A mathematical framework for the semantics of symbolic languages representing periodic time

Lavinia Egidi and Paolo Terenziani

Dipartimento di Informatica
Università del Piemonte Orientale *A. Avogadro*
Spalto Marengo, 33 - 15100 Alessandria - Italy
{lavinia.egidi, paolo.terenziani}@mfn.unipmn.it

Abstract

In several areas, including Temporal DataBases, Presburger Arithmetic has been chosen as a standard reference to express the semantics of languages representing periodic time, and to study their expressiveness. On the other hand, the proposal of most symbolic languages in the AI literature has not been paired with an adequate semantic counterpart, making the task of studying the expressiveness of such languages and of comparing them a very complex one. In this paper, we first define a representation language which enables us to handle each temporal point as a complex object enriched with all the structure it is immersed in, and then we use it in order to provide a Presburger semantics for classes of symbolic languages coping with periodicity.

Keywords: user-defined periodicity, symbolic languages, semantics, Presburger Arithmetic

1. Introduction

In many application areas, including planning, scheduling, process control, multimedia, active databases, banking and law, the treatment of periodic events is an essential task. Thus, many approaches in the area of AI and of temporal databases (TDB henceforth) focused on periodicity and user-defined calendars. Concerning the TDB area, Baudinet [1] distinguished between the approaches dealing with periodicity using *deductive rules* and those using *constraints*. As regards the (logical) approaches based on deductive rules, consider e.g. [4] which dealt with periodicity via the introduction of the successor function in Datalog. More generally, also classical temporal logics can be conceived as deductive-rule approaches, in Baudinet's terms. On the other hand, constraint-based approaches use mathematical formulae, and constraints on the variables in the formulae, in order to express (user-defined) calendars and periodicities (e.g. [8]). A third possible approach is the one of symbolic languages, that provide a human-oriented way of handling time, allowing users to define periodicities in a natural, incremental and compositional way (see, e.g., the discussions in [10, 11, 12]).

But the semantics of symbolic languages has been rarely provided in a clean and rigorous way (and, in some cases, just relying on user intuition). In contrast, in general, the expressive power and the semantics of TDB constraint-based approaches have been specified in a clean and formal way on the basis of Presburger Arithmetic (i.e., the first order theory of addition and ordering over integers, henceforth PA). PA has been chosen because of its simplicity and its expressive power. In fact, it turns out that all sets definable in PA are finite, periodic, or eventually periodic [7], which makes such a theory a natural reference to evaluate the expressiveness of languages representing periodicities.

Therefore, a semantics of symbolic languages based on PA, would not only allow a clear analysis of those languages, but also offer a powerful tool for comparison with totally different approaches (like constraint-based ones).

Actually, a uniform semantical analysis has a broader relevance in that it enables to reason on the significance and usefulness of the virtually infinite temporal structures that are formally definable, to expose shortcomings of existing approaches (limited expressiveness, side-effects, definition of structures with no clear semantical content), and thus to define operators (and languages) in a more focused way.

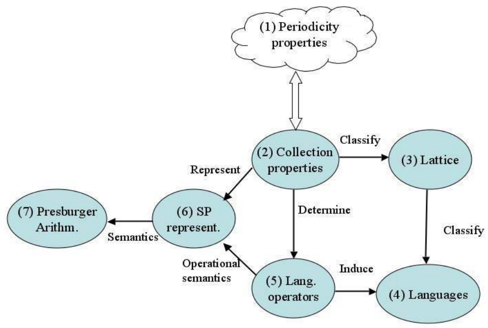

Figure 1. A comprehensive approach to user-defined periodicities

Our comprehensive approach to user-defined symbolic periodicities is schematized in Fig. 1. The starting point is the abstract concept of periodicity (Fig.1-(1)).

In all symbolic approaches, the underlying assumption is that a periodicity can be identified with the temporal extent it denotes, and is usually conceived as a collection of time points or time intervals, which, in some approaches, may also have additional structure (Fig.1-(2)).

We structured our study by defining five orthogonal properties of periodicities. We construct a lattice of classes of periodicities, based on these properties. The bottom class contains periodicities for which none of the properties hold, and the top class contains all periodicities (Fig.1-(3)).

We noticed that classes of periodicities can be matched with languages (Fig.1-(4)), and exploited this for a uniform analysis and classification of languages in the literature. It turned out that none of the languages we considered covers exactly any of our classes, nor is powerful enough to express the top class. This led us to the idea of designing a family of languages, each one matching precisely one class in the lattice. Of course the possible languages corresponding to each class are virtually infinite. We chose to favor languages whose operators clearly deal with one of our orthogonal properties with no side-effects and no exceptions. This way, the operators themselves (Fig.1-(5)) can be superimposed to the lattice.

All of this framework is supported by the semantical analysis. But it is necessary to define a meaningful and uniform way of interpreting the different approaches with PA. We propose an intermediate mathematical representation language (that we call the *structured point representation* (SPR)) which allows us to represent time points together with the structure in which they are embedded; the SPR is a sort of *interlingua* to express the temporal extent denoted by user-defined periodicities (Fig.1-(6).) Then, we provide a Presburger semantics of such a language (Fig.1-(7)).

The definition of the interlingua and of the Presburger semantics, enabled us to:

1. give a mathematical representation of the temporal extent of periodicities of symbolic approaches in the literature, and through that a semantics to those languages, as a starting point for the design of new languages; in [5] we use the above framework in order to propose a classification of some symbolic approaches in the literature, based on their expressive power;

2. compare existing approaches via the uniform semantics;

3. fully understand the interplay between operators and the five properties on which the lattice is based; in [6] the five properties guide us in the definition of five operators for a maximally expressive language;

4. study the *operational* semantics of our operators and thus to provide a formal semantics to each one of the languages in the *family*; in [6]) we analyze the expressiveness of our operators, and prove that they are orthogonal with respect to the five properties, in that each one introduces a single property, independently of the others; in [6] we also argue that the analysis can be carried over to any languages using those operators, in a compositional way;

5. analyze the expressive power of our languages with respect to periodic (and eventually periodic) sets.

This paper focuses on issues (6) and (7) in Fig.1. First, it contributes the definition of the interlingua and of the semantical interpretation in PA (Section 2). In Section 3 we show how some basic structures and features from symbolic approaches can be represented by structured points and thus defined in PA (cf. Item 1 above). Finally, in Section 4, we compare some prominent formalisms from the literature, using the uniform Presburger semantics we have defined (cf. Items 1 and 2 above).

Aspects (2)-(5) from Fig.1 are discussed in two companion papers [5, 6].

2. Structured points and Presburger semantics

We adopt *discrete*, *linearly ordered* and *unbounded* time; *time points* are the basic temporal primitives. The basic structures on time points are *time intervals*, i.e. non-empty sets of time points. A *convex* time interval is the set of all points between two endpoints; a *gap* interval is a non-convex set of time points. Time intervals may be *left* and/or *right-infinite*.

Our interest is focused solely on the *temporal extent* of periodic events (and not on other aspects such as agent, location, etc.), i.e. on structural properties. Therefore we follow [10] and identify *periodicities* with order-n collections, for $n > 0$, where an order-n collection is defined as follows:

Definition 1 (Order-n collections) *For each $n > 0$, an* order-n collection *is a multiset of order-$(n-1)$ collections. (For uniformity of exposition we refer to intervals as* order-0 *collections.)*

For instance, "days grouped by months" is an order-2 collection.

Precisely, in contrast with [10], we define collections as *multi*sets. Also, in the following, we implicitly assume an ordering on multisets. Any ordering can be chosen: we use the lexicographic order on the endpoints of extents, where the *extent* of an order-n collection ($n \geq 0$) is the interval whose endpoints are the minimum and maximum points belonging to the collection [2].

A *periodic* event is an order-n collection that can be defined giving a non-empty repeating pattern RP, and a positive period specifying the spacing of different occurrences of RP. An *eventually periodic* event is (using the terminology of [7]) a generalization of a periodic one, in which we admit the existence of a finite aperiodic part, and of left (or right) bounds.

Definition 2 (SPR) *The* structured point representation (SPR) *of an order-n collection is a set of sequences of length $n+1$, $(id_0, id_1, \ldots, id_n)$, such that*

- *the first element of the sequence denotes a temporal point, as displacement from the origin of the (discrete) temporal axis;*
- *the subsequence $(id_k, \ldots, id_n)$ identifies the order-$(k-1)$ subcollection in which the point is embedded.*

As an example,

$$(1,1,1,1),(2,1,1,1),(3,1,1,1),(5,2,1,1),\\ (6,2,1,1),(3,3,2,1),(4,3,2,1),(4,4,3,2) \quad (1)$$

models the set

$$\{\{\{[1,3],[5,6]\},\{[3,4]\}\},\{\{[4,4]\}\}\}. \quad (2)$$

We guide the reading of SPR (1) with some remarks. The order-3 collection (2) consists of two order-2 collections, and therefore is represented by 4-tuples. Index id_3 takes values 1 and 2. The first order-2 collection contains two order-1 collections, identified by $(id_2, id_3) = (1,1)$ and $(id_2, id_3) = (2,1)$. The order-1 collection $(1,1)$ contains two intervals, the first of which has points 1, 2 and 3: the tuples $(1,1,1,1)$ through $(3,1,1,1)$ express that points $1-3$ belong to interval $(1,1,1)$ nested in order-1 collection $(1,1)$, in turn nested in order-2 collection 1.

Similarly, $(4,4,3,2)$ describes point 4 as belonging to interval $[4,4]$ (identified by $(id_1, id_2, id_3) = (4,3,2)$), nested in order-1 collection $(3,2)$, which in turn is nested in order-2 collection $id_3 = 2$. Notice that, on the other hand, the structured point $(4,3,2,1)$ is the same point 4 but viewed as belonging to interval $[3,4]$, identified by $(3,2,1)$, which is nested in different order-1 and order-2 collections.

Indeed, the name "structured point representation" stems from the observation that each sequence represents a point along with the structure it lives in.

Presburger Arithmetic (PA) is the first order theory of addition and ordering over the integers (see, e.g., [7]).

The choice of PA is based on two essential features: its simplicity and its expressive power. It turns out that the latter is sufficient for all eventually periodic temporal languages since the following holds:

Theorem 1 ([7], Theorem 32F) *A set of natural numbers is definable in Presburger Arithmetic if and only if it is finite, periodic, or eventually periodic.*

As mentioned, in our approach, order-n collections are interpreted in PA via their SPR. The latter are simply defined as sets of sequences $(x_0, x_1, \ldots, x_n)$.

By the theorem above, all patterns of eventually periodic sequences of values can be defined for each variable. This, combined with the capabilities of structured points to describe structure, implies that order-n collections representing eventually periodic events are all definable. We illustrate with some examples the definitional process, and refer the reader to Section 3 for further details and applications.

The simplest structure that we are interested in, is a finite set of intervals (an order-1 collection). We give a general form for it and show the

PA formula defining it. Consider the order-1 collection $\{[s_1, e_1], [s_2, e_2], \ldots, [s_t, e_t]\}$, with SPR

$$(s_1, 1), (s_1 + 1, 1), ..., (e_1, 1), (s_2, 2), ..., (e_2, 2), ..., (e_t, t).$$

It is defined by the formula

$$(s_1 \leq x_0 \leq e_1 \wedge x_1 = 1) \vee (s_2 \leq x_0 \leq e_2 \wedge x_1 = 2) \vee \\ \vee \ldots \vee (s_t \leq x_0 \leq e_t \wedge x_1 = t) \quad (3)$$

where s_j, e_j are integers such that $s_j \leq e_j$.

A more complex situation arises in the case of a periodic structure. Consider, for instance, the order-1 collection $\{\ldots [-8, -7], [-5, -3], [2, 3], [5, 7], [12, 13], [15, 17], \ldots\}$. It is represented by the following SPR:

$$\ldots (-8, -1), (-7, -1), (-5, 0), (-4, 0), (-3, 0), \\ (2, 1), (3, 1), (5, 2), (6, 2), (7, 2), (12, 3), \\ (13, 3), (15, 4), (16, 4), (17, 4), \ldots \quad (4)$$

The formula defining (4) is

$$(\exists y)(\exists z) \\ [(2 \leq y \leq 3 \wedge x_0 \equiv y \bmod 10 \wedge \\ \wedge x_0 = 10 \cdot z + y \wedge \ x_1 = 2 \cdot z + 1) \vee \\ \vee (5 \leq y \leq 7 \wedge x_0 \equiv y \bmod 10 \wedge \\ \wedge x_0 = 10 \cdot z + y \wedge \ x_1 = 2 \cdot z + 2)]. \quad (5)$$

Here 10 is the length of the period, and 2 is the number of intervals in the repeating pattern (since the repeating pattern is a sequence of intervals). In Formula (5), y is used to define the periodic portion of the sequence, z is used to identify the z-th periodic repetition of that portion, and $x_1 = 2 \cdot z + i$ defines the numbering of the index id_1 in terms of the occurrence of the repeating pattern that the interval belongs to, and its position (1 or 2 in this simple case) in the repeating pattern.

Periodicities can be explicitly bounded with conditions of the form $x_0 \geq b_L$ and $x_0 \leq b_R$.

3. Interpretation of symbolic periodicities

We show the SPRs of basic structures as a first example of what SPRs look like, and how they can be defined in PA. Then, we present with two examples how structure ([10]), gaps ([3]) and repetitions (exact overlaps, [5]) are expressed in our framework, to give a feeling of the expressiveness of our approach.

Symbolic languages are typically built starting from the concept of chronon and the most basic periodicities, calendars.

We call *chronon calendar* the order-1 collection

$$\{\ldots, [-1, -1], [0, 0], [1, 1], [2, 2], \ldots\}.$$

SPR 1 (Chronon Calendar) *The chronon calendar has SPR* $\{(i, i) | i \in \mathbb{N}\}$.

Then, trivially,

PA 1 (Chronon Calendar) *The PA formula representing a chronon calendar is* $x_0 = x_1$.

Calendars [10] are order-1 collections consisting of convex adjacent intervals spanning the whole time axis, in which a repeating pattern made of a sequence of intervals (possibly of length one, i.e. a single interval) occurs periodically—again, see e.g. [10]. Therefore, the first elements of the tuples in the SPRs of calendars cover the whole time axis, and a single index id_1 is sufficient, since the collection has no structure; Given the width of each interval in the repeating pattern, the SPR can be described in a clean way as follows:

SPR 2 (Calendars) *Given a calendar C, of period p and whose repeating pattern consists of intervals of widths $w_1, \ldots, w_n$ such that $p = \sum_{h=1}^{n} w_h$, its SPR is (letting $w_0 = 0$)*

$$\left\{ (i, j) | i \in \mathbb{N} \ \wedge \ j = n \left\lfloor \frac{i}{p} \right\rfloor + l \ \text{ where } l \text{ is such that} \\ \sum_{h=0}^{l} w_h \leq i \ mod \, p < \sum_{h=0}^{l+1} w_h \right\}.$$

With the notation above,

PA 2 (Calendars) *The formula of PA defining the SPR of a calendar has the form:*

$$(\exists y)(\exists z) \ [(0 \leq y \leq w_1 \wedge x_0 \equiv y \ mod \, p \wedge \\ \wedge x_0 = p \cdot z + y \wedge x_1 = n \cdot z) \vee \\ \vee (w_1 + 1 \leq y \leq w_1 + w_2 \wedge x_0 \equiv y \ mod \, p \wedge \\ \wedge x_0 = p \cdot z + y \wedge x_1 = n \cdot z + 1) \vee \\ \vee \ \ldots \ \vee (\sum_{i=1}^{n-1} w_i + 1 \leq y \leq \sum_{i=1}^{n} w_i \wedge \\ \wedge \ x_0 \equiv y \ mod \, p \ \wedge \ x_0 = p \cdot z + y \ \wedge \\ \wedge x_1 = n \cdot z + n - 1)].$$

Structure was introduced in [10], using, as mathematical formalism, order-n collections. An example is the order-3 collection in (2):

PA 3 (Structure) *The PA formula*

$$(1 \leq x_0 \leq 3 \wedge x_1 = 1 \wedge x_2 = 1 \wedge x_3 = 1) \vee \\ \vee (5 \leq x_0 \leq 6 \wedge x_1 = 2 \wedge x_2 = 1 \wedge x_3 = 1) \vee \\ \vee (3 \leq x_0 \leq 4 \wedge x_1 = 1 \wedge x_2 = 2 \wedge x_3 = 1) \vee \\ \vee (4 \leq x_0 \leq 4 \wedge x_1 = 1 \wedge x_2 = 1 \wedge x_3 = 2)$$

defines the order-3 collection

$$\{\{\{[1, 3], [5, 6]\}, \{[3, 4]\}\}, \{\{[4, 4]\}\}\},$$

with SPR (1) in Sec. 2.

It can be seen how indexes describe the structure in which each temporal point is nested.

In [3], the formalism of [10] was enriched with the introduction of gaps. We take a step further, defining collections as multi-sets (see Def. 1), as opposed to sets, and thus allowing repetitions. An example of a collection (in our sense) with gap intervals is $\{[-1,1],[2,3],[2,3],[9,9]\bigcup[50,52]\}$, where $[9,9]\bigcup[50,52]$ denotes a unique interval with a hole in it.

SPR 3 (Gaps and overlaps) *The SPR of the set*

$$\{[-1,1],[2,3],[2,3],[9,9]\bigcup[50,52]\}$$

is

$$(-1,1),(0,1),(1,1),(2,2),(3,2),(2,3)(3,3),\\ (9,4),(50,4),(51,4),(52,4).$$

It can be defined in PA this way:

PA 4 (Gaps and overlaps) *The SPR of the set*

$$\{[-1,1],[2,3],[2,3],[9,9]\bigcup[50,52]\}$$

is defined by the PA formula

$$(-1\le x_0\le 1\wedge x_1=1)\vee(2\le x_0\le 3\wedge x_1=2)\vee\\ \vee(2\le x_0\le 3\wedge x_1=3)\vee(9\le x_0\le 9\wedge x_1=4)\vee\\ \vee(50\le x_0\le 52\wedge x_1=4).$$

Notice that here the second and third clause differ only for the value of the second variable, to yield an exact overlap, and that on the other hand the same variable is instantiated in the same way in the last two clauses to obtain a gap.

4. Comparing approaches

We review very briefly three approaches that we wish to compare looking at the shape of the formulae of PA that express the temporal structures through SPRs.

Kabanza et al. [8] define a mathematical formalism based on the concept of *linear repeating points* (**lrps**). An lrp is a set of points $\{x(n)\}$ defined by an expression of the form $x(n)=c+kn$ where k and c are integer constants and n ranges over the integers. The temporal structures are sequences of intervals, whose endpoints are lrps possibly bound by linear equalities or inequalities ($X_1\le aX_2+b, X_1=aX_2+b$, with a,b integers). Actually, a constraint of the form $X_1\le X_2$ is always necessary in order to define objects that can be interpreted as intervals. In [8] it is proven that this language is equivalent to the language of PA. (We consider the pure language of possibly constrained lrps, with no union operator, for the purposes of this analysis.)

A first observation about the language of Kabanza et al. is that the endpoints of each interval can be defined independently of one another. But, on the other hand, each endpoint belongs to a sequence that being defined by a linear relation $c+kn$ defines points that are at the same distance k each from the next: $c+k(n+1)-(c+kn)=k$.

This implies that a temporal structure $[a+n_1m_1, b+n_2m_2]$, with constraint $X_1\le X_2$, can be expressed by a formula:

$$(\exists y)(\exists z)y\le x_0\le z\wedge y\equiv a \bmod m_1\wedge\\ \wedge z\equiv b \bmod m_2\wedge x_1=y, \quad (6)$$

which is obtained simply making the congruence relations explicit.

The form discussed above, with no other constraints than the minimal ones $X_1\le X_2$, actually defines collections of intervals that are too general to have a clear meaning. And in fact concrete examples in [8] exhibit constraints of the form $X_2=X_1+d$. When such a constraint is specified, next to an interval $[a+n_1m_1, b+n_2m_2]$ like the one above, intervals are limited to a fixed width d. Notice that not all the intervals starting at points $y\equiv a \bmod m_1$ belong to the structure: since the above constraint binds the two endpoints of the interval, the starting point of intervals must also satisfy a congruence derived from the one for the ending point. It is much simpler to express this with a formula of PA:

$$(\exists y)y\le x_0\le y+d\ \wedge y\equiv a \bmod m_1\wedge\\ \wedge y\equiv b-d \bmod m_2\wedge x_1=y. \quad (7)$$

By the Chinese Remainder Theorem (see e.g. [9]), if m_1 and m_2 are relatively prime, the system of congruences for y has a unique solution modulo $m_1\cdot m_2=\mathrm{lcm}(m_1,m_2)$, where $\mathrm{lcm}(m_1,m_2)$ is the least common multiple of the two. Otherwise, if m_1 and m_2 are not coprime, no solution is guaranteed, but if one exists it is unique modulo $\mathrm{lcm}(m_1,m_2)$. Therefore, unless (7) defines the empty set, there exists a non negative integer e such that formula (7) is equivalent to

$$(\exists y)\ y\le x_0\le y+d\quad\wedge y\equiv e \bmod \mathrm{lcm}(m_1,m_2)\wedge\\ \wedge\ \ x_1=y. \quad (8)$$

Niezette and Stevenne [11] define a periodic set of *starting points* which are essentially the startpoints of calendars. The temporal structures are sequences of intervals defined attaching a duration to the set of starting points. Calendars themselves are built starting from a basic one, that defines "the tick of the system".

In Niezette and Stevenne's formalism, starting points of intervals are more general, since they can be starting points of intervals of arbitrary calendars (i.e. they need not be all at the same distance each from the next), but the width of all intervals is the same—we choose to analyze here the basic version of the language, with no union operator. This implies that starting and ending point of each interval are related, and that a multi-interval pattern can be detected in the periodicity. Therefore, the general shape of a formula of PA for SPRs of Niezette and Stevenne's temporal structures can be (for some natural numbers d and a_i):

$$(\exists y) y \leq x_0 \leq y + d \ \wedge (y \equiv a_1 \bmod p \ \vee \ \dots \ \vee \vee y \equiv a_n \bmod p) \wedge x_1 = y. \quad (9)$$

Notice that Niezette and Stevenne's formalism can't define formulae of the general shape (6), since interval widths are fixed here.

Leban et al. [10] propose a richer framework that allows to group temporal intervals in sets. We disregard this aspect in the present analysis, and concentrate on the forms of periodicities allowed in instances with no structure. It is possible to recognize in a periodicity a repeating pattern, which is a sequence of intervals that occurs again and again at regular intervals.

Let us look at the endpoints of intervals in the formalism of Leban et al. Here, as we noticed, there is a repeating pattern, in which interval widths can be different. This implies formulae of the form (5) in Sect. 2.

Formulae like (8) are special cases of formulae (9), that in turn are special cases of formulae of the shape (5). These inclusions induce an ordering on the expressiveness of the related languages.

In contrast, formulae of the shape (6) define a set of temporal structures that is incomparable with the others. They allow to define structures that are far richer, but they are not expressive enough to select from the latter patterns with a clearer structure, like those definable in the other three cases. This suggests that a comparison of the "unconstrained" formalism from [8] with the symbolic ones we considered is not very significant. The conclusions above are in tune with the fact, noticed before, that the meaning of the structures definable by (6) is unclear, and concrete examples in [8] add constraints.

We summarize the above in the following table in which the entry in the row labelled X and column labelled Y is marked '=' if X is expressive enough to describe exactly structures defined by formulae of the shape Y, '>' if it is more expressive, '<' if it is not enough expressive, "nc" if it doesn't compare. We distinguish the two cases of [8]'s formalism that we studied above, indicating whether a constraint of the shape $X_2 = X_1 + d$ is present ('w/ c.') or not ('general').

	(6)	(8)	(9)	(5)
Kabanza et al., general [8]	=	nc	nc	nc
Kabanza et al., w/ c. [8]	nc	=	<	<
Niezette et al. [11]	nc	>	=	<
Leban et al. [10]	nc	>	>	>

(We indicated that the formalism of Leban et al. has more expressive power than formulae like (5), since the latter are not sufficient to express structure—see later).

Intuitively, the classification in the table above, distinguishes formalisms based on the following parameters: formulae like (6) allow the definition of families of intervals whose endpoints are taken in all possible combinations from two lists of equidistant points; formulae of the shape (8) allow to define collections of intervals of constant widths whose starting points have regular distances; formulae like (9) enable to define collections in which intervals start at points that have distances repeating periodically, but fixed widths; finally formulae like (5) allow to define sets of intervals in which both the distances between starting points and the widths repeat periodically.

We show the practical impact of the difference of expressiveness by giving examples for the entries in the left-top to right-bottom diagonal of the table above (so to speak).

An example of a temporal structure that can be expressed in the formalism of Kabanza et al. with only the basic constraint, but not in any other formalism we considered, is (from [8], Ex.2.2) the set $\{[1,1],[1,3],[1,5],\dots\}$, defined as $[1, 1+2n] \wedge X_2 \geq 1$, and expressed in PA as

$$(\exists y)\ 0 \leq x_0 \leq y \quad \wedge \quad y \equiv 1 \bmod 2 \quad \wedge \ x_1 = y.$$

On the other hand, $[3+2n_1, 5+2n_2] \wedge X_1 = X_2 - 2$ (Ex. 2.3. from [8]) that represents the infinite set $\{\dots,[1,3],[3,5],[5,7],[7,9],\dots\}$, is definable also by formalisms [11] and [10] and the corresponding formula of PA (of the form (8)) is

$$(\exists y)\ y \leq x_0 \leq y + 2 \ \wedge \ y \equiv 1 \bmod 2 \ \wedge \ x_1 = y,$$

where $y \equiv 1 \bmod 2$ is obviously equivalent to $y \equiv 3 \bmod 2 \ \wedge \ y \equiv 7 \bmod 2$.

Consider now the calendar *school-terms* from Ex. 3.6 in [11], defined as $\{9 : months : [4, 6, 2]\}$ (i.e., three terms, starting from September, of lengths 4,6 and 2 months respectively). The first two months of each school term of each year, is defined as *years+school-terms+1.months*$\triangleright$*2.months* in the formalism of [11]. Using *months* as the chronon calendar, it denotes the infinite set $\{\ldots[9, 10], [13, 14], [19, 20], [21, 22], [25, 26], \ldots\}$. This set is expressed by the formula (of the form (9)):

$$(\exists y)\ y \leq x_0 \leq y + 1\ \wedge (y \equiv 1 \bmod 12\ \vee \\ \vee\ y \equiv 9 \bmod 12\ \vee\ y \equiv 13 \bmod 12) \wedge x_1 = y.$$

An example of a periodicity that is definable only in the formalism of Leban et al. is at the end of Sect. 2, yielding formula (5).

5. Conclusions

In this paper, we introduce an intermediate layer (the SPR) to fill the gap between symbolic languages and arithmetic, and we use it to provide a Presburger-based analysis of different symbolic approaches in the literature. Using such a Presburger-based homogeneous framework, we could consider, in our comparison, also TDB constraint-based languages (we took [8] as a significant example). Elsewhere, we also used such a framework to model the semantics of a family of new languages we defined, on the basis of expressiveness criteria.

Bettini et al.'s approach [2, 3] is the approach in the literature most closely related to ours. In [2], Bettini et al. proposed a mathematical characterization of granularity, trying to introduce a standard for Temporal Databases. In [3], Bettini and De Sibi show that, given such a characterization, granularity is a superset of periodicity. They defined when a granularity G is periodical with respect to another granularity G1, pointed out a set of basic relations holding between granularities, and used their mathematical framework in order to compare the approaches of Leban et al. [10] and Niezette and Stevenne [11]. However, their mathematical framework explicitly aims to capture granularities, so that features such as "structure" (see, e.g., [10]) were deliberately not taken into account. Moreover, the expressiveness of such a framework with respect to eventually periodic sets (and/or Presburger Arithmetics) was not taken into account.

As future work, we plan to extensively apply our framework in order to extend the comparisons in this paper to other approaches, considering both the AI and the TDB literature.

Also, we will investigate how a formula of Presburger Arithmetic can be constructed starting from the SPR it defines, in an automatic fashion.

Acknowledgements

We thank the referees for their very constructive critics.

References

[1] M. Baudinet, J. Chomicki, P. Wolper, Temporal Databases: Beyond Finite Extensions, *Proc. Int'l Workshop on an Infrastructure for Temporal Databases*, 1993.

[2] C. Bettini, C. Dyreson, W. Evans, R. Snodgrass, X. Wang, A Glossary of Time Granularity Concepts, in *Temporal Databases: Research and Practice*, Springer Verlag, 1998.

[3] C. Bettini, R. De Sibi, Symbolic Representation of User-defined Time Granularities, *Proc. TIME'99*, IEEE Computer Society, 17-28, 1999.

[4] J. Chomicki, T. Imielinsky, Finite Representation of Infinite Query Answers, *ACM ToDS* 18(2), 181-223, 1993.

[5] L. Egidi, P. Terenziani, A lattice of classes of user-defined symbolic periodicities, *Proc. TIME'04*, IEEE Computer Society, 2004.

[6] L. Egidi, P. Terenziani, Orthogonal constructors for user-defined symbolic periodicities, *Tech. Rep.* TR-INF-2004-04-06-UNIPMN, Università del Piemonte Orientale, Italy, 2004.

[7] H. Enderton, *A Mathematical Introduction to Logic*, Academic Press, New York, 1972.

[8] F. Kabanza, J.-M. Stevenne, P. Wolper, Handling Infinite Temporal Data, *Journal of Computer and System Sciences* 51, 3-17, 1995.

[9] N. Koblitz, *A coourse in number theory and cryptography*, GTM 114, Springer-Verlag, 1994.

[10] B. Leban, D.D. McDonald, D.R. Forster, A representation for collections of temporal intervals, *AAAI'86*, 367-371, 1986.

[11] M. Niezette, J.-M. Stevenne, An Efficient Symbolic Representation of Periodic Time, *Proc. first Int'l Conf. Information and Knowledge Management*, 1992.

[12] P. Terenziani, Symbolic User-defined Periodicity in Temporal Relational Databases, *IEEE TKDE* 15(2), 489-509, 2003.