Journal of Universal Computer Science, vol. 18, no. 20 (2012), 2798-2831
submitted: 11/7/11, accepted: 15/10/12, appeared: 1/12/12

Crossing the Undecidability Border with Extensions of Propositional Neighborhood Logic over Natural Numbers

Dario Della Monica
(Reykjavik University, Reykjavik, Iceland
dariodm@ru.is)

Valentin Goranko
(Technical Institute of Denmark Kongens Lyngby, Denmark
and University of Johannesburg, Johannesburg, South Africa
vfgo@imm.dtu.dk)

Angelo Montanari
(University of Udine, Udine, Italy
angelo.montanari@uniud.it)

Guido Sciavicco
(University of Murcia, Murcia, Spain
guido@um.es)

Abstract: Propositional Neighborhood Logic (PNL) is an interval temporal logic featuring two modalities corresponding to the relations of right and left neighborhood between two intervals on a linear order (in terms of Allen's relations, *meets* and *met by*). Recently, it has been shown that PNL interpreted over several classes of linear orders, including natural numbers, is decidable (NEXPTIME-complete) and that some of its natural extensions preserve decidability. Most notably, this is the case with PNL over natural numbers extended with a limited form of metric constraints and with the future fragment of PNL extended with modal operators corresponding to Allen's relations *begins*, *begun by*, and *before*. This paper aims at demonstrating that PNL and its metric version MPNL, interpreted over natural numbers, are indeed very close to the border with undecidability, and even relatively weak extensions of them become undecidable. In particular, we show that (i) the addition of binders on integer variables ranging over interval lengths makes the resulting hybrid extension of MPNL undecidable, and (ii) a very weak first-order extension of the future fragment of PNL, obtained by replacing proposition letters by a restricted subclass of first-order formulae where only one variable is allowed, is undecidable (in contrast with the decidability of similar first-order extensions of point-based temporal logics).
Key Words: interval neighborhood logics, undecidability, hybrid logics, interval length binders, first-order logic
Category: F.2, F.4.1, F.4.3

1 Introduction

Reasoning about time arises naturally in various fields of computer science, including artificial intelligence, temporal databases, and software specification and verification, and temporal logics provide a natural framework for it. Among them, *Linear Temporal Logic* (LTL) [Pnu77] and *Computation Tree Logic* (CTL) [CE81] turned out to be particularly well suited for a variety of applications. In particular, they have been extensively applied to satisfiability and model checking. LTL and CTL model checking has proved itself to be a tremendously successful technology to verify requirements and design for a variety of systems, ranging from hardware systems to real-time, embedded, and safety-critical systems [CGP99, BBF$^+$01]. A number of automatic verification tools have been developed, including SMV [McM93], COSPAN [HHK96], SPIN [Hol03], and PSL [Var08]. Even though "durationless" time points are assumed to be the basic ontological temporal entities in many applications, often they are not suitable to properly reason about real-world events, which have an intrinsic *duration*, as well as to deal with essential aspects of temporality like accomplishments and temporal aggregations. These temporal features are definitely better modeled and dealt with if the underlying temporal ontology is based on time intervals (periods), rather than points, as the primitive entities. Interval temporal logics, possibly extended with a metric dimension, have been successfully applied, for instance, to the specification and verification of hardware [Mos83] and of real-time systems [ZH04, ZHR91]. In this paper, we focus our attention on one of the most interesting (decidable) interval temporal logics, namely, Propositional Neighborhood Logic.

1.1 Halpern-Shoham interval logic and fragments: undecidability rules

A systematic analysis of the variety of relations between two intervals on a linear order was initiated by Allen [All83]. He studied the distinctive properties of the thirteen binary relations that may hold between any pair of intervals on a linear order, since called Allen's relations, and he proposed the use of interval reasoning in systems for time management and planning. In [HS91], Halpern and Shoham introduced a multimodal logic, thereafter called HS, featuring modal operators for all Allen's interval relations, and they showed that such a logic is undecidable under very weak assumptions on the class of interval structures in which it is interpreted. Since then, much effort has been devoted to the search of expressive enough, decidable interval logics (see, e.g. [BMSS11a, BMSS11b, Del11, MPS10]). Unfortunately, the overwhelming majority of the interval logics studied in the literature turned out to be undecidable [Ven91, Lod00]. The underlying technical reason for the robust undecidability of HS and most of its fragments is rooted in the very nature of purely interval-based temporal reasoning, where all proposition letters, and therefore all formulae, are interpreted as true or false on intervals, rather than points in the model. This amounts to say that the set-theoretic interpretation of an HS formula in an interval model is a set of abstract intervals, that is, a set of pairs of points. Thus, all HS formulae translate into *binary relations* over the underlying linear orders, and consequently the validity (resp., satisfiability) problem for HS translates into the respective problem for the universal (resp., existential) dyadic fragment of second-order logic over linear orders. Consequently, none of the well known and widely applied decidability results for fragments of monadic second-order logic following from Rabin's theorem applies here.

1.2 An island of decidability: Propositional Neighborhood Logic (PNL) and its metric extensions

In some cases, decidability has been recovered by imposing strong syntactic and/or semantic restrictions on the logic [Mon08]. One of the few known cases of decidable interval logics with a genuine interval semantics – that is, not reducible to point-based semantics – is Propositional Neighborhood Logic (PNL) [GMS03, BGMS09]. PNL is a proper fragment of HS with two modal operators only, corresponding to Allen's relations *meets* and its inverse *met by*. It can be viewed as the propositional fragment of Zhou and Hansen's Neighborhood Logic [ZH98]. PNL has been first studied in [GMS03], and further investigated in [BGMS09]. The language of PNL is built on a set of proposition letters, the standard logical connectives, and two modal operators that allow one to move from the current interval to a right (resp., left) neighboring interval. PNL formulae take truth values over intervals, that is, ordered pairs of time points $[a, b]$, with $a \leq b$. In some interpretations (*strict* semantics), the set of admissible intervals excludes zero-length intervals (often called point-intervals), that is, intervals of the form $[a, a]$; here, we adopt the so-called *non-strict* semantics, including them. To deal with point-intervals, a modal constant π can be incorporated into the language of PNL such that π holds true over an interval if and only if it is a point-interval.

The satisfiability problem for PNL has been proved to be decidable (NEXPTIME-complete) for a number of classes of linear orders [BGMS09, BMSS11b]. In this paper, we confine ourselves to models over the natural numbers. Results in [BGMS09, MPS10] show that the addition to PNL of any other modality for Allen's relations (with the exception of the modalities for Allen's relations *before* and *after*, which are definable in PNL) makes the logic undecidable when interpreted over natural numbers. As a matter of fact, it has been shown that the addition of other modalities from Allen's repository to PNL yields undecidability for most classes of linear orders, and not only for natural numbers. The only known exception is the extension of PNL with modalities for Allen's relation *begins* and its inverse *begun by*, interpreted over finite linear orders, which turns out to be decidable, but non-primitive recursive [MPS10].

In [BDG+10, BDG+11, BMSS11a], a 'metric' extension of PNL, called *Metric* PNL (MPNL for short), has been investigated. MPNL makes use of special proposition letters expressing equality or inequality constraints on the length of the current interval with respect to fixed positive integer constants. In [BDG+10, BDG+11], Bresolin et al. prove that the satisfiability problem for MPNL over natural numbers is decidable. More precisely, they show that it is NEXPTIME-complete, when the positive integer constraints occurring in formulae are constant or represented in unary, and in between EXPSPACE and 2NEXPTIME, when they are represented in binary. EXPSPACE-completeness of the satisfiability problem for MPNL, with a binary representation of constraints, over the class of finite linear orders, natural numbers, and integer numbers, has been recently proved in [BMSS11a].

1.3 The contributions of this paper

In quest of more expressive interval logics, in this paper we explore two different, but related, ways of extending the expressive power of (metric) PNL. First, we consider the idea of adding features from *hybrid logics* to metric PNL. Hybrid modal logics (see e.g., [BS95, BT99, BdRV02]) expand modal logic with some first-order features that enable a more direct reference to the possible worlds of the model in the language. The most common features of this kind are: *nominals*, a special sort of propositional variables ranging over single possible worlds and thus serving as their names; *universal* and *difference* modalities, that refer to all (resp., all other than the current one) possible worlds; and *binders* (first introduced as *reference pointers* [Gor96]), which, by using an additional type of variables over possible worlds, called *state variables*, allow for storing the current possible world where the formula is being evaluated in the memory and later referring to it. The advantage of hybrid modal languages is that they extend naturally and considerably the expressive power of modal logic, while preserving its syntactic and semantic simplicity and often its decidability. As discussed in Section 1.4, some quite non-trivial decidability results have also been obtained for some hybrid extensions of the interval logic Duration Calculus.

What hybrid features can be meaningfully added to interval logics and, in particular, to PNL? Since the *difference* modality is already definable in PNL [GMS03], nominals are essentially definable there as well, and thus their addition is unproblematic with regards to decidability, but also uninteresting. On the other hand, it is not difficult to show that the addition of binders over state variables, even to the plain PNL, immediately leads to undecidability. However, the most natural and useful binders in the context of metric interval logics are *not* those on variables ranging over intervals, but those on variables ranging over non-negative numbers representing *lengths of intervals*. Such binders make it possible to store not the current interval itself, but only its length, and to later refer to it. However, as we show in the paper, adding to PNL only binders for interval lengths suffices to cross the undecidability border even without any metric constraints. Indeed, the first main result of this paper is a proof of the undecidability of the extension of PNL, interpreted over natural numbers, with length binders (PNL+LB for short). We note that PNL+LB is not expressively comparable with MPNL, as MPNL does not involve length binders, but it allows one to constrain the length of the current interval to be equal to (resp., less than, greater than) a certain positive integer k.

The second natural extension of PNL we consider here is a (restricted) *first-order extension* obtained by replacing propositional variables by first-order formulae, called PNL+FO. Since first-order logics are usually undecidable, one cannot expect otherwise if extending an interval logic to a full-blown first-order logic. Suitable restrictions, however, can still preserve decidability in the case of point-based temporal logics, as demonstrated e.g., in [HWZ00] (see more details in Section 1.4), which gives some hope for the case of PNL too. It turns out, however, that even a *very* limited first-order extension of PNL – with a single individual variable – interpreted over a finite first-

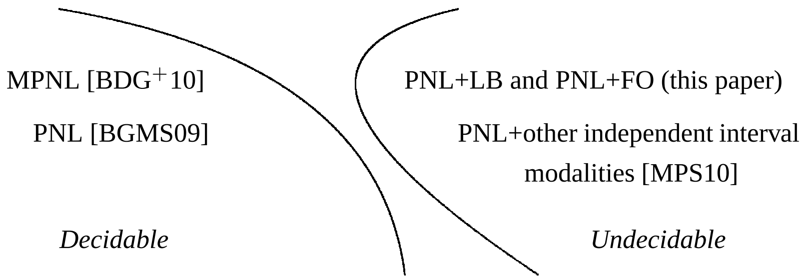

Figure 1: Decidable and undecidable extensions of PNL over natural numbers.

order domain or over $\mathbb{N}$, is already undecidable. In fact, we show that even a single modal operator is sufficient for undecidability, by proving that for purely-future fragment Right PNL (RPNL for short) [BMS07b]. Moreover, it turns out to be that the same undecidability result holds for other classes of linear orders, such as dense orders and finite orders [BMSS11b].

The rest of the paper is organized as follows. In Subsection 1.4, we briefly discuss related work. In Section 2, we give syntax and semantics of PNL and we recall basic decidability and complexity results for it. In Section 3, we introduce hybrid and first-order extensions of (metric) PNL. We provide their syntax and semantics, and we show that undecidability of most hybrid extensions of PNL can be easily derived from that of HS. Next, in Section 4, we introduce the basic features of the method we will later exploit to prove undecidability of hybrid and first-order extensions of (metric) PNL, namely, a reduction from the finite tiling problem. In Section 5, we prove the undecidability of the extension of PNL with equality constraints over length variables. A minor decidable variant of hybrid (metric) PNL is considered in Section 6. Finally, undecidability of first-order extensions of PNL is the subject of Section 7. The concluding section provides an assessment of the work and outlines future research directions.

For the reader's convenience we have depicted graphically in Figure 1 the variety of PNL extensions considered here and their (un)decidability status.

1.4 Related work

Here we briefly survey related work about hybrid and first-order extensions of propositional (metric) temporal logics. Some of the cited papers are directly, technically related to the present research, whereas other are related in a more conceptual, indirect way, by treating similar phenomena but with different languages and semantic structures.

The length variables and binders we deal with in this paper bear a natural resemblance with the interval length variables used in Duration Calculus (DC/ITL) [ZH04, HZ97] – an extension of Moszkowski's Interval Temporal Logic (ITL) [Mos83] developed by Zhou, Hoare, and Ravn [ZHR91]. The original version of ITL involves only one, binary modal operator C, called *chop*, where $\varphi C \psi$ states that the current interval $[a, b]$ can be split (chopped) into two consecutive intervals $[a, c]$ and $[c, b]$ such that $[a, c]$ satisfies φ and $[c, b]$ satisfies ψ. DC/ITL is a real-time extension of ITL that adds *state expressions* to the language of ITL to make it possible to model the states of the system; moreover, it allows one to associate a *duration* with state expressions, in order to constrain the length of the time period during which the system remains in the given state. In [ZH98], Zhou and Hansen propose a version of DC based on Neighborhood Logic (NL), denoted by DC/NL, which involves the two interval neighborhood modalities $\Diamond_r$ and $\Diamond_l$ of NL, and which, as the authors point out, subsumes the original DC/ITL. The satisfiability/validity problem for ITL, and thus those for DC/ITL and DC/NL, turns out to be undecidable over all relevant classes of linear orders.

A lot of work has been done in the search for decidable variants and fragments of DC/ITL (and of ITL). A variant of DC/ITL, called *Interval Duration Logic* (IDL), has been developed by Pandya in [Pan02]. In its full generality, such a logic is undecidable. However, it admits interesting fragments, such as LIDL$-$, which can be proved to be decidable by exploiting an automata-theoretic argument [Pan02]. Checking IDL formulae for validity has been further investigated in [CP03]. In that work, Chakravorty and Pandya provide a syntactic characterization of the proper subset of IDL-formulae that satisfy the property of *strong closure under inverse digitalization*, and they show that the problem of checking the validity of formulae belonging to such a subset can be reduced to that for *Discrete Time Duration Calculus*, a discrete-time logic whose validity problem has been shown to be decidable following an automaton-based approach in [Pan01] (a complexity improvement to such a decidability result has been given in [KP05]). In [FH07], Fränzle and Hansen prove the decidability of a quite expressive fragment of DC/ITL, properly extending the work on linear duration invariants by Zhou et al. [ZJLX94]. Other fragments of DC/ITL have been studied in [HH07]. In particular, the Restricted Duration Calculus, abbreviated RDC$_1$, allows one to constrain the length of the current interval to be equal to a given constant value. RDC$_1$ is decidable over discrete linear orders and undecidable over dense ones. Richer fragments, such as, for instance, RDC$_3$, that allows one to quantify over the variable denoting the length of the current interval, turn out to be undecidable over both discrete and dense linear orders.

Since the original formulation of ITL by Moszkowski [Mos83], an alternative path to decidability has been the enforcement of *locality*: all proposition letters are pointwise and truth over an interval is defined as truth at its initial point. The assumption of locality has also been exploited in DC/ITL to recover decidability. In [BHH07], Bolander et al. describe a hybrid extension of local DC/ITL, introducing interval binders (over intervals, not over their lengths) and nominals, that allow one to refer to specific intervals, and prove its decidability over natural numbers. This does not come as a surprise, as the locality assumption essentially reduces the logic to a point-based one, and therefore reduces its satisfiability problem to the one for monadic second-order logic (over the same linear order).

Some work has also been devoted to the model checking problem for duration calculi. For instance, in [BLR95], Bouajjani at al. address the problem of specifying and verifying hybrid systems in the framework of duration calculi, exploiting techniques borrowed from hybrid automata. Model checking algorithms for duration calculi have been developed in [Frä04, FH08, MFHR08].

First-order point-based temporal logics have been systematically studied by Hodkinson et al. in [HWZ00]. They show that (un)decidability of such logics depends on both the classical (first-order) and the temporal components of the language. In particular, they prove that the two-variable fragment of first-order Linear Temporal Logic (LTL), with Since and Until, interpreted over $\mathbb{N}$ and $\mathbb{Z}$, is undecidable. The same results hold for LTL with Next and Future modalities only. Then, they show that decidability can be recovered by restricting the first-order component to a decidable fragment of first-order logic and the temporal one to *monodic* formulae, that is, formulae whose sub-formulae with a temporal operator as their outermost operator have at most one free variable. In particular, they prove that the two-variable fragment of monodic first-order LTL (without equality and function symbols, and with constant first-order domains) is decidable over various linear time structures, including $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$, and $\mathbb{R}$ (the latter holds for finite first-order domains only). In the following, we will show that there is not a counterpart of these decidability results in the setting of first-order extensions of PNL.

2 Preliminaries

The language of Propositional Neighborhood Logic (PNL) consists of a set $\mathcal{AP}$ of proposition letters, the propositional connectives $\neg$ and $\vee$, and the modal operators $\Diamond_r$ and $\Diamond_l$, corresponding to Allen's interval relation *meets* and its inverse *met-by*, respectively [All83]. The other propositional connectives, as well as the logical constants $\top$ (*true*) and $\bot$ (*false*), and the dual modal operators $\Box_r$ and $\Box_l$, are defined as usual ($\Box_r = \neg\Diamond_r\neg$ and $\Box_l = \neg\Diamond_l\neg$). PNL has been studied both by assuming the so-called *strict semantics*, which excludes point-intervals, and by assuming the *non-strict* one, which includes them. In the latter case, it is natural to include in the language a special proposition letter (modal constant), usually denoted by π, that is true over all and only the point-intervals. A systematic analysis and comparison of the expressiveness of the various PNL instances can be found in [GMS03, BGMS09].
PNL formulae, denoted by $\varphi, \psi, \ldots$, are generated by the following grammar:

$$\varphi ::= \pi \mid \mathsf{p} \mid \neg\varphi \mid \varphi \vee \varphi \mid \Diamond_r\varphi \mid \Diamond_l\varphi.$$

The future fragment of PNL (Right PNL, RPNL for short) is obtained by removing the past modality $\Diamond_l$.

Given a linearly ordered domain $\mathbb{D} = \langle D, < \rangle$, a (*non-strict*) *interval* over $\mathbb{D}$ is any ordered pair $[a, b]$, with $a \leq b$. We denote by $\mathbb{I}(\mathbb{D})$ the set of all intervals over $\mathbb{D}$. The semantics of PNL is given in terms of models of the form $M = \langle \mathbb{D}, V \rangle$, where

$V : \mathcal{AP} \to 2^{\mathbb{I}(\mathbb{D})}$ is a valuation function assigning to every proposition letter the set of those intervals over which it is true (notice that no conditions on the valuation, such as locality or homogeneity, are imposed). We recursively define the satisfiability relation $\Vdash$ as follows:

- $M, [a, b] \Vdash \pi$ iff $a = b$;
- $M, [a, b] \Vdash \mathsf{p}$ iff $[a, b] \in V(\mathsf{p})$, for any $\mathsf{p} \in \mathcal{AP}$;
- $M, [a, b] \Vdash \neg\psi$ iff it is not the case that $M, [a, b] \Vdash \psi$;
- $M, [a, b] \Vdash \psi \vee \tau$ iff $M, [a, b] \Vdash \psi$ or $M, [a, b] \Vdash \tau$;
- $M, [a, b] \Vdash \Diamond_r \psi$ iff there exists $c \geq b$ such that $M, [b, c] \Vdash \psi$;
- $M, [a, b] \Vdash \Diamond_l \psi$ iff there exists $c \leq a$ such that $M, [c, a] \Vdash \psi$.

The satisfiability problem for (various instances of) PNL has been shown to be decidable in [BMS07a, BGMS09, BMSS11b, MS12].

Theorem 1. *The satisfiability problem for PNL, over the classes of all linear orders, well-orders, dense linear orders, discrete linear orders, and finite linear orders, as well as over* $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, *and* $\mathbb{R}$, *is NEXPTIME-complete.*

In [BDG$^+$11], Bresolin et al. develop a metric extension of PNL interpreted over finite linear orders (resp., natural numbers, integers), called MPNL. In this paper, we restrict our attention to the class of interval structures over the ordering of the natural numbers, that is, we assume $\mathbb{D} = \mathbb{N}$. As a matter of fact, most results also hold for $\mathbb{Z}$ as well as for various other linear orders on which a distance function is definable. MPNL extends PNL by featuring *proposition letters for length constraints*. These (pre)interpreted proposition letters allow one to refer to the length of the current interval, and can be viewed as the metric generalization of the modal constant π. Formally, let $\mathbb{C} = \{<, \leq, =, \geq, >\}$. For each $\mathcal{C} \in \mathbb{C}$ and $k \in \mathbb{N}$, the length constraint $\mathsf{len}_{\mathcal{C}k}$ is defined as follows:

$$M, [a, b] \Vdash \mathsf{len}_{\mathcal{C}k} \ \text{ iff } \ \delta(a, b)\mathcal{C}k,$$

where $\delta : \mathbb{N} \times \mathbb{N} \to \mathbb{N}$ is the *distance* function on $\mathbb{N}$, defined as $\delta(a, b) = |a - b|$.

Decidability and complexity of the satisfiability problem for MPNL, interpreted over $\mathbb{N}$, have been investigated in [BDG$^+$10, BDG$^+$11, BMSS11b].

Theorem 2. *The satisfiability problem for MPNL, interpreted over* $\mathbb{N}$, *is NEXPTIME-complete, if length constraints are either constants or represented in unary, and it is EXPSPACE-complete, if they are represented in binary.*

Expressiveness and (potential) applications of MPNL have been extensively discussed in [BDG$^+$10, BDG$^+$11]. We refer the reader to such a publication for details.

Here, we provide a short summary of its outcomes. First, it has been shown that MPNL is expressive enough to encode classical point-based temporal operators, like 'sometimes in the future' and 'sometimes in the past', as well as to define a metric version of the 'until' operator. Moreover, it has been proved that almost all Allen's relations (all but the relation *during*), over bounded intervals, can be expressed in MPNL, by making use of the universal modality $[G]$ ($[G]\phi$ ensures that ϕ holds over every interval of the model) and the difference modality $[\neq]$ ($[\neq]\phi$ ensures that ϕ holds over every interval of the model but the current one), which can be defined in pure PNL (see next section). Finally, a set of application examples, ranging from formal specification of complex systems (like a gas burner or a railway signaling system) to medical guidelines and ambient intelligence, has been given.

3 Hybrid and First-Order Extensions of (Metric) PNL

In this section, we introduce hybrid and first-order extensions of (metric) PNL. In the first part, we focus our attention on hybrid extensions. We start by showing that some typical components of hybrid logics, such as nominals, can actually be defined in PNL. Then, we discuss the effects of the addition of binders to (metric) PNL. In the second part, we take into consideration first-order extensions of PNL. Both extensions will be investigated in detail in the following sections.

3.1 Metric Hybrid Extensions

Despite its simplicity, PNL makes it possible to define significant hybrid features such as nominals. Let $[G]$ be the *universal modality*, that is, given a formula φ, $[G]\varphi$ is true over an interval if and only if φ is true over all intervals. The universal modality can be defined in all variants of PNL. For instance, if the non-strict semantics is assumed, it can be defined as follows: $[G]\varphi \equiv \Box_l\Box_l\Box_r\Box_r\varphi$. The same holds for the *difference modality* $[\neq]$. In the strict semantics, it can be defined as follows [GMS03]:

$$[\neq]\varphi \equiv \Box_l\Box_l\Box_r\varphi \wedge \Box_l\Box_r\Box_r\varphi \wedge \Box_r\Box_l\Box_l\varphi \wedge \Box_r\Box_r\Box_l\varphi.$$

Such a formula can be easily modified to define the modality $[\neq]$ when non-strict semantics is assumed (the revised formula makes an essential use of the modal constant π for point-intervals). Thus, nominals over intervals can be simulated in PNL, and therefore this (basic) hybrid extension of PNL remains decidable over a large family of linear orders, including $\mathbb{N}$. However, it is quite easy to see that the addition of stronger hybrid features, such as binders or quantifiers over intervals, immediately leads to undecidability, even under very weak assumptions about the class of linear orders.

In this paper, we focus our attention on the addition of length binders to MPNL. Besides binders on state variables, ranging over intervals, one may introduce binders on integer variables, ranging over *interval lengths*. In its "classical" version, MPNL involves metric constraints expressed by constants. As an example, $\Diamond_r(\mathsf{len}_{=5} \wedge \mathsf{p} \rightarrow \Diamond_l\Diamond_r\mathsf{q})$ is a well-formed MPNL formula, while $\Diamond_r(\mathsf{len}_{=x} \wedge \mathsf{p})$, for some variable x, is not. This means that, despite the fact that MPNL can be considered a quite expressive interval logic (a number of meaningful examples of temporal conditions that can be specified in MPNL are given in [BDG$^+$10, BDG$^+$11]), there are simple and natural properties that it cannot express, such as, for instance, *the right neighbor interval, whose length is equal to the length of the current interval, satisfies the property* q. To deal with properties like this one, we extend the language of MPNL with a sort of hybrid machinery making it possible to store the length of the current interval and to use it further in formulae.

Let us denote by MPNL+LB the hybrid extension of MPNL with length binders, which is defined as follows. First, we introduce a binder $\downarrow$, called *length binder*, a countable set of *length variables* $DVar = \{x, y, \ldots\}$, where $DVar \cap \mathcal{AP} = \emptyset$, and a set of *hybrid metric constraints* of the form $\mathsf{len}_{\mathcal{C}x}$, for each $\mathcal{C} \in \mathbb{C}$ and $x \in DVar$. Semantics of MPNL+LB formulae is defined as usual, pairing the classical valuation function for proposition letters with a *length assignment* $g : DVar \rightarrow \mathbb{N}$. An MPNL+LB model over $\mathbb{N}$ is a triplet $M = \langle \mathbb{N}, V, g\rangle$, where $V : \mathcal{AP} \rightarrow 2^{\mathbb{I}(\mathbb{N})}$ is the valuation function for proposition letters and g is the length assignment. For any pair of length assignments g, g' and any variable x, we write $g' \sim_x g$ to mean that g' possibly differs from g on the value of x only. Formally, formulae are defined by the following grammar:

$$\varphi ::= \mathsf{p} \mid \mathsf{len}_{\mathcal{C}k} \mid \mathsf{len}_{\mathcal{C}x} \mid \neg\varphi \mid \psi \vee \varphi \mid \Diamond_r\varphi \mid \Diamond_l\varphi \mid \downarrow_x \varphi,$$

where $k \in \mathbb{N}$ and $x \in DVar$.

Let $M = \langle \mathbb{N}, V, g\rangle$. The semantic rules for MPNL+LB consist of those for MPNL plus the following clauses:

– $M, [a, b] \Vdash \mathsf{len}_{\mathcal{C}x}$ iff $\delta(a, b)\mathcal{C}g(x)$;

– $M, [a, b] \Vdash\downarrow_x \varphi$ iff $M', [a, b] \Vdash \varphi$ for $M' = \langle \mathbb{N}, V, g'\rangle$, where g' is a length assignment such that $g' \sim_x g$ and $g'(x) = \delta(a, b)$.

It is worth pointing out that a universal analogue of the hybrid operator @, with the following semantics:

– $M, [a, b] \Vdash @_x\varphi$ iff for every interval $[c, d]$ such that $\delta(c, d) = g(x)$ it is the case that $M, [c, d] \Vdash \varphi$,

can be easily defined in MPNL+LB as follows:

$$@_x\varphi := [G](\mathsf{len}_{=x} \rightarrow \varphi).$$

The same holds for the existential analogue of @.

We show now that undecidability of almost all extensions of MPNL with length binders can be easily proved by a reduction from the satisfiability problem for undecidable fragments of HS, the only difficult case being that of the extension of MPNL with equality constraints over length variables ($\mathsf{len}_{=x}$, with $x \in DVar$), which will be dealt with in Section 5. As a matter of fact, we prove a stronger result showing that the fragment PNL+LB of MPNL+LB, devoid of proposition letters for length constraints over constants ($\mathsf{len}_{\mathcal{C}k}$, with $k \in \mathbb{N}$), is already undecidable.

Unlike what happens with proposition letters for length constraints over constants, that is, $\mathsf{len}_{=k}$, $\mathsf{len}_{>k}$, $\mathsf{len}_{\geq k}$, $\mathsf{len}_{<k}$, and $\mathsf{len}_{\leq k}$, with $k \in \mathbb{N}$, which are known to be definable in terms of each other, no general interdefinability rules are known for constraints over length variables. As an example, it can be easily shown that $\mathsf{len}_{\geq x}$ is equivalent to $\neg\mathsf{len}_{<x}$, but we are not aware of any way of expressing $\mathsf{len}_{\leq x}$ or $\mathsf{len}_{<x}$ in terms of $\mathsf{len}_{=x}$ (since the value of x is unknown, $\mathsf{len}_{\leq x}$ and $\mathsf{len}_{<x}$ cannot be expressed as logical disjunctions of equalities, as in the case of constant metric constraints). The undecidability of PNL+LB immediately follows from that of HS, as HS operators $\langle B\rangle$, $\langle E\rangle$, $\langle\overline{B}\rangle$, and $\langle\overline{E}\rangle$, that is, the modalities corresponding to Allen's relations *starts*, *finishes*, *started-by*, and *finished-by*, respectively, which suffice to define all other HS operators when non-strict semantics is assumed, can be easily defined in it as follows:

$$\begin{aligned}
\langle B\rangle \mathsf{p} &:= \downarrow_x \Diamond_l \Diamond_r (\mathsf{p} \wedge \mathsf{len}_{<x}),\\
\langle E\rangle \mathsf{p} &:= \downarrow_x \Diamond_r \Diamond_l (\mathsf{p} \wedge \mathsf{len}_{<x}).\\
\langle\overline{B}\rangle \mathsf{p} &:= \downarrow_x \Diamond_l \Diamond_r (\mathsf{p} \wedge \mathsf{len}_{>x}),\\
\langle\overline{E}\rangle \mathsf{p} &:= \downarrow_x \Diamond_r \Diamond_l (\mathsf{p} \wedge \mathsf{len}_{>x}).
\end{aligned}$$

Theorem 3. *The satisfiability problem for PNL+LB over* $\mathbb{N}$ *is undecidable.*

Proof. As shown in [BDG+08, Del11], the HS fragment featuring the pair of modalities $\langle B\rangle$ and $\langle E\rangle$ (resp., $\langle B\rangle$ and $\langle\overline{E}\rangle$, $\langle\overline{B}\rangle$ and $\langle E\rangle$, $\langle\overline{B}\rangle$ and $\langle\overline{E}\rangle$) only, interpreted over $\mathbb{N}$, is undecidable. Hence, the fragments of PNL+LB featuring only one length variable and only one (type of) constraint among $\{<, \leq, >, \geq\}$ are already undecidable. The only remaining case is that of the PNL+LB fragment with one length variable and length constraints of the form $\mathsf{len}_{=x}$. We will complete the proof of the theorem by proving its undecidability in Section 5.

Undecidability of the satisfiability problem for MPNL+LB over $\mathbb{N}$ immediately follows.

3.2 First-Order Extensions

We now consider a completely different extension of PNL over $\mathbb{N}$, which is obtained by lifting it to the first-order setting. We call the resulting logic PNL+FO. PNL+FO is obtained from PNL by replacing proposition letters by *predicate symbols* $P, Q, \ldots$ of fixed arity (proposition letters can be recovered as 0-ary predicate symbols) and by adding a set of *individual variables* $x, y, \ldots$, a set of *individual constants* $c_1, c_2, \ldots$, that is, functions of arity 0 (for the sake of simplicity, we exclude function symbols of arity greater than 0), and the *universal (first-order) quantifier* $\forall$. The *terms* $\tau_1, \tau_2, \ldots$ are either individual variables or individual constants. As usual, the existential (first-order) quantifier can be defined in terms of the universal one: $\exists x\varphi(x) \equiv \neg\forall x\neg\varphi(x)$.

A *first-order interval model* is a tuple $M = \langle \mathbb{N}, \mathfrak{D}, \mathcal{I} \rangle$, where $\mathfrak{D}$ is the *first-order domain* of M and $\mathcal{I}$ is the *interpretation function* that maps every interval of $\mathbb{I}(\mathbb{N})$ into a *first-order structure*:

$$\mathcal{I}([a,b]) = \langle \mathfrak{D}, P^{\mathcal{I}([a,b])}, Q^{\mathcal{I}([a,b])}, \ldots \rangle.$$

For every interval $[a,b]$ and predicate symbol P, $P^{\mathcal{I}([a,b])}$ is a relation on $\mathfrak{D}$ with the same arity as P (for proposition letters, it is simply *true* or *false*).

An *assignment* λ is a function that maps terms into elements of $\mathfrak{D}$. We assume constants to be *rigid*, that is, we assume that each constant refers to the same element of $\mathfrak{D}$ regardless of which is the current interval.

The set of semantic clauses for PNL+FO is obtained from that for PNL by adding the assignment as an additional parameter, by replacing the clause for proposition letters by a clause for predicates, and by introducing a clause for the universal quantifier:

– for each predicate symbol P, $M, [a,b], \lambda \Vdash P(\tau_1, \ldots, \tau_n)$ iff $P^{\mathcal{I}([a,b])}(\lambda(\tau_1), \ldots, \lambda(\tau_n))$;

– $M, [a,b], \lambda \Vdash \forall x \psi$ iff $M, [a,b], \lambda' \Vdash \psi$ for every assignment λ' that differs from λ at most for the value of x.

PNL+FO can thus be viewed as a *limited* first-order generalization of PNL: it allows one to move along the time domain by applying the modalities and to formulate specific statements about what is true over a given interval by using first-order constructs.

What can we say about the (un)decidability of PNL+FO? On the one hand, first-order modal and temporal logics are usually undecidable. On the other hand, there are at least two important decidability results in the first-order setting, that are relevant here: (i) the decidability of the two-variable fragment of first-order logic [BGG97], and (ii) the decidability of the two-variable fragment of first-order logic interpreted over various classes of linear orders, in particular, over the class of all linear orders and over $\mathbb{N}$ [Ott01]. In the framework of temporal logics, as we have already pointed out, it has been shown that a first-order extension of LTL (with Since and Until, but the result also applies to the fragment with Future and Next only) where two distinct variables may be used yields undecidability [HWZ00]. To recover decidability, one must restrict the language by allowing one variable only. We will show that in the interval setting the situation is way worse: the addition of very elementary first-order ingredients to PNL suffices to cross the undecidability border.

4 A General Path to Undecidability

The undecidability of both hybrid and first-order extensions of (metric) PNL will be proved by exploiting reduction from the *Finite Tiling Problem* (FTP for short). The method for proving undecidability via tiling is very common, but we have adapted it here in a quite non-trivial way, building on our previous developments of that technique (see, for instance, [BDG+08, BDG+09, BDMG+11]). In this section, we provide the reader with a gentle introduction to the method.

The FTP is the problem of establishing whether, given a finite set of tile types $\mathcal{T} = \{t_1, \ldots, t_k\}$, there exists a finite rectangle $\mathcal{R} = [0, X] \times [0, Y] = \{(i, j) : i, j \in \mathbb{N} \wedge 0 \leq i \leq X \wedge 0 \leq j \leq Y\}$, for some $X, Y \in \mathbb{N}$, such that $\mathcal{T}$ can correctly tile $\mathcal{R}$ with the entire border colored by the same designated color $\$$, also called *side color*. More precisely, for every tile type $t_i \in \mathcal{T}$, let *right*(t_i), *left*(t_i), *up*(t_i), and *down*(t_i) be the colors of the corresponding sides of t_i. To solve the FTP for $\mathcal{T}$, one must find two natural numbers X and Y, and a mapping $f : \mathcal{R} \rightarrow \mathcal{T}$ such that

$$\begin{array}{l} \forall\, 0 \leq i < X, 0 \leq j \leq Y \ (\mathit{right}(f(i,j)) = \mathit{left}(f(i+1,j))); \\ \forall\, 0 \leq i \leq X, 0 \leq j < Y \ (\mathit{up}(f(i,j)) \quad = \mathit{down}(f(i,j+1))), \end{array}$$

and

$$\begin{array}{l} \forall\, 0 \leq j \leq Y \ (\mathit{left}(f(0,j)) = \$); \quad \forall\, 0 \leq j \leq Y \ (\mathit{right}(f(X,j)) = \$); \\ \forall\, 0 \leq i \leq X \ (\mathit{down}(f(i,0)) = \$); \forall\, 0 \leq i \leq X \ (\mathit{up}(f(i,Y)) = \$). \end{array}$$

The FTP has been first introduced and shown to be undecidable in [LPRT95].

In order to perform the reduction from the FTP for a given set of tile types $\mathcal{T} = \{t_1, \ldots, t_k\}$ to the satisfiability problem for the logic under consideration, we will make use of some special proposition letters, that depend on the logic. For every proposition letter p, we call any interval satisfying p a p-*interval*. The reduction consists of three main steps:

(i) the encoding of the rectangle by means of a suitable chain of so-called 'unit' intervals (u-intervals for short);

(ii) the (difficult) encoding of the 'above-neighbor' relation by means of a suitable family of so-called Up_rel-intervals;

(iii) the (easy) encoding of the 'right-neighbor' relation.

The idea of the encoding is as follows. First, we introduce a set of proposition letters $\mathtt{T} = \{\mathtt{t}_1, \mathtt{t}_2, \ldots, \mathtt{t}_k\}$ corresponding to the set of tile types $\mathcal{T} = \{t_1, t_2, \ldots, t_k\}$. Next, we set our framework by forcing the existence of a unique finite chain of u-intervals on the linear ordering (u-*chain*), u-intervals being used as cells to arrange the tiling. Then, we define a chain of Id-intervals (Id-*chain*), each one consisting of a sequence of u-intervals and representing a row of the rectangle (Id is a shorthand for identifier borrowed from [HS91]). Finally, the above-neighbor relation that connects each tile with its above neighbor in $\mathcal{R}$ is encoded by means of the proposition letter Up_rel. The last step is the definition of a formula Φ which is satisfiable if and only if there is a finite rectangle $\mathcal{R}$ for some $X, Y \in \mathbb{N}$ and a proper tiling of $\mathcal{R}$ by $\mathcal{T}$, that is, a tiling that satisfies the color constraints on the border tiles and between vertically- and horizontally-adjacent tiles.

5 Undecidability of PNL + equality constraints over length variables

As we have already shown in Section 3.1, the fragments of PNL+LB with only one length variable and only one type of constraint from the set $\{<, \leq, >, \geq\}$ are undecidable. In this section, we provide a reduction of the above-described FTP to the satisfiability problem for the fragment of PNL+LB with only one length variable and length constraints of the form $\mathsf{len}_{=x}$ only, thus proving its undecidability. Together with those given in Section 3.1, this result allows one to conclude that extending PNL with length binders always leads to undecidability, even when only one length variable is used and regardless of the kind of length constraints allowed. As a matter of fact, we will show that the use of π in the undecidability proof is inessential: we first give an encoding of FTP that makes use of π, and then we show how to get rid of it.

The u-chain is defined by the following set of formulae:

$\mathsf{Start} \wedge \Box_r \neg\mathsf{Start} \wedge \Diamond_r \Diamond_r \mathsf{Stop}$	(1)	*beginning / ending the* u*-chain*
$[G]((\mathsf{Start} \vee \mathsf{Stop} \rightarrow \mathsf{u}) \wedge (\mathsf{u} \rightarrow \mathsf{len}_{=x}))$	(2)	u, Start, *and* Stop *same length*
$[G](\Diamond_r \mathsf{Start} \rightarrow \Box_r(\neg\pi \rightarrow \Box_r \neg\mathsf{Start}))$	(3)	Start *is unique*
$[G](\Diamond_r \mathsf{Stop} \rightarrow \Box_r(\neg\pi \rightarrow \Box_r \neg\mathsf{Stop}))$	(4)	Stop *is unique*
$[G](\mathsf{u} \wedge \neg\mathsf{Stop} \rightarrow \Diamond_r \mathsf{u})$	(5)	u*-chain to the right*
$[G]((\mathsf{Start} \rightarrow \Box_l \Box_l \neg\mathsf{u}) \wedge (\mathsf{Stop} \rightarrow \Box_r \Box_r \neg\mathsf{u}))$	(6)	*no* u *out of the chain*
$(1) \wedge \ldots \wedge (6)$	(7)	

Lemma 4. *Let $M = \langle \mathbb{N}, V, g \rangle$ be a PNL+LB model and $[a, b] \in \mathbb{I}(\mathbb{N})$ such that*

$$M, [a, b] \Vdash \downarrow_x (7).$$

Then, there exists a finite sequence of points $b_0 < b_1 < \ldots < b_k$, with $k > 0$, such that $b_0 = a$, $b_1 = b$, and

1. *all intervals $[b_i, b_{i+1}]$, for $0 \leq i \leq k-1$, have the same length $b - a > 0$;*
2. *$M, [b_i, b_{i+1}] \Vdash \mathsf{u}$ for each $0 \leq i \leq k-1$;*
3. *no other interval satisfies* u.

Proof. First of all, by (1), the interval $[a, b] (= [b_0, b_1])$ satisfies Start and it is not a point-interval, as it satisfies $\Box_r \neg\mathsf{Start}$ as well; moreover, there exists a Stop-interval to the right of it. By (2), Start- and Stop-intervals are u-intervals, and all u-intervals have the same length (equal to $b - a > 0$). Hence, two different Start-intervals (resp., Stop-intervals, u-intervals) cannot start at the same point. Then, from (3) (resp., (4)), it follows that the interval satisfying Start (resp., Stop) is unique.

Next, by (1), (2), (5), and (6), the interval $[b_0, b_1]$ starts a finite chain of u-intervals $[b_i, b_{i+1}]$, with $0 \leq i \leq k-1$. The finiteness follows from the fact that, by (1) and (2), some future u-interval satisfies Stop and, by (6), there are no u-intervals starting to the right of it. Moreover, the (unique) Stop-interval must belong to the u-chain, otherwise, by (5), the u-chain would go beyond the Stop-interval, that is, there would be a u-interval following the Stop-interval, thus contradicting (6). Hence, the (unique) Stop-interval must be the last u-interval of the u-chain, that is, the interval $[b_{k-1}, b_k]$.

To conclude the proof, suppose, by contradiction, that there exists a u-interval $[c, d]$ not belonging to the u-chain. By (6), it can be neither to the left of the Start-interval $[b_0, b_1]$ nor to the right of the Stop-interval $[b_{k-1}, b_k]$. Thus, it must lie in between b_0 and b_k, and it must start another chain of u-intervals, all of the same length $b - a$ (by (2)), whose u-intervals overlap those of the first u-chain. However, the unique interval satisfying Stop cannot belong to this second u-chain, and thus it will be crossed by it, leading to a contradiction with (6). □

We now define the Id-chain with the following formulae:

$\mathsf{IdStart} \wedge \Diamond_r \Diamond_r \mathsf{IdStop}$	(8)	*beginning / ending the* Id*-chain*
$[G]((\mathsf{IdStart} \vee \mathsf{IdStop} \rightarrow \mathsf{Id}) \wedge (\mathsf{Id} \rightarrow \mathsf{len}_{=x}))$	(9)	Id, IdStart, IdStop *same length*
$[G](\Diamond_r \mathsf{Start} \leftrightarrow \Diamond_r \mathsf{IdStart})$	(10)	IdStart *is unique*
$[G](\Diamond_l \mathsf{Stop} \leftrightarrow \Diamond_l \mathsf{IdStop})$	(11)	IdStop *is unique*
$[G]((\mathsf{u} \leftrightarrow \mathsf{Tile} \vee *) \wedge (* \rightarrow \neg \mathsf{Tile}))$	(12)	u *is either* Tile *or* $*$
$[G]((\Diamond_r \mathsf{Id} \leftrightarrow \Diamond_r *) \wedge (\mathsf{Id} \rightarrow \Diamond_r \Diamond_l \mathsf{Tile}))$	(13)	Id*s begin / end with* * / Tile
$[G](\mathsf{Id} \wedge \neg \mathsf{IdStop} \rightarrow \Diamond_r \mathsf{Id})$	(14)	Id*-chain to the right*
$(8) \wedge \ldots \wedge (14)$	(15)	

Lemma 5. *Let $M = \langle \mathbb{N}, V, g \rangle$ be a PNL+LB model and $[a, b] \in \mathbb{I}(\mathbb{N})$ such that*

$$M, [a, b] \Vdash \downarrow_x (7) \wedge \Diamond_l \Diamond_r \downarrow_x (15).$$

Then, there exist two positive integers h, v and a finite sequence of points $b_1^0 = a < b_1^1 = b < \ldots < b_1^h = b_2^0 < \ldots < b_2^h = b_3^0 < \ldots < b_{v-1}^h = b_v^0 < \ldots < b_v^h$ such that for each $1 \leq j \leq v$, we have:

1. $M, [b_j^0, b_j^1] \Vdash *$;
2. $M, [b_j^i, b_j^{i+1}] \Vdash \mathsf{Tile}$ *for each* $0 < i < h$;
3. $M, [b_j^0, b_j^h] \Vdash \mathsf{Id}$.

Moreover, no other interval satisfies $*$, Tile, or Id.

Proof. First of all, by Lemma 4, there is a finite sequence of points $b_0 = a < b_1 = b < \ldots < b_k$, which defines a finite chain of u-intervals. By (12), each of these u-intervals is either a $*$-interval or a Tile-interval and no other interval is a $*$-interval or a Tile-interval.

By (9), IdStart- and IdStop-intervals are Id-intervals, and all Id-intervals have the same length. By (10) and (11), IdStart-intervals can only start where the unique Start-interval (u-interval $[b_0, b_1]$) starts, and IdStop-intervals can only end where the unique Stop-interval (u-interval $[b_{k-1}, b_k]$) ends. Since all IdStart-intervals (resp., IdStop-intervals) have the same length, it immediately follows that there is a unique IdStart-interval (resp., IdStop-interval).

We prove now that Id-intervals (including the IdStart- and the IdStop-interval) are neither point-intervals nor u-intervals. By contradiction, let us assume Id-intervals to be point-intervals. By (13) (second conjunct), it follows that there exists a Tile-interval to the left of IdStart, and thus, by (10), to the left of Start as well, which contradicts (6) (no u-interval before Start). Together with (8), this allows us to conclude that the IdStart-interval and the IdStop-interval are distinct (the IdStop-interval is to the right of the IdStart-interval). Now, we show that Id-intervals cannot be u-intervals. Let us assume, by contradiction, that they are u-intervals. By (13) (first conjunct), it follows that the IdStart-interval is a $*$-interval and by (13) (second conjunct) it follows that it is a Tile-interval, which contradicts (12) (a u-interval is either a $*$-interval or a Tile-interval).

Now, we show that each Id-interval spans a round number of u-intervals. By (13) (first conjunct), every Id-interval starts with a $*$-interval, and, by (13) (second conjunct), it ends with a Tile-interval. By Lemma 4 and (12), this implies that every Id-interval spans a round number, greater than one, of u-intervals. By (9) (second conjunct), this round number is the same, say h, for every Id-interval.

We prove now that Id-intervals are arranged to form a unique (finite) chain. Let us consider the unique IdStart-interval. By (10), it starts with a Start-interval. By Lemma 4, there are not u-intervals to the left of such a unique Start-interval, and thus, by (13), there are no Id-intervals starting (strictly) to the left of it. Since the IdStart-interval is not an IdStop-interval, by (14), it follows that it starts a chain of Id-intervals. Such a chain must be finite, as there are only finitely many u-intervals, and its last Id-interval must be the unique IdStop-interval. Indeed, if this was not the case, there would be an Id-interval crossing the IdStop-interval, and thus an u-interval to the right of the IdStop-interval (contradiction, by (11), Lemma 4, and (13)). Thus, such a sequence of Id-intervals define a partition of the sequence of u-intervals. Let v be the total number of u-intervals divided by h. The sequence $b_0 < b_1 < \ldots < b_k$ can be rewritten as $b_1^0 = b_0 < b_1^1 = b_1 < \ldots < b_1^h = b_2^0 < \ldots < b_2^h = b_3^0 < \ldots < b_{v-1}^h = b_v^0 < \ldots < b_v^h = b_k$, as required. Points 1, 2, and 3 of the lemma immediately follow.

To complete the proof, we only need to show that no other interval in between b_1^0 and b_v^h satisfies Id. To this end, it suffices to observe that any such Id-interval would start a second chain of Id-intervals overlapping those of the first Id-chain. Since the unique IdStop-interval cannot belong to this second Id-chain, it will be crossed by it (contradiction). $\square$

The above lemma guarantees the existence of a unique Id-chain. Now, we want to constrain the proposition letter $\mathsf{Up\_rel}$ to correctly encode the relation that connects pairs of tiles of the rectangle that are vertically adjacent. We force such a condition by means of the following set of formulae:

$$[G](\mathsf{Up\_rel} \to \mathsf{len}_{=x} \wedge \Diamond_l \Diamond_r \mathsf{Tile}) \quad (16) \qquad \mathsf{Up\_rel} \textit{ and } \mathsf{Id} \textit{ same length}$$

$$[G](\mathsf{Tile} \to (\Diamond_r \Diamond_r \mathsf{IdStop} \leftrightarrow \Diamond_l \Diamond_r \mathsf{Up\_rel})) \quad (17) \qquad \mathsf{Tile} \textit{ begins } \mathsf{Up\_rel}$$

$$(16) \wedge (17) \quad (18)$$

Lemma 6. *Let $M = \langle \mathbb{N}, V, g \rangle$ be a PNL+LB model and $[a, b] \in \mathbb{I}(\mathbb{N})$ such that*

$$M, [a, b] \Vdash \downarrow_x (7) \wedge \Diamond_l \Diamond_r \downarrow_x ((15) \wedge (18))$$

and let $b_1^0 = a < b_1^1 = b < \ldots < b_1^h = b_2^0 < \ldots < b_2^h = b_3^0 < \ldots < b_v^0 < \ldots < b_v^h$ be the sequence of points whose existence is guaranteed by Lemma 5. Then, for each $1 \leq j < v$ and $1 \leq i < h$, the interval $[b_j^i, b_{j+1}^i]$ satisfies $\mathsf{Up\_rel}$, and no other interval satisfies $\mathsf{Up\_rel}$.

Proof. As (15) and (18) are in the scope of the same length binder $\downarrow_x$, by (16), $\mathsf{Up\_rel}$-intervals have the same length as Id-intervals. Moreover, by (17), each Tile-interval, but the ones belonging to the last Id-interval, starts an $\mathsf{Up\_rel}$-interval. Finally, by (16), each $\mathsf{Up\_rel}$-interval is started by a Tile-interval. Given that the length of all u-intervals is the same and every Id-interval spans the same number of u-intervals, the claim immediately follows from Lemma 5. $\square$

Finally, we can force all color-matching conditions to be fulfilled by means of the following set of formulae, where $\mathcal{T}_r$ (resp., $\mathcal{T}_l$, $\mathcal{T}_t$, $\mathcal{T}_b$) is the subset of $\mathcal{T}$ containing all and only those tiles whose right (resp., left, up, down) side is colored with \$.

$$[G]((\mathsf{Tile} \leftrightarrow \bigvee_{t_q \in \mathcal{T}} \mathtt{t_q}) \wedge \bigwedge_{t_q, t_{q'} \in \mathcal{T}, t_q \neq t_{q'}} \neg(\mathtt{t_q} \wedge \mathtt{t_{q'}})) \quad (19) \qquad \mathsf{Tile}\textit{s are tiles}$$

$$[G](\mathsf{Tile} \wedge \Diamond_r \mathsf{Tile} \to \bigvee_{\textit{right}(t_q) = \textit{left}(t_{q'})} (\mathtt{t_q} \wedge \Diamond_r \mathtt{t_{q'}})) \quad (20) \qquad \textit{right-left constraint}$$

$$[G](\mathsf{Up\_rel} \to \bigvee_{\textit{up}(t_q) = \textit{down}(t_{q'})} (\Diamond_l \Diamond_r \mathtt{t_q} \wedge \Diamond_r \mathtt{t_{q'}})) \quad (21) \qquad \textit{up-down constraint}$$

$$[G](\Diamond_l \mathsf{IdStart} \to \Box_l \Box_l (\mathsf{Tile} \to \bigvee_{t_q \in \mathcal{T}_b} \mathtt{t_q})) \quad (22) \qquad \textit{bottom side constraint}$$

$$[G](\Diamond_r \mathsf{IdStop} \to \Box_r \Box_r (\mathsf{Tile} \to \bigvee_{t_q \in \mathcal{T}_t} \mathtt{t_q})) \quad (23) \qquad \textit{top side constraint}$$

$$[G](((\mathsf{Tile} \wedge \Diamond_r *) \vee (\mathsf{Tile} \wedge \mathsf{Stop})) \leftrightarrow \mathsf{Rtile}) \quad (24) \qquad \textit{right side } \mathsf{Tiles}$$

$$[G](\mathsf{Tile} \wedge \Diamond_l * \leftrightarrow \mathsf{Ltile}) \quad (25) \qquad \textit{left side } \mathsf{Tiles}$$

$$[G](\mathsf{Ltile} \to \bigvee_{t_q \in \mathcal{T}_l} \mathtt{t_q}) \quad (26) \qquad \textit{left side constraint}$$

$$[G](\mathsf{Rtile} \rightarrow \bigvee_{t_q \in \mathcal{T}_r} \mathtt{t_q}) \qquad (27) \qquad \textit{right side constraint}$$

$$(24) \wedge \ldots \wedge (27) \qquad (28)$$

Theorem 7. *Given a finite set of tile types* $\mathcal{T} = \{t_1, \ldots, t_k\}$ *and a side color* \$*, the formula*

$$\Phi := \downarrow_x (7) \wedge \Diamond_l \Diamond_r \downarrow_x ((15) \wedge (18) \wedge (28))$$

is satisfiable in $\mathbb{N}$ *if and only if* $\mathcal{T}$ *can tile some finite rectangle* $\mathcal{R} = \{(x, y) \mid 1 \leq x \leq X$ *and* $1 \leq y \leq Y\}$*, for some* $X, Y \in \mathbb{N}$*, with side color* \$*.*

Proof. (Only if:): Suppose that $M, [a, b] \Vdash \Phi$. By Lemma 5, there is a sequence of points $b_1^0 = a < b_1^1 = b < \ldots < b_1^h = b_2^0 < \ldots < b_2^h = b_3^0 < \ldots < b_v^0 < \ldots < b_v^h = b_k$ such that $M, [b_r^s, b_r^{s+1}] \Vdash \mathsf{Tile}$ if and only if $s > 0$. We put $X = h - 1$ and $Y = v$. Then, by (19), it holds that, for every $s > 0$, $M, [b_r^s, b_r^{s+1}] \Vdash \mathtt{t_q}$ for a unique $\mathtt{t_q}$. We define a function $f : \mathcal{R} \rightarrow \mathcal{T}$, with $\mathcal{R} = \{(x, y) \mid 1 \leq x \leq X$ and $1 \leq y \leq Y\}$, such that $f(s, r) = t_q$, for all s, r, with $1 \leq s \leq X$, $1 \leq r \leq Y$, if and only if $M, [b_r^s, b_r^{s+1}] \Vdash \mathtt{t_q}$. By exploiting Lemmas 4, 5, and 6, as well as the conditions imposed by formulae (20-27), it can be easily shown that f defines a correct tiling of $\mathcal{R}$.

(If:) Let $f : \mathcal{R} \rightarrow \mathcal{T}$ be a correct tiling of the rectangle $\mathcal{R} = \{(x, y) \mid 1 \leq x \leq X$ and $1 \leq y \leq Y\}$ for some X and Y, and a given border color \$. We show that there exist a model M and an interval $[a, b]$ such that $M, [a, b] \Vdash \Phi$. Let $n = (X + 1) \cdot Y$. We define a model $M = \langle \mathbb{N}, V, g \rangle$ such that $M, [0, 1] \Vdash \Phi$. Since the only length variable occurring in Φ is x and it has no free occurrences there, any possible valuation of x is as good as any other, so we put $g(x) = 1$. The valuation function V is defined as follows.

$$\begin{aligned} V(\mathsf{u}) &:= \{[i, i+1] \mid 0 \leq i < n\}; \\ V(\mathsf{Start}) &:= \{[0, 1]\}; \\ V(\mathsf{Stop}) &:= \{[n-1, n]\}. \end{aligned}$$

This guarantees that $\downarrow_x (7)$ is satisfied. Now, in order to satisfy the remaining part of Φ on $[0, 1]$, it suffices to show that the formula $\downarrow_x ((15) \wedge (18) \wedge (28))$ can be satisfied on the interval $[0, X + 1]$, i.e., $(15) \wedge (18) \wedge (28)$ can be satisfied on $[0, X + 1]$ by a valuation assigning value $X + 1$ to the length variable x. In the following, we define the valuation for the remaining proposition letters:

$$\begin{aligned} V(\mathsf{Id}) &:= \{[i \cdot (X+1), (i+1) \cdot (X+1)] \mid 0 \leq i < Y\}; \\ V(*) &:= \{[i \cdot (X+1), i \cdot (X+1) + 1] \mid 0 \leq i < Y\}; \\ V(\mathsf{Tile}) &:= V(\mathsf{u}) \setminus V(*); \\ V(\mathsf{IdStart}) &:= \{[0, X+1]\}; \\ V(\mathsf{IdStop}) &:= \{[(X+1) \cdot (Y-1), (X+1) \cdot Y]\}; \\ V(\mathsf{Up\_rel}) &:= \{[i, j] \mid \delta(i, j) = X+1, [i, j] \notin V(\mathsf{Id}), 0 \leq i, j < n\}; \\ V(\mathsf{Ltile}) &:= \{[i \cdot (X+1) + 1, i \cdot (X+1) + 2] \mid 0 \leq i < Y\}; \\ V(\mathsf{Rtile}) &:= \{[i \cdot (X+1) - 1, i \cdot (X+1)] \mid 0 < i \leq Y\}. \end{aligned}$$

Finally, the valuation of the proposition letters in $\mathtt{T} = \{\mathtt{t}_1, \dots \mathtt{t_k}\}$ (*tile-variables*) is defined as follows. For each $\mathtt{t_i} \in \mathtt{T}$:

$$V(\mathtt{t_i}) := \{[i + (j-1) \cdot (X+1), i + (j-1) \cdot (X+1) + 1)] \mid f(i,j) = t_i\}.$$

It is straightforward to check that $M, [0,1] \Vdash \Phi$, hence the claim. $\Box$

Corollary 8. *The satisfiability problem for PNL+LB with one length variable and length constraints of the form* $\mathsf{len}_{=x}$ *only, over* $\mathbb{N}$*, is undecidable.*

The above reduction can be easily adapted to the case of PNL+LB with *strict semantics* (which excludes point-intervals). To this end, it suffices to replace formulae of the form $\Box_r\Box_r\psi$ (resp., $\Box_l\Box_l\psi$) by formulae of the form $\Box_r\psi \wedge \Box_r\Box_r\psi$ (resp., $\Box_l\psi \wedge \Box_l\Box_l\psi$) and formulae of the form $\Diamond_r\Diamond_r\psi$ (resp., $\Diamond_l\Diamond_l\psi$) by formulae of the form $\Diamond_r\psi \wedge \Diamond_r\Diamond_r\psi$ (resp., $\Diamond_l\psi \wedge \Diamond_l\Diamond_l\psi$). The rest of the reduction is basically the same, apart from the fact that some complications coming from point-intervals disappear, because nominals can be defined directly when the strict semantics is assumed.

We conclude the section by showing that the removal of the modal constant π does not suffice to recover decidability. First, we observe that the modal constant π is used in formulae (3) and (4) only, to force the uniqueness of the u-intervals satisfying Start and Stop, and, consequently, the uniqueness of the u-chain. We prove that the uniqueness of Stop can also be forced by the following formulae that make no use of π:

$$[G](\mathsf{Stop} \rightarrow \mathsf{u} \wedge \Box_r\Box_r\neg\mathsf{Stop}) \tag{29}$$

$$[G](\mathsf{Stop} \rightarrow \Diamond_l(\mathsf{u} \wedge \Box_l\Box_r(\Diamond_r\mathsf{Stop} \leftrightarrow \mathsf{u}))) \tag{30}$$

By contradiction, let us assume that there exist two distinct Stop-intervals, say, $[a,b]$ and $[c,d]$. Since Stop-intervals have the same length, it must hold that $a \neq c$. Without loss of generality, we assume $a < c$, and thus $b < d$. Two cases are possible. If $b \leq c$, then formula (29), over $[a,b]$, is false. Otherwise, if $c < b$, then $[a,b]$ overlaps $[c,d]$ ($a < c < b < d$), and formula (30), over $[c,d]$, is false. Indeed, consider the u-interval immediately to the left of the Stop-interval $[c,d]$, say, $[c',c]$. Since $\delta(c',c) = \delta(a,b)$, $c' < a$. Moreover, by (30), $[c',c]$ must satisfy $\Box_l\Box_r(\Diamond_r\mathsf{Stop} \leftrightarrow \mathsf{u})$. In particular, $[c',a]$ must satisfy $\Diamond_r\mathsf{Stop} \leftrightarrow \mathsf{u}$. However, $[c',a]$ satisfies $\Diamond_r\mathsf{Stop}$, but it is not a u-interval as it is shorter than $[c',c]$, which is a u-interval (contradiction). The case with Start-intervals is completely symmetric. Thus, we obtain the following.

Theorem 9. *The satisfiability problem for PNL+LB with one length variable and length constraints of the form* $\mathsf{len}_{=x}$ *only, devoid of the modal constant* π*, over* $\mathbb{N}$*, is undecidable.*

6 A decidable variant of PNL + equality constraints over length variables

As we have seen so far, even the addition of a single length variable and a binder over it to PNL yields undecidability. A natural question is whether decidability can be regained by imposing suitable restrictions on length binders, e.g., by limiting the amount of memory at their disposal. In the following, we show that this can be done by limiting the range of binders over length variables. As a matter of fact, the resulting logic improves succinctness, but no increase in expressiveness is achieved.

The idea is to replace the binder $\downarrow_x$ by a hierarchy of restricted binders $\{\downarrow_x^k \mid k \in \mathbb{N}\}$ and to properly define their semantics when the length of the current interval exceeds the limit of the binder. Formally, let $M = \langle \mathbb{N}, V, g' \rangle$. The semantics of $\downarrow_x^k$, called *truncated semantics*, is defined as follows:

$M, [a, b] \Vdash^t \downarrow_x^k \varphi$ iff

i) $\delta(a, b) \leq k$ and $M', [a, b] \Vdash \varphi$, for some $M' = \langle \mathbb{N}, V, g' \rangle$, where $g' \sim_x g$ and $g'(x) = \delta(a, b)$, or

ii) $\delta(a, b) > k$ and $M', [a, b] \Vdash \varphi$, for some $M' = \langle \mathbb{N}, V, g' \rangle$, where $g' \sim_x g$ and $g'(x) = k + 1$.

The rationale behind the above definition is evident: the binder $\downarrow_x^k$ can store the length of the current interval only if it does not exceed k; otherwise, it stores $k + 1$.

Consider now the *truncated* fragment of PNL+LB, where only restricted length binders may occur and all hybrid metric constraints occur inside the scope of a binder $\downarrow_x^k$ (as in the case of PNL + equality constraints over length variables, the only metric constraints are of the form $\mathsf{len}_{=x}$), interpreted according to the semantics given above.

Let τ^t be the translation defined as follows:

$$
\begin{aligned}
&\tau^t(p) := p \\
&\tau^t(\mathsf{len}_{=x}) := \mathsf{len}_{=x} \\
&\tau^t(\neg\psi) := \neg\tau^t(\psi) \\
&\tau^t(\psi_1 \vee \psi_2) := \tau^t(\psi_1) \vee \tau^t(\psi_2) \\
&\tau^t(\Diamond_r \psi) := \Diamond_r \tau^t(\psi) \\
&\tau^t(\Diamond_l \psi) := \Diamond_l \tau^t(\psi) \\
&\tau^t(\downarrow_x^k \psi) := (\mathsf{len}_{>k} \wedge \tau^t(\psi)[\mathsf{len}_{=k+1}/\mathsf{len}_{=x}]) \vee \bigvee_{j=0}^{k} (\mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}])
\end{aligned}
$$

The translation function τ^t maps every formula ψ of the truncated fragment of PNL+LB into a formula $\tau^t(\psi)$ of MPNL with length at most exponential in $|\psi|$. However, when applied to a sub-formula of a formula of such a fragment, τ^t does not necessarily produce an MPNL formula (the fragment is not closed under sub-formulae).

It is worth pointing out that the condition that every hybrid metric constraint of the form $\mathsf{len}_{=x}$ occurs inside the scope of a binder $\downarrow_x^k$ plays a fundamental role. Indeed, it can be easily checked that the application of τ^t to a formula with an occurrence of $\mathsf{len}_{=x}$ which is not in the scope of any binder produces a formula not belonging to MPNL.

Lemma 10. *The translation τ^t preserves the truth value of every formula ψ of the truncated fragment of PNL+LB when interpreted over $\mathbb{N}$.*

Proof. The proof is by structural induction on the input formula. All cases, but the one of formulae of the form $\downarrow_x^k \psi$, are straightforward. Thus, we focus our attention on such a case. In the following, we will denote models of MPNL formulae as pairs $\langle \mathbb{N}, V \rangle$ and models of formulae of the truncated fragment of PNL+LB as triples $\langle \mathbb{N}, V, g \rangle$. Moreover, we will denote by M both kinds of model whenever we do not need to refer to model components. We will prove that, given a formula φ of the truncated fragment of PNL+LB and a model $\langle \mathbb{N}, V, g \rangle$ for it, a model for its translation $\tau^t(\varphi)$ can be obtained by removing its third component g.

We now show that, for every valuation function V, length assignment g, and interval $[a, b]$, the following equation holds:

$$\langle \mathbb{N}, V \rangle, [a, b] \Vdash \tau^t(\downarrow_x^k \psi) \Leftrightarrow \langle \mathbb{N}, V, g \rangle, [a, b] \Vdash \downarrow_x^k \psi.$$

$\Rightarrow$ (*only if* direction) Suppose that $\langle \mathbb{N}, V \rangle, [a, b] \Vdash \tau^t(\downarrow_x^k \psi)$. Then, $\langle \mathbb{N}, V \rangle, [a, b] \Vdash (\mathsf{len}_{>k} \wedge \tau^t(\psi)[\mathsf{len}_{=k+1}/\mathsf{len}_{=x}]) \vee \bigvee_{j=0}^{k}(\mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}])$. We must prove that, for any g, $\langle \mathbb{N}, V, g \rangle, [a, b] \Vdash \downarrow_x^k \psi$. We distinguish two cases:

1. If $b - a > k$, then $\langle \mathbb{N}, V \rangle, [a, b] \nVdash \bigvee_{j=0}^{k}(\mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}])$, and thus $\langle \mathbb{N}, V \rangle, [a, b] \Vdash \mathsf{len}_{>k} \wedge \tau^t(\psi)[\mathsf{len}_{=k+1}/\mathsf{len}_{=x}]$. The following chain of implications holds:

   $\langle \mathbb{N}, V \rangle, [a, b] \Vdash \mathsf{len}_{>k} \wedge \tau^t(\psi)[\mathsf{len}_{=k+1}/\mathsf{len}_{=x}]$
   $\Rightarrow \langle \mathbb{N}, V \rangle, [a, b] \Vdash \mathsf{len}_{>k} \wedge \tau^t(\psi')$, *where ψ' is obtained from ψ by replacing every free occurrence of x by $k + 1$*
   $\Rightarrow \langle \mathbb{N}, V, g \rangle, [a, b] \Vdash \mathsf{len}_{>k} \wedge \psi'$, *by the inductive hypothesis*
   $\Rightarrow \langle \mathbb{N}, V, g \rangle, [a, b] \Vdash \mathsf{len}_{>k} \wedge \downarrow_x^k \psi'$, *as there are no free occurrences of x in ψ'*
   $\Rightarrow \langle \mathbb{N}, V, g \rangle, [a, b] \Vdash \downarrow_x^k \psi$, *by $\mathsf{len}_{>k}$ and the semantics of $\downarrow_x^k$*

2. If $b - a \leq k$, then $\langle \mathbb{N}, V \rangle, [a, b] \nVdash (\mathsf{len}_{>k} \wedge \tau^t(\psi)[\mathsf{len}_{=k+1}/\mathsf{len}_{=x}])$, and thus $\langle \mathbb{N}, V \rangle, [a, b] \Vdash \bigvee_{j=0}^{k}(\mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}])$. The following chain of

implications holds:

$$\begin{aligned}
&\langle \mathbb{N}, V\rangle, [a,b] \Vdash \bigvee_{j=0}^{k} (\mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}])\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}], \textit{ for } 0 \leq j \leq k\\
&\qquad \textit{such that } b-a=j\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \mathsf{len}_{=j} \wedge \tau^t(\psi'), \textit{ where } \psi' \textit{ is obtained from } \psi\\
&\qquad \textit{by replacing every free occurrence of } x \textit{ by } j\\
&\Rightarrow \langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \mathsf{len}_{=j} \wedge \psi', \textit{ by the inductive hypothesis}\\
&\Rightarrow \langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \mathsf{len}_{=j} \wedge \downarrow_x^k \psi', \textit{ as there are no free occurrences of } x \textit{ in } \psi'\\
&\Rightarrow \langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \downarrow_x^k \psi \textit{ by } \mathsf{len}_{=j} \textit{ and the semantics of } \downarrow_x^k
\end{aligned}$$

$\Leftarrow$ (*if* direction) Suppose that $\langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \downarrow_x^k \psi$. We show that $\langle \mathbb{N}, V\rangle$, $[a,b] \Vdash \tau^t(\downarrow_x^k \psi)$. We distinguish two cases:

1. If $b-a > k$, we have that:

$$\begin{aligned}
&\langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \downarrow_x^k \psi\\
&\Rightarrow \langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \psi', \textit{ where } \psi' \textit{ is obtained from } \psi\\
&\qquad \textit{by replacing every free occurrence of } x \textit{ by } k+1\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \tau^t(\psi'), \textit{ by the inductive hypothesis}\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \mathsf{len}_{>k} \wedge \tau^t(\psi)[\mathsf{len}_{=k+1}/\mathsf{len}_{=x}]\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \tau^t(\downarrow_x^k \psi), \textit{ by the definition of } \tau^t
\end{aligned}$$

2. If $b-a = j \leq k$, we have that:

$$\begin{aligned}
&\langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \downarrow_x^k \psi\\
&\Rightarrow \langle \mathbb{N}, V, g\rangle, [a,b] \Vdash \psi', \textit{ where } \psi' \textit{ is obtained from } \psi\\
&\qquad \textit{by replacing every free occurrence of } x \textit{ by } j\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \tau^t(\psi'), \textit{ by the inductive hypothesis}\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \mathsf{len}_{=j} \wedge \tau^t(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}]\\
&\Rightarrow \langle \mathbb{N}, V\rangle, [a,b] \Vdash \tau^t(\downarrow_x^k \psi), \textit{ by the definition of } \tau^t
\end{aligned}$$

□

Corollary 11. *The satisfiability problem for the truncated fragment of PNL+LB, over* $\mathbb{N}$*, is decidable in 2EXPSPACE, when length constraints in the formulae are represented in binary, and in 2NEXPTIME, when they are constant or represented in unary.*

Corollary 11 immediately follows from Theorem 2.

We conclude the section by pointing out that one must be extremely careful in the search for decidable variants of PNL+LB. Let us consider, for instance, the following restricted semantics for length binders proposed in [DGS10]:

$M(= \langle \mathbb{N}, V, g\rangle), [a,b] \Vdash^r \downarrow_x^k \varphi$ iff

i) $\delta(a,b) \leq k$ and $M', [a,b] \Vdash \varphi$, for all $M' = \langle \mathbb{N}, V, g'\rangle$, where $g' \sim_x g$ and $g'(x) = \delta(a,b)$, or

ii) $\delta(a,b) > k$ and $M', [a,b] \Vdash \varphi$, for all $M' = \langle \mathbb{N}, V, g'\rangle$, where g' is an assignment such that $g' \sim_x g$ and $g'(x) > k$.

Unlike the case of truncated semantics, in such a case, when the length of the current interval exceeds k, the binder $\downarrow_x^k$ stores the constraint $\mathsf{len}_{>k}$.

In [DGS10], the authors consider a variant of PNL+LB with such a restricted semantics for length binders and an additional constraint imposing that variable length constraints of the form $\mathsf{len}_{=x}$ may only occur *positively*, that is, once the formula has been transformed into the negation normal form, the logic does not allow sub-formulae of the form $\neg\mathsf{len}_{=x}$. They claim that every formula ψ of such a restricted fragment of PNL+LB can be effectively translated into a formula $\tau^r(\psi)$ of MPNL which is equisatisfiable with ψ when interpreted over $\mathbb{N}$ and has length at most exponential in $|\psi|$. The translation rule for $\downarrow_x^k$ they propose is the following one:

$$\tau^r(\downarrow_x^k \psi) := (\mathsf{len}_{>k} \wedge \tau^r(\psi)[\mathsf{len}_{>k}/\mathsf{len}_{=x}]) \vee \bigvee_{j=0}^{k} (\mathsf{len}_{=j} \wedge \tau^r(\psi)[\mathsf{len}_{=j}/\mathsf{len}_{=x}]).$$

In [DGS10], it has been shown that the translation τ^r would not properly work with variable length constraint $\mathsf{len}_{=x}$ occurring negatively, as $\neg\mathsf{len}_{=x}$ is not equivalent to $\neg\mathsf{len}_{>k}$ when $x > k$. As an example, $\downarrow_x^k (\mathsf{len}_{>k} \wedge \Diamond_r(\neg\mathsf{len}_{=x} \wedge \mathsf{len}_{>k}))$ is satisfiable according to the restricted semantics, but $\tau^r(\downarrow_x^k (\mathsf{len}_{>k} \wedge \Diamond_r(\neg\mathsf{len}_{=x} \wedge \mathsf{len}_{>k}))) = (\mathsf{len}_{>k} \wedge \Diamond_r(\neg\mathsf{len}_{>k} \wedge \mathsf{len}_{>k})) \vee \bigvee_{j=0}^{k}(\mathsf{len}_{=j} \wedge (\mathsf{len}_{>k} \wedge \Diamond_r(\neg\mathsf{len}_{=j} \wedge \mathsf{len}_{>k})))$ is not.

It is possible to show that forcing variable length constraints of the form $\mathsf{len}_{=x}$ to occur only *positively* does not suffice to guarantee that τ^r preserves (un)satisfiability (and thus the claim in [DGS10] turns out to be incorrect). As an example, consider the formula $\Diamond_r \downarrow_x^k (\mathsf{len}_{=x} \wedge p \wedge \Diamond_l\Diamond_r(\mathsf{len}_{=x} \wedge \neg p))$. Such a formula is clearly unsatisfiable according to the restricted semantics, while its translation in MPNL $\Diamond_r((\mathsf{len}_{>k} \wedge p \wedge \Diamond_l\Diamond_r(\mathsf{len}_{>k} \wedge \neg p)) \vee \bigvee_{j=0}^{k}(\mathsf{len}_{=j} \wedge p \wedge \Diamond_l\Diamond_r(\mathsf{len}_{=j} \wedge \neg p)))$ is satisfiable.

7 Undecidability of single-variable fragments of (R)PNL+FO

In this section, we focus our attention on first-order extensions of (R)PNL, (R)PNL+FO for short. As it is clear from our discussion so far, there are a number of parameters to be set for (R)PNL+FO. Besides the usual alternatives about the nature of the temporal domain, which can be finite or infinite, dense or discrete, bounded or unbounded, and so on, we have different options for the first-order domain, which can be finite or infinite, constant, variable, or expanding, and so on (as a matter of fact, we can also constrain the first-order domain to be ordered and to satisfy additional properties such as, for instance, linearity, discreteness, or denseness). Moreover, we can impose suitable syntactic restrictions to formulae, such as, for instance, limiting the number of distinct variables that may occur in a formula. Since we are interested in tight undecidability results, in contrast with decidability results for first-order point-based temporal logic, we will make quite restrictive assumptions, thus showing that the addition of a very weak first-order flavor to (R)PNL immediately yields undecidability. More precisely, we consider a first-order extension of RPNL, which involves only one variable, no first-order constants, and no free variables (this means that the variable assignment λ plays no role, and thus it can be safely omitted). Moreover, we assume a (countable) constant first-order domain $\mathfrak{D}$, and, to make the undecidability proof simpler, we choose a finite linear order as the temporal domain. At the end of the section, we will show how to adapt the proof to the case of $\mathbb{N}$.

The proof hinges on the fact that the addition of first-order features makes it possible to express properties like: "*if an interval satisfies φ, then all its beginning intervals (resp., ending intervals, strict sub-intervals) satisfy ψ*" (the set of strict sub-intervals of an interval $[a, b]$ contains all and only the intervals $[c, d]$ such that $a < c < d < b$). To this end, we make use of a unary predicate P that associates a distinct element (possibly more than one) of the first-order domain with each element of the temporal domain. More precisely, we constrain the predicate P to satisfy the following condition: for any $e \in \mathfrak{D}$, if $P(e)$ holds over an interval $[a, b]$, then $P(e)$ does not hold over any interval $[c, d]$, with $c \neq a$. Such a condition can be formally stated as follows:

$$[G](\exists x \Diamond_r P(x) \wedge \forall x(\Diamond_r P(x) \rightarrow \Box_r(\neg\pi \rightarrow \Box_r \neg P(x)))) \tag{31}$$

where $[G]$, by an abuse of notation, is the universal modality for RPNL. Given a formula φ, $[G]\varphi$ holds over an interval $[a, b]$ if and only if φ holds over $[a, b]$ and over all intervals $[c, d]$, with $c \geq b$. Formally, $[G]\varphi$ is defined as: $\varphi \wedge \Box_r\varphi \wedge \Box_r\Box_r\varphi$.

Lemma 12. *Let $M = \langle \mathbb{D}, \mathfrak{D}, \mathcal{I} \rangle$ be an RPNL+FO model and $[a, b] \in \mathbb{I}(\mathbb{D})$. If*

$$M, [a, b] \Vdash (31),$$

then, for each $c \in D$, with $c \geq b$, there exists $d \geq c$ such that $P^{\mathcal{I}([c,d])}(c_1)$ holds for some $c_1 \in \mathfrak{D}$ and, for any $e, f \geq b$, with $e \neq c$ and $e \leq f$, $P^{\mathcal{I}([e,f])}(c_1)$ does not hold.

The proof is straightforward, and thus omitted.

By using the unary predicate P above, we can define a formula $[B^{\varphi}_{\psi}]$ (resp., $[E^{\varphi}_{\psi}]$, $[D^{\varphi}_{\psi}]$) expressing the condition "*if an interval $[b, e]$ satisfies φ, then all its beginning* *intervals (resp., ending intervals, strict sub-intervals) satisfy* ψ" as follows:

$$[B^{\varphi}_{\psi}] \equiv [G]\forall x(\Diamond_r(\varphi \wedge \Diamond_r P(x)) \rightarrow \Box_r(\Diamond_r(\neg\pi \wedge \Diamond_r P(x)) \rightarrow \psi))$$
$$[E^{\varphi}_{\psi}] \equiv [G]\forall x(\Diamond_r(\varphi \wedge \Diamond_r P(x)) \rightarrow \Box_r(\neg\pi \rightarrow \Box_r(\Diamond_r P(x) \rightarrow \psi)))$$
$$[D^{\varphi}_{\psi}] \equiv [G]\forall x(\Diamond_r(\varphi \wedge \Diamond_r P(x)) \rightarrow \Box_r(\neg\pi \rightarrow \Box_r(\Diamond_r(\neg\pi \wedge \Diamond_r P(x)) \rightarrow \psi)))$$

The three formulae $[B^{\varphi}_{\psi}]$, $[E^{\varphi}_{\psi}]$, and $[D^{\varphi}_{\psi}]$ closely resemble the modalities $[B]$, $[E]$, and $[D]$ of HS for Allen's relations *begins*, *ends*, and *during*, respectively. However, they are not equivalent to them, as they do not (allow one to) refer to the beginning intervals, ending intervals, and strict subintervals of the current interval.

The following formulae allow us to define the u-chain (see Section 4):

$\neg\mathsf{u} \wedge \Diamond_r(\neg\pi \wedge \mathsf{u})$	(32)	*starts the* u*-chain*
$[G](\mathsf{u} \rightarrow (\neg\pi \wedge (\Diamond_r\mathsf{u} \vee \Box_r\pi)))$	(33)	*completes the* u*-chain*
$[B^{\mathsf{u}}_{\neg\mathsf{u}}] \wedge [B^{\mathsf{u}}_{\neg\pi\rightarrow\Box_r\neg\mathsf{u}}]$	(34)	*makes the* u*-chain unique*
$(31) \wedge (32) \wedge (33) \wedge (34)$	(35)	

Lemma 13. *Let* $M = \langle \mathbb{D}, \mathfrak{D}, \mathcal{I}\rangle$ *be an RPNL+FO model and* $[a, b] \in \mathbb{I}(\mathbb{D})$ *such that*

$$M, [a, b] \Vdash (35).$$

Then, there is a finite sequence of points $b_0 = b < b_1 < .. < b_n$*, with* $n > 0$*, such that:*

1. $M, [b_l, b_{l+1}] \Vdash \mathsf{u}$ *for each* $0 \leq l \leq n-1$*;*
2. $M, [c, d] \Vdash \mathsf{u}$ *holds for no other interval* $[c, d]$*, unless* $c < b$*.*

Proof. If $M, [a, b] \Vdash (35)$, then, by (32), the interval $[b, c]$, for some $c > b$, is a u-interval. Hence, by (33), $b(= b_0)$ starts a finite chain of u-intervals $[b_l, b_{l+1}]$, with $l \geq 0$ and $b_l < b_{l+1}$. (The satisfiability of (33) over finite temporal domains follows from the fact that the last point of the temporal domain satisfies $\Box_r\pi$.) Now suppose, by contradiction, that for some interval $[c, d]$, with $c \geq b$, it is the case that $[c, d]$ is a u-interval, but $[c, d] \neq [b_l, b_{l+1}]$ for any $l \geq 0$. Then, either $c = b_l$ for some l, thus contradicting the first conjunct of (34), or $b_l < c < b_{l+1}$, thus contradicting the second conjunct of (34). $\Box$

Notice that, unlike the case of the fragment of PNL+LB analyzed in Section 5, we cannot force the length of u-intervals to be the same.

The ld-chain can be expressed by the following formulae:

$\neg\mathsf{ld} \wedge \Diamond_r\mathsf{ld} \wedge [G]((\Diamond_r\mathsf{ld} \rightarrow \Diamond_r\mathsf{u})\wedge$		
$(\mathsf{ld} \rightarrow \neg\pi \wedge \neg\mathsf{u} \wedge (\Diamond_r\mathsf{ld} \vee \Box_r\pi)))$	(36)	*constructs the* ld*-chain*
$[B^{\mathsf{ld}}_{\neg\mathsf{ld}}] \wedge [B^{\mathsf{ld}}_{\neg\pi\rightarrow\Box_r\neg\mathsf{ld}}]$	(37)	*makes the* ld*-chain unique*
$(36) \wedge (37)$	(38)	

Lemma 14. *Let $M = \langle \mathbb{D}, \mathfrak{D}, \mathcal{I} \rangle$ be an RPNL+FO model and $[a, b] \in \mathbb{I}(\mathbb{D})$ such that*

$$M, [a, b] \Vdash (35) \wedge (38).$$

Then, there exist a positive integer v, a finite sequence of positive integers $m_1, m_2, \ldots, m_v$, and a finite sequence of points $b_0^1 < b_1^1 < \ldots < b_{m_1}^1 = b_0^2 < \ldots < b_{m_2}^2 = \ldots = b_0^{v-1} < \ldots < b_{m_{v-1}}^{v-1} = b_0^v < \ldots < b_{m_v}^v$ such that for each $1 \leq s \leq v$, $M, [b_0^s, b_{m_s}^s] \Vdash \mathsf{Id}$ and, for each $0 \leq l < m_s$, $M, [b_l^s, b_{l+1}^s] \Vdash \mathsf{u}$, and no other interval $[c, d]$ satisfies Id, unless $c < b$.

Proof. First, by Lemma 13, there is a finite sequence of points $b_0 = b < b_1 < \ldots < b_n$, with $n > 0$, which defines a finite chain of u-intervals. By (36), b_0 begins an Id-interval, each Id-interval must start at some b_l and spans several u-intervals, and each Id-interval that does not end at the last point of the finite linear order is followed by another Id-interval. Moreover, since the linear order is finite, there are finitely many Id-intervals. Let v be the number of Id-intervals. The sequence $b_0 < b_1 < \ldots < b_n$ can be rewritten as $b_0^1 = b_0 < b_1^1 < \ldots < b_{m_1}^1 = b_0^2 < \ldots < b_{m_2}^2 = \ldots = b_0^{v-1} < \ldots < b_{m_{v-1}}^{v-1} = b_0^v < \ldots < b_{m_v}^v = b_n$, as required. To conclude the proof, we must show that there are no other Id-intervals apart from those of the form $[b_0^s, b_{m_s}^s]$. This can be proved exactly as in Lemma 13, by using (37). $\square$

The above lemma guarantees the existence of an Id-chain (but it does not constrain the number of u-intervals each Id-interval consists of).

$[G](\mathsf{u} \wedge \Box_r \neg \mathsf{Id} \wedge \Box_r \Box_r \neg \mathsf{Id} \leftrightarrow \mathsf{Final})$	(39)	*set* Final
$[G](\mathsf{u} \rightarrow (\neg \mathsf{Final} \leftrightarrow \Diamond_r \mathsf{Up\_rel}))$	(40)	*start the* Up_rel-*chain*
$\neg \mathsf{Up\_rel} \wedge \neg \Diamond_r \mathsf{Up\_rel} \wedge [G](\Diamond_r \mathsf{Up\_rel} \rightarrow \Diamond_r \mathsf{u})$	(41)	Up_rel *starts with a* u
$[G](\mathsf{Up\_rel} \rightarrow \neg \mathsf{Id} \wedge \neg \pi \wedge \neg \mathsf{u} \wedge \Diamond_r \mathsf{u})$	(42)	Up_rel *spans various* u
$\forall x(\Diamond_r(\mathsf{Id} \wedge ((\Diamond_r \mathsf{u} \wedge \Diamond_r P(x)) \vee \Diamond_r(\Diamond_r \mathsf{u} \wedge \Diamond_r P(x)))) \rightarrow$ $\Diamond_r \Diamond_r(\mathsf{Up\_rel} \wedge \Diamond_r P(x)))$	(43)	*no* u *is skipped*
$[B^{\mathsf{Up\_rel}}_{\neg \mathsf{Up\_rel}}] \wedge [E^{\mathsf{Up\_rel}}_{\neg \mathsf{Up\_rel}}] \wedge [D^{\mathsf{Up\_rel}}_{\neg \mathsf{Up\_rel}}]$	(44)	Up_rel*s are unique*
$[B^{\mathsf{Up\_rel}}_{\neg \mathsf{Id}}] \wedge [E^{\mathsf{Up\_rel}}_{\neg \mathsf{Id}}] \wedge [D^{\mathsf{Up\_rel}}_{\neg \mathsf{Id}}]$	(45)	*the* Up_rel-*chain...*
$[D^{\mathsf{Id}}_{\neg \mathsf{Up\_rel}}]$	(46)	*...overlaps the* Id-*chain*
$(39) \wedge \ldots \wedge (46)$	(47)	

Now, we force the proposition letter Up_rel to correctly encode the relation that connects pairs of tiles which are vertically adjacent. On the basis of the correspondence between tiles and u-intervals, we express the relation of vertical adjacency between pairs of tiles in terms of the truth of Up_rel over intervals whose endpoints are the endpoints of suitable u-intervals. Formally, we say that two u-intervals $[b_l, b_{l+1}]$ and $[b_{l'}, b_{l'+1}]$ are *above-connected* if (and only if) $[b_{l+1}, b_{l'}]$ is an Up_rel-interval. We must guarantee that

(i) each Id-interval spans the same number of u-intervals (tiles), (ii) each u-interval of an Id-interval is connected to exactly one u-interval of the next Id-interval (if any), and to exactly one u-interval of the previous Id-interval (if any). To this end, we first label u-intervals belonging to the last Id-interval with the proposition letter Final (formula (39) below). Next, we force each $\mathsf{Up\_rel}$-interval to start with a u-interval (formula (41)) and to span several u-intervals (formula (42)). Then, making use of $\mathsf{Up\_rel}$-intervals, we constrain each u-interval not belonging to the last Id-interval to be connected to at least one u-interval in the future (formula (40)) and each u-interval not belonging to the first Id-interval to be connected to at least one u-interval in the past (formula (43)). To guarantee that each u-interval is connected to at most one u-interval in the future and to at most one u-interval in the past, we force any $\mathsf{Up\_rel}$-interval not to be a beginning interval, an ending interval, or a strict sub-interval of another $\mathsf{Up\_rel}$-interval (formula (44)). Finally, to guarantee that $\mathsf{Up\_rel}$-intervals connect u-intervals belonging to consecutive Id-intervals, we constrain any Id-interval not to be a beginning interval (resp., ending interval, strict sub-interval, strict super-interval) of an $\mathsf{Up\_rel}$-interval (formulae (45) and (46)).

Lemma 15. *Let $M = \langle \mathbb{D}, \mathfrak{D}, \mathcal{I} \rangle$ be a PNL+FO model and $[a,b] \in \mathbb{I}(\mathbb{D})$ such that*

$$M, [a,b] \Vdash (35) \wedge (38) \wedge (47),$$

and let $b_0^1 < b_1^1 < \ldots < b_{m_1}^1 = b_0^2 < \ldots < b_{m_2}^2 = \ldots = b_0^{v-1} < \ldots < b_{m_{v-1}}^{v-1} = b_0^v < \ldots < b_{m_v}^v$ be the sequence of points whose existence is guaranteed by Lemma 14. Then, for each $1 \leq s < v$ and $0 \leq l < m_s$, $M, [b_{l+1}^s, b_l^{s+1}] \Vdash \mathsf{Up\_rel}$, and, for each $1 \leq s, s' \leq v$, $m_s = m_{s'}$. Moreover, no other interval $[c,d]$ satisfies $\mathsf{Up\_rel}$, unless $c < b$.

Proof. By (39) and (40), no u-interval belonging to the last Id-interval meets an $\mathsf{Up\_rel}$-interval. Let $[b_i^s, b_{i+1}^s]$ be a u-interval not belonging to the last Id-interval. By (40), b_{i+1}^s starts an $\mathsf{Up\_rel}$-interval, that, by (42), is neither a point-interval nor a u-interval, and it meets a u-interval. Hence, it must span more than one u-interval, ending at some $b_{i'}^{s'} > b_{i+2}^s$. By (46), $b_{i'}^{s'} \geq b_0^{s+1}$, and, by (45), $b_{i'}^{s'} < b_{m_{s+1}}^{s+1}$. We prove by induction on i that $[b_i^s, b_{i+1}^s]$ is above-connected to $[b_i^{s+1}, b_{i+1}^{s+1}]$. Let us consider $[b_0^s, b_1^s]$ and suppose, by contradiction, that it is above-connected to $[b_{i'}^{s+1}, b_{i'+1}^{s+1}]$, for some $i' > 0$. Now, by (43), there must be an $\mathsf{Up\_rel}$-interval ending at b_0^{s+1}. By the second conjunct of (45), such an $\mathsf{Up\_rel}$-interval cannot start at a point $b_h^{s'}$, with $s' < s$ and $0 \leq h \leq m_{s'}$, and, by (42), it cannot start at b_0^s. Moreover, by the third conjunct of (44), it cannot start at a point b_h^s, with $1 < h < m_{s'}$ (in such a case, the $\mathsf{Up\_rel}$-interval $[b_h^s, b_0^{s+1}]$ would be a strict sub-interval of the $\mathsf{Up\_rel}$-interval $[b_1^s, b_{i'}^{s+1}]$). Finally, by the first conjunct of (44), it cannot at start at b_1^s (there would be two distinct $\mathsf{Up\_rel}$-intervals starting at the same point). Hence, $[b_0^s, b_1^s]$ is above-connected to $[b_0^{s+1}, b_1^{s+1}]$. Now we suppose the thesis to hold for $1, 2, \ldots, i$, with $i < m_s$, and we prove it for $i+1$. The argument is quite similar to the one for the base case, provided that b_0^{s+1} is replaced by b_i^{s+1}. By contradiction, let us assume $[b_i^s, b_{i+1}^s]$ to be above-connected to $[b_{i'}^{s+1}, b_{i'+1}^{s+1}]$, for some $i' > i$ (the case $i' < i$ is excluded by the second disjunct of (44) making use of the inductive hypothesis). By (43), there must be an $\mathsf{Up\_rel}$-interval ending at b_i^{s+1}. By the third conjunct of (44), this $\mathsf{Up\_rel}$-interval can start neither at some point before b_i^s nor at some point after b_{i+1}^s, and, by the first conjunct of (44), it can start neither at b_i^s nor at b_{i+1}^s. Hence, $[b_i^s, b_{i+1}^s]$ is above-connected to $[b_i^{s+1}, b_{i+1}^{s+1}]$.

To prove that, for each $1 \leq s, s' \leq v$, $m_s = m_{s'}$, it suffices to observe that, by (43), the left endpoint of every u-interval $[b_i^s, b_{i+1}^s]$, with $s > 1$ (that is, to the right of the first Id-interval), is the right endpoint of an $\mathsf{Up\_rel}$-interval. Such an $\mathsf{Up\_rel}$-interval cannot start before b_0^{s-1} (by the second and the third conjunct of (45)), at b_0^{s-1} (by (42) and by the first conjunct of (45)), or after b_0^s (by 46). The claim immediately follows.

Finally, since, by (41) any $\mathsf{Up\_rel}$-interval (to the right of b) starts with a u-interval, by the first conjunct of (44), we can conclude that no other interval $[c, d]$, with $c \geq b$, satisfies $\mathsf{Up\_rel}$. $\square$

Finally, we can force all color-matching conditions to be respected, by means of the following set of formulae, where $\mathcal{T}_r$ (resp., $\mathcal{T}_l$, $\mathcal{T}_t$, $\mathcal{T}_b$) is the subset of $\mathcal{T}$ containing all and only those tiles whose right (resp., left, up, down) side is colored with \$.

$$[G](\mathsf{u} \rightarrow \bigvee_{t_q \in \mathcal{T}} \mathtt{t_q} \wedge \bigwedge_{t_q \neq t_{q'}} \neg(\mathtt{t_q} \wedge \mathtt{t_{q'}})) \tag{48}$$

put the tiles

$$[G](\bigvee_{t_q \in \mathcal{T}} \mathtt{t_q} \rightarrow \mathsf{u}) \tag{49}$$

tiles are only u

$$[G](\mathsf{u} \wedge \neg\Diamond_r \mathsf{Id} \wedge \neg\Box_r \pi \rightarrow \bigvee_{\textit{right}(t_q)=\textit{left}(t_{q'})} (\mathtt{t_q} \wedge \Diamond_r \mathtt{t_{q'}})) \tag{50}$$

right-left constraint

$$[G](\mathsf{u} \wedge \Diamond_r \mathsf{Up\_rel} \rightarrow \bigvee_{\textit{up}(t_q)=\textit{down}(t_{q'})} (\mathtt{t_q} \wedge \Diamond_r(\mathsf{Up\_rel} \wedge \Diamond_r \mathtt{t_{q'}}))) \tag{51}$$

up-down constraint

$$[G](\Diamond_r \mathsf{Id} \rightarrow \Diamond_r \bigvee_{t_q \in \mathcal{T}_l} \mathtt{t_q}) \tag{52}$$

left side

$$[G](((\mathsf{u} \wedge \Diamond_r \mathsf{Id}) \vee (\mathsf{u} \wedge \Box_r \pi)) \rightarrow \bigvee_{t_q \in \mathcal{T}_r} \mathtt{t_q}) \tag{53}$$

right side

$$\forall x(\Diamond_r(\mathsf{Id} \wedge \Diamond_r P(x)) \rightarrow [G](\mathsf{u} \wedge (\Diamond_r P(x) \vee \Diamond_r\Diamond_r P(x)) \rightarrow \bigvee_{t_q \in \mathcal{T}_b} \mathtt{t_q})) \tag{54}$$

bottom side

top side

$$[G](\mathsf{u} \wedge \mathsf{Final} \rightarrow \bigvee_{t_q \in \mathcal{T}_t} \mathtt{t_q}) \tag{55}$$

$$(48) \wedge \ldots \wedge (55) \tag{56}$$

Theorem 16. *Given a finite set of tile types* $\mathcal{T} = \{t_1, \ldots, t_k\}$ *and a side color* \$, *the* *formula*

$$\Phi := (35) \wedge (38) \wedge (47) \wedge (56)$$

is satisfiable in a finite linearly-ordered temporal domain if and only if $\mathcal{T}$ *can tile a finite rectangle* $\mathcal{R} = \{(x, y) \mid 1 \leq x \leq X$ *and* $1 \leq y \leq Y\}$, *for some* $X, Y \in \mathbb{N}$, *with side color* \$.

Proof. (Only if:): Suppose that $M, [a, b] \Vdash \Phi$, and let $b_0^1 < b_1^1 < \ldots < b_m^1 = b_0^2 < \ldots < b_m^2 = \ldots = b_0^{v-1} < \ldots < b_m^{v-1} = b_0^v < \ldots < b_m^v$ be the sequence of points whose existence is guaranteed by Lemmas 14 and 15. We put $X = m$ and $Y = v$. Then, we define a function $f : \mathcal{R} \to \mathcal{T}$, with $\mathcal{R} = \{(x, y) \mid 1 \leq x \leq X$ and $1 \leq y \leq Y\}$, such that, for all s, r, with $0 \leq s \leq X - 1$, $1 \leq r \leq Y$, $f(s + 1, r) = t_q$ if and only if $M, [b_s^r, b_{s+1}^r] \Vdash \mathtt{t_q}$. By exploiting Lemmas 13, 14, and 15, it can be easily shown that f defines a correct tiling of $\mathcal{R}$.

(If:) Let $f : \mathcal{R} \to \mathcal{T}$ be a correct tiling of the rectangle $\mathcal{R} = \{(x, y) \mid 1 \leq x \leq X$ and $1 \leq y \leq Y\}$ for some X and Y, and a given border color \$. We show that there exist a model M and an interval $[a, b]$ such that $M, [a, b] \Vdash \Phi$. Let $D = \mathfrak{D} = \mathbb{N}\mid_{X \cdot Y+1}$ $(= \{0, 1, \ldots, X \cdot Y + 1\})$, and let M be the RPNL+FO model built on these two domains. We define an interpretation $\mathcal{I}$ such that $M, [0, 1] \Vdash \Phi$ (we tacitly assume that all non-listed interpretations are false). First, to guarantee that (35) is satisfied, we put

$$\mathsf{u}^{\mathcal{I}([i,i+1])} \quad \text{for all } 1 \leq i \leq X \cdot Y$$

Then, to satisfy the other conjuncts of Φ on $[0, 1]$, we define the interpretation of the remaining proposition letters and of the predicate symbol P as follows:

$$\begin{array}{ll} P^{\mathcal{I}([i,j])}(i) & \text{for all } i, j > 0 \\ \mathsf{Id}^{\mathcal{I}([i \cdot X+1, (i+1) \cdot X+1])} & \text{for all } 0 \leq i \leq Y - 1 \\ \mathsf{Up\_rel}^{\mathcal{I}([i, i+X-1])} & \text{for all } 2 \leq i \leq X \cdot (Y - 1) + 1 \\ \mathsf{Final}^{\mathcal{I}([i,i+1])} & \text{for all } X \cdot (Y - 1) + 1 \leq i \leq X \cdot Y \end{array}$$

Finally, for each $\mathtt{t_q} \in \mathtt{T}$, we put:

$$\mathtt{t_q}^{\mathcal{I}([i,i+1])} \Leftrightarrow f(s + 1, r) = t_q \quad \begin{array}{l} \text{for all } i = X \cdot (r - 1) + s + 1), \\ \text{with } 1 \leq r \leq Y \text{ and } 0 \leq s \leq X - 1. \end{array}$$

$\square$

To adapt the proof to structures based on the linear order $\mathbb{N}$, we need to force the u-chain and the Id-chain to be finite. In the finite case, we can directly exploit the finiteness of the linear order (intervals whose right endpoint is the last point of the domain are all and only those intervals that satisfy the formula $\Box_r \pi$). In the case of $\mathbb{N}$, we can use a proposition letter Stop to label all and only those intervals starting at a given point a of the linear order, and take a as the last point of the finite linear order (which can be detected by means of formulae of the form $\Box_r\mathsf{Stop}$ or, equivalently, $\Diamond_r\mathsf{Stop}$). The following formulae define the properties of Stop.

$$\Diamond_r(\neg\pi \wedge \Diamond_r\mathsf{Stop}) \tag{57}$$

$$\Box_r(\Diamond_r\mathsf{Stop} \rightarrow \Box_r\mathsf{Stop} \wedge \Box_r\neg\mathsf{u} \wedge \Box_r\Box_r\neg\mathsf{u} \wedge \Box_r(\neg\pi \rightarrow \Box_r\neg\mathsf{Stop})) \tag{58}$$

Now, it suffices to substitute $\Diamond_r\mathsf{Stop}$ by $\Box_r\pi$ in formulae (33), (36), (50), and (53).

Corollary 17. *The satisfiability problem for RPNL+FO with only one first-order variable, interpreted over finite domains or the natural numbers, is undecidable.*

It is worth pointing out that the above proof does not make any assumption on the semantics being strict or non-strict, and thus the corollary holds in both cases. On the other hand, the proof makes a massive use of the modal constant π (when the strict semantic is assumed, the constant π can simply be interpreted as the constant $\perp$) and its elimination does not seem to be trivial. Therefore, the satisfiability problem for RPNL+FO, devoid of π, when the non-strict semantics is assumed, is still open.

8 Conclusions

Point-based temporal logics have been successfully exploited in various computer science areas, ranging from well-established areas, like program specification and verification, knowledge representation and reasoning, and temporal databases, to emerging ones, like multi-agent systems and bioinformatics. We argue that interval temporal logics are particularly well suited for a number of applications, including natural language processing, constraint management, planning and synthesis of plan controllers, and temporal data aggregation. Unfortunately, for a long time, undecidability or high complexity of most of them has discouraged their extensive study in computer science. The recent discovery of some quite expressive, yet decidable, interval temporal logics will hopefully change the scenario in the years to come and give a boost to the search for tractable fragments and their application in the above-mentioned domains.

PNL is one of the most significant examples of a genuine interval-based temporal logic which has been shown to be decidable. Decidability is preserved when PNL is enriched with metric features that allow one to constrain the length of an interval over natural numbers (Metric PNL). In this paper, we have shown that the extension of (Metric) PNL with variables and binders over interval lengths is quite natural, but yields undecidability, even in the case of very restricted fragments. Furthermore, we have proved that another natural extension of propositional temporal logics, namely, the one obtained by replacing proposition letters by first-order formulae, oversteps the decidability barrier even in a very restricted case like that of monodic first-order formulae interpreted over finite linear orders or natural numbers.

Since our main goal is to find more expressive, yet decidable, extensions of (Metric) PNL, at a first glance, these results may appear interesting but discouraging, taking into account that we already know that the addition of any other independent modality from Allen's repository makes PNL undecidable over natural numbers. However, a careful look at the undecidability proofs shows that there is still room for improvements. In the undecidability proof for PNL+LB, the use of both modalities (future and past) turns out to be essential for the reduction, and, in the undecidability proof for PNL+FO, one modality suffices, but the modal constant π plays a crucial role in the reduction. Hence, there are still interesting sub-fragments to be investigated to further refine the boundary between decidable and undecidable variants and extensions of PNL. Moreover, in order to recover decidability, it is worth considering the possibility of constraining the interactions between the modal and first-order components of (R)PNL+FO by means of suitable syntactic rules. It would also be interesting to investigate whether interval logics can be shown to be decidable (via the application of Rabin's theorem) when interpreted in monadic second-order theories of suitable structures.

Finally, the problem of providing a sound and complete axiomatization of the considered logics deserves to be investigated as well. Such a problem is apparently quite difficult for PNL and MPNL over natural numbers. In the case of PNL+LB and (R)PNL +FO, thanks to the higher expressive power of the logics, it might be easier.

Acknowledgements

We would like to acknowledge the financial support from the HYLOCORE project, funded by the Danish Natural Science Research Council (V. Goranko), the Spanish MEC projects *TIN2009-14372-C03-01* and *RYC-2011-07821* (G. Sciavicco), the Icelandic Research Fund project 100048021 *Processes and Modal Logics* (D. Della Monica), and the Italian PRIN project *Innovative and multi-disciplinary approaches for constraint and preference reasoning* (A. Montanari and D. Della Monica).

References

[All83] Allen, J. F. "Maintaining knowledge about temporal intervals". Communications of the ACM, 26, 11, 832–843, 1983.

[BBF+01] Berard, B., Bidoit, M., Finkel, A., Laroussinie, F., Petit, A., Petrucci, L., Schnoebelen, P., McKenzie, P. Systems and Software Verification: Model-Checking Techniques and Tools. Springer, 2001.

[BDG+08] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G. "Decidable and Undecidable Fragments of Halpern and Shoham's Interval Temporal Logic: Towards a Complete Classification". In Proc. of 15th International Confergernce on Logic for Programming, Artificial Intelligence, and Reasoning (LPAR), volume 5330 of *LNCS*, 590–604. Springer, 2008.

[BDG+09] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G. "Undecidability of interval temporal logics with the *Overlap* modality". In Proc. of 16th International Symposium on Temporal Representation and Reasoning (TIME), 88–95. IEEE Computer Society Press, 2009.

[BDG+10] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G. "Metric propositional neighborhood logic: Expressiveness, decidability, and undecidability". In Proc. of the 19th European Conference on Artificial Intelligence (ECAI), 695–700. IOS Press, 2010.

[BDG+11] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G. "Metric propositional neighborhood logics on natural numbers". Software and Systems Modeling, 1–20, 2011. ISSN 1619-1366. URL `http://dx.doi.org/10.1007/s10270-011-0195-y`. 10.1007/s10270-011-0195-y.

[BDMG+11] Bresolin, D., Della Monica, D., Goranko, V., Montanari, A., Sciavicco, G. "The dark side of interval temporal logic: sharpening the undecidability border". In Proc. of the 18th International Symposium on Temporal Representation and Reasoning (TIME), 131–138. IEEE Computer Society Press, 2011.

[BdRV02] Blackburn, P., de Rijke, M., Venema, Y. Modal Logic. Cambridge University Press, 2002.

[BGG97] Börger, E., Grädel, E., Gurevich, Y. The Classical Decision Problem. Perspectives of Mathematical Logic. Springer, 1997.

[BGMS09] Bresolin, D., Goranko, V., Montanari, A., Sciavicco, G. "Propositional interval neighborhood logics: Expressiveness, decidability, and undecidable extensions". Annals of Pure and Applied Logic, 161, 3, 289–304, 2009.

[BHH07] Bolander, T., Hansen, J., Hansen, M. R. "Decidability of a Hybrid Duration Calculus". Electronic Notes in Theoretical Computer Science, 174, 6, 113–133, 2007.

[BLR95] Bouajjani, A., Lakhnech, Y., Robbana, R. "From Duration Calculus to linear hybrid automata". In Proc. of the 7th International Conference on Computer Aided Verification (CAV), volume 939 of *LNCS*, 196–210. Springer, 1995.

[BMS07a] Bresolin, D., Montanari, A., Sala, P. "An optimal tableau-based decision algorithm for Propositional Neighborhood Logic". In Proc. of the 24th International Symposium on Theoretical Aspects of Computer Science (STACS), volume 4393 of *LNCS*, 549–560. Springer, 2007.

[BMS07b] Bresolin, D., Montanari, A., Sciavicco, G. "An optimal decision procedure for Right Propositional Neighborhood Logic". Journal of Automated Reasoning, 38, 1-3, 173–199, 2007.

[BMSS11a] Bresolin, D., Montanari, A., Sala, P., Sciavicco, G. "An optimal decision procedure for MPNL over the integers". In Proc. of the 2nd International Symposium on Games, Automata, Logics and Formal Verification (GandALF), volume 54, 192–206. EPTCS, 2011.

[BMSS11b] Bresolin, D., Montanari, A., Sala, P., Sciavicco, G. "Optimal tableau systems for propositional neighborhood logic over all, dense, and discrete linear orders". In Proc. of TABLEAUX 2011, volume 6793 of *LNAI*, 73–87. Springer, 2011.

[BS95] Blackburn, P., Seligman, J. "Hybrid languages". Journal of Logic, Language and Information, 4, 3, 251–272, 1995.

[BT99] Blackburn, P., Tzakova, M. "Hybrid languages and temporal logic". Logic Journal of the IGPL, 7, 1, 27–54, 1999.

[CE81] Clarke, E. M., Emerson, E. A. "Design and synthesis of synchronization skeletons using branching time temporal logic". In Proc. of Workshop on Logics of Programs, volume 131 of *LNCS*, 52–71. Springer, 1981.

[CGP99] Clarke, E., Grumberg, O., Peled, D. Model Checking. MIT Press, 1999.

[CP03] Chakravorty, G., Pandya, P. K. "Digitizing Interval Duration Logic". In Proc. of the 15th International Conference on Computer Aided Verification (CAV), volume 2725 of *LNCS*, 167–179. Springer, 2003.

[Del11] Della Monica, D. Expressiveness, Decidability, and Undecidability of Interval Temporal Logic. Ph.D. thesis, Dept. of Mathematics and Computer Science, University of Udine, 2011.

[DGS10] Della Monica, D., Goranko, V., Sciavicco, G. "Hybrid Metric Propositional Neighborhood Logics with Interval Length Binders". In Proc. of 2010 International Workshop on Hybrid Logic and Applications (HyLo). 2010.

[FH07] Fränzle, M., Hansen, M. R. "Deciding an interval logic with accumulated durations". In Proc. of the 13th International Conference on Tools and Algorithms for the Construction and Analysis of Systems (TACAS), volume 4424 of *LNCS*, 201–215. Springer, 2007.

[FH08] Fränzle, M., Hansen, M. R. "Efficient model checking for Duration Calculus based on branching-time approximations". In Proc. of the 6th Conference on Software Engineering and Formal Methods (SEFM), 63–72. 2008.

[Frä04] Fränzle, M. "Model-checking dense-time Duration Calculus". Formal Aspects of Computing, 16, 2, 121–139, 2004.

[GMS03] Goranko, V., Montanari, A., Sciavicco, G. "Propositional interval neighborhood temporal logics". Journal of Universal Computer Science, 9, 9, 1137–1167, 2003.

[Gor96] Goranko, V. "Hierarchies of modal and temporal logics with reference pointers". Journal of Logic, Language and Information, 5, 1, 1–24, 1996.

[HH07] Hansen, M., Hung, D. V. "A theory of Duration Calculus with application". In Domain Modeling and the Duration Calculus, volume 4710 of *LNCS*, 119–176. Springer, 2007.

[HHK96] Hardin, R., Har'El, Z., Kurshan, R. "COSPAN". In Proc. of the 8th International Conference on Computer Aided Verification (CAV), volume 1102 of *LNCS*, 423–427. Springer, 1996.

[Hol03] Holzmann, G. J. The SPIN Model Checker. Addison-Wesley, 2003.

[HS91] Halpern, J., Shoham, Y. "A propositional modal logic of time intervals". Journal of the ACM, 38, 4, 935–962, 1991.

[HWZ00] Hodkinson, I., Wolter, F., Zakharyaschev, M. "Decidable fragment of first-order temporal logics". Annals of Pure and Applied Logic, 106, 1-3, 85–134, 2000.

[HZ97] Hansen, M. R., Zhou, C. "Duration Calculus: Logical foundations". Formal Aspects of Computing, 9, 283–330, 1997.

[KP05] Krishna, S. N., Pandya, P. K. "Modal Strength Reduction in Quantified Discrete Duration Calculus". In Proc. of the 25th Conference on Foundations of Software Technology and Theoretical Computer Science (FSTTCS), volume 3821 of *LNCS*, 444–456. 2005.

[Lod00] Lodaya, K. "Sharpening the undecidability of interval temporal logic". In Proc. of 6th Asian Computing Science Conference, volume 1961 of *LNCS*, 290–298. Springer, 2000.

[LPRT95] Lodaya, K., Parikh, R., Ramanujam, R., Thiagarajan, P. S. "A logical study of distributed transition systems". Information and Computation, 119, 1, 91 – 118, 1995.

[McM93] McMillan, K. L. Symbolic Model Checking. Kluwer Academic Publishers, 1993.

[MFHR08] Meyer, R., Faber, J., Hoenicke, J., Rybalchenko, A. "Model checking Duration Calculus: a practical approach". Formal Aspects of Computing, 20, 4-5, 481–505, 2008.

[Mon08] Montanari, A. "Back to interval temporal logics". In Proc. of the 24th International Conference on Logic Programming (ICLP), volume 5366 of *LNCS*, 11–13. Springer, 2008.

[Mos83] Moszkowski, B. Reasoning about digital circuits. Tech. rep. stan-cs-83-970, Dept. of Computer Science, Stanford University, Stanford, CA, 1983.

[MPS10] Montanari, A., Puppis, G., Sala, P. "Maximal decidable fragments of Halpern and Shoham's modal logic of intervals". In Proc. of the 37th International Colloquium on Automata, Languages and Programming - Part II (ICALP), volume 6199 of *LNCS*, 345–356. 2010.

[MS12] Montanari, A., Sala, P. "An optimal tableau system for the logic of temporal neighborhood over the reals". In Proc. of the 19th International Symposium on Temporal Representation and Reasoning (TIME), 39–46. IEEE Computer Society Press, 2012.

[Ott01] Otto, M. "Two variable first-order logic over ordered domains". Journal of Symbolic Logic, 66, 2, 685–702, 2001.

[Pan01] Pandya, P. K. "Specifying and deciding quantified discrete-time duration calculus formulae using DCVALID". In Proc. of the International Workshop on Real-time Tools (co-located with CONCUR 2001). 2001.

[Pan02] Pandya, P. K. "Interval duration logic: Expressiveness and decidability". Electronic Notes in Theoretical Computer Science, 65, 6, 254 – 272, 2002.

[Pnu77] Pnueli, A. "The temporal logic of programs". In Proc. of the 18th Annual Symposium on Foundations of Computer Science (FOCS), 46–57. IEEE Computer Society, 1977.

[Var08] Vardi, M. Y. "From monadic logic to PSL". In Pillars of Computer Science, volume 4800 of *LNCS*, 656–681. Springer, 2008.

[Ven91] Venema, Y. "A modal logic for chopping intervals". Journal of Logic and Computation, 1, 4, 453–476, 1991.

[ZH98] Zhou, C., Hansen, M. "An adequate first order interval logic". In Proc. of the 1997 International Symposium on Compositionality: The Significant Difference, volume 1536 of *LNCS*, 584–608. Springer, 1998.

[ZH04] Zhou, C., Hansen, M. R. Duration Calculus: A Formal Approach to Real-Time Systems. EATCS Monographs in Theoretical Computer Science. Springer, 2004.

[ZHR91] Zhou, C., Hoare, C. A. R., Ravn, A. P. "A calculus of durations". Information Processing Letters, 40, 5, 269–276, 1991.

[ZJLX94] Zhou, C., Jingzhong, Z., Lu, Y., Xiaoshan, L. "Linear Duration Invariants". In Proc. of the 8th International Symposium on Formal Techniques in Real-Time and Fault-Tolerant Systems (FTRTFT), volume 863 of *LNCS*, 86–109. Springer, 1994.