ( $\left.\left.\square^{3} \sqcup^{2}\right\urcorner ; \square^{5} \sqcup^{2} \ell\right)$ | $\left(\left\llcorner; \square^{3} \sqcup^{2} \square^{2}\right)\right.$ |
|  | $\left(\varepsilon ; \square^{2} \^{3} \square^{3}\right)$ |

The proof of the following theorem is given in the appendix. It states the desired result that every regular granularity is represented by a unique canonical granspec.
Theorem 8 (Uniqueness.) Let $\alpha$ and $\beta$ be two canonical granspecs. If $\alpha \equiv_{\mathrm{G}} \beta$ then $\alpha=\beta$.
Consequently, it can be easily checked whether two granspecs are gran-equal: Rewrite both granspecs in canonical form and test whether these canonical representations are identical.

How does the length of a canonical granspec compare to its non-canonical but gran-equal representations? The alignment requirement may lengthen certain granspecs. For example, ( $\sqcup$; 1 ■) is not aligned; each aligned granspec that is gran-equal to this granspec,
 It is the case, however, that a canonical granspec is the shortest among all gran-equal aligned granspecs.
Theorem 9 Let $(v ; w)$ be a canonical granspec, and ( $v^{\prime} ; w^{\prime}$ ) an aligned granspec. If $(v ; w) \equiv_{G}\left(v^{\prime} ; w^{\prime}\right)$ then $|v| \leq\left|v^{\prime}\right|$ and $|w| \leq\left|w^{\prime}\right|$.
To conclude, if an aligned granspec is not canonical, then its offset can be shortened by moving one or more of its rear-end symbols to the repeating part, or its repeating part can be shortened because it is itself a repetition of some shorter pattern. Interestingly, our results show that there is no other possibility.

## Symbolic Computation

Theorem 1 states that the set of all granularities ordered by "finer than" is a lattice. This lattice turns out to be significant in reasoning about temporal functional dependencies (Wang et al. 1997; Wijsen 1999). Other reasoning problems may also involve the computation of lower and upper bounds in this lattice. For example, suppose we are given, for two production processes $P$ and $Q$, the number of failures. For production process $P$, the number of failures is reported for every business week, for example,

| business week index | 1 | 2 | 3 | 4 | $\ldots$ |
| :--- | :---: | :---: | :---: | :---: | :---: |
| number of failures | 2 | 1 | 3 | 0 | $\ldots$ |

For process $Q$, on the other hand, the number of failures is given for every month. We want to know, given some time interval, the total number of failures during this interval, i.e., the number of failures in $P$ during this interval plus the number of failures in Q during this interval. Clearly, for most time intervals, it is impossible to know the exact total. Nevertheless, even though months do not divide evenly into business weeks, there may still be time intervals that evenly divide into months and business weeks, and for which an exact sum can thus be obtained. These intervals are exactly the granules of the supremum of business week and month. How can we effectively construct this supremum?

One can show that the set of all regular granularities ordered by set inclusion, which coincides with finer
than semantics, is a sublattice of the lattice shown in Theorem 1. The bottom and top elements are represented by $(\varepsilon ; \sqcup)$ and $(\varepsilon ; \mathbb{C})$ respectively. Theorem 1 does not provide us with an effective way to construct the supremum. We show how our formalism allows the computation of the supremum of two regular granularities by symbolic manipulation of their granspecs.

## Combining Infinite Strings

We observe that an occurrence of $u$ can be of two kinds: it can occur within a granule, or between two granules. For example, consider the string
which also appears in Figure 3. In the above string, we can insert a separator in front of any occurrence of $\Delta$, except for the two left-most ones, without changing the granularity represented. The following string, where such separators have been inserted, represents indeed the same granularity:

We therefore say that the first two occurrences of $u$ occur within a granule, while all subsequent occurrences of $\sqcup$ occur between two granules. This observation is generalized in the following definition.
Definition 6 Let $t$ be a string containing infinitely many non-separator symbols. Let $j \in \mathbb{N}$. We say that $t$ is separable at $j$ iff $\operatorname{gran}(t)=\operatorname{gran}\left(t^{\prime}\right)$ where $t^{\prime}$ is the string obtained from $t$ by inserting in $t$ a separator right in front of the $j^{\text {th }}$ non-separator symbol; that is,

1. $t^{\prime}[1, j-1]=t[1, j-1]$,
2. $t^{\prime}(j)=$ ?, and
3. $t^{\prime}(i)=t(i-1)$ for all $i>j$.

Hence, the string $t_{1}$ shown in Figure 3 is separable at 5, $6,8,9, \ldots$, but neither at 2 nor at 3 .
Figure 3 shows the calculation of the supremum of $\operatorname{gran}\left(t_{1}\right)$ and $\operatorname{gran}\left(t_{2}\right)$ for two infinite strings $t_{1}$ and $t_{2}$. The supremum is represented by the string $s$. In this figure, the $i^{\text {th }}$ non-separator symbols of $t_{1}, t_{2}$, and $s$ are aligned in the same column, for each $i$. The $i^{\text {th }}$ nonseparator symbol of $s$ is a filler if the $i^{\text {th }}$ non-separator symbol of $t_{1}$ or $t_{2}$ is a filler; otherwise it is a gap. Moreover, a separator in $t_{1}$ is copied in $s$ if $t_{2}$ is separable at the corresponding position, and vice versa. For example, the first separator in $t_{1}$ is not copied in $s$ as $t_{2}$ is not separable at 5 . The third separator in $t_{2}$ is copied in $s$ as $t_{1}$ is separable at 7 .
Definition 7 Let $t$ be a string with infinitely many non-separator symbols. We write $t^{\square}$ for the string obtained from $t$ by removing all separators. That is,

1. $(\lambda t)^{\square}=t^{\square}$,
2. $(\Sigma t)^{\square}=\llbracket s$ with $s=t^{\square}$, and
3. $(\sqcup t)^{\square}=\sqcup s$ with $s=t^{\square}$.


Figure 3: Supremum construction.

Theorem 10 Let $t_{1}$ and $t_{2}$ be two strings with infinitely many non-separator symbols. Let $s$ be an infinite string. For the lattice of granularities ordered by $\subseteq$, we have $\sup \left\{\operatorname{gran}\left(t_{1}\right), \operatorname{gran}\left(t_{2}\right)\right\}=\operatorname{gran}(s)$ iff for all $i \in \mathbb{N}$,

- $s^{\square}(i)= \begin{cases}\sqcup & \text { if } t_{1}{ }^{\square}(i)=t_{2} \\ \square & \text { otherwise }\end{cases}$
- $s$ is separable at $i$ iff both $t_{1}$ and $t_{2}$ are separable at $i$.


## Combining Granspecs

We now discuss the computation of the supremum of two granularities that are given under the form of two granspecs, say $\alpha$ and $\beta$. Basically, we can combine $\alpha^{\infty}$ and $\beta^{\infty}$ as explained in the previous section. The string resulting from this combination will be ultimately periodic. The only difficulty concerns delimiting in this string the offset and the repeating part. For example, in Figure 3, the strings $t_{1}$ and $t_{2}$ are produced by the
 spectively. Note that both granspecs contain a single non-separator occurrence in their offsets. The supremum granspec, on the other hand, contains three nonseparator occurrences in its offset. It is produced by the canonical granspec (■! $\boldsymbol{\|} \cup \boldsymbol{U}$ ).

We assume that the two granspecs whose supremum has to be computed, correspond on the number of non-separator symbols in their offsets and repeating parts respectively. This is without loss of generality, as for any two granspecs $\alpha$ and $\beta$, there exist aligned granspecs $\alpha^{\prime}$ and $\beta^{\prime}$ such that $\alpha^{\prime}$ and $\beta^{\prime}$ correspond on the number of non-separator symbols in their offsets and repeating parts respectively, and such that $\alpha^{\prime}$ and $\beta^{\prime}$ are trace-equal to $\alpha$ and $\beta$ respectively. For example,

The granspecs $\alpha^{\prime}$ and $\beta^{\prime}$ both have one non-separator symbol in their offsets, and 6 non-separator symbols in their repeating parts.
Theorem 11 Let $\alpha_{1}=\left(v_{1} ; w_{1}\right)$ and $\alpha_{2}=\left(v_{2} ; w_{2}\right)$ be two granspecs. There exist two aligned granspecs $\alpha_{1}^{\prime}=\left(v_{1}^{\prime} ; w_{1}^{\prime}\right)$ and $\alpha_{2}^{\prime}=\left(v_{2}^{\prime} ; w_{2}^{\prime}\right)$ such that $\rrbracket v_{1}^{\prime} \rrbracket=\rrbracket v_{2}^{\prime} \rrbracket$, $\rrbracket w_{1}^{\prime} \rrbracket=\rrbracket w_{2}^{\prime} \rrbracket, \alpha_{1} \equiv \mathrm{~T} \alpha_{1}^{\prime}$, and $\alpha_{2} \equiv \mathrm{~T} \alpha_{2}^{\prime}$.

Given two granspecs $\alpha$ and $\beta$, the following theorem tells how to distinguish the offset and the supremum in the string representing the supremum of $\operatorname{gran}\left(\alpha^{\infty}\right)$ and $\operatorname{gran}\left(\beta^{\infty}\right)$.
Theorem 12 Let $\alpha=(v ; w)$ and $\beta=(x ; y)$ be two granularities where both $w$ and $y$ contain at least one occurrence of $\mathbb{1}$, and $\llbracket v \rrbracket=\rrbracket x \rrbracket$ and $\rrbracket w \rrbracket=$ \y\|. Let $s$ be an infinite string such that gran(s) $=$ $\sup \left\{\operatorname{gran}\left(\alpha^{\infty}\right), \operatorname{gran}\left(\beta^{\infty}\right)\right\}$. Let a be the smallest number such that $w(a)=\mathbb{1}$, and let $b$ be the smallest number such that $y(b)=\mathbb{1}$. Let $m=\|v\|+$ $\max \{\square(w, a), \square(y, b)\}$. Let $k \in \mathbb{N}$ such that $s(k)=\square$ and $\square(s, k)=m$. That is, $k$ is the position in $s$ of the $m^{\text {th }}$ non-separator. Likewise, let $l \in \mathbb{N}$ such that $s(l)=\|$ and $\square(s, l)=m+\| w \rrbracket$. Then $\operatorname{gran}(s)=$ $\operatorname{gran}\left((s[1, k-1] ; s[k, l-1])^{\infty}\right)$.
Note that Theorem 12 implies that for the computation of the granspec of the supremum, it suffices to know the first $(l-1)$ symbols of $s$.

For example, let $\alpha=(v ; w)=(\boldsymbol{\bullet} ; \mathrm{U} \boldsymbol{\bullet} \ell)$ and $\beta=(x ; y)=(\sqcup \ell ; \cup 1 \mathbf{l})$; see Figure 3. Then the first filler in $w$ is the third non-separator symbol of $w$, and the first filler in $y$ is the second non-separator symbol of $y$. Hence, the value for $m$ in Theorem 12 is $1+\max \{3,2\}=4$. The $4^{\text {th }}$ non-separator symbol of $s$ occurs at position 4 of $s$, hence $k=4$. For the value of $l$, note that $m+\rrbracket w \rrbracket=4+3=7$, and the $7^{\text {th }}$ nonseparator symbol of $s$ occurs at position 8 of $s$ (for $s(7)$ is a separator), hence $l=8$. Consequently, the offset for the supremum can be found as $s[1,4-1]=\llbracket \sqcup!$, and the repeating part as $s[4,8-1]=\| \sqcup!$.

## Extensions

So far, we have represented a granularity by an infinite string. We now sketch two natural extensions to infinite strings that can be useful for granularity modeling.

The first extension is concerned with infinite trees. In our framework, all granularities are expressed in terms of one single base granularity. It is often more natural to define new granularities in terms of previously defined ones. For example, one may take "day" as the base granularity, define "month" in terms of "day," and then define "year" in terms of "month." It seems natural to use trees, rather than flat strings, to model such granularity layers. Significantly, in the area of temporal logic, multi-layered temporal structures have been
used as the underlying structures of logics that, unlike linear time temporal logic, can deal with different time granularities (Montanari, Peron, \& Policriti 1999).

The second extension deals with multiple dimensions. These extra dimensions can be temporal or spatial. For example, air traffic statistics may subdivide the controlled airspace in regular tridimensional corridors and compute the number of daily flight traversals per corridor. This involves a four-dimensional granularity, with one temporal (day) and three spatial dimensions.
For two dimensions, it suffices to replace strings by "rectangular arrays" in granspecs. An example with an empty offset:

$$
\alpha=\left(\begin{array}{ccccc} 
& 1 & \sqcup & l & U \\
\varepsilon ; & 1 & 1 & l & 1 \\
& l & l & l & l \\
& \sqcup & 1 & l & 1
\end{array}\right)
$$

The produced two-dimensional trace is shown next:


One
lence classes is $\{(6,2),(6,4),(7,2),(7,3),(8,2),(8,4)\}$. Extensions to higher dimensions can be defined along the same lines.

## Conclusion

We have presented a new formalism for the finite representation of infinite granularities with linearly repeating patterns, and possibly gaps within and between granules. Granularities are expressed as strings over a three-symbol alphabet. We have introduced a unique canonical form which turns out to be the most concise representation for a particular granularity. We showed how certain reasoning tasks can be performed by simple string manipulation. The formalism can be extended in a natural way to multiple dimensions.

## Acknowledgment

The anonymous referees provided a number of very valuable comments.

## References

Bettini, C., and De Sibi, R. 1999. Symbolic representation of user-defined time granularities. In Proc.

6th Int. Workshop on Temporal Representation and Reasoning (TIME99), 17-28. IEEE Computer Society Press.
Bettini, C.; Dyreson, C.; Evans, W.; Snodgrass, R.; and Wang, X. 1998. A glossary of time granularity concepts. In Etzion, O.; Jajodia, S.; and Sripada, S., eds., Temporal Databases: Research and Practice, number 1399 in LNCS State-of-the-art Survey. Springer-Verlag. 406-413.
Bettini, C.; Wang, X.; and Jajodia, S. 1998. A general framework for time granularity and its application to temporal reasoning. Annals of Mathematics and Artificial Intelligence 22(1,2):29-58.
Clifford, J., and Rao, A. 1987. A simple, general structure for temporal domains. In Rolland, C.; Bodart, F.; and Leonard, M., eds., Temporal Aspects in Information Systems. Amsterdam, The Netherlands: North-Holland/IFIP. 17-28.
Dyreson, C. E.; Evans, W. S.; Lin, H.; and Snodgrass, R. T. 1998. Efficiently supporting temporal granularities. Technical Report TR-31, TimeCenter.
Montanari, A.; Peron, A.; and Policriti, A. 1999. The way to go: Multi-level temporal logics. In Gerbrandy, J.; Marx, M.; de Rijke, M.; and Venema, Y., eds., Liber Amicorum for the Fiftieth Birthday of Johan van Benthem. ILLC.
Wang, X.; Bettini, C.; Brodsky, A.; and Jajodia, S. 1997. Logical design for temporal databases with multiple granularities. ACM Trans. on Database Systems 22(2):115-170.
Wijsen, J., and Ng, R. T. 1999. Temporal dependencies generalized for spatial and other dimensions. In Proc. Int. Workshop on Spatio-Temporal Database Management (STDBM'99), LNCS 1678, 189-203. Springer.
Wijsen, J. 1998. Reasoning about qualitative trends in databases. Information Systems 23(7):469-493.
Wijsen, J. 1999. Temporal FDs on complex objects. ACM Trans. on Database Systems 24(1):127-176.

## Proof of Theorem 6

Proof. Assume $\alpha$ and $\beta$ are two aligned granspecs such that $\alpha \not \equiv \mathrm{T} \beta$. Let $n \in \mathbb{N}$ such that $\alpha^{\infty}(n) \neq$ $\beta^{\infty}(n)$ and $\alpha^{\infty}(i)=\beta^{\infty}(i)$ for every $i<n$. Let $m=\square\left(\alpha^{\infty}, n-1\right)=\square\left(\beta^{\infty}, n-1\right)$. We consider the following cases (the roles of $\alpha$ and $\beta$ can be interchanged):

1. $\alpha^{\infty}(n)=$ and $\beta^{\infty}(n)=1$. Since $\beta^{\infty}$ is aligned, $\beta^{\infty}(n-1)==\alpha^{\infty}(n-1)$. Then $(m, m+1) \epsilon$ $\operatorname{gran}\left(\alpha^{\infty}\right)$ while $(m, m+1) \notin \operatorname{gran}\left(\beta^{\infty}\right)$, hence $\alpha \not \equiv_{\mathrm{G}}$ $\beta$.
2. $\alpha^{\infty}(n)=$ ? and $\beta^{\infty}(n)=\cup$. Suppose $\alpha \equiv$ G $\beta$. We show by induction on $i$ that $\alpha^{\infty}(i)=\beta^{\infty}(i)=U$ for all $i \geq n+1$. Base $i=n+1$. The situation is depicted in Figure 4. Clearly, $\alpha^{\infty}(i) \neq$ l because otherwise $\alpha$ would not be aligned. Likewise, $\beta^{\infty}(i) \neq$ ?

| $\square\left(\alpha^{\infty}, j\right):$ | $m$ | $m$ |  |
| :--- | :---: | :---: | :---: |
| $\square\left(\beta^{\infty}, j\right):$ | $m$ | $m+1$ |  |
| $\alpha^{\infty}(j):$ | $\vdots$ | $l$ | $?$ |
| $\beta^{\infty}(j):$ | $\vdots$ | $\stackrel{y}{c}$ | $?$ |
| $j:$ | $n-1$ | $n$ | $n+1$ |

Figure 4: Construction used in the base of Theorem 6.


Figure 5: Construction used in the inductive step of Theorem 6.
because otherwise $\beta$ would not be aligned. Suppose $\alpha^{\infty}(i)=\llbracket$. Then $(m+1, m+1) \in \operatorname{gran}\left(\alpha^{\infty}\right)$ while $(m+1, m+1) \notin \operatorname{gran}\left(\beta^{\infty}\right)$, hence $\alpha \not \equiv_{G} \beta$, a contradiction. We conclude $\alpha^{\infty}(i)=\amalg$. Suppose $\beta^{\infty}(i)=$. Then $(m, m+2) \in \operatorname{gran}\left(\beta^{\infty}\right)$ while $(m, m+2) \notin \operatorname{gran}\left(\alpha^{\infty}\right)$, hence $\alpha \not \equiv_{\mathrm{G}} \beta$, a contradiction. We conclude $\beta^{\infty}(i)=\sqcup$. Inductive step. By the induction hypothesis, $\alpha^{\infty}(j)=\beta^{\infty}(j)$ for all $j \in[n+1, i]$. The situation is depicted in Figure 5. Obviously, since $\alpha$ and $\beta$ are aligned, $\alpha^{\infty}(i+1) \neq 1$ and $\beta^{\infty}(i+1) \neq 1$. If $\alpha^{\infty}(i+1)=$ - then $(m+(i-n)+1, m+(i-n)+1)$ belongs to $\operatorname{gran}\left(\alpha^{\infty}\right)$ but not to $\operatorname{gran}\left(\beta^{\infty}\right)$, hence $\alpha \not \equiv \mathrm{G} \beta$, a contradiction. Hence, $\alpha^{\infty}(i+1)=ப$. If $\beta^{\infty}(i+1)=\square$ then ( $m, m+(i-n)+2$ ) belongs to $\operatorname{gran}\left(\beta^{\infty}\right)$ but not to $\operatorname{gran}\left(\alpha^{\infty}\right)$, hence $\alpha \equiv_{\mathrm{G}} \beta$, a contradiction. Hence, $\beta^{\infty}(i+1)=\sqcup$. This concludes the induction step. Hence, $\alpha^{\infty}(i)=\beta^{\infty}(i)=\sqcup$ for all $i \geq n+1$. But then $\alpha$ is not aligned, a contradiction. We conclude by contradiction that $\alpha \not \equiv{ }_{\mathrm{G}} \beta$.
3. $\alpha^{\infty}(n)=$ and $\beta^{\infty}(n)=\sqcup$. Then $(m+1, m+1) \in$ $\operatorname{gran}\left(\alpha^{\infty}\right)$ while $(m+1, m+1) \notin \operatorname{gran}\left(\beta^{\infty}\right)$, hence $\alpha \not \boldsymbol{F}_{\mathrm{G}} \beta$.

## Proof of Theorem 8

Before we give the proof, we recall the following result.
Theorem 13 (Fine and Wilf, 1965) Let $x, y \in A^{*}$, $n=|x|, m=|y|, d=\operatorname{gcd}(n, m)$. If two powers $x^{p}$ and $y^{q}$ of $x$ and $y$ have a common left factor of length at least equal to $n+m-d$, then $x$ and $y$ are powers of the same word.
The proof Theorem 8 is given next.
Proof. Let $\alpha$ and $\beta$ be two canonical granspecs such that $\alpha \equiv_{\mathrm{G}} \beta$. Let $\alpha=(v ; w)$ and $\beta=(x ; y)$. Since $\alpha$ and $\beta$ are aligned, $\alpha \equiv_{\mathrm{T}} \beta$ by Theorem 6. Hence $v w w w \ldots=x y y y \ldots$ Three cases can occur: $|v|<|x|$, $|v|=|x|$, or $|v|>|x|$. The third case is similar to the first one.

1. $|v|<|x|$. Hence, $x=v z$ for some string $z \neq \varepsilon$. Hence, $v w w w \ldots=v z y y y \ldots$.... Hence, $w w w \ldots=z y y y \ldots$ Let $t=w w w \ldots$ Hence, $\operatorname{tail}(y)=$ $t(|z|+k|y|)$ for any $k \geq 1$, and $\operatorname{tail}(z)=t(|z|+l|w|)$ for any $l \geq 0$. Since $\alpha$ and $\beta$ are canonical, $\operatorname{tail}(z)=$ $\operatorname{tail}(x) \neq \operatorname{tail}(y)$. Hence, $t(|z|+k|y|) \neq t(|z|+l|w|)$ for any $k \geq 1, l \geq 0$. Hence, $|z|+k|y| \neq|z|+l|w|$ for any $k \geq 1, l \geq 0$. Hence, $k|y| \neq l|w|$ for any $k \geq 1, l \geq 0$. Let $k=|w|$ and $l=|y|$. So $|w| \cdot|y| \neq|\bar{y}| \cdot|w|$, a contradiction.
2. $|v|=|x|$. Hence, $w w w \ldots=y y y \ldots$ By Theorem 13 , there is some string $r$ such that $w=r^{k}$ and $y=r^{l}$ for some $k, l \geq 1$. If $k=l=1$ then $r=w=y$ and $\alpha=\beta$. If $k>1$ or $l>1$ then $\alpha$ or $\beta$ is not canonical, a contradiction.

[^0]:    ${ }^{1}$ Of course, the repeating part must contain at least one occurrence of $\boldsymbol{\|}$ or $\downarrow$.

